Typing and genome comparison of *Neisseria meningitidis* by DNA-microarrays

Dissertation zur Erlangung des

naturwissenschaftlichen Doktorgrades

der Bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Halina Świderek

aus Tomaszów Lubelski in Polen

Würzburg 2005

Eingereicht am:

Mitglieder der Promotionskommission:

Vorsitzender: Prof. Dr. U. Scheer

Gutachter: Prof. Dr. U. Vogel

Gutachter: Prof. Dr. R. Gross

Tag des Promotionskolloquiums:

Doktorurkunde ausgehändigt am:

Erklärungen

Hiermit versichere ich, die vorliegende Dissertation selbständig angefertigt und nur die angegeben on Quellen und Hilfsmittel benutzt zu haben.

Diese Arbeit hat bisher in gleicher oder ähnlicher Form keinem anderen Prüfungsverfahren vorgelegen.

1990 habe ich den akademischen Grad einer Diplom-Biologin erworben.

Halina Świderek

Würzburg, den 31.01.2005

Acknowledgements

Table of contents

1. Zusammenfassung

In der vorliegenden Doktorarbeit wurden zwei Projekte zum Einsatz der *microarray*-Technologie in der molekularen Epidemiologie von *Neisseria meningitidis* bearbeitet. Im Rahmen des ersten Projektes wurde die Einsatzmöglichkeit der *microarray*-Technologie unter Verwendung von Oligonukleotiden, die von bekannten Polymorphismen abgeleitet waren, für die Typisierung von *N. meningitidis* nach dem Prinzip der Multilokus-Sequenztypisierung (MLST) untersucht. Um alle bekannten Polymorphismen abzudecken, wurde im Vergleich zu der so genannten *tiling*-Methode eine deutlich geringere Anzahl an Oligonukleotiden benötigt. Vorversuche mit Oligonukleotiden einer Länge von 28-32 Nukleotiden zeigten, dass unter den angewandten Hybridisierungsbedingungen eine hohe Spezifität erzielt werden konnte. Dennoch lag die durchschnittliche Fehlidentifikationsrate der Oligonukleotide von *microarrays*, die die genetische Diversität von drei MLST-Loci repräsentierten, trotz mehrerer Optimierungsschritte der Oligonukleotide über 1,8%. Es ist anzunehmen, dass die hohe Dichte an Polymorphismen und die große Variation des GC-Gehaltes der MLST-Loci von *N. meningitidis* den erfolgreichen Einsatz eines MLST-*microarrays* mit Oligonukleotiden, die von bekannten Polymorphismen abgeleitet waren, verhinderten.

In den achtziger Jahren des letzten Jahrhunderts tauchte der ET-15-Klon als Abkömmling des Sequenztyp (ST-) 11-Komplexes von *N. meningitidis* auf. Dieser neue Klon ist assoziiert mit schweren Meningokokkenerkrankungen und weltweiten Erkrankungsausbrüchen. Deshalb war es das Ziel des zweiten Projektes dieser Arbeit, genetische Unterschiede zwischen ET-15-Stämmen und anderen ST-11-Stämmen mit Hilfe von Gesamtgenom-*microarrays* zu identifizieren. Drei Gene, die hypothetische Proteine kodieren, waren in allen ET-15-Stämmen vorhanden, während sie in den anderen ST-11-Stämmen fehlten. Dieses Ergebnis, zusammen mit weiteren bisher nicht publizierten Beobachtungen aus unserer Arbeitsgruppe, deutet daraufhin, dass vor der klonalen Ausbreitung des ET-15-Klons mehrere Veränderungen im Genom auftraten. Die Bedeutung dieser drei Gene für die Pathogenität des ET-15-Klons ist unklar. Die Genomvergleiche zeigten weiterhin, dass Stämme des ET-15-Klons eine nur halb so große genetische Variabilität aufwiesen wie die ST-11-Stämme, die nicht zum ET-15-Klon gehörten. Dieses Ergebnis steht in Einklang mit der erst kürzlich erfolgten klonalen Expansion der ET-15-Meningokokken.

2. Abstract

In the present thesis, two projects on the use of microarray technology for molecular epidemiology of *Neisseria meningitidis* have been followed. The first one evaluated microarrays based on polymorphism-directed oligonucleotide design for typing of *N. meningitidis* adopting the multilocus sequence typing (MLST) concept. The number of oligonucleotides needed to cover all known polymorphisms was much lower compared to the number needed if a tiling strategy would have been chosen. Initial experiments using oligonucleotides 28-32 nucleotides in length, revealed that the applied hybridisation protocols were highly specific. However, despite of several optimisation steps, the rate of misidentification of oligonucleotides remained >1.8% in consecutive validation experiments using arrays representing the genetic diversity at three MLST loci. This finding led to the assumption that the high density of polymorphic sites and extensive GC-content variations at *N. meningitidis* MLST loci hindered the successful implementation of MLST microarrays based on polymorphism-directed oligonucleotide design.

In the 1980s, the ET-15 clone emerged within the ST-11 complex of *N. meningitidis*. This new clone was associated with severe meningococcal disease and outbreaks world-wide. Therefore, the goal of the second project was to identify genetic differences between ET-15 strains and other ST-11 strains using whole genome microarray technology. Three genes encoding hypothetical proteins were identified to be present in all ET-15 strains but absent in other ST-11 strains. This finding together with unpublished observation from our group suggested that several genome alterations occurred before the clonal expansion of the ET-15 clone started. The role that these three genes play in the pathogenicity of the ET-15 clone is unclear. The genome comparisons revealed furthermore that studies of the ET-15 clone displayed approximately two-fold less gene content variation than ST-11 strains not belonging to the ET-15 clone. This finding is in accordance with the recent emergence and clonal expansion of the ET-15 variant.

3. Introduction

3.1. *Neisseria meningitidis*

The genus *Neisseria* together with the genera *Moraxella* and *Kingella* are classified in the family *Neisseriaceae.* The genus *Neisseria* includes two species of pyogenic cocci that are pathogenic for man, and that have no other known reservoir: the meningococcus (*Neisseria meningitidis*) and the gonococcus (*Neisseria gonorrhoeae*) (reviewed in Davies *et al*., 1980). Other species of this genus: *N. lactamica, N. cinerea, N. flavescens* or *N. sicca* are considered non-pathogenic and are normal inhabitants of oro- and nasopharyngeal mucous membranes of humans.

Neisseria meningitidis is a gram-negative bacterial species with a typical diplo-coccoid appearance (0.6x0.8 μm). The bacterium was first isolated by Weichselbaum in 1887 from the spinal fluid of a patient with a purulent form of meningitis (reviewed in Davies *et al*., 1980). The species is aerobic and optimal growth conditions are achieved in a moist environment at 35-37°C under an atmosphere of 5-10% carbon dioxide. The bacteria are oxidase and catalase positive.

The organism produces a polysaccharide capsule, which is the basis of the serogroup typing system. Menincococci can be devided by seroagglutination into at least 13 serogroups: A, B, C, D, X, Y, Z, 29E, W-135, H, I, K and L (reviewed in Apicella, 1985). Among the different serogroups of *Neisseria meningitidis*, serogroup A, B and C account for 90% of the disease. In the Americas and Europe serogroup B and C are predominant agents causing systemic disease. Serogroup A meningococcus has been the main cause of epidemic meningococcal disease globally and still predominates in Africa and Asia (Tikhomirov *et al*., 1997).

Meningococci usually inhabit the human nasopharyngeal area without causing any symptoms. This carrier state may last for a few days to months; it provides the reservoir for the meningococcus, and it enhances the immunity of the host. The carrier rate tends to be considerable at all times in the general population (up to 10%), and in military populations it may achieve higher rate (reviewed in Gotschlich, 1980).

Because meningococci do not survive well outside their human host (Berger and Döring, 1969) and have no alternate host, transmission of meningococcal strains usually occurs by direct contact with contaminated respiratory secretions or air-borne droplets. After dissemination from the nasopharynx, meningococci may cause meningococcemia and/or meningitis. In Germany, meningococcal disease occurs with an endemic incidence of 0.94 cases per 100.000 inhabitants (Epidemiologisches Bulletin, 2004, Robert Koch Institute in Berlin). Disease occurs if the balance of the host's immune system and meningococcal

virulence factors is disturbed and bacteria enter the blood stream. Meningococcemia is usually characterized by profound vascular effects including a petechial or purpuric skin rash that occur in about 75% of patients. In fulminant infections (Waterhouse-Friderichsen syndrome), disseminated intravascular coagulation occurs, resulting in shock and, ultimately, in death. Meningococci also cause arthritis and pneumonia. Mortality of meningococcal meningitis without treatment is approximately 85% (reviewed in Gotschlich, 1980). The mortality in Germany is currently about 5-10% over all age groups (Epidemiologisches Bulletin, 2004, Robert Koch Institute, Berlin).

3.2. Population biology

3.2.1. Development of genetic diversity in *N. meningitidis*

Neisseria meningitidis is constitutively competent for transformation, more than for transduction or conjugation, and the transformation is believed to be the most important recombinational mechanism (Vazquez *et al*., 1995). Additionally, sequence changes are more often due to recombination after import than to mutation, both in housekeeping genes (Feil *et al*., 1999) and in highly polymorphic genes (Zhu *et al*., 2001).

Nucleotide sequence determination of multiple alleles of various genes has demonstrated widespread occurrence of mosaic genes, which is consistent with horizontal genetic exchange occurring regularly throughout the meningococcal genome (Bowler *et al*., 1994). This process involves the exchange of small segments of genetic material among organisms what it is probably a consequence of autolysis of meningococcal cells, which releases DNA that can be taken up by other cells and incorporated into their genomes by homologues recombination (Maiden and Frosch, 2001). Homologous recombination requires as little as 70% nucleotide sequence identity between the recipient and donor bacteria, although recombinational exchanges are far more likely to occur between closely related DNA sequences. A consequence of this is that bacteria can recruit variation from other members of the same species and, at lower frequency, from related species (reviewed in Spratt and Maiden, 1999). Morelli *et al*. (1997) described imports of *opa* (opacity) and *iga* (IgA1 protease) alleles during pandemic spread of subgroup III *N. meningitidis*. Recombinational imports of diverged sequences from other named *Neisseria* species have been noted in housekeeping genes, e.g. *aroE* and *glnA* (Zhou *et al*., 1997). Sequences of *tbpB* from other strains of *N. meningitidis* as well as *Neisseria lactamica* and commensal *Neisseria* spp. isolated from the same geographical area revealed that these species share a common *tbpB* gene pool and identified several examples of interspecific genetic exchange (Linz *et al*., 2000). These interspecies

recombinational exchanges can be a consequence of the coexistence in the same ecological niche (nasopharynx) of pathogenic and non-pathogenic *Neisseria* species. Therefore, different *Neisseria* species cannot be considered as independent from each other, but as a network where horizontal gene transfer takes place within a global *Neisseria spp.* gene pool (Maiden *et al.,* 1996). Recombination can also result in significant evolutionary changes over a short-term that was observed by the emergence horizontal distribution of some chromosomally encoded antibiotic resistance genes from different sources to create multi-antibiotic-resistant strains (reviewed in Feil and Spratt, 2001). Point mutations (single nucleotide changes) or recombinational imports (multiple nucleotide changes) lead to the variability of the *Neisseria* population. The new alleles, which arise in this way, will be confined in the lineage in which they arose.

3.2.2. Measurement of genetic diversity

Molecular typing techniques for *N. meningitidis* are used for epidemiological and scientific purposes. The choice of typing methods depends on the epidemiological questions to be answered and on the population genetics of the organism under investigation. If a thorough evaluation of the overall population is sought to determine the relationship between strains and clonal lineages, then multilocus approaches such as MLEE and MLST are necessary (reviewed in Vogel *et al.*, 2004).

Multilocus enzyme electrophoresis (MLEE) (Selander *et al.*, 1986) has been the reference method for global epidemiology and population structure of *Neisseria meningitidis* identifying the major disease-associated lineages. The allelic variants of 10 to 15 enzymes were inferred by the differential mobility during gel electrophoresis. Numbers were assigned to enzyme alleles on the basis of enzyme mobility, and each unique set of alleles was defined as an electrophoretic type (ET). The differential mobility of the variant proteins relied on charge differences in the allelic variants, so truly neutral mutations, those which generated synonymous changes, were not detected. Iso-enzymes with the same electrophoretic mobility might be encoded by very different gene sequences.

To overcome the problems associated with MLEE, multilocus sequence typing (MLST) was designed (Maiden 1998). MLST is based on the sequencing of 450-500 bp fragments belonging to seven housekeeping genes. For each locus, every unique sequence is defined as a different allele, and for each isolate, the alleles at the seven loci define an allelic profile or sequence type (ST). Examination of data of STs revealed that many STs are present at low frequency, whereas others are more prevalent. The prevalent STs are usually isolated over many years and diverse geographical locations. When data are analysed by the BURST

algorithm, these STs occupy a central position, in that they have numerous relatives that appear to be derived from them by a limited number of genetic events. These founders form the basis of clonal complexes to which they give their name. The isolates that share a minimum of four out of seven identical alleles with the central genotype are assigned to clonal complexes. A clonal complex comprises genetically related, but not identical, bacteria. The MLST approach built on the success of MLEE by targeting the variation present at multiple housekeeping loci, however, in the case of MLST definitive identification of this variation is achieved by nucleotide sequence determination of gene fragments (Urwin and Maiden, 2003). The data of MLST can be used to estimate the relative contributions of recombination and point mutation to clonal diversification. MLEE has long been the reference standard for subtyping *N. meningitidis*, but several features of MLST, such as electronic portability and the unambiguous nature of DNA sequence, are advantageous.

3.2.3 Clonality of *N. meningitidis*

In clonal species, genetic differences between individual isolates reflect changes that have accumulated sequentially since the descent from ancestral bacteria. Clonal populations can only exist if recombination and horizontal genetic exchange are rare. Consequently, clonal populations are characterized by linkage disequilibrium and relatively low diversity (Maiden and Frosch, 2001). van Belkum *et al*. (2001) define linkage disequilibrium as non-random reassortment of alleles occurring at different loci due to physical lineage. Recombination disrupts the clonal structure because the newly introduced DNA does not share the same evolutionary history as other genes within the genome. If recombination is frequent, phylogenetic trees will differ depending on the locus under analysis and alleles at different loci will be in linkage equilibrium (Suerbaum and Achtman, 2001). The extent of recombination varies among bacteria, so that some species have rather stable clones (*Salmonella enterica,* Boyd *et al.,* 1996), whereas in other species (*Helicobacter pylori,* Suerbaum *et al*., 1998) clones may be so transient that they can not readily be discerned.

MLEE indexed the allelic variation in multiple chromosomal genes and has been highly successful in generating large data sets for statistical analysis of *Neisseria* populations, and has shown that these bacteria exhibit strong linkage disequilibrium (the non-random association of alleles), with the frequent recovery of only a few of all the possible multilocus genotypes. This observation suggested that the population structure of *N. meningitidis* is clonal (Smith *et al*., 1993). However, if a particular clone within a recombining species expand and spread very rapidly, they may transiently posses a clonal population structure. Wang *et al*., (1992), on the basis of MLEE, interpreted the serogroup A meningococci as truly

clonal and possessing only limited genetic variability. This strong linkage disequilibrium was observed between the alleles at different loci in populations of meningococci recovered from cases of disease. Studies of isolates form asymptomatic nasopharyngeal carriage done in Norway (Caugant *et al.*, 1988), the Czech Republic (Jolley *et al.*, 2000) and Germany (Claus *et al.,* 2004) revealed higher diversity of isolates than those recovered from cases of disease. Isolates of two of the known hypervirulent meningococcal clonal complexes (ET-5 and ET-37) comprised 91% of the population from cases of invasive disease in Norway (Caugant *et al.*, 1988), but were only found at low abundance among the carried population. The clonal complex that was most frequently represented by isolates from carriers (19%) has never been recovered from patients with meningococcal disease in Norway or elsewhere, which suggests that these isolates have a low virulence potential. This clonal complex probably represents the emergence and diversification of an ecologically successful clone that is highly transmissible to new hosts (reviewed in Feil and Spratt, 2001).

Smith *et al.*, (1993) suggested that the meningococcal population is composed of two parts: a basic net in which there is frequent recombination within all members of the population and occasionally highly successful clones. These highly successful individuals arise and increase rapidly in frequency to produce an epidemic lineage. Variant clones are regularly purified from the population by sequential bottlenecks, which occur when only a few bacteria spread between geographical areas (Achtman, 1995).

3.3. Clonal lineages of *N. meningitidis*

The species *N. meningitidis* comprises a large variety of genetically different clones, which are defined by multilocus enzyme electrophoresis or by multilocus sequence typing of housekeeping genes. Only a few genetic groupings of meningococci have been associated with epidemic disease. The ST-32 complex has been associated with a global hyper-endemic epidemiology from the mid 1970s through to the 1990s. The lineage is mainly associated with serogroup B, although serogroup C strains are also found (Caugant *et al.*, 1988). Since the mid 1990s, cases of disease caused by ST-32 complex strains have declined, replaced largely by strains belonging to the ST-41/44 complex, showing a similar epidemiology and serogroup distribution (Jolley *et al.*, 2000). The ST-11 complex has been responsible for outbreaks, often causing severe septicaemia and forms of meningococcal disease. Isolates belonging to this complex are also found to cause large numbers of sporadic cases. Most strains belonging to the ST-11 complex are serogroup C (Jelfs *et al.*, 2000). The ST-8 complex shares a similar epidemiology with the ST-11 complex but the ST-8 complex shows a more equal distribution of serogroup B and C strains (Jolley *et al.*, 2000).

3.3.1. Structure and characteristics of the ST-11 complex

The clonal lineage ST-11 (ET-37) is one of the several hyperinvasive lineages. Most strains belonging to the ST-11 complex are serogroup C, but serogroups B, Y and W-135 are also found. The oldest ST-11 isolate deposited at the *Neisseria* MLST Isolate Database (http://pubmlst.org/perl/mlstdbnet/mlstdbnet.pl?file=pub-nm_isolates.xml) was retrieved in 1964, at present (August 2004) the ST-11 complex consists of 159 STs.

It has been demonstrated that the hypervirulent complex ST-11 can be distinguished by the presence or absence of several genomic loci. The genomic loci 37/1-7 (putative R-M system) and 37/1-26 (putative prophage) were present in all tested ST-11 complex strains and more than 90% of the ST-8 complex strains (Claus *et al.,* 2000). The identification of a differentially distributed restriction-modification system enabled the generation of monoclonal antibodies available for typing (Claus *et al*., 2003).

A new hypervirulent variant within the ST-11 complex appeared first time in Canada in 1986. Physical mapping of genes in strains of the ST-11 complex demonstrated that large genetic rearrangements of the genome have occurred in association with the appearance of this new variant (Ashton *et al*., 1991). MLEE revealed that the strains responsible for these outbreaks belonged to the ET-37 complex, but were characterized by a rarely occurring allele at the fumarase locus. The fumarate hydratase mutation elucidated by MLEE of ET-15 strains was due to a single non-synonymous mutation of the *fumC* at the position 640 (Vogel *et al*., 2000), not identified by the classical MLST scheme of *Neisseria.* This point mutation permits the destination of ET-15 from other ST-11 complex strains.

3.3.2. Epidemiology of the ST-11 complex

The ET-15 clone has been the most frequently isolated meningococcal genotype in Canada during next years from the appearance in 1986. The percentage of invasive meningococcal disease increased from 2.0% to 51.8% (Kertesz *et al*., 1997). ET-15 was also subsequently associated with meningococcal disease in parts of the United States, Israel, the Czech Republic, Iceland, Finland, Norway, England, Germany and Australia. Epidemiologically, ET-15 was associated with a significant increase in invasive meningococcal disease incidence among persons 15-19 years old in contrast to the usual pattern of higher rates of meningococcal disease in age groups under 4 years. Morbidity and mortality were greater than they have been previously seen with other strains of the ET-37 complex (Ashton *et al*., 1991).

3.4. Microarray technology

As a "Father of Microarray Technology" Mark Schena is recognised who trained with Dan Koshland at U.C.Berkeley, Keith Yamamoto at UCSF, and Ron Davis at Stanford University. Schena published the first paper on microarrays in 1995 entitled "Quantitative monitoring of gene expression patterns with a complementary DNA microarray" (Schena *et al*., 1995).

The DNA microarray (or microchip) technology is having a significant impact on genomics study. It is a powerful tool for studying gene expression and regulation on a genomic scale (Khodursky *et al*., 2000) and detecting genetic polymorphisms (Hacia, 1999) in both eukaryotes and prokaryotes. Arrays have been used in such applications as drug discovery (Chicurel *et al*., 2002), evolutionary studies (Hacia *et al*., 1999) and genome mapping (Shalon *et al*., 1996). The application of DNA microarray technology promises to be extremely powerful as a diagnostic tool, since microarrays theoretically permit a simultaneous screen for any tens of thousands of nucleic acid sequences (reviewed in Li *et al*., 2000).

Microarray analysis differs from traditional research in a number of many ways, one of which is the relationship between the amount of experimental time required and the amount of data obtained. Traditional experimental approaches based on gels and filter blots required a relatively large amount of experimental time to obtain a small volume of data, whereas microarray analysis affords vast quantities of data with relatively little experimental time. Small spots enable high density (>5000 elements/cm²), rapid reaction kinetics and the analysis of entire genomes on a single chip. Experiments that examine all of the genes in a given genome provide a comprehensive insight of a biological phenomenon that is not possible with technologies limited to gene subsets (Schena, 2002). Compared to conventional membrane-based hybridisation, glass slide-based microarrays offer the additional advantage of lower cost, automation and low background levels (Shalon *et al*., 1996).

3.4.1. Technical aspects of microarray technology

Microarray probe material can be derived from whole genes or parts of genes, and may include genomic DNA, cDNA, mRNA, small molecules, tissues, or any other type of molecule that allows quantitative gene analysis. Probe molecules include natural and synthetic derivatives obtained from a variety of sources, such as cells, enzymatic reactions, and machines that carry out chemical synthesis. Also synthetic oligonucleotides, short single-stranded molecules made by chemical synthesis, provide an excellent source of probe material (Schena, 2002). According to the nomenclature recommended by *Nature Genetics* (January 1999 Supplement), a "probe" is the tethered nucleic acid with known sequence, whereas a "target" is the free nucleic acid sample, whose identity/abundance is being detected. DNA

arrays are fabricated by high-speed robotics on a nylon substrates or a solid support, such as a coated glass surface, for which labelled targets are used to determine complementary binding allowing massively parallel gene expression and gene discovery studies. The generic term "DNA chips" refers to miniaturised arrays of nucleic acid segments anchored on glass supports no larger than a microscope slide (Hacia and Collins, 1999).

The use intention of an array dictates its design. The major types of DNA microarrays currently in use can be distinguished by the length of their probes. Larger segments, usually greater than 100 bp are commonly created using PCR products or cDNA clones, and are referred to as cDNA arrays (Duggan *et al.*, 1999). The use of PCR products for hybridisation-based studies has a number of advantages, a main one being that no gene sequence information is required before microarray analysis. Another advantage of PCR products is that their relatively large size (~500 base pairs) provides extensive complementarity for hybridisation, generating intense fluorescent signals in virtually every experiment. On the other hand, PCR-generated probes can provide unwanted hybridisation or cross-hybridisation when studying a family of genes that share considerable sequence identity. Cross-hybridisation occurs when a given probe shares significant sequence identity (>70%) with more than one species in the target mixture, leading to a loss of gene specificity in microarray experiment (Schena, 2002).

Microarrays that possess shorter DNA sequences are termed oligonucleotide microarrays. Oligonucleotide microarrays are fabricated either by in situ light-directed combinational synthesis or by conventional synthesis followed by immobilization on glass substrates. Using photolithography (a light-directed synthesis technique used in semiconductor industry), it is possible to synthesize millions of different oligonucleotides at specific locations on a glass substrate, in order to create high-density oligonucleotide microarrays (Lipshutz *et al.*, 1999). Large collections of oligonucleotides provide a rapid means of genotyping, by hybridising targets to the oligonucleotide arrays and reading the fluorescence intensities at each location on the chip. The GeneChip technology (Affymetrix, Santa Clara, CA) provides an extremely powerful platform for examining small variations in DNA sequence, including mutations and single nucleotide polymorphisms (SNP) applying the tiling strategy (Chee *et al.*, 1996). Each position in the target sequence is queried by a set of four probes on the chip, identical except at a single position exactly in the middle. The central position of each set of four oligonucleotides contains one of the four possible bases (A, G, C, T) in the genetic code. By comparing the fluorescence intensity at the four oligonucleotides on the chip, the presence of a SNP can be detected with extremely high accuracy. The massively parallel format for SNP

analysis can provide detailed information about all point mutations within bacteria. Although, oligonucleotide-chip-based mutation analysis is limited by a lack of sensitivity for mutations in regions with high local A/T or C/G content (reviewed in Lucchini *et al.*, 2001), this method as a diagnostic tool also shows great promise. Targets can be derived from many different biological and chemical sources, although nucleic acid targets (PCR products, chromosomal DNA, RNA) are the most common at present. The quality of the target solution is a key determinant of the quality of the microarray data. Fluorescence is the dominant target label. The target labelling can be done direct or indirect.

Hybridisation reactions between single-stranded probe and target molecules occur by hydrogen bond formation between the bases of complementary nucleic acids sequences. Sequence composition, target and probe length, hybridisation temperature, secondary structure, degree of homology, salt concentration, pH, and a number of other factors influence hybridisation efficiency and the strength of the duplex. The specificity and affinity of target-probe interactions determine the usefulness of microarray assays.

3.4.2. Microarray application in bacteriology

3.4.2.1. Transcriptome analyses

Microarrays containing probe elements for all of the genes in a genome reveal comprehensive genetic and biochemical relationships that are impossible to obtain by any other means. This method has rapidly become a central platform for functional genomics enabling to perform high-quality experiments, which have improved our understanding of microbial environmental responses and global gene expression.

Organisms express their genes at a relatively constant rate until the products encoded by those genes are needed for a specific function. When needed, genes are activated or repressed rapidly, changing by 10-, 100- or even 1000-fold or more, depending on the particular gene and the strength of the regulatory signal. The expression of genes changes in response to a wide spectrum of signals, including chemicals, nutrients, stress, changes in cell division and development providing a gene expression "fingerprint" that is characteristic for a given physiological state. Because gene expression correlates specifically and tightly with function, it is possible to infer the function of genes and the interaction of pathways by documenting which genes are turned up or down in a given physiological state (Schena, 2002).

The first global transcriptional profile was obtained at the resolution of individual genes for *Saccharomyces cerevisiae* (Spellman, 1998). The first completion of the microbial genomes sequencing facilitated the applications of microarrays involving the whole genomes:

Mycobacterium tuberculosis (Behr *et al*., 1999) and *Escherichia coli* (Richmond *et al*., 1999). From that time the explosive growth in numbers of reviews discussing microarray technology is followed by many papers describing results obtained from gene expression microarray profiling. Transcriptome analysis and the pathogenicity mechanisms of *N. meningitidis* during infection are also already done employing whole genome DNA microarrays. Dietrich *et al.* (2003) identified 72 differentially expressed genes as important candidates for explanation of the pathogenicity mechanisms and the interaction of meningococci with their human hosts. Transcriptome study on *N. meningitidis* serogroup B during different stages of infection (exposure to human serum, interaction with human epithelial and endothelial cells) revealed membrane and surface proteins as vaccine candidates against serogroup B (Kurz *et al.,* 2003).

3.4.2.2. Genome comparisons

Genetic variations within genomes are consequences of bacterial adaptation to environment conditions. Each difference may be important for an organism's specific life style and virulence potential (Zhang *et al*., 2004).

Genotypic characteristics became the essential classification of bacteria replacing classification based on phenotypic markers. Genotypic information exceeds the categorisation achievable with phenotypic markers, giving deeper insights into evolutionary relationships between species that are indistinguishable on the basis of phenotypic traits. Additionally, much of the interesting genome information that determines phenotypic traits, such as degree of virulence, is encoded in small differences between strains and serovars, rather than species (Kim *et a*l., 2002). Phylogenetic classification based on rRNA/rDNA gene detection providing an accurate classification of microorganisms to the species level was already described (Anthony *et al*., 2000; Small *et al*., 2001; Loy *et al*., 2002). However, lateral gene transfer is an important mechanism of evolution for prokaryotes, which complicate phylogenetic analysis based on a small number of genes. Therefore, whole-genome based methods are required to determine the repertoire of the genes found or not found in bacteria. Strain comparison by hybridising genomic DNA to microarrays (genomotyping) is a more realistic approach than the whole-genome sequencing of dozens of strains (Lucchini *et al*., 2001). Gene-specific microarrays have been used to assess the genomic diversity of the species: *Mycobacterium bovis* (Behr *et al*., 1999), *Helicobacter pylori* (Salama *et al.*, 2000), *Campylobacter jejuni* (Dorrell *et al*., 2001), *Salmonalla enterica* and *Salmonella bongori* (Chan *et al*., 2002), *Listeria monocytogenes* (Doumith *et al*., (2003), *Escherichia coli* (Anjum *et al*., (2003), *Neisseria gonorrhoeae* (Snyder *et al*., 2004), *Porphyromonas gingivalis* (Chen *et al*., 2004). Hybridising DNA microarrays or "mixed genome" arrays with genomic DNA

from closely related species facilitates analysis of many genes (or complete microbial genomes) and permits comparisons of organisms in a functional genomic framework. The ability to see the relationships between closely related microbes by using data visualization tools such as cluster analysis provides an unmatched view of how different microorganisms are affiliated with each other (Murray *et al.*, 2001).

3.4.2.3. Molecular strain typing by microarrays

The capacity to distinguish single base differences between targets and probes forms the basis of genotyping applications of microarrays. Mutant alleles or single nucleotide polymorphisms (SNPs) can be detected with high-density oligonucleotide microarrays applying tiling strategy, which carry shorter probes (typically about 25 nt on Affymetrix gene chips). A single nucleotide difference between target and probe can be sufficient to prevent hybridisation, and has been used to identify 3000 polymorphisms between two strains of *Saccharomyces cerevisiae*. These polymorphisms were used as markers to map a multidrug-resistance locus and other four loci with high resolution within 11-64 kb (Winzeler *et al.*, 1998). Furthermore, MLST has recently been successfully adapted to the tiling strategy, using *Staphylococcus aureus* as a model (van Leeuwen *et al.*, 2003). The DNA array identified sequence variations in seven MLST loci with high correct score. Additionally, the results obtained were reproducible and epidemiologically concordant with the results from conventional MLST. The first example of an array-based database generated for the purpose of correlating genotypic changes with species and isolate identification as well as drug resistance phenotypes was reported by Gingeras *et al.*, (1998). The database contained the array-generated sequences and mutations in the *rpoB* gene, associated with rifampicin resistance in clinical isolates of *M. tuberculosis*. It was possible to carry out the species-specific identification and detection of sequence variation correlated with rifampicin-resistance in the same hybridisation experiments.

The oligonucleotide microarray for identification of SNP offers an attractive option for the identification and epidemiologic monitoring of TEM beta-lactamases in the routine clinic diagnostic laboratory. The array for the genotyping of TEM beta-lactamases within 3.5 h has been shown to be a reproducible and sensitive means for identification of various gene variants of ESBL-type beta-lactamases, which cannot be identified by simple cultivation-based assays (Grimm *et al.*, 2004).

The use of high-density oligonucleotide arrays enables to identify a large variety of bacterial species. Anthony *et al.*, (2000) were able to distinguish correctly 125 positive blood cultures and other clinical specimens by designing an array based on 23S rDNA oligonucleotides. 26

different species of *Mycobacterium* were distinguished by designing an array that evaluates 82 unique 16S rRNA sequences (Troesch *et al.*, 1999). The use of an array based on 16S rRNA oligonucleotides were also utilized to detect different species of bacteria from soil extracts (Small *et al.*, 2001) and all recognized lineages of sulphate-reducing Procaryotes in the environment (Loy *et al.*, 2002).

Microarrays can also be used to find genes involved in certain biological processes, such as pathogenicity. Unambiguous genotyping of *E.coli* O157:H7 based on sequence variations of four virulence loci (intimin, Shiga-like toxinI and II, and hemolysin A) revealed microarray technology as a meaningful tool for rapid detection and characterisation of food borne pathogens (Call *et al.*, 2000).

Objectives of the work

One goal of the project was to evaluate the use of microarrays for meningococcal sequence typing adopting the MLST concept. The second aim was to define genetic differences between the ET-15 clone and other members of the ST-11 complex by gene content comparisons using whole genome microarrays.

4. Materials and methods

4.1. Laboratory tools and equipment

Affymetrix 418 TM Scanner (MWG-Biotech AG)

Spotter OmniGrid 100 (GeneMachines, Huntingdon, UK)

Temperature incubators
  BB 6220 CU (Heraeus, Hanau)
  BB 6200 (Heraeus, Hanau)

Thermoblock
  2099/OA (Liebisch, Bielefeld)

Shaking incubator
  Certomat H (Braun, Melsungen)

FastPrep® FP120A Instrument (Qbiogene, Heidelberg)

Speed Vac® Plus SC110A (Savant, New York, USA)

Spectrophotometers
  GeneQuant *pro* (Amersham Pharmacia Biotech, Uppsala, Sweden)
  U-2000 (Hitachi, Schwäbisch Gmünd)

Thermocyclers
  TRIO (Biometra, Göttingen)
  Personal Cycler (Biometra, Göttingen)

Vortex Mixer
  VM-300 (neo Lab Migge, Heidelberg)

Centrifuges
  Biofuge 15R (Heraeus, Hanau)
  RC-5B Plus (Sorvall Heraeus, Hanau)

Megafuge 1.0R (Heraeus, Hanau)

Mikro Rapid/K 1306 (Hettich, Tuttlingen)

Electrophoresis chamber

H6-SET (BioRad, München)

Magnetic stirrer (A.Hartenstein, Würzburg)

Hybridisation chambers (Corning, NY, USA)

Glass utensils for slide washing (A.Hartenstein, Würzburg)

4.2. Consumables

15-ml and 50-ml centrifuge BD Falcon™ tubes (Becton Dickinson, Heidelberg)

12-ml, 14 ml-sterile plastic tubes (Greiner Labortechnik, Essen)

0,5-ml, 1,5-ml and 2-ml reaction tubes (Sarstedt, Nümbrecht)

Cover slips: 24x40, 24x50, 24x60 (A.Hartenstein, Würzburg)

AutoSeq™ G-50 columns (Amersham Biosciences, Freiburg)

4.3. Chemicals, kits, enzymes

from Amersham Biosciences, Freiburg

CyDye 3-dCTP

CyDye 5-dCTP

λ Hind III marker

from Roche, Mannheim

Klenow enzyme

RNase, DNase-free

DNase I, RNase-free

Ampli*Taq* DNA polymerase

Random hexanucleotide primers

Lysozym

from Schott Nexterion AG, Mainz

Nexterion™ Slides E

2xNexterion Spotting Solution

from Sigma, München

Nona Random Primer

Ethanolamine

Salmon Sperm for Hybridisation

Parafin mineral oil

from Eurogentec, Seraing, Belgium

PCR products spotted onto aldehyde glass slides

Smart-Ladder

Hybridisation solution

from bioMerieux, Marcy I'Etoile, France

PolyVitex (Vitamins for fluid medium)

from Qiagen, Hilden

QIAquick PCR Purification Kit

MinElute PCR Purification Kit

Genomic-tip 100/G

RNeasy Midi KIT (50)

from Invitrogen, Carlsbad, CA, USA

RNase OUT™ Ribonuclease Inhibitor (recombinant)

SuperScript™ III RNase H⁻ Reverse Transcriptase

1 Kb DNA Ladder

dATP

dCTP

dGTP

dTTP

DTT (dithiothreitol)

5xFirst Strand Buffer

<u>from Aldrich Chemical Company, Inc., Milwaukee, Wi, USA</u>

Betaine

<u>from AppliChem GmbH, Darmstadt</u>

Formamide deionised

EDTA (disodium salt dihydrate)

$MgSO_4x7H_2O$ (magnesium sulfate heptahydrate)

NaCl (sodium chloride)

$C_6H_5Na_3O_7$ $2H_2O$ (*tri*-sodium citrate dihydrate)

<u>from New England BioLabs, Frankfurt</u>

BSA 100x

<u>from Roth GmbH, Karlsruhe</u>

Tris

Glycerol

dNTPs, lyophilised

<u>from Merck, Schwalbach</u>

2-Mercaptoethanol

$MgCl_2$ (magnesium chloride)

KCl (potassium chloride)

25% HCl (hydrochloric acid)

<u>from Riedel-de Haen, Seelze</u>

Ethanol absolute

Triton® X-100

4.4. Buffers and solutions

<u>GEBS</u>

20 % Glycerol

50 mM EDTA

0.05 % Bromphenol blue

0.5 % N-Lauroylsarcosine

<u>PBS</u>

10 mM Na-phosphate buffer, pH 7.4

140 mM NaCl

<u>1 x TBE</u>

100 mM Tris-HCl

100 mM Boric acid

2.5 mM EDTA

pH 8.3

<u>1 x TAE</u>

40 mM Tris-HCl

1 mM EDTA

pH 8.1 (with CH_3COOH)

<u>TE (10/1)</u>

10 mM Tris-HCl, pH 8

1 mM EDTA

pH 8.0

<u>20 x SSC</u>

3 M NaCl

0,3 M Na-citrate

pH 7

<u>10 x PCR-Buffer</u>

      500 mM KCl

      100 mM Tris-HCl, pH 8.3

      25 mM $MgCl_2$

<u>10 x dNTP-Mixture for Cy-labelling</u>

      1 mM dATP

      0.5 mM dCTP

      1 mM dGTP

      1 mM dTTP

      10 mM Tris

      1 mM EDTA, pH 8

<u>Nonanucleotide mixture</u>

      5µg/µl nonanucleotides

      1 M Tris

      1 M $MgCl_2$

      100 mM DTT

      10% BSA

<u>Reaction mixture for reverse transcription</u>

      0.2 mM dCTP

      0.5 mM dATP

      0.5 mM dGTP

      0.5 mM dTTP

      1 mM DTT

      1xFirst Strand buffer

      water

1 U RNase OUT™ Ribonuclease Inhibitor (Recombinant)

200 U SuperScript™ III RNase H‾Reverse Transcriptase

1.25 nM Cy3 or/and Cy5

The Qiagen Genomic-tip 100/G Kit for DNA isolation:

Buffer B1

> 50 mM Tris Cl, pH 8.0
>
> 50 mM EDTA
>
> 0.5% Tween®-20
>
> 0.5% Triton®X-100

Buffer B2

> 3 M Guanidine HCl
>
> 20% Tween-20

4.5. Strain collection and growth conditions

4.5.1. *Neisseria spp.*

Table 1 shows the strains used in this study.

Strain	Serological characterization	MLST[a]	Country	Year	MLEE[b]	fumC-640[c]
MC58	B:15:P1.7,16	74	UK	1985	ET-5	
Z2491	A:4,21:P1.7b,13a	4	Gambia	1983	subgroup IV-1	
FAM18		11	USA	1983	ET-37	
171	P1.10	11	USA		not available	G
172	P1.2,5	166	USA		not available	G
531			Norway		ET-37	G
532			Norway		ET-37	G
533			Norway		ET-37	G
534			Norway		ET-37	G
535			Norway		ET-37	G
536			Norway		ET-37	G
537			Norway		ET-37	G
538			Norway		ET-37	G
539			Norway		ET-37	G
540			Norway	1986	ET-37	G
1110			Argentina	1994	ET-37	G
1111			Argentina	1994	ET-37	G
1112			Australia	1994	ET-37	A
1128			Austria	1990	ET-37	G
1133			Norway	1995	ET-37	A
1137		11	Austria	1991	ET-15	A
2018			Italy	1984	ET-37	G
2120	P1.2/P1.5	11	Germany	1997	ET-37	G
2121	P1.2/P1.5	11	Germany	1997	not available	G
2140	P1.2,5	11	UK	1997	not available	G
2141	P1.2,5	11	UK	1996	not available	G
2209			Germany	1998	ET-15	A
2210			Germany	1998	ET-15	A
2211			Germany	1998	ET-15	A
2212			Germany	1998	ET-15	A

Strain	Serological characterization	MLST[a]	Country	Year	MLEE[b]	fumC-640[c]
2213			Germany	1998	ET-15	A
2214			Germany	1998	ET-15	A
2225	P1.2,5	11	Germany	1994	not available	G
2240	P1.2,5		Canada	1997	ET-4	G
2241	P1.5,2		Czech Republic	1993	ET-15	A
2246			Germany		ET-15	A
2359	P1.2		Czech Republic	1993	ET-15	A
2360	P1.2/P1.5	11	Czech Republic	1993	ET-15	A
2361	P1.2,5		Czech Republic	1993	ET-15	A
2362	P1.2		Czech Republic	1993	ET-15	A
2363	P1.2(5)		Czech Republic	1994	ET-15	A
2364	P1.2(5)		Czech Republic	1994	ET-15	A
2365	P1.2(5)		Czech Republic	1994	ET-37	G
2366	P1.2		Czech Republic	1995	ET-15	A
2367	P1.2,5		Czech Republic	1995	ET-15	A
2368	P1.2,5	11	Czech Republic	1995	ET-15	A
2369	P1.2		Czech Republic	1995	ET-15	A
2370	P1.2,5		Czech Republic	1995	ET-15	A
2371	P1.2		Czech Republic	1995	ET-37	G
2372	P1.2,5		Czech Republic	1995	ET-15	A
2373	P1.2		Czech Republic	1996	ET-15	A
2374	P1.2		Czech Republic	1996	ET-15	A
2375	P1.2,5		Czech Republic	1996	ET-15	G
2376	P1.2,5		Czech Republic	1997	ET-15	A
2377	P1.2,5		Czech Republic	1997	ET-15	A
2378	P1.2,5		Czech Republic	1998	ET-15	A
2379	P1.2,5	11	Czech Republic	1998	ET-15	A
2524			Mali	1998	ET-37	G
2525	(P1.2)	11	Mali	1995	ET-37	G
2698			Germany		ET-37	G
2699			Germany		ET-15	A
2700			Germany		ET-37	G
2977	P1.5,2	11	Canada	1988	ET-15	A
2978	P1.5		Australia	1996	ET-15	A
2979	P1.5,2	11	Australia	1994	ET-15	A
2980	P1.5		Australia	1996	ET-15	A
2981	NST	11	Iceland	1991	ET-15	A
2982			Iceland	1993	ET-15	A
2983	P1.2	11	Finland	1994	ET-15	A
2984	P1.5		Finland	1994	ET-15	A
2985	P1.5		Finland	1995	ET-15	A
2986	P1.5,2		UK	1995	ET-15	A
2987	P1.5		UK	1993	ET-15	A
2988	P1.5	11	Norway	1994	ET-15	A
2989	P1.5,2		Greece	1989	ET-15	A
2990	P1.2	11	Israel	1993	ET-15	A
2991	P1.5,2		Israel	1993	ET-15	A
2992	P1.5,2		Israel	1993	ET-15	A
2993	P1.5,2	11	Netherlands	1990	ET-15	A
DE9246		11	Germany	2003	ET-15	A
DE9260		11	Germany	2003	ET-15	A
94N266	P1.5		Australia	1994	ET-15	A
86/94	P1.5,2		Norway	1994	ET-15	A

Strain	Serological characterization	MLST[a]	Country	Year	MLEE[b]	fumC-640[c]
343/93	P1.2		Czech Republic	1993	ET-15	A
67			Finland	1994	ET-15	A
97N005	P1.5		Australia	1997	ET-15	A
91297			Canada	1991	ET-15	A
95N477	P1.2		Australia	1995	ET-15	A
89486			Canada	1989	ET-15	A
393	P1.5,2		Austria	1991	ET-37	G
4575	P1.5,2		Australia	1989	ET-37	G
93-01	P1.5,2		Australia	1993	ET-37	G
93-45	P1.5,2		Australia	1993	ET-37	G
M710	P1.5		Israel	1991	ET-37	G
24580	P1.5,2		South Africa	1996	ET-37	G
PB8	P1.2		Norway	1986	ET-37	G
8732	P1.10		Chile	1985	ET-37	G
53/94	P1.5,2		Norway	1993	ET-37	G
88/19730	P1.2		Scotland	1988	ET-37	G
88048			Canada	1988	ET-37	G
901256	P1.2		Netherlands	1990	ET-37	A
α4		19	Germany	1999	not available	
α19		44	Germany	1999	not available	
α51		820	Germany	1999	not available	
α52		965	Germany	1999	not available	
α55		39	Germany	1999	not available	
α66		914	Germany	1999	not available	
α110		1139	Germany	1999	not available	
α122		11	Germany	1999	not available	G
α124		32	Germany	1999	not available	
α125		1136	Germany	1999	not available	
α126		808	Germany	1999	not available	
α141		23	Germany	1999	not available	
α147		35	Germany	1999	not available	
α162		23	Germany	1999	not available	
α184		939	Germany	1999	not available	
α188		784	Germany	1999	not available	
α198		8	Germany	1999	not available	
α223		11	Germany	2000	not available	G
α259		768	Germany	2000	not available	
α270		191	Germany	2000	not available	
α272		92	Germany	2000	not available	
α273		178	Germany	2000	not available	
α290		53	Germany	2000	not available	
α292		444	Germany	2000	not available	
α306		11	Germany	2000	not available	G
α308		11	Germany	2000	not available	A
α320		198	Germany	2000	not available	
α324		11	Germany	2000	not available	G
α349		890	Germany	2000	not available	
α359		162	Germany	2000	not available	
α360		162	Germany	2000	not available	
α364		1117	Germany	2000	not available	
α449		24	Germany	2000	not available	
α471		44	Germany	2000	not available	
α492		60	Germany	2000	not available	

Strain	Serological characterization	MLST[a]	Country	Year	MLEE[b]	fumC-640[c]
α516		857	Germany	2000	not available	
α518		860	Germany	2000	not available	
α578		198	Germany	2000	not available	
α579		337	Germany	2000	not available	
α587		53	Germany	2000	not available	
α630		235	Germany	2000	not available	
α637		919	Germany	2000	not available	
α641		875	Germany	2000	not available	
α656		136	Germany	2000	not available	
α657		939	Germany	2000	not available	
α669		802	Germany	2000	not available	
α675		908	Germany	2000	not available	
α734		951	Germany	2000	not available	
α737		11	Germany	2000	not available	A
α738		11	Germany	2000	not available	A
α749		11	Germany	2000	not available	A
α756		964	Germany	2000	not available	
α758		839	Germany	2000	not available	
α800		813	Germany	2000	not available	
α820		817	Germany	2000	not available	
α822		957	Germany	2000	not available	
α828		44	Germany	2000	not available	
β147[1]		1611[2]	Germany	2000	not available	
β271[1]		1615[2]	Germany	2000	not available	

[a] MLST-multilocus sequence typing

[b] MLEE-multilocus enzyme electrophoresis, data from D. Caugant, Oslo

[c] ET-15, A; ST-11/non ET-15, G (Vogel *et al.*, 2000)

[1] *N.lactamica*

[2] Information obtained from MLST website

4.5.2. Growth media

<u>GC-Agar</u>

      1.5 % Proteose-Peptone

      0.1 % starch

      0.4 % K-phosphate

      0.5 % NaCl

      1.0 % Bacto-Agar

      pH 7.2

<u>Proteose-Pepton-Medium (PPM)</u>

      1.5 % Proteose-Peptone

      0.5 % NaCl

      0.05 % starch

      0.4 % KH_2PO_4

      0.1 % K_2HPO_4

      pH 7.8 (with KOH)

<u>PPM, supplemented</u>

PPM-Medium + 0.01 M $MgCl_2$

           + 0.042 % $NaHCO_3$

           + 1 % (v/v) PolyVitex

4.5.3. Growth conditions

All *Neisseria* strains were grown on GC agar at 37°C under 5% CO_2.

In supplemented PPM bacteria were incubated at 37°C with shaking.

4.6. Amplification of MLST loci and ET-15 specific genes

4.6.1. Primers used for PCR

abcZ
abcZ-P1   5'-AATCGTTTATGTACCGCAGG-3'

abcZ-P2   5'-GTTGATTTCTGCCTGTTCGG-3'

adk
adk-P1   5'-ATGGCAGTTTGTGCAGTTGG-3'

adk-P2   5'-GATTTAAACAGCGATTGCCC-3'

aroE
aroE-P1   5'-ACGCATTTGCGCCGACATC-3'

aroE-P2   5'-ATCAGGGCTTTTTTCAGGTT-3'

abcZ
abcZ-P1   5'-AATCGTTTATGTACCGCAGG-3'

abcZ-P2   5'-GTTGATTTCTGCCTGTTCGG-3'

adk
adk-P1   5'-ATGGCAGTTTGTGCAGTTGG-3'

adk-P2   5'-GATTTAAACAGCGATTGCCC-3'

aroE
aroE-P1   5'-ACGCATTTGCGCCGACATC-3'

aroE-P2   5'-ATCAGGGCTTTTTTCAGGTT-3'

NMA1373
NMA 1373 F   5'-GGACTCAACAAGAAAATTTGGG-3'

NMA 1373 R   5'-CAACTGTACCTTTCTTATACCTG-3'

NMA1622
NMA 1622 F   5'-GGTATATGTGGATCTATTCATTG-3'

NMA 1622 R   5'-CAAATACTCTCTCTTAACTTATTG-3'

NMA0522
NMA 0522 F   5'-CGTTCAACCTTTTAGACGCAC-3'

NMA 0522 R   5'-GCAAACCGCCAAGTTCAGTT-3'

4.6.2. PCR with AmpliTaq®DNA-Polymerase

PCR was done according to the following scheme:

   6 µl 25 mM $MgCl_2$ (final concentration 150 µM)

   10 µl 10 x PCR-Buffer

   10 µl 2 mM dNTP-mixture (dATP, dCTP, dGTP and dTTP each in final concentration of
   200 µM)

   1 µl 20 µM primer A (final concentration 0.2µM)

   1 µl 20 µM primer B (final concentration 0.2µM)

   0.4 µl (2U) *Taq*®DNA-Polymerase

   10-20 ng DNA was used as a template

   ddH_2O ad 100 µl

PCR reaction conditions:

1. denaturation	300 - 600 sec at 94 °C	
2. hybridisation	60 sec at annealing temperature	
3. polymerase reaction	x sec at 72 °C	35 x
4. denaturation	60 sec at 94 °C	
5. hybridisation	60 sec at annealing temperature	
6. polymerase reaction	600 sec with 72 °C	

4.7. Chromosomal DNA isolation from bacteria

Overnight grown bacteria harvested from 2 GC-agar plates were resuspended in 5 ml 1xPBS and the optical density was determined at OD_{600}. For *N. meningitidis* $OD_{600} =1$ corresponds to $1x10^9$ cells. 1.5×10^{10} bacteria were used for DNA isolation. The bacterial suspension was centrifuged for 10 min at 4000 rpm. Chromosomal DNA was isolated with the Genomic-tip 100/G Kit (Qiagen) according to the manufacturer's instructions. The DNA pellet was dried at 37°C and resuspended in 100 µl TE. The amount and the quality of the DNA were estimated by gel electrophoresis and optical density at 260 nm.

4.8. RNA isolation

In order to avoid contaminations during RNA isolation all steps were done with disposable gloves. During preparation only sterile plastic tubes, pipettes and RNase–free pipette tips were used. Water was treated overnight with 0.1% DEPC and autoclaved thereafter to eliminate residual DEPC.

The bacteria grown overnight on GC-agar were resuspended in 5 ml supplemented PPM and incubated for 30 min with shaking at 37°C. Thereafter, the bacteria were diluted to an OD_{600} 0.1 and grown for another 75 min. 5×10^9 bacteria were used for RNA isolation. The bacterial suspension was centifugated 5 min at 4000 rpm, the pellet was resuspended in 1 ml RLT buffer + 2-mercaptoethanol (100:1) and bacterial cells were destroyed by centrifugation in the FastPrep® FP120A Instrument. A short centrifugation step allowed separating the glass beads. The supernatant transferred from the tube was mixted with 1 ml of RLT buffer + 2-mercaptoethanol and applied onto an RNeasy column. After that buffer RW1 was added and centrifugated. The washing step was done 2 times with RPE buffer and followed by elution with 500 µl RNase-free water. The quantity of RNA was estimated by gel electrophoresis on 1% agarose gel containing ethidium bromide. DNA degradation was done in the presence of $MgSO_4$, natrium sodium acetate and RNase-free DNase I for 1 hour at 37°C. Purification of RNA and elution was performed as written above (beginning with the step of mixing with RLT buffer + 2-mercaptoethanol and applying onto RNeasy column). PCR with the primers of housekeeping genes (*aroE*) allowed to control the complete degradation of DNA. The amount and the final quality of the RNA were estimated by gel electrophoresis and optical density.

4.9. Labelling of nucleic acids

4.9.1. Labelling of chromosomal DNA

4 µg chromosomal DNA (in 14 µl TE) was denatured by incubation for 10 min at 100°C and cooled on ice. Thereafter, 2 µl 10xdNTPs mixture, 2 µl nonanucleotide random primer mixture, 1 µl Klenow enzyme and 1 µl Cy3 or Cy5 were added. After overnight incubation at 37°C in darkness, the reaction was stopped by adding 2 µl EDTA (0.2 M, pH 8.0). Precipitation was done with 2 µl Na-acetate (3M, pH 4.5) and 75 µl ethanol for 30 min at – 80°C. After centrifugation for 15 min at 4°C (13000 rpm) DNA was washed with cold 70% ethanol. Labelled DNA was dried in a Speed Vac and resuspended in 10 µl TE. To the labelled DNA 5 µl (10 mg/ml) heat denaturated salmon sperm DNA was added. The mixture was incubated 10 min at 100°C and cooled on ice. Thereafter, 35µl EGT hybridisation buffer (Eurogentec) was added, mixed and used as a target.

4.9.2. Labelling of PCR products

The procedure for labelling of PCR products was the same as for chromosomal DNA with the following exceptions: 50 ng of PCR products were used; instead of nonanucleotide primers, 2 µl of hexanucleotide mixture was used; the mixture was incubated at 37°C overnigth; labelled

DNA was purified with the MinElute PCR Purification Kit (Qiagen) according to the manufacturer's instructions. Labelled PCR product resuspended in hybridisation buffer (50% formamide, 3xSSC [0.45 M NaCl, 0.045 M Na citrate], 0.2% sodium dodecyl sulfate) was denatured by incubation at 100°C for 5 min, and thereafter immediately cooled on ice.

4.9.3. Reverse transcription with Cy-labelling

25 µg RNA, 0.9 µl Nonanucleotide Mixture and 2µl Cy3/Cy5 were incubated 5 min at 70°C to separate the strands and cooled on ice. Thereafter, the mastermix was added to the solution and kept for 10 min at room temperature for binding of the primers. The mixture was incubated for 2 hours at 50°C to synthesise complementary DNA (cDNA). The process was followed by enzyme inactivation for 15 min at 70°C. The RNA was destroyed by adding DNase-free RNase and incubation for 45 min at 37°C. Purification of the cDNA was done on AutoSeq™ G-50 columns. 25 µg of each 2 different cDNA labelled with Cy3 and Cy5 were mixed together and resuspended in hybridisation buffer (50% formamide, 3xSSC, 0.2% sodium dodecyl sulfate), and denatured by incubation at 100°C for 5 min, and thereafter immediately cooled on ice.

4.10. Microarrays used in this study

4.10.1. Custom-made DNA microarrays used for MLST typing

Oligonucleotides with C6 amino linker at the 5'-end were bought from MWG Biotech as lyophilizates. All oligonucleotides sequences are provided at http://www.hygiene.uni-wuerzburg.de/Swiderek_supplement_2004. They were dissolved in water and 2xSpotting Solution II (Quantifoil) to final concentration of 20 µM. Oligonucleotides were spotted under controlled humidity and temperature environment onto epoxy slides in triplicate.

4.10.2. *N. meningitidis* DNA array from Operon used for gene expression analyses

2872 oligonucleotides with a length of 70 bp were designed by Qiagen from ORFs based on the genome sequences of 4 different strains:

Z2491 (serogroup A) (Parkhill *et al*., 2000)

MC58 (serogroup B) (Tettelin *et al*., 2000)

FAM18 (serogroup C) (unpublished, website of the Sanger Centre)

α14 (capsule null locus) (unpublished, Claus *et al*.)

One oligonucleotide (70mer) represented one gene. Oligonucleotides dissolved in spotting solution containing betaine (1.5 M betaine in 3x SSC) were spotted in final concentration of 30 µM with OmniGrid spotter onto epoxy slides in triplicate.

Immobilization of the oligonucleotides was done in 2 steps:

▪ 30 min incubation in 100% humidity at RT

▪ 60 min incubation at 120°C

After spotting and immobilization, the slides were stored under dry and dark conditions at RT.

4.10.3. *N. meningitidis* DNA array from Eurogentec used for genome comparisons

The Eurogentec DNA array comprised 2058 ORFs of strain Z2491, and 65 specific ORFs of strain MC58. PCR products included fragments with an average size of between 300 to 800 bp with a few exceptions. Each probe with $C6-NH_2$ modification at the 5'-end for covalent coupling to solid support was spotted onto aldehyde glass slides in duplicate. All slides from Eurogentec were bought ready to use.

4.11. Washing and blocking of epoxy slides before hybridisation

All washing steps were done in glass utensils with gentle rotation of washing solutions:

▪ 1x 5 min in 0,1% Triton-X100 at RT

▪ 2x 2 min in HCl, pH 4 at RT

▪ 1x 10 min in 100 mM KCl at RT

▪ 1x 1 min in ddH_2O at RT

▪ 1x 15 min in blocking solution (50 mM Ethanolamine, 0.1% SDS in 0.1 M Tris, pH 9) at
   50°C

▪ 1x 1 min in ddH_2O at RT

Slides were dried by centrifugation with 1300 rpm / 3 min.

4.12. Hybridisation conditions

The targets were given onto the slides, covered with coverslips and inserted into hybridisation chambers. Hybridisations with PCR products as target were performed at 42°C, with DNA and cDNA as a target at 36°C. All hybridisations were done overnight in a water bath in darkness. After hybridisation the slides were washed under stringent conditions at RT.

4.13. Washing of epoxy slides after hybridisation

All washing steps were done in glass utensils with gentle rotation of washing solutions at RT. Washing of the epoxy slides after hybridisation was as follows:

▪ 1x10 min solution 1 (2x SSC + 0.2% SDS)

▪ 2x10 min solution 2 (2x SSC)

▪ 1x10 min solution 3 (0.25x SSC)

Slides were dried by centrifugation at 1300 rpm for 3 min at RT.

Washing of Eurogentec slides after hybridisation was performed according to the manufacturer's instructions.

4.14. Data analysis

Fluorescence images of signals were recorded by the Affymetrix 418 TM Scanner (MWG-Biotech) with 75 photomultiplier tube voltage setting. Signal as well as background intensities were obtained for each spot and processed signal medians were generated by the ImaGene 4.0 software (Biodiscovery Inc., Los Angeles, CA, USA). In the case of discrepant results visual inspection of the scanned images was performed. Cluster analysis was done with the computer program Hierarchical Clustering Explorer 2.0 available at www.cs.umd.edu/hcil/hce.

4.15. Sequencing of DNA fragments

DNA sequencing was performed at the molecular diagnostics department of the Institute for Hygiene and Microbiology. Sequence data were analysed with the LASERGENE sequence-analysis software (DNASTAR, Madison, WI, USA).

4.16. Statistics

The statistic data analysis was performed with 2-tailed Fisher's exact test (http://www.matforsk.no/ola/fisher.htm).

5. Results

5.1. DNA arrays for MLST typing

The aim of the first project of this thesis was to develop microarrays for multilocus sequence typing of *N.menigitidis*. The disadvantage of the tiling strategy (Chee *et al*., 1996) is the large number of oligonucleotides required, e.g. to interrogate 3,282 nucleotides sequenced in the *N. meningitidis* MLST scheme (7 loci), more than 12,000 oligonucleotides would be needed. Therefore, we decided to design custom-made oligonucleotides based on polymorphisms deposited at the *Neisseria* MLST database.

5.1.1. Evaluation of the DNA-probe array

In the first experiments, the optimal experimental conditions were established. All hybridisations were performed with 50% formamide at 42°C overnight.

5.1.1.1. Optimalization of the oligonucleotide length

Probes with different lengths of oligonucleotide (18-30 nt) were designed to find the optimal length which resulted in both satisfying hybridisation intensities and discriminatory power. Oligonucleotides were derived from *abcZ* allele no 1 (*abcZ-1*), positions 128-157 (Figure 1A+1B). As shown at the Figure 1C, oligonucleotide probes with 26 nt and 30 nt gave the best hybridisation results.

(A)

HS45 AGATCGAAGCGAAGGACGGCTGGAAGTTGG

HS62   ATCGAAGCGAAGGACGGCTGGAAGTT

HS65   CGAAGCGAAGGACGGCTGGAAG

HS68   AAGCGAAGGACGGCTGGA

(B)

Oligonucleotide	Length (nt)	GC content (%)	Melting temperature (°C)
HS 45	30	56	94
HS 62	26	53	80
HS 65	22	63	72
HS 68	18	61	58

(C)
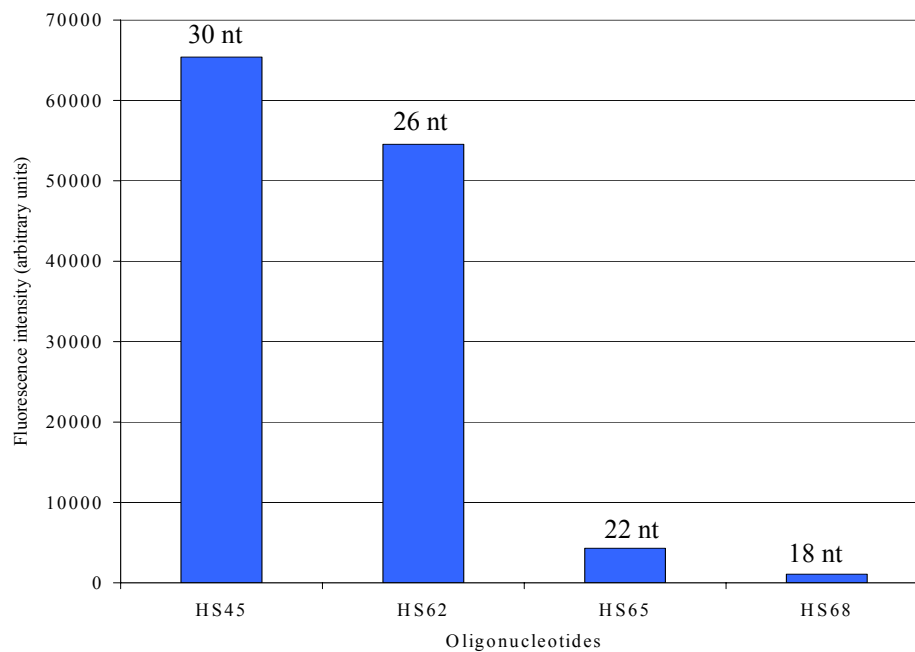
Figure 1. Influence of oligonucleotide length on hybridisation. (A+B) Oligonucleotides with different length (HS45 through HS68) were derived from *abc*Z-1. (C) Signal intensities after hybridisation with Cy3-labelled *abc*Z-1.

5.1.1.2. Influence of single and multiple polymorphisms on hybridisation

We wondered whether the protocol used for DNA/DNA hybridisation of 30-mer oligonucleotides allowed to discriminate single nucleotide polymorphisms. For this purpose, 30 nt oligonucleotides derived from *abc*Z-1 (position 128-157) with 1 to 4 polymorphisms at different positions were hybridised to an *abc*Z-1 PCR product (Figure 2A).

(A)

```
HS45   AGA TCG AAG CGA AGG ACG GCT GGA AGT TGG
HS46   AGA TCG AAG CGA AGG ACG GCT GGA AGT TGC
HS47   AGA TCG AAG CGA ACG ACG GCT GGA AGT TGG
HS48   TGA TCG AAG CGA AGG ACG GCT GGA AGT TCC
HS49   AGA TCG AAG CGA TGC ACG GCT GGA AGT TGG
HS50   AGA TCG AAG CGA ACG ACG GCT GGA AGT TGC
HS51   TGA TCG AAG CGA ACG ACG GCT GGA AGT TGG
HS52   AGA TCG AAG GGA AGG ACG GGT GGA AGT TGG
HS53   TGA TCG AAG CGA TGG ACG CCT GGA AGT TGG
HS54   AGA TCG AAG CGA TGG ACG CCT GGA AGT TGC
HS55   AGA TCG AAG CGA TGC TCG GCT GGA AGT TGG
HS56   TGA TCG AAG CGA ACG ACG GCT GGA AGT TGC
HS57   AGA TCG ATG CGA ACG ACG GCA GGA AGT TGG
HS58   TGA TCG AAC CGA AGG AGG GCT GGA AGT TGC
HS59   AGA TGG AAG CGA TGG AGG GCT GCA AGT TGG
HS60   AGA TCG AAG CCT ACC ACG GCT GGA AGT TGG
HS59   AGA TGG AAG CGA TGG AGG GCT GCA AGT TGG
HS60   AGA TCG AAG CCT ACC ACG GCT GGA AGT TGG
```

(B)
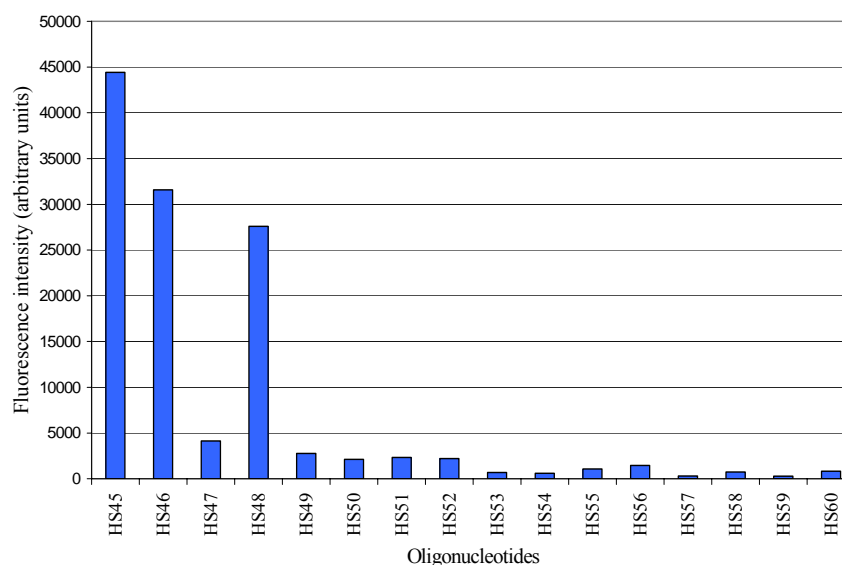
Figure 2. Influence of one or more polymorphisms on DNA/DNA hybridisation.

(A) Oligonucleotides with 1, 2, 3, and 4 polymorphisms at different locations within 30 nt *abc*Z-1 fragment.

(B) Signal intensities after hybridisation with Cy3-labelled *abc*Z-1.

Essentially, all mismatches destabilised DNA/DNA hybridisation except of mismatches at the 5'-and 3'-ends (oligonucleotides HS46 and HS48). I wondered how single polymorphisms at different positions affected the hybridisation. To answer this question, labelled PCR product of *aroE*-9 was hybridised to oligonucleotides corresponding to the positions 1-26 of *aroE*-9 with SNPs at different locations of the sequence (with the exception of the two terminal nucleotides). As shown at the Figure 3 all SNPs would be detected by the experimental protocol.

(A)

HS72   TAT CGG TTT GGC CAA CGA CAT CAC GC

HS76   TAT CGG TTT GGC **G**AA CGA CAT CAC GC

HS77   TAT CGG TTT GGC CAA CGA CAT CAC **C**C

HS78   TAT CGG TTT GGC CAA CGA CAT C**T**C GC

HS79   TAT CGG TTT GGC CAA CGA CA**A**CAC GC

HS80   TAT CGG TTT GGC CAA CGA **G**AT CAC GC

HS81   TAT CGG TTT GGC CAA C**C**A CAT CAC GC

HS82   T**T**T CGG TTT GGC CAA CGA CAT CAC GC

HS83   TAT **G**GG TTT GGC CAA CGA CAT CAC GC

HS84   TAT CG**C** TTT GGC CAA CGA CAT CAC GC

HS85   TAT CGG T**A**T GGC CAA CGA CAT CAC GC

HS86   TAT CGG TTT **C**GC CAA CGA CAT CAC GC

(B)
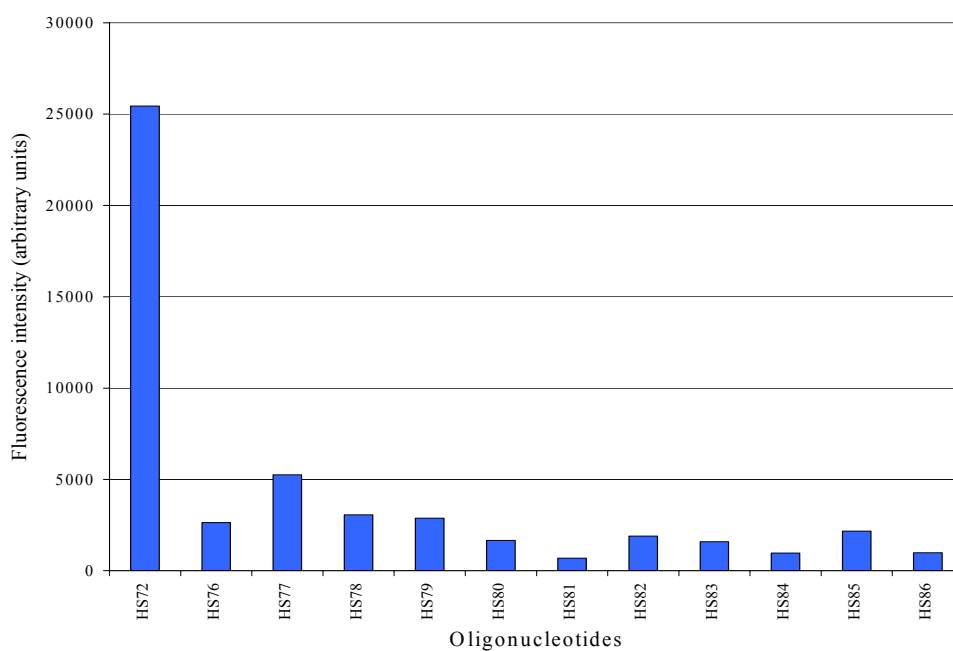
Figure 3. Influence of one polymorphism at different positions on the hybridisation. (A) Oligonucleotide HS72 represents *aroE*-9, position 1-26. The oligonucleotides HS76-HS86 were constructed with one mismatch at different positions. (B) Signal intensities after hybridisation with Cy3-labelled *aroE*-9.

5.1.1.3. Influence of melting temperature on hybridisation

I wondered if the same melting temperature could be the basis for oligonucleotide construction. The probes with identical melting temperature were investigated according to the equation of Meinkoth and Wahl (1984)[*], but different length and GC content. The experiment showed that independent of the GC content there was a critical length of oligonucleotides beyond which no hybridisation signal was obtained (Figure 4B). Probes at least with 22 nucleotides yielded high hybridisation signals, whereas probes with smaller numbers of nucleotides did not bind the fluorescence-labelled product.

(A)

		length (nt)	GC content(%)	Tm (°C)
HS87	ATTTACGCCCAATATCCGCCCTGTTATCGAATACATGAAA	40	40	59
HS88	GGTTTGGCCAACGACATCACGCAGGTTAA	29	51.7	59
HS89	TTTTACTTTTGGGCGCGGGCGGTGCG	26	61.5	60
HS90	TGCGGTGCGCGGCGTGATTCCT	22	68.1	59
HS91	CCGCGACTGCCGCCTTGCC	19	78.9	59
HS92	GGGCGGTGCGGTGCGCG	17	88.2	59

(B)
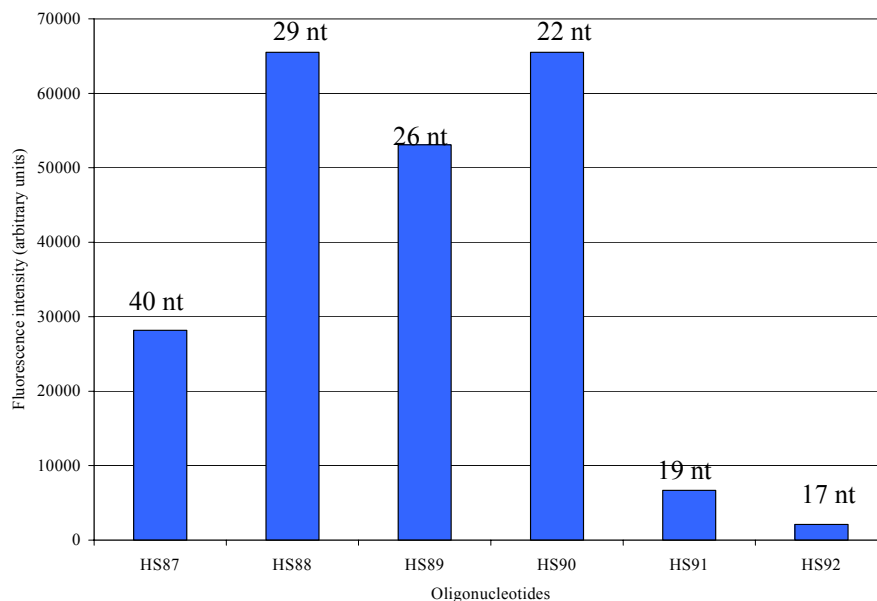
Figure 4. Influence of melting temperature on hybridisation intensity. (A) oligonucleotides with the same melting temperature (59°C) but with different length from different parts of *aroE*-9. (B) Signal intensities after hybridisation with Cy3-labelled *aroE*-9.

---

[*] The melting temperature (T$_m$) of oligonucleotides was calculated using the following equation: T$_m$=81.5°C + 16.6 log$_{10}$[C] + 0.41(%GC)-500/n-0.61(formamide). C, concentration of sodium ions in the final washing buffer used for microarray hybridisation (41.25 mM); % GC, G+C content of oligonucleotide (%-value); n, length of oligonucleotide; formamide concentration was 50%.

5.1.1.4. Influence of GC content and GC distribution on hybridisation

The GC content of a DNA segment is of importance for the strength of DNA/DNA hybridisation. I wondered whether there were major fluctuations of the GC content within alleles of MLST loci of *N. meningitidis*. To answer these question two alleles (*abcZ*-1 and *aroE*-1) were divided into 30 nt fragments and GC content was calculated. Figure 5 shows abundant GC content variation within the *aroE*-1 and *aroE*-1 alleles, which was calculated for consecutive 30 nucleotide fragments.
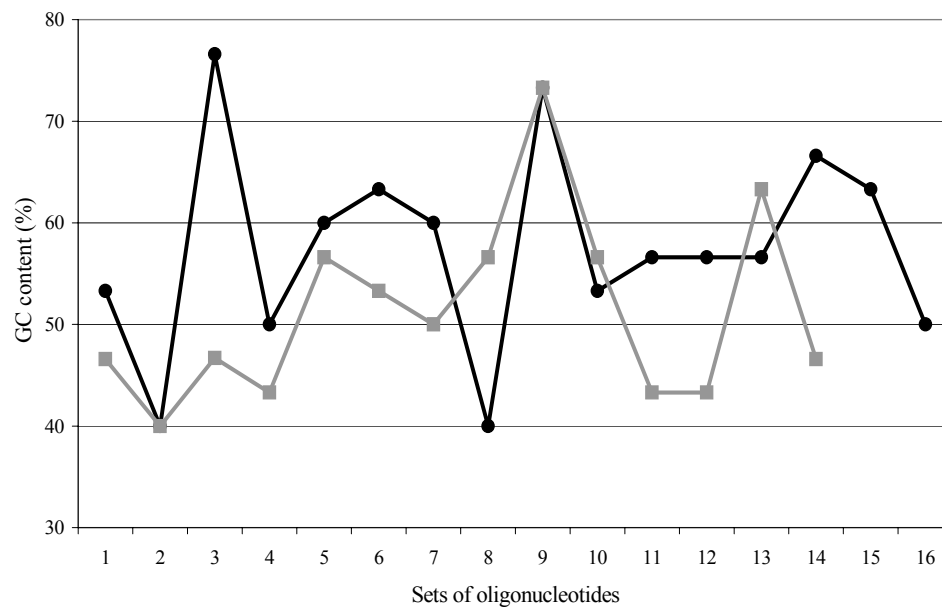
Figure 5. GC content variation within *abcZ*-1 (grey line) and *aroE*-1 (black line). The GC content was calculated for consecutive 30 nt.

To examine the influence of GC content on hybridisation, oligonucleotides with the same length (n=32), but different GC content (34-53%), derived from different parts of the *abcZ*-4 allele (Figure 6A), were tested for their DNA/DNA hybridisation to the *abcZ*-4 allele by using microarrays. As shown in Figure 6 (B) there was an almost linear relationship of GC content and hybridisation intensity.

(A)

Oligonucleotides	Length (nt)	GC content (%)
abcZ 35a2	32	34
abcZ 72a	32	40
abcZ 2a	32	46
abcZ 88	32	50
abcZ 96	32	53

(B)
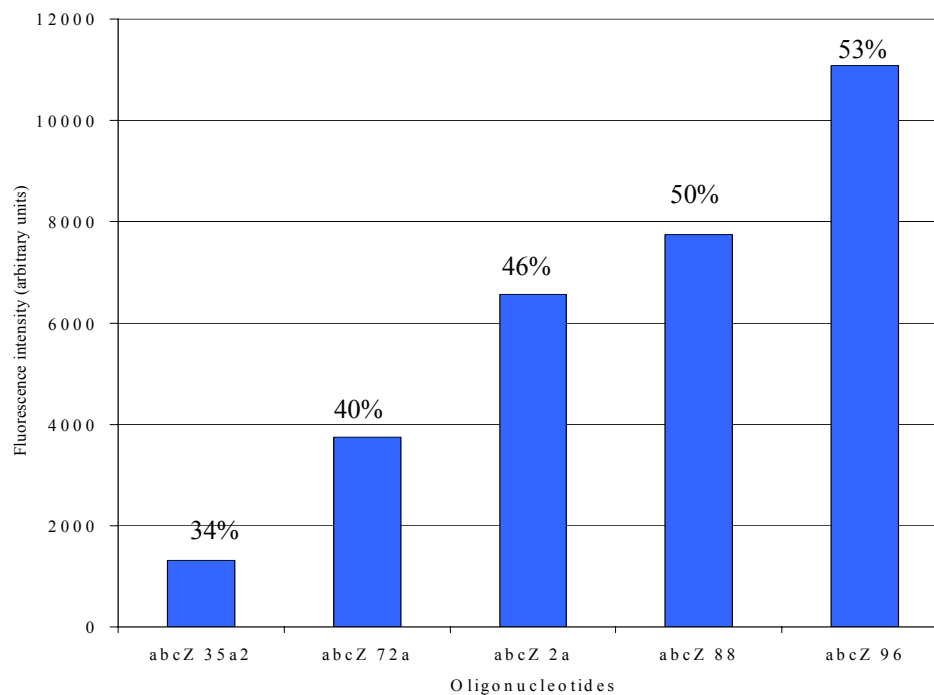
Figure 6. Influence of GC content on hybridisation. (A) Oligonucleotides of *abcZ*-4 with the same length but with different GC content. (B) Signal intensities of hybridisation with Cy3-labelled *abcZ*-4.

In summary, the data shown in Fig. 1- Fig. 6 demonstrated that the experimental set-up enabled us to discriminate DNA sequences with SNP with the exception of terminal SNPs of the probes. Signal intensities were dependent on length of oligonucleotides and GC content. The initial experiments were the basis for the further experimental strategy, i.e. oligonucleotides with 28-32 nucleotides in length were used; interpretation of hybridisation results was based on comparison of signal intensities of homologous oligonucleotides, and not on thresholds of signal intensities.

5.1.2. DNA arrays for sequence typing of *abcZ*, *adk* and *aroE* loci.

For the design of the MLST array, alleles of the two loci *abcZ* and *aroE* of invasive strains were obtained from the MLST website (Juli, 2001). 40 different alleles of *abcZ* and 57 of *aroE* were aligned with MegAlign (DNASTAR) to define the number and location of polymorphisms and polymorphic sites (Table 2).

Consequently, *abcZ* and *aroE* sequences were divided into fragments of 28-32 nucleotides with overlapping ends (Figure 7A). 16 and 19 sets of oligonucleotides of *abcZ* and *aroE* comprised 103 and 144 oligonucleotides, respectively. Each unique oligonucleotide was assigned a number (Figure 7B). The strategy chosen for microarray design is shown in Figure 7A.

Table 2. The number of polymorphisms and polymorphic sites within *abcZ* and *aroE* MLST loci of *N. meningitidis*

Locus	Length (nt)	Allele (n)	Polymorphisms (n)	Polymorphic sites (n)	Polymorphic sites/ 10 positions
abcZ	433	40	182	88	2
aroE	490	57	390	182	4
adk	465	23	82	40	1

(A)
MLST locus/Cy3-labelled PCR product

Hybridisation to oligonucleotide probes representing all polymorphisms

1

10

16

28

Assess numeric code of oligonucleotide numbers from maximum hybridisation intensities

Oligo #   1   10   16   28   30   37   42   52   60   63   69   78   87   90   95   105   115   124   129   140

Deduce MLST allele number of respective locus
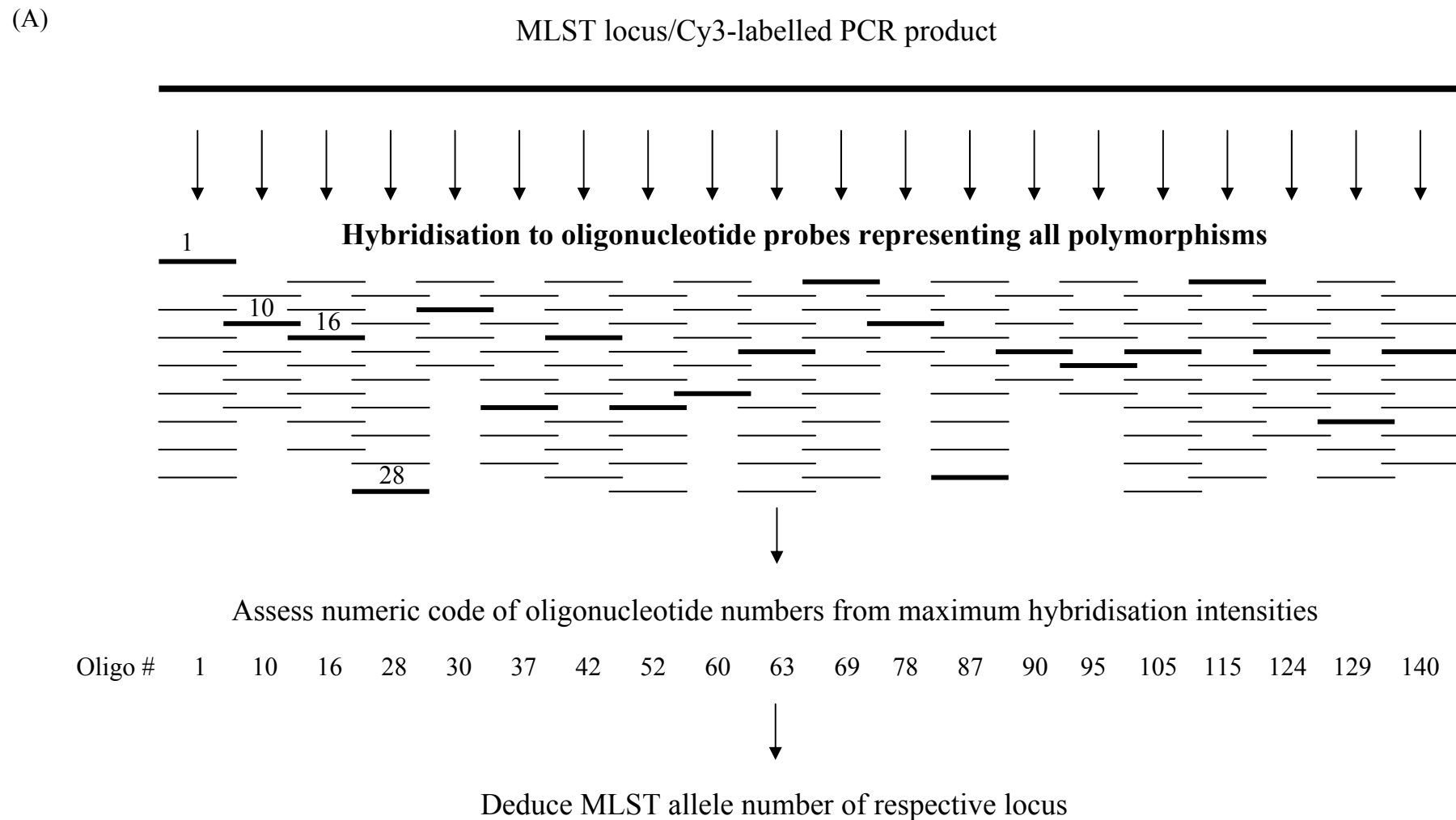
Figure 7A. Schematical depiction of the experimental design of the typing microarray. A labelled PCR product was hybridised to oligonucleotide probes representing the genetic diversity at a MLST locus. Sets of oligonucleotides represent all polymorphisms occuring at a distinct 28-32 nt fragment of the locus. Sets of oligonucleotides overlapped in average 3.6 oligonucleotides. Each oligonucleotide was assigned a number. The oligonucleotide of a set yielding the highest fluorescence intensity was considered as positive.

(B)

Sets of oligo-nucleotides / Allele number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	7	13	22	31	40	44	51	55	58	63	66	70	78	87	92
11	2	8	13	23	32	40	44	51	55	59	63	66	71	79	88	93
2	3	9	14	24	33	41	45	52	56	59	63	66	71	79	89	94
27	4	10	15	24	33	40	46	53	55	60	64	66	72	80	89	94
3	2	9	16	25	34	40	47	53	57	60	64	67	73	81	89	95
4	5	11	13	26	35	42	47	51	57	61	65	67	74	82	90	96
5	5	12	17	26	36	43	48	54	57	62	63	68	75	83	91	97
6	6	12	18	27	37	43	44	54	57	62	63	69	75	84	90	98
7	1	12	19	28	38	40	49	54	57	62	63	66	76	84	91	99
8	5	12	20	29	39	40	50	54	57	60	63	66	77	85	91	100
9	5	12	21	30	36	40	50	51	57	60	64	66	77	86	87	100

Figure 7B. The numerical patterns for each MLST-allele at the *abcZ* locus. The number within each set represents the oligonucleotide expected to yield the highest hybridisation intensity for the given allele. The identification of the allele is based on the numerical pattern of oligonucleotides which gave the highest hybridisation signals within each set of oligonucleotides.

The numerical pattern for each allele was deduced on the basis of number assignments for each unique oligonucleotide (Figure 7B). The highest hybridisation intensity within a set was considered as positive. The accuracy of the typing was assessed using different strains with different *abcZ* and *aroE* alleles. 39 hybridisations were performed with 14 different *abcZ* alleles and 36 hybridisations with 14 different *aroE* alleles. 87% and 89.9% of the oligonucleotide sets interrogated for *abcZ* and *aroE*, respectively, were correctly identified. False reacting oligonucleotides were reanalysed considering the location of the polymorphisms. In contrast to the initial evaluation with test oligonucleotides (Figure 3), some oligonucleotides with single nucleotide polymorphisms located at the 2nd and 3rd position of the 5'- and 3'-ends, respectively, gave false positive results. Therefore, some oligonucleotides were shifted to the 5'- or 3'-direction, others were changed in length (still keeping length 28-32 nt) to avoid SNPs at the three positions located at the 5'- or the 3'-ends, respectively. 128 of 247 oligonucleotides, 55 of *abcZ* and 73 of *aroE*, were modified. 18 oligonucleotides yielding false positive results could not be re-designed without affecting the adjacent blocks. The modification of the oligonucleotide design reduced the error rates from 13% to 7.5% for *abcZ* (10 hybridisations performed with 9 different alleles),

and from 10.1% to 8% for *aroE* (19 hybridisations done with 12 different alleles) (Table 3). All of these findings were applied to the design of the *adk* oligonucleotides. The locus was divided into 17 blocks comprising 50 oligonucleotides. 13 hybridisations were done with 13 different *adk* alleles to test 33 *adk* oligonucleotides of all 50. 98.2% of the oligonucleotide sets interrogated were correctly identified (Table 3). I assume that the smaller number of mismatches within the *adk* locus (1 polymorphic site/10 positions) in comparison to the other loci facilitated the design of oligonucleotides.

Table 3. Validation of microarrays for typing at the *abcZ, aroE* and *adk* loci of *N. meningitidis*

Locus	Round of validation	Oligonucleotides (n)	Alleles (n)	Hybridisations (n)	Number of oligonucleotide sets interrogated (n)	Misidentified oligonucleotide sets (%)
abcZ	#1	103	14	39	624	81 (13.0)
	#2	100	9	10	160	12 (7.5)
aroE	#1	144	14	36	684	69 (10.1)
	#2	140	12	19	361	29 (8.0)
adk	#2	50	13	13	221	4 (1.8)

5.1.3. Hybridisation with more than one locus as a target

For efficient MLST typing by microarrays, all MLST alleles of a strain must be amplified in a single or two PCR reactions. Two multiplex PCRs were designed for the amplification of *abcZ, pgm, adk* and *aroE, fumC, gdh, pdhC,* respectively. Six different strains were tested for multiplex PCR, in all cases the same results were achieved. A representative result of the multiplex PCR for two strains is shown in Figure 8. I was able to amplify three and four targets, respectively, in one reaction. There were, however, some variations in the DNA concentration of single bands as suggested by visual inspection.
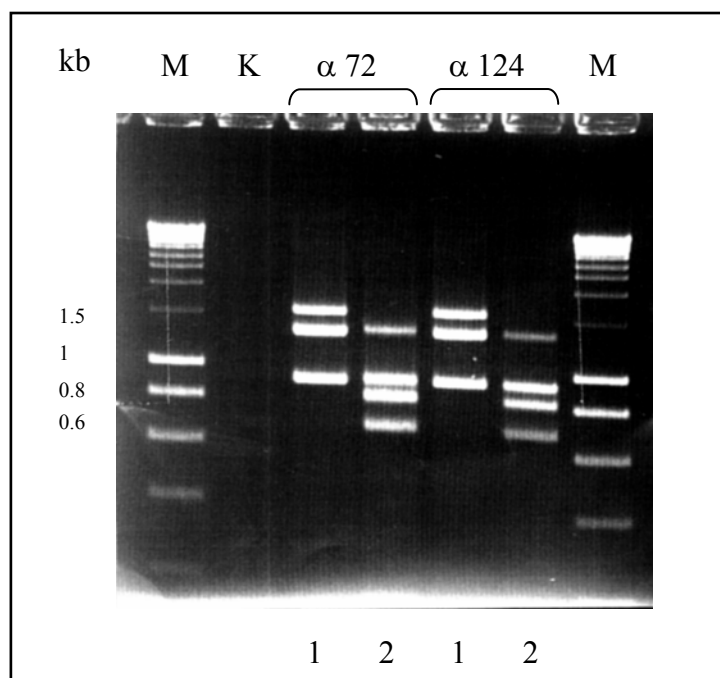

Figure 8: Multiplex PCR of 7 MLST loci. 1, amplification of *abcZ, pgm* and *adk* loci; 2, amplification of *aroE, fumC, gdh* and *pdhC*. Chromosomal DNAs of α72 and α124 meningococcal strains were used for the multiplex PCRs. M, marker.

5.3. Genomic comparisons of ET-15 and non ET-15 strains by microarrays.

The second project of this thesis was dedicated to the study of the genetic diversity of the ET-15 clone, which represents a successful and highly virulent variant of the ST-11 complex of *Neisseria meningitidis*. This clone has been responsible for serogroup C meningococcal disease worldwide. Microarray with complete genomes enabled us to analyse differences in genome content between ET-15 isolates and other members of the ST-11 complex at a genomic scale.

5.3.1. Cut-off validation for the presence and absence of genes within tested strains

The whole genome DNA microarray was purchased from Eurogentec and contained PCR products of 2058 ORFs of the serogroup A strain Z2491, and 65 ORFs of the serogroup B strain MC58. 12 ORFs were omitted for final analysis because of their length (72-117 nt) resulting in false negative signals. Chromosomal DNA of ET-15 and non ET-15 strains was labelled with Cy3-dCTP and hybridised to the slides. 12 strains belonging to the ET-15 clone, 3 non ET-15 but ST-11 complex strains, as well as strains Z2491 (serogroup A) and MC58 (serogroup B) were analysed. At least two independent experiments were done with each strain. The repeat experiment confirmed the absence of any differences. The visual inspection of the scanned images was performed, which in every case solved the discrepancy between experiments. The data of each experiment were sorted by rank abundance curves based on the fluorescence intensities. Then a binary assignment was applied to categorise genes into present (1) or absent (0) based on the level of hybridisation signals. The cut-off was empirically determined. For this purpose, we selected 75 genes, which were known to be present in strains Z2491 (serogroup A) and/or MC58 (serogroup B), respectively, but which were lacking in strains of the ST-11 complex (Table 4). I defined the cut-off value as the hybridisation intensity of the 90% percentile of those 75 genes.

Table 4. Genes lacking in the genome sequence of FAM18 (ST-11 complex).

Genes lacking in ST-11 complex	Characteristic	Number of lacking ORFs on the chip
sacA,B,C,D	capsule synthesis genes of serogroup A	4 ORFs
pheS/T	lineage specific insertion between *pheS* and *pheT*	2 ORFs
opcA	absent in ST-11	1 ORF
NMA ORFs	phage associated ORFs	42 ORFs
NMB ORFs	ORFs specific only for serogroup B strains	26 ORFs

Figure 9 demonstrates the position of cut-offs in rank abundance curves obtained by plotting of hybridisation intensities in a few experiments with different ST-11 strains. In all cases, 175-195 genes were assigned as negative. Furthermore, the cut-offs in all cases were located in the part of the curves representing a rapid decline of hybridisation.

(A)   strain 2979

(B)   strain α737
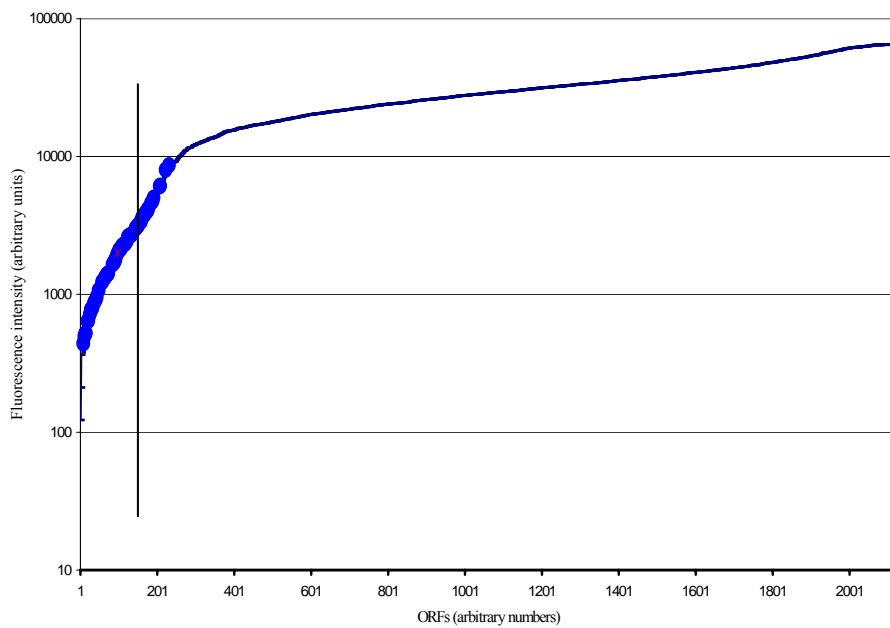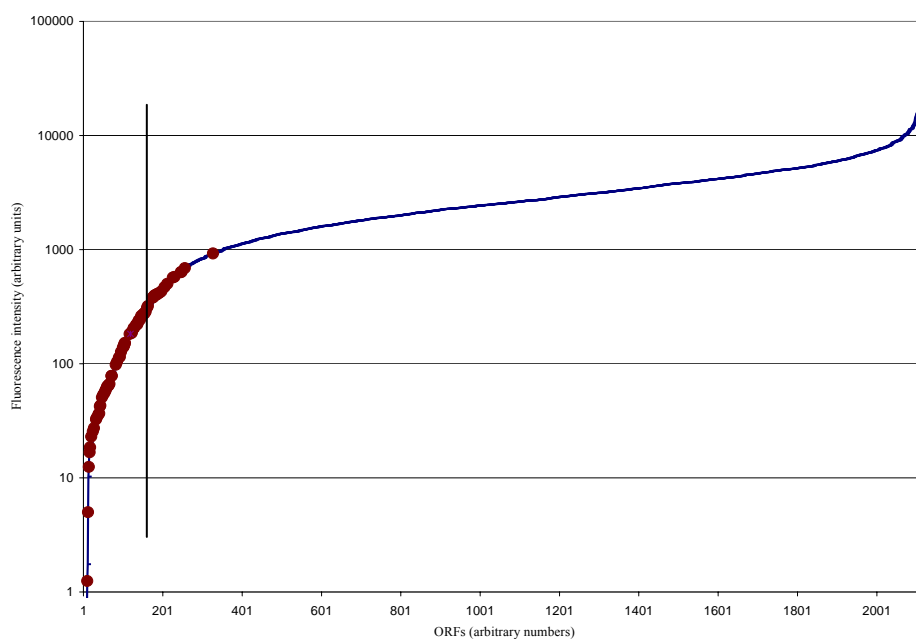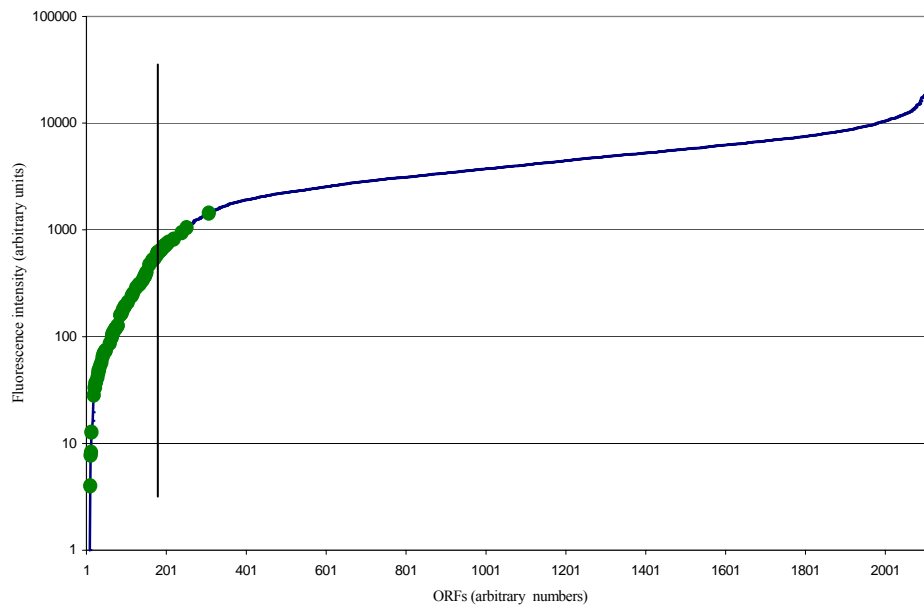
(C)   strain DE9260

(D)   strain DE9246
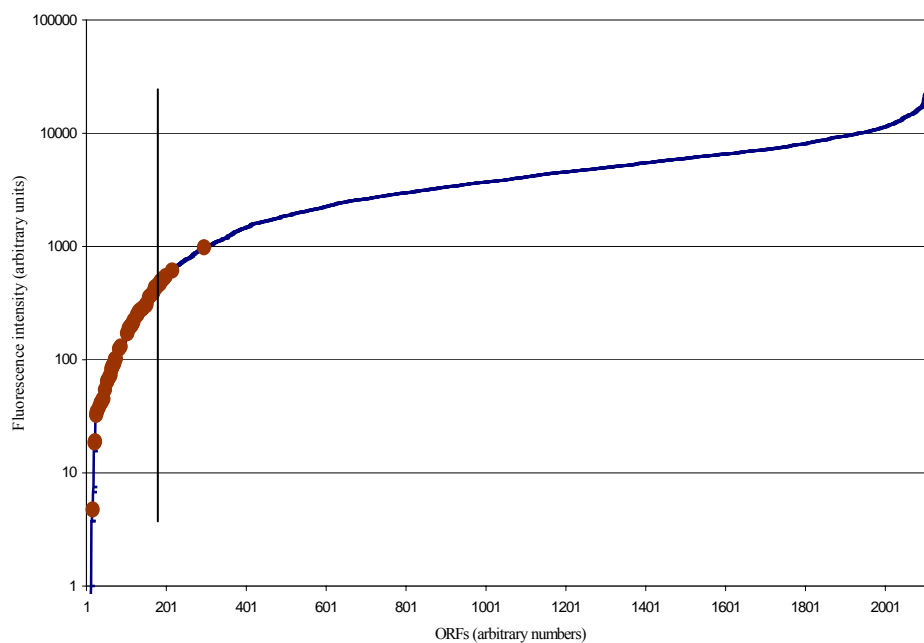
Figure 9 (A-D). The rank abundance curves of DNA/DNA hybridisation intensities for cut-off determination. Chromosomal DNAs of four meningococcal strains of the ST-11 complex were hybridised to whole genome DNA microarrays purchased from Eurogentec. The following strains were used: (A) strain 2979; (B) strain α737; (C) strain DE9260; (D) strain DE9246. The Y-axis presents the fluorescence intensity in the logarithmic scale, while the X-axis shows all ORFs tested on the slide. Points represent the population of genes known to be lacking in the genome sequence of strain FAM18 (ST-11). These genes were used for cut-off definition. The vertical line shows the position of the cut-off.

The cut-off definition proved to be reliable and was applied to all strains tested with the Eurogentec microarrays. The results of experiments are summarised in Figure 10. Between 175 (8.5%) and 195 (9.4%) of all ORFs represented on the array did not hybridise with chromosomal DNA of ST-11 strains according to the algorithm applied. The value was much lower (1.7%) for the strain Z2491 (serogroup A). This finding was expected because the genome sequence of this strain was the basis for array construction and because only a minor portion of MC58 (serogroup B) –specific genes is represented on the slide.

Figure 10. Proportions of present and absent ORFs in ST-11 strains. Grey bar shows the presence of ORFs while white bar the absence of ORFs within each strain.

5.3.2. Genomic diversity within tested strains

1836 ORFs represented on the Eurogentec microarray were shared by all analysed strains. 85 ORFs were differentially distributed within all ST-11 complex strains. Differences between gene contents are shown in Table 5. The genetic differences within all strains belonging to the ET-15 clone ranged between 0.4 and 1.3% (9-26 genes). The gene content difference between ET-15 str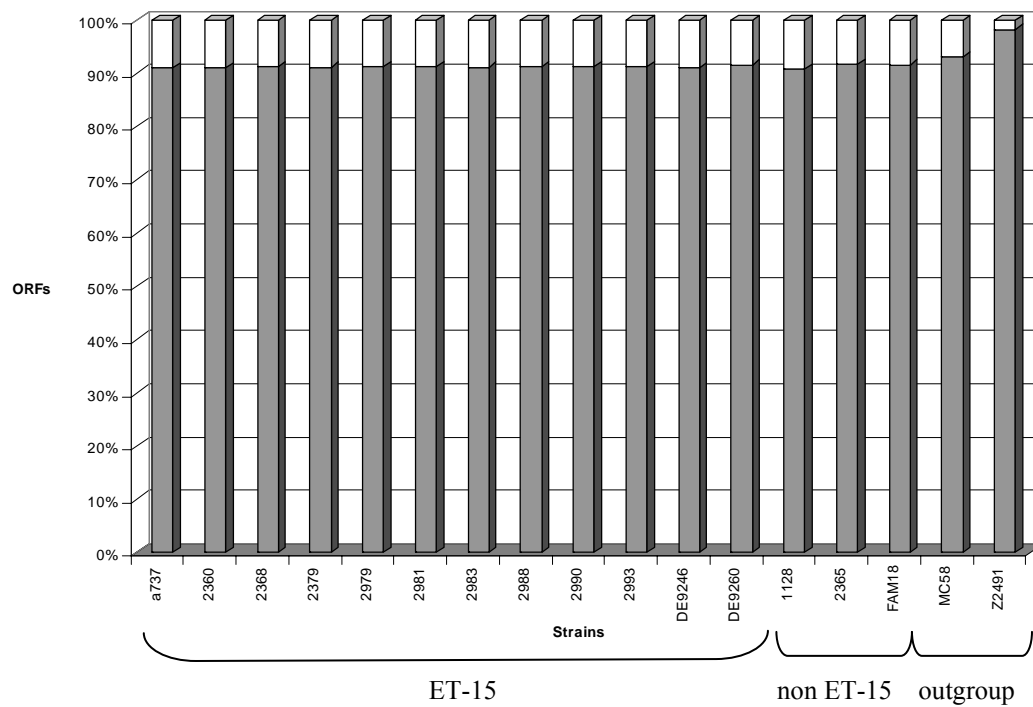ains and non ET-15 strains of the ST-11 complex was shown to be almost twice as high as the genetic difference within the ET-15 clone (1.5% vs. 0.8% in average). Additionally, differences of genome composition within non ET-15 but ST-11 complex strains ranged between 1 and 2.1%. The highest genetic diversity of analysed genomes was observed between Z2491 or MC58 and all ET-15 strains with an average difference of 8.2% and 6%, respectively (Table 5):

Table 5. Genomic differences (%) between tested strains. The values show the genetic diversity of all tested strains compared to each other based on their gene content.

	a737	2360	2368	2379	2979	2981	2983	2988	2990	2993	DE9246	DE9260	1128	2365	FAM18	MC58	Z2491
a737		0,8	0,8	1,1	0,9	0,6	1	0,9	0,9	1,2	0,8	0,8	2	1,6	1,9	6,3	8,2
2360			0,9	1	0,8	0,4	1	0,9	0,6	0,9	0,9	1,2	2	1,6	2,1	6,4	8,4
2368				0,6	0,6	0,6	0,7	0,7	0,9	0,9	0,8	0,8	1,5	1,2	1,5	6,4	8,1
2379					0,8	0,9	0,6	0,7	1	0,8	0,6	0,9	1,3	1,3	1,4	6	8,3
2979						0,5	0,7	0,8	0,6	0,8	0,6	1	1,7	1,4	1,6	6	8
2981							0,8	0,6	0,4	0,6	0,5	1	1,8	1,2	1,7	6,1	9
2983								0,6	0,9	0,8	1	1	1,3	1,6	1,5	6	8,2
2988									0,7	0,8	0,6	1	1,3	1	1,5	5,8	8,1
2990										0,9	0,7	1	1,8	1,4	1,8	6,1	8,1
2993											0,5	1,3	1,5	1,2	1,6	6,1	8,1
DE9246												1	1,5	1,4	1,6	6	8,3
DE9260													1,6	1,4	1,8	5,9	7,8
1128														1,5	1,7	6,3	8,6
2365															1,6	6,7	7,6
FAM18																6	7,9
MC58																	8,5
Z2491																	

ET-15   non ET-15   outgroup

5.3.3. Cluster analysis

The computer program Hierarchical Clustering Explorer 2.0 (www.cs.umd.edu/hcil/hce) was applied to the lineage assignments indicating presence or absence of genes in each strain. By hierarchical clustering a small group of genes is identified that shares a common pattern of results. A dendrogram is constructed in a sequential manner using a ranked series of such clusters. Figure 11 demonstrated the cluster analysis. The program grouped together ET-15 strains and separated them from other ST-11 strains. Non ST-11 strains served as outgroup strains.
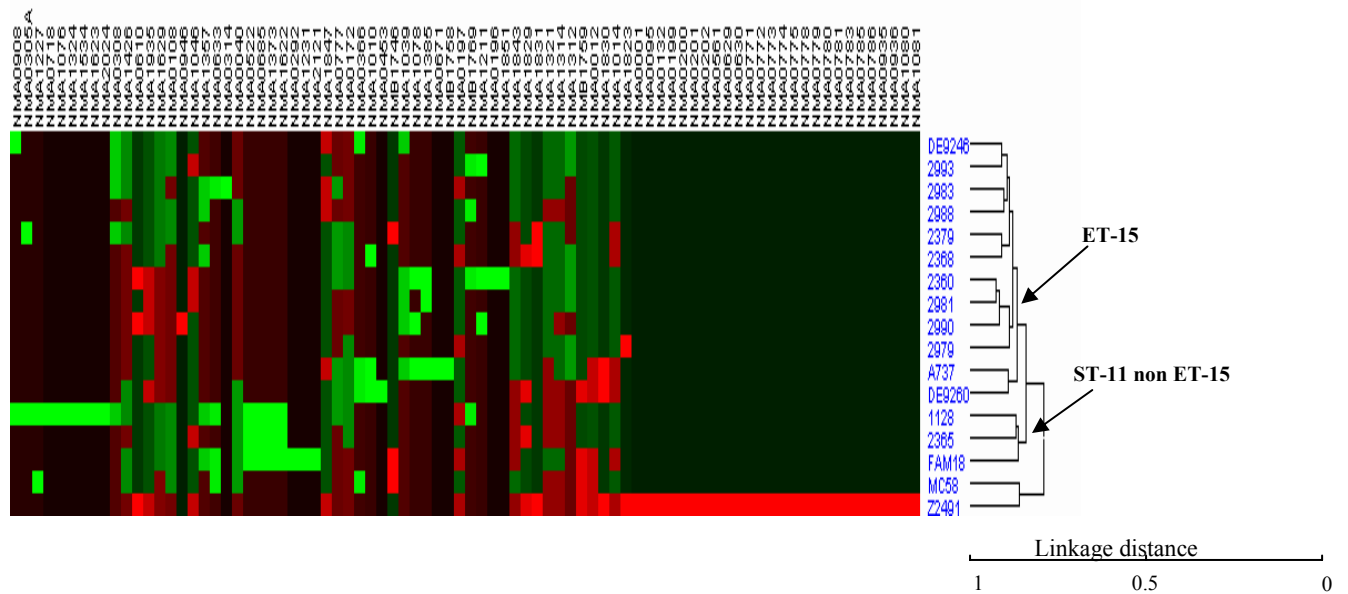

Figure 11. (Left panel) Cluster analysis of all tested strains. Strains and genes were grouped by average hierarchical clustering using Hierarchical Clustering Explorer 2.0. (Right panel) Dendrogram deduced from the cluster analysis.

5.3.4. ORFs specific to the ET-15 clone

As elucidated by microarray analysis, 3 ORFs (NMA1373, NMA1622 and NMA0522) were associated to ET-15 strains and only rarely occurred in other ST-11 strains. These 3 genes are classified as genes encoding hypothetical proteins. The differential distribution of 3 genes specific for the ET-15 clone was confirmed by PCR reactions using the 17 strains. A representative example is provided in Figure 12:
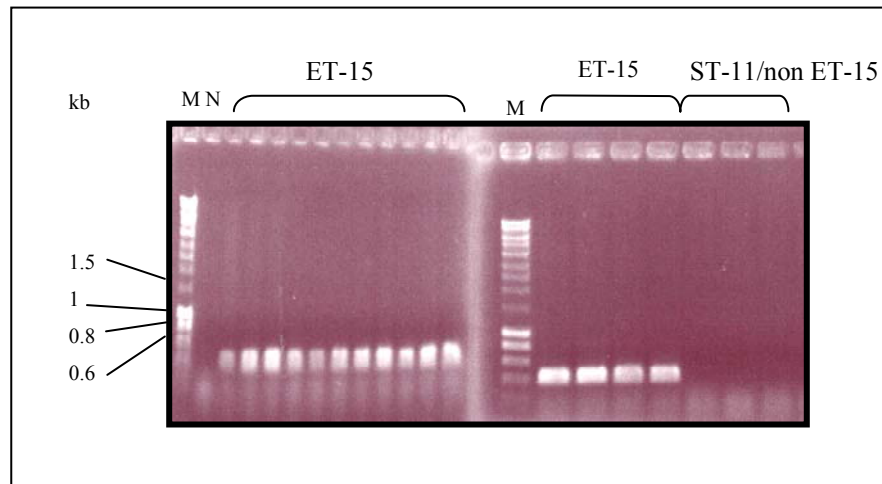
Figure 12. ORF NMA1622 was shown by microarrays to be specific to the ET-15 clone. The presence of NMA1622 in ET-15 strains and its absence in other ST-11 strains was confirmed by PCR using the primers NMA 1622 F and NMA 1622 R. M, marker; N, negative control.

To further prove the differential distribution of NMA1373, NMA1622 and NMA0522, additional strains of the ET-15 clone and non ET-15 strains were analysed by PCR. The strains were obtained at various time points from different geographical regions (Table 1). As shown in Table 6, the three ORFs were significantly associated with the ET-15 clone.

Table 6. Distribution of NMA1373, NMA1622 and NMA0552 genes within tested strains. *2-tailed Fisher's exact test.

Genomic locus	ET-15	non ET-15	p-value*
NMA1373	65/65	1/44	1×10^{-29}(s)
NMA1622	64/65	4/44	5×10^{-24}(s)
NMA0522	74/76	5/34	3×10^{-19}(s)

6. Discussion

6.1. Microarray technology as an alternative to MLST

MLST was first described by Maiden *et al.*, (1998) who applied the method to a collection of *N. meningitidis* isolates that had been characterized previously by MLEE. From that time, MLST has become the standard method for typing of several bacterial species and *C.albicans*. Sequences of internal fragments of at least seven housekeeping genes are determined for each isolate. 450-500bp fragments are sequenced providing the identification of many different alleles within a bacterial population. The method has several advantages: the method is highly accurate; it can be automated; it can be conducted in all laboratories with access to PCR and DNA sequencing technology; the data are phylogenetically meaningful; they are portable and easy to analyse. Despite the reduction of costs of sequencing reagents in the recent years, there is still the open question of the necessity of the application of MLST in routine molecular epidemiology with regard to costs of 7 PCRs and 14 sequencing reactions. At the Institute for Hygiene and Microbiology of the University of Würzburg, the MLST of a single bacterial isolate costs approximately € 50, if only consumables are taken into account. Oligonucleotide array based sequence analysis is suitable for several applications. One is the detection of commonly occurring sequence changes for which specific hybridisation patterns are known. Another is the detection of unknown changes (Wang et al., 1998; Hacia, 1999). In contrast, high density microarrays utilizing the tiling strategy, which requires the synthesis of a large number of oligonucleotides (Chee *et al.*, 1996) may be considered as a substitute for DNA sequence analysis, because each position of a reference sequence is interrogated using all four nucleotides This approach has been successfully applied for MLST analysis by van Leeuwen *et al.*, (2003), who used *Staphylococcus aureus* as a model. In this study, we constructed the custom-made microarrays for *N. meningitidis* MLST analysis with moderate numbers of oligonucleotides interrogating only polymorphisms published on the *Neisseria* website. We chose this approach because: 1) polymorphisms in *N. meningitidis* are phylogenetically old and the likelihood of appearance of novel polymorphisms is low (Jolley *et al.*, 2000); 2) the allelic profile of bacterial isolates is sufficiently stable over the past decades (reviewed in Vogel *et al.*, 2004). The advantage of handling a limited number of oligonucleotides seemed to outbalance a disadvantage in comparison to the tiling technology, i.e. the impossibility to discover novel polymorphisms. Unfortunately, the strategy chosen could not be conducted with satisfying accuracy. There are several possible reasons. The first seems to be the high density of polymorphisms which complicated the oligonucleotide design. It was not surprising that the highest accuracy was achieved for the *adk* locus with the

smallest number of polymorphisms. The second reason was that mismatches in the centre of the hybrid exert more destabilizing effect than those located near the ends. The tiling strategy overcomes the problem of mismatch location by interrogating nucleotides exactly in the middle position of the oligonucleotide. Thirdly, we observed significant undulations of the GC content within single alleles up to 37%. Unfortunately, in our experiments, nucleotide composition played an important role in hybridisation (Figure 6). In summary, our strategy of polymorphism-directed oligonucleotide design was not suitable for *N. meningitidis*, probably due to the large number of polymorphisms and the excessive undulations of the GC content. Considering the fact that in the MLST loci of *S. aureus* the proportion of polymorphic sites and polymorphisms among alleles as well as GC content fluctuations are much lower than in *N. meningitidis*, this organism might have been a better choice for our approach. A comparison of the *S. aureus* and *N. meningitidis* MLST data is provided in Tables 7 and 8. The number of polymorphisms and polymorphic sites in all seven MLST loci of *S. aureus* is 3-4 fold lower than in the loci analysed in *N. meningitidis*. It may be assumed that the lower number of polymorphic sites and thus the lower number of reference sequences made *S .aureus* also a suitable tool for tiling strategy (van Leeuwen *et al*., 2003).

Table 7. Polymorphic sites at seven MLST loci of (A) *S. aureus* (data from Enright *et al*., 1999) and (B) *N. meningitidis* (data from Maiden *et al*., 1998 and Jolley *et al*., 2000), respectively.

(A)

S. aureus (53 STs)

Gene	Sequence length (bp)	No. of alleles	No. of polymorphic sites	P.s./ 10 positions
arcC	456	17	19	0,4
aroE	456	17	23	0,5
glpF	465	11	14	0,3
gmk	429	11	13	0,3
pta	474	15	18	0,4
tpi	402	14	18	0,5
yqiL	516	16	19	0,4

(B)

N. meningitidis (74 STs)

Gene	Sequence length (bp)	No. of alleles	No. of polymorphic sites	P.s./ 10 positions
abcZ	433	15	75	1,7
adk	465	10	17	0,4
aroE	490	18	166	3,4
fumC	465	19	38	0,8
gdh	501	16	28	0,6
pdhC	480	24	80	1,7
pgm	450	21	77	1,7

Table 8. GC content fluctuation (%) among ten different alleles of the *aroE* locus of *S. aureus* (A) and of *N. meningitidis* (B).

(A) *S. aureus*

aroE-allele (No)	GC (%)[*]													
1	23,3	13,3	23,3	33,3	36,6	33,3	40	36,6	23,3	50	23,3	40	26,6	33,3
2	23,3	10	23,3	36,6	36,6	33,3	43,3	36,6	23,3	46,6	23,3	40	26,6	36,6
3	23,3	13,3	23,3	33,3	33,3	33,3	43,3	36,6	23,3	50	23,3	40	26,6	33,3
4	23,3	13,3	23,3	33,3	33,3	33,3	43,3	36,6	23,3	50	23,3	40	26,6	36,6
5	23,3	10	23,3	36,6	36,6	33,3	43,3	40	23,3	50	23,3	40	26,6	33,3
6	23,3	13,3	23,3	36,6	36,6	33,3	43,3	36,6	23,3	50	23,3	40	26,6	33,3
7	23,3	10	23,3	33,3	36,6	33,3	43,3	36,6	23,3	53,3	23,3	40	26,6	36,6
8	23,3	13,3	23,3	36,6	36,6	33,3	43,3	36,6	23,3	50	23,3	40	26,6	36,6
9	23,3	13,3	23,3	33,3	33,3	33,3	43,3	36,6	23,3	50	23,3	40	26,6	33,3
10	23,3	10	23,3	36,6	36,6	30	43,3	36,6	23,3	46,6	23,3	40	26,6	36,6
differences of GC content (%) among alleles	**0**	**3,3**	**0**	**3,3**	**3,3**	**3,3**	**3,3**	**3,4**	**0**	**6,7**	**0**	**0**	**0**	**3,3**

(B) *N. meningitidis*

aroE-allele (No)	GC (%)[*]														
1	40	76,6	50	60	63,3	60	40	73,3	53,3	56,6	56,6	56,6	66,6	63,3	50
2	40	70	50	56,6	66,6	60	40	63,3	40	60	46,6	50	63,3	63,3	52,5
3	40	76,6	50	56,6	66,6	60	40	63,3	40	60	46,6	50	63,3	63,3	52,5
4	40	70	53,3	60	66,6	60	40	63,3	40	60	46,6	50	63,3	63,3	52,5
5	40	70	50	56,6	66,6	60	43,3	63,3	40	60	46,6	50	63,3	63,3	54,8
6	33,3	80	53,3	56,6	66,6	60	36,6	63,3	40	60	46,6	50	63,3	63,3	50
7	40	73,3	53,3	66,6	63,3	60	40	73,3	53,3	56,6	56,6	56,6	66,6	63,3	50
8	40	73,3	53,3	66,6	63,3	60	40	70	53,3	56,6	56,6	56,6	66,6	63,3	50
9	36,6	66,6	50	56,6	60	60	40	66,6	53,3	60	46,6	66,6	70	63,3	46,6
10	43,3	70	50	56,6	66,6	60	40	63,3	40	56,6	50	56,6	66,6	63,3	53,3
differences of GC content (%) among alleles	**10**	**13,4**	**3,3**	**10**	**6,6**	**0**	**6,7**	**10**	**13,3**	**3,4**	**10**	**16,6**	**6,7**	**0**	**8,2**

[*] The GC content was calculated for consecutive 30 nt fragments of the *aroE* sequence used in the MLST scheme.

In general, it is our assumption that a broad use of such microarrays will be hampered by the availability of specific equipment in diagnostic laboratories. Furthermore, it is also unlikely that microarrays could significantly save operator time and costs in comparison to DNA sequencing, because several experimental steps are required, i.e. DNA isolation, performance of a reliable multiplex PCR on seven targets, PCR product purification, target labeling, hybridisation, and data analysis (Table 9).

Table 9. Comparison of technical procedures and cost calculations for MLST and microarrays

	MLST-7 loci	Microarrays-7 loci
Experimental steps	PCR-7 loci	Multiplex PCR
	PCR product purification	PCR product purification
	Sequencing	Target labeling
	Data analysis	Hybridisation
		Data analysis
Duration	≥ 2 days	3 days
Estimated cost per strain (€)[*]	50	75

[*]) Costs are calculated exclusively for consumables.

We assume that MLST will remain the typing system of choice, especially because future technical improvements will reduce costs and throughput-time even further.

6.2. Genetic diversity of the ET-15 clone.

The ET-15 clone was first observed in Canada and spread worldwide consecutively. It was reported that the ET-15 clone tended to cause severe meningococcal disease at high frequency (Ashton et al., 1991). The reasons for the appearance of a novel fitness peak within the ST-11 complex are unknown. The ET-15 clone provides the unique opportunity to study the abundance of genetic alterations of a recently emerged successful meningococcal clone. In this study comparative genome scanning of ET-15 strains and non ET-15 strains of the ST-11 complex was applied. Of 2111 genes analysed by microarrays, 1837 (87%) were common to all strains of *N. meningitidis*. Within 84 divergent ORFs of ST-11 strains, 3 ORFs (NMA1373, NMA1622 and NMA0522) were found to be specific to the ET-15 clone of the ST-11 complex. These 3 genes encode hypothetical proteins. NMA1373 and NMA1622 exibit very low GC content frequently observed in acquired DNA. The role of these ORFs in biology of the ET-15 clone is difficult to estimate now, but as they are significantly associated with ET-15 strains they could be used as markers for clone identification. NMA1373, NMA1622 and NMA0522 are not the only alterations of the ET-15 clone identified by our group until now: (i) a mutation at position 640 of the fumarate hydratase gene was shown to be a clone-specific characteristic (Vogel *et al.*, 2000); (ii) ET-15 meningococci consistently exhibit an alternative allele of an iron-related protein not observed in non ET-15 strains belonging to the ST-11 complex (Caugant, personal communication); (iii) ET-15 meningococci harbour the IS*1301*, which is absent from other ST-11 (Claus *et al.*, manuscript in preparation).

The following schematic scenario for the recent evolution of the ET-15 clone can be proposed (Figure 13). ST-11 meningococci underwent a considerable number of genetic alterations by DNA uptake from other *Neisseria*, until a clone emerged which displayed the ability to increase carriage and transmission within human population. The successful epidemiology of the new clone is reflected by rapid intercontinental spread. The reason for the high pathogenicity and tendency to cause outbreaks is unknown and not explained by genetic changes elucidated until now. As demonstrated for other meningococcal lineages, the ET-15 clone underwent diversification following clonally expansion (Jelfs *et al.*,2000; Krizowa and Musilek, 1995; Krizowa *et al.*,1997).

NMA 1373, 1622, 0522

tbpB II

ST-11 C:2a:P1.5,2 (*non ET-15*) → time → ST-11 C:2a:P1.5,2 (ET-15)

IS*1301*

fumC position 640 A

Diversification
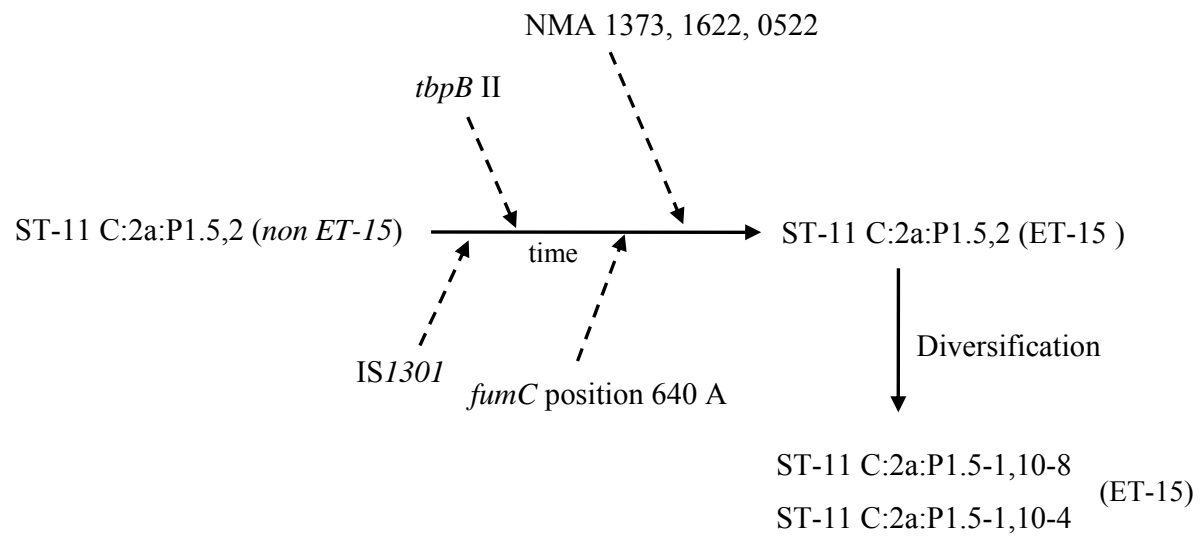
ST-11 C:2a:P1.5-1,10-8
ST-11 C:2a:P1.5-1,10-4   (ET-15)

Figure 13. Acquisition of genetic traits in the evolution of the ET-15 clone.

7. Bibliography

M. Achtman. Epidemic spread and antigenic variability of Neisseria meningitidis. *Trends Microbiol.* 3 (5):186-192, 1995.

M. F. Anjum, S. Lucchini, A. Thompson, J. C. Hinton, and M. J. Woodward. Comparative genomic indexing reveals the phylogenomics of Escherichia coli pathogens. *Infect.Immun.* 71 (8):4674-4683, 2003.

R. M. Anthony, T. J. Brown, and G. L. French. Rapid diagnosis of bacteremia by universal amplification of 23S ribosomal DNA followed by hybridization to an oligonucleotide array. *J.Clin.Microbiol.* 38 (2):781-788, 2000.

A. Apicella. Chapter 171: Neisseria meningitidis. Principles and Practice of Infectious Disease. Second Edition. John Wiley&Sons.Inc., 1985, 1186-1190.

F. E. Ashton, J. A. Ryan, A. Borczyk, D. A. Caugant, L. Mancino, and D. Huang. Emergence of a virulent clone of Neisseria meningitidis serotype 2a that is associated with meningococcal group C disease in Canada. *J.Clin.Microbiol.* 29 (11):2489-2493, 1991.

M. A. Behr, M. A. Wilson, W. P. Gill, H. Salamon, G. K. Schoolnik, S. Rane, and P. M. Small. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 284 (5419):1520-1523, 1999.

U.Berger, M.-T. Döring. Zur Resistenz der humanen Neisseria-Arten gegen Austrocknung. *Arch. Hyg.* 153 (6):556-559, 1969.

L. D. Bowler, Q. Y. Zhang, J. Y. Riou, and B. G. Spratt. Interspecies recombination between the penA genes of Neisseria meningitidis and commensal Neisseria species during the emergence of penicillin resistance in N. meningitidis: natural events and laboratory simulation. *J.Bacteriol.* 176 (2):333-337, 1994.

E. F. Boyd, F. S. Wang, T. S. Whittam, and R. K. Selander. Molecular genetic relationships of the salmonellae. *Appl.Environ.Microbiol.* 62 (3):804-808, 1996.

D. R. Call, F. J. Brockman, and D. P. Chandler. Detecting and genotyping Escherichia coli O157:H7 using multiplexed PCR and nucleic acid microarrays. *Int.J.Food Microbiol.* 67 (1-2):71-80, 2001.

D. A. Caugant, B. E. Kristiansen, L. O. Froholm, K. Bovre, and R. K. Selander. Clonal diversity of Neisseria meningitidis from a population of asymptomatic carriers. *Infect.Immun.* 56 (8):2060-2068, 1988.

D. A. Caugant. Population genetics and molecular epidemiology of Neisseria meningitidis. *APMIS* 106 (5):505-525, 1998.

K. Chan, S. Baker, C. C. Kim, C. S. Detweiler, G. Dougan, and S. Falkow. Genomic comparison of Salmonella enterica serovars and Salmonella bongori by use of an S. enterica serovar typhimurium DNA microarray. *J.Bacteriol.* 185 (2):553-563, 2003.

M. Chee, R. Yang, E. Hubbell, A. Berno, X. C. Huang, D. Stern, J. Winkler, D. J. Lockhart, M. S. Morris, and S. P. Fodor. Accessing genetic information with high-density DNA arrays. *Science* 274 (5287):610-614, 1996.

T. Chen, Y. Hosogi, K. Nishikawa, K. Abbey, R. D. Fleischmann, J. Walling, and M. J. Duncan. Comparative whole-genome analysis of virulent and avirulent strains of Porphyromonas gingivalis. *J.Bacteriol.* 186 (16):5473-5479, 2004.

M. E. Chicurel and D. D. Dalma-Weiszhausz. Microarrays in pharmacogenomics--advances and future promise. *Pharmacogenomics.* 3 (5):589-601, 2002.

H. Claus, A. Friedrich, M. Frosch , U. Vogel. Differential Distribution of Novel Restriction-Modification Systems in Clonal Lineages of Neisseria meningitidis. *Journal of Bacteriology*, Mar. 2000:1296-1303.

H. Claus, J. Stoevesandt, M. Frosch, and U. Vogel. Genetic isolation of meningococci of the electrophoretic type 37 complex. *J.Bacteriol.* 183 (8):2570-2575, 2001.

H. Claus, H. Weinand, M. Frosch, and U. Vogel. Identification of the hypervirulent lineages of Neisseria meningitidis, the ST-8 and ST-11 complexes, by using monoclonal antibodies specific to NmeDI. *J.Clin.Microbiol.* 41 (8):3873-3876, 2003.

H. Claus, H. Swiderek, D. Caugant, J. Elias, M. Frosch, U. Vogel. Specific genetic traits maintained during global spread of the electrophoretic type 15, an emerging clone within the epidemic species Neisseria meningitidis (manuscript in preparation).

B. Davis et al., Chapter: The Genus Neisseria. Microbiology. Third Edition. Harper&Row, Maryland, 1980, 636-641.

G. Dietrich, S. Kurz, C. Hubner, C. Aepinus, S. Theiss, M. Guckenberger, U. Panzner, J. Weber, and M. Frosch. Transcriptome analysis of Neisseria meningitidis during infection. *J.Bacteriol.* 185 (1):155-164, 2003.

N. Dorrell, J. A. Mangan, K. G. Laing, J. Hinds, D. Linton, H. Al Ghusein, B. G. Barrell, J. Parkhill, N. G. Stoker, A. V. Karlyshev, P. D. Butcher, and B. W. Wren. Whole genome comparison of Campylobacter jejuni human isolates using a low-cost microarray reveals extensive genetic diversity. *Genome Res.* 11 (10):1706-1715, 2001.

M. Doumith, C. Cazalet, N. Simoes, L. Frangeul, C. Jacquet, F. Kunst, P. Martin, P. Cossart, P. Glaser, and C. Buchrieser. New aspects regarding evolution and virulence of Listeria monocytogenes revealed by comparative genomics and DNA arrays. *Infect.Immun.* 72 (2):1072-1083, 2004.

D. J. Duggan, M. Bittner, Y. Chen, P. Meltzer, and J. M. Trent. Expression profiling using cDNA microarrays. *Nat.Genet.* 21 (1 Suppl):10-14, 1999.

Epidemiologisches Bulletin, Robert Koch Institute in Berlin, 27/2004:211-214.

E. J. Feil, M. C. Maiden, M. Achtman, and B. G. Spratt. The relative contributions of recombination and mutation to the divergence of clones of Neisseria meningitidis. *Mol.Biol.Evol.* 16 (11):1496-1502, 1999.

E. J. Feil and B. G. Spratt. Recombination and the population structures of bacterial pathogens. *Annu.Rev.Microbiol.* 55:561-590, 2001.

T. R. Gingeras, G. Ghandour, E. Wang, A. Berno, P. M. Small, F. Drobniewski, D. Alland, E. Desmond, M. Holodniy, and J. Drenkow. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic Mycobacterium DNA arrays. *Genome Res.* 8 (5):435-448, 1998.

E. Gotschlich. Chapter 26: Neisseria. Microbiology. Fouth Edition. J. B. Lippincott Company, Philadelphia, 1990, 551-560.

V. Grimm, S. Ezaki, M. Susa, C. Knabbe, R. D. Schmid, and T. T. Bachmann. Use of DNA microarrays for rapid genotyping of TEM beta-lactamases that confer resistance. *J.Clin.Microbiol.* 42 (8):3766-3774, 2004.

J. G. Hacia and F. S. Collins. Mutational analysis using oligonucleotide microarrays. *J.Med.Genet.* 36 (10):730-736, 1999.

J. G. Hacia. Resequencing and mutational analysis using oligonucleotide microarrays. *Nat.Genet.* 21 (1 Suppl):42-47, 1999.

J. Jelfs, R. Munro, F. E. Ashton, and D. A. Caugant. Genetic characterization of a new variant within the ET-37 complex of Neisseria meningitidis associated with outbreaks in various parts of the world. *Epidemiol.Infect.* 125 (2):285-298, 2000.

K. A. Jolley, J. Kalmusova, E. J. Feil, S. Gupta, M. Musilek, P. Kriz, and M. C. Maiden. Carried meningococci in the Czech Republic: a diverse recombining population. *J.Clin.Microbiol.* 38 (12):4492-4498, 2000.

D. A. Kertesz, M. B. Coulthart, J. A. Ryan, W. M. Johnson, and F. E. Ashton. Serogroup B, electrophoretic type 15 Neisseria meningitidis in Canada. *J.Infect.Dis.* 177 (6):1754-1757, 1998.

A. B. Khodursky, B. J. Peter, N. R. Cozzarelli, D. Botstein, P. O. Brown, and C. Yanofsky. DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in Escherichia coli. *Proc.Natl.Acad.Sci.U.S.A* 97 (22):12170-12175, 2000.

C. C. Kim, E. A. Joyce, K. Chan, and S. Falkow. Improved analytical methods for microarray-based genome-composition analysis. *Genome Biol.* 3 (11):RESEARCH0065, 2002.

P. Krizova, M. Musilek. Changing epidemiology of meningococcal invasive disease in the Czech Republic caused by new clone Neisseria meningitidis C:2a: P1.2(P1.5), ET-15/37. *Cent.Eur.J.Public Health* 3:189-194, 1995.

P. Krizova, M. Musilek, J. Kalmusowa. Development of the epidemiological situation in invasive meningococcal disease in the Czech Republic caused by Neisseria meningitidis clone ET-15/37. *Cent.Eur.J.Public Health* 5:214-218, 1997.

S. Kurz, C. Hubner, C. Aepinus, S. Theiss, M. Guckenberger, U. Panzner, J. Weber, M. Frosch, and G. Dietrich. Transcriptome-based antigen identification for Neisseria meningitidis. *Vaccine* 21 (7-8):768-775, 2003.

J. Li, S. Chen, and D. H. Evans. Typing and subtyping influenza virus using DNA microarrays and multiplex reverse transcriptase PCR. *J.Clin.Microbiol.* 39 (2):696-704, 2001.

B. Linz, M. Schenker, P. Zhu, and M. Achtman. Frequent interspecific genetic exchange between commensal Neisseriae and Neisseria meningitidis. *Mol.Microbiol.* 36 (5):1049-1058, 2000.

R. J. Lipshutz, S. P. Fodor, T. R. Gingeras, and D. J. Lockhart. High density synthetic oligonucleotide arrays. *Nat.Genet.* 21 (1 Suppl):20-24, 1999.

A. Loy, A. Lehner, N. Lee, J. Adamczyk, H. Meier, J. Ernst, K. H. Schleifer, and M. Wagner. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl.Environ.Microbiol.* 68 (10):5064-5081, 2002.

S. Lucchini, A. Thompson, and J. C. Hinton. Microarrays for microbiologists. *Microbiology* 147 (Pt 6):1403-1414, 2001.

M. C. Maiden, J. A. Bygraves, E. Feil, G. Morelli, J. E. Russel, R.Urwin, Q. Zhang, J. Zhou, K. Zurth, D. Caugant, I. M. Feavers, M. Achtman, and B. Spratt. Multilocus sequence typing: A portable approach to the identification of clones within population of pathogenic microorganisms. *Microbiology.* Vol. 95, pp. 3140-3145. 1998.

M. C. Maiden, B. Malorny, and M. Achtman. A global gene pool in the neisseriae. *Mol.Microbiol.* 21 (6):1297-1298, 1996.

M. C. Maiden. Horizontal genetic exchange, evolution, and spread of antibiotic resistance in bacteria. *Clin.Infect.Dis.* 27 Suppl 1:S12-S20, 1998.

M. C. Maiden and M. Frosch. Molecular techniques for the investigation of meningococcal disease epidemiology. *Mol.Biotechnol.* 18 (2):119-134, 2001.

J. Meinkoth, G. Wahl. Hybridization of nucleic acids immobilized on solid supports.. 1984 May 1;138("):267-84.

G. Morelli, B. Malorny, K. Muller, A. Seiler, J. F. Wang, J. del Valle, and M. Achtman. Clonal descent and microevolution of Neisseria meningitidis during 30 years of epidemic spread. *Mol.Microbiol.* 25 (6):1047-1064, 1997.

E. Murray, D. Lies, G. Li, K. Nealson, J. Zhou, and J. M. Tiedje. DNA/DNA hybridization to microarrays reveals gene-specific differences between closely related microbial genomes. *Proc.Natl.Acad.Sci.U.S.A* 98 (17):9853-9858, 2001.

J. Parkhill, M. Achtman, K. D. James, S. D. Bentley, C. Churcher, S. R. Klee, G. Morelli, D. Basham, D. Brown, T. Chillingworth, R. M. Davies, P. Davis, K. Devlin, T. Feltwell, N. Hamlin, S. Holroyd, K. Jagels, S. Leather, S. Moule, K. Mungall, M. A. Quail, M. A. Rajandream, K. M. Rutherford, M. Simmonds, J. Skelton, S. Whitehead, B. G.Spratt, B. G. Barrell. Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491. *Nature*. 2000 Mar 30;404(6777):502-6.

C. S. Richmond, J. D. Glasner, R. Mau, H. Jin, and F. R. Blattner. Genome-wide expression profiling in Escherichia coli K-12. *Nucleic Acids Res.* 27 (19):3821-3835, 1999.

N. Salama, K. Guillemin, T. K. McDaniel, G. Sherlock, L. Tompkins, and S. Falkow. A whole-genome microarray reveals genetic diversity among Helicobacter pylori strains. *Proc.Natl.Acad.Sci.U.S.A* 97 (26):14668-14673, 2000.

M. Schena. Microarray Analysis. John Wiley&Sons, Inc. New Jersey, 2003.

M. Schena, D. Shalon, R. W. Davis, and P. O. Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270 (5235):467-470, 1995.

R. K. Selander, D. A. Caugant, H. Ochman, J. M. Musser, M. N. Gilmour, T. Whittam. Methods of Multilocus Enzyme Electrophoresis for Bacterial Population Genetics and Systematics. *Applied and Environmental Microbiology*. May 1986:873-884.

D. Shalon, S. J. Smith, and P. O. Brown. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* 6 (7):639-645, 1996.

J. Small, D. R. Call, F. J. Brockman, T. M. Straub, and D. P. Chandler. Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Appl.Environ.Microbiol.* 67 (10):4708-4716, 2001.

J. M. Smith, N. H. Smith, M. O'Rourke, and B. G. Spratt. How clonal are bacteria? *Proc.Natl.Acad.Sci.U.S.A* 90 (10):4384-4388, 1993.

L. A. Snyder, J. K. Davies, and N. J. Saunders. Microarray genomotyping of key experimental strains of Neisseria gonorrhoeae reveals gene complement diversity and five new neisserial genes associated with Minimal Mobile Elements. *BMC.Genomics* 5 (1):23, 2004.

P. T. Spellman, G. Sherlock, M. Q. Zhang, V. R. Iyer, K. Anders, M. B. Eisen, P. O. Brown, D. Botstein, and B. Futcher. Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. *Mol.Biol.Cell* 9 (12):3273-3297, 1998.

B. G. Spratt and M. C. Maiden. Bacterial population genetics, evolution and epidemiology. *Philos.Trans.R.Soc.Lond B Biol.Sci.* 354 (1384):701-710, 1999.

S. Suerbaum, J. M. Smith, K. Bapumia, G. Morelli, N. H. Smith, E. Kunstmann, I. Dyrek, and M. Achtman. Free recombination within Helicobacter pylori. *Proc.Natl.Acad.Sci.U.S.A* 95 (21):12619-12624, 1998.

S. Suerbaum , M. Achtman. *Population Genetics*. Helicobacter pylori: Physiology and Genetics. Chapter 32:355-361, 2001.

H. Tettelin, N. J. Saunders, J.Heidelberg, A. C. Jeffries, K. E. Nelson, J. A.Eisen, K. A. Ketchum, D. W. Hood, J. F. Peden, R. J. Dodson, W. C. Nelson, M. L. Gwinn, R. DeBoy, J. D. Peterson, E. K. Haft, S. L. Salzberg, O. White, R. D. Fleischmann, B. A. Dougherty, T. Mason, A. Ciecko, D. S. Parksey, E. Blair, H Cittone, E. B. Clark, M. D. Cotton, T. R. Utterback, H. Khouri, H. Qin, J. Vamathevan, J. Gill, V. Scarlato, V. Masignani, M. Pizza, G. Grandi, L. Sun, H. O. Smith, C. M. Fraser, E. R. Moxon, R. Rappuoli, J. C. Venter. Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. *Science*. 2000 Mar 10;287(5459):1809-15.

E. Tikhomirov, M. Santamaria, and K. Esteves. Meningococcal disease: public health burden and control. *World Health Stat.Q.* 50 (3-4):170-177, 1997.

A. Troesch, H. Nguyen, C. G. Miyada, S. Desvarenne, T. R. Gingeras, P. M. Kaplan, P. Cros, and C. Mabilat. Mycobacterium species identification and rifampin resistance testing with high-density DNA probe arrays. *J.Clin.Microbiol.* 37 (1):49-55, 1999.

R. Urwin and M. C. Maiden. Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol.* 11 (10):479-487, 2003.

A. van Belkum, M. Struelens, A. de Visser, H. Verbrugh, M. Tibayrenc. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology. *Clin. Microbiol. Rev.* Jul; 14(3):547-60, 2001.

W. B. van Leeuwen, C. Jay, S. Snijders, N. Durin, B. Lacroix, H. A. Verbrugh, M. C. Enright, A. Troesch, and A. van Belkum. Multilocus sequence typing of Staphylococcus aureus with DNA array technology. *J.Clin.Microbiol.* 41 (7):3323-3326, 2003.

J. A. Vazquez, S. Berron, M. O'Rourke, G. Carpenter, E. Feil, N. H. Smith, and B. G. Spratt. Interspecies recombination in nature: a meningococcus that has acquired a gonococcal PIB porin. *Mol.Microbiol.* 15 (6):1001-1007, 1995.

U. Vogel, H. Claus, M. Frosch, and D. A. Caugant. Molecular basis for distinction of the ET-15 clone within the ET-37 complex of Neisseria meningitidis. *J.Clin.Microbiol.* 38 (2):941-942, 2000.

U. Vogel, H. Claus, and M. Frosch. Genetic lineages and their traits in Neisseria meningitidis. *Int.J.Med.Microbiol.* 294 (2-3):75-82, 2004.

D. G. Wang, J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, E. S. Lander, and . Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 280 (5366):1077-1082, 1998.

J. F. Wang, D. A. Caugant, X. Li, X. Hu, J. T. Poolman, B. A. Crowe, and M. Achtman. Clonal and antigenic analysis of serogroup A Neisseria meningitidis with particular reference to epidemiological features of epidemic meningitis in the People's Republic of China. *Infect.Immun.* 60 (12):5267-5282, 1992.

J. F. Wang, D. A. Caugant, G. Morelli, B. Koumare, and M. Achtman. Antigenic and epidemiologic properties of the ET-37 complex of Neisseria meningitidis. *J.Infect.Dis.* 167 (6):1320-1329, 1993.

E. A. Winzeler, D. R. Richards, A. R. Conway, A. L. Goldstein, S. Kalman, M. J. McCullough, J. H. McCusker, D. A. Stevens, L. Wodicka, D. J. Lockhart, and R. W. Davis. Direct allelic variation scanning of the yeast genome. *Science* 281 (5380):1194-1197, 1998.

L. Zhang, U. Srinivasan, C. F. Marrs, D. Ghosh, J. R. Gilsdorf, B. Foxman. *BMC Microbiol.* 2004 Mar 22;4(1):12.
J. Zhou, L. D. Bowler, and B. G. Spratt. Interspecies recombination, and phylogenetic distortions, within the glutamine synthetase and shikimate dehydrogenase genes of Neisseria meningitidis and commensal Neisseria species. *Mol.Microbiol.* 23 (4):799-812, 1997.

P. Zhu, Ende A. van der, D. Falush, N. Brieske, G. Morelli, B. Linz, T. Popovic, I. G. Schuurman, R. A. Adegbola, K. Zurth, S. Gagneux, A. E. Platonov, J. Y. Riou, D. A. Caugant, P. Nicolas, and M. Achtman. Fit genotypes and escape variants of subgroup III Neisseria meningitidis during three pandemics of epidemic meningitis. *Proc.Natl.Acad.Sci.U.S.A* 98 (9):5234-5239, 2001.

8. Abbreviations

et al.	and others (et alii)
bp	base pair
BSA	bovine serum albumine
cDNA	complementary DNA
dATP	2'-deoxyadenosine 5'-triphosphate
dCTP	2'-deoxycytidine 5'-triphosphate
ddH$_2$O	double distilled H$_2$O
DEPC	diethyl pyrocarbonate
dGTP	2´-deoxyguanosine 5´-triphosphate
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dTTP	2'-deoxythymidine 5'-triphosphate
E.coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
e.g.	for example
ET	electrophoretic type
Fig.	figure
i.e.	that is to say (it means)
kb	kilo base
m	mili
M	molar
min	minute
MLEE	multilocus enzyme electrophoresis
MLST	multilocus sequence typing
MOPS	3-(N-Morpholino)propanesulfonic acid
M. tuberculosis	*Mycobacterium tuberculosis*
n	nano
N.	*Neisseria*
N. meningitidis	*Neisseria meningitidis*
nt	nucleotide
ORF	open reading frame
OD$_{xxx}$	optical density
p	pico
PCR	polymerase chain reaction
rpm	resolutions per minute
primer	oligonucleotide
pH	the negative logarithm (base 10) of the H$_3$O+ concentration
rDNA	ribosomal deoxyribonucleic acid
RNA	ribonucleic acid
RNase	ribonuclease
rRNA	ribosomal ribonucleic acid
RT	room temperature
SNP	single nucleotide polymorphism
spp.	species
ST	sequence type
Taq	*Thermus aquaticus*
Tab.	table
TBE	Tris–Borate–EDTA
TE	Tris-HCl/EDTA
Tm	melting temperature

Tris	tris(hydroxymethyl)-aminomethan
μ	micro
μm	micrometer
vs.	versus

Lebenslauf:

Name: Halina Świderek

geboren am   1.7.1966 in Tomaszów Lub., Polen

Schul- und Universitätsbildung:

1973-1981   Grundschule in Tomaszów Lub., Polen

1981-1985   B. Glowacki Gymnasium in Tomaszów Lub., Polen

1985   Abitur

1985-1990   Studium der Biologie an der Marie Curie-Sklodowska-Universität in
                               Lublin, Polen
                               Schwerpunktfach Mikrobiologie

12.7.1990   Abschluss mit dem Titel Master of Science,
                               Masterarbeit zum Thema "Concurrence of marked strains of
                               Rhisobium trifolia in infecting clover".

Berufliche Tätigkeiten:

1990 - 6/2001   Lehrerin für Medizinische Mikrobiologie und Parasitologie am
                               Medical Training College in Lublin, Polen

8.1. - 29.3.2001   FEMS-Stipendium an der Julius-Maximilians-Universität
                               Würzburg, Institut für Hygiene und Mikrobiologie

seit 7/2001   Doktorandenstelle (BAT IIa/2) am Institut für Hygiene und
                               Mikrobiologie der Julius-Maximilians-Universität Würzburg
                               gefördert durch das EU-Projekt "MenNet-EU" ("Impact of
                               meningococcal epidemiology and population biology on public
                               health in Europe"; Projektreferenznummer: QLK2-CT-2001-01436).

Publikationen von Daten aus dieser Arbeit:

H. Swiderek, H. Claus, M. Frosch, U. Vogel. Evaluation of custom-made DNA microarrays for multilocus sequence typing of Neisseria meningitidis. *International Journal of Medical Microbiology* (available online since 20.01.2005).

H. Claus, H. Swiderek, D. Caugant, J. Elias, M. Frosch, U. Vogel. Specific traits maintained during global spread of the electrophoretic type 15, an emerging clone within the epidemic species Neisseria meningitidis (manuscript in preparation).

Posters and abstracts

Claus H., Swiderek H., Frosch M., Vogel U. (2002). Menigococcal sequence typing using microarrays. Poster presentation at the 13th International Pathogenetic *Neisseria* Conference in Oslo, Norway.

Swiderek H., Vogel U.,Frosch M., Claus H. (2003). Application of microarray technology to meningococcal sequence typing. Abstract at the meeting of the 55th DGHM in Dresden, Germany.

Claus H., Swiderek H., Elias J., Frosch M. and Vogel U. (2004). Identification of ET-15 meningococci. Poster presentation at the meeting of the 56th DGHM in Münster, Germany.