

Design and application of a DNA microarray for the identification of intestinal pathogens during gastroenteritis and monitoring of the resident intestinal microbiota

Von der Fakultät Energie-, Verfahrens- und Biotechnik der Universität Stuttgart zur
Erlangung der Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

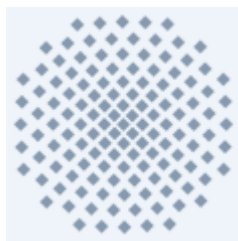
vorgelegt von

Kristina Hänel

aus Potsdam

Hauptberichter: Prof. Dr. Rolf D. Schmid
Mitberichter: PD Dr. Till T. Bachmann
Prüfungsvorsitzender: Prof. Dr. Peter Scheurich

Tag der mündlichen Prüfung: 16.02.2012



Erklärung/ Declaration

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und nur unter Zuhilfenahme der angegebenen Hilfsmittel verfasst habe.

Herewith I declare that the presented work was accomplished independently and only by the use of the specified resources and literature.

Stuttgart, November 2010

Kristina Hänel

Table of contents

| | |
|--|-----------|
| Table of contents | 3 |
| Preface | 5 |
| Summary | 6 |
| Deutsche Zusammenfassung | 8 |
| 1 Introduction | 19 |
| 1.1 The digestive tract..... | 19 |
| 1.2 Molecular techniques to monitor the intestinal community | 21 |
| 1.3 Bacterial gastroenteritis | 25 |
| 1.4 Diagnostic methods for clinical pathogen identification | 30 |
| 1.4.1 Phenotypic pathogen identification | 30 |
| 1.4.2 Genotypic pathogen identification..... | 31 |
| 1.5 DNA-microarrays | 33 |
| 1.2 Objectives | 38 |
| 2 Material and methods | 39 |
| 2.1 Material | 39 |
| 2.1.1 Reagents and kits | 39 |
| 2.1.2 Technical equipment..... | 40 |
| 2.1.3 Biological material..... | 41 |
| 2.1.4 Primers..... | 42 |
| 2.2 Methods..... | 44 |
| 2.2.1 Capture probe design | 44 |
| 2.2.2 Oligonucleotide array fabrication | 45 |
| 2.2.3 DNA extraction..... | 48 |
| 2.2.4 DNA target amplification | 49 |
| 2.2.5 Purification of amplified DNA and quality control..... | 50 |
| 2.2.6 DNase I digestion of amplified DNA | 51 |
| 2.2.7 Hybridization | 51 |
| 2.2.8 Silver staining..... | 51 |
| 2.2.9 Data acquisition and processing..... | 51 |
| 2.2.10 Limit of detection (LOD) determination..... | 52 |
| 2.2.11 Sanger DNA sequencing..... | 52 |
| 2.2.12 Denaturing gradient gel electrophoresis (DGGE) | 53 |
| 2.2.13 Multivariate Analysis..... | 54 |
| 2.2.14 DNA target quantification on DNA microarrays..... | 55 |
| 3 Results | 56 |
| 3.1 Design of the identification microarray..... | 56 |
| 3.2 Array preparation – theoretical considerations | 61 |
| 3.3 Extraction of DNA reference material..... | 62 |
| 3.4 Amplification of the DNA material | 65 |
| 3.5 Influence of DNA digestion on hybridization..... | 70 |
| 3.6 Influence of relative probe position in the target on hybridization..... | 72 |
| 3.7 Probe verification with reference strains..... | 75 |
| 3.8 Sanger sequencing of the reference strains | 82 |
| 3.9 Array redesign | 85 |
| 3.10 Limit of detection of the DNA microarray | 91 |

| | |
|---|------------|
| 3.11 Application of the identification microarray | 93 |
| 3.11.1 Clinical faecal isolates from gastroenteritis patients and healthy individuals | 93 |
| 3.11.2 Clinical faecal samples from rotavirus-infected and healthy children | 100 |
| 3.11.3 Establishment of human and pig intestinal flora in piglets | 103 |
| 3.12 DNA target quantification with DNA microarrays | 110 |
| 3.12.1 Hybridization of PCR product to E. coli-specific probes with a quencher | 111 |
| 4 Discussion | 117 |
| 4.1 Genotyping of intestinal pathogens | 117 |
| 4.2 Design and performance of the Gastroenteritis-Chip | 119 |
| 4.2.1 Choice of the target gene | 120 |
| 4.2.2 Target amplification..... | 121 |
| 4.2.3 Steric and kinetic effects on the microarray surface..... | 122 |
| 4.2.4 Probe design..... | 125 |
| 4.2.5 Probe verification and redesign | 125 |
| 4.2.6 Limit of detection of the Gastroenteritis-Chip | 127 |
| 4.3 Application of the Gastroenteritis-Chip | 129 |
| 4.3.1 Clinical samples from gastroenteritis patients and healthy volunteers | 130 |
| 4.3.2 Children with rotavirus infection and healthy individuals | 132 |
| 4.3.3 HFA and PFA piglets | 134 |
| 4.4 DNA target quantification with DNA microarrays | 135 |
| 4.5 Market relevance and aspects of future development | 137 |
| 4.6 Conclusion | 139 |
| 5 Literature..... | 141 |
| Acknowledgment..... | 156 |
| Abbreviations and international units..... | 157 |
| Curriculum vitae | 159 |

Preface

This work was part of the BMBF (German Federal Ministry of Education and Research) funded project “Genome Research on Microorganisms” (PathoGenoMik) and was accomplished in cooperation with the Institute for Medical Microbiology at the University Hospital of Giessen (Dr. Domann, Prof. Chakraborty) and the industrial partner Eppendorf AG. The University Hospital of Giessen provided reference material and information on clinical relevant pathogens. Eppendorf AG supported the project with laboratory equipment and application-oriented professional advice.

Parts of this work, namely the investigation of faecal samples from piglets and from healthy and rotavirus-infected children with microarray and DGGE, were carried out at the School of Life Sciences at Shanghai Jiao Tong University in the group of Prof. Liping Zhao during a five-month research stay.

From November, 2004 until November, 2006, and from February, 2007 until January, 2008 this scientific work was kindly supported by a scholarship from the Federal State of Baden-Württemberg (Landesgraduiertenförderung).

Summary

Bacterial gastroenteritis is a commonly occurring disorder with three to five billion cases of acute diarrhea annually worldwide, particularly in developing countries where it is a leading cause of childhood morbidity and mortality. Diarrheal diseases can quickly reach epidemic dimensions, especially in the case of food-borne pathogens. Fast identification of the etiologic agent remains a key task in clinical diagnosis. This is of major concern regarding the choice of therapy, the clinical hygiene management, the patient's confidence in the clinician, and acceptance of and satisfaction with the medical treatment. Currently, it is mainly based on phenotypic classification using stool culture and a set of morphological, physiological, and serological tests. However, cultivation of the pathogen, which can often be challenging in the case of intestinal bacteria, is a complex and time-consuming task, which does not meet the requirements of an evidence-based therapy and a cost effective medical care system. Therefore, genotyping methods would be highly appreciated to replace classical laboratory methods.

Microarrays can fulfil the need for a rapid diagnostic tool that generates reliable information with a minimum of laboratory effort and do not rely on cultivation of the organisms. They enable parallel identification of many species, which would clearly facilitate the search for the etiologic agent. Moreover, microarrays can be integrated into small, microfluidic devices, which could be applied in point-of-care diagnostics.

Here, a diagnostic oligonucleotide microarray was developed to detect the most common bacterial species associated with gastroenteritis. As a new concept, the array aimed at also providing information about the residential flora and potential probiotics. This combination allows not only pathogen identification but also a monitoring of the patients gut flora recovery during therapy. The developed array comprised probes against the genes coding for the small and large ribosomal RNA for 23 species and 11 genera following a multiple-probe concept. This comprised 13 common bacterial gastroenteritis-related pathogens, namely *Aeromonas hydrophila*, *Campylobacter jejuni*, *Clostridium difficile*, *Mycobacterium avium* complex, *Mycobacterium avium* subsp. *paratuberculosis*, *Plesiomonas shigelloides*, *Vibrio cholerae*, *Vibrio parahaemolyticus*, *Yersinia enterocolitica*, *Listeria monocytogenes*, *Salmonella enterica* subsp. *enterica* serovar *Enteritidis*, *Yersinia enterocolitica*, and *E. coli* O157:H7 (EHEC). The DNA array performance was initially verified with pure cultures from clinical isolates and DSM strains and confirmed by standard Sanger sequencing. The final array allowed unambiguous identification of the target species by combining the multiple-probe concept with a robust cut-off using the negative hybridization control plus three times its standard deviation and by accepting only positive signals below 50% standard deviation. In most cases, the obligatory detection on two taxonomic levels provided additional robustness. The limit of detection was determined to be 10^3 genome equivalents. The whole assay took 5.5 hours after DNA extraction from a clinical sample.

The clinical applicability was examined by processing stool samples from 58 patients with gastroenteritic symptoms and six healthy individuals using the diagnostic microarray. For 30 samples where the etiologic pathogen had been detected by a combination of culture and specific real-time PCR assays, 67% of these samples were correctly identified. It was assumed that a higher recovery rate would require a lower detection limit of the array, which might be reached by using more sensitive labels, array surfaces with higher probe capacity, pathogen enrichment strategies, pre-hybridization steps in solution, reduction of the reaction volumes by microfluidic systems, and RNA as the target for detection. Furthermore, highly individualized communities of the detected resident bacteria were found in the analysed samples. A difference of this microbiota between healthy and infected subjects was not observed by multivariate analysis of the array data. This was attributed to the inhomogeneous sample group in terms of patient age and clinical diagnose.

Additionally, the array was applied to investigate the influence of a viral, intestinal infection on the composition of the infantile microbiota and the establishment of human and porcine

intestinal community in inoculated piglets. In both trials, also bacterial pathogens could be identified. Multivariate analysis of the microarray data revealed a significant difference of the resident flora between healthy and infected children and between the pig and human intestinal flora. Partial least squares analysis and one-way analysis of variance allowed identification of resident species or genera, which significantly differed between these groups. According to literature survey, this was the first application of principle component and partial least squares analysis for statistical evaluation of identification microarray data. The results indicated the applicability of the present microarray for studying the resident microbiota under a certain question. Nevertheless, it was concluded that more species should be included in the array to allow in-depth analysis of faecal microbiota.

To address the problem of reliable quantification of intestinal bacteria and pathogens with a concentration range over several orders of magnitude using fluorescently labelled microarrays, a FRET-based system was developed. A black hole quencher was used to gradually quench the Cyanine 3 reporter signal allowing detection of a target species in the linear scanner range independent from its initial concentration. The proof of concept based on *E. coli* and its specific probes showed that an effective quenching of the fluorophore label in the target was possible upon hybridization and dependent on the amount of quencher immobilized in the spot. It was concluded that two spots, one without quencher and one with 8 μM probe-coupled quencher, would be sufficient to detect 10^4 up to $2 \cdot 10^6$ genome equivalents in the linear range of the scanner. This setup partly overcomes the technical limitation of linear fluorescence signal acquisition for different amounts of bacteria. The upper limit represents the maximum applied DNA amount in PCR.

In summary, in this study solutions for two problems of molecular pathogen detection in the clinic were developed: (1) the time for pathogen detection and (2) the assay coverage. The analytical sensitivity for clinical application was not yet sufficient due to limitations of the pre-analytical procedures. A general concept to expand the linear range of a fluorescence intensity based detection system was achieved by a FRET-based system with immobilized quencher on the array surface.

Deutsche Zusammenfassung

Einführung

Bakterielle Gastroenteritis ist mit weltweit 3-5 Millionen akuten Fällen pro Jahr eine verbreitete Erkrankung. Als ein Hauptgrund für Kindersterblichkeit betrifft sie besonders stark Entwicklungsländer, ist aber auch in Industrieländern von großer Bedeutung, da sie hier für erhebliche Arbeitsausfälle sorgt und durch Auftreten von epidemieartigen Ausbrüchen in Kindergärten, Kantinen, Krankenhäusern oder Altersheimen für Kinder oder ältere bzw. geschwächte Personen auch lebensbedrohliche Ausmaße annehmen kann. Diarrhöe kann insbesondere im Falle von Lebensmittelvergiftungen relativ schnell die Dimension einer Epidemie erlangen. Die schnelle Identifizierung des auslösenden Pathogens bleibt nach wie vor die Kernaufgabe in der klinischen Diagnostik dieser Erkrankung. Dies ist von besonderer Bedeutung im Hinblick auf die Therapiewahl, das klinische Hygienemanagement, das Vertrauen des Patienten zum behandelnden Arzt und dessen Akzeptanz und Zufriedenheit mit der medizinischen Behandlung. Gegenwärtig basiert die Diagnostik überwiegend auf der phänotypischen Klassifikation von Stuhlkulturen unter Verwendung von morphologischen, physiologischen und serologischen Tests. Die Pathogenkultivierung, die bei intestinalen Bakterien schwierig sein kann, ist jedoch eine komplexe und zeitaufwändige Prozedur, die nicht den Anforderungen einer wissenschaftsbasierten Therapie und eines kosteneffizienten Gesundheitssystems gerecht wird. Daher wären genotypische Methoden hochwillkommen um klassische labor diagnostische Nachweisverfahren zu verbessern.

Mikroarrays können den Wunsch nach einem schnellen, diagnostischen Werkzeug erfüllen, das vertrauenswürdige Informationen mit minimalem Laboraufwand und unabhängig von der Kultivierung der Organismen liefert. Sie erlauben eine parallele Identifikation von vielen Spezies, was die Suche nach dem pathogenen Agens deutlich vereinfachen würde. Darüber hinaus gibt es bei Mikroarrays die Möglichkeit sie in miniaturisierte, mikrofluidische Systeme zu integrieren, die in patientennahen Diagnostika Anwendung finden können.

In der vorliegenden Arbeit wurde ein diagnostischer Oligonukleotid-Mikroarray für die Detektion der verbreitetsten bakteriellen Gastroenteritiserreger entwickelt und getestet. Das Ziel mit dem Array auch Informationen über die residente Darmflora sowie potentiell probiotische Bakterien zu generieren, stellt ein neues Konzept für diagnostische DNA-*chips* dar. Diese Kombination erlaubt nicht nur die Pathogenidentifizierung sondern auch die Beobachtung der Beeinträchtigung sowie Regeneration der intestinalen mikrobiellen Gemeinschaft während der Therapie.

Ergebnisse

Auswahl der Zielspezies

Zunächst wurde eine Auswahl der zu detektierenden Erreger und residenten Darmflora getroffen. Dies erfolgte auf Basis von Literaturrecherchen sowie Informationen zur klinischen Relevanz der Erreger bzw. Häufigkeit der Bakterien im Darmtrakt. Bei den Pathogenen wurde eine möglichst umfassende Abdeckung aller bakteriellen Gastroenteritiserreger angestrebt. Als Zielgene für den Nachweis der Bakterien wurden die Gene der ribosomalen RNA ausgewählt, die ubiquitär in Bakterien vorkommen und deren Sequenzen von zahlreichen Bakterien bereits in vielfacher Abdeckung vorliegen. Zudem enthalten diese Gene neben spezifischen Regionen, die für die Identifikation relevant sind, auch konservierte Bereiche, welche für die DNA-Vervielfältigung mit universellen Primern unverzichtbar sind. Da das RNA-Gen der kleinen ribosomalen Untereinheit (SSU, 16S rRNA-Gen) nicht für die Differenzierung aller Zielarten ausreichte, wie in-silico-Untersuchungen mit der ARB Software vorab zeigten, wurde auch das RNA-Gen der großen ribosomalen Untereinheit (LSU, 23S rRNA-Gen) ins Sondendesign einbezogen.

Sondendesign

Beim Sondendesign wurde ein Mehrsonden-Konzept verfolgt, das den Versuch beschreibt, möglichst zwei oder mehr Sonden auf Art- bzw. Gattungsebene zu entwickeln, um eine ausreichende Robustheit des Assays zu erzielen. Aufgrund der eingeschränkten Verfügbarkeit spezifischer Regionen, die ihre Ursache in der komplexen Zusammensetzung der intestinalen Gemeinschaft hat, war dies jedoch nicht in allen Fällen möglich. Das Sondendesign war in einigen Fällen auch durch nicht Verfügbarkeit von Sequenzdaten bzw. identischen Sequenzen limitiert. Letzteres wurde bei *Shigella* spp. und *E. coli* im 16S rRNA Gen beobachtet. Die Abdeckung und Spezifität der Sonden wurde in-silico (ARB Datenbank) überprüft. Auf diese Weise konnte die Qualität der Sonden hinsichtlich potentieller Kreuzreaktivität mit anderen Vertretern der komplexen, intestinalen Gemeinschaft bereits vorab weitgehend beurteilt werden. In einigen Fällen war es jedoch aufgrund fehlender Sequenzdaten mehrerer Vertreter einer Gattung nicht möglich zu entscheiden, ob es sich bei Sonden um Spezies-spezifische oder Gattungs-spezifische Sonden handelte. Die Abdeckung der Sonden betrug in den meisten Fällen 100% der in-silico verfügbaren Sequenzen der zu detektierenden Spezies bzw. Gattungen. Mithilfe des OligoAnalyzer (Webresource) konnten die thermodynamischen Parameter der Sonden auf Basis der *nearest neighbour*-Methode berechnet und weitestgehend aufeinander abgestimmt werden. Die Schmelztemperatur der Sonden lag in einem Bereich von 48,2 bis 59,5°C, wobei 70 von 87 Sonden dem vorher definierten Bereich $T_M > 50^\circ\text{C}$ und $< 56^\circ\text{C}$ entsprachen. Für die Stabilität von Haarnadelkonformationen wurde vorab ein Akzeptanzbereich von $\geq 0 \text{ kcal}\cdot\text{mol}^{-1}$ definiert, dem 73 von 87 Sonden entsprachen. Hinsichtlich der Stabilität von Selbst-Dimeren wurde ein Bereich $\geq -7 \text{ kcal}\cdot\text{mol}^{-1}$ als tolerierbar definiert, dem 72 Sonden gerecht wurden. Die höchste Haarnadelstabilität wurde für eine *Vibrio parahaemolyticus*-Sonde mit $-1.17 \text{ kcal}\cdot\text{mol}^{-1}$, aufgrund nur zweier benachbarter GC-Paare, gefunden. Die höchste Selbst-Dimerisierung mit einer Stabilität von $-16.03 \text{ kcal}\cdot\text{mol}^{-1}$ einer *Shigella sonnei*-Sonde basierte auf sechs benachbarten GC-Paaren. Der GC-Gehalt der Sonden variierte zwischen 32% und 66,7%.

DNA Extraktion aus Reinkulturen

Zur Überprüfung der Funktionalität der Sonden musste Referenz-DNA aus bakteriellen Reinkulturen extrahiert werden. Dies erfolgte durch Phenol-Chloroform-Extraktion oder bei unseren klinischen Partnern mithilfe eines Extraktions-Kits. Die DNA-Konzentration variierte zwischen 33 und 413 ng/ μL . Eine Überprüfung der DNA-Qualität mittels Agarosegelelektrophorese offenbarte eine starke Degradation der DNA in den klinischen Extrakten, die mit dem Extraktionskit gewonnen wurden. Dies hatte Einfluss auf die nachfolgend entwickelte Amplifikationsstrategie, denn eine Anpassung an die klinischen Gegebenheiten des Probenmaterials war unerlässlich für die geplante Applikation des Arrays.

Entwicklung einer Amplifikationsstrategie

Für die Detektion auf einem Mikroarray ist eine Vervielfältigung des genetischen Materials notwendig, um eine ausreichende Sensitivität zu erreichen. Daher wurde im Folgenden eine Amplifikationsstrategie mittels Polymerasekettenreaktion (PCR) entwickelt, die alle Genbereiche vervielfältigen sollte, die mit Sonden zu detektieren waren. Da die entwickelten Sonden über beide ribosomalen RNA-Gene (SSU, LSU) verteilt waren, mussten diese demnach in voller Länge amplifiziert werden. Zudem sollte die Vervielfältigung der Ziel-DNA für alle Zielspezies universell sein. Die ersten Versuche, beide Gene mit einem oder zwei Primerpaaren zu amplifizieren, scheiterten am hohen Grad der Fragmentierung innerhalb der DNA-Extrakte. Folglich wurde ein Satz von fünf Primerpaaren entwickelt, der universell für alle Zielspezies war und maximal drei degenerierte Nukleotide pro Primer enthielt. Dabei umspannten zwei Primerpaare das 16S rRNA-Gen und drei das 23S rRNA-Gen in nahezu kompletter Länge. Die erhaltenen PCR-Produkte hatten eine Länge von 794, 616 bp (SSU-

Gen), 1.000, 850, und 736 bp (LSU-Gen) bei *E. coli*, wobei die Längen zwischen den verschiedenen Arten leicht variierten. Die Amplifikationsstrategie wurde mit 31 Referenz-DNA-Extrakten von Ziel- und Nicht-Zielspezies validiert. Die Vervielfältigung der fünf DNA-Bereiche war für alle Spezies erfolgreich, abgesehen von Fragment F bzw. DE von *Veillonella parvula* und *Atopobium minutum*, die jedoch für deren Detektion keine Rolle spielten. Die Amplifikationseffizienz der verschiedenen Fragmente variierte um den Faktor fünf. Die entwickelte PCR war demnach in dieser Form für den Assay anwendbar.

Zusätzlich wurde auch eine Multiplex-PCR mit den fünf Primerpaaren getestet und mit den 31 Referenz-Extrakten durchgeführt. Die Analyse der Amplifikate mittels Kapillargel-elektrophorese zeigte jedoch, dass die Effizienz der Vervielfältigung für die verschiedenen Spezies und Fragmente sehr unterschiedlich war. Teilweise konnten DNA-Abschnitte nicht oder nur schwach detektiert werden. Insbesondere das lange DE-Fragment zeigte eine schlechte Amplifikationseffizienz. Dies wurde auf die starke Überlappung der Primer benachbarter Fragmente zurück geführt, deren Heterodimere teilweise hohe Stabilitäten von $-14,07$ bis $-21,78$ kcal* mol^{-1} erreichten. Als Konsequenz wurde die Singleplex-PCR beibehalten, die durch die hohe Informationsdichte des nachgeschalteten Arrays ihre Berechtigung behält.

Optimierung der Hybridisierungsparameter

Um ein optimales Hybridisierungsergebnis zu erhalten müssen in der Regel die Hybridisierungsparameter optimiert werden. In dieser Arbeit konnte hinsichtlich der Hybridisierungstemperatur und –zeit auf Erfahrungswerte aus der Arbeitsgruppe zurück gegriffen werden. Dennoch musste der Einfluss des Ziel-DNA-Verdau auf das Hybridisierungsergebnis experimentell überprüft werden. Anhand von vier Referenz-DNA-Isolaten von *E. coli*, *R. intestinalis*, *P. shigelloides* und *C. coli*, konnte exemplarisch gezeigt werden, dass der partielle Verdau der amplifizierten DNA-Abschnitte vor der Hybridisierung in der Regel zu höheren absoluten Fluoreszenzintensitäten führte. Der Verdau mit DNase I wurde hinsichtlich der einzusetzenden Enzymkonzentration optimiert, wobei eine Fragmentgröße von 50-100 nt angestrebt wurde. Die optimale DNase I-Konzentration betrug dafür 0,4 mU/ng DNA. Die Hybridisierungssignale von den 26 analysierten Sonden gegen die oben genannten vier Spezies zeigten nach PCR-Produkt-Verdau durchschnittlich 1,8-fach höhere Intensitäten als ohne Verdau.

Einfluss der Sondenposition im Zielfragment auf das Hybridisierungssignal

Anhand der vier benannten Spezies, *E. coli*, *R. intestinalis*, *P. shigelloides* und *C. coli*, wurde auch der Einfluss der relativen Sondenposition in der Ziel-DNA auf das Hybridisierungsergebnis untersucht. Es wurde die Hybridisierung mit unverdauter Ziel-DNA analysiert. Zunächst wurde die relative Position jeder Sonde im Ziel-DNA-Strang auf Basis der Fragmentlänge sowie der Startposition der Sonde berechnet und als Prozent vom Beginn des dem Zielfragment komplementären Stranges (5'→3') ausgedrückt. Eine *sense*-Sonde war dabei sequenzidentisch mit dem *sense*-Strang und hybridisierte mit dem *antisense*-Strang. Entgegengesetztes galt für *antisense*-Sonden. Außerdem wurden die theoretischen Hybridisierungskonformationen für eher 3'- bzw. eher 5'-lokalisierte Sonden im Hinblick auf die Ausrichtung des freien Überhangs des Zielfragments aufgezeigt. Es war zu erwarten, dass 5'-lokalisierte Sonden, die mit dem 3'-Ende des Zielfragments hybridisieren, zu einem in die Lösung gerichteten Überhang der Ziel-DNA führen. Dagegen sollten 3'-lokalisierte Sonden, die das 5'-Ende der Ziel-DNA binden, zu einem in Richtung Chipoberfläche gerichteten Überhang führen. Die Analyse der Hybridisierungssignale zeigte mit unverdauten Ziel-DNA-Fragmenten keinen ausgeprägten Trend zu stärkeren Signalen bei der Sondenposition am einen oder anderen Fragmentende. Lediglich zwei deutlich 3'-lokalisierte Sonden erreichten höhere Hybridisierungssignale als die anderen Sonden. Bei zwei weiteren 3'-lokalisierten Sonden war dies aber nicht zu beobachten. Demnach war die

Ausrichtung des Fragmentüberhangs zur Chipoberfläche oder in die Lösung relativ unwichtig für die Höhe der Hybridisierungssignale bzw. hatte nur einen marginalen Einfluss. Die Untersuchung zeigte dementsprechend ganz deutlich, dass ein zur Chipoberfläche gerichteter Überhang des hybridisierenden DNA-Fragments beim hier bestehenden Arraydesign keinen negativen Einfluss auf die Fluoreszenzsignale hatte. Es war zu erwarten, dass ein DNase-Verdau der Zielfragmente zu einer noch stärkeren Angleichung der Fluoreszenzsignale führen würde und auch die leichte Tendenz zu stärkeren Intensitäten bei den zwei 3'-lokalisierten Sonden nivellieren würde. Dies wurde auch so beobachtet. Die meisten Sonden wiesen deutlich erhöhte Fluoreszenzsignale auf, während die drei 3'-lokalisierten Sonden mit den ursprünglich höchsten Fluoreszenzintensitäten nun geringere Intensitäten hatten. Die beobachteten, unterschiedlichen Hybridisierungssignale von korrespondierenden *sense*- und *antisense*-Sonden konnten auf diese Weise nicht erklärt werden. Es wurde vermutet, dass hier andere Effekte, wie Sekundärstrukturen und die individuelle Basenabfolge in und neben der Zielsequenz, einen vergleichsweise größeren Einfluss auf die Hybridisierungssignale ausüben, als die relative Position der Sonde im Zielfragment.

Verifizierung der Sonden mit DNA-Extrakten aus Reinkulturen von Referenzspezies

Die Funktionalität der entwickelten Sonden wurde mit isolierter DNA aus Reinkulturen (DSM-Stämme und klinische Isolate) überprüft. Es wurden 32 Spezies, darunter Ziel- und Nichtzielspezies, für die Sondenverifizierung herangezogen. Aufgrund des Fehlens von Referenzmaterial für *Mycobacterium avium* Complex, *Vibrio cholerae*, *Eubacterium bifforme*, *Fusobacterium prausnitzii*, *Acholeplasma laidlawii* und *Lactobacillus acidophilus* konnten deren Spezies-spezifische Sonden nicht mit DNA aus Reinkulturen verifiziert werden. Die Hybridisierungsversuche mit den vorhandenen und amplifizierten DNA-Extrakten ergaben individuelle Hybridisierungsmuster für jede Spezies, was eine eindeutige Identifizierung ermöglichte. Dennoch wiesen einige Sonden trotz der stringenten Bedingungen beim Design Kreuzhybridisierungen auf. Solche Kreuzhybridisierungen waren aufgrund der komplexen Zusammensetzung des Probenmaterials nicht akzeptabel, denn es bestand die Gefahr, dass es bei Realproben zu Akkumulationen von falschpositiven Signalen hätte kommen können. Zur besseren Beurteilung dieser Kreuzhybridisierungen und um ein Sonden-Redesign zu ermöglichen, wurden die 16S und 23S rRNA-Gene der Referenz-DNA-Extrakte mit dem Standard-Sanger-Verfahren sequenziert. Die Sequenzierung diente zudem der Bestätigung der Hybridisierungsergebnisse.

Sequenzierung

Die Sequenzierung des 16S rRNA-Gens konnte für alle Spezies außer *Mycobacterium avium* subsp. *paratuberculosis* (MAP) erfolgreich durchgeführt werden. Im Falle von MAP wurde nur das Fragment J auswertbar sequenziert. Das längere 23S rRNA-Gen wurde bei 25 Spezies vollständig sequenziert. Im Falle von *Y. enterocolitica* und *C. haemolyticum* fehlte dagegen der erste Abschnitt des DE-Fragments. Bei den Spezies MAP, *A. minutum* und *B. fragilis* konnte das gesamte DE-Fragment nicht sequenziert werden. Bei EHEC fehlte das Ende des 23S rRNA-Gens. Die Sequenzen wurden bei GenBank hinterlegt.

Mithilfe einer BLAST (basic local alignment search tool) Analyse wurden Spezies mit größtmöglicher Sequenzhomologie in der GenBank-Datenbank identifiziert. Dadurch konnten alle Hybridisierungsergebnisse bestätigt werden, obwohl diese in drei Fällen im Konflikt mit der klinischen Bestimmung lagen. Ein DNA-Extrakt aus Reinkultur stellte sich bei der Mikroarray-Hybridisierung und Sequenzierung als *Enterococcus faecalis* anstatt *Enterococcus faecium* heraus. Ein *Aeromonas trota*-Extrakt war laut Mikroarray-Hybridisierung und Sequenzierung *Atopobium minutum*. Ein DNA-Extrakt zeigte nicht die erwartete Hybridisierung mit den *Aeromonas hydrophila*-Sonden und konnte anhand der Sequenzdaten als *Aeromonas bestiarum* identifiziert werden. Bei einigen Sequenzen war

anhand des Datenbank-Vergleichs keine eindeutige Identifizierung möglich. *Mycobacterium tuberculosis*, zum Beispiel, war in der 16S-Sequenz identisch mit *M. tuberculosis* wie auch mit weiteren Mykobakterien und das 23S-Gen konnte nicht zwischen *M. tuberculosis* und *M. bovis* unterscheiden. *Salmonella enteritidis* war im 16S rRNA-Gen nicht unterscheidbar von *Salmonella paratyphi*, einem anderen *S. enterica*-Serovar, zeigte aber deutliche Unterschiede im 23S rRNA-Gen. Hier konnte das Serovar Enteritidis jedoch nicht von weiteren *S. enterica*-Serovaren differenziert werden. Ähnliche Konstellationen gab es bei weiteren Spezies, wobei teilweise auch der Mangel an hinterlegten Sequenzen, insbesondere des 23 rRNA-Gens, eine eindeutige Identifizierung der Referenz-Stämme verhinderte. Die Sequenzierungsergebnisse standen jedoch nie im Widerspruch zu den Hybridisierungsergebnissen mit dem Mikroarray.

Die Sequenzierung offenbarte, dass viele der Kreuzhybridisierungen trotz vorhandener Basenfehlpaarungen auftraten. Vor allem Sonden mit nur ein bis zwei Fehlpaarungen zu einigen Spezies konnten mit diesen noch kreuzhybridisieren, wobei die Signale jedoch schwächer waren als von richtig-positiven Spezies mit vollständiger Sequenzübereinstimmung. In vielen Fällen reichten aber auch ein bis zwei Basenfehlpaarungen für die vollständige Diskriminierung von Spezies. Die Position der Fehlpaarungen in der Sonde war dabei nicht ausschließlich ausschlaggebend, wie der Vergleich zweier kreuzhybridisierender Sonden und deren Fluoreszenzsignale zeigte. In einem Fall wurden auch Kreuzhybridisierungen trotz höherer Anzahl an Basenfehlpaarungen beobachtet. Die Ursache lag in der einseitigen Verteilung der Basenfehlpaarungen in der Sonde. Die vier Fehlpaarungen kumulierten in den ersten acht Basen der Sonde, so dass weitere 16 zusammenhängende Basen für eine Hybridisierung zur Verfügung standen. Eine weitere Nichtzielspezies hatte vier anders verteilte Fehlpaarungen, die nur acht zusammenhängende Basen für die Hybridisierung verfügbar machten und somit eine Kreuzhybridisierung verhinderten. Nur bei drei Kreuzhybridisierungen waren diese auf vollständige Sequenzidentität der Sonde mit der Spezies zurückzuführen.

Array-Redesign

Auf Basis der Hybridisierungsergebnisse mit den Referenz-DNA-Extrakten sowie den Sequenzierungsergebnissen wurde der Array und dessen Sonden überarbeitet. Generell wurden Sonden mit Kreuzhybridisierungen mit einem MM/PM-Verhältnis von unter 2% als ausreichend spezifisch akzeptiert. Sonden, die stärkere Kreuzhybridisierungen zeigten, wurden in der Regel vom Array ausgeschlossen. Das wurde vor allem durch die mehrfache Abdeckung jeder Spezies mit verschiedenen Sonden ermöglicht. Aufgrund dieser Mehrfachabdeckung konnten aber auch einige Sonden mit MM/PM > 2% erhalten werden, wenn dieser Sonde mindestens zwei absolut spezifische Sonden gegenüber standen. Insgesamt wurden im Laufe des Validierungsprozesses 66 von 164 Sonden aus dem Arraypanel entfernt.

Bei einigen Sonden war es möglich durch eine Abänderung der Sonde zu der gewünschten Spezifität zu gelangen. Dies wurde zum Beispiel im Falle einer *L. monocytogenes*-Sonde erreicht, deren Spezifität gegenüber *L. innocua* durch eine Verkürzung der Sonde um vier Nukleotide zustande kam. Auf diese Weise konnten die zwei bestehenden Fehlpaarungen zwischen *L. monocytogenes*-Sonde und *L. innocua* PCR-Produkt ihre Wirkung effektiver entfalten. Die Kreuzhybridisierungssignale wurden dadurch um 57% (sense) bzw. 83% (antisense) reduziert. Die antisense-Sonde konnte nun für die spezifische Detektion von *L. monocytogenes* erhalten bleiben. Im Falle einer Sonde für *C. jejuni*, die eine unerwartet hohe Unspezifität gegenüber *C. coli* zeigte, offenbarte die Sequenzierung vollständige Sequenzübereinstimmung beider Spezies mit der Sonde. Im Gegensatz dazu zeigten die *C. coli*-Sequenzen in der Datenbank ARB, die für das Sondendesign herangezogen worden waren, sieben Fehlpaarungen und zwei Insertionen in diesem Bereich. Auf der Grundlage der neuen Sequenzen wurde eine neue Sonde (sense + antisense) entworfen, die zwischen beiden Arten unterscheiden konnte.

Alternativ zur Entfernung oder Abänderung wurden auch einige Sonden zu Sonden mit multipler Spezifität umdeklariert. Dies war der Fall bei einer *E. coli*-Sonde, die auch mit EHEC (enterohaemorrhagischer *E. coli*), *S. enteritidis* und *S. typhimurium* hybridisieren konnte, sowie einer *Y. enterocolitica*-Sonde, die mit *Y. pseudotuberculosis* kreuzhybridisierte. Zwei *V. cholerae*-Sonden mussten ebenso als Gruppensonden umdeklariert werden, da sie auch *V. fischeri* und *V. parahaemolyticus* binden konnten. Eine weitere Kreuzhybridisierung der EHEC-Sonde mit *S. enteritidis* führte leider zu einer Einschränkung der Nachweismöglichkeit von EHEC. Diese können nicht in Gegenwart von Salmonellen, also einer potentiellen Co-Infektion, nachgewiesen werden.

Der finale Array beinhaltete 4 Kontrollsonden und 108 Sonden zur Detektion der 16S und 23S ribosomalen RNA-Gene von 23 Spezies und 11 Gattungen. Diese deckten die 13 verbreitetsten bakteriellen Gastroenteritiserreger ab, nämlich *Aeromonas hydrophila*, *Campylobacter jejuni*, *Clostridium difficile*, *Mycobacterium avium* complex, *Mycobacterium avium* subsp. *paratuberculosis*, *Plesiomonas shigelloides*, *Vibrio cholerae*, *Vibrio parahaemolyticus*, *Yersinia enterocolitica*, *Listeria monocytogenes*, *Salmonella*, *Yersinia enterocolitica*, und *E. coli* O157:H7 (EHEC). Desweiteren waren folgende residente Darmbakterien und potentiell probiotische Bakterien auf dem finalen Chip vertreten: *Acholeplasma laidlawii*, *Atopobium minutum*, *Bacteroides fragilis*, *Bifidobacterium bifidum*, *Enterococcus faecalis*, *Escherichia coli* DH5 α , *Eubacterium bifforme*, *Fusobacterium prausnitzii*, *Lactobacillus acidophilus*, *Lactobacillus delbrueckii*, *Lactococcus lactis*, *Roseburia intestinalis*, und *Veillonella*.

Der endgültige Mikroarray ermöglichte die zweifelsfreie Bestimmung der Zielspezies durch eine Kombination des Mehrsonden-Konzepts mit einem robusten *cut-off* unter Verwendung der negativen Hybridisierungskontrolle plus drei Mal deren Standardabweichung, wobei nur positive Signale mit einer Standardabweichung <50% akzeptiert wurden. In den meisten Fällen lieferte die Detektion auf zwei taxonomischen Ebenen zusätzliche Robustheit.

Nachweisgrenze des Mikroarrays

Die Nachweisgrenze des Arrays wurde in Abhängigkeit vom Grad des DNase-Verdau und in einem weiteren Versuch vor dem Hintergrund eines Überschuss an Fremd-DNA bestimmt. Ohne Fremd-DNA lag das Detektionslimit von *E. coli* bei 10^3 Genomäquivalenten nach Verdau mit 0,4 oder 4 mU DNase/ng DNA. Nach Verdau mit 4 mU DNase I pro ng DNA zeigten drei von vier Sonden die höchsten Signalintensitäten. Im Vergleich zu unverdautem PCR-Produkt waren die Signale 3 bis 35-fach erhöht. Dagegen waren bei Verdau mit 0,4 mU DNase I pro ng DNA die Signalintensitäten der drei Sonden nahezu vergleichbar, aber die vierte Sonde zeigte eine fünffach höhere Signalintensität als bei dem stärkeren Fragmentverdau vor Hybridisierung.

Zur Bestimmung der Nachweisgrenze vor dem Hintergrund von Fremd-DNA wurde *Campylobacter jejuni* in eine *C. jejuni*-negative Stuhlprobe in verschiedenen Konzentrationen eingespiked. Die Nachweisgrenze nach Verdau mit 4 mU DNase pro ng DNA betrug ebenfalls 10^3 Genomäquivalente. Der Überschuss an Fremd-DNA an der Nachweisgrenze war im Versuch 2.770fach gegenüber *C. jejuni*.

Klinische Validierung des Mikroarrays

Die klinische Anwendbarkeit des DNA-Mikroarrays wurde durch Untersuchung von Stuhlproben von 58 Patienten mit gastroenteritischen Symptomen und 6 gesunden Individuen überprüft. Von den 58 Patienten-Proben wurden in der Klinik bereits 30 durch kulturbasierte Methoden bzw. *real-time* PCR positiv auf Pathogene getestet. Diese Pathogen-positiven Proben umfassten 17 *Salmonella*-positive, acht *Clostridium*-positive, vier *Campylobacter*-positive und eine *Yersinia*-positive Probe. Alle weiteren Patientenproben waren negativ für die getesteten Pathogene. Die DNA-Extrakte wurden mittels des

entwickelten PCR-Protokolls amplifiziert. Bei vier DNA-Extrakten war es nicht möglich die ribosomalen RNA-Gene zu vervielfältigen, was auf die Faktoren lange Lagerzeit und starke Fragmentierung durch die DNA-Extraktion zurückgeführt wurde. Die korrekte Identifizierung von Pathogenen in den Patientenisolaten (ohne Gesunde) lag bei 77%. Von den 30 Proben bei denen das ätiologische Agens bereits mit einer Kombination aus Stuhlkultur und spezifischer *real-time* PCR detektiert worden war, wurden 67% mit dem Mikroarray korrekt identifiziert. Die klinische Sensitivität der Mikroarraydetektion bezogen auf die Referenzmethoden betrug demnach 67%, wobei hier nicht auf mögliche Falschpositive oder Falschnegative durch die klinische Diagnostik untersucht und korrigiert wurde. Die Spezifität der Arraydetektion lag bei 89%. Es wurde geschlussfolgert, dass eine höhere Wiederfindungsrate eine niedrigere Nachweisgrenze des Assays erfordern würde, die durch sensitivere Markierungsreagenzien, Arrayoberflächen mit einer höheren Sondenbindungskapazität, Anreicherungsstrategien für das Pathogen, Vorhybridisierung in Lösung, Reduktion des Reaktionsvolumens durch mikrofluidische Systeme und durch RNA als Zielstruktur erreicht werden könnte.

Darüber hinaus wurden in den analysierten Proben hoch individualisierte mikrobielle Gemeinschaften der detektierten residenten Bakterien gefunden. Dominant, das heißt besonders häufig auffindbar, waren hierbei die Spezies bzw. Gattungen *Bacteroides* spp., *B. fragilis*, *F. prausnitzii*, *Veillonella* spp., *Roseburia* spp. und *R. intestinalis*. Ein Unterschied in der mikrobiellen Zusammensetzung zwischen gesunden und erkrankten Individuen wurde unter Verwendung von Multivarianzanalyse der Arraydaten jedoch nicht beobachtet. Für die Multivarianzanalyse wurden die Hauptkomponentenanalyse (PCA) und die *partial least squares*-Pfadanalyse (PLSA) herangezogen. Die einfaktorielle Varianzanalyse (*one-way* ANOVA) wurde als Referenzmethode genutzt. Laut einer Literaturrecherche zum Thema war dies die erste Anwendung der Hauptkomponentenanalyse und *partial least squares*-Pfadanalyse für die statistische Auswertung von Identifikationsmikroarraydaten.

In der PCA konnten die zwei Komponenten, die die beste Separation der beiden Gruppen ermöglichten, nur 16% der Varianz erklären. In der PLSA ließen sich beide Gruppen ebenfalls nicht trennen. Die Fehlerrate der Klassifizierung der Proben zu gesunden bzw. kranken Individuen lag hier bei 10,3% für ein Ein- bzw. Zweikomponentenmodell wie mit *leave-one-out cross-validation* (LOOCV) gezeigt wurde. Obwohl durch die PLSA *Bacteroides fragiles* als die wichtigste Variable für die Trennung der beiden Gruppen identifiziert wurde, konnte dies durch *one-way* ANOVA nicht bestätigt werden. Die nicht erfolgreiche Zuordnung der Proben zu gesunden bzw. erkrankten Individuen anhand der residenten Intestinalmikrobiota wurde auf die inhomogene Zusammensetzung der Proben bezüglich des Patientenalters und der klinischen Diagnose zurückgeführt. Eine Analyse der Daten nach Trennung in drei Gruppen, wobei die erkrankten Individuen nochmals in Personen mit bzw. ohne exakte Pathogendiagnose unterteilt wurden, brachte keine Verbesserung bei der Klassifizierung.

Die Fluoreszenzsignale der Hybridisierungsversuche mit Realproben wurden auch auf ihre Standardabweichung innerhalb von zwei hybridisierten Arrays (je drei Spots) hin untersucht. In 90% der Fälle lag diese bei bzw. unter 20%. Insgesamt 63% der Sonden wiesen Standardabweichungen von $\leq 10\%$ auf. Die durchschnittliche Standardabweichung lag bei 11%.

Die Gesamtassayzeit betrug 5,5 h nach DNA-Extraktion aus dem klinischen Material.

Anwendung des Arrays für Pathogendetektion und Analyse der residenten Mikrobiota

Nach Validierung des Mikroarrays wurde dieser zur Untersuchung des Einflusses einer viralen, intestinalen Infektion auf die Zusammensetzung der kindlichen Intestinalmikrobiota sowie der Etablierung von menschlicher und Schweine-Intestinalmikrobiota in inokulierten Mini-Schweinen genutzt. Für diese Untersuchungen wurde erfolgreich von Fluoreszenzdetektion auf Detektion von Silberabscheidung umgestellt, da die Versuche im Rahmen

eines Forschungsaufenthalts in Shanghai/China stattfanden. Die DNA-Extraktion erfolgte hier anhand eines Protokolls mit mechanischem Aufschluss und Phenol-Chloroform-Extraktion.

Die erste Probengruppe umfasste zehn Proben von gesunden Kindern und zehn Proben von Rotavirus-infizierten Kindern. Das eigentliche Pathogen, Rotavirus, konnte mit den Sonden des Arrays nicht detektiert werden. Dennoch wurden überraschend in drei Proben der erkrankten Kinder auch bakterielle Pathogene gefunden. Alle drei DNA-Extrakte wiesen eine nachweisbare Konzentration von *Clostridium difficile* auf. Zusätzlich enthielt eine Probe nachweislich *Campylobacter jejuni*. Die residente Intestinalmikrobiota wies deutliche Unterschiede zwischen gesunden und infizierten Kindern auf, wie die Multivarianzanalyse der Arraydaten belegte. Das PCA-Modell konnte 50% der vorhandenen Varianz in den Daten erklären und beide Gruppen wurden anhand zweier PCA-Komponenten klar separiert. Ein Zweikomponenten-PLSA-Modell machte die Gruppenzuordnung noch deutlicher und erreichte laut LOOCV 95% korrekte Zuordnung der Proben zu den gesunden bzw. kranken Kindern. Es wurde vermutet, dass sich eine intestinale Infektion auf die im Vergleich zu Erwachsenen wesentlich geringer ausgeformte Intestinalflora von Kindern stärker auswirken könnte und hier insbesondere auch die dominanten Spezies betrifft. Dies könnte als Erklärung dienen, dass hier eine Trennung von erkrankten und gesunden Individuen auf Basis der residenten Intestinalflora möglich war, während dies bei der gemischten Probenkohorte nicht gelungen war. Die Anwendung von PLSA und einfaktorieller Varianzanalyse ermöglichte anschließend die Offenlegung der Spezies bzw. Gattungen, die sich in ihrer Präsenz in den jeweiligen Gruppen signifikant unterschieden. In der PLSA waren dies *Roseburia* spp., *E. coli*, *F. prausnitzii* und *Atopobium* spp., wobei der signifikante Unterschied ($p < 0,05$) mit *one-way* ANOVA nur für *Roseburia* spp. und *F. prausnitzii* bestätigt wurde. Zusätzlich wurde mit *one-way* ANOVA noch *R. intestinalis* als signifikant verschieden zwischen den beiden Gruppen identifiziert. Basierend auf den Mikroarraydaten von *Roseburia* spp., *E. coli*, *F. prausnitzii* und *Atopobium* spp. wurde ein neues PLSA-Modell etabliert. Das Einkomponentenmodell hatte eine kreuzvalidierte korrekte Klassifizierungsrate von 95%, während das Zweikomponentenmodell nur 90% erreichte. Wurde nur mit *Roseburia* spp. und *F. prausnitzii* ein PLS-Modell berechnet, erreichten sowohl Ein- als auch Zweikomponentenmodell 95% korrekte Klassifizierung. Durch Einbeziehung von *R. intestinalis*, wie es durch *one-way* ANOVA empfohlen war, wurde die Falschklassifizierung deutlich erhöht.

Die zweite Versuchsreihe umfasste Stuhlproben von Schweinen mit inokulierter humaner (HFA) bzw. Schweineintestinalmikrobiota (PFA). Mit beiden Versuchsgruppen sollte die Etablierung der Humanflora im Vergleich zur Schweineflora untersucht werden. Unerwartet trat in beiden Gruppen eine Diarrhöe auf, von der die Gruppe mit inokulierter Schweineflora stärker betroffen war. In dieser Gruppe starben zwei Schweine während des Versuchsverlaufs. Von beiden Gruppen wurden Stuhlproben an Tag 14 sowie Tag 21 nach Geburt genommen und mit dem Gastroenteritis-Mikroarray und denaturierender Gelelektrophorese (DGGE) (nur Tag 14) untersucht. Die Mikroarrayanalyse zeigte, dass nach 14 Tagen jedes Individuum bereits eine individuelle Intestinalmikrobiota hatte, die sich klar vom Donor unterschied. Die etablierte Schweinemikrobiota in den inokulierten Schweinen unterschied sich auch von der Schweineflora der konventionell aufgezogenen (CV) Tiere. Ein bakterielles Pathogen, nämlich EHEC, wurde nur in zwei HFA-Schweinen gefunden. Sieben Tage später hatten sich die bakteriellen Profile teilweise verändert. Die HFA-Profile hatten sich deutlich diversifiziert während die PFA-Profile ausgedünnt wurden. Dies entsprach dem Krankheitsverlauf in seiner symptomatischen Ausprägung. Bei den HFA-Schweinen hatte sich das Intestinalprofil dem Donor stärker angeglichen. Bei den Pathogenen wurde EHEC nun in zwei anderen HFA-Tieren detektiert, während die ursprünglichen zwei Individuen keine nachweisbaren Pathogene mehr enthielten.

Mittels DGGE konnten keine Spezies identifiziert, aber die intestinale Diversität beurteilt werden. In der DGGE wurde die V3-Region untersucht, die eine konservierte Region des 16S rRNA-Gens darstellt. In allen Proben wurde ein ausgeprägtes Fragmentprofil gefunden,

wobei in der HFA-Gruppe entsprechend der Mikroarrayanalyse eine stärkere Fragment-homologie zwischen den einzelnen Individuen zu beobachten war, als dies bei den PFA-Tieren der Fall war. Die Ähnlichkeit des Fragmentprofils der HFA-Schweine und ihres Donors war ebenfalls stärker als zwischen den PFA-Tieren und ihrem Donor. Der PFA-Donor hatte deutlich mehr Ähnlichkeit mit dem Fragmentprofil der konventionell aufgezogenen Schweine. Offensichtlich hatte sich die humane Intestinalflora schneller und stabiler in den inokulierten Schweinen etabliert als die Schweineintestinalflora, was hier im wesentlichen auf die Infektion und deren symptomatische Auswirkungen zurückgeführt wurde.

Mit Hilfe einer Multivarianzanalyse der Arraydaten wurde ein signifikanter Unterschied in der Zusammensetzung der etablierten Intestinalmikrobiota zwischen PFA- und HFA-Schweinen festgestellt. Die Proben wurden zwei Gruppen zugeordnet, wobei eine Gruppe die HFA-Schweine und deren Donor umfasste und die andere Gruppe die PFA-Schweine, deren Donor und die CV-Schweine. Die Analyse der Arraydaten von Proben Tag 14 zeigte noch keine vollständige Trennung der beiden Gruppen und die zwei Komponenten der PCA erklärten nur 29% der Varianz der Daten. Die PLSA resultierte in einer klareren Trennung beider Gruppen. Ein Einkomponentenmodell erreichte 84% korrekte Klassifizierung. Es wurden vier Spezies bzw. Gattungen identifiziert die maßgeblich zu einer Trennung beider Gruppen beitragen: *Fusobacterium prausnitzii*, *Veillonella* spp., *Bifidobacterium bifidum* und *Lactobacillus delbrueckii*. Anhand der einfaktoriellen Varianzanalyse konnte der signifikante Unterschied nur für *L. delbrueckii* und *F. prausnitzii* bestätigt werden. Ein PLS-Modell auf Basis dieser beiden Spezies resultierte in 89% korrekter Klassifizierung mit einer Komponente und 84% mit zwei Komponenten. Die Verwendung aller vier Spezies verbesserte die Probenzuordnung nicht.

Die Multivarianzanalyse der Arraydaten von Tag 21 nach Geburt zeigte eine klare Trennung beider Gruppen, wobei die zwei Hauptkomponenten der PCA 64% der Varianz in den Daten erfassten. In der PLSA erzielte ein Einkomponentenmodell 100% korrekte Klassifizierung der Proben. Es wurden die fünf Spezies *Bacteroides fragilis*, *Enterococcus faecalis*, *Escherichia coli*, *Bifidobacterium bifidum* und *Lactobacillus delbrueckii* als die Spezies identifiziert, die am meisten zur Gruppentrennung beitragen. Dieses Ergebnis wurde in der *one-way* ANOVA für alle Spezies außer *L. delbrueckii* bestätigt. Ein PLS-Modell auf Basis der vier verbliebenen Variablen erreichte mit einer Komponente ebenso eine 100% korrekte Klassifizierung.

Die Ergebnisse bestätigten, dass sich sowohl Schweine- als auch humane Intestinalmikrobiota bereits innerhalb von 14 Tagen im Intestinaltrakt von inokulierten Schweinen ansiedeln kann und nach 21 Tagen der Donorflora ähnlich ist. Außerdem wurde deutlich, dass die inokulierte Intestinalflora vom Schwein nach 21 Tagen auch den konventionell aufgezogenen Schweinen ähnelte. Beide Mikrobiota, human und vom Schwein, unterschieden sich nach 21 Tagen deutlich voneinander, wie die Multivarianzanalyse zeigte. Die identifizierten Spezies die zur Unterscheidung von beiden Mikrobiota beitragen, können jedoch nicht als Markerspezies gelten, da die Versuchsreihe von einer Intestinalinfektion betroffen war. Das Ziel das auslösende Pathogen zu identifizieren gelang nicht mit absoluter Sicherheit. Es wurde zwar EHEC in vier Humanflora-assoziierten Schweinen detektiert, aber gerade diese Tiere zeigten weniger Symptome der Infektion als die Schweineflora-assoziierten Tiere. Dies war möglicherweise auf eine unzureichende Nachweisgrenze zurückzuführen, aber es könnte auch ein zweites Pathogen, z.B. viraler Natur, eine Rolle gespielt haben.

Die Ergebnisse dieser Versuche legen die Anwendbarkeit des entwickelten Arrays für Studien an der residenten Intestinalflora unter verschiedenen Fragenstellungen nahe. Dennoch wurde geschlossen, dass eine tiefergehende Analyse der intestinalen Gemeinschaft eine Erweiterung des Arrays mit anderen Spezies erfordern würde.

Ziel-DNA-Quantifizierung mit DNA-Mikroarrays

Um das Problem einer verlässlichen Quantifizierung der intestinalen Bakterien und Pathogene auf dem Mikroarray über die Spanne mehrerer Zehnerpotenzen zu lösen, wurde ein FRET-basiertes System evaluiert. Ein *black hole quencher* wurde zur graduellen Löschung des Cyanin 3-Reportersignals eingesetzt, wodurch die Detektion der Zielspezies im linearen Bereich des Scanners unabhängig von deren Ausgangskonzentration ermöglicht wurde. Dazu wurden von jeder Sonde mehrere *spots* mit unterschiedlichen, abgestuften Mengen an gebundenem *quencher* auf der Arrayoberfläche aufgebracht und mit einer komplementären Ziel-DNA in verschiedenen Konzentrationen hybridisiert. Der Nachweis des Wirkprinzips wurde anhand von *E. coli* und dessen spezifischen Sonden erbracht und zeigte, dass ein effizientes *quenching* des Fluorophores in der Ziel-DNA nach deren Hybridisierung in Abhängigkeit vom der im *spot* immobilisierten *quencher*-Menge möglich ist. Es wurde geschlussfolgert, dass zwei *spots*, einer ohne und einer mit 8 μM Sonden-gekoppeltem *quencher*, ausreichend sind um 10^4 bis $2 \cdot 10^6$ Genomäquivalente im linearen Scannerbereich zu detektieren. Diese Anordnung mit zwei *spots* überwindet teilweise die technische Einschränkung bei der linearen Fluoreszenzsignalerfassung für verschiedene Bakterienmengen. Die obere Grenze von $2 \cdot 10^6$ Genomäquivalenten wurde hier durch die maximale DNA-Menge von 10 ng bestimmt, die nach dem bestehenden Protokoll des Assays in einer PCR eingesetzt wurde.

Diskussion und Schlussfolgerung

Zusammenfassend kann festgestellt werden, dass die Entwicklung des Gastroenteritis-Chips zwei Hauptprobleme der klinischen Pathogendetektion, die benötigte Detektionszeit und Assayabdeckung möglicher Pathogene, gelöst hat. Die wesentlichen Faktoren waren hierbei die Ablösung der in der klinischen Diagnostik üblichen Kultivierung der Pathogen durch molekularbiologische Detektion sowie die Möglichkeit von Mikroarrays zum parallelen Nachweis von vielen Spezies. Für die klinische Anwendung ist jedoch die klinische Sensitivität und Spezifität noch verbesserungswürdig um mit anderen molekular-diagnostischen Verfahren konkurrieren zu können. Hier dürfte vor allem die Nachweisgrenze des Arrays von Bedeutung sein, die allerdings in der Gegenüberstellung mit anderen publizierten Mikroarrays absolut vergleichbar war.

Neben der Anwendung des Arrays als Diagnostikum stand auch die Nutzung als Therapiebegleitendes Werkzeug im Fokus der Entwicklung. Der Array wurde mit einem Panel an Sonden zum Nachweis residenter Intestinalbakterien und potentieller Probiotika ausgestattet. Die Untersuchung einer Rotavirus-infizierten Kinderkohorte und einer gesunden Vergleichsgruppe brachte dabei erste Erkenntnisse zur Anwendbarkeit des Arrays in diesem Bereich. Hier wurden deutliche Unterschiede zwischen der Mikrobiota beider Gruppen in Abhängigkeit von einer Infektion gefunden. Ob und inwieweit diese Ergebnisse auch auf Erwachsene übertragbar sind, müsste in weiteren Untersuchungen gezeigt werden. Dafür und auch für die Anwendung in der Grundlagenforschung hinsichtlich der Etablierung der intestinalen Gemeinschaft im Verdauungstrakt wäre eine Erweiterung des Spektrums der Ziel-Spezies in Erwägung zu ziehen.

Darüber hinaus konnte in dieser Arbeit ein Konzept für ein FRET-basiertes System zur DNA-Quantifizierung auf Mikroarrays über mehrere Zehnerpotenzen erfolgreich entwickelt werden, wodurch ein zentrales Problem in klinischen und wissenschaftlichen Mikroarrayanwendungen aufgegriffen wurde.

Für die Weiterentwicklung des Arrays in Richtung klinischem Diagnostikum wurde die Marktrelevanz, das Potential zur Integration des Arrays in *point-of-care* Diagnostika sowie potentielle Erweiterungen diskutiert. Von Interesse könnte hier vor allem die Kombination Speziesidentifikation mit Resistenztestung sein, da zunehmend auch eine Verbreitung von Resistenzdeterminanten unter Intestinalbakterien und -pathogenen beobachtet wird. Generell ist ein großer Markt für Assays basierend auf Nukleinsäureamplifikationstechniken

vorhanden. Kalkulationen haben aber gezeigt, dass neue Methoden die klassischen ersetzen können müssen, damit ihr Einsatz im Rahmen einer kostenfokussierten Gesundheitspolitik möglich ist. Es ist demnach zu erwarten, dass diagnostische Mikroarrays den Spagat zwischen Multiplexingkapazität, Sensitivität, Spezifität, Probendurchsatz und Kosten schaffen müssen.

Neben der klinischen Applikation ließe sich der Array anhand des bestehenden Sondenpanels auch für die Anwendung in der Lebensmittel- oder Wasserkontrolle einsetzen. Bei Ersterem stellt jedoch die DNA-Extraktion aus den verschiedenen, komplexen Matrices eine wesentliche Herausforderung dar, während bei Letzterem die Nachweisgrenze aufgrund der niedrigen Pathogenkonzentration in kontaminiertem Wasser besonders zum Tragen kommt.

Der entwickelte Gastroenteritis-Chip, der Pathogene in Kombination mit residenter Intestinalmikrobiota und potentiellen Probiotika erfasst, stellt mit seinem Sondenspektrum eine deutliche Erweiterung und Verbesserung gegenüber bisherigen Entwicklungen in der Gastroenteritis-Diagnostik dar. Ergänzungen des Detektionsspektrums hinsichtlich der intestinalen Diversität sowie dem sich ständig erweiternden Pathogenspektrum sind jedoch wünschenswert. Ein wichtiger Vorteil der Mikroarraytechnologie liegt in der Möglichkeit begründet, solche Erweiterungen mit relativ geringem Aufwand und in kurzer Zeit realisieren zu können.

1 Introduction

1.1 The digestive tract

The largest surface of the human body that is exposed to its surrounding and interacts with the environment is the digestive tract lying within the body. The digestive tract consists of the upper intestinal tract with mouth, pharynx, oesophagus, and stomach, and the lower intestinal tract with the bowel or intestine comprised of small and large intestine, and the anus. The small intestine can be additionally separated into duodenum, jejunum, and ileum, while the large intestine consists of cecum, colon, and rectum. The primary functions of the gastrointestinal tract (GI tract) are ingestion, digestion, absorption, and defecation but a lot of other functions, like shaping of the immune system, are assigned to it and especially to the gut microbiota, whose members outnumber the human cells by the factor of ten (Savage 1977). Concerning the composition of the intestinal flora, our knowledge goes back about 40 years. Until now, more than 400 species of the GI tract are known (Hill 1975; Guarner 2006) but it is estimated that there exist far more. Culturing techniques, which gave first insight into the intestinal system (Eller 1971; Moore 1974), have been replaced by molecular methods based on sequencing of the bacterial ribosomal RNA genes. Nevertheless, new species are rarely identified, which is most probably due to the low abundance of these bacteria in the intestinal community that is dominated by the two groups *Bacteroidetes* and *Firmicutes*. The Low GC Gram-positive *Firmicutes* account for 46-58% of bacteria including the huge *Clostridium leptum* (~25%) and *Clostridium coccooides-Eubacterium rectale* (~28%) group, the *Eubacterium cylindroides* (~1.5%), the *Lactobacillus-Enterococcus* (~1.8%) and the *Veillonella* (~1.3%) group. The Gram-negative *Bacteroides-Prevotella* group (*Bacteroidetes*) accounts for up to 28% whereas the Gram-negative *Enterobacteriaceae* that comprise species like *E. coli* and *Klebsiella* spp. is only represented by about 0.2%. The *Akkermansia* group has a low abundance of 1.3% as well. Another small group, the High GC Gram-positives, with 8-17 percentage includes the bifidobacteria and the *Atopobium* group (Zoetendal 2006) (Fig. 1.1).

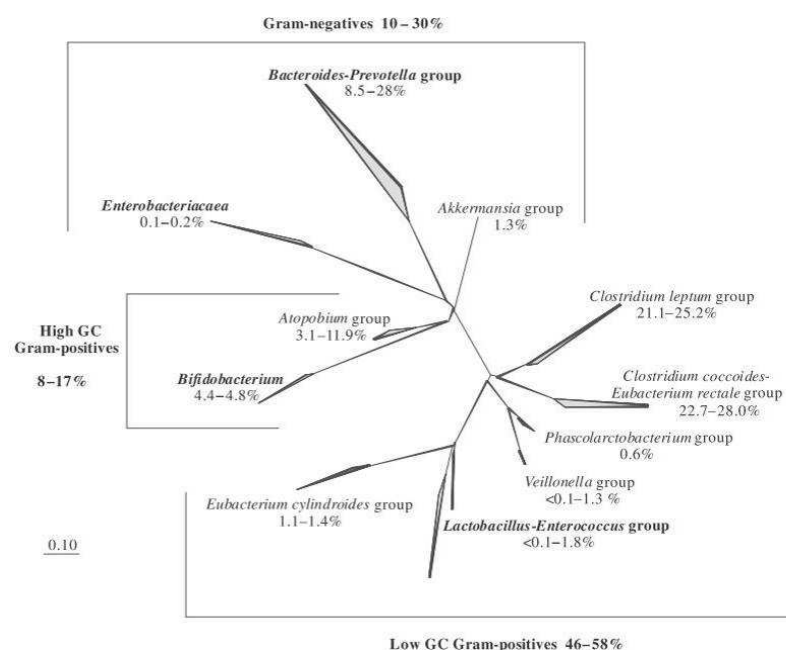


Fig. 1.1: Phylogenetic tree of 16S rRNA gene sequences representing different groups of bacteria, which are most frequently detected in human faeces. The scale bar represents the calculated distance between the sequences. From Zoetendal et al. (Zoetendal 2006).

In spite of the huge amount of bacteria colonizing the GI tract, the distribution along its length is uneven. Due to the acidic environment in the stomach and the pancreatic excretions and bile salts in the duodenum, this part is only inhabited by 10^1 - 10^3 CFU/mL, while the amount of bacteria reaches 10^4 - 10^8 CFU/mL in jejunum and ileum. Compared to the colon with 10^{10} to 10^{13} CFU/mL this is still a low number of bacteria, which can be ascribed to the shorter transit time through the small intestine of 3-4 hours compared to 8-72 hours through the colon (Kararli 1995). Moreover, large spacial variations of the microbiota composition exist within the GI tract (Hayashi 2005).

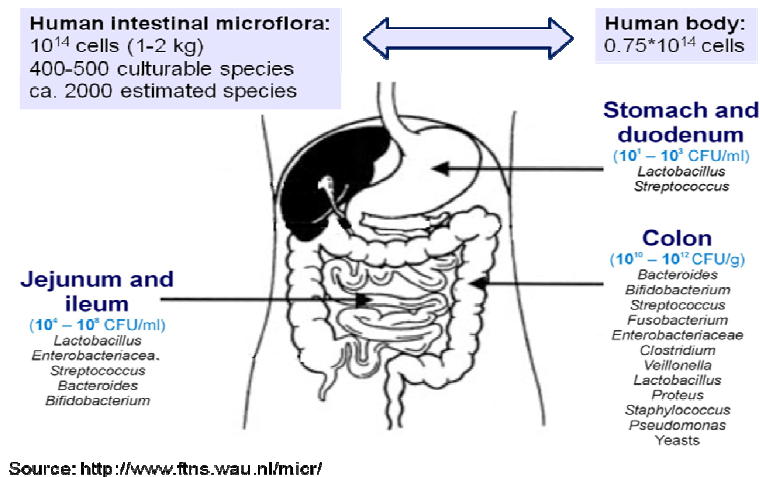


Fig. 1.2: Spatial variations of microbiota composition and density in the human GI tract.

Colonization of the GI tract starts with birth and undergoes major changes during the first weeks and years of life from a *Bifidobacterium*- to a *Bacteroides*-dominated community (Favier 2003; Bezirtzoglou 2006). From a certain age this community remains remarkably stable (Zoetendal 1998; Shen 2006), although transient bacteria are continuously ingested to the gut via food intake. It is nearly possible to call this an intestinal fingerprint, as it is individual to different subjects (Zoetendal 1998). Recent research also focuses on the changes of human gut flora during ageing (D'Souza 2007; Woodmansey 2007). The heavy colonization with bacteria and the permanent contact with ingested, external material including transient microorganisms require a well-modulated defence system, which nevertheless allows for interaction with the necessary commensals (Round 2010; Hooper 2010). An unbalanced colonization of the intestinal mucosa due to an antimicrobial activity deficiency is, for example, supposed to be one characteristic of inflammatory Crohn's disease (Nuding 2007). On the other hand, intestinal bacteria contribute to defence against pathogens through stimulation of the development of the host immune system (Cebra 1999; Moreau 2001) and several antagonistic mechanisms between bacteria (Fons 2000). Bacterial antagonism can occur through competition for substrates, competition for mucin-adhesion receptor sites, creation of a restrictive physiological environment (pH, redox potential, hydrogen sulphide production, toxic metabolites) or production of bacteriocines. Apart from some morphological and physiological changes germ-free animals survive in a sterile surrounding, but in a non-sterile environment the gut flora is essential to prevent pathogens from entering the mucosal tissue (Bourlioux 2003). Also transient bacteria, like probiotics, were shown to prevent pathogenic infection (Corr 2007). Anyhow, there is a permanent risk for attacks to the human body via the GI tract by bacteria, viruses, and protozoa.

1.2 Molecular techniques to monitor the intestinal community

With increasing knowledge about the functions and importance of the intestinal flora, a number of molecular techniques to analyze the intestinal community have been developed (Tab. 1.1). Molecular approaches allow describing bacterial communities on the bases of their 16S gene composition. This has led to a new way of investigating bacterial relationship and a new taxonomy based on 16S gene sequences. The methods can be classified into four groups according to the major principle for the identification or monitoring of species: (I) methods that use physical properties of the target strands based on its base composition and length, (II) methods that use restriction enzyme activity to generate species- or community-specific fragment patterns, (III) methods that use hybridization mediated enzymatic activity to randomly or selectively amplify DNA, and (IV) methods that use allele-specific hybridization with DNA or RNA target regions. All methods of group IV (FISH, dot-blot hybridization, microarray) as well as some of group III (real-time PCR, RT-PCR, Sanger sequencing) require previous knowledge about the target and, therefore, are applied mainly for monitoring of the intestinal microbiota under a particular question. For *de-novo* identification of intestinal inhabitants the other techniques, which are listed in Tab. 1.1, are favourable. The basic principle of each method is also described in this table.

A standard method, which yielded much new information about the GI-tract inhabitants, is the sequencing of clone-libraries (Bik 2006; Ozutsumi 2005; Hayashi 2002; Bibiloni 2006). Ley et al., for example, studied the correlation of obesity with the composition of the intestinal flora (Ley 2005). The cloning approach was compared with a classical culture-based method by Wilson and Blitchington, showing a good agreement between both methods if a PCR-produced bias was prevented (Wilson 1996). Later investigations even revealed a better coverage of the microbiota by molecular techniques (Suau 1999). The cloning approach has also led to a huge amount of sequences referred to as 'uncultured bacterium', which cannot be linked to known species but to phylogenetic groups or genera, only. Large fractions of the intestinal community have not been cultured yet but can be found in clone-libraries. To fully describe those bacteria it is still necessary to isolate and culture them for physiological analysis. Unfortunately, our lack of knowledge concerning physiology of and communication between single cell organisms disables us to culture them in the laboratory. Some attempts have been made to address this problem. Zengler and colleagues published a high-throughput cultivation technique based on micro-encapsulation of single cells in an agarose-in-oil emulsion combined with microbial cultivation under low nutrient-flux conditions based on the environmental matrix. This approach has the plausible advantage that the microbial ecosystem and communication relations are not disrupted during cultivation. Flow cytometry was used to detect growth within the capsules and to separate them into microwell plates for the second cultivation phase in rich medium. Subsequent PCR of the ribosomal RNA genes and sequencing of a shotgun library provides information on the composition of the microbiota (Zengler 2002).

Genetic fingerprinting techniques, such as denaturing or temperature gradient gel electrophoresis (DGGE/TGGE), terminal restriction fragment length polymorphism (T-RFLP), randomly amplified polymorphic DNA (RAPD), and amplified ribosome DNA restriction analysis (ARDRA), enable the cloning-independent comparative investigation of microbial communities and identification of new rRNA gene species. These methods rely on electrophoresis to resolve DNA fragments having different length. In principle, PCR or RT-PCR amplicons are assigned to anonymous groups, which are taxonomically described after subsequent excision of bands, cloning, and sequencing. The obtained taxonomic knowledge about the diversity patterns can be used to monitor intestinal community changes (Shen 2006; Sepehri 2007). The application of denaturing high performance liquid chromatography (DHPLC) for the identification of bacteria was introduced by Hurtle et al., who used it for the identification of *Y. pestis* and *B. anthracis* (Hurtle 2002). This technique is widely employed for SNP (single nucleotide polymorphism) detection due to its high resolution power, but rarely for species identification purposes (Domann 2003; Goldenberg 2005). Goldenberg et al.

recently showed that the applicability of this method in combination with group-specific PCR is comparable to DGGE for the monitoring of the gut microbiota (Goldenberg 2007).

Tab. 1.1: Molecular techniques to monitor the intestinal community.

| Molecular technique | Principle | References |
|---|--|--|
| Using physical properties of the DNA | | |
| Denaturing/temperature gradient gel electrophoresis (DGGE/TGGE) | Separates amplified fragments of equal length in an agarose gel with a temperature or denaturing agent gradient due to its different base composition and sequence (possible resolution: one nucleotide); a GC clamp is added during PCR, which prevents the strands from completely melting and improves resolution; nevertheless, a single band does not necessarily represent a single species; method is often combined with subsequent cloning and sequencing | <i>Bacteroides</i> (Pang 2005), <i>Bifidobacterium</i> (Satokari 2001; Requena 2002), <i>Clostridium leptum</i> group (Shen 2006), <i>Lactobacillus</i> related species (Walter 2001; Heilig 2002), (Vanhoutte 2006) |
| Denaturing high performance liquid chromatography (DHPLC) | Universal primers are used to amplify a highly variable region of the ribosomal DNA; before separation on an ion-pair reverse-phase HPLC column, the double stranded PCR products are partly heat-denatured; as single-stranded DNA has lower binding efficiency than double-stranded DNA and the degree of strand-melting is dependent on base composition, sequence, and the chosen temperature, it is possible to differentiate between species | <i>Y. pestis</i> and <i>B. anthracis</i> (Hurtle 2002), urinary tract pathogens (Domann 2003), 7 <i>Candida</i> species (Goldenberg 2005), intestinal community (Goldenberg 2007) |
| Single strand conformation polymorphism (SSCP) | Separation of denatured, single stranded DNA of equal length by agarose gel electrophoresis; single stranded DNA undergoes a unique 3-dimensional folding depending on its base composition and migrates differently in an electrophoresis gel | Spatial and temporal bacterial community variations (Ott 2004a; Wache 2009; Michelland 2009) |
| Using restriction enzyme activity | | |
| Amplified ribosome DNA restriction analysis (ARDRA) | Uses differences in the enzymatic fragmentation pattern of amplified rRNA genes; the discrimination power depends on the chosen enzyme and the length of the amplified fragment; due to high similarities between some species, it might be difficult to find satisfying conditions | Lactobacilli (Guan 2003; Ventura 2000; Moreira 2005), Bifidobacteria (Ventura 2001; Krizova 2006), Bacillus spp. (Wu 2006) |
| Pulse field gel electrophoresis (PFGE) | DNA fragmentation pattern derived from an enzymatic digestion by rare cutters; in the alternating electric field the extremely large fragments (>30-50 kb) can be separated according to their size; reorientation is more quickly for smaller DNA fragments, resulting in a faster net migration to the opposite gel end; quite laborious | (Tynkkynen 1999; Fugett 2007; Kohara 2006; Scott 2006; Vali 2007) |
| Restriction fragment length polymorphism (RFLP) | Polymorphism identification based on site-specific restriction enzyme digestion of DNA templates followed by gelelectrophoretic analysis | (Suau 1999) |
| Terminal-restriction fragment length polymorphism (T-RFLP) | Selective amplification of a certain gene, mainly the 16S-rRNA gene, with a primer set including one fluorescently end-labelled primer and subsequent enzymatic restriction ideally resulting in an individual fragment pattern | (Clement 1998; Sepehri 2007; Dinoto 2006; Fairchild 2005; Hayashi 2005; Jernberg 2005) |
| Automated ribosomal intergenic spacer analysis (ARISA) | The intergenic spacer region (ITS) between 16S and 23S rRNA genes is amplified with primers targeting conserved regions in both adjacent genes; PCR products differ in length already due to potentially encoded tRNAs but analysis is based on the terminal restriction fragment pattern | GI-tract (Sepehri 2007); widely used for aquatic and soil ecosystems |
| Using hybridization mediated enzymatic activity | | |
| Randomly amplified polymorphic DNA (RAPD) | Uses short, universal primers, often decamers, that randomly amplify sequences from the whole genome target DNA, which result in a fingerprint composed of unknown PCR products | Lactobacilli (Du Plessis 1995; Tynkkynen 1999) and Bifidobacteria (Vincent 1998) specific marker for a <i>Lactococcus lactis</i> strain (Maruo 2006) |

Table continued on next page

Continuation of Tab. 1.1

| Molecular technique | Principle | References |
|--|--|--|
| Using hybridization mediated enzymatic activity | | |
| Real-time polymerase chain reaction (real-time PCR) | Amplification of a target sequence using specific primers, nucleotides and a polymerase; double-stranded PCR products are detected via intercalation of a fluorophore like SYBR® Green, incorporation of a self-quenching double-labelled TaqMan® probe (hydrolysis probe) or incorporation of two individually labelled LightCycler® probes (hybridization probes) during PCR that recognise two adjacent sequences; the TaqMan® probe is degraded by the exonuclease function of the polymerase and the released fluorophore can display its fluorescence; LightCycler® probes are incorporated into the PCR-product in close proximity allowing fluorescence resonance energy transfer (FRET) from one fluorophore to the other | TaqMan assay (Heid 1996; Holland 1991; Lee 1993); LightCycler assay (Wittwer 1997); community monitoring (Armougom 2009) |
| Reverse transcriptase PCR (RT-PCR) | Reverse transcription of RNA to single-stranded DNA by reverse transcriptase and selective amplification of a certain DNA strand | Pro-/probiotics (De Preter 2007); gut microbiota (Matsuda 2009); <i>C. difficile</i> (Iizuka 2004) |
| Dye/Chain-termination sequencing (Sanger) | DNA-polymerase synthesizes the complementary strand of a single stranded DNA template by incorporation of the four desoxynucleotides and differently labelled dideoxynucleotides (lacking a 3'-OH group), which terminate the elongation; the resulting DNA strands of different length are labelled with one of the four fluorophores at the 3' end depending on the last incorporated nucleotide and can be separated by capillary electrophoresis; the sequence of the four fluorophores in the chromatogram represents the base sequence; read length ~800 bp | Gut microbiota (Hayashi 2006; Bibiloni 2006; Bik 2006; Ozutsumi 2005; Hayashi 2002; Lan 2002; Delgado 2006); novel species (Mohan 2006) |
| Pyrosequencing (sequencing by synthesis) | Visualization of incorporation of individual nucleotides at discrete positions in the sequence utilizing DNA polymerase, an ATP sulfurylase, a luciferase, and an apyrase; primer extension in presence of the appropriate nucleotide by DNA polymerase releases a pyrophosphate that in the presence of adenosine 5'-phosphosulfate is converted to ATP, which in turn drives a chemiluminescent reaction with the luciferase; read length ~100 bp | Method (Ronaghi 1998; Ronaghi 2002); influence of antibiotics on gut flora (Dethlefsen 2008); gut microbiota in obesity and gastric bypass (Zhang 2009a); susceptibility to pathogens (Stecher 2010) |
| Cloning with subsequent sequencing | A gene of interest (e.g. 16S rRNA gene) is PCR-amplified from an environmental source and ligated into vectors; the vectors are taken up by suitable host bacteria, which are cultivated and selected for vector uptake; single clones are picked and the introduced DNA strand is sequenced | Gut microbiota (Hayashi 2006; Bibiloni 2006; Bik 2006; Ozutsumi 2005; Hayashi 2002; Lan 2002; Delgado 2006) |
| Using allele-specific hybridization | | |
| FISH | Combines identification of bacteria with fluorescently labelled probes in solution and enumeration of whole cells, avoiding the problem of correct re-calculation of cell numbers from fluorescence signals like in array- and PCR-based techniques; the combination with flow cytometry has proven to be most suitable for analysis and quantification (Amann 1990); an alternative to the combination with flow cytometry is the microscopy-based enumeration of fluorescently labelled bacteria, which is going to be automated as well (Thiel 2005) | predominant groups of bacteria in human intestine (Rigottier-Gois 2003b; Lay 2005; Fallani 2006; Langendijk 1995; Manz 1996; Hold 2003; Harmsen 2000; Franks 1998; Schwirtz 2000; Aminov 2006) |
| Dot-blot hybridization | The sample is blotted to a nylon membrane and hybridized with a labelled probe, generally a radioactive label; detection of rRNA gives higher sensitivities than rRNA gene probes (Malinen 2003) | study human faecal flora (Dore 1998; Hopkins 2001; Marteau 2001; Sghir 2000; Suau 2001) |
| DNA-Microarray | DNA oligonucleotides (=probes) are immobilized to a planar surface (often glass) in distinct microscopic spots; labelled target DNA from a sample is applied to the array surface, where hybridization with complementary probes occurs; acquisition of the signals results in a specific hybridization pattern, which allows identification of the target DNA | study human faecal flora (Boesten 2009; Paliy 2009; Rajilic-Stojanovic 2009; Palmer 2007; Wang 2004a) |

Of major interest for the investigation of the intestinal community are quantitative methods. Fluorescence in-situ hybridization (FISH), dot-blot hybridization, microarrays, and real-time PCR have been advanced to detect bacteria quantitatively based on allele-specific hybridization. Therefore, these techniques require sequence information to serve as tools for this purpose. FISH probes, which detect complementary genetic sequences directly *in-vivo* or in biopsies, were developed for all predominant groups of bacteria in the human intestine (Langendijk 1995; Manz 1996; Franks 1998; Schwirtz 2000; Harmsen 2002; Hold 2003; Lay 2005; Fallani 2006; Aminov 2006). Franks et al. enumerated the predominant human faecal bacteria covering two third of the faecal flora (Franks 1998). Quantitation of bacteria can be realized by combination with flow cytometry. An alternative to this is the microscopy-based enumeration of fluorescently labelled bacteria, which is going to be automated as well (Thiel 2005). The specificity of FISH probes in complex intestinal samples is difficult to achieve. By applying mismatch competitor probes Lay et al. improved specificity of several group-specific probes for members of the *Clostridium leptum* subgroup (Lay 2005). The validation of FISH probes is often a problem due to a lack of pure cultures. Therefore, a Clone-FISH approach was developed by Schramm et al., which allows probe validation with clones containing the desired target as insert. This approach is also useful for screening of clone libraries (Schramm 2002).

Probe-based detection of sequence signatures was transferred to solid surfaces in dot-blot hybridization. Dot-blot hybridization was intensively used to study the human faecal flora (Dore 1998; Hopkins 2001; Marteau 2001; Suau 2001). A relative quantification of cell numbers from signal intensities is often done using an universal probe as reference, but the results may be influenced by hybridization efficiencies of the specific probes (Sghir 2000). Quantitative inconsistencies were also found in a comparison of dot-blot hybridization and FISH/flow cytometry. For two bacterial groups, *Bacteroides* and *Atopobium*, significant different values were measured, whereas other bacterial groups gave same results. This finding was ascribed to the different parameter, which are measured by both methods. Dot-blot hybridization reflects metabolic activity by detecting rRNA levels, whereas FISH measures proportions of cells (Rigottier-Gois 2003a).

Although PCR-based methods showed to be biased, real-time PCR has developed to a reliable identification and quantification method. For quantification, the exponential phase is used instead of an end-point determination, which reduces the typical bias due to different amplification efficiencies. Real-time PCR is widely applied for species identification in research and clinical routine. It was applied to detect resident intestinal bacteria (Haarman 2005; Matsuki 2004; Malinen 2003; Huijsdens 2002), for analysis of the microbial colonization of humans (Malinen 2005; Tanaka 2009) and animals (Skanseng 2006), as well as pathogen identification in food (de Oliveira 2010; Hruskova 2009; Omiccioli 2009; Fu 2005) and on food contact surfaces (Mafu 2009). The limited multiplexing capacity has been partly overcome by introducing technical devices with higher multiplexing capacity. Three detection methods are frequently used: intercalation of a fluorophore in the DNA double strand, the quenching-based TaqMan® probe approach, or the FRET-based LightCycler® probes approach. The latter two techniques enable parallel detection of different species by using more than one fluorophore. Moreover, this concept has the advantage of a higher specificity due to the usage of additional probes than two primers only. On the other hand, intercalating fluorophores allow subsequent melting curve analysis of the PCR-product for unspecific formation of product. Recently, a method was published that combines specific, fluorescently labelled probes with an intercalating dye without interfering PCR efficiency (Lind 2006). A comparison of real-time PCR and dot-blot hybridization showed a similar specificity of both techniques but higher sensitivity for real-time PCR with a detection limit of $4 \cdot 10^2$ compared to 10^7 genome equivalents, for pure cultures respectively. Quantification of single species with real-time PCR in a background of faecal DNA was successful for concentrations of 10^6 cells. Dot-blot hybridization with rRNA genes as target was limited to the detection of a 3% subpopulation of a faecal sample (Malinen 2003).

Today's state-of-the-art for intestinal community analysis is high-throughput sequencing. As described before, sequencing of clone libraries is a preferred method to gain new information about the composition of the intestinal microbiota. Standard Sanger (dideoxy) sequencing was the common method, by which most of the current data were gathered. Now, pyrosequencing (Ronaghi 1996) (sequencing by synthesis) has introduced a new era of sequencing. Pyrosequencing allows real-time detection of the synthesis of a DNA strand and by this, its sequence is determined. The main advantage is the opportunity to perform multiple sequencing reactions in parallel in high-density microwell plates, which dramatically increases the throughput in sequencing. This omits the main drawback of this method, which is the short read length of 100-200 nt compared to Sanger sequencing (800-1,000 nt). Two principle approaches are followed to compile a microbial map of the intestinal tract, i.e. microbial identification via single genetic targets (e.g. 16S rRNA gene) or whole-genome methodologies (Petrosino 2009). The most challenging and most attended project in this regard is the Human Microbiome Project that was approved in May 2007 by United States National Institute of Health (Turnbaugh 2007). The aim of this project, to which high-throughput sequencing will significantly contribute, is to map the complete intestinal community with its genome and the intestinal metabolome.

Microarrays are frequently used for analysis of the intestinal ecosystem and have much contributed to our current knowledge about the composition of the microbiota. Especially phylogenetic assays, like HITChip (Rajilic-Stojanovic 2009), allow in-depth analysis of spacial, temporal, and individual differences in the GI tract microbiota. Kajander et al. investigated the influence of probiotic supplementation on the gut flora in IBS (irritable bowel syndrome) patients (Kajander 2008). Smaller low-density arrays were developed to examine only dominant members of the intestinal flora (Wang 2004a; Lehner 2005; Boesten 2009). A ligase detection reaction in combination with universal cZipCodes was used by Candela et al. to compile a high taxonomic level fingerprint of the human intestinal microbiota (Candela 2010). Additionally, expression microarrays were used to study intestinal transcriptional responses to colonization with bacteria (Hooper 2001), use of pre- and probiotics (Shima 2008), or stress factors (Schumann 2005). Microarray analysis was shown to reveal greater diversity in environmental samples than sequencing of a typically sized clone library (Desantis 2007). However, a comparison of a phylogenetic array with pyrosequencing showed a strong correlation between these two culture-independent methods indicating their robustness relative to each other, as well as their capacity for in-depth profiling of microbial communities (Claesson 2009). Microarrays are described in detail in chapter 1.5.

1.3 Bacterial gastroenteritis

Gastroenteritis is a general term for a pathological state that manifests in diarrhea, nausea, vomiting, and fever and can quickly become epidemic. Gastroenteritis can be caused by a variety of pathogens including viruses, bacteria, and protozoa and is boosted by food intolerance, antibiotic treatment, and intestinal diseases like inflammatory bowel disease or irritable bowel syndrome (Farthing 2002). Bacteria have a main relevance concerning children's diarrhea (Koletzko 2009), traveller's diarrhea (Gascon 2006; Shah 2009), local outbreaks evoked by food or water contamination, and antibiotic-associated bacterial overgrowth. Bacteria are responsible for about 5% of all reported cases in adults (Baldi 2009) and 20% in children (Koletzko 2009). Between 12 and 20%, depending on living region, of all death among children aged <5 years in 2004 were caused by diarrheal diseases, which means 1.5 million children each year (WHO 2009). Diarrhea caused by contaminated water is one of the major reasons for death in the developing world and is occurring regularly as concomitant effect on natural phenomena like monsoon floods. Epidemical cases have been reported in 2006 for Ethiopia, Sudan, and Angola with case fatality rates between 1 and 3.8% all caused by a *Vibrio cholerae* strain and involving thousands of people (WHO Epidemic and Pandemic Alert and Response). In industrialised countries diarrheal disease is one of the

major reasons for sickness work absence and in public facilities, like hospitals, nursing homes, school canteens, and kindergartens enteric pathogens can cause outbreaks which are life-threatening to young children and the elderly (Kanazawa 2007; Much 2009; Kirk 2010; Robert-Koch-Institut 2009). Due to changes in the consumer needs regarding the availability of products throughout the year and the global food supply, food-borne infections are of increasing concern (DuPont 2007). A study of the European Commission, for example, revealed an average prevalence of 31% for the occurrence of *Salmonella* in laying hen stocks in 25 European countries (Robert-Koch-Institut 2007b). *Campylobacter* spp., pathogenic *Escherichia coli*, *Salmonella* serovars, *Yersinia*, and *Shigella* spp. represent the most common bacterial causes of diarrhea in humans (Fig. 1.3). In 2006 the incidence (reported cases/100,000 persons) of infections with these pathogens in Germany accounted for 63.1, 63.8, 7.8, 6.3, and 1, respectively (Robert-Koch-Institut 2007a). However, the rate of underreporting is substantial in case of acute watery diarrhea, as it is caused by *Salmonella* and *Campylobacter* for example. By population survey it was estimated that in the United States only 10% of all *Salmonella* infections with bloody diarrhea and 1.5% of infections with non-bloody diarrhea were reported to health care institutions (Voetsch 2004).

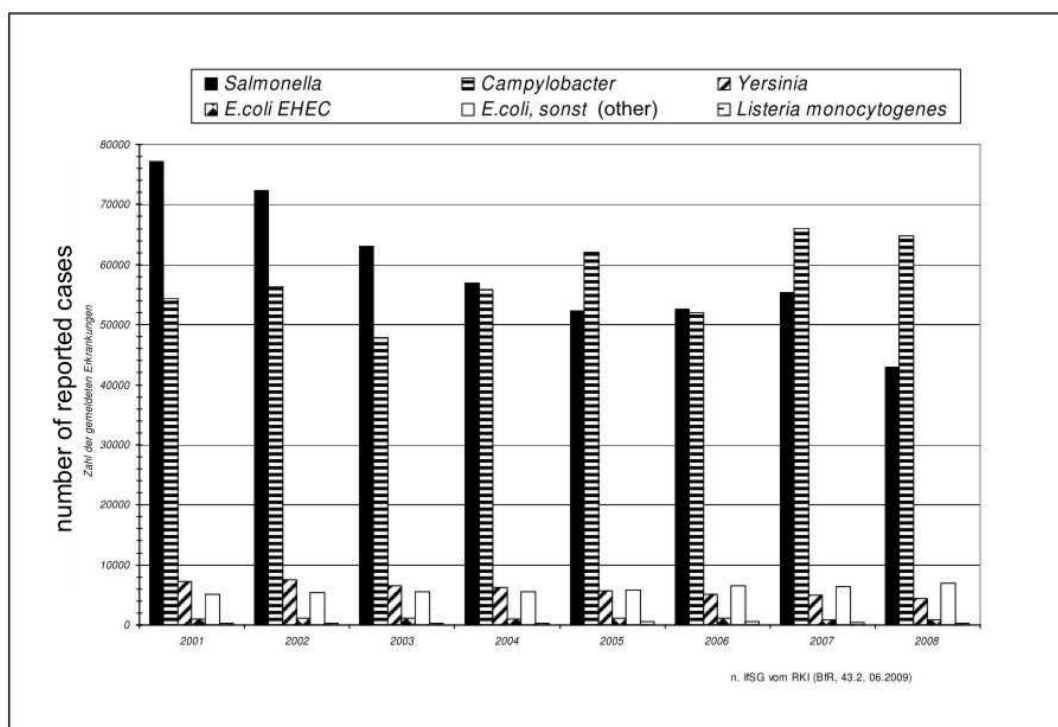


Fig. 1.3: Gastroenteritic infections in humans 2001-2008 in Germany. From *Bundesinstitut für Risikobewertung* Germany (BfR 2010).

In the last decade one pathogen, *C. difficile*, has attracted much attention as a nosocomial pathogen of increasing importance due to new resistant populations and increasing incidence (Fordtran 2006; McFarland 2008; Kelly 2008). The colonization rate in hospitalized individuals with *C. difficile* is much higher than generally in adults with 20-30% compared to 3%, respectively (Bartlett 2002). The pathogen is often associated with antibiotic therapy. Spores are found frequently in hospitals and can proliferate in the human intestine if the normal gut flora is suppressed by antibiotics, but also elderly and immunocompromised patients are susceptible. In intensive care units approximately one third of patients develop diarrhea during their hospitalization (Ringel 1995), which is often attributed to *C. difficile*. The financial impact of *C. difficile*-associated disease (CDAD) on the healthcare system has been estimated to be €3,000 million/year for Europe and \$1.1 billion/year in the US and is expected to double over the next four decades (Kuijper 2006).

Tab. 1.2: Most common bacterial enteropathogens that cause enteric disease and their pathogenic features. (continued on next two pages)

| Pathogen | Infective dose/ incubation time/ duration | Pathogenesis | Clinical symptoms | Source | Therapy |
|------------------------------|--|---|---|---|--|
| <i>Aeromonas hydrophila</i> | unknown/ unknown/ Up to several weeks (dysenteric) | Enterotoxin Act is haemolytic and destroys the intestinal epithelium via cytokines and activation of the arachidonic acid pathway; enterotoxin Alt and Ast cause increased cAMP and prostaglandin levels; various haemolysins like AerA, HlyA, Ahh1 and Asa1 increase the intestinal cAMP level | Acute watery diarrhea or dysenteric bloody diarrhea with mucus | Surface water, brackish water, food in contact with contaminated surface water (fish, shellfish), beef, swine, lamb, poultry | Treat as shigellosis |
| <i>Campylobacter spp.</i> | 400-500 cells/ 2-5 d/ 7-10 d | Little is known about the pathogenicity and associated factors; only rare expression of classical virulence factors; invasion in and translocation across the epithelial cell barrier of the intestine; disruption of tight junctions; may also induce translocation of commensal microbiota; produces cytolethal distending toxin (CDT) with three subunits; CdtB enters the nucleus and causes cell cycle arrest; CDT causes production of IL-8, which recruits macrophages, dendritic cells, and neutrophils leading to inflammation | Acute watery diarrhea (sometimes bloody), often with fever or dysenteric characteristics, abdominal pain | 80% foodborne, often poultry, contaminated drinking water, international travel | None; in severe cases azithromycin or erythromycin |
| <i>Clostridium difficile</i> | ?/ Difficult to determine due to possible previous colonisation, few days/ 2-3 d | Toxin A causes an inflammatory reaction, hypersecretion of fluid, and haemorrhagic necrosis of the ileal gut mucosa by triggering the cytokine release of neutrophils; toxin B depolymerises actin, a major protein of the cytoskeleton, and by this causes extensive tissue destruction | Watery, self-limiting diarrhea, often with fever (28%) or dysenteric characteristics, and leukocytosis (50%) after administration of antibacterial drugs in elderly patients with coexisting conditions, abdominal pain (22%), pseudomembranous colitis, mortality 1-2% (higher in the elderly) | Faecal-oral via person-to-person contact or contact with contaminated surfaces, often nosocomial, after administration of antibiotics | Metronidazole; in severe cases vancomycin; stop of antibiotic administration |

Continuation Tab. 1.2

| Pathogen | Infective dose/ incubation time/ duration | Pathogenesis | Clinical symptoms | Source | Therapy |
|--|---|--|---|--|--|
| Enterohemorrhagic <i>Escherichia coli</i> (EHEC) | 10 cells/ 2-10 d HUS 7 d/ 3-8 d | toxins verotoxin (VT), shiga-like toxin are closely related or identical to the toxin produced by <i>Shigella dysenteriae</i> | Watery diarrhea, vomiting, fever, abdominal pain, 10-20 % haemorrhagic colitis with bloody diarrhea, 5-10 % haemolytic-uremic syndrome (HUS) with 3-5% mortality (especially children), ischemic colitis in the elderly | Often foodborne (52%) (ground beef), person-to-person spread, contaminated water, contact with animals | None; anti-bacterial therapy can stimulate toxin production; symptomatic therapy |
| <i>Listeria monocytogenes</i> | ≤1,000 cells/ Unknown (>12h)/ 2-4 weeks | may invade the gastrointestinal epithelium; enters the host's monocytes, macrophages, or polymorphonuclear leukocytes; Listeriolysin (toxin) production | Fever, nausea, vomiting, diarrhea; septicemia, meningitis (or meningoenzephalitis), encephalitis | Food (raw milk, cheeses, ice cream, raw vegetables, raw meats (all types), and raw and smoked fish) | Penicillin, ampicillin, trimethoprim-sulfamethoxazole |
| <i>Mycobacterium avium</i> (paratuberculosis /complex) | | Attaches (fibronectin attachment proteins) intestinal tissue and enters through M cells in the Peyer's patches; phagocytosis by subepithelial macrophages; induces TNF α -mediated apoptosis of macrophages; carried to the local lymph nodes by lymphatics | MAP seems to be associated with Crohn's disease and occurs often in AIDS patients; diarrhea, abdominal pain, fever, fatigue, weight loss, anaemia and neutropenia, lymphadenitis | Drinking water (especially in hospitals), perhaps milk, raw fish and shellfish, indoor swimming pools, | Surgical excision of lymph nodes; clarithromycin and rifabutin |
| <i>Plasiomonas shigelloides</i> | >1,000,000 cells/ 20-24 h/ 1-7 d | A heat-labile and a heat stable enterotoxin have been described | usually mild and self-limiting with fever, chills, abdominal pain, nausea, or vomiting; watery, non-mucoid, and non-bloody diarrhea; infections in western countries rare; occur preferentially in immunocompromised (here massive watery, sometimes bloody diarrhea) | Contaminated water, raw shellfish | Treat as shigellosis |

Continuation Tab. 1.2

| Pathogen | Infective dose/ incubation time/ duration | Pathogenesis | Clinical symptoms | Source | Therapy |
|--------------------------------|---|--|---|--|---|
| <i>Salmonella enterica</i> | 10-20 cells/ 6-48 h (-7d)/ 1-2 d | Penetration and passage from gut lumen into epithelium of small intestine, mediated by type III secretion system (T3SS), where inflammation occurs; enters non-phagocytic enterocytes and M-cells; may disrupt tight junctions; recruiting of neutrophils to the lumen leading to inflammation | gastroenteritis (75 %), fever (50 %), bacteraemia (5-10 %), bloody diarrhea (sometimes) Acute watery, sometimes bloody diarrhea, often with fever, occasionally with dysenteric characteristics | Contaminated food (eggs, milk and dairy products, raw meats, poultry, fish, shrimp, sauces and salad dressing) | None or ceftriaxone or azithromycin (children); none or levofloxacin or azithromycin (adults) |
| <i>Shigella spp.</i> | 10 cells/ 12-50 h/ 3 d – 1 (-4) week | Attach to, and penetrate, epithelial cells of the intestinal mucosa; invasion; multiply intracellularly; spread to contiguous epithelial cells; tissue destruction; some strains produce enterotoxin and Shiga toxin | Severe, bloody diarrhea, often with fever or dysenteric characteristics, abdominal pain; seldom haemolytic-uremic syndrome (HUS) | Faecal-oral, contam. food (salads, raw vegetables, milk, dairy products, poultry) or water | Azithromycin or ceftriaxone (children), ciprofloxacin or azithromycin (adults) |
| <i>Vibrio cholerae</i> | 1,000,000 cells/ 6 h – 5 d/ 6-7 d | Cholera toxin (CT) binds to GM1 gangliosid, a glycolipid found in the apical surface of intestinal epithelial cells; internalization; activation of adenylate cyclase; cAMP increase; imbalanced electrolyte transport forces increased water secretion | Acute dehydrating diarrhea, abdominal cramps, nausea, vomiting, dehydration, and shock | Contaminated water, raw shellfish | |
| Noncholeraic <i>Vibrio</i> | >1,000,000 cells/ 4-96 h/ 1-5 d | attaches itself to an individuals' small intestine and uses type III secretion system to excrete virulence proteins into host cells; | Mild or moderate, watery diarrhea often with dysenteric characteristics, abdominal cramps, nausea, vomiting, headache, fever, and chill | Shellfish and seafood | None or treat as shigellosis |
| <i>Yersinia enterocolitica</i> | unknown/ 24-48 h/ 1-3 weeks/ | Penetration of the mucosa and cell invasion in the ileum mediated by invasin (<i>inv</i> -gene); multiplication in Peyer patches; production of an enterotoxin | Acute watery diarrhea, may cause fever and dysentery and a pseudo-appendicitis condition, abdominal pain | Water (ponds, lakes), food sources (meats, ice cream, milk) | |

The severity of gastroenteritis varies depending on the volume of fluid loss and the treatment. In the United States the hospitalization rate is 20% of the confirmed cases (Voetsch 2004). The mortality rate is especially high in very young children and the elderly (Much 2009; Kirk 2010). The infectious dose is highly variable depending on the health status of the host and the way of infection and ranges between 10^1 and 10^6 cells (Lamps 2007). Especially, infection via fat and sugar rich food requires lower germ numbers, as pathogens are protected from the acids and bile during transport through the intestinal system (Waterman 1998).

Three types of infections can be described: toxin-associated, superficial, and systemic gastrointestinal infections. The former ones are caused by bacteria that directly affect the transport of fluid and electrolytes by toxins, such as *C. difficile*, *V. cholerae*, *A. hydrophila*, and *P. shigelloides*. Superficial gastrointestinal infections are characterized by the ability of the pathogen to invade the superficial layers of the intestine, including the epithelial cell layer and the underlying lamina propria. These pathogens only uncommonly invade systemically and result in bacteraemia. Superficial infections are caused by *Shigella* spp., *V. parahaemolyticus*, and *Campylobacter* spp.. In systemic intestinal infections, as caused by *Salmonella* spp., *Y. enterocolitica*, *L. monocytogenes*, and EHEC, the pathogen invades through the mucosa to reach the systemic circulation and can infect other organs. The most common bacterial enteropathogens are listed in Tab. 1.2.

1.4 Diagnostic methods for clinical pathogen identification

In clinical routine detection of bacterial pathogens is a key task. The information on the causative agent of a disease influences the choice of therapeutics and answers questions on hygiene. This is of importance in cases of severe diarrhea, which is not self-limiting, and in case of infected children, who are more compromised by loss of fluids and electrolytes. Regarding intestinal pathogens, it is mainly based on stool studies and cultivation techniques in combination with morphological and physiological analysis and enzyme immunoassays. However, also genotyping methods, like real-time PCR and microarray, have made their way into clinical diagnostics as fast detection techniques, but up to now only few have the necessary in-vitro diagnostics (IVD) approval to serve as alone standing diagnostic tool. Nevertheless, genotyping methods have the potential once to replace phenotypic methods due to their speed in generating a diagnostic answer and attempts to miniaturize assays. This may result in new point-of-care (POC) diagnostic devices, which can shift the place of generating a diagnose from the laboratory to the physician and, finally, to the patient (Ince 2009).

1.4.1 Phenotypic pathogen identification

Phenotyping methods based on stool enrichment and culture are the gold standard for intestinal pathogen identification in clinical routine. It is a complex and time-consuming task, which requires well-trained personnel. Stool studies for leucocytes, blood and mucus give first information if a bacterial pathogen is likely to be the cause of diarrhea (Hoshiko 1994). Especially bloody diarrhea is a major diagnostic challenge for the clinician, because of the importance of distinguishing infection from non-specific inflammatory bowel disease and other inflammatory conditions of the colon (Farthing 2002). The definite identification generally requires time-consuming stool enrichment and selective culture. Indications for stool culture are the passage of six or more unformed stools per day, diarrhea for longer than three days, fever $\geq 38.5^\circ\text{C}$, dysentery, bloody stool, and multiple cases that suggest an outbreak (DuPont 2009; Chan 2003). Pathogen identification by stool culture usually takes between 12 and 48 hours or even longer depending on the pathogen. The usage of selective enrichment media and growth conditions allows first restriction of the potential etiologic agent. Identification comprises then (I) morphological, (II) biochemical, (III) physiological, and

(IV) serological characteristics. The characterization of the morphology includes colony shape, colony dimension, pigmentation, cell shape, Gram reaction, flagellation, and others. Biochemical properties are defined by the presence of catalyzing enzymes like catalase (*Campylobacter* spp.), oxidase (*Campylobacter* spp.), nitrate reductase, or glutamate dehydrogenase (*C. difficile*). Physiological properties are e.g. the carbohydrate fermentation, indole production, urea splitting and growth under miscellaneous physiological features (temperature, pH, salt, gaseous environment). Serological tests, using immunological methods, for the presence of surface antigens, e.g. somatic (O-antigen), flagellar (H-antigen), or capsular (K-antigen) antigens, give valuable information for the identification of a pathogen and allow epidemiological conclusions. For salmonellae, more than 2,000 serovars based on H- and O-antigens were described. *E. coli* is classified into ~250 serovars by O-, H-, and K-antigens.

This phenotypic search for and characterization of a pathogen requires a proper algorithm to reduce the number of required tests and by this the cost of a diagnostic result. Moreover, phenotypic pathogen identification is accompanied by many difficulties. Some pathogens, such as *Vibrio* species, require special media and are more difficult to culture and routine stool cultures will for example not distinguish between EHEC and non-pathogenic *E. coli* from the normal intestinal flora. Slow growers, such as enteropathogenic mycobacteria, make the diagnostic process lengthy. Bacteria such as *Campylobacter* spp. require a rapid transport to the laboratory to ensure vitality and some bacteria, such as *Yersinia enterocolitica*, require a several days lasting cold enrichment. Problems in disease diagnosis will also occur in case of failed cultivation of the potential pathogen. Although the histological diagnosis of the infectious process plays a valuable role, acute infectious-type colitis is often indistinguishable from other inflammatory conditions of the gut such as ischaemia or chronic idiopathic inflammatory bowel disease (Lamps 2007). Furthermore, many cases of gastroenteritis caused by *Vibrio* spp., a pathogen that is mainly related to seafood, are under-recognized because culture of this species is not routinely done in clinical laboratories (Chan 2003).

The culture-based diagnostic process is facilitated by automated systems, such as VITEK 2 (BioMérieux, France), Phoenix (Becton Dickinson, USA), and Dynal eAIMS (Invitrogen), which allow fully automated cultivation, pathogen identification based on colorimetric tests, antibiotic susceptibility testing and analysis of resistance mechanisms, but they still suffer from the problems mentioned above.

1.4.2 Genotypic pathogen identification

Genotyping methods are a promising alternative to the current culture-based diagnostic methods for pathogen detection. The long time, which is required for identification of the causative agent by phenotyping, generally forces the physician to start an empirical therapy, since acute diarrhea strongly affects the life quality and can become life threatening. Therefore, the results of stool culture are often of no therapeutic consequence anymore (Koletzko 2009). This was also described in a study, where 88.5% of gastroenteritis outpatients had recovered by the time that culture positive results were available, and thus a change of antimicrobial therapy was not required anymore (Chan 2003). For different bacterial pathogens varying antibiotics are recommended and non-adequate therapy can increase the risk of developing the haemolytic-uremic syndrome (DuPont 2009). Additionally, the mentioned study found a higher level of Ciprofloxacin resistance in *Campylobacter* spp. in the 130 stool culture positive samples, underpinning the need for an evidence-based therapy. This is supported by other studies and a permanent increase of resistance among enteric pathogens can be observed (Garcia 2009; Senok 2007; NARMS 2009; Threlfall 2006). Evidence-based therapy is, therefore, of main importance in terms of a cautious usage of antibiotics. Prescription of unnecessary or unspecific antibiotics promotes the spread of resistance determinants (Donskey 2006). Moreover, an early identification of a bacterial pathogen allows changes in clinical management. Cost-intensive separation of a patient, which is for example required in case of a potential norovirus infection, is not

necessary anymore, if a bacterium is identified as causative agent. This can reduce hospitalization time and the cost of medical care. In one study, genotyping methods have also proven to increase the detection rates and by this to close a diagnostic gap (Ajjampur 2008).

To overcome the limitations of classical identification methods, that is low speed and accuracy, poor reproducibility, and intensive labour of trained personnel, new techniques that identify bacteria based on their genetic information without prior cultivation have been developed, such as FISH (fluorescence in-situ hybridization), DHPLC (denaturing high performance liquid chromatography), microarrays, real-time PCR and high-throughput sequencing. In principle, all methods that were used for the investigation of the intestinal flora (chap. 1.2) could be adopted for pathogen detection. An all-embracing overview cannot be given in this work, due to the amount of published methods for gastroenteritis-related pathogens. From the clinical point of view, however, only some methods have practical applicability, which is mainly a question of the methods cost, multiplexing capacity, handling efficiency, and speed. The increased sensitivity of nucleic acid-based tests requires improved contamination prevention and quality control in the clinical laboratory.

The adaption of amplification strategies based on PCR (Ke 1999), such as real-time PCR, RT-PCR (Zheng 2008; Zeng 2008), nested PCR (Chen 2000), PCR-ELISA (Sails 2001), and multiplex-PCR (Farfan 2010; Espineira 2010) was the first step towards intestinal pathogen genotyping in clinical laboratories. Assays have not only been described for clinical specimens but also for food and water, which can be the source of an intestinal infection. An isothermal amplification strategy, nucleic acid sequence-based amplification (NASBA), has been described for detection of viruses from faecal specimens (Lamhoujeb 2009) and *Mycobacterium avium* subsp. *paratuberculosis* from water and milk (Rodriguez-Lazaro 2004). Real-time PCR, which monitors the amplification reaction in real time, is to date the most relevant method in clinical laboratories. It has already made its way into medical laboratories as a fast identification method with reliable quantification ability, using specific primers for selective amplification of expected pathogens. In LightCycler technology it is combined with melting curve analysis of the PCR products (Lyon 2009). The main advantage of real-time PCR is its capability of quantifying the pathogen and its high sensitivity down to 10 or fewer copies of target (Liu-Stratton 2004). Assays have been developed for *C. jejuni* (Skanseng 2006), *C. coli* (Keramas 2004), EHEC (Hsu 2005; Fu 2005), *Listeria* (Huijsdens 2003), *C. difficile* (Penders 2005), and some pathogenic protozoa (Blessmann 2002; Limor 2002; Verweij 2004; Ng 2005). It is mainly restricted by its limited multiplexing capacity, which is due to two principle reasons: First, the interference between multiple primer pairs or probes and, second, the number of different fluorophores that can be used for simultaneous non-overlapping detection in a single tube, which is determined by the number of available channels (Petrik 2006). New devices, however, improved the multiplexing capacity by offering many PCR reactions in parallel using low sample volumes (LightCycler 1,536 Real-Time PCR System, ABI PRISM 7900HT Fast Real-Time PCR System 384-well, Fluidigm BioMark System 9,216-plex). Assays for the detection of *Campylobacter*, *Salmonella*, *L. monocytogenes*, *Listeria*, *E. coli* O157 (Roche Applied Science 2010), and *C. difficile* (Prodesse ProGastro Cd/ GenProbe; Xpert *C. difficile*/ Cepheid) are commercially available.

LUMINEX's xMAP technology is a method combining flow cytometry with colour-coded microspheres that allow up to 100 parallel bioassays by coating the beads with different capturing molecules (www.luminexcorp.com). The advantage of this bead-based assay is the favourable reaction kinetics in solution. Based on the LUMINEX technology detection assays were described for the most common food-borne pathogens (Dunbar 2003; Fitzgerald 2007). However, IVD approved assays are not yet available for these pathogens.

DHPLC was used for the identification of genitourinary tract pathogens (Domann 2003) or *Candida* infections of the blood or GI tract (Goldenberg 2005). However, a parallel identification of more than 11 species based on the retention time only, was not yet shown. When analysing complex faecal samples, Goldenberg et al. were not able to assign

individual peaks to species without subsequent sequencing due to day-to-day variations of retention times (Goldenberg 2007).

Mass spectrometry has also been shown to detect bacterial pathogens reliably (Gielen 2007; Mandrell 2005). Very fast detection times of few minutes could be realized already by the MALDI Biotyper (Bruker Daltonics), which identifies bacteria on the spectra of 16S rRNA genes. Nevertheless, the method still relies on over-night culturing of the pathogen, which makes it a good supplementation of culture-based diagnostics without replacing it. The main disadvantages are the high cost of the instrumentation and the impossible integration in small point-of-care devices.

Sequencing methods have a high potential for fundamental research on the composition of the intestinal microbiota (chap. 1.2). Although several studies indicated that sequencing of the ribosomal genes could provide an effective diagnostic tool (Kolbert 1999), this method has not yet established in clinical diagnostic due to its user-intensive highly technical nature. Although high-throughput sequencing has overcome limitations of Sanger sequencing, e.g. low multiplexing capacity, it is still not suitable for most clinical applications due to high cost and required time, which may be overcome in future. Clinical applications of sequencing for pathogen identification are usually based on previous cultivation and/or 16S amplification or amplification of a structural gene (Hou 2008; Gleesen 2008; Justesen 2010; Luna 2007). It was assumed that the most probable application will be a combined one with established methods of pathogen identification and cultivation (Luna 2007). However, sequencing technology is underway to serve as full diagnostic tool.

Various nucleic acid hybridization-mediated methods have been developed to detect intestinal pathogens. A peptide nucleic acid-FISH (PNA-FISH) assay was published for *Salmonella* spp. identification from different samples (Almeida 2010). Although the assay showed high specificity, it was still lengthy, because an over-night enrichment step was required. A line-probe assay, which uses immobilized probes on a paper test strip, is available for the detection of mycobacteria (Innogenetics, Belgium) but not yet for intestinal pathogens.

Microarrays have huge potential for pathogen genotyping in clinical diagnostics, due to their ability to generate much genetic information in one experiment, in short time, and with relatively low technical effort. Several microarrays for cancer marker analysis and pathogen detection have been marketed yet. A concept for on-chip sequencing with overlapping probes, which cover the entire diagnostic sequence, was published for pathogen detection (Wilson 2002). Microarray technology and its application are described in detail in chapter 1.5.

1.5 DNA-microarrays

Microarrays are a genotyping method that uses immobilized, short DNA oligonucleotides for the detection of genes, genetic modifications, or expression levels of genes by hybridization with complementary sequences. It combines nucleic acid amplification strategies with the multiplexing capability of a structured, surface-coupled detection. One array may contain up to 100,000 distinct spots. Two types of microarrays can be defined: genomic arrays and arrays for genetic testing. The former include gene expression arrays and splicing arrays. The latter designates the determination of presence or absence of a certain DNA sequence for identification purposes or detection of genes and single point mutations leading to increased pathogenicity. The focus in this chapter will be on the former. Numerous arrays for genetic testing have been published to date (Leinberger 2005; Castiglioni 2004; Palmer 2006; Palacios 2007; Barl 2008; Leinberger 2009; Wilson 2002; Li 2006) including microarrays for the detection of intestinal bacteria (Wang 2004a; Lehner 2005; Boesten 2009; Paliy 2009; Rajilic-Stojanovic 2009) and pathogens (Jin 2006; Majtan 2007; Kostic 2007; You 2008; Mao 2008; Li 2006).

Three platforms are frequently used for microarray technology: glass, synthetic membrane, and gel-pad arrays on glass (Pozhitkov 2007). DNA oligonucleotides, so-called probes, are immobilized to the surface and serve as specific recognition elements for complementary DNA, which is the target. Alternatively, PCR products (Schena 1995), cDNA, plasmid DNA (Call 2006), bacterial artificial chromosomes (Ishkanian 2004), and whole genome DNA (Bae 2005) can serve as probe. Probe and target should have high affinity and specificity for each other. This need is reduced for individual probes in resequencing arrays, which contain tiled probes covering a diagnostic region (Lin 2006). Usually, the target is labelled and detected upon hybridization with a complementary probe. The hybridization pattern allows identification of the target. Additionally, the observed signal provides a measure of the amount of bound target and indirectly the amount of target DNA in the sample.

Two main techniques for the production of the arrays are in use, i.e. direct in-situ synthesis of the probes using a lithographic procedure or deposition and immobilization of pre-synthesized probes on the array surface. The former principle was developed by Affimetrix (GeneChip®; <http://www.affymetrix.com>) and allows production of high-density arrays. Affimetrix's GeneChip® was the first approved by the FDA for genetic testing. High-density arrays are mainly applied for expression microarrays or phylogenetic arrays, which have outstanding importance for clinical science. In clinical diagnostics, however, low-density arrays printed with pre-synthesized oligonucleotides are more convenient, because they are relatively easy to prepare, inexpensive, require simpler equipment, and can be adapted to new targets without contacting the manufacturer (Mikhailovich 2008).

A set of labelling and detection methods has been developed, in order to address problems of sensitivity, signal stability, cost, and miniaturization. The event of a DNA double strand formation can be detected using (I) optical, (II) electrical, and (III) physical properties of the DNA itself or incorporated labels. Optical hybridization detection is the most common detection principle in microarray technology. In most cases, a fluorophore is incorporated into the target DNA and sensed by a fluorescence scanner. The most common dyes are Cyanine dyes, Alexa dyes, BODIPY dyes, fluorescein, and rhodamine derivatives. New powerful labels are the luminescent semiconductor quantum dots, which are 20 times as bright and 100 times as stable as organic dyes (Chan 2002). Indirect labelling is possible by incorporation of biotin or haptens into the target DNA, which can be detected by streptavidin or antibodies coupled to a fluorophore, an enzyme, or other label. Direct or indirect labelling of the target DNA with gold nanoparticles allows silver precipitation on the gold, which is detectable with a standard flatbed scanner (Taton 2000). Alternatively, silver precipitation can also be detected electrically measuring the conductivity changes at the electrodes (Park 2002). Electronic DNA hybridization detection comprises amperometry, voltammetry, potentiometry, and impedimetry. Redox-active molecules that are incorporated in the target DNA can be detected by amperometry or voltammetry (Lin 2008; Hajdukiewicz 2010). Electrochemical impedance spectroscopy has been demonstrated to be an effective method for label-free detection of DNA hybridization (Hang 2004; Guiseppi-Elie 2005). Direct detection of the charged DNA strand was shown using a field-effect sensor (Fritz 2002). Additionally, physical properties of the DNA were successfully applied for hybridization detection. Surface plasmon resonance (SPR) (Piliarik 2009) and quartz crystal microbalance (QCM) (Mao 2006) have been used to detect a hybridization event in real-time, and can be used for label-free detection in principle. By measuring the mass-to-charge ratio of species-specific probes using MALDI-TOF (matrix-assisted laser desorption ionization-time of flight) mass spectrometry, the specific hybridization with a target can be detected (Isola 2001). In this approach, however, the principle of a microarray with immobilized probes in distinct spots is abrogated.

The main influence on specificity and hybridization efficiency emanates from the probe itself, which requires thorough design and optimization. The hybridization efficiency of probes on a microarray was technically improved by using spacer between probe and array surface (Shchepinov 1997) or changing the chemical structure of the probes. Peptide nucleic acid (PNA) probes (Braasch 2002), for example, have an artificial peptidic backbone, which is

more stable under stringent hybridization conditions and is not easily recognized by nucleases or proteases. The neutral nature of the backbone allows hybridization at low ionic strength, which is favourable, because secondary structures in the target molecule are eliminated. LNA probes are modified RNA molecules with a bridge that locks the ribose in the 3'-endo conformation, which increases thermal stability of duplexes and allows hybridization at higher temperatures (Braasch 2001). Molecular beacons are hairpin shaped probes with an internally quenched fluorophore. Only upon correct hybridization the quenching is revealed, which increases specificity of the detection. Molecular beacons can be applied in solution but also surface bound (Yao 2004). The principle of ligase detection reaction (LDR) was transferred to the array surface by the universal ZipCode approach (Busti 2002). Two probes, one containing a label and the other coupled with a unique sequence, bind the target in adjacent position and are ligated, if the match is perfect. The ligated product contains both, the label and the zip-sequence, and is directed to the array surface, which contains probes complementary to the different ZipCodes. This approach increases specificity of the target detection.

For further review of the basic principles of microarray technology see Dufva (Dufva 2009a; Dufva 2009b).

In principle, a microarray experiment comprises following steps: (I) nucleic acid purification, (II) optional target gene amplification, (III) target labelling, (IV) hybridization to the microarray, (V) signal read out, and (VI) data processing comprising quantification and analysis (Fig. 1.4). Although amplification of the target is not required for detection, it is usually done to guaranty a sufficient sensitivity, which is most important for clinical applications. Additionally, the PCR amplification allows incorporation of labelling molecules for detection of the hybridization event. Besides the choice of target sequences, this step is often the bottleneck in designing microarray-based test for the detection of multiple targets in parallel, because ideally a multiplex PCR would be required to keep the full multiplexing capacity of the method.

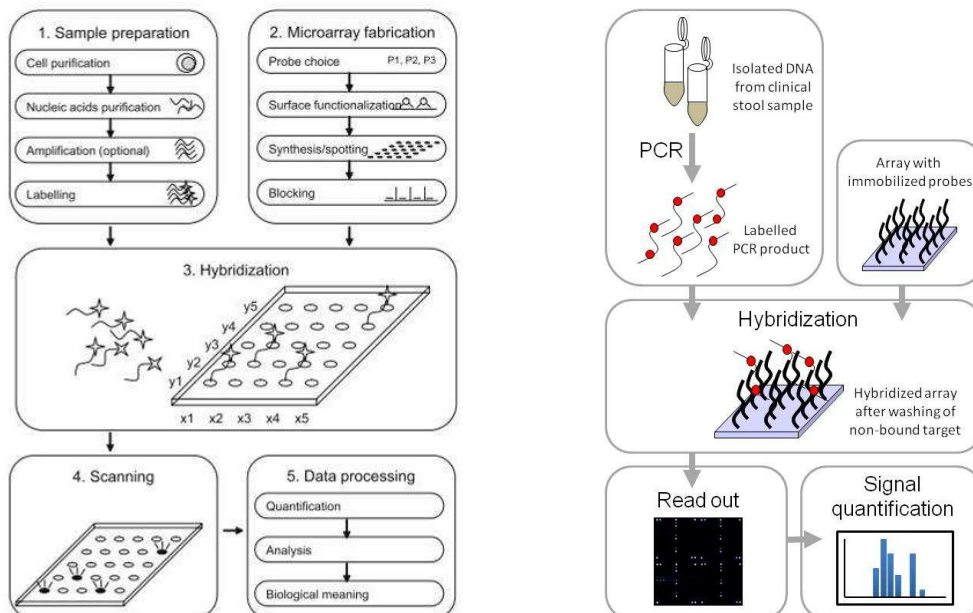


Fig. 1.4: (left) Steps of microarray production and usage. From (Dufva 2009b). (right) Principle of pathogen detection from clinical faecal samples using a DNA microarray.

The application spectrum of microarrays for genetic testing is diverse including arrays for environmental analysis, food and water control, animal and plant pathogens, clinical pathogen detection, and detection of pathogenicity determinants. Ecological community studies including the intestinal tract have been performed using large phylogenetic arrays based on 16S rRNA gene (Desantis 2005; Palmer 2006; Rajilic-Stojanovic 2009). Smaller arrays detect restricted groups of organisms like cyanobacteria (Castiglioni 2004) or the endodontic microbiota of necrotic root canals (Vianna 2005).

Recently, Mikhailovich and colleagues argued that the most-promising area for microarray application in near future is the clinical diagnostics and monitoring of infectious diseases (Mikhailovich 2008). The rising interest in this technology is reflected by the increasing amount of publications on this topic. Former limitations of microarrays in the clinical application like the relatively high costs are gradually overcome by more cost effective array support materials and alternative labelling strategies. The quality of obtained data from microarray experiments is currently actively discussed but mainly in case of expression analysis (Wilkes 2007). The genetic testing with low-density arrays and limited sets of oligonucleotides generally gives clear answers and is, therefore, highly accepted (Mikhailovich 2008). Intensive studies on the reproducibility and accuracy of microarray technology have resolved previous doubts (Shi 2008).

Expression microarrays, e.g. for the detection of cancer biomarkers (Bertucci 2008), have already longer tradition in clinical application and a new emerging field are protein microarrays for the same purpose (Lin 2010). The most promising development, however, is to expect from arrays for pathogen genotyping in infectious diseases. This covers identification of the etiologic agent as well as SNP typing (Leinberger 2009; Yu 2007; Barl 2008), which allows determining resistances and increased pathogenicity of the infective agent. Identification DNA microarrays were, for example, published for the detection of pathogens affecting the respiratory tract (Cannon 2010; Lin 2007), the intestinal tract, and the blood (Huang 2006; Leinberger 2005).

The detection of intestinal pathogens is a well-attended topic regarding microarrays for clinical purposes, which is very probably due to the current, non-satisfying diagnostic procedures and the clearly defined targets. For most of the etiologic agents of gastroenteritis arrays were constructed, which were meant for food processing industry (Myers 2006; Liu-Stratton 2004) as well as for clinical diagnostics. With respect to intestinal pathogens, food and water control and clinical pathogen detection are closely related, as many intestinal pathogens have their infection source in these media. Microarrays were published for pathogenic bacteria (Kostic 2007; Jin 2006; Mao 2008; Li 2006; Kakinuma 2003), protozoa (Wang 2004c; Wang 2005), and viruses (Jaaskelainen 2006). While the multiplexing capacity of microarrays attracts to design comprehensive arrays, such as the panmicrobial microarray for the detection of infectious diseases of vertebrates published by Palacios et al. (Palacios 2007), smaller arrays have proven to be more relevant in the clinical laboratory. Most marketed pathogen typing arrays (Tab. 1.1) do not cover more than 100 features. This facilitates the diagnostic answer and reduces assay costs.

Tab. 1.1: Overview on marketed microarrays for clinical applications.

| Company | Product name | Targets | Application | Webpage |
|-------------------------------------|--|--|---|--------------------------------|
| Affimetrix | GeneChip | Microarray platform for molecular diagnostics | | www.affymetrix.com |
| Clondiag/ Inverness (Germany) | Arraytube/ Arraystrip | Platform: custom micro probe array integrated into a micro reaction vial or a microplate compatible format | | www.clondiag.com |
| Roche (Swiss/ Germany) | AmpliChip CYP450® (CE-IVD) (based on Affimetrix Gene Chip) | SNPs in CYP2D6 and CYP2C19 genes | Drug metabolism profiling | www.roche.de |
| Chipron (Germany) | Myco Direct | <i>Mycobacterium tuberculosis</i> , atypical mycobacteria | Mycobacteria diagnostics | www.chipron.com |
| | Myco Resist | Resistance associated point mutations (12 rpoB, 2 katG), <i>M. tuberculosis</i> (6 rpoB, 1 katG) | Mycobacteria diagnostics | |
| | BacDent | 11 bacterial species | Periodontitis | |
| Mobidiag (Finland) | Prove-it™ (applied for Bacteria CE-IVD approval) | 50 bacteria, mecA gene | Sepsis | www.mobidiag.com |
| | Prove-it™ Herpes | 8 herpes viruses | Herpes | |
| Lambda / Greiner Bio-One | CytoCheck | 40 mycoplasma species | Cell culture quality control | www.greinerbioone.com |
| | ParoCheck® (CE-IVD) PapilloCheck® (CE-IVD) | 10/20 periodontal pathogens 24 genital papillomaviruses | Periodontitis Cervical cancer | |
| TessArae (USA) | RPM-Flu | 48 viral, 22 bacterial upper respiratory pathogens | Influenza | www.tessarae.com |
| | RPM-TEI | 80 CDC threat agents (bacteria, viruses, toxins) | Tropical and emerging pathogens | |
| | RPM-HFV | Genome resequencing of 61 viruses | Haemorrhagic fever | |
| CheckPoints (Holland) | Check&Trace <i>Salmonella</i> | 98 <i>Salmonella</i> serotypes | Salmonellosis | www.check-points.com |
| | Check-ESBL | 25 SNPs in TEM, SHV, CTX-M, several <i>Mycobacteria</i> | Resistance control | |
| Dr. Chip (Taiwan) | Dr. Food Kit | nine most common food pathogens <i>S. aureus</i> , <i>E. coli</i> , <i>Y. enterocolitica</i> , <i>B. cereus</i> , <i>C. perfringens</i> , <i>L. monocytogenes</i> , <i>Salmonella</i> spp., <i>Shigella</i> spp., and <i>Vibrio</i> spp. | Food quality control | www.bio-drchip.com.tw/home2eng |
| | Dr. RV IVD Kit | 10 major respiratory viruses incl. SARS | Influenza | |
| | Dr. Milk Kit | 7 mastitis causative agents | Dairy farm monitoring | |
| | Dr. EV IVD Kit | Enterovirus, two serotypes enterovirus 71 and coxsackie A16 | Hand, foot and mouth disease (HFMD), encephalitis | |
| CapitalBio (China) | Mycobacteria Identification Detection Array Kit | 17 mycobacterial species | Mycobacteria diagnostics | www.capitalbio.com |
| | <i>M. tuberculosis</i> Drug Resistance Detection Array Kit | 14 SNP related to rifampicin and isoniazid resistance | Tuberculosis | |
| GeneIn (Korea) | CombiChip | 21 human pathogenic mycobacteria, rifampicin and isoniazid resistance in <i>M. tuberculosis</i> | Mycobacteria genotyping | www.genein.com |
| Genetic Analysis AS (Norway) | G-MAP™ infant (research use) | Gut microbiota | Necrotizing Enterocolitis (NEC) | www.genet-analysis.com |

1.2 Objectives

The central aim of this work was the development and application of a DNA-microarray for the rapid, simultaneous detection of the predominant bacterial pathogens that cause gastroenteritis. As a new concept, the array was designed to provide also information about the residential flora and potential probiotics, which may be used as therapeutics. This combination should allow not only pathogen identification but also a monitoring of the patients gut flora recovery during therapy. The developed array should improve pathogen identification time and the coverage of routinely detected pathogens compared to classical culture-based techniques. An important feature was the adaptation of the array to clinical demands regarding DNA extraction from faecal specimens. The sensitivity of the Gastroenteritis-Chip should be comparable to current state-of-the-art phenotyping methods.

The research plan comprised the following elements:

- Establishment of a robust amplification strategy, which comprises all relevant bacterial pathogens and is applicable to real clinical isolates
- Design of specific probe sets, which detect the targets on two taxonomic levels
- Validation of the developed microarray with reference DNA isolates and selection of the specific probes for the final array version
- Investigation of clinical stool samples with the developed microarray and assessment of the array performance
- Investigation of the applicability of the array for the monitoring of the resident intestinal flora
- Development of a concept for the quantification of intestinal bacteria and pathogens with a microarray over several orders of magnitude, which can solve the problem of a low dynamic range of current microarray-scanner systems

2 Material and methods

2.1 Material

2.1.1 Reagents and kits

| Product | Company |
|--|---|
| Ethylenediamine tetraacetic acid (EDTA) | Fluka Chemie, Buchs/CH |
| Ethylenediamine tetraacetic acid disodium salt | Fluka Chemie, Buchs/CH |
| Potassium chloride (KCl) | Fluka Chemie, Buchs/CH |
| Sodium chloride (NaCl) | Fluka Chemie, Buchs/CH |
| Sodium citrate dihydrate | Sigma-Aldrich Chemie GmbH, Munich/GER |
| Sodium dihydrogen phosphate NaH ₂ PO ₄ | Fluka Chemie, Buchs/CH |
| Sodium hydrogen phosphate Na ₂ HPO ₄ | Fluka Chemie, Buchs/CH |
| Sodium dodecyl sulfate (SDS) | Roth, Karlsruhe/GER |
| Acetic acid (HCl; 37% v/v) | Fluka Chemie, Buchs/CH |
| Triton X-100 | Fluka Chemie, Buchs/CH |
| Tris base | Riedel de Haën, Seelze/GER |
| Ethylene glycol (25% vol/vol) | Fluka Chemie, Buchs/CH |
| EGTA (ethylene glycol tetraacetic acid) | Fluka Chemie, Buchs/CH |
| HPLC- H ₂ O LiChrosolv | Merck, Darmstadt/GER |
| Nitrogen gas | |
| 2x Spotting buffer (160 mM Na ₂ SO ₄ , 130 mM Na ₂ HPO ₄) | Eppendorf, Hamburg/GER |
| Nexterion 2x spotting buffer Spot I and Spot III | Schott Technical Glass Solutions, Jena/GER |
| Taq DNA polymerase; reaction buffer 1.5 mM MgCl ₂ ; Mg(OAc) ² | Eppendorf, Hamburg/GER |
| Expand™ High Fidelity Taq DNA polymerase, 1x reaction buffer | Roche Applied Science, Penzberg/GER |
| dATP, dGTP, dTTP, and dCTP | Amersham Biosciences, Freiburg/GER |
| Cy3-dCTP | Amersham Biosciences, Freiburg/GER |
| Biotin-11-dUTP | Fermentas, Burlington, Ontario/Can |
| DNase I, 1x reaction buffer, stop buffer with EGTA | Promega, Madison, WI/USA |
| Agarose | Sigma-Aldrich Chemie GmbH, Munich/GER |
| Xylene cyanol FF | Fluka Chemie, Buchs/CH |
| Orange G | Fluka Chemie, Buchs/CH |
| 87% Glycerol | Riedel de Haën, Seelze/GER |
| DNA ladder TrackIt™ 1kb | Invitrogen |
| China: | <i>(unknown - manufacturers of basic chemicals were not identifiable)</i> |
| Proteinase K | unknown |
| Phenol/chloroform | unknown |
| Sodium acetate | unknown |
| Isopropanol | unknown |
| Ethanol | unknown |
| RNase | unknown |
| Zirconia/silica beads 0.1 mm | BioSpec Products, Bartlesville, OK/USA |
| Zirconia/silica beads 3 mm | BioSpec Products, Bartlesville, OK/USA |
| Phenol | unknown |
| Chlorophorm | unknown |
| Isoamyl alcohol | unknown |
| rTaq DNA polymerase; 1x Mg ²⁺ -free reaction buffer; MgCl ₂ | Takara Bio Inc., Shiga/JAP |
| dATP, dGTP, dTTP, and dCTP | Fermentas; Burlington, Ontario/CAN |
| Biotin-11-dUTP | Fermentas; Burlington, Ontario/CAN |
| DNase I; 1x reaction buffer, stop buffer with EGTA | Fermentas; Burlington, Ontario/CAN |
| Agarose | unknown |
| Ladder GeneRuler 100bp Plus | Fermentas, Burlington, Ontario/CAN |
| Lambda DNA/EcoRI+HindIII Marker 3 | Fermentas, Burlington, Ontario/Can |

| Product | Company |
|------------------------------------|------------------------|
| Glycerol | unknown |
| Dimethylbenzene xylene | unknown |
| Bromophenol blue | unknown |
| Polyacrylamide | unknown |
| Urea | unknown |
| Deionized formamide | unknown |
| 10% APS (ammonium persulfate) | unknown |
| TEMED (Tetramethylethylenediamine) | unknown |
| SybrGreen I | Amresco, Solon, OH/USA |

Kits

| Product | Company |
|--|---|
| QIAquick Spin PCR purification kit | Qiagen, Hilden/GER |
| 2100 Bioanalyzer DNA 7500 LabChip kit | Agilent Technologies, Santa Clara, CA/USA |
| 2100 Bioanalyzer DNA 1000 LabChip kit | Agilent Technologies, Santa Clara, CA/USA |
| Silverquant Detection Kit | Eppendorf, Hamburg/GER |
| BigDye Terminator cycle sequencing kit | Applied Biosystems, Darmstadt/GER |
| Stool extraction kit | Sigma-Aldrich Chemie GmbH, Munich/GER |

2.1.2 Technical equipment

DNA extraction and processing

| Product | Company |
|---|---|
| Reaction tubes, various sizes | Eppendorf, Hamburg/GER |
| Falcon tubes | Becton Dickinson, Franklin Lakes, NJ/USA |
| Vortex-Genie 2™ | Bender und Hobein AG, Zürich/CH |
| Centrifuge 5810 R (rotor: A-4-62) | Eppendorf, Hamburg/GER |
| MiniSpin® | Eppendorf, Hamburg/GER |
| ND-1000 spectrophotometer | NanoDrop Technologies, Rockland, ME/USA |
| DyNA Quant 200 fluorometer | Amersham Biosciences, Freiburg/GER |
| Mastercycler gradient | Eppendorf, Hamburg/GER |
| Electrophoresis Power Pac 300, Sub Cell GT | Bio-Rad Laboratories, Munich/GER |
| Tetrad2 thermocycler | Bio-Rad Laboratories, Munich/GER |
| Bioanalyzer 2100, lab-on-a-chip electrophoresis | Agilent Technologies, Santa Clara, CA/USA |
| Gel documentation system Las-1000 Plus | Fujifilm, Düsseldorf/GER |

Array fabrication and processing

| Product | Company |
|---|--|
| MicroGrid II 610 microarrayer | BioRobotics, Cambridge/UK |
| MicroSpot 2500 pins | BioRobotics, Cambridge/UK |
| Accelerator PT 3000 split pins (70 µm) | Point Technologies, Boulder/USA |
| Microarray scanner ScanArray Express | Perkin Elmer, Waltham, MA/USA |
| Thermomixer comfort with slide adapter | Eppendorf, Hamburg/GER |
| OV5 heating compartment | Biometra, Göttingen/GER |
| drying compartment (model U) | Memmert, Schwabach/GER |
| Gene Frame®, 65 µL; 15 x 16 mm | ABgene, Hamburg/GER |
| Gene Frame® coverslips | ABgene, Hamburg/GER |
| 384-well polystyrol (PS) microwell plates (U-shape) | Greiner Bio-One, Frickenhausen/GER |
| Nexterion® Slide E (75,6 x 25,0 x 1,0 mm) | Schott Technical Glass Solutions, Jena/GER |
| Epoxy-coated glass slides (3D) | Eppendorf, Hamburg/GER |
| Glass container 10x7 cm | unknown |
| Magnetic stirrer IKAMAG® RET-GS; IKA® | Janke und Kunkel, Staufen/GER |
| ScanArray Express array scanner | Perkin Elmer, Waltham, MA/USA |
| Silverquant scanner | Eppendorf, Hamburg/GER |
| Digital pH-meter pH525 | WTW, Weinheim/GER |

Sequencing

| Product | Company |
|---|-----------------------------------|
| MasterCycler Gradient, MasterCycler ep Gradient, MasterCycler ep Gradient S | Eppendorf, Hamburg/GER |
| ABI Prism 377 DNA sequencer | Applied Biosystems, Darmstadt/GER |

DGGE

| Product | Company |
|------------------------------|----------------------------------|
| Dcode System apparatus | Bio-Rad Laboratories, Munich/GER |
| UVI gel documentation system | UVIttec, Cambridge/UK |

Software

| Product | Company |
|--|--|
| ARB software environment | Freeware (Ludwig 2004) |
| ScanArray® Express 3.0 | Perkin Elmer, Waltham, MA/USA |
| Microsoft Excel | Microsoft, Richmond, WA/USA |
| Microsoft Word | Microsoft, Richmond, WA/USA |
| Microcal Origin 5.0 | Microcal, Milton Keynes/UK |
| Corel PhotoPaint 12.0 | Corel Corporation, Ottawa, Ontario/CAN |
| SeqMan II 5.0 | DNASstar, Madison, Wis/USA |
| OligoAnalyzer 3.0 | Online tool, Integrated DNA Technologies |
| MatLab & Simulink R2006a | MathWorks GmbH, Ismaning/GER |
| PAST (Paleontological Statistics Software) | Freeware (Hammer 2001) |

2.1.3 Biological material

Isolated DNA from 31 bacterial reference strains was obtained from the Institute for Medical Microbiology at the University of Giessen (Germany) for the verification of the microarray. Nine of them were strains from the German Collection of Microorganisms and Cell Cultures (DSMZ): *Aeromonas trota* DSM7312 (later identified as *A. minutum*), *Aeromonas hydrophila* (later identified as *Aeromonas bestiarum* DSM30019), *Atopobium minutum* DSM20586, *Bacteroides fragilis* DSM9671, *Bifidobacterium bifidum* DSM20456, *Clostridium haemolyticum* DSM5565, *Atopobium vaginae* DSM15829, *Bacteroides ureolyticus* DSM20703, *Campylobacter coli* DSM4689, *Campylobacter jejuni* DSM4688, *Campylobacter lari* DSM11375, *Clostridium difficile* DSM12056, *Lactobacillus delbrueckii* DSM20074, *Lactococcus lactis* DSM4644, *Mycobacterium avium* subsp. *paratuberculosis* DSM44133, *Plesiomonas shigelloides* DSM8224, *Roseburia intestinalis* DSM14610, *Veillonella parvula* DSM2008, *Vibrio fischeri* DSM9499, *Vibrio parahaemolyticus* DSM11058, *Yersinia pseudotuberculosis* DSM8992. All others were clinical isolates from the University Hospital of Giessen: *Enterococcus faecium* UR13873 (later identified as *E. faecalis*), *Enterococcus faecium* UR15676, *Enterococcus faecalis*, *Listeria innocua* 6b, *Listeria monocytogenes* EGD-e, *Mycobacterium tuberculosis*, *Salmonella enterica* subsp. *enterica* serovar *Typhimurium* (in the following called *S. typhimurium*), *Salmonella enterica* subsp. *enterica* serovar *Enteritidis* (in the following called *S. enteritidis*), *Yersinia enterocolitica*. All strains were identified by culture and specific real-time PCR assays. *Escherichia coli* DH5 α was cultivated in-house and DNA was isolated by a phenol/chloroform extraction protocol. An *E. coli* O157:H7 (EHEC) DNA isolate was derived from the collection of the Institute for Technical Biochemistry.

Fifty-eight clinical faecal isolates of patients with gastroenteritic symptoms and six isolates from healthy volunteers were obtained from clinical routine of the Institute for Medical Microbiology at the University of Giessen (Germany). The samples of patients were pre-characterized by culture-based methods and real-time PCR.

Sixteen faecal samples from a trial with germ-free piglets delivered via caesarean section into a SPF-level barrier system, which were inoculated orally with a whole faecal suspension from a healthy 10-year old boy (human flora-associated piglets, HFA), were obtained from the School of Life Sciences and Biotechnology at Shanghai Jiao Tong University (China) (Pang 2007). The faecal samples were taken at two different points in time after birth, eight at day 14 and another eight at day 21. Additionally, seven faecal samples from pig-flora associated (PFA) piglets of day 14 and six samples of day 21 were obtained. Both trial groups were affected by an intestinal pathogen, which caused death of one PFA piglet (P1) at day 10 and a second one (P4) at day 21 after birth. Two faecal samples from the pig donor and the human donor and four faecal samples from two conventionally raised piglets at day 14 and day 21 were obtained for analysis.

Ten faecal specimens from rotavirus-infected children (group R) and ten samples from healthy children (group H) were obtained from the School of Life Sciences and Biotechnology at Shanghai Jiao Tong University (China) for investigation. Samples of rotavirus-infected children were derived from patients of Shanghai Xinhua Hospital. Samples of healthy children were randomly collected from children of a local kindergarten. Samples of group R were taken before therapy was started and confirmed rotavirus-positive by the hospital using a commercial kit (Diarlex-Rota, Orion-Diagnostics, Finland). None of the children involved in this study had received any antibiotic medication within three months before collection of their faeces. The study was performed with their parents' consent.

2.1.4 Primers

For the amplification of the target genes, 16S and 23S rRNA genes, primer pairs were designed or derived from literature (see probeBase (Loy 2007)) and modified(*) according to own purposes, which are directed against conserved regions of both genes regarding the target species (Tab. 2.1). Positions are numbered according to *E. coli* base numbering. Primer pair 616V/985R was used to amplify both, the 16S as well as the 23S ribosomal gene in one fragment. Two additional primers, Lo180Va and Lo110Ra, located at the beginning of the 23S rRNA gene, were used to amplify the 16S and 23S genes separately. For the amplification of both genes in seven shorter fragments, eleven new primers (Tab. 2.1) were combined with the primers 616V, Lo180Va, and 985R, amplifying the 16S ribosomal gene in three fragments and the 23S rRNA gene in four fragments in nearly full length. For amplification of the genes in five fragments, three new primers (Tab. 2.1) were designed and seven ones were rejected. The new primer pairs amplified the 16S rRNA gene in two fragments and the 23S rRNA gene in three ones. In few cases, the fragment DE was amplified with the primer combination Lo180Va/1084R instead of 114V/1084R.

Three more primers were designed for sequencing only, to receive overlapping fragments: 855V, 1608V-G, and 1608V-A for the 23S rRNA gene. Primer 855V in combination with 1084R amplified the end of the DE fragment together with the F fragment. Primers 1608V-G and 1608V-A in combination with primer 985R amplified the end of fragment F together with fragment G. Primer 1608V-G was used for *E. coli*, mycobacteria, *Atopobium*, *Campylobacter*, *Salmonella*, *Yersinia*, *Plesiomonas*, *Roseburia*, *Vibrio*, and *Aeromonas*, while primer 1608V-A was used for enterococci, lactococci, *Listeria*, *Veillonella*, and clostridia.

For amplification of the V3 region of the 16S rRNA gene for DGGE analysis, a primer pair was derived from literature (Tab. 2.1). The V3 region was amplified with primers P2 and P3 containing a 40 nt-GC-clamp (Muyzer 1993).

All oligonucleotide primers, besides those for DGGE, were obtained from Metabion international AG (Martinsried/GER). The source of DGGE primers is unknown (work in Shanghai).

Tab. 2.1: Primers for amplification of the 16S and 23S ribosomal genes. (*) primers previously published in ProbeBase and modified according to own purposes.

| Gene | Fragment ^{a)} | Primer name | Sequence 5'→3' | <i>E. coli</i> position | T _M °C ^{b)} |
|--|------------------------|-----------------------------|---|----------------------------|---------------------------------|
| Amplification of both genes in one 4.5 kb fragment | | | | | |
| 16S | 4.5 kb | 616V | AGAGTTTGATYMTGGCTCAG | 8-27 | 52.7 |
| 23S | 4.5 kb | 985R | CCGGTCCTCTCGTACT | 2654-2669 | 53.6 |
| Amplification of both genes in two fragments of 2.1 and 2.5 kb | | | | | |
| 16S | 2.1 kb | 616V | AGAGTTTGATYMTGGCTCAG | 8-27 | 52.7 |
| 23S | 2.1 kb | Lo110Ra | TBCCCCATTTCRGANATC | 110-126 | 51.1 |
| 23S | 2.5 kb | Lo180Va | CYGAATGGGGVAACC | 115-129 | 51.2 |
| 23S | 2.5 kb | 985R | CCGGTCCTCTCGTACT | 2654-2669 | 53.6 |
| Amplification of both genes in seven fragments | | | | | |
| 16S | A | 616V | AGAGTTTGATYMTGGCTCAG | 8-27 | 52.7 |
| 16S | A | 515R | CGGCTGYTGGCAC | 515-527 | 53.5 |
| 16S | B | 524V | GCCGCGGTAATACG | 524-537 | 50.8 |
| 16S | B | 907R* | CCCGTCAATTYMTTGGAGTTT | 907-927 | 51.9 |
| 16S | C | 926V | GGGRCCCGCACAA | 926-938 | 54.9 |
| 16S | C | 1390R* | GGGCGGTGWGTACAA | 1390-1404 | 48.8 |
| 23S | D | Lo180Va | CYGAATGGGGVAACC | 115-129 | 51.2 |
| 23S | D | 458R | CCWTTCCYTCACRGTAC | 458-474 | 50.2 |
| 23S | E | 463V | GTGARGGAAWGGYGAAAAG | 463-481 | 52.0 |
| 23S | E | 1084R | GAGCTRRTACGCWYTCTTT | 1082-1102 | 50.6 |
| 23S | F | 1091V | GCGTAAAYAGCTCACTRGT | 1091-1109 | 52.4 |
| 23S | F | 1923R | GAATTTTCGCYACSTTAGGA | 1923-1941 | 52.0 |
| 23S | G | 1934V | CGAAATTCCTTGTTCRGKTA | 1934-1953 | 51.0 |
| 23S | G | 985R | CCGGTCCTCTCGTACT | 2654-2669 | 53.6 |
| Amplification of both genes in five fragments | | | | | |
| 16S | H | 616V | AGAGTTTGATYMTGGCTCAG | 8-27 | 52.7 |
| 16S | H | 781R | CCAGGGTATCTAATCCTGTT | 781-800 | 51.6 |
| 16S | J | 790V | AGATACCCTGGTAGTCC | 790-806 | 50.0 |
| 16S | J | 1390R | GGGCGGTGWGTACAA | 1390-1404 | 48.8 |
| 23S | DE | 114V | TCYGAATGGGGVAAC | 114-128 | 48.6 |
| 23S | DE | 1084R | GAGCTRRTACGCWYTCTTT | 1084-1102 | 50.6 |
| 23S | F | 1091V | GCGTAAAYAGCTCACTRGT | 1091-1109 | 52.4 |
| 23S | F | 1923R | GAATTTTCGCYACSTTAGGA | 1923-1941 | 52.0 |
| 23S | G | 1934V | CGAAATTCCTTGTTCRGKTA | 1934-1953 | 51.0 |
| 23S | G | 985R | CCGGTCCTCTCGTACT | 2654-2669 | 53.6 |
| Additional primers for sequencing | | | | | |
| 23S | EFü | 855V | GGVGGTAGAGCACTG | 855-869 | 51.5 |
| 23S | FGü | 1608V-G | AAACCGACACAGGTGG | 1608-1623 | 52.9 |
| 23S | FGü | 1608V-A | AAACCGACACAGGTAG | 1608-1623 | 48.9 |
| Primers for DGGE analysis | | | | | |
| 16S | V3 | P2 | ATTACCGCGGCTGCTGG | | 59.5 |
| 16S | V3 | P3 with 40nt 5'-GC-clamp | CGCCCGCCGCGCGGCGGG CGGGGCGGGGCACGGGGGC CTACGGGAGGCAGCAG | | 56.2 |

a) fragment (internal identifier), which is amplified with the primer pair

b) calculated with OligoAnalyzer (<http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/>); concentration 0.4 µM; for primers with degenerated bases the average T_M is given; for primers with GC-clamp the clamp was not included in the calculation.

2.2 Methods

2.2.1 Capture probe design

The capture probes were designed using ARB software (Ludwig 2004) and the available database *ssu_jan04.arb* (www.arb-home.de) and a LSU database. Both databases were updated with new sequences from GenBank in respect of the target species. By using the internal probe search tool, species- and genus-specific capture probes with length between 17 and 27 nt were identified. To obtain similar hybridization efficiencies the thermodynamic parameters were also taken into consideration. If possible probes were chosen according to or adjusted by length variation to pre-defined ranges of the melting temperature (50-56°C), self-dimerization efficiency ($dG > -7 \text{ kcal} \cdot \text{mol}^{-1}$), and hairpin stability ($dG > 0 \text{ kcal} \cdot \text{mol}^{-1}$) calculated by the free online tool OligoAnalyzer 3.0 (Integrated DNA Technologies) with predefined parameters and for 45°C hybridization temperature. All probes were checked for their *in-silico* specificity using the ARB probe match function with the option “weighted mismatch”. Each probe was designed as sense and antisense. The sense probe was a sequence derived from the sense strand and was complementary to the antisense strand, while the antisense probe was a sequence from the antisense strand complementary to the sense strand. All probes were named according to the following principle: [target organism]_[target gene]_[position]_[length]. The [target organism] was an abbreviation for the genus or species name. The [target gene] was abbreviated with SSU or LSU. The [position] of the probe was a relative position towards the *E. coli* gene sequence. All antisense probes were indicated with an “i” in the end. Probe names ending with “/i” comprised sense and antisense probe.

2.2.1.1 Array V1-V6 ^(V01-V03D)

During the process of probe design and verification comprising six different array versions, 86 sense and 86 antisense probes were designed, which targeted following species and genera: (pathogens) *Plesiomonas shigelloides*, *Mycobacterium* spp., *Mycobacterium avium* complex, *Mycobacterium avium* subsp. *paratuberculosis*, *Listeria* spp., *Listeria monocytogenes*, *Salmonella* spp., *Shigella flexneri*, *Shigella sonnei*, *Shigella dysenteriae*, *Aeromonas* spp., *Aeromonas hydrophila*, *Escherichia coli* O157:H7, *Vibrio* spp., *Vibrio cholerae*, *Vibrio parahaemolyticus*, *Vibrio parahaemolyticus* and relatives, *Campylobacter* spp., *Campylobacter jejuni*, *Campylobacter coli*, *Clostridium difficile*, *Yersinia* spp., *Yersinia enterocolitica*, (residents) *Escherichia coli*, *Roseburia* spp., *Roseburia intestinalis*, *Acholeplasma laidlawii*, *Atopobium* spp., *Atopobium minutum/fossor*, *Veillonella* spp., *Fusobacterium prausnitzii*, *Bacteroides* spp., *Bacteroides fragilis*, *Enterococcus* spp., *Enterococcus faecalis*, (probiotics) *Lactococcus lactis*, *Lactobacillus delbrueckii*, *Lactobacillus acidophilus*, *Bifidobacterium bifidum*. The final probe set V6f comprised 108 probes (plus 4 control probes). The oligonucleotide probes including a 5'-amino-poly(T)₁₁ modification were purchased from Metabion international AG (Martinsried/GER) in desalted purity and quality controlled by matrix-assisted laser desorption ionization-time of flight (MALDI-TOF) mass spectrometry.

2.2.1.2 Array V7 ^(V04)

The probe set used in Shanghai for array application comprised 106 probes targeting following species and genera: (pathogens) *Plesiomonas shigelloides*, *Mycobacterium* spp., *Mycobacterium avium* complex, *M. avium* subsp. *paratuberculosis*, *Listeria* spp., *Listeria monocytogenes*, *Salmonella* spp., *Aeromonas* spp., *Aeromonas hydrophila*, enterohemorrhagic *Escherichia coli* O157:H7, *Vibrio* spp., *Vibrio cholerae*, *Vibrio parahaemolyticus*, *V. parahaemolyticus* and relatives, *Campylobacter* spp., *Campylobacter jejuni*, *Clostridium difficile*, *Yersinia enterocolitica*, (residents) *Eubacterium bifforme*, *Escherichia coli*, *Roseburia* spp., *Roseburia intestinalis*, *Acholeplasma laidlawii*, *Atopobium* spp., *Atopobium minutum/fossor*, *Veillonella* spp., *Fusobacterium prausnitzii*, *Bacteroides*

spp., *Bacteroides fragilis*, *Enterococcus* spp., *Enterococcus faecalis*, (probiotics) *Lactococcus lactis*, *Lactobacillus acidophilus*, *Lactobacillus delbrueckii*, *Bifidobacterium bifidum*. The oligonucleotide probes including a 5'-amino-poly(T)₁₁ modification were purchased from Metabion international AG (Martinsried/GER) in desalted purity and quality controlled by MALDI-TOF mass spectrometry.

2.2.1.3 Controls

In addition to the species- and genus-specific capture probes, each array also included several controls: a Cyanine 3- (V6, V8) or Biotin- (V7) pre-labelled spotting control (5'-amino-(T)₁₁- TCTAGACAGCCACTCATA-Cy3/Biotin-3'), a positive hybridization control (5'-amino-(T)₁₁- GATTGGACGAGTCAGGAGC-3') complementary to a Cy3- (V6, V8) or Biotin- (V7) labelled oligonucleotide target (5'-Cy3/Biotin-GCTCCTGACTCGTCCAATC-3'), which was spiked during hybridization, and a negative hybridization control (5'-amino- (T)₁₁- TCTAGACAGCCACTCATA-3'). All these control sequences were unrelated to bacterial sequences. The universal bacterial probe EUB338 (5'-amino- (T)₁₁-GCTGCCTCCCG-TAGGAGT-3') was used as process control.

2.2.2 Oligonucleotide array fabrication

2.2.2.1 Arrays for species identification (Eppendorf slides)

The capture and control probes were laid out in four subarrays and were grouped in array V3-V6 by their affiliation to pathogens, residents, or potential probiotics. Seven different layouts were spotted with proceeding development and redesign of probes. The given layouts only represent layout V6 (Fig. 2.1), which contained all designed and modified probes, layout V6f (Fig. 2.2) comprising the final probe set, and layout V7 (Fig. 2.3), which was a reduced probe set used for faecal sample analysis in Shanghai (chap. 1.1.1 and 3.11.3). The spotting control was placed at the corner positions of each subgrid and all other controls were distributed over the array. The probes were dissolved in 1x Eppendorf spotting buffer (2x: 160 mM Na₂SO₄, 130 mM Na₂HPO₄) to a final concentration of 20 μM (pre-labelled spotting control 10 μM) in a microwell plate in a volume of 50 μL. They were spotted in triplicates with a MicroGrid II microarrayer with four MicroSpot 2500 pins (BioRobotics) on epoxy-coated glass slides (Eppendorf) at an air humidity of 55%±5%. The spot size and the spot-to-spot distance were estimated to be 120 μm and 320 μm, respectively. Covalent immobilization was achieved by incubating the oligonucleotide arrays at 60°C for 30 min in a drying compartment (Mettler). Shortly before hybridization, the slides were rinsed for 5 min in 0.1% (vol/vol) Triton X-100 in ddH₂O, 4 min in 0.5 μL of concentrated HCl per mL of ddH₂O, and 10 min in a 100 mM KCl solution with constant stirring. Subsequently, the slides were incubated in blocking solution (25% [vol/vol] ethylene glycol, 0.5 μl of concentrated HCl per mL of ddH₂O), with the spotted side facing upwards, at 50°C in a heating compartment (OV5; Biometra); rinsed in ddH₂O for 1 min; and finally, dried under a flow of nitrogen.

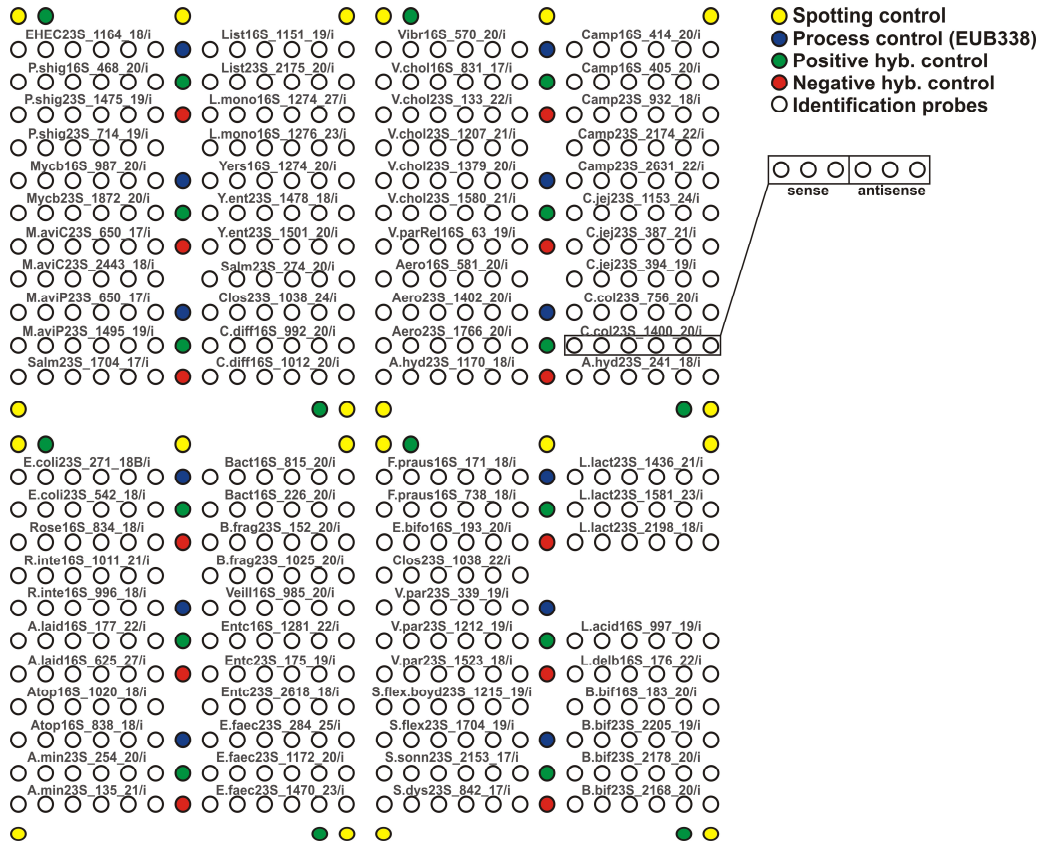


Fig. 2.1: Microarray layout V6^(V03D) of the Gastroenteritis-Chip comprising 172 probes (86 sense + 86 antisense probes).

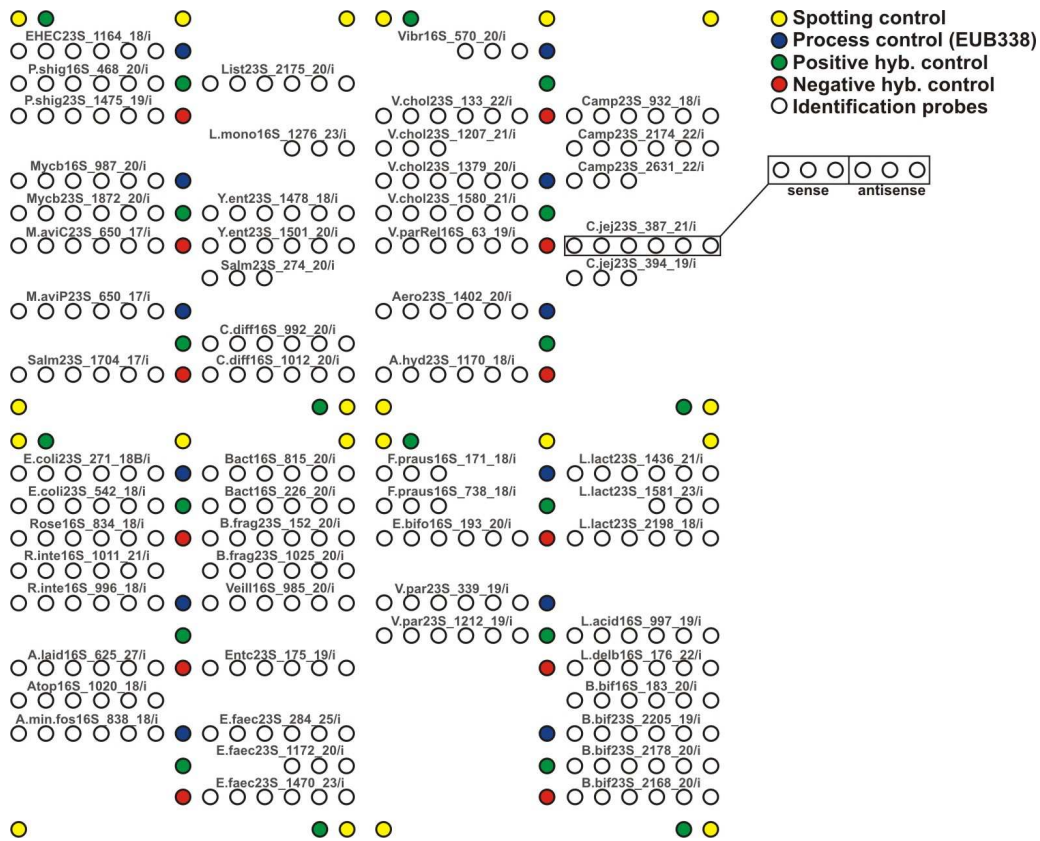


Fig. 2.2: Microarray layout V6^(final probe set) of the Gastroenteritis-Chip comprising only the 108 positively verified probes. The array was not printed in this form.

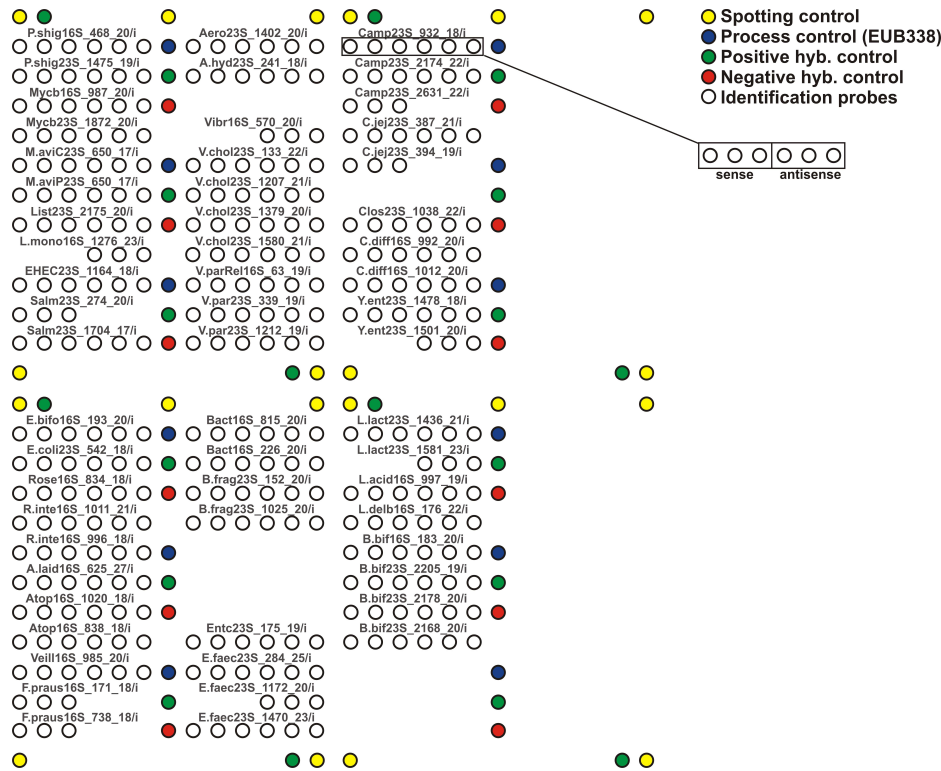


Fig. 2.3: Microarray layout V7^(V04) of the Gastroenteritis-Chip comprising 106 probes.

2.2.2.2 Arrays for quantification experiments (Nexterion slides)

The protocol described in chapter 2.2.2.1 was slightly modified. The probes were dissolved in 1x Nexterion spotting buffer (2x: 1+3 Spot I/Spot III) to a final concentration of 20 μM (pre-labelled spotting control 5 μM) in a microwell plate in a volume of 50 μL . They were spotted according to layout Fig. 2.4 on epoxy-coated glass slides (Nexterion) at an air humidity of 55% \pm 5%. Three arrays were spotted on each slide. The following procedures were the same as in 2.2.2.1.

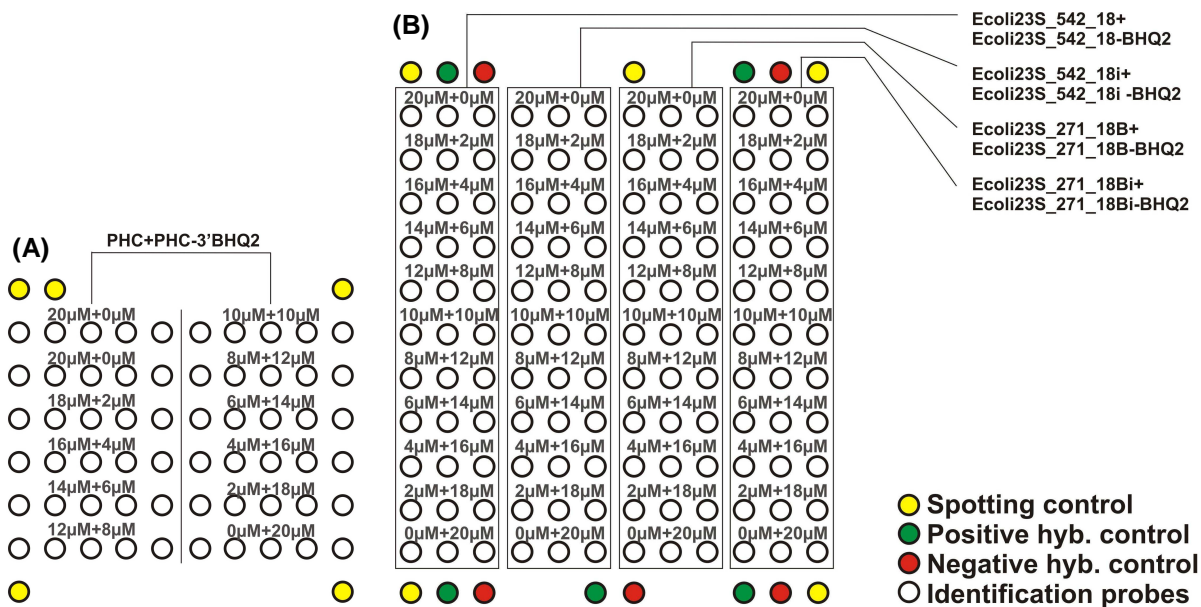


Fig. 2.4: (A) Microarray layout for preliminary quantification experiments by Hörmann (Hörmann 2006). (B) Microarray layout V8 for quantification experiments.

2.2.3 DNA extraction

2.2.3.1 Extraction with Qiagen Stool Kit

DNA from pure cultures (reference species) was isolated as described by Domann et al. (Domann 2003). DNA from faecal samples was isolated using the Sigma Stool Extraction Kit (Sigma) according to the manufactures instructions manual. For each extraction, about 2 mL of suspended faecal sample were utilized.

The DNA content was determined by measurement of the optical density using the ND-1000 spectrophotometer (NanoDrop Technologies). Exemplarily, the DNA integrity of two isolates (*C. jejuni* and *C. difficile*) was checked in a 1.5% agarose gel run in 1x TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA-Na₂-salt) for 30 min at 120 V using 500 ng DNA and 3 µL DNA ladder solution (TrackIt™ 1 kb DNA ladder, Invitrogen).

2.2.3.2 Extraction with proteinase K and phenol/chloroform

For the extraction of DNA from *E. coli* DH5α 1.5 mL of an overnight culture were centrifuged at 20,800 x g for 2 min and the pellet was resuspended in 467 µL TE buffer (10 M Tris, 1 M EDTA, pH 8.0). Thirty µL of a 10% (m/v) SDS solution and 3 µL proteinase K (20 mg/mL) were added to pulp the cells and denature and degrade the proteins. The mixture was incubated for 1 hour at 37°C. By adding the same volume of phenol/chloroform the proteins were extracted. The merged solutions were shaken until both phases were fully mixed and were then centrifuged at 20,800 x g at 4°C for 2 min. The upper liquid phase containing the DNA was transferred into a new reaction tube and mixed with the same volume of phenol/chloroform. After another centrifugation the upper liquid phase was again transferred to a new reaction tube and the previous extraction step, centrifugation and transfer of the upper phase were repeated. Then, 1/10 of the volume of 3 M sodium acetate pH 5.2 and 0.6 times the volume of isopropanol were added and gently mixed to precipitate the DNA. After centrifugation at 2,700 x g at 4°C for 10 min the supernatant was removed and the pellet was washed two times with 1 mL 70% ethanol by gentle mixing followed by centrifugation at 2,700 x g at 4°C for 10 min and removal of the supernatant. The pellet was dried for 3-4 min in a vacuum centrifuge and then resuspended in 100 µL ddH₂O. DNA quality control and concentration determination was performed as described above (2.2.3.1). The RNA was digested with 0.28 µg RNase/µg DNA for two hours at room temperature. The DNA was precipitated by adding 1/10 of the volume of 3 M sodium acetate pH 5.2 and 2.5 times the volume of 100% ethanol and incubation for two hours on ice. After centrifugation for 15 min at 4°C and 17,900 x g the pellet was washed with 70% ethanol and centrifuged again for 10 min at 4°C and 10,000 x g. The pellet was dried and resuspended in 50 µL HPLC-H₂O. The DNA concentration was measured again as described above (2.2.3.1).

2.2.3.3 Extraction with mechanical pulping and phenol/chloroform/iso-amyl-alcohol

DNA from faecal samples from piglets and from healthy and infected children (Shanghai) was extracted using the bead beater method. Two times one gram of faecal sample were weighed each into a 50 mL falcon tube and 15 mL ice cold 0.1 M sodium phosphate buffer pH 7 (57.7 mL 1 M Na₂HPO₄, 42.3 mL 1 M NaH₂PO₄ ad 1 L with H₂O) was added. In case of low sample amount only 0.5 gram were used. Five glass beads with a diameter of 3 mm were added per tube and the samples were spun thoroughly for several minutes until they were completely suspended. Again, 15 mL of 0.1 M sodium phosphate buffer were added and the samples were vortexed. Subsequently, all specimens were centrifuged for five minutes at 200 x g to remove non-bacterial material like mucus together with the beads and the supernatant was transferred to a new falcon tube. This was repeated two more times. The supernatant was then again transferred to a new falcon tube, vortexed and centrifuged for five minutes at 9,000 x g. The pellet was dissolved in 30 mL phosphate buffer and centrifuged for five minutes at 9,000 x g. The supernatant was discarded and the step was repeated. Finally, the pellet was dissolved in five millilitre of phosphate buffer (in case of

0.5 g faecal sample only 3 mL) and the samples were aliquoted and stored as 1 mL fractions at -20°C.

For the extraction of DNA the thawed faecal samples were vortexed and centrifuged for five minutes at 9,000 x g and the supernatant was discarded. The pellets were resuspended in one millilitre of ice cold buffer Z (10 mM Tris-HCl pH 8, 150 mM NaCl). If two aliquots had to be pooled due to low cell amount in the sample, only 0.5 mL of buffer was added per tube. The samples were transferred to screw tubes containing 0.3 g beads (0.1 mm zirconia/silica beads, Biospec Inc.), while sample aliquots were pooled if necessary. Then, 150 µL ice cold Tris-buffered phenol pH 7-8 were added and the samples were shaken in the bead beater for 60 seconds. Shaking was repeated three times, while in-between the samples were put on ice for one minute each time to cool down. After adding of 100 µL of 10% SDS the samples were gently mixed twice followed by five minutes cooling on ice. Then, 150 µL ice cold chloroform/iso-amyl-alcohol (24:1) were added and gently mixed. This was followed by centrifugation for ten minutes at 15,000 x g and transferring of the supernatant to new 2 mL reaction tubes. Now, 1/10 volume of ice cold 3 M sodium acetate buffer and one volume of ice cold phenol were added and gently mixed. The samples were centrifuged for ten minutes at 16,100 x g and the supernatant was transferred to new 2 mL tubes. Now, ½ volume of ice-cold phenol and ½ volume of ice-cold chloroform/iso-amyl-alcohol (24:1) were mixed with each sample. Subsequently, the samples were centrifuged for ten minutes at 16,100 x g and the supernatant was transferred to new 2 mL tubes. This time, only one volume of ice-cold chloroform/iso-amyl-alcohol (24:1) was added and gently mixed with the sample. After another centrifugation for ten minutes at 16,100 x g and transfer of the supernatant to new 2 mL tubes (maximum 600 µL/tube), the DNA was precipitated with two volumes of ice cold ethanol at -20°C for about 12 h. After thawing and centrifugation for 15 min at 15,000 x g the pellets were completely dried in a vacuum centrifuge. The pellets were then resuspended in 100 µL ddH₂O and two belonging samples were pooled. In order to remove residual RNA, samples were incubated with 150 µg RNase/200 µL DNA-extract at 37°C for 30 min. The quality of the extracted DNA was checked in a 0.8% agarose gel as described in 2.2.3.1 but using 3 µL sample and 3 µL ladder (Lambda DNA/EcoRI+HindIII Marker 3, Fermentas). The DNA amount was determined by measurement of the optical density using the DyNA Quant 200 fluorometer (Amersham Pharmacia Biotech).

2.2.4 DNA target amplification

2.2.4.1 Amplification of long fragments without labelling

For amplification of long DNA fragments >2 kb, 100 ng of template DNA (if not stated otherwise) was added to 0.4 µM (each) forward and reverse primer; 100 µM (each) dATP, dGTP, dTTP, and dCTP (Amersham Biosciences); 1 mM Mg(OAc)₂; 1x *Taq* polymerase reaction buffer containing 1.5 mM MgCl₂; and 5 U of *Taq* DNA polymerase (Eppendorf) in a total volume of 50 µL. The amplification was performed in a Mastercycler gradient (Eppendorf) with following cycle profile: 3 min of initial denaturation at 94°C; 30 cycles of DNA denaturation at 94°C for 1 min, primer annealing at 52°C (50°C for Duplex-PCR) for 1 min, and elongation at 72°C for 4 min; and a final extension step at 72°C for 10 min. A no-template negative control was included in each PCR run. All PCR products were analyzed in an agarose gel electrophoresis as described in 2.2.5.

2.2.4.2 Amplification of short fragments without labelling

If not stated otherwise, for amplification of short fragments ≤1,000 nt 100 ng of template DNA was added to 0.4 µM (each) forward and reverse primer; 100 µM (each) dATP, dGTP, dTTP, and dCTP (Amersham Biosciences); 1 mM Mg(OAc)₂; 1x *Taq* polymerase reaction buffer containing 1.5 mM MgCl₂; and 2.5 U of *Taq* DNA polymerase (Eppendorf) in a total volume of 25 µL. The amplification was performed in a Mastercycler gradient (Eppendorf) with following cycle profile: 2 min of initial denaturation at 94°C; 30 cycles of DNA denaturation at 94°C for 30 s, primer annealing at 50°C for 30 s, and elongation at 72°C for 1 min; and a final

extension step at 72°C for 4 min. A no-template negative control was included in each PCR run. All PCR products were analyzed in an agarose gel electrophoresis as described in 2.2.5.

2.2.4.3 Amplification of short fragments with Cy3-labelling

The target genes of bacterial reference DNA and clinical faecal samples were amplified and fluorescently labelled by PCR prior to hybridization. Each fragment was amplified separately. For amplification and labelling, 10 ng of template DNA was added to 0.4 µM (each) forward and reverse primer; 100 µM (each) dATP, dGTP, and dTTP; 60 µM dCTP; 40 µM Cy3-dCTP (ratio of unlabelled dCTP/labelled dCTP, 3:2; Amersham Biosciences); 1 mM Mg(OAc)₂; 1x *Taq* polymerase reaction buffer containing 1.5 mM MgCl₂; and 2.5 U of *Taq* DNA polymerase (Eppendorf) in a total volume of 25 µL, if not described otherwise. The amplification was performed in a Mastercycler gradient (Eppendorf) with following cycle profile: 2 min of initial denaturation at 94°C; 30 cycles of DNA denaturation at 94°C for 45 s, primer annealing at 50°C for 45 s, and elongation at 70°C for 1.5 min; and a final extension step at 70°C for 4 min. A no-template negative control was included in each PCR run. All PCR products were analyzed in an agarose gel electrophoresis as described in 2.2.5.

2.2.4.4 Amplification with Biotin-labelling

The target genes were amplified and Biotin-labelled by PCR prior to hybridization. Each fragment was amplified separately. For amplification and labelling, 10 ng of template DNA from children or alternatively 4 ng (DE, F, G) or 2 ng (H, J) of template DNA from piglets were added to 0.4 µM (each) forward and reverse primer; 100 µM (each) dATP, dGTP, and dCTP; 60 µM dTTP; 40 µM Biotin-11-dUTP (ratio of unlabelled dTTP/labelled dUTP, 3:2; Fermentas); 2.5 mM MgCl₂; 1x *Taq* polymerase Mg²⁺-free reaction buffer; and 2.5 U of *rTaq* DNA polymerase (Takara Bio Inc.) in a total volume of 25 µL. The amplification was performed in a Tetrad2 thermocycler (BioRad) with following cycle profile: 2 min of initial denaturation at 94°C; 30 cycles of DNA denaturation at 94°C for 45 s, primer annealing at 50°C for 45 s, and elongation at 70°C for 1.5 min; and a final extension step at 70°C for 4 min. A no-template negative control was included in each PCR run. All PCR products were analyzed in an agarose gel electrophoresis as described in 2.2.5.

2.2.5 Purification of amplified DNA and quality control

All PCR products were purified with the QIAquick Spin PCR purification kit (Qiagen) according to the manufacturer's instruction. The DNA was eluted with 30 µL HPLC-H₂O. In case of pure culture isolates and patients isolates from the University Hospital of Giessen the DNA amount and the rate of incorporation of Cy3-dCTP, expressed as number of nucleotides/number of incorporated fluorescent dyes (NT/F), was determined by measurement of the optical density using the ND-1000 spectrophotometer (NanoDrop Technologies). The size of each amplification product was checked in a 1.5% agarose gel run in 1x TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA-Na₂-salt) for 30 min at 120 V and using 1 µL sample with 3 µL loading buffer (60 mM EDTA, 0.05% (m/v) Xylene cyanol, 0.2% (m/v) Orange G, 60% glycerol), and 3 µL DNA ladder solution (0.1 µg/µL TrackIt™ 1 kb DNA ladder in 10 mM Tris-HCl pH 8.5; 1:1 with loading buffer). In case of faecal isolates from piglets and from children volunteers in Shanghai the DNA amount was determined by measurement of the optical density using the DyNA Quant 200 fluorometer (Amersham Biosciences). The size of each amplification product was checked in a 1.2% agarose gel run in 1x TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA-Na₂-salt) for 30 min at 120 V using 1 µL sample and 3 µL ladder (GeneRuler 100bp Plus, Fermentas). Multiplex-PCR products were additionally analyzed with Bioanalyzer 2100 capillary gel electrophoresis (Agilent Technologies) and the DNA 7500 kit according to the manufacturer's protocol.

2.2.6 DNase I digestion of amplified DNA

In case of reference DNA from pure cultures and faecal DNA from gastroenteritis patients, the PCR products of the five fragments for each reference sample or faecal isolate were mixed in equimolar amounts. For a second hybridization of faecal DNA from gastroenteritis patients and in case of faecal samples from piglets and children, the whole PCR product of each fragment was mixed with the others. If not described otherwise, the fragment mix was digested with 0.4 mU DNase I (Promega) per ng DNA for 5 min at room temperature in 1x DNase I reaction buffer and a variable end volume between 30 and 65 μ L. The reaction was stopped by adding an EGTA (ethylene glycol tetraacetic acid) containing stop buffer to 1x end concentration and incubation at 65°C for 10 min. In case of faecal samples from piglets or children the PCR products were digested with 0.04 mU DNase I (Fermentas) per ng DNA due to higher specific activity of the enzyme. The degree of digestion was determined with capillary gel electrophoresis using the Agilent Bioanalyzer 2100 and a DNA 1000 kit (Agilent Technologies) according to the manufacturer's instructions.

2.2.7 Hybridization

Two hundred ng of digested PCR product from reference species or 500 ng of amplified, digested DNA from faecal isolates were hybridized to the array within a gene frame (15 by 16 mm) that was closed with a coverslip (Abgene House). Hybridization took place in a volume of 70 μ L containing 6x SSPE (1x SSPE is 0.18 M NaCl, 10 mM NaH₂PO₄, and 1 mM EDTA [pH 7.7]) and 0.7 nM of a labelled oligonucleotide target complementary to the positive hybridization control (Cy3-Comp or Biotin-Comp). For nine faecal isolates a second hybridization using the whole PCR product for two arrays (2-3 μ g) was performed. The digested PCR product of faecal DNA from piglets and children was completely hybridized to two arrays, as well. After one hour of incubation at 48°C and 1,400 rpm in the Eppendorf Thermomixer with slide adapter (Eppendorf), the slides were washed with 2x SSC (20x SSC: 3 M NaCl plus 0.3 M sodium citrate, pH7 with 1M HCl) containing 0.1% sodium dodecyl sulfate, then 2x SSC, and finally, 0.2x SSC for 10 min, respectively. The washing procedure was performed at room temperature with agitation in a glass container. The slides were subsequently dried with N₂.

2.2.8 Silver staining

The Biotin-labelled PCR products of the faecal samples from children and piglets were detected using the Silverquant Detection Kit (Eppendorf) according to the manufactures instructions. In brief, the slides were washed several times in Tween-20-containing buffer and were incubated for 10 min in a blocking buffer. This was followed by 45 min binding of a gold-labelled antibody to the DNA-incorporated Biotin and subsequent washing steps. Then, a silver salt containing solution was incubated with the slides resulting in silver deposition at the gold nanoparticles. After final washing steps with water, the slides were dried with N₂.

2.2.9 Data acquisition and processing

The fluorescence hybridization signals of the oligonucleotide arrays were read out by using the ScanArray Express Array Scanner (Perkin Elmer) with a resolution of 10 μ m/pixel and an excitation wavelength of 543 nm. The laser power and gain were adjusted to 90%. Photomultiplier Tubes (PMT) enhancement was varied between 50 and 70. The silver hybridization signals were read out by the Silverquant Scanner (Eppendorf) with a resolution of 6 μ m/pixel.

Image processing and calculation of signal intensities were performed with ScanArray Express software (version 3.0) using the adaptive circle-mode for slides with Cy3-labelling and the fixed circle-mode for slides with silver staining. The black-on-white images derived from the Silverquant Scanner were previously transformed from .eat-files into .tif-files and inverted using Corel PhotoPaint 12.0 (Corel Corporation). For calculation of the individual net

signal intensities, the local background was subtracted from the raw spot intensity value. Further data processing was performed by using Excel software (Microsoft). Average values were calculated for each triplicate if one array and for six spots if two arrays were hybridized with the same sample. Standard deviations were calculated from three or six spots, respectively. Only net signal intensities, which were significantly higher than the background, were considered. Therefore, the average signal of the negative control plus three times its standard deviation (NHC+3SD) was applied as a cut off. By definition, all hybridization signals less than NHC+3SD were set equal to zero. Species or genera were considered positive, if more than 50% of probes displayed a signal above zero with a standard deviation below 50%. The standard deviation was calculated for the net signals, in order not to distort the result in case of high values for NHC+3SD. Additionally, signals from species-specific probes were only regarded as positive signals, if the appropriate genus-specific probes showed a positive hybridization signal. Only in case of *C. difficile* the species was not supposed to hybridize with the genera-specific probe, due to sequence differences. The process control EUB338 was only used for normalization of the net signal intensities during verification with reference strains. The background corrected signal of each capture probe was expressed as percentage of the mean signal of the twelve EUB338 probes (set as 100%), resulting in the normalized net signal intensities referred to as the "relative signal intensity" [RI] of one probe. Average values of the relative intensity were again calculated for each probe. Standard deviations were calculated from three or six spots, respectively. The degree of cross-hybridization was expressed as ratio between mismatch and perfect match (MM/PM) of the relative signal intensities.

2.2.10 Limit of detection (LOD) determination

The limit of detection (LOD) of the microarray was determined by two independent hybridization experiments using *E. coli* reference DNA from a pure culture and in a second approach *Campylobacter jejuni* DNA from pure culture spiked into a *Campylobacter*-negative stool DNA extract (internal number S6). In the first experiment, *E. coli* DNA was used as a template for PCR in a dilution series of 0, 1, 10, 10², 10³, 10⁴, and 10⁵ genome equivalents. In case of *E. coli*, only the 114V/1084R fragment of the LSU was amplified, because *E. coli* probes are directed against this fragment. Additionally, four different DNase I amounts, 0.004, 0.04, 0.4, and 4 mU/ng DNA, were used for digestion of the DNA before hybridization to analyze the influence of digestion on the signal intensities and the detection limit. Each PCR product was digested and hybridized according to protocol 2.2.6-2.2.9.

To investigate the detection limit of the array for bacteria in the background of foreign DNA, *Campylobacter jejuni* DNA was spiked into 5 ng *Campylobacter*-negative stool DNA extract in the same gradations and applied to PCR amplifying the 616V/781R fragment of the SSU. The foreign DNA exceeded *C. jejuni* by 27.7 fold (10⁵ *C. jejuni* genome equivalents) to 2.7*10⁶ fold (1 *C. jejuni* genome equivalent). All PCR products were digested with 4 mU DNase I/ng DNA and fully hybridized according to the hybridization protocol. The probes targeting the 16S rRNA gene of *Campylobacter*, Camp16S_414_20/-i (5'-CACTTTTCGGAG-CGTAAACT-3'/ 5'-AGTTTACGCTCCGAAAAGTG-3') and Camp16S_405_20/-i (5'-GGAGG-ATGACACTTTTCGGA-3'/ 5'-TCCGAAAAGTGTCATCCTCC-3'), were excluded from the array in a later phase due to unspecific binding of *B. ureolyticus*.

2.2.11 Sanger DNA sequencing

For standard Sanger DNA sequencing the 16S gene was amplified in three overlapping fragments using primer combinations 616V/781R, 524V/907R and 790V/1390R. For the 23S gene following primer pairs were applied: 114V/1084R, 855V/1923R, 1091V/1923R, and 1934V/985R for all reference species, 1608V-G/985R for *E. coli*, *B. bifidum*, *R. intestinalis*, *B. ureolyticus*, *Mycobacterium*, *Atopobium*, *Campylobacter*, *Salmonella*, *Yersinia*, *Plesiomonas*, *Vibrio*, *Aeromonas*, and 1608V-A/985R for *Enterococcus*, *L. lactis*, *L. delbrueckii*, *Listeria*, *Clostridium*, and *V. parvula*. Ten ng of template were amplified in a 25

μL reaction volume containing 0.4 μM of (each) forward primer and reverse primer, 100 μM of each dNTP, 1 mM $\text{Mg}(\text{OAc})_2$, 1x *Taq* polymerase reaction buffer containing 1.5 mM $\text{Mg}(\text{OAc})_2$, 2.5 U *Taq* polymerase (Eppendorf) and HPLC- H_2O . The amplification was performed in a Mastercycler gradient (Eppendorf). An initial denaturation step (94°C for 2 min) was followed by 30 cycles consisting of DNA denaturation at 94°C for 45 s, primer annealing at 50°C for 45 s, and elongation at 72°C for 1.5 min, and a final extension step at 72°C for 4 min. A no-template negative control was included in each PCR run. The PCR product was purified with a QIAquick Spin PCR purification kit (Qiagen), according to the protocol of the manufacturer. The DNA was eluted in 30 μL of dd H_2O . The purified PCR products were sequenced in both directions with the same primers as above by GATC Biotech AG (Konstanz/GER). Some samples were sequenced in-house using the BigDye Terminator cycle sequencing kit (Applied Biosystems) and the ABI Prism 377 DNA sequencer (Applied Biosystems). The sequence data were assembled by using SeqMan II software (version 5.00; DNASTar) and have been submitted to GenBank.

2.2.12 Denaturing gradient gel electrophoresis (DGGE)

2.2.12.1 Amplification of specific fragments of the 16S gene

The denaturing gradient gel electrophoresis (DGGE) was performed with the amplified, variable region V3 of the 16S rRNA gene. Twenty ng of template DNA were added to 0.25 μM (each) forward (P2 5'-ATTACCGCGGCTGCTGG-3') and reverse primer containing a 40 bp-GC-clamp (P3 5'-CGCCCGCCGCGCGCGGGCGGGGCGGGGGCACGGGGGGCC-TACGGGAGGCAGCAG-3') (Muyzer 1993); 200 μM (each) dATP, dGTP, dTTP, and dCTP; 2 mM MgCl_2 ; 1x *Taq* polymerase reaction buffer containing Mg^{2+} -free; and 0.75 U of *rTaq* DNA polymerase (Takara) in a total volume of 25 μL . The amplification was performed in a Tetrad2 Thermocycler (Biorad) with following touch down cycle profile: 3 min of initial denaturation at 94°C; 30 cycles of DNA denaturation at 94°C for 1 min, primer annealing between 65 and 55°C for 1 min with a 1°C temperature decrease every second cycle, and elongation at 72°C for 1 min; and a final extension step at 72°C for 6 min according to Muyzer (Muyzer 1993). A no-template negative control was included in each PCR run. To minimize heteroduplex-formation and single-stranded DNA contamination from each PCR product four separate re-amplifications were performed using 5 μL PCR product as template in 50 μL total volume containing the same concentrations of reactants as before. The PCR had following cycle profile: 3 min of initial denaturation at 95°C; 5 cycles of DNA denaturation at 94°C for 1 min, primer annealing between at 55°C for 1 min, and elongation at 72°C for 1 min; and a final extension step at 72°C for 10 min. After each PCR the products were checked by electrophoresis in a 1% (wt/vol) agarose gel. The four re-amplified PCR products of each sample were pooled and concentrated to about 80 μL in a vacuum centrifuge. The DNA concentration was determined in a DyNA Quant 200 fluorometer (Amersham Pharmacia Biotech) and the samples were stored at -20°C before DGGE analysis.

2.2.12.2 DGGE analysis

For DGGE analysis 350 ng or 500 ng of PCR product of HFA-, PFA-, CV-piglets, or donor piglets, respectively, were transferred to a new tube and completely dried in a vacuum centrifuge. Then, 7 μL dd H_2O and 7 μL 2x DGGE loading buffer (70% glycerol, 0.05% dimethylbenzene xylene, 0.05% bromophenol blue in dd H_2O) were added and the DNA was resuspended. The DGGE was performed in an 8% (wt/vol) polyacrylamide gel containing a linear denaturant gradient from 27% to 52% (100% denaturant corresponds to 7 M urea and 40% deionized formamide) which was polymerised using 30 μL APS and 15 μL TEMED per 12.5 mL of low (27%) and high (52%) denaturant polyacrylamide gel solution, respectively. The gel was topped by an 8% (wt/vol) non-denaturing gel (4 mL 8% polyacrylamide gel + 40 μL APS + 8 μL TEMED). The gel was run in 1x TAE (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA- Na_2 -salt) buffer at a constant voltage of 200 V and a temperature of 60°C for 240 min in the Dcode System apparatus (Bio-Rad, Munich/GER). The gel was stained for

each 15 min with three times 10 mL 1xTAE-buffer containing 1 μ L SybrGreen I (Amresco) and photographed with UVI gel documentation system (UVItec).

2.2.13 *Multivariate Analysis*

Microarray data were mathematically analysed by different methods of multivariate analysis using the MATLAB R2006a software environment (MathWorks GmbH).

Principle component analysis (PCA), which was developed by Karl Pearson (Pearson 1901), is the simplest of the true eigenvector-based multivariate analysis and is applied to reduce a high dimensional set of attributes to few attributes that reflect the maximum variance. PCA, also named Karhunen-Loève transform, creates attributes that are more significant as a linear combination of the initial attributes. The data are transformed to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (first principle component), the second greatest variance on the second coordinate (second principle component), and so on. PCA does not take group classification of samples into account. Therefore, the obtained variables (PC1, PC2...) may be weakly related to a classification of the samples based on biological information.

PCA analysis was applied to the microarray data of faecal isolates. The results of the microarray experiments were transformed to a 0/1 matrix, where 0 stood for 'not detected' bacterium and 1 for 'detected' bacterium, including only the data for resident and probiotic bacteria. This two-dimensional matrix was subjected to PCA in MATLAB®. The principle components were calculated and, subsequently, sorted according to their correlation with the biological groups or according to their ability to predict those groups. These groups were defined as follows:

- (I) Hospitalized adults ($y=0$) and healthy adults ($y=1$);
- (II) Rotavirus-infected children ($y=0$) and healthy children ($y=1$)
- (III) Conventionally raised piglets, pig donor and PFA piglets ($y=0$), and human donor and HFA piglets ($y=1$)

The calculated data were plotted in a 2D-graph showing the principle components on the axes, which were most suitable to separate the defined groups.

Partial least squares (PLS) regression, a statistical method developed by Herman and Svante Wold (Wold 2001), was used to predict the real cases of the causal model that was calculated by PCA. Prior knowledge about the different biological groups is applied to the model. PLS decomposes the matrix by maximizing the group separation. By projecting both the predicted variables Y (the defined groups) and the observable variables X (the detected species/genera) to a new space, PLS finds a linear regression model to elucidate the fundamental relation between them. PLS regression also models the structure of X and Y , which gives richer results than traditional multiple regression. First, the samples were classified as described above for PCA including two additional classifications:

- (IV) Hospitalized adults with confirmed pathogen ($y=0$), hospitalized adults with unknown pathogen ($y=1$), and healthy adults ($y=2$);
- (V) Conventionally raised piglets ($y=0$), pig donor and PFA piglets ($y=1$), and human donor and HFA piglets ($y=2$).

The detected species/genera (X) were then related to the corresponding classification index (Y) by a PLS model. The data were centred before performing the PLS analysis. The PLS components were projected to a 2D-plot.

Leave-one-out cross-validation (LOOCV) (Wold 1978) was subsequently used to estimate the prediction ability of the PLS model with different numbers of PLS components. Each time, a sample is left out and is predicted by the sub-model built on the remaining samples. A model that gives higher prediction accuracy and lower component number ('model size') is considered optimal. The model accuracy, which is the mean of the prediction accuracy of the

sub-models, was expressed as the cross-validated correct classification rate, i.e. the percentage of correctly classified samples from the total number of samples.

The variables that mostly contributed to group discrimination were identified by variable importance analysis according to the method proposed by Westad and Martens (Westad 2000) at a significance of $p=0.01$ and $p=0.05$. With the optimal 'model size' giving best prediction accuracy, different sub-PLS models were constructed by leaving out one subject each time. Based on the obtained regression coefficients of each sub-model, the uncertainty of each variable (species/genus) was then checked.

One-way analysis of variance (one-way ANOVA) was used as an alternative method to PLS regression to verify the significance of the different distribution of the identified variables within the groups. One-way ANOVA is a statistical method that compares means of two or more groups to test whether the samples in these groups were drawn from the same population or not. A ratio of the variance calculated among the group means to the variance within the samples is produced. If all samples were drawn from the same population, i.e. the groups do not differ significantly, the variance between the group means should be lower than the variance of the samples. If the difference between both groups is significant, the variance between the groups should be higher than within the groups. To perform one-way ANOVA, the 0/1 matrix of the microarray results, as described for PCA, regarding the identified important variables was transferred to the software PAST (Paleontological Statistics Software) (Hammer 2001). The microarray data of each group, e.g. healthy vs. infected children, regarding one variable (species/genus) were compared and a p -value was calculated as a measure of significant difference.

2.2.14 DNA target quantification on DNA microarrays

2.2.14.1 Hybridization of *E. coli* PCR product to quencher-coupled probes

A microarray containing a series of spots with a constant amount of 20 μM of a probe mixture containing the *E. coli*-specific probes E.coli23S_542_18 and its 3' black hole quencher-labelled counterpart (3'BHQ2- E.coli23S_542_18) in a variable ratio of 20 μM + 0 μM , 18 μM + 2 μM , 16 μM + 4 μM up to 0 μM + 20 μM was spotted according to the modified protocol (chap. 2.2.2.2) on epoxy-coated glass slides (Nexterion, layout V8, Fig. 2.4 B). A second, third and fourth series was spotted with a mixture of E.coli23S_542_18i, E.coli23S_271_18B, E.coli23S_271_18Bi and its quencher-labelled counterparts. Three arrays were spotted on each slide. Spotting control, positive and negative hybridization controls were included in the array layout. Immobilization, blocking, and washing were carried out as described before (chap. 2.2.2.1).

Different amounts of *E. coli*-DNA corresponding to 2×10^6 , 10^6 , 10^5 , 10^4 , 10^3 , 10^2 , 10, 1, and 0 genome equivalents were applied to a PCR amplifying the DE-Fragment with primers 114V and 1084R as described above (chap. 2.2.4.3). One *E. coli* genome equivalent accounts for 5.09 fg DNA. Each amount of DNA was amplified and Cy3-labelled in six separate reactions in a volume of 25 μL . The PCR products were analyzed in a 1% agarose gel and purified with the Qiagen PCR purification kit (Qiagen). Each PCR product was eluted in 30 μL ddH₂O and analyzed for its DNA content and Cy3 incorporation (chap. 2.2.5).

Nine slides were hybridized with undigested *E. coli* PCR product derived from 2×10^6 , 10^6 , 10^5 , 10^4 , 10^3 , 10^2 , 10, 1, and 0 genome equivalents. The three arrays on each slide were hybridized in separated frames with three separately amplified target DNA samples (chap. 2.2.7). The whole, purified PCR product (~26-28 μL) was used for hybridization of each array. Another nine slides were hybridized with DNase I digested *E. coli* PCR product derived from 2×10^6 , 10^6 , 10^5 , 10^4 , 10^3 , 10^2 , 10, 1, and 0 genome equivalents as described before. For digestion 100 mU DNase I in 40 μL end volume were used according to protocol 2.2.6. Data acquisition followed the previously described protocol 2.2.9. All slides were scanned at PMT 33, 40, 45, 50, 55, 60, 65, and 70.

3 Results

3.1 Design of the identification microarray

A microarray for the detection of intestinal, bacterial pathogens regarding diarrheal disease and predominant intestinal residents including lactic acid bacteria (LAB) was developed. The target organisms were chosen according to clinical and literature information (Janka 2001) about their relevance in gastrointestinal disease and the healthy gut. The primary selection of target organisms is listed in Tab. 3.1. The bacteria were grouped by their affiliation to resident intestinal bacteria, potentially probiotic bacteria, and intestinal pathogens. Probes were designed (chap. 2.2.1) targeting the ribosomal RNA genes, which are universal to all bacteria and provide conserved regions for primer annealing as well as variable regions, which allow specific identification. Specificity is defined in this chapter as *in-silico* specificity. The small subunit (SSU) RNA gene alone was not sufficient to differentiate between all target species as *in-silico* investigations indicated (ARB). Therefore, the large subunit (LSU) RNA gene was also taken into consideration. As far as possible, probes were designed following a multiple probe concept. It was aimed to detect each species with at least two probes and, if possible, on the species and genus level. Due to a limitation in available sequences (Tab. 3.1) and specific regions, this could not be achieved in all cases. The probe sets of the different array versions (chap. 2.2.1) are depicted in Tab. 3.2. Probe design was in some cases limited due to a lack of available sequences. Full-length sequences of the 16S ribosomal gene were not available for *C. coli*. Likewise, for *Acholeplasma* spp., *Anaeroplasma* spp., *C. difficile*, *F. prausnitzii*, *Prevotella* spp., *Roseburia* spp., and *Veillonella* spp. no 23S ribosomal gene was sequenced at that time.

Despite availability of sequences, no genus-specific 16S probes could be designed for *Acholeplasma*, *Clostridium*, *Eubacterium*, *Lactobacillus*, and *Salmonella* due to low homology between members of the genus. This also applied for *Bifidobacterium*, *Fusobacterium*, *Lactococcus*, *Prevotella*, and *Shigella* due to non-specificity towards other genera. For *Plesiomonas* no genus probes were designed, because only *P. shigelloides* sequences were available. For clostridia, also sub-clustering of the sequences into OTUs (operational taxonomic units) did not result in probes with clearly defined specificity. In case of *Clostridium* and *Bifidobacterium*, 23S probes were selected instead, but on a much smaller data basis of only 16 and 2 sequences, respectively. The *Clostridium* genus-probe did not cover *C. difficile*. For *Salmonella*, two 23S probes were selected based on 34 sequences covering all *Salmonella* except for *S. bongori*.

Otherwise, dominant members of the group were chosen and species-specific probes were developed. Species-specific 16S probes could not be designed for *A. hydrophila*, *C. jejuni*, *E. faecalis*, *E. coli*, *L. lactis*, *M. avium* complex and *paratuberculosis*, *S. enterica*, several *Shigella* species, *V. parahaemolyticus*, and *Y. enterocolitica* due to low discriminative power towards other species of the same or different genus. *Escherichia coli* and *Shigella* spp., or the *M. avium* subspecies, for example, do not distinguish in the 16S gene. In all cases, except for *Salmonella enterica*, also 23S species-specific probes were selected, but in general from a smaller sequence basis.

The coverage and the predicted specificity of the designed probes were variable between the targets. In most cases, the coverage of target sequences was 100% of the available full-length sequences in the database, but for some probes, it was lower between 75% and 98% (Tab. 3.2). In some cases, the specificity of a probe could not be predicted, if only one member of a genus was already sequenced. For *Y. enterica*, *L. lactis*, and *B. fragilis*, for example, no sequences of other species of the genus were in the 23S database and *P. shigelloides* was the only representative of its genus in both, 16S and 23S database. Therefore, the specificity of these probes towards their relatives remained unclear. The probes might have been species- or genus-specific.

Tab. 3.1: Target organisms, which were selected for detection by the microarray according to clinical information and literature data about their relevance in gastrointestinal disease and the healthy gut, availability of their sequence data for probe design, and the operon copy number.

| | 16S operon copy number | 16S full length sequences | Designed 16S probes | 23S operon copy number | 23S full length sequences | Designed 23S probes | |
|--|---|------------------------------|------------------------|---------------------------|------------------------------|------------------------|----------|
| Species/genus | | | | | | | |
| residents | Acholeplasma ssp. | 2 | 9 | 0 | 2 | 0 | |
| | <i>Acholeplasma laidlawii</i> | 2 | 2 | 2 | 2 | 0 | |
| | Anaeroplasmia ssp. | <i>no data</i> | 0 | 0 | <i>no data</i> | 0 | 0 |
| | Atopobium ssp. | 1 | 4 | 2 | 1 | 1 | 2 |
| | Bacteroides ssp. | 5-7 | 38 | 2 | 4-7 | 1 | 0 |
| | <i>B. fragiles</i> | 6 | x | 0 | 6 | 1 | 2 |
| | Clostridium ssp. | 4-15 | 323 | 0 | 4-15 | 16 | 2 |
| | Enterococcus ssp. | 4 | 117 | 1 | 4 | 66 | 2 |
| | <i>Enterococcus faecalis</i> | 4 | 21 | 0 | 4 | 4 | 3 |
| | Escherichia ssp. | 7 | 64 | 0 | 7 | 51 | 0 |
| | <i>Escherichia coli</i> K12+EHEC | 7 | 61 | 0 | 7 | 51 | 1 |
| | <i>Escherichia coli</i> K12 | 7 | x | 0 | 7 | 37 | 1 |
| | Eubacterium ssp. | <i>no data</i> | 78 | 0 | <i>no data</i> | 8 | 0 |
| | <i>Eubacterium bifforme</i> | <i>no data</i> | 1 | 1 | <i>no data</i> | x | 0 |
| | Fusobacterium ssp. | 5 | 27 | 0 | 5 | 1 | 0 |
| | <i>Fusobacterium prausnitzii</i> | <i>no data</i> | 5 | 2 | <i>no data</i> | 0 | 0 |
| | Prevotella ssp. | <i>no data</i> | 86 | 0 | <i>no data</i> | 0 | 0 |
| | Roseburia ssp. | <i>no data</i> | 5 | 1 | <i>no data</i> | 0 | 0 |
| | <i>Roseburia intestinalis</i> | <i>no data</i> | 2 | 2 | <i>no data</i> | 0 | 0 |
| | Veillonella ssp. | <i>no data</i> | 8 | 1 | <i>no data</i> | 0 | 0 |
| | probiotics | Bifidobacterium ssp. | 2-5 | 54 | 0 | 2-5 | 2 |
| <i>Bifidobacterium bifidum</i> | | <i>no data</i> | 5 | 1 | <i>no data</i> | 1 | 2 |
| Lactobacillus ssp. | | 4-9 | 305 | 0 | 4-9 | 24 | 0 |
| <i>Lactobacillus delbrueckii</i> | | 9 | 8 | 1 | 9 | 1 | 0 |
| <i>Lactobacillus acidophilus</i> | | 4 | 4 | 1 | 4 | 1 | 0 |
| Lactococcus ssp. | | 6 | 43 | 0 | 6 | 4 | 0 |
| <i>Lactococcus lactis</i> | | 6 | 39 | 0 | 6 | 4 | 3 |
| pathogens | Aeromonas ssp. | 9-10 | 91 | 1 | 9-10 | 26 | 2 |
| | <i>Aeromonas hydrophila</i> | 10 | 25 | 0 | 10 | 8 | 2 |
| | Campylobacter ssp. | 2-3 | 35 | 2 | 3 | 13 | 3 |
| | <i>Campylobacter coli</i> | 3 | 0 | 0 | 3 | 3 | 3 |
| | <i>Campylobacter jejuni</i> | 3 | 6 | 0 | 3 | 4 | 3 |
| | <i>Clostridium difficile</i> | 11 | 4 | 2 | 11 | 0 | 0 |
| | <i>Escherichia coli</i> O157:H7 | 7 | x | 0 | 7 | 14 | 1 |
| | Listeria ssp. | 6 | 22 | 1 | 6 | 22 | 1 |
| | <i>Listeria monocytogenes</i> | 6 | 9 | 2 | 6 | 8 | 0 |
| | Mycobacterium ssp. | 1-2 | 224 | 1 | 1-2 | 26 | 1 |
| | <i>M. avium</i> complex | <i>no data</i> | 3 | 0 | <i>no data</i> | 2 | 2 |
| | <i>M. avium</i> subsp. paratuberculosis | 1 | 2 | 0 | 1 | 2 | 2 |
| | <i>Mycobacterium avium</i> | <i>no data</i> | 5 | 0 | <i>no data</i> | x | 0 |
| | Plesiomonas ssp. | <i>no data</i> | 4 | 0 | <i>no data</i> | 3 | 0 |
| | <i>Plesiomonas shigelloides</i> | <i>no data</i> | 4 | 1 | <i>no data</i> | 3 | 2 |
| | Salmonella ssp. | 7 | 51 | 0 | 7 | 34 | 2 |
| | <i>Salmonella enteritidis</i> | <i>no data</i> | 16 | 0 | <i>no data</i> | 16 | 0 |
| | Shigella ssp. | 7 | 21 | 0 | 7 | 5 | 0 |
| | <i>Shigella flexneri</i> | 7 | x | 0 | 7 | 2 | 1 |
| | <i>Shig. flexneri</i> + <i>Shig. boydii</i> | 7 | x | 0 | 7 | 3 | 1 |
| | <i>Shigella sonnei</i> | 7 | x | 0 | 7 | 1 | 1 |
| | <i>Shigella dysenteriae</i> | 7 | x | 0 | 7 | 1 | 1 |
| | Vibrio ssp. | 8-13 | 246 | 1 | 7-12 | 30 | 0 |
| | <i>Vibrio cholerae</i> | 8 | 16 | 1 | 8-10 | 12 | 4 |
| | <i>Vibrio parahaemolyticus</i> | 11 | 22 | 0 | 11 | 1 | 3 |
| <i>Vibrio parahaemolyticus et rel.</i> | | 100 | 1 | | x | 0 | |
| Yersinia ssp. | 6-7 | 66 | 2 | 6-7 | 3 | 0 | |
| <i>Yersinia enterocolitica</i> | 7 | 6 | 0 | 7 | 3 | 2 | |

(x) not checked

Tab. 3.2: Identification oligonucleotide probes on the microarray with thermodynamical parameters.

| target | array ^{a)} | in-silico coverage | probe name ^{b)} | PM _{wm} (0) of (n) ^{c)} | PM _{wm} (1) | PM _{wm} (2) | probe 5'-->3' (sense) ^{d)} | final array ^{e)} s i | GC (%) | Tm (ARB) ^{e)} | Tm (OIAn) ^{e)} | dG Hairpin ^{f)} (kcal*mol-1) | dG Selfdimer ^{g)} (kcal*mol-1) |
|----------------------------------|---------------------|--------------------|--------------------------|---|----------------------|----------------------|-------------------------------------|----------------------------------|--------|---------------------------|----------------------------|--|--|
| <i>Acholeplasma laidlawii</i> | V1-V6 | 100% | A.laid16S_177_22 | 2 (2) | 3 | 3 | GGATAGGATGTGTGCATGAAAA | | 40,9 | n.d. | 53,2 | 0,3 | -7,05 |
| | V1-V7 | 100% | A.laid16S_625_27 | 2 (2) | 2 | 9 | CTGTGAGGCTATGAAAACATAATAACT | x x | 33,3 | n.d. | 52,6 | 1,69 | -5,83 |
| <i>Aeromonas hydrophila</i> | V1-V6 | 75% | A.hyd23S_1170_18 | 6 (8) | 11 | 30 | CTAAGCGTGAGTGGTAGG | x x | 55,6 | 56,0 | 52,0 | 0,95 | -3,61 |
| | V5-V7 | 11% | A.hyd23S_241_18 | 1 (9) | 1 | 4 | AGTAGCGGCAGAGCGAAC | | 61,1 | n.d. | 57,7 | 0,47 | -3,61 |
| <i>Aeromonas ssp.</i> | V1-V6 | 98% | Aero16S_581_20 | 167 | 180 | 275 | CAGCGGGTTGGATAAGTTAG | | 50,0 | 60,0 | 53,0 | 2,34 | -3,61 |
| | V1-V7 | 100% | Aero23S_1402_20 | 26 (26) | 26 | 26 | ACGACTTGTAATTGCGATGG | x x | 45,0 | 58,0 | 53,2 | 0,52 | -5,36 |
| | V1-V6 | 100% | Aero23S_1766_20 | 26 (26) | 35 | 110 | GCTGGGACTGTTTATCAAAA | | 40,0 | 56,0 | 50,6 | 1,53 | -3,89 |
| <i>Atopobium ssp.</i> | V1-V7 | 100% | A.min.fos16S_838_18 | 5 (5) | 5 | 10 | GATGTACCTCCGTGCGC | x x | 61,1 | 58,0 | 55,8 | 1,02 | -3,65 |
| | V3-V7 | 100% | Atop16S_1020_18 | 25 (17) | 47 | 99 | GTGGCCGAAAGGAGCCTA | x x | 61,1 | 58,0 | 57,3 | -0,87 | -9,28 |
| | V2-V6 | 100% | A.min23S_254_20 | 1 (1) | 1 | 1 | AACGGGGATATAGGTCAAAC | | 45,0 | 58,0 | 51,7 | 1,54 | -3,91 |
| | V2-V6 | 100% | A.min23S_135_21 | 1 (1) | 1 | 1 | GGACTCATCTCTGATACTTG | | 47,6 | n.d. | 51,9 | 0,63 | -4,64 |
| <i>Bacteroides ssp.</i> | V3-V7 | 81% | Bact16S_815_20 | 111 (64) | 316 | 1475 | ACGATGAATACTCGCTGTTT | x x | 40,0 | 56,0 | 51,8 | 1,09 | -5,19 |
| | V3-V7 | 81% | Bact16S_226_20 | 103 (64) | 109 | 307 | GATGGGGATGCGTTCCATTA | x x | 50,0 | 60,0 | 55,0 | 0,13 | -6,5 |
| <i>Bacteroides fragilis</i> | V1-V7 | 100% | B.frag23S_152_20 | 1 (1) | 1 | 1 | AACTTGGTTGGAGGCTAACG | x x | 50,0 | 60,0 | 55,2 | 1,22 | -3,61 |
| | V1-V7 | 100% | B.frag23S_1025_20 | 1 (1) | 1 | 1 | AACTAACGAAGTCAGATTGC | x x | 40,0 | 56,0 | 50,3 | -0,12 | -3,61 |
| <i>Bifidobacterium ssp.</i> | V4-V7 | 100% | B.bif23S_2205_19 | 3 (3) | 3 | 3 | ACAGTCATCTGTTCCAGGG | x x | 52,6 | 58,0 | 54,1 | 0,04 | -6,62 |
| <i>Bifidobacterium bifidum</i> | V4-V7 | 100% | B.bif16S_183_20 | 7 (5) | 7 | 15 | ATGATCGCATGTGATTGTGG | x x | 45,0 | 58,0 | 53,4 | 0,97 | -5,38 |
| | V1-V7 | 100% | B.bif23S_2168_20 | 1 (1) | 1 | 1 | AAATACCGCTCTGTTCTCGT | x x | 45,0 | 58,0 | 53,8 | 1,57 | -3,61 |
| | V4-V7 | 100% | B.bif23S_2178_20 | 1 (1) | 1 | 1 | TCTGTTCTCGTTGATGTCT | x x | 40,0 | n.d. | 51,2 | 1,16 | -3,61 |
| <i>Campylobacter coli</i> | V1-V6 | 100% | C.col23S_756_20 | 3 (3) | 3 | 3 | CAGGGTAGACTTGTGGATAG | | 50,0 | 60,0 | 51,7 | 1,11 | -1,95 |
| | V1 | 100% | C.col23S_755_20 | 3 (3) | 3 | 3 | CCAGGGTAGACTTGTGGATA | | 50,0 | 60,0 | 53,2 | 0,79 | -5,02 |
| | V1-V6 | 100% | C.col23S_1400_20 | 3 (3) | 3 | 3 | ATACCAACATTAGTCGTCGC | | 45,0 | 58,0 | 52,4 | 0,93 | -3,61 |
| <i>Campylobacter jejuni</i> | V3-V7 | 100% | C.jej23S_387_21 | 5 (5) | 9 | 24 | AGGAATCCTGTCTGAATCCGG | x x | 52,4 | n.d. | 56,4 | 1,19 | -9,75 |
| | V3-V6, V7s | 100% | C.jej23S_394_19 | 5 (5) | 9 | 35 | CTGTCTGAATCCGGGTCGA | x | 57,9 | n.d. | 56,5 | 1,04 | -9,75 |
| | V1-V6 | 100% | C.jej23S_1153_24 | 4 (4) | 4 | 4 | GAAGCTGTAGACTTAGTTACTAA | | 33,3 | n.d. | 49,9 | 0,97 | -6,34 |
| <i>Campylobacter ssp.</i> | V1-V6 | 100% | Camp16S_414_20 | 142 (136) | 155 | 334 | CACTTTTCGGAGCGTAAACT | | 45,0 | 58,0 | 53,1 | 1,31 | -3,89 |
| | V1-V6 | 100% | Camp16S_405_20 | 141 (136) | 145 | 510 | GGAGGATGACACTTTTCGGA | | 50,0 | 60,0 | 54,4 | 1,79 | -3,61 |
| | V1-V6, V7s | 100% | Camp23S_2631_22 | 13 (13) | 13 | 13 | AAGATTGAAGAGATTTGACCCT | x | 36,4 | n.d. | 51,6 | 0,53 | -1,94 |
| | V1-V7 | 100% | Camp23S_932_18 | 13 (13) | 13 | 13 | AATCACAGCAGTCAGGCG | x x | 55,6 | n.d. | 55,5 | 0,41 | -3,61 |
| | V1-V7 | 100% | Camp23S_2174_22 | 13 (13) | 13 | 13 | CACTCTTTCTATTTGGGTAGC | x x | 40,9 | n.d. | 51,2 | 1,98 | -3,14 |
| <i>Clostridium difficile</i> | V1-V7 | 100% | C.diff16S_992_20 | 5 (4) | 6 | 37 | GACATCCCAATGACATCTCC | x x | 50,0 | 60,0 | 52,9 | 0,04 | -3,43 |
| | V1-V7 | 100% | C.diff16S_1012_20 | 4 (4) | 4 | 40 | TAATCGGAGAGTCCCTTCG | x x | 50,0 | 60,0 | 53,7 | 0,32 | -4,64 |
| <i>Clostridium ssp.</i> | V1-V6 | 100% | Clos23S_1038_24 | 21 (21) | 22 | 45 | GATTTCTAAGACAAGTGGATGTT | | 33,3 | n.d. | 50,4 | 0,85 | -4,16 |
| | V6-V7 | | Clos23S_1038_22 | n.d. | n.d. | n.d. | GATTTCTAAGACAAGTGGATG | | 36,4 | n.d. | 48,2 | 1,3 | -4,16 |
| <i>Enterococcus ssp.</i> | V2-V6 | 86% | Entc16S_1281_22* | 131 (101) ^{fsd} | 182 | 297 | CTCTTAAAGCTTCTCTCAGTTC | | 40,9 | n.d. | 50,7 | 1,2 | -10,23 |
| | V2-V6 | 89% | Entc23S_2618_18 | 59 (63) | 160 | 293 | GTCCGGGGCGTTGAAAT | | 61,1 | n.d. | 59,2 | 0,75 | -10,36 |
| | V2-V7 | 91% | Entc23S_175_19 | 63 (63) | 77 | 101 | GAGGTAGACGCAGAGAACT | x x | 52,6 | 58,0 | 53,4 | 1,68 | -3,61 |
| <i>Enterococcus faecalis</i> | V1-V7 | | E.faec23S_1470_23 | | | | GCAATGAGTCTTGAGTAGAGTTA | x x | 39,1 | n.d. | 52,0 | 0,83 | -3,9 |
| | V1-V6, V7i | 100% | E.faec23S_1172_20 | 4 (4) | 4 | 25 | ATTAGGTGTAGTGGTAGGAG | x | 45,0 | 58,0 | 50,3 | 2,38 | -1,47 |
| | V1-V7 | 100% | E.faec23S_284_25 | 4 (4) | 4 | 5 | TGTTAGTATAGTTGAAGGATTTGGA | x x | 32,0 | n.d. | 51,7 | 1,17 | -3,4 |
| <i>Escherichia coli K12+EHEC</i> | V1-V6 | | E.coli23S_271_18B | 47 (51) | 47 | 56 | AGCCTGAATCAGTGTGTG | x x | 50,0 | n.d. | 52,3 | 1,15 | -5,13 |
| <i>Escherichia coli K12</i> | V1-V7 | 92% | E.coli23S_542_18 | 35 (34) | 35 | 44 | GCTTAGCGGTGTGACTGC | x x | 61,1 | n.d. | 56,6 | 0,98 | -3,61 |
| <i>Escherichia coli O157:H7</i> | V1-V7 | 100% | EHEC23S_1164_18 | 14 (14) | 16 | 31 | AGCGACACTGTGTGTTGT | x x | 50,0 | 54,0 | 54,6 | -0,4 | -4,64 |

Tab. 3.2: Continuation from previous page.

| target | array ^{a)} | in-silico coverage | probe name ^{b)} | PM _{wm} (0) of (n) ^{c)} | PM _{wm} (1) | PM _{wm} (2) | probe 5'→3' (sense) ^{d)} | final array ^{e)} s i | GC (%) | Tm (ARB) ^{e)} | Tm (OIAn) ^{e)} | dG Hairpin ^{f)} (kcal/mol-1) | dG Selfdimer ^{g)} (kcal/mol-1) |
|---|---------------------|--------------------|--------------------------|---|----------------------|----------------------|-----------------------------------|----------------------------------|--------|------------------------|-------------------------|---------------------------------------|---|
| <i>Eubacterium bifforme</i> | V4-V7 | 100% | E.bifo16S_193_20 | 1 (1) | 1 | 3 | GCTATTTAAAGTGGCTAAGGC | x x | 45,0 | 58,0 | 51,1 | 0,83 | -4,85 |
| <i>Fusobacterium prausnitzii</i> | V1-V6, V7s | 100% | F.praus16S_171_18 | 51 (7) | 401 | 1348 | ATACCGCATAGCCACG | x | 55,6 | 56,0 | 55,1 | 0,51 | -3,61 |
| | V1-V6, V7s | 100% | F.praus16S_738_18 | 63 (7) | 102 | 1389 | CTGGGCACCAACTGACGC | x | 66,7 | 60,0 | 59,5 | -0,98 | -5,02 |
| <i>Lactobacillus delbrueckii</i> | V3-V7 | 100% | L.delb16S_176_22 | 17 (8) | 17 | 21 | GGATAACAACATGAATCGCATG | x x | 40,9 | n.d. | 52,3 | -0,2 | -5,38 |
| <i>Lactobacillus acidophilus</i> | V4-V7 | 100% | L.acid16S_997_19 | 3 (4) | 13 | 93 | CTAGTGCAATCCGTAGAGA | x x | 47,4 | 56,0 | 50,8 | 1,79 | -7,05 |
| <i>Lactococcus lactis</i> | V2-V6, V7i | 100% | L.lac23S_1581_23 | 7 (7) | 7 | 7 | CGTAAAGTCATATCTACCCGTAC | x | 43,5 | n.d. | 52,3 | 1,91 | -3,91 |
| | V2-V6 | 100% | L.lac23S_2198_18 | 7 (7) | 7 | 15 | AACCCGCTGGCATAATCG | x x | 55,6 | n.d. | 55,6 | 0,89 | -3,61 |
| | V2-V7 | 100% | L.lac23S_1436_21 | 7 (7) | 7 | 7 | GCTAAGAGATGCCAGTTAATG | x x | 42,9 | n.d. | 51,1 | 1,38 | -4,85 |
| <i>Listeria monocytogenes</i> | V1-V6 | 79% | L.mono16S_1274_27 | 22 (19) | 51 | 808 | GAGCTAATCCATAAAACTATTCTCAG | | 37,0 | n.d. | 53,4 | 1,46 | -6,34 |
| | V3-V6, V7i | | L.mono16S_1276_23 | 24 (16) | 127 | 1473 | GCTAATCCATAAAACTATTCTC | x | 34,8 | n.d. | 49,0 | 1,76 | -3,42 |
| <i>Listeria ssp.</i> | V2-V6 | 100% | List16S_1151_19 | 45 (43) | 55 | 443 | AAGTAGCTGCCGTGCAAGC | | 57,9 | n.d. | 58,4 | 0,00 | -9,75 |
| | V2-V7 | 100% | List23S_2175_20 | 22 (22) | 22 | 22 | ACCCTGGCTGTATGACCATT | x x | 50,0 | 60,0 | 56,2 | 1,20 | -5,02 |
| <i>Mycobacterium avium</i> complex | V2-V7 | 100% | M.aviC23S_650_17 | 2 (2) | n.d. | 8 | CGCATCCCCTTTGGGGT | | 64,7 | n.d. | 58,3 | -0,79 | -9,21 |
| | V2-V6 | 100% | M.aviC23S_2443_18 | 2 (2) | 2 | 614 | GGGGATAACGGGCTGATC | x x | 61,1 | 58,0 | 55,2 | 1,08 | -4,62 |
| <i>Mycobacterium avium</i> subsp. <i>paratuberculosis</i> | V2-V7 | 100% | M.aviP23S_650_17 | 2 (2) | 4 | 4 | CGCATCCCCTTTGGGGT | | 58,8 | n.d. | 55,7 | -0,86 | -6,14 |
| | V2-V6 | 100% | M.aviP23S_1495_19 | 2 (2) | 2 | 3 | ATCCGTCGGTCACTAATCC | x x | 52,6 | 58,0 | 53,9 | 2,00 | -3,61 |
| <i>Mycobacterium ssp.</i> | V2-V7 | 100% | Mycb16S_987_20 | 426 (n) | 455 | 568 | GGTTTGACATGCACAGGACG | x x | 55,0 | n.d. | 56,7 | 1,56 | -7,05 |
| | V2-V7 | 100% | Mycb23S_1872_20 | 21 (17) | 21 (17) | 27 (17) | GGGTGAAGCGGAGAATTTAA | x x | 45,0 | 58,0 | 52,7 | 1,94 | -5,36 |
| <i>Plesiomonas shigelloides</i> | V1-V7 | 100% | P.shig16S_468_20 | 4 (4) | 4 | 19 | CCTAGTGGCATTGACGTTAC | x x | 50,0 | 60,0 | 53,2 | 1,19 | -6,3 |
| | V2-V6 | 100% | P.shig23S_714_19 | 3 (3) | 3 | 3 | AACACGTAACCTGGAGGACC | | 52,6 | 58,0 | 54,3 | 1,07 | -6,3 |
| | V2-V7 | 100% | P.shig23S_1475_19 | 3 (3) | 3 | 3 | AAGGTGGAGAGACTAGGCA | x x | 52,6 | 58,0 | 54,9 | 1,78 | -4,16 |
| <i>Roseburia intestinalis</i> | V1-V7 | 100% | R.inte16S_996_18 | 12 (2) | 20 | 36 | TCCCGATGACAGAACATG | x x | 50,0 | 54,0 | 51,7 | 0,75 | -5,38 |
| | V3-V7 | 100% | R.inte16S_1011_21 | 9 (3) | 15 | 15 | TGTAATGTGTTTTCTCTTCGG | x x | 38,1 | n.d. | 50,6 | 1,98 | -3,61 |
| <i>Roseburia ssp.</i> | V3-V7 | 100% | Rose16S_834_18 | 25 (6) | 41 | 46 | GGGGAGCATTTGCTCTTCG | x x | 61,1 | 58,0 | 56,2 | -0,15 | -8,27 |
| <i>Salmonella ssp.</i> | V4-V6, V7s | 82% | Salm23S_274_20 | 29 (28) | 97 | 103 | CTGAATCAGCATGTGTGTTA | x | 40,0 | 56,0 | 50,4 | 0,98 | -5,38 |
| | V1-V7 | 85% | Salm23S_1704_17 | 32 (29) | 40 | 109 | ACGCTGACACGTAGGTG | x x | 58,8 | n.d. | 54,7 | -0,11 | -6,3 |
| <i>Shigella flexneri</i> | V4-V6 | 100% | S.flex23S_1704_19 | 2 (2) | 102 | 116 | ACGCTGATACGTAGGTGAA | | 47,4 | 56,0 | 53,1 | 0,33 | -8,22 |
| <i>Shig. flexneri</i> + <i>Shig. boydii</i> | V4-V6 | 100% | S.flex.boyd23S_1215_19 | 3 (3) | 3 | 74 | GTGTGGTGTGAGGTATGCT | | 52,6 | 58,0 | 54,9 | 1,3 | -3,14 |
| <i>Shigella sonnei</i> | V4-V6 | 100% | S.sonn23S_2153_17 | 1 (1) | 1 | 156 | CATGGAGCCGGCCTTGA | | 64,7 | n.d. | 58,2 | 1,44 | -16,03 |
| <i>Shigella dysenteriae</i> | V4-V6 | 100% | S.dys23S_842_17 | n.d. | n.d. | n.d. | AACTCATCTCGGGGGGT | | 58,8 | n.d. | 55,6 | 2,12 | -3,61 |
| <i>Veillonella ssp.</i> | V1-V7 | 100% | Veill16S_985_20 | 33 (27) | 35 | 76 | AGGTCTTGACATTGATGGAC | x x | 45,0 | 58,0 | 52,2 | -0,12 | -3,43 |

Tab. 3.2: Continuation from previous page.

| target | array ^{a)} | in-silico coverage | probe name ^{b)} | PM _{wm} (0) of (n) ^{c)} | PM _{wm} (1) | PM _{wm} (2) | probe 5'-->3' (sense) ^{d)} | final array ^{e)} s i | GC (%) | Tm (ARB) ^{e)} | Tm (OIAn) ^{e)} | dG Hairpin ^{f)} (kcal ^o mol ⁻¹) | dG Selfdimer ^{g)} (kcal ^o mol ⁻¹) | |
|--|---------------------|--------------------|--------------------------|---|----------------------|----------------------|-------------------------------------|----------------------------------|--------|------------------------|-------------------------|---|---|--------|
| <i>Vibrio cholerae</i> | V1-V6 | 100% | V.chol16S_831_17 | 18 (15) | 764 | 837 | AGGTTGTGCCCTAGAGG | | 58,8 | n.d. | 53,7 | 0,55 | -4,67 | |
| | V1-V7 | 83% | V.chol23S_1379_20 | 10 (10) | 11 | 17 | CCCGTACTTCTGACTATTGC | x x | 50,0 | 60,0 | 52,7 | 2,35 | -3,65 | |
| | V1-V7 | 83% | V.chol23S_1207_21 | 10 (10) | 14 | 64 | CGTTGAAGGTGAATCGTAAGG | x | 47,6 | n.d. | 53,5 | 0,46 | -3,61 | |
| | V1-V7 | 83% | V.chol23S_1580_21 | 10 (10) | 10 | 10 | GCTTCAGATAGTCAGGAATCG | x x | 47,6 | n.d. | 52,7 | 1,53 | -3,61 | |
| | V1-V7 | 83% | V.chol23S_133_22 | 10 (10) | 10 | 10 | TGCATAAGCAGTTACTGTTAAC | x x | 36,4 | n.d. | 50,9 | -0,28 | -7,53 | |
| <i>Vibrio parahaemolyticus</i> | V4-V7 | 100% | V.par23S_1212_19 | 1 (1) | 1 | 1 | AAGGTGTGTGGTAACGCAT | x x | 47,4 | 56,0 | 54,6 | -0,03 | -3,61 | |
| | V4-V7 | 100% | V.par23S_339_19 | 1 (1) | 1 | 4 | AGTTGACGACGTGTGTTC | x x | 47,4 | 56,0 | 54,3 | 0,79 | -6,3 | |
| | V4-V6 | 100% | V.par23S_1523_18 | 1 (1) | 1 | 4 | CGTCGAGCATCTACGGAT | | 55,6 | n.d. | 54,0 | 1,23 | -6,76 | |
| <i>Vibrio parahaemolyticus et rel.</i> | V3-V7 | n.d. | V.parRel16S_63_19 | 109 | 135 | 165 | CGAGCGGAAACGAGTTATC | x x | 52,6 | n.d. | 53,3 | -1,17 | -3,61 | |
| <i>Vibrio ssp.</i> | V3-V6, V7i | 80% | Vibr16S_570_20 | 392 (375) | 559 | n.d. | TAAAGCGCATGCAGGTGGTT | | x | 50,0 | 60,0 | 57,8 | 0,12 | -11,66 |
| <i>Yersinia enterocolitica</i> | V1-V7 | 100% | Y.ent23S_1478_18 | 3 (3) | 3 | 4 | GCGGAGTGACCAGGTA | x x | 55,6 | n.d. | 54,5 | 0,31 | -4,41 | |
| | V1-V6, V7i | 100% | Y.ent23S_1501_20 | 3 (3) | 3 | 4 | TTGCTTATCAACGCTGAGGT | x x | 45,0 | 58,0 | 54,4 | 0,97 | -3,09 | |
| <i>Yersinia ssp.</i> | V1-V6 | 100% | Yers16S_1274_20 | 88 (87) | 122 | 363 | AGCGGACCACATAAAGTCTG | | 50,0 | 60,0 | 54,7 | 0,81 | -3,61 | |

a) array version on which probe is present; i - only antisense probe is present, s - only sense probe is present

b) target_E.coli-position_length; antisense probes are labelled with "i" in the end

c) number of perfect matches of the probe with zero, one, or two weighted mismatches towards the target against the whole database (except for 'fsd' = against full sequence database)

d) plus 5'-polyT spacer with a length of 11 nucleotides and amino modification

e) °C; ARB: 2AT+4GC; Oligoanalyzer: nearest neighbor

f) at 45°C and Na+ 50mM; determined for the sense probe

g) probe presence on final array, s - sense, i - antisense

(n.d.) not determined, (*) wrong sequence ordered and spotted; corresponds to Entc16S_1283_22

For all selected species and genera, besides *Prevotella* and *Anaeroplasma*, a group- or species-specific probe based on the 16S or 23S ribosomal gene was finally found. The melting temperature according to OligoAnalyzer (nearest neighbour method) varied between 48.2 and 59.5°C. In total, 70 out of 87 probes matched the predefined T_M parameter of >50°C and <56°C. Regarding the hairpin and selfdimer stability, 73 and 72 probes matched the defined parameters of ≥ 0 kcal* mol^{-1} and ≥ -7 kcal* mol^{-1} , respectively. The highest hairpin stability was found for a *Vibrio parahaemolyticus* probe (V.parRel16S_63_19/i) with -1.17 kcal* mol^{-1} due to only two neighbored GC pairs. The highest selfdimer stability was -16.03 kcal* mol^{-1} for a *Shigella sonnei* probe (S.sonn23S_2153_17/i) due to six neighbored GC pairs. The GC content of all probes varied between 32% (E.faec23S_284_25/i) and 66.7% (F.praus16S_738_18/i). The distribution of the probes of the final array V6f across both target genes is depicted in Fig. 3.1.

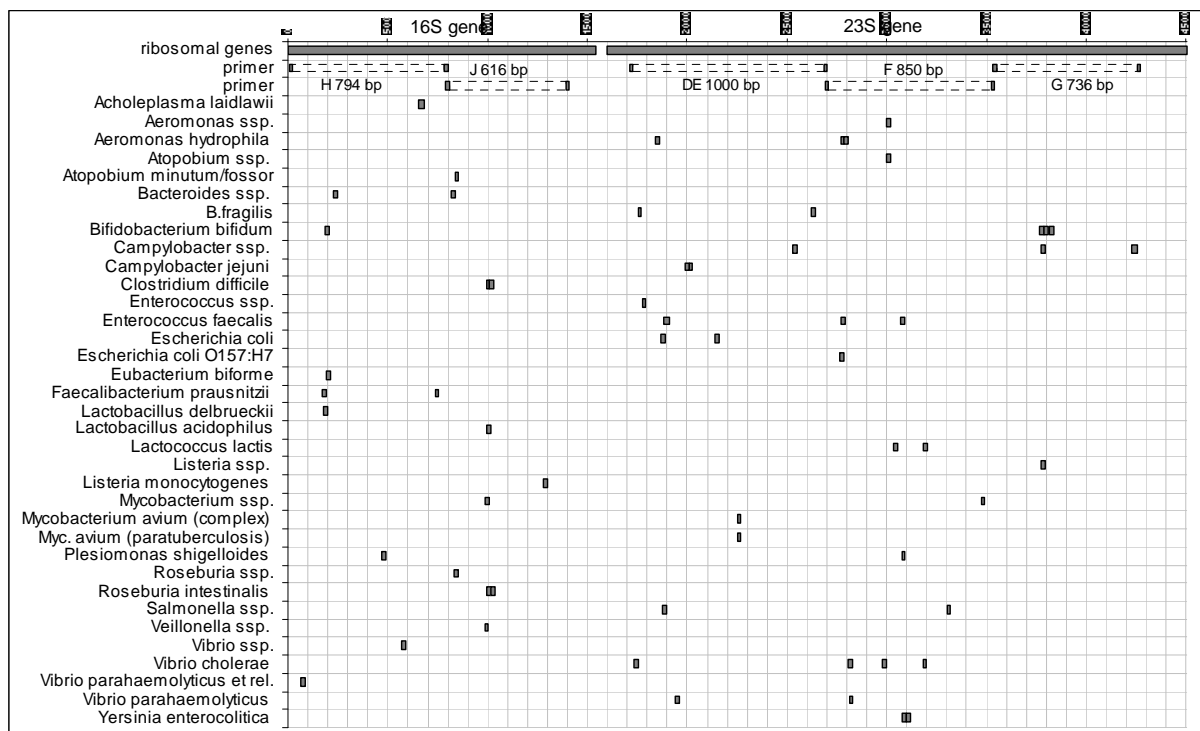


Fig. 3.1: Distribution of detection probes (array V6f) for all species on the ribosomal genes and position of the primers used for amplification.

3.2 Array preparation – theoretical considerations

The probes were spotted according to chapter 2.2.2. The theoretical probe density per spot was calculated under assumption that the probe solution is spotted as a hemisphere. The diameter of one spot was measured as 120 μm . The spotted probe solution had a concentration of 20 μM . The probe DNA per spot was calculated to be 9.04 fmol by multiplying the volume of a hemisphere above the spot area with the probe concentration in the spotting solution using equation (1).

$$\text{probe density [fmol DNA/spot]} = \frac{4/3 * \pi * r^3}{2} * 10^{-6} * [\text{probe}] \quad (1)$$

$$r = 60 \mu\text{m}; 1 \mu\text{m}^3 = 1 \text{fL}; [\text{probe}] = 20 \text{fmol/nL}$$

Using equation (2) and an average molecular weight of $327 \text{ g} \cdot \text{mol}^{-1}$ for a nucleotide (Dolezel 2003), the mass of probe DNA per spot was calculated by multiplying the molecular weight of one probe with the probe density per spot in fmol. The theoretical probe density was $91,639 \text{ fg/spot}$.

$$\text{probe density [fg / spot]} = \langle \text{probe length} \rangle * 327 \text{ g} * \text{mol}^{-1} * \langle \text{probe density fmol / spot} \rangle \quad (2)$$

Average probe length with linker = 31 nt

The density of immobilized probe molecules in one spot was then calculated by equation (3) by dividing the probe mass per spot by the mass of one probe with 31 nucleotides.

$$\text{probe density [molecules / spot]} = \frac{\langle \text{probe density fg / spot} \rangle}{\langle \text{probe length} \rangle * 327 \text{ g} * \text{mol}^{-1} / \text{AVOGADRO}} * 10^{-15} \quad (3)$$

Average probe length with linker = 31 nt; AVOGADRO constant = $6.022 * 10^{23} \text{ mol}^{-1}$

The calculated probe density was $5.4 * 10^9$ probes/spot, which is for a spot diameter of $120 \mu\text{m}$ $4.8 * 10^{13}$ probe molecules/cm².

3.3 Extraction of DNA reference material

DNA extractions from bacterial pure cultures and faecal samples from gastroenteritis patients were performed at the University Hospital of Giessen by our project partners according to protocol 2.2.3.1, as work with human isolates was not allowed at the Institute for Technical Biochemistry. The DNA content varied between 33 and 413 ng/ μL in pure culture isolates (Tab. 3.3) and 2.4 and 65.3 ng/ μL in faecal isolates (Tab. 3.5). The DNA quality control by agarose gel electrophoresis revealed a high degree of fragmentation within the DNA isolates. In contrast, the DNA extraction from an apathogenic *E. coli* DH5 α according to protocol 2.2.3.2 yielded high molecular DNA (Fig. 3.2 (A)).

The DNA extraction from porcine faecal samples and one pure culture of *E. coli* according to protocol 2.2.3.3 yielded high molecular DNA with an average fragment size above 5,000 bp (Fig. 3.2 (B)) and a DNA content between 34 and 155 ng/ μL (Tab. 3.4).

Extracted DNA from healthy and rotavirus-infected children was provided by the group of Prof. Liping Zhao at Shanghai Jiao Tong University after extraction according to protocol 2.2.3.3. The DNA content varied between 6 and 23 ng/ μL (Tab. 3.4).

Tab. 3.3: DNA isolates from bacterial pure cultures extracted for this work.

| Pure culture isolates | Internal number | Reference | [DNA] ng/μL ^{a)} | 260/280 | 260/230 |
|--|-----------------|-----------------|---------------------------|---------|---------|
| <i>Aeromonas hydrophila</i> (*) | 24 | DSM30019 | 642,0 | 1,79 | 0,86 |
| <i>Aeromonas trota</i> (*) | 23 | DSM7312 | 248,0 | 1,25 | 0,26 |
| <i>Atopobium vaginiae</i> | 4 | DSM15829 | 44,6 | 2,19 | n.d. |
| <i>Atopobium minutum</i> | 22 | DSM20586 | 231,0 | 1,23 | 0,25 |
| <i>Bacteroides ureolyticus</i> | 21 | DSM20703 | 227,0 | 1,58 | 0,47 |
| <i>Bacteroides fragiles</i> | 20 | DSM9671 | 198,0 | 1,13 | 0,21 |
| <i>Bifidobacterium bifidum</i> | 19 | DSM20456 | 131,0 | 1,14 | 0,19 |
| <i>Campylobacter coli</i> | 8 | DSM4689 | 66,1 | n.d. | n.d. |
| <i>Campylobacter jejuni</i> | 7 | DSM4688 | 65,2 | n.d. | n.d. |
| <i>Campylobacter lari</i> | 9 | DSM11375 | 38,3 | 2,60 | n.d. |
| <i>Clostridium difficile</i> | 11 | DSM12056 | 112,3 | 1,90 | n.d. |
| <i>Clostridium haemolyticum</i> | 31 | DSM5565 | 387,0 | 1,48 | 0,56 |
| <i>Enterococcus faecium</i> (*) | 17A | UR13873 | 45,1 | 2,62 | 0,43 |
| <i>Enterococcus faecium</i> | 17B | UR15676 | 53,4 | 3,01 | 0,70 |
| <i>Enterococcus faecalis</i> | 14 | clin. isol. | 45,6 | 2,72 | 0,62 |
| <i>Escherichia coli</i> DH5 α | 12 | ITB | 35,6 | 2,02 | n.d. |
| <i>Escherichia coli</i> DH5 α | 12B | Shanghai | 124,0 | n.d. | n.d. |
| <i>Escherichia coli</i> O157:H7 (EHEC) | 33 | ITB clin. isol. | 180,0 | n.d. | n.d. |
| <i>Lactobacillus delbrueckii</i> | 30 | DSM20074 | 191,0 | 1,38 | 0,28 |
| <i>Lactococcus lactis</i> | 29 | DSM4644 | 127,0 | 1,27 | 0,21 |
| <i>Listeria monocytogenes</i> EGD-e | 6A | | 42,7 | n.d. | n.d. |
| <i>Listeria monocytogenes</i> EGD-e | 6B | | 46,0 | 1,97 | 0,20 |
| <i>Listeria innocua</i> 6b | 1 | | 33,0 | n.d. | n.d. |
| <i>Mycobacterium avium avium</i> | 27 | DSM44158 | 340,0 | 1,03 | 0,31 |
| <i>Mycobacterium avium silvaticum</i> | 28 | DSM44175 | 413,0 | 1,05 | 0,72 |
| <i>Myc. avium paratuberculosis</i> | 15 | DSM44133 | 42,4 | 2,76 | 0,08 |
| <i>Mycobacterium tuberculosis</i> | 16 | clin. isol. | 35,3 | 1,75 | 0,32 |
| <i>Plesiomonas shigelloides</i> | 10 | DSM8224 | 38,6 | 2,12 | n.d. |
| <i>Roseburia intestinalis</i> | 2 | DSM14610 | 96,1 | 2,01 | n.d. |
| <i>Salmonella typhimurium</i> | 18 | 5275 | 305,0 | 1,76 | 0,58 |
| <i>Salmonella enteritidis</i> | 13 | | 48,2 | 2,58 | 0,49 |
| <i>Veillonella parvula</i> | 3 | DSM2008 | 32,9 | 2,94 | n.d. |
| <i>Vibrio parahaemolyticus</i> | 25 | DSM11058 | 150,0 | 1,23 | 0,20 |
| <i>Vibrio fischeri</i> | 26 | DSM9499/7151 | 186,0 | 1,33 | 0,25 |
| <i>Yersinia enterocolitica</i> | 5 | | 66,7 | 2,30 | n.d. |
| <i>Yersinia paratuberculosis</i> | 32 | DSM8992 | 371,0 | 1,54 | 0,43 |

a) blue values - determined by University Hospital of Giessen; results at the ITB were comparable
n.d. - not determined

Tab. 3.4: DNA isolates from Rotavirus-infected and healthy children and from porcine faecal samples extracted for this work (Shanghai/China).

| Faecal isolates from children | Internal number | [DNA] ng/μL | Faecal isolates from piglets | Internal number | [DNA] ng/μL | Internal number | [DNA] ng/μL |
|-------------------------------|-----------------|-------------|------------------------------|-----------------|-------------|-----------------|-------------|
| Rotavirus-infected | R2 | 18,0 | | | | | |
| Rotavirus-infected | R5 | 20,0 | | | | | |
| Rotavirus-infected | R15-3 | 7,0 | | | | | |
| Rotavirus-infected | R21-1 | 7,0 | | | | | |
| Rotavirus-infected | R22-1 | 8,0 | PFA piglet P2 | P2d14 | 96.0 | P2d21 | 296.0 |
| Rotavirus-infected | R23-1 | 23,0 | PFA piglet P3 | P3d14 | 37.0 | P3d21 | 40.0 |
| Rotavirus-infected | R24-1 | 6,0 | PFA piglet P4 | P4d14 | 115.0 | | |
| Rotavirus-infected | R25-1 | 15,0 | PFA piglet P5 | P5d14 | 70.0 | P5d21 | 249.0 |
| Rotavirus-infected | R26-1 | 8,0 | PFA piglet P6 | P6d14 | 41.0 | P6d21 | 550.0 |
| Rotavirus-infected | R27-1 | 10,0 | PFA piglet P7 | P7d14 | 98.0 | P7d21 | 167.0 |
| healthy child | H5 | n.d. | PFA piglet P8 | P8d14 | 155.0 | P8d21 | 695.0 |
| healthy child | H6-3 | n.d. | HFA piglet H1 | H1d14 | 119.0 | H1d21 | 481.0 |
| healthy child | H7 | n.d. | HFA piglet H2 | H2d14 | 125.0 | H2d21 | 768.0 |
| healthy child | H12 | n.d. | HFA piglet H3 | H3d14 | 131.0 | H3d21 | 402.0 |
| healthy child | H14-1 | n.d. | HFA piglet H4 | H4d14 | 122.0 | H4d21 | 209.0 |
| healthy child | H17-3 | n.d. | HFA piglet H5 | H5d14 | 142.0 | H5d21 | 202.0 |
| healthy child | H19 | n.d. | HFA piglet H6 | H6d14 | 127.0 | H6d21 | 559.0 |
| healthy child | H20 | n.d. | HFA piglet H7 | H7d14 | 102.0 | H7d21 | 213.0 |
| healthy child | H23 | n.d. | HFA piglet H8 | H8d14 | 131.0 | H8d21 | 354.0 |
| healthy child | H24 | n.d. | conventionally raised piglet | CV1d14 | 143.0 | CV1d21 | 101.0 |
| healthy child | H28 | n.d. | conventionally raised piglet | CV2d14 | 34.0 | CV2d21 | 48.0 |

n.d. - not determined

Tab. 3.5: DNA isolates from human faecal samples extracted for this work.

| Faecal isolates from patients | Internal number | Reference | [DNA] ng/ μ L | 260/280 | 260/230 |
|---------------------------------|-----------------|-----------|-------------------|---------|---------|
| PC 4300 <i>Clos.</i> (-) | S1 | PC 4300 | 21,3 | 1,77 | 1,64 |
| PC 4300 <i>Yers.</i> (-) | S2 | PC 4300 | 15,5 | 2,00 | 2,56 |
| ST 572 <i>Clos.</i> (-) | S3 | ST 572 | 8,1 | 2,55 | 0,22 |
| ST 14568 <i>Camp.</i> (-) | S4 | ST 14568 | 5,8 | 2,26 | 0,17 |
| ST 14568 <i>Salm.</i> (-) | S5 | ST 14568 | 6,4 | 1,70 | 0,26 |
| ST 14569 <i>Camp.</i> (-) | S6 | ST 14569 | 7,8 | 2,26 | 0,30 |
| ST 14569 <i>Salm.</i> (-) | S7 | ST 14569 | 8,1 | 1,96 | 0,30 |
| ST 14571 <i>Yers.</i> (-) | S8 | ST 14571 | 14,2 | 2,10 | 0,40 |
| ST 14565 <i>Yers.</i> (+) | S9 | ST 14565 | 29,7 | 1,97 | 0,97 |
| ST 14371 <i>Yers.</i> (+) | S10 | ST 14371 | 25,7 | 1,86 | 0,55 |
| ST 14256 <i>Camp.</i> (+) | S11 | ST 14256 | 5,8 | 2,12 | 0,18 |
| ST 13986 <i>Camp.</i> (+) | S12 | ST 13986 | 10,0 | 1,84 | 0,25 |
| ST 14496 <i>C.diff.</i> (+) | S13 | ST 14496 | 3,3 | 2,34 | 0,14 |
| ST 14501 <i>C.diff.</i> (+) | S14 | ST 14501 | 7,9 | 2,37 | 0,31 |
| ST 14429 <i>Salm.</i> (+) | S15 | ST 14429 | 4,0 | 3,19 | 0,21 |
| ST 14066 <i>Salm.</i> (+) | S16 | ST 14066 | 6,7 | 2,10 | 0,22 |
| ST 14672 <i>Salm.</i> (+) A | S17 | ST 14672 | 2,4 | 1,82 | 0,15 |
| ST 14672 <i>Salm.</i> (+) B | S18 | ST 14672 | 8,2 | 1,88 | 0,38 |
| 16114 <i>Salm.</i> (+) | S19 | 16114 | 17,3 | 2,22 | 0,52 |
| 16125 <i>Clost.-Tox A+B</i> (+) | S20 | 16125 | 9,3 | 2,08 | 0,44 |
| 16148 <i>Clost.-Tox A+B</i> (+) | S21 | 16148 | 6,1 | 1,92 | 0,22 |
| 16170 | S22 | 16170 | 7,0 | 1,65 | 0,26 |
| 16171 <i>Salm.</i> (+) | S23 | 16171 | 7,2 | 1,46 | 0,20 |
| 16172 <i>Clost.-Tox A+B</i> (+) | S24 | 16172 | 6,0 | 2,18 | 0,20 |
| 16174 <i>Salm.</i> (+) | S25 | 16174 | 2,8 | 1,77 | 0,11 |
| 16178 | S26 | 16178 | 6,8 | 1,56 | 0,18 |
| 16184 <i>Salm.</i> (+) | S27 | 16184 | 7,3 | 1,56 | 0,29 |
| 16190 | S28 | 16190 | 13,6 | 1,70 | 0,34 |
| 16198 | S29 | 16198 | 2,8 | 1,14 | 0,13 |
| 16208 <i>Camp.</i> (+) | S30 | 16208 | 4,8 | 2,26 | 0,14 |
| 16209 | S31 | 16209 | 6,6 | 1,45 | 0,21 |
| 16211 <i>Clost.-Tox A+B</i> (+) | S32 | 16211 | 11,3 | 1,41 | 0,29 |
| 16212 | S33 | 16212 | 10,6 | 1,54 | 0,37 |
| 16213 | S34 | 16213 | 7,1 | 1,68 | 0,10 |
| 16214 | S35 | 16214 | 11,8 | 1,99 | 0,26 |
| 16220 <i>Clost.-Tox A+B</i> (+) | S36 | 16220 | 24,7 | 1,78 | 0,47 |
| 16224 | S37 | 16224 | 3,7 | 2,54 | 0,11 |
| 16270 | S38 | 16270 | 11,4 | 2,13 | 0,33 |
| 16283 <i>Salm.</i> (+) | S39 | 16283 | 26,9 | 1,83 | 0,58 |
| 16319 <i>Camp.</i> (+) | S40 | 16319 | 13,1 | 2,05 | 0,46 |
| 16320 | S41 | 16320 | 4,0 | 1,70 | 0,22 |
| 16321 | S42 | 16321 | 12,7 | 1,62 | 0,38 |
| 16325 <i>Salm.</i> (+) | S43 | 16325 | 42,1 | 1,68 | 1,13 |
| 16338 | S44 | 16338 | 11,2 | 1,50 | 0,49 |
| 16341 | S45 | 16341 | 65,3 | 1,41 | 0,41 |
| 16369 <i>Salm.</i> (+) | S46 | 16369 | 4,7 | 1,33 | 0,22 |
| 16378 <i>Salm.</i> (+) | S47 | 16378 | 10,0 | 1,33 | 0,30 |
| 16384 | S48 | 16384 | 5,9 | 1,40 | 0,19 |
| 16386 <i>Salm.</i> (+) | S49 | 16386 | 10,5 | 1,55 | 0,47 |
| 16387 | S50 | 16387 | 8,9 | 1,32 | 0,30 |
| 16392 | S51 | 16392 | 18,2 | 1,28 | 0,32 |
| 16397 | S52 | 16397 | 3,7 | 1,25 | 0,13 |
| 16407 <i>Salm.</i> (+) | S53 | 16407 | 17,5 | 1,33 | 0,35 |
| 16449 <i>Salm.</i> (+) | S54 | 16449 | 21,9 | 1,45 | 0,39 |
| 16552 | S55 | 16552 | 20,2 | 1,69 | 0,51 |
| 16538 <i>Salm.</i> (+) | S56 | 16538 | 12,7 | 1,65 | 0,44 |
| 16452 <i>Salm.</i> (+) | S57 | 16452 | 10,3 | 1,54 | 0,32 |
| 16557 <i>Salm.</i> (+) | S58 | 16557 | 22,0 | 1,48 | 0,53 |
| Healthy individual | S59 | | 7,9 | n.d. | n.d. |
| Healthy individual | S60 | 20496 | n.d. | n.d. | n.d. |
| Healthy individual | S61 | 20503 | n.d. | n.d. | n.d. |
| Healthy individual | S62 | 20505 | n.d. | n.d. | n.d. |
| Healthy individual | S63 | 20522 | n.d. | n.d. | n.d. |
| Healthy individual | S64 | 20523 | n.d. | n.d. | n.d. |

n.d. - not determined, (+) - positive and (-) - negative for the pathogen

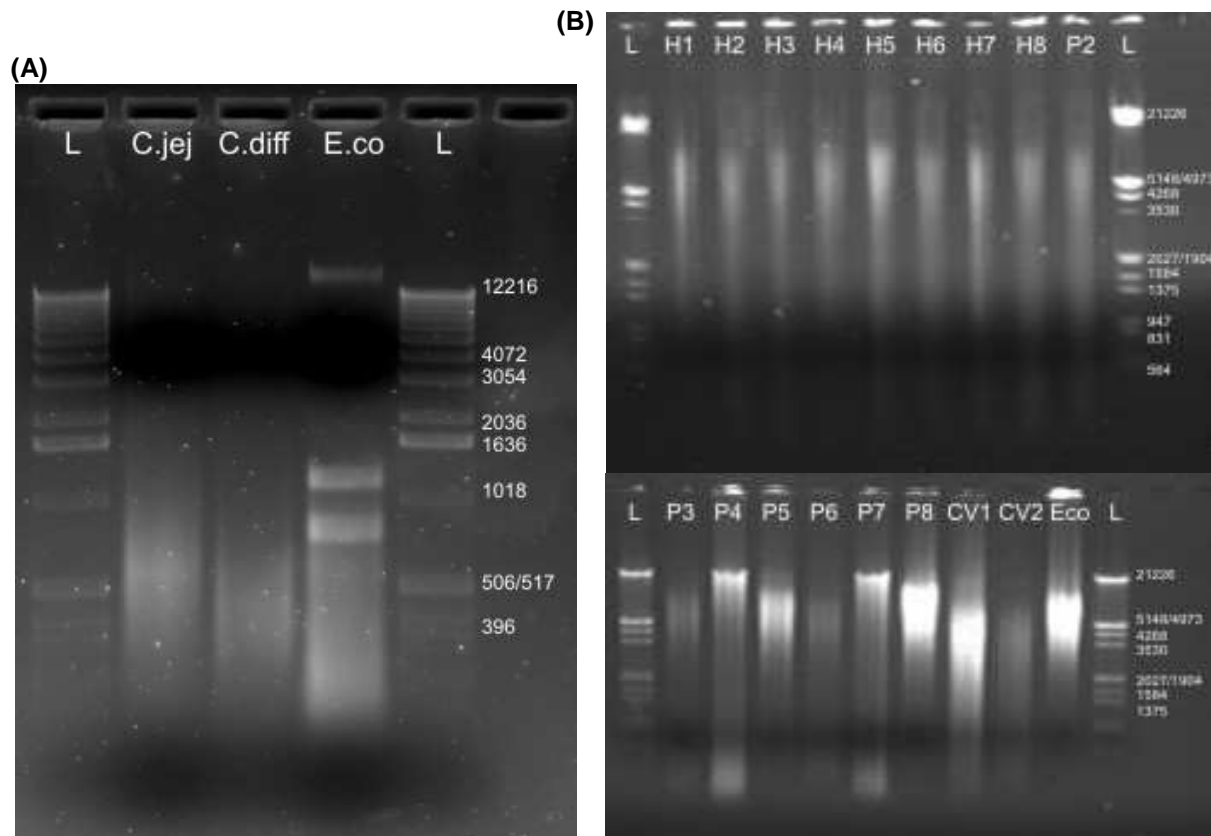


Fig. 3.2: (A) Agarose gel analysis of the isolated DNA from clinical samples (C.jej = *Campylobacter jejuni*; C. diff = *Clostridium difficile*) in contrast to an in-house *E. coli* DH5 α -isolate (E.co). The *E. coli*-isolate still contains ribosomal RNA (bands between 500 and 1,600 bp). (B) Agarose gel electrophoresis of isolated DNA from porcine faecal samples (HFA piglets H1-H8; PFA piglets P2-P8; conventionally raised piglets CV1, CV2) and one *E. coli* DH5 α overnight culture (Eco). L – ladder

3.4 Amplification of the DNA material

The first approach to amplify both ribosomal genes with one (Fig. 3.3 (A)) or two primer pairs (Fig. 3.3 (B)) according to protocol 2.2.4.1 failed in case of the clinical DNA isolates from pure cultures of *Campylobacter jejuni*, *Listeria monocytogenes* and *Yersinia enterocolitica*. The expected PCR products would have had a length of 4.5 kb (primers 616V/985R) or 2.1 kb (primers 616V/Lo110Ra) and 2.5 kb (primers Lo180Va/985R), respectively. Only for the *E. coli* DH5 α DNA isolate (extraction protocol 2.2.3.2) all expected fragments could be amplified. Additionally, a 4.5 kb band and a second band around 2.1 kb were observed in the duplex-PCR of *E. coli* (Fig. 3.3 (B)). For *C. jejuni* only the amplification of the 2.5 kb fragment in the duplex-PCR was successful. Duplex-amplification with reduced amount of template DNA (*C. jejuni*) of 50 and 20 ng in order to reduce the amount of potential inhibitors did not improve the result (data not shown). An agarose gel analysis of the template DNA revealed a high degree of fragmentation within the clinical DNA isolates in contrast to the *E. coli* DH5 α isolate (Fig. 3.2 (A)). Degradation fragments of the clinical DNA isolates ranged between 100 and 1,000 bp.

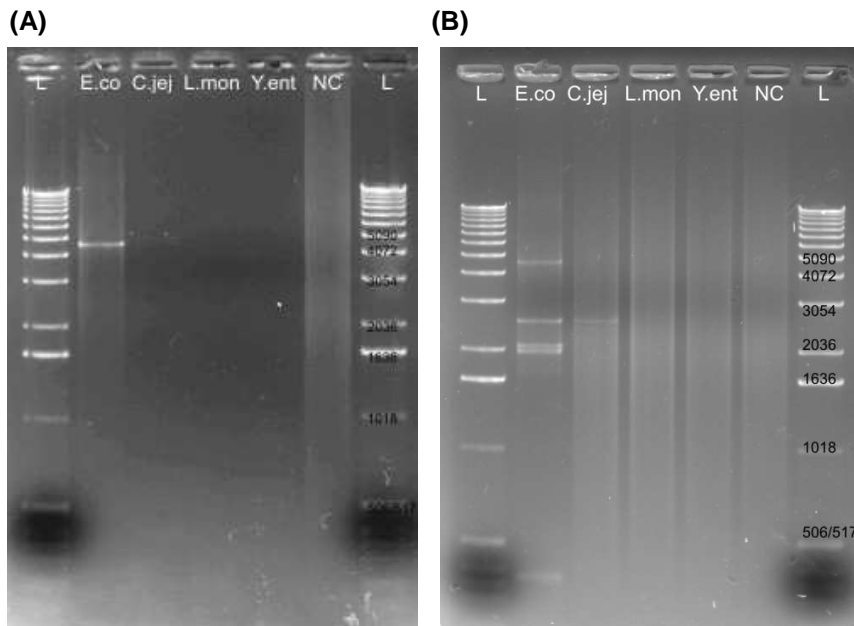


Fig. 3.3: (A) Agarose gel analysis of the PCR products using primer pair 616V/985R to amplify a 4.5 kb fragment containing 16S, ITS, and 23S genes from DNA isolates (*C.jej-Campylobacter jejuni*, *L.mon-Listeria monocytogenes*, *Y.ent-Yersinia enterocolitica*; Qiagen kit extraction) and one DNA isolate from *E. coli* DH5 α (*E.co*; phenol/ chloroform extraction). (B) Agarose gel analysis of the PCR products with primer pairs 616V/Lo110Ra and Lo180Va/985R to amplify a 2.1 kb and 2.5 kb fragment containing 16S+ITS and 23S, respectively, from the same isolates. NC - negative control, L - ladder

Due to the failed amplification of long DNA fragments from clinical DNA isolates, a set of seven primer pairs was designed, which was universal for all target species but contained three degenerated nucleotides per primer at maximum (Tab. 2.1). Three primer pairs span the 16S ribosomal gene and four ones the 23S ribosomal gene in nearly full length. The amplification of the fragment from *E. coli* according to protocol 2.2.4.2 using 1 mM Mg²⁺ and 5 U of *Taq* polymerase was successful, but fragment A and C contained spurious side products and fragment D was only weakly amplified (Fig. 3.4). Some weak contamination was seen in the negative controls for some fragments.

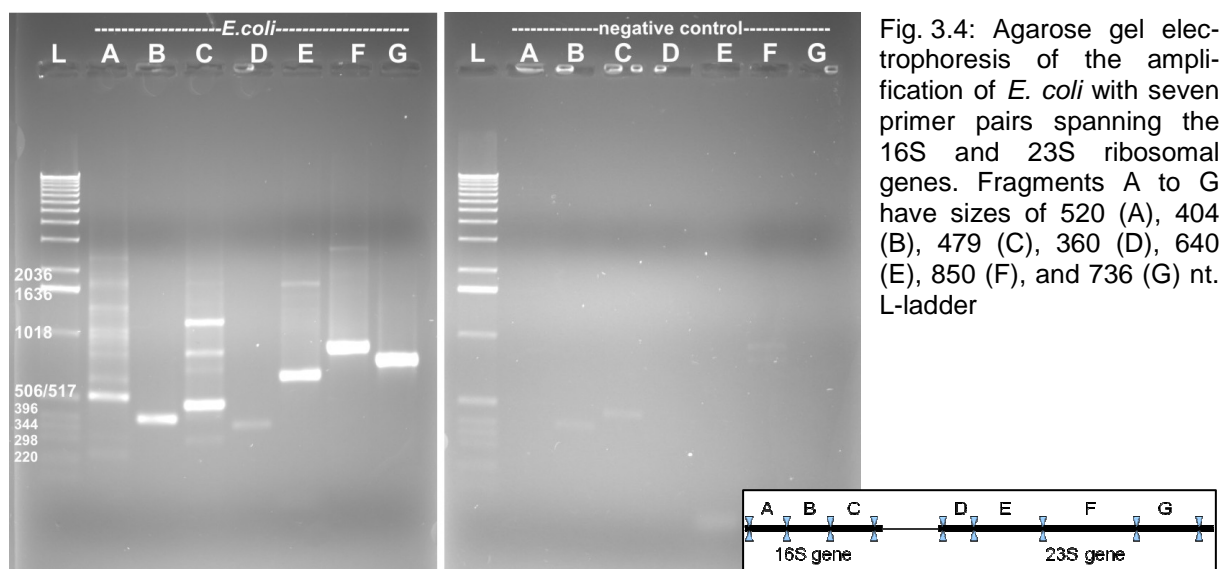


Fig. 3.4: Agarose gel electrophoresis of the amplification of *E. coli* with seven primer pairs spanning the 16S and 23S ribosomal genes. Fragments A to G have sizes of 520 (A), 404 (B), 479 (C), 360 (D), 640 (E), 850 (F), and 736 (G) nt. L-ladder

To reduce the amount of fragments to be amplified and to solve the problems of side products and weak amplification of one fragment, new primers for the 16S gene were designed and the fragments D and E were combined. Additionally, primer 114V was developed as alternative for Lo180Va, which did not amplify *C. difficile* due to a 3' mismatch,

but was partly used for amplification of the DE fragment of other species. Finally, primer 114V completely replaced primer Lo180Va. In the final primer combination, two primer pairs span the 16S rRNA gene and three ones the 23S rRNA gene in nearly full length. Primers of adjacent fragments slightly overlapped. Primers 781R and 790V had eleven base pairs in common resulting in a hetero dimer stability of $-19.71 \text{ kcal} \cdot \text{mol}^{-1}$. Primers 1084R and 1091V shared twelve nucleotides that gave a dimer stability of $-21.78 \text{ kcal} \cdot \text{mol}^{-1}$. The shared, eight nucleotides of primers 1923R and 1934V stabilized the dimer by $-14.07 \text{ kcal} \cdot \text{mol}^{-1}$. All other possible hetero dimers had stabilities below $-10 \text{ kcal} \cdot \text{mol}^{-1}$, in average around $-6 \text{ kcal} \cdot \text{mol}^{-1}$. The obtained PCR products had a length of 794, 616 bp (SSU gene), 1,000, 850, and 736 bp (LSU gene) regarding *E. coli*, but length slightly differed between species. The amplification result according to protocol 2.2.4.2 for the *E. coli* DH5 α isolate is depicted in Fig. 3.5 (A).

Fig. 3.5 (B) shows the successful amplification of the same five fragments in a multiplex-PCR. The DE fragment was amplified less effective than the other four fragments. In both cases for DE the forward primer, Lo180Va, was still used.

Amplification was tested with 31 reference strains, including target and some related non-target strains (Tab. 3.6). Only for *Veillonella parvula* and *Atopobium minutum* there could be no PCR product detected for primer pairs 1091V/1923R and 114V/1084R, respectively, whose fragments did not harbour probe regions for these species. Minor side products and some strong side products were found for some fragments, merely for fragment 790V/1390R (J), 114V/1084R (DE) and 1091V/1923R (F), in several species, which did not affect species detection on the microchip. Amplification efficiency was different for each fragment and species, varying by up to a factor of five, but enough PCR product was generated from all reference strains for hybridization.

Thus, a functional amplification strategy was developed, which could be implemented in the assay procedure.

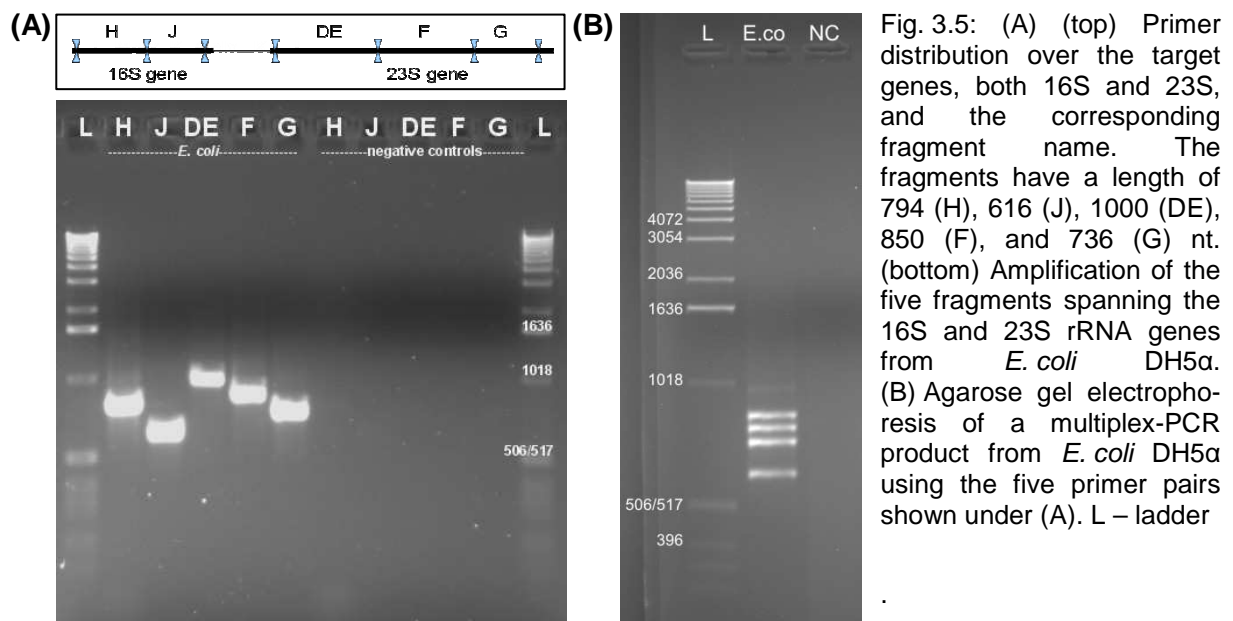


Fig. 3.5: (A) (top) Primer distribution over the target genes, both 16S and 23S, and the corresponding fragment name. The fragments have a length of 794 (H), 616 (J), 1000 (DE), 850 (F), and 736 (G) nt. (bottom) Amplification of the five fragments spanning the 16S and 23S rRNA genes from *E. coli* DH5 α . (B) Agarose gel electrophoresis of a multiplex-PCR product from *E. coli* DH5 α using the five primer pairs shown under (A). L – ladder

Tab. 3.6: Amplification results for 31 bacterial reference DNA isolates with five primer pairs amplifying the ribosomal genes in five fragments H, J, DE, F and G. (white) successful amplification, (grey) successful amplification with primer Lo180Va, (orange) no amplification, (+) genus-specific probe in this fragment located, (#) species-specific probe in this fragment located, (ssp/sp) strong sp/side product, (w) weak PCR product, (*) species later identified as different one.

| Nb. | Species | H | J | DE | F | G |
|-----|----------------------------------|---|------|-------|-----|---|
| 1 | <i>L. innocua</i> | | | | w | + |
| 2 | <i>R. intestinalis</i> | | + # | | | |
| 3 | <i>V. parvula</i> | | + | | | |
| 4 | <i>A. vaginae</i> | | + | | | |
| 5 | <i>Y. enterocolitica</i> | | | | # | |
| 6A | <i>L. monocytogenes</i> | | # | | w | + |
| 7 | <i>C. jejuni</i> | | | + # | | + |
| 8 | <i>C. coli</i> | | | + | | + |
| 9 | <i>C. lari</i> | | | + | | + |
| 10 | <i>P. shigelloides</i> | # | | | # | |
| 11 | <i>C. difficile</i> | | # | | | |
| 12 | <i>E. coli</i> | | | # | | |
| 13 | <i>S. enteritidis</i> | | | + | + | |
| 14 | <i>E. faecalis</i> | | | + # | # | |
| 15 | <i>M. avium paratuberculosis</i> | w | + w | # w | + w | w |
| 16 | <i>M. tuberculosis</i> | | + | | + | |
| 17A | <i>E. faecium</i> (*) | | | + | sp | |
| 17B | <i>E. faecium</i> | | | + | sp | |
| 18 | <i>S. typhimurium</i> | | sp | + | + | |
| 19 | <i>B. bifidum</i> | # | | | | # |
| 20 | <i>B. fragilis</i> | + | + sp | # ssp | | |
| 21 | <i>B. ureolyticus</i> | + | + | | | |
| 22 | <i>A. minutum</i> | | + sp | sp | | |
| 23 | <i>A. trota</i> (*) | | sp | sp | + | |
| 24 | <i>A. hydrophila</i> (*) | | sp | # | + # | |
| 25 | <i>V. parahaemolyticus</i> | + | sp | # sp | # | |
| 26 | <i>V. fischeri</i> | + | | | | |
| 29 | <i>L. lactis</i> | | | | # | |
| 30 | <i>L. delbrueckii</i> | # | | | | |
| 31 | <i>C. haemolyticum</i> | | | + ssp | | |
| 32 | <i>Y. pseudotuberculosis</i> | | ssp | sp | | |
| 33 | <i>E. coli</i> O157:H7 | | sp | | # | |

Additionally, a multiplex-PCR with the five primer pairs spanning the ribosomal genes was performed with all bacterial reference DNA isolates. For the DE fragment primer 114V was applied and the PCR protocol (chap. 2.2.4.2) was slightly changed. Only 10 ng of each template were subjected to amplification with all primers in equimolar amounts. The PCR cycling was prolonged taking 2 min of initial denaturation at 94°C; 30 cycles of DNA denaturation at 94°C for 45 s, primer annealing at 50°C for 45 s, and elongation at 72°C for 1.5 min; and a final extension step at 72°C for 4 min. The PCR products were analyzed by capillary gel electrophoresis (Agilent Bioanalyzer) (chap. 2.2.5). The amplification efficiency was highly variable for the different fragments depending on the amplified species and not all fragments could be clearly identified (0). Especially, the long DE fragment had poor amplification efficiency in most species. Only in *Campylobacter* species, it was amplified in comparable amounts compared to the other fragments. The highest amplification efficiency showed the J and G fragment. In some cases, the F and H fragment could not be distinguished, as can be seen in *L. innocua*, *L. monocytogenes*, *E. faecalis*, *B. fragilis*, *L. lactis*, and *L. delbrueckii*. Both fragments may have same length or one of the fragments was not amplified. In *Roseburia intestinalis*, the two overlapping peaks from fragments H and F still can be distinguished. In singleplex-PCR, the F fragment of *R. intestinalis* showed to be shorter compared to the F fragment of *E. coli*. In *Yersinia enterocolitica*, the F and DE fragment seemed to overlap. In singleplex-PCR, the F fragment showed increased length

compared to *E. coli*, while the DE fragment was slightly shorter. This was also observed for both mycobacteria, for *S. typhimurium*, *B. bifidum*, and *B. ureolyticus*. Therefore, in case of *B. ureolyticus* the DE fragment might be missing or completely overlap with the peak of the F fragment. Increased fragment length was also seen in singleplex-PCR for both mycobacteria regarding the DE fragment, which also proved true in the multiplex PCR. The DE fragments had a length of 1,142 nt and 1,222 nt. In contrast to the unsuccessful single amplification of the F fragment from *Veillonella parvula*, in the multiplex-PCR peak 4 could be assigned to the F fragment, which seemed to be amplified. In contrast, the amplification of fragment H was weak. In *Atopobium minutum*, the DE fragment was not seen in the multiplex-PCR like in single amplification.

Thus, a multiplexing strategy could not be developed for the detection of intestinal bacteria and pathogens.

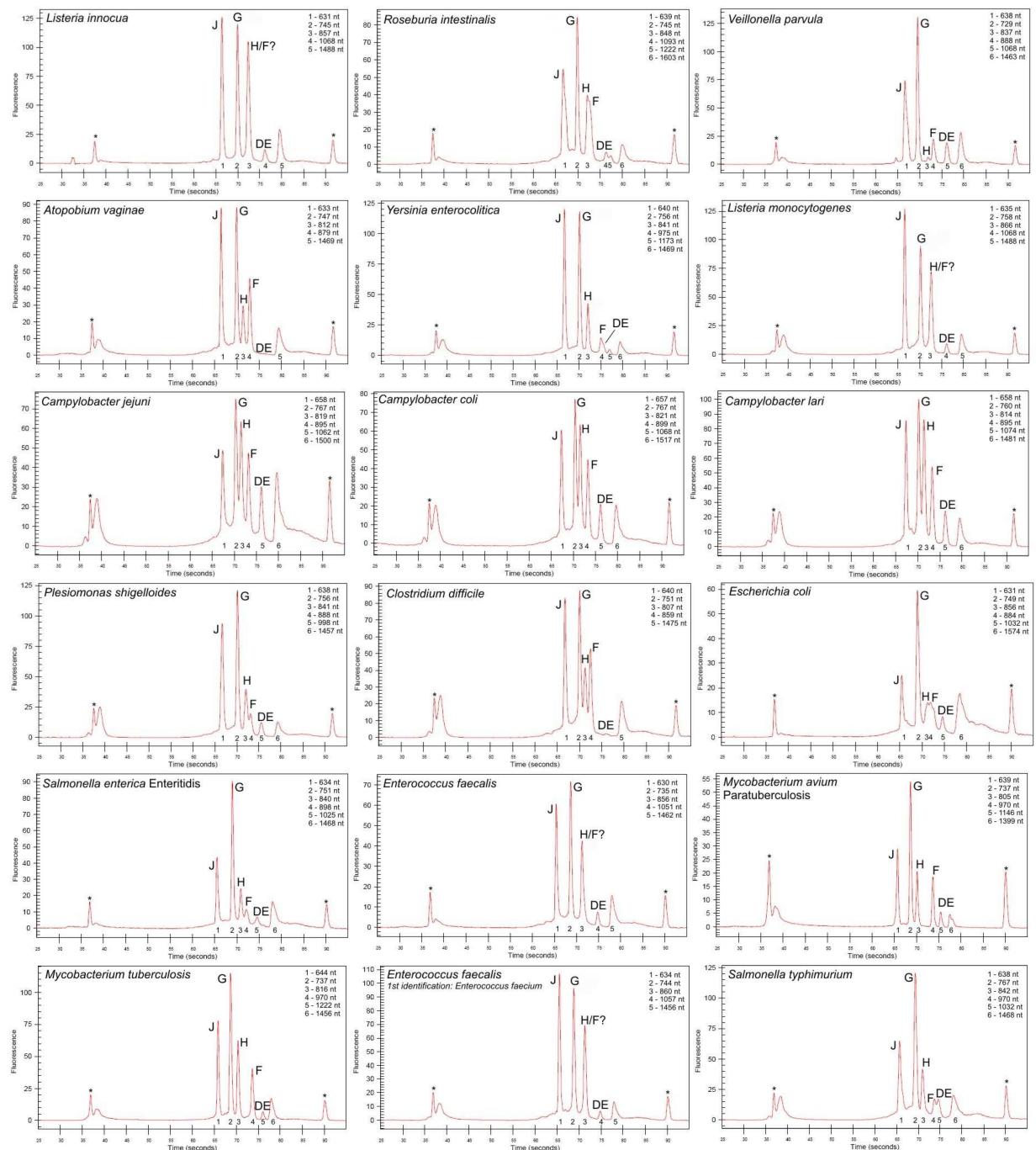


Fig. 3.6: continued on next page

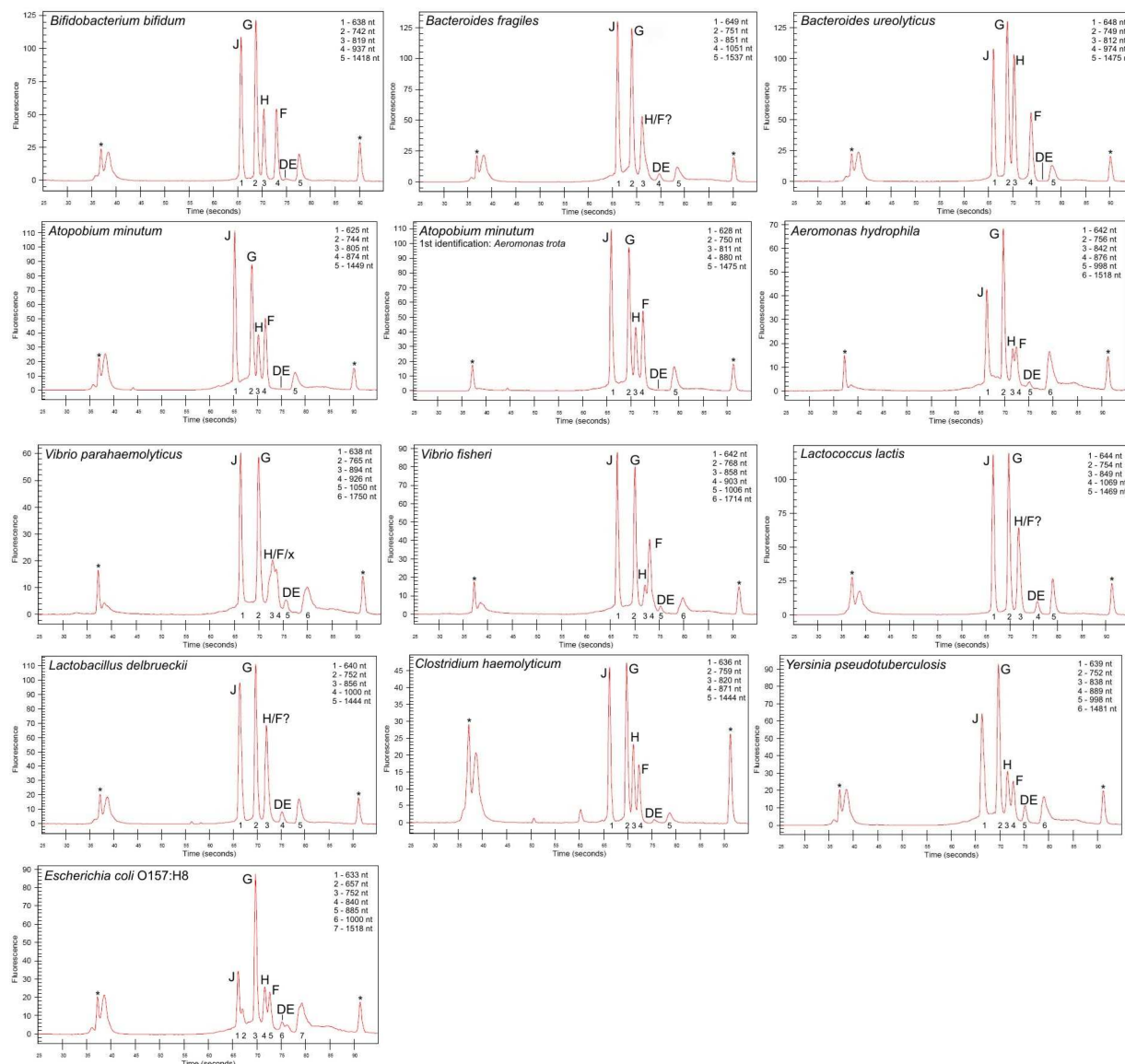


Fig. 3.6: Capillary gel electrophoresis (Agilent Bioanalyzer with DNA7500 kit) of multiplex-PCR products of bacterial reference DNA isolates with five primer pairs amplifying the ribosomal genes in five fragments H (794 nt), J (616 nt), DE (1,000 nt), F (850 nt) and G (736 nt). Average fragment length according to *E. coli*. (*) lower and upper marker of 50 and 10,380 kb.

3.5 Influence of DNA digestion on hybridization

By means of four reference strains, *E. coli*, *R. intestinalis*, *P. shigelloides*, and *C. coli*, it was exemplarily shown that a digestion of the PCR products prior to hybridization leads in most probes to higher absolute fluorescence signals compared to non-digested DNA (Fig. 3.7 and Tab. 3.7). The ribosomal genes of isolated DNA from these four reference strains were amplified in five fragments (2.2.4.3). Digestion and hybridization of 200 ng equimolarly mixed PCR products were performed according to 2.2.5-2.2.7. The same PCR products were also hybridized without prior digestion. Data acquisition and processing followed protocol 2.2.9, but no cut-off was applied to the fluorescence data. Each sample was hybridized to two arrays and standard deviations belong to six spots. All in all, 26 probes were analyzed of which only three displayed lower signals, nine similar or slightly higher signals and fourteen showed a signal increase between 3 and 28 times when hybridized with digested complementary DNA. In average, the signal intensity was 1.8 times higher, when digested

instead of undigested DNA was hybridized. Therefore, for future experiments the DNA digestion was always undertaken before hybridization.

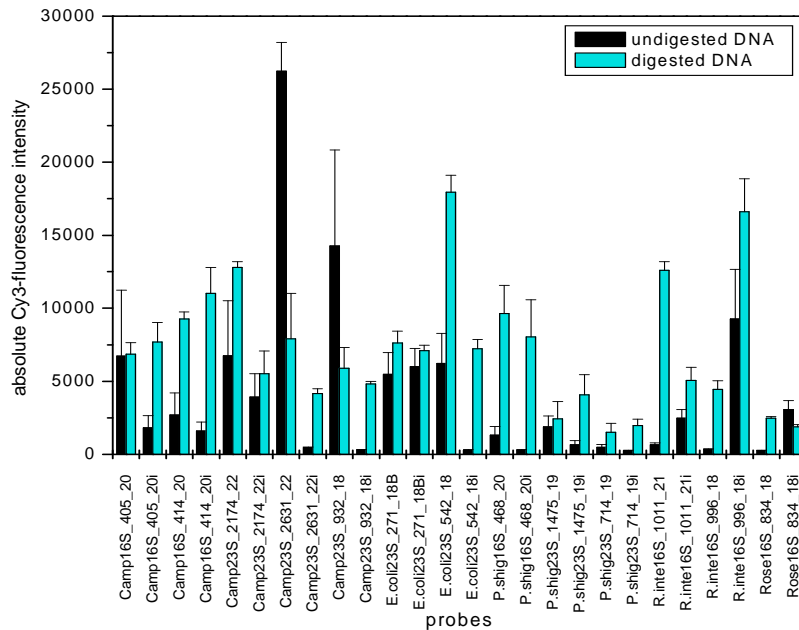


Fig. 3.7: Signal response at PMT 65 of 26 specific probes upon hybridization of 200 ng undigested and DNase I-digested (0.4 mU/ng DNA) PCR-product of *R. intestinalis*, *C. coli*, *E. coli* and *P. shigelloides*. All values are average values of six spots from two array hybridizations.

Tab. 3.7: Signal response of 26 specific probes at PMT 65 upon hybridization of undigested and DNase I-digested (0.4 mU/ng DNA) PCR-product of *R. intestinalis*, *C. coli*, *E. coli* and *P. shigelloides* and their relative position in the target complementary strand. Data corresponding to Fig. 3.7.

| Probe name | FU _{abs} undigested ^{a)} | FU _{abs} digested ^{a)} | Relative position ^{b)} | FUabs dig/ FUabs undig ^{c)} |
|--------------------|---|---|------------------------------------|---|
| Camp16S_405_20 | 6731 | 6859 | 50 | 1.0 |
| Camp16S_405_20i | 1814 | 7680 | 50 | 4.2 |
| Camp16S_414_20 | 2689 | 9271 | 51 | 3.4 |
| Camp16S_414_20i | 1589 | 11017 | 49 | 6.9 |
| Camp23S_2174_22 | 6742 | 12800 | 33 | 1.9 |
| Camp23S_2174_22i | 3913 | 5521 | 67 | 1.4 |
| Camp23S_2631_22 | 26232 | 7906 | 95 | 0.3 |
| Camp23S_2631_22i | 467 | 4162 | 5 | 8.9 |
| Camp23S_932_18 | 14258 | 5898 | 82 | 0.4 |
| Camp23S_932_18i | 309 | 4816 | 18 | 15.6 |
| E.coli23S_271_18B | 5479 | 7621 | 16 | 1.4 |
| E.coli23S_271_18Bi | 6007 | 7087 | 84 | 1.2 |
| E.coli23S_542_18 | 6226 | 17951 | 43 | 2.9 |
| E.coli23S_542_18i | 285 | 7222 | 57 | 25.3 |
| P.shig16S_468_20 | 1306 | 9635 | 60 | 7.4 |
| P.shig16S_468_20i | 284 | 8031 | 40 | 28.3 |
| P.shig23S_1475_19 | 1884 | 2439 | 45 | 1.3 |
| P.shig23S_1475_19i | 669 | 4081 | 55 | 6.1 |
| P.shig23S_714_19 | 486 | 1510 | 60 | 3.1 |
| P.shig23S_714_19i | 267 | 1958 | 40 | 7.3 |
| R.inte16S_1011_21 | 652 | 12591 | 36 | 19.3 |
| R.inte16S_1011_21i | 2463 | 5066 | 64 | 2.1 |
| R.inte16S_996_18 | 345 | 4452 | 33 | 12.9 |
| R.inte16S_996_18i | 9268 | 16610 | 67 | 1.8 |
| Rose16S_834_18 | 260 | 2463 | 7 | 9.5 |
| Rose16S_834_18i | 3070 | 1882 | 93 | 0.6 |
| average value | 3988 | 7174 | | 1.8 |

a) absolute fluorescence units at PMT65 after hybridization with 200 ng target corresponding to the probe

b) relative position of the probe in the target strand complementary expressed as percentage of the 5'→3' distance

c) signal change factor expressed as division of hybridization signals from digested by undigested PCR product

To optimize the DNA digestion an *E. coli* multiplex-PCR product was digested with different amounts of DNase I of 0.01, 0.05, 0.1, 0.3, 0.5 and 0.7 mU/ng DNA according to protocol 2.2.6. The results determined by capillary gel electrophoresis (Agilent Bioanalyzer) for 0.1, 0.3 and 0.5 mU/ng are shown in Fig. 3.8. Digestion with 0.1 mU/ng or lower amounts (data for lower amounts not shown) of DNase I was incomplete, whereas the fragment sizes were equally distributed between 100 and 1,000 nucleotides after digestion with 0.3 mU/ng. After digestion with 0.5 mU/ng, a fragment accumulation between 50 and 100 bp was observed. With higher DNase I concentration there was a fragment distribution between <15 kb and 100 kb observed (not shown).

Thus, the DNase I digestion of the DNA targets before hybridization was ascertained mandatory and the applicable DNase I-concentration was 0.4 mU/ng DNA.

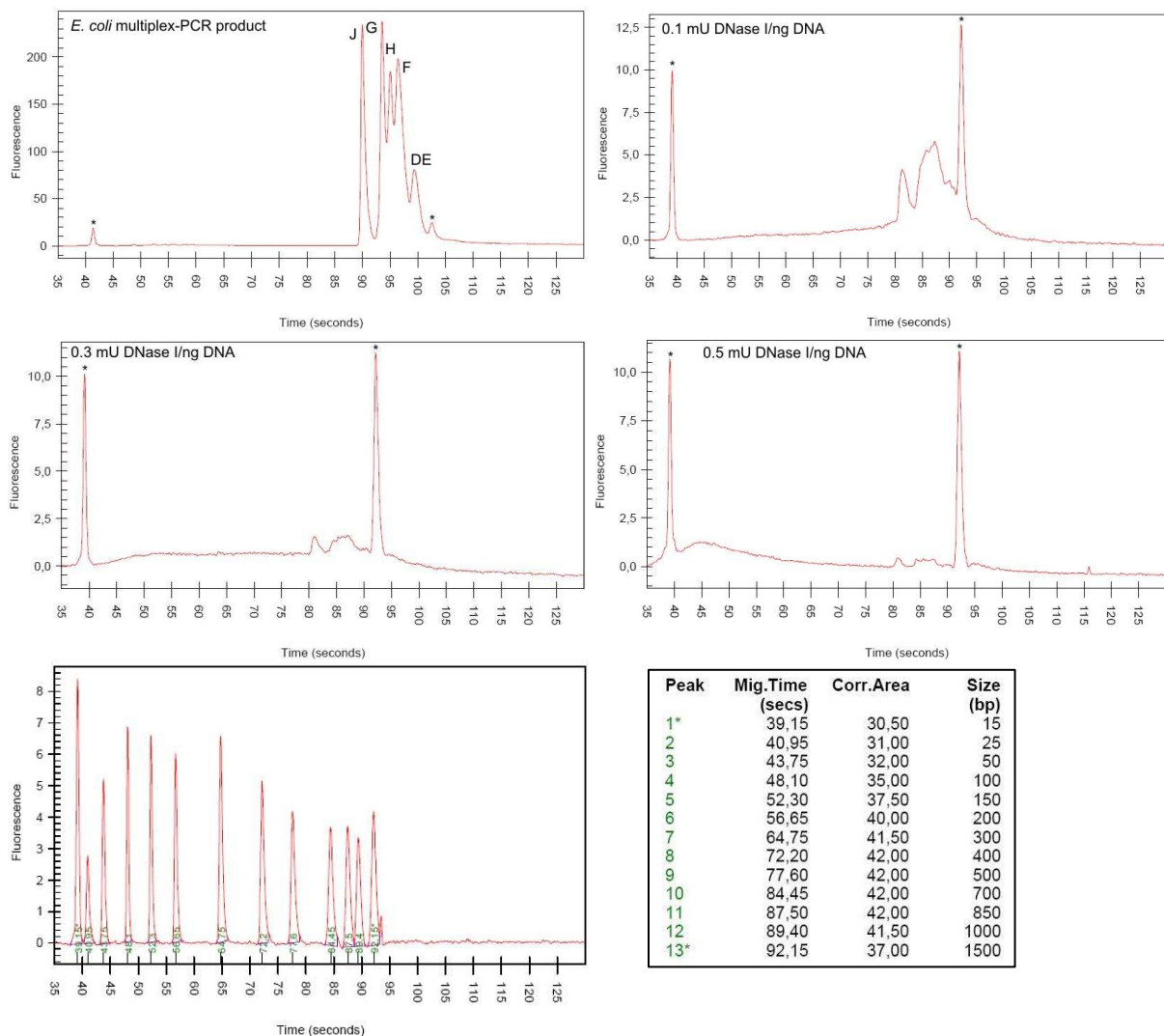


Fig. 3.8: Digestion of a multiplex-PCR product from *E. coli* (Fig. 3.5 (B)) with three different amounts of DNase I. Capillary gel electrophoresis with Agilent Bioanalyzer and DNA1000 kit. (*) lower and upper marker of 15 and 1,500 kb. (bottom) Ladder with corresponding migration time and size.

3.6 Influence of relative probe position in the target on hybridization

To analyze the influence of the probe position within the target strand, hybridization experiments of four reference strains, *E. coli*, *R. intestinalis*, *P. shigelloides* and *C. coli*, were evaluated. In all cases 200 ng PCR product (chap. 2.2.4.3) were hybridized (chap. 2.2.7)

after DNase I digestion with 0.4 mU/ng DNA (chap. 2.2.5 and 2.2.6) or without prior DNase I digestion. Only the response signals of probes corresponding with the hybridized species were taken into account and all signals were quantified at a scan enhancement of PMT 65. The relative position of each probe in its target strand was calculated using the absolute length of the target fragment H, J, DE, F or G and the starting nucleotide of the probe and was expressed as percentage of the target complementary strand. This means, a sense probe located at the 5' end of the target complementary strand, which is the sense strand, will hybridize with the 3' end of the target strand, which is the antisense strand. The position of all antisense probes was also expressed as percentage of the target complementary strand. The corresponding antisense probe to a 5' located sense probe is, therefore, located at the 3' end of the target complementary strand, which is the antisense strand, and will hybridize with the 5' end of the target strand, which is the sense strand. Fig. 3.9 shows the theoretical conformation of probe-target hybrids after hybridization to the array. A 5'-surface immobilized probe located at the 5' end of the target complementary strand, would hybridize with the target so that its short 3' end is directed towards the array surface leaving a long 5' overhang directed towards the solution. For the opposite case, steric interference of a target's long 3' overhang with the array surface was expected. In a probe pair with sense and antisense probe, one displays the first and one the second conformation. For probe pairs directed towards the middle of a target strand, the surface directed target overhang has similar length for sense and antisense probe. In case of digested PCR product used as the target, the length of the target overhangs should be clearly reduced.

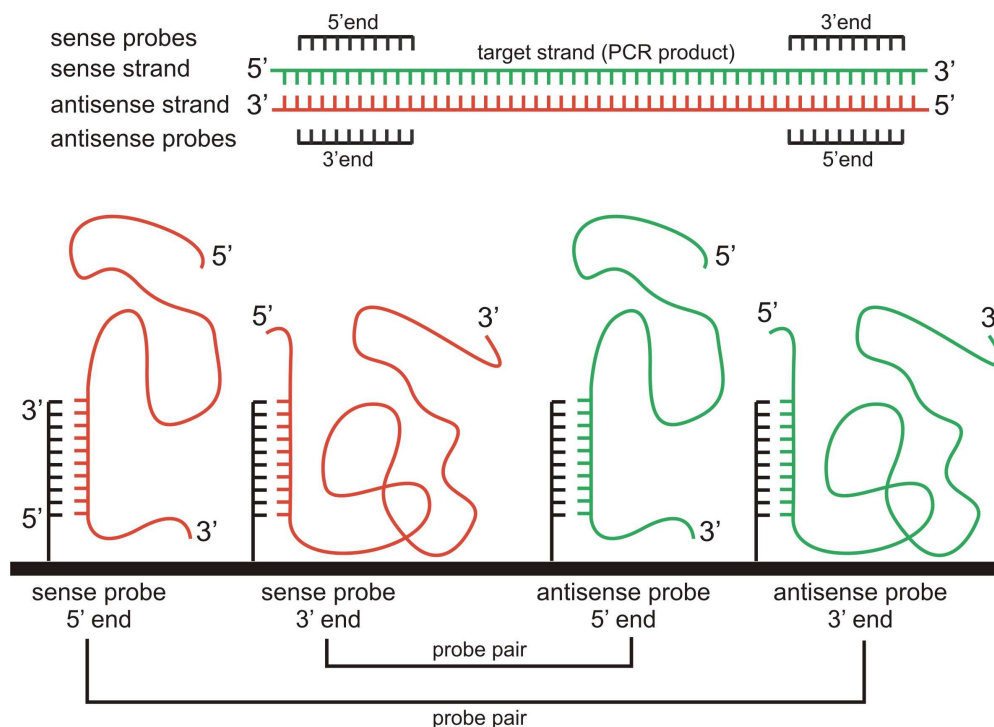


Fig. 3.9: Hybridization conformation depending on the relative position of a probe towards its target strand. For 5' located probes binding the 3' end of the target, the main target overhang is directed towards the hybridization solution, while for 3' located probes binding the 5' end of the target the overhang directs towards the array surface.

The analysis of hybridization results from 26 probes targeting four species revealed that the relative position of the probes within the DNA target strands had no significant influence on the hybridization efficiency (Fig. 3.10). Only two probes, Camp23S_932_18 and Camp23S_2631_22, located at 82 and 95% of the relative target length (5'→3') have slightly higher hybridization signals of 14,258 and 26,232 FU compared to the other probes (260-9,268 FU) when the PCR product is hybridized undigested (Tab. 3.7). These two probes

together with Rose16S_834_18i at position 93% were the only probes, which showed reduced fluorescence intensity after hybridization with digested PCR product (Fig. 3.10 (B)). In some cases probes targeting the middle region of a target strand seemed to display slightly higher fluorescence signals than end located probes after target strand digestion. However, there was no significant behaviour observable.

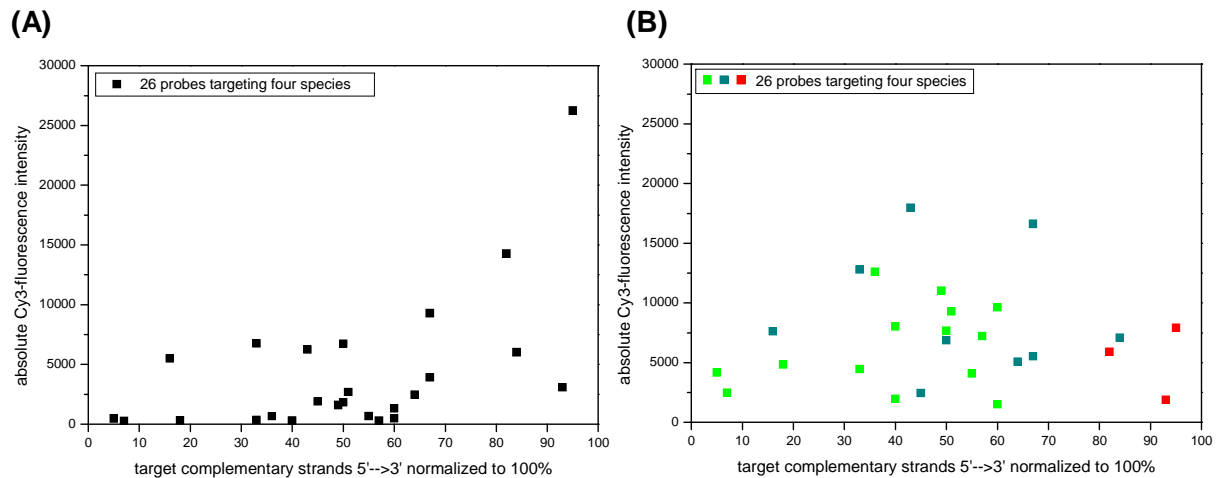


Fig. 3.10: Signal response at PMT 65 of 26 specific probes upon hybridization of 200 ng (A) undigested and (B) DNase I-digested (0.4 mU/ng DNA) PCR-product of *R. intestinalis*, *C. coli*, *E. coli* and *P. shigelloides* depending on the probe position within the different target complementary strands normalized to 100% of length. Colours in (B) indicate relative intensities in comparison of (B) with (A): (green) probes with signal increase >3 times, (cyan) probes with signal increase between 0 and 3 times, (red) probes with signal decrease.

The relative signal intensity of the sense probe vs. the corresponding antisense probe depending on the relative distance to the target middle was analyzed as well (Fig. 3.11). Only the 26 probes, which were hybridized with undigested PCR product, were investigated. No clear correlation between the distance from the target middle and the difference in the signal intensity of sense and antisense probe was observed. Sense probes close to the 3' end of the target displayed higher signals than their antisense counterparts, while sense probes at the 5' end displayed similar or lower signals than the corresponding antisense probe. Probes located in the middle displayed, with one exception, similar signal intensity.

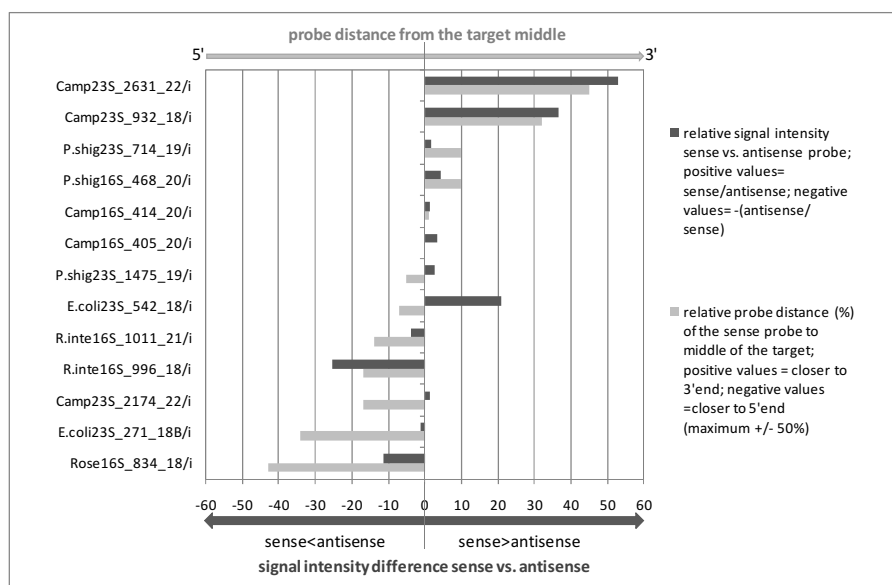


Fig. 3.11: Analysis of the relative signal intensity of the sense probe vs. the corresponding antisense probe depending on the relative distance to the target strand middle.

3.7 Probe verification with reference strains

The functionality of the probes for the specific detection of bacteria was verified using 32 reference strains (chap. 2.1.3) including target and non-target species. For chip validation, the five fragments H, J, DE, F, and G were separately amplified and Cy3-labelled (chap. 2.2.4.3 and 2.2.5). The purified fragments were equimolarly mixed to make fluorescence intensities from different probes comparable independent from amplification efficiency of the belonging fragments. Each verification experiment comprised the hybridization of two independent microarrays with each 200 ng DNase I-digested PCR product (chap. 2.2.6, 2.2.7). Data acquisition and processing followed protocol 2.2.9.

The complete assay accounted for about 5:20 h after DNA extraction from the sample material (0). The time-limiting steps were the target amplification and the hybridization.

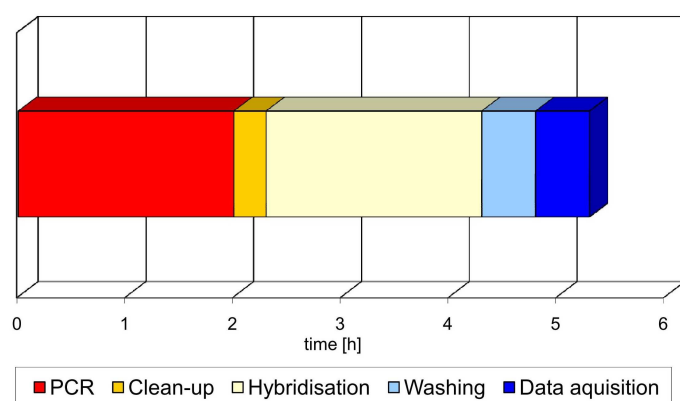


Fig. 3.12: Required time for each step of species identification with the Gastroenteritis-Chip.

Each species showed an individual hybridization pattern with its specific probes, which enables identification of the bacteria (Tab. 3.8). However, despite the stringent conditions for probe design including a comprehensive database of ribosomal genes some probes showed minor or even intense cross-hybridizations.

Plesiomonas shigelloides was correctly identified by two sense/antisense probe pairs. P.shig23S_714_19/i cross-hybridized with *E. coli*, *Yersinia*, *Salmonella*, and *B. fragilis*. The hybridization signals were comparable or even higher than those derived from hybridization with *Plesiomonas shigelloides*. As the two other probe positions showed reliable species identification, P.shig23S_714_19/i was dispensable for species identification.

The four genus probes allowed reliable detection of mycobacteria. Only minor cross-reactivity was found for probe Mycb16S_987_20i. In contrast, the probes for the species level were less specific. M.aviC23S_2443_18/i nearly behaved like a universal probe and, additionally, hybridized with MAP and *M. tuberculosis*. M.aviP23S_1495_19/i showed weak cross-hybridization with *M. tuberculosis*. Therefore, only one probe position could discriminate between *M. avium* complex (MAC) and *M. avium paratuberculosis* (MAP). These probes differ in one nucleotide (C/T) only, which is located in the exact middle of the probes. Unfortunately, for MAC no reference DNA was available. The successful hybridization of MAP with its probes M.aviP23S_650_17/i without cross-hybridization with M.aviC23S_650_17/i can only indicate that MAC will behave oppositely enabling identification of both species.

Two of the genus probes for Listeria, List16S_1151_19/i, displayed weak cross-hybridizations with many other species. Although the relative values were not high, such a probe is not stressable under real sample conditions. It is likely that other and even more

intense cross-hybridizations with non-target species occur. The second genus-probe set for *Listeria* only cross-hybridized with MAP in the sense version and can be used for identification. *Listeria monocytogenes* should be identified by two species-level probes (sense and antisense). However, the sense probe heavily cross-hybridized with *Listeria innocua*, *Clostridium difficile*, and *Lactococcus lactis* and the antisense probe less intensely with the former two species. As no other probe for this species could fill the gap, a redesign of this probe was requisite. The large difference in signal intensity between the correct positive and the wrong positives in combination with the probe length held the potential for a successful redesign.

The EHEC/Salm23S_1164_18/i hybridized with both, EHEC and Salmonella enteritidis, and very weakly with *E. coli*. An identification of EHEC was only possible in the absence of *Salmonella enteritidis*. *Salmonella typhimurium*, on the other hand, did not hybridize with these probes. Therefore, the EHEC/Salm23S_1164_18/i probes may distinguish between *S. enteritidis* and *S. typhimurium*, which were reliably detected by two more probe positions for the genus. However, in case of the presence of *S. typhimurium* and EHEC together this may lead to misidentification. The antisense probe Salm23S_274_20i weakly cross-hybridized with EHEC, *E. coli*, and *Yersinia* spp., and was dispensable for the array.

The Shigella probes all displayed strong cross-hybridization with EHEC and *E. coli*, some with *P. shigelloides*, *Salmonella*, *Yersinia*, *A. hydrophila*, and *V. parahaemolyticus*. *Shigella* spp. could not be identified by the present array.

The genus probes for Aeromonas hybridized with *A. hydrophila* but, unexpectedly, not with sample 23 (*Aeromonas trota*). Cross-hybridizations were not observed. Additionally, the specific probes for *Aeromonas hydrophila*, A.hyd23S_1170_18/i, displayed no hybridization signals with the reference DNA of *A. hydrophila*. Sequencing should reveal the reason. In case of sample 23, the hybridization result was in conflict with the prior clinical identification or the labelling of the reference strain. The sample displayed positive hybridization signals with *Atopobium minutum* genus- and species-specific probes. An *Atopobium minutum* sample (22) showed the same hybridization pattern. Sequencing had to reveal whether sample 23 was really *Aeromonas trota* or *Atopobium minutum*, instead.

The Atopobium minutum species-specific probes did not hybridize with both samples (22 and 23). Only the less specific A.min.fos16S_838_18/i probes resulted in hybridization signals. The A.min.fos16S_838_18/i probes could not distinguish between *A. minutum* and *A. fossor* but did not bind to other targets. The *Atopobium* genus probes, however, hybridized with sample 23 (*Aeromonas trota*), *A. minutum* (sample 22) and *A. vaginae* and cross-hybridized only very weakly with *E. faecium* (sample 17B). Interestingly, sample 17A, which should also contain *E. faecium*, did not show this cross-hybridization. This, together with the fact that 17A hybridized with all *E. faecalis*-probes but 17B did not, led to the assumption that here another conflict between clinical identification/labelling and the array result had to be solved by sequencing.

The performance of the Enterococcus probes was not satisfying. Entc16S_1281_22/i and Entc23S_2618_18/i cross-hybridized with several species, of which *C. difficile*, *Listeria*, *P. shigelloides*, *L. lactis*, and *L. delbrueckii* showed highest signals. Although the signal intensity was not comparable to hybridization with *Enterococcus*, it was too high for a successful redesign. The *E. faecalis* species probes displayed only few, weak cross-hybridizations. Sense probe E.faec23S_1172_20, however, hybridized with *E. faecium* (sample 17B), while the corresponding antisense probe displayed only a weak signal with *E. faecium*. If sample 17A was *E. faecalis* instead of *E. faecium*, the specific probes for *E. faecalis* were suitable to identify this species.

The Vibrio sense probe recognized many other species, while the antisense probe was specific for the genus. The *Vibrio cholerae* probes could only be tested for their cross-reactivity. Due to a lack of reference material, the probes could not be verified with pure

culture DNA. Three probes presented cross-hybridization with several other species and were not suitable for the identification array. Three *V. cholerae*-probes also detected *V. parahaemolyticus* and *V. fischeri* and can be regarded as genus-probes. Four probes did not hybridize with any reference DNA and are supposed to be specific for *V. cholerae*. The high coverage of this species with different probes on several taxonomic levels may guaranty a correct identification of this species without prior verification of the probes with reference DNA. *Vibrio parahaemolyticus* could be detected by six specific probes. These probes were not verified with MAP, *L. innocua*, and *C. lari*, but their close relatives did not show cross-reactivity with them. Additionally, two probes covering *V. parahaemolyticus* and its close relatives detected this species together with *V. fischeri*.

For the genus *Campylobacter* ten probes were designed, which exhibited only few, weak cross-hybridizations with some other bacteria, except for *Bacteroides ureolyticus*. This species heavily cross-hybridized with probes Camp16S_405_20/i and Camp16S_414_20/i presenting comparable signal intensity to *Campylobacter*. These probes cannot be used for detection of *Campylobacter*. In addition, probe Camp23S_2631_22i displayed weak cross-reactivity with *B. ureolyticus* and *Y. pseudotuberculosis* and should, therefore, be excluded from the array. Despite minor cross-hybridizations of the five residual probes, the 50% positive-concept allows reliable detection of *Campylobacter*. Unfortunately, the two *C. jejuni* probes on the array were not specific for *C. jejuni*. *Campylobacter coli* displayed even higher hybridization signals and *C. lari* was detected by these probes as well. For the identification of *C. jejuni* it was necessary to find new probes. Moreover, the four *C. coli*-probes did not detect *C. coli* but weakly cross-hybridized with the other *Campylobacter* species. Here, the hybridization result also indicated a conflict between the clinical identification result and the microarray result, which had to be solved by sequencing. In this case, however, single sequences for *C. coli* were found in the 16S-database, which differed completely from most other *C. coli* sequences but did not equal *C. jejuni*.

Clostridium difficile could be detected on the species level with high specificity. Only one minor cross-hybridization with *R. intestinalis* was observed for one of the four probes. The genus probe did not cover *C. difficile* but other clostridia. Though the genus probe matched *C. haemolyticum*, it also displayed cross-reactivity with *R. intestinalis*, *B. fragilis*, enterococci, and *L. lactis*, and had to be redesigned or replaced.

Yersinia enterocolitica could not be specifically detected on the genus level. The two genus probes also reacted on *E. coli* including EHEC and *Salmonella*. Although, the signal intensity was only around 10% of the correct positive, these probes were not suitable for pathogen detection, because the 50% positive-concept will not prevent wrong identification and *E. coli* is a natural inhabitant of the intestine. The *Y. enterocolitica* species-specific probes were also tested for cross-reaction with the close relative *Y. pseudotuberculosis*. Three probes also showed intense signals upon hybridization with *Y. pseudotuberculosis*. Only one probe, Y.ent23S_1501_20i, really was specific for *Y. enterocolitica*, thus identification of this species requires signals from all probes.

The probes for *Eubacterium bifforme* could not be verified with reference material, but cross-reactivity was not found with all other tested reference species.

E. coli could be successfully identified with the present probe set. Nonetheless, probes E.coli23S_271_18B/i did not discriminate between pathogenic and non-pathogenic *E. coli* and detected *Salmonella*, a hybridization signal from more than 50% of the probes identifies non-pathogenic *E. coli*.

Roseburia intestinalis, *Bacteroides fragilis*, and *Veillonella parvula* were perfectly identified on genus level and the former two additionally on species level by the respective probes. Only one *B. fragilis* probe showed a weak cross-hybridization with *B. bifidum*. *Bifidobacterium bifidum* could be clearly identified with eight species-specific probes.

For *Acholeplasma laidlawii*, no reference DNA was available, thus the probes were only analyzed for cross-reactivity. Two of the four probes cross-hybridized with *C. haemolyticum* and weakly with many other species. The two remaining probes may be capable of detecting specifically *A. laidlawii*.

The two *Fusobacterium prausnitzii* sense probes did not show unspecific signals with other bacteria, but due to a lack of reference material, they could not be positively verified. The antisense probes also detected mycobacteria, *R. intestinalis*, *A. hydrophila*, and slightly recognized other bacteria as well. Thus, only the sense probes were useful for species identification.

Lactococcus lactis was identified by six probes. Two probes showed minor cross-hybridization with *C. lari*, while one probe heavily hybridized with *B. fragilis*. Therefore, the latter probe was regarded dispensable for the array.

For *Lactobacillus acidophilus* and *Lactobacillus delbrueckii*, only two probes each were present on the array. Not any displayed cross-hybridization, but only those probes for *L. delbrueckii* could be verified with the appropriate reference DNA. The *L. acidophilus* probes were not checked against MAP and *C. lari* but their close relatives.

3.8 Sanger sequencing of the reference strains

Sanger sequencing of the ribosomal genes (chap. 2.2.11) was used to verify the microarray hybridization results, to explain certain cross-hybridizations, and enable probe redesign. Tab. 3.9 lists all reference strains with the length of their sequenced ribosomal genes. The 16S gene was successfully sequenced and aligned for all reference strains besides *Mycobacterium avium* subsp. *paratuberculosis* (MAP). In case of MAP, sequencing gave satisfactory results only for fragment J. The longer 23S gene was successfully sequenced in full length from 25 reference strains. In case of *Y. enterocolitica* and *C. haemolyticum*, the first part of the DE fragment is missing. The whole DE fragment could not be sequenced for MAP, *A. minutum*, and *B. fragilis*. EHEC was not completely sequenced in the end of the 23S gene. The aligned sequences were submitted to GenBank (Benson 2010) under the given accession numbers.

A BLAST (basic local alignment search tool) search was performed with each sequence to identify the next relatives in the GenBank database. All results were in agreement with previous microarray identification and, therefore, confirmed the accuracy of the microarray performance. As three reference species, sample 17A, 23, and 24, showed conflicting results between microarray hybridization and previous clinical identification by culture based methods and PCR (or a wrong tube labelling), sequencing was used as the gold standard method to identify these species correctly. By sequencing, the microarray results were verified. Sample 17A was identified as *Enterococcus faecalis* instead of *Enterococcus faecium* by 16S and 23S sequences, sample 23 as *Atopobium minutum* instead of *Aeromonas trota* by 16S sequence, and sample 24 as *Aeromonas bestiarum* instead of *Aeromonas hydrophila* by 23S sequence. A revisal of the DSM numbers of sample 24 confirmed that this sample was *A. bestiarum*, while the DSM number of sample 23 belonged to *A. trota*. These results explained the hybridization pattern of the three reference isolates. This meant that the *A. hydrophila* species-specific probes were not positively verified, but showed no cross-hybridization with a closely related species.

Some species could not be clearly identified by sequencing the ribosomal genes, but the results were not contrary to the clinical identification. Sample 16, *Mycobacterium tuberculosis*, was identical with *M. tuberculosis* in GenBank but also with other mycobacteria in the 16S gene sequence and even the 23S gene could not separate between *M. tuberculosis* and *M. bovis*. *Salmonella enteritidis* (sample 13) was indistinguishable from *Salmonella paratyphi*, which is another *S. enterica* serovar, based on the 16S gene but differed clearly in the 23S gene. However, the serovar Enteritidis could not be distinguished from further *S. enterica* serovars. The 16S sequence of sample 18, containing *S. typhimurium*, had 100% identity with *S. typhimurium* but also with some other *S. enterica* serovars. The 23S sequence was identical with one *S. typhimurium* sequence, but the different *rrn* operons displayed higher sequence variability than the difference was between *S. typhimurium* and, for example, *S. heidelberg*. Therefore, *Salmonella* was detected by microarray only on the genus level. *Clostridium haemolyticum* (sample 31) could not be distinguished from *C. novyi* by 16S gene sequence. However, the 23S sequence of this sample was only 98% identical with *C. novyi*. A 23S sequence of *C. haemolyticum* was not yet available. Therefore, it can be assumed, that the isolate was indeed *C. haemolyticum*. The 16S sequence of sample 32 was identical with *Yersinia pseudotuberculosis* and *Yersinia pestis*, but the 23S gene sequence matched with some *rrn* operons of *Y. pseudotuberculosis*, while at least one mismatch was found for *Y. pestis*. This strongly supports the clinical identification as *Y. pseudotuberculosis* isolate. The 16S and 23S sequences of sample 33 displayed only 99% identity with *E. coli* O157:H7, but also with other *E. coli* strains and *Shigella*. However, most operons of *E. coli* O157:H7 matched perfectly with both gene sequences and only few had sequence variations. Regarding the 16S gene, also other *E. coli* strains matched the sequence, which was not the case for the 23S gene. Here, other strains had at least one mismatch towards the query sequence. Sample 33 differed in both genes from the sequences of *Shigella* species, which possessed two or more mismatches towards the 16S sequence and five or

more mismatches towards the 23S sequence. Despite the operon sequence variation of *E. coli*, the result supports the clinical identification result.

Tab. 3.9: Reference strains with length and accession numbers of the sequenced 16S and 23S rRNA genes. A BLAST search was performed to identify the next relative of each sequence. (continued on next page)

| Nb. | Species | Length (nt) | Acc. Nb. | 16S rRNA gene | | 23S rRNA gene | | BLAST relative | next |
|-----|--|-------------|----------|---|------------|---------------|---|----------------|------|
| | | | | BLAST relative | next | Length (nt) | Acc. Nb. | | |
| 1 | <i>L. innocua</i> | 1370 | HM007562 | 100% <i>L. innocua</i> | 2541 | HM007626 | 99% <i>L. innocua</i> / 4 other Listeria | | |
| 2 | <i>R. intestinalis</i> | 1329 | HM007565 | 99% <i>R. intestinalis</i> | 2482 | HM007596 | 93% <i>Eubacterium rectale</i> (no <i>R. intestinalis</i> sequenced) | | |
| 3 | <i>V. parvula</i> | 1345 | HM007566 | 99% <i>V. parvula</i> | 2531 | HM007597 | 100% <i>Desulfotomaculum reducens</i> / 99% <i>V. parvula</i> | | |
| 4 | <i>A. vaginae</i> | 1331 | HM007563 | 100% <i>A. vaginae</i> | 2565 | HM007594 | 93% <i>Desulfotomaculum acetoxidans</i> (no <i>A. vaginae</i> sequenced) | | |
| 5 | <i>Y. enterocolitica</i> | 1349 | HM007567 | 100% <i>Y. enterocolitica</i> | 2254 (EFG) | HM007598 | 99% <i>Y. enterocolitica</i> / other <i>Yersinia</i> and <i>Edwardsiella</i> | | |
| 6A | <i>L. monocytogenes</i> | 1368 | HM007564 | 99% <i>L. monocytogenes</i> (0 MM)/ other listeria (>4 MM) | 2504 | HM007595 | 99% <i>L. monocytogenes</i> (0 MM)/ other listeria (>2 MM) | | |
| 7 | <i>C. jejuni</i> | 1327 | HM007568 | 99% <i>C. jejuni</i> (1 MM)/ 99% <i>C. coli</i> (4 MM) | 2516 | HM007599 | 100% <i>C. jejuni</i> | | |
| 8 | <i>C. coli</i> | 1342 | HM007569 | 99% <i>C. coli</i> (1-11 MM)/ 98% <i>C. jejuni</i> (≥ 20 MM) | 2555 | HM007600 | 99% <i>C. coli</i> (≥ 0 MM)/ <i>C. jejuni</i> (≥ 1 MM) | | |
| 9 | <i>C. lari</i> | 1338 | HM007571 | 100% <i>C. lari</i> | 2482 | HM007601 | 99% <i>C. lari</i> | | |
| 10 | <i>P. shigelloides</i> | 1348 | HM007572 | 99% <i>P. shigelloides</i> | 2459 | HM007602 | 98% <i>P. shigelloides</i> | | |
| 11 | <i>C. difficile</i> | 1303 | HM007570 | 99% <i>C. difficile</i> | 2509 | HM007603 | 99% <i>C. difficile</i> | | |
| 12 | <i>E. coli DH5α</i> | 1333 | HM007573 | 99% <i>E. coli</i> (0 MM)/ <i>Shigella dysenteriae</i> (0 MM) | 2470 | HM007604 | 99% <i>E. coli</i> (0 MM)/ <i>Shigella dysenteriae</i> (2 MM) | | |
| 13 | <i>S. enteritidis</i> | 1332 | HM007577 | 100% <i>S. enteritidis</i> / other <i>S. enterica</i> serovars/ <i>S. paratyphi</i> | 2490 | HM007608 | 99% <i>S. enteritidis</i> / other <i>S. enterica</i> serovars (1-3 MM)/ <i>S. paratyphi</i> (14 MM) | | |
| 14 | <i>E. faecalis</i> | 1378 | HM007574 | 100% <i>E. faecalis</i> | 2479 | HM007605 | 99% <i>E. faecalis</i> | | |
| 15 | <i>M. avium</i> subsp. <i>paratuberculosis</i> | 595 (J) | HM007575 | 100% <i>M. avium</i> subsp. <i>paratuberculosis</i> / 4 other mycobacteria | 1699 (FG) | HM007606 | 99% <i>M. avium</i> subsp. <i>paratuberculosis</i> | | |
| 16 | <i>M. tuberculosis</i> | 1346 | HM007576 | 100% <i>M. tuberculosis</i> / 3 other mycobacteria | 2700 | HM007607 | 100% <i>M. tuberculosis</i> / <i>M. bovis</i> | | |
| 17A | <i>E. faecium</i> (*) | 1374 | HM007580 | 100% <i>E. faecalis</i> | 2493 | HM007610 | 99% <i>E. faecalis</i> | | |
| 17B | <i>E. faecium</i> | 1374 | HM007579 | 100% <i>E. faecium</i> | 2507 | HM007611 | 99% <i>E. faecium</i> | | |
| 18 | <i>S. typhimurium</i> | 1365 | HM007581 | 100% <i>S. typhimurium</i> / other <i>S. enterica</i> serovars | 2572 | HM007612 | 99% <i>S. typhimurium</i> (0-5) MM/ other <i>S. enterica</i> serovars (≥ 1 MM) | | |

| Nb. | Species | Length (nt) | Acc. Nb. | BLAST relative | next | Length (nt) | Acc. Nb. | BLAST relative | next |
|-----|------------------------------|-------------|----------|--|------|---------------------|-----------------------|--|------|
| 19 | <i>B. bifidum</i> | 1344 | HM007578 | 100% <i>B. bifidum</i> | | 2558 | HM007609 | 91% <i>B. animalis</i> (no <i>B. bifidum</i> sequenced) | |
| 20 | <i>B. fragilis</i> | 1335 | HM007585 | 99% <i>B. fragilis</i> | | 805 (F), 704 (G) | HM007616, HM007617 | 99% <i>B. fragilis</i> | |
| 21 | <i>B. ureolyticus</i> | 1312 | HM007586 | 99% <i>B. ureolyticus</i> | | 2610 | HM007618 | 99% <i>B. ureolyticus</i> | |
| 22 | <i>A. minutum</i> | 1345 | HM007583 | 99% <i>A. minutum</i> | | 1529 (FG) | HM007614 | 97% <i>Desulfotomaculum reducens</i> (no <i>A. minutum</i> sequenced) | |
| 23 | <i>A. trota</i> (*) | 1340 | HM007584 | 100% <i>A. minutum</i> | | 1525 (FG) | HM007615 | 97% <i>Desulfotomaculum reducens</i> (no <i>A. minutum</i> sequenced) | |
| 24 | <i>A. hydrophila</i> (*) | 1375 | HM007582 | 100% <i>A. bestiarum</i> / <i>A. hydrophila</i> / <i>A. salmonicida</i> | | 2453 | HM007613 | 99% <i>A. bestiarum</i> (2-3 MM)/ <i>A. hydrophila</i> (≥8 MM)/ other aeromonads (≥2 MM) | |
| 25 | <i>V. parahaemolyticus</i> | 1351 | HM007587 | 99% <i>V. parahaemolyticus</i> / 3 other <i>Vibrio</i> | | 2472 | HM007619 | 99% <i>V. parahaemolyticus</i> | |
| 26 | <i>V. fischeri</i> | 1346 | HM007592 | 99% <i>V. fischeri</i> | | 2402 | HM007624 | 100% <i>V. fischeri</i> | |
| 29 | <i>L. lactis</i> | 1344 | HM007591 | 100% <i>L. lactis</i> | | 2476 | HM007623 | 100% <i>L. lactis</i> | |
| 30 | <i>L. delbrueckii</i> | 1365 | HM007590 | 99% <i>L. delbrueckii</i> | | 2455 | HM007622 | 99% <i>L. delbrueckii</i> | |
| 31 | <i>C. haemolyticum</i> | 1339 | HM007588 | 100% <i>C. haemolyticum</i> / <i>C. novyi</i> | | 1990 (EFG) | HM007620 | 98% <i>C. novyi</i> (no <i>C. haemolyticum</i> sequenced) | |
| 32 | <i>Y. pseudotuberculosis</i> | 1356 | HM007593 | 100% <i>Y. pseudotuberculosis</i> / <i>Y. pestis</i> | | 2491 | HM007625 | 99% <i>Y. pseudotuberculosis</i> (0-5 MM)/ <i>Y. pestis</i> (1-4 MM) | |
| 33 | <i>E. coli</i> O157:H7 | 1353 | HM007589 | 99% <i>E. coli</i> O157:H7 (0-7 MM)/ other <i>E. coli</i> (≥0 MM)/ <i>E. coli</i> K12 (≥1 MM)/ <i>Shigella</i> (≥2 MM) | | 2105 (DEFGv) | HM007621 | 99% <i>E. coli</i> O157:H7 (0-10 MM)/ <i>E. coli</i> (≥1 MM)/ <i>Shigella</i> (≥5 MM) | |

(*) conflict between clinical species identification/clinical labelling and identification by microarray and sequencing

(MM) mismatches between query and database sequence; only real mismatches excluding mismatches due to degenerated bases (if they can potentially match); therefore, a query can have only 99% identity with a database sequence according to BLAST, but zero real mismatches; a range of mismatches can exist due to multiple sequenced isolates in the database or due to multiple operon copies with sequence variation

All sequences were compared with the designed probes of the microarray (data not shown) to explain certain cross-hybridizations or false negative results. The analysis showed that in particular probes with only one or two mismatches towards the non-target organism displayed high cross-hybridization signals. This was the case for, M.aviC23S_2443_18/i, the *Shigella*-probes, several *V. cholerae*-probes, *Y. enterocolitica*-probes, and some *Enterococcus*-probes. However, in many cases one or two mismatches were enough to discriminate the bacteria with a probe. The relative position of the mismatch was not always the crucial factor, as it could be seen in a comparison of probes S.flex23S_1704_19/i with M.aviC23S_2443_18/i. The latter probe had one mismatch at position 10 (sense 5'→3') towards e.g. *B. fragilis* leaving a strand of eight and nine nucleotides on both sides for hybridization. Cross-hybridization of 31 and 24 RI were observed for the sense and anti-sense probe (Tab. 3.8). The *S. flexneri* probe had one mismatch at position 8 towards e.g. *Salmonella* leaving seven and eleven consecutive nucleotides for hybridization on both sides

of the mismatch. This resulted in case of the antisense probe in a cross-hybridization signal of <1 RI and for the sense probe in a signal of 9 RI towards *S. enteritidis* and a signal towards *S. typhimurium* of 17 RI. Although, both positions of the mismatch were comparable, the latter displayed lower cross-hybridization signals.

Probes with more than two mismatches towards the non-target species also hybridized with it, if the mismatches were not equally distributed over the probe length. The clostridia probe Clos23S_1038_24/i, for example, had four mismatches with the sequences of *E. faecalis*, *E. faecium*, and *L. lactis*. They were located at position 2, 5, 7, and 8 (sense 5'→3') of the probe leaving a 16 nucleotides long sequence for hybridization. *C. difficile*, which was not a target for this probe, also had four mismatches with this probe, but did not hybridize at all. The mismatches were located at position 5, 7, 8, and 16 of the probe leaving only eight consecutive nucleotides for hybridization. The most promising attempt to increase specificity of this probe was to shorten it at the 3' end. In few cases, sequencing revealed complete identity between a probe and a non-target. This was found for the 16S *Campylobacter* genus probes and *B. ureolyticus*, the *C. jejuni* probe at position 1,153 and *C. coli*, and the *E. coli* probe at position 271 and *E. coli* O157:H7. This clearly explained cross-hybridization between these probes and species. However, some cross-hybridizations could not be explained by sequencing.

All sequencing results supported the microarray hybridization results and explained cross-hybridizations, which partly enabled probe redesign.

3.9 Array redesign

The first array versions comprised probes, which cross-hybridized with other species or did not match the target DNA as described in chapter 3.7. During development of the array, the probe set was optimized by editing probes or excluding them from the array. Results from sequencing of the reference strains enabled sequence-based assessment of the probes. The array optimization was a process spanning several array versions and cannot be illustrated by a before-and-after array layout.

Specificity was the most important feature of the array due to the dominant DNA background in faecal samples, which accompanies each identification assay. At first, probes that did not exclusively match the target bacteria were excluded from the array with some exceptions. Minor cross-hybridizations below 2% MM/PM ratio were accepted. Moreover, few probes with cross-hybridizations above 2% MM/PM ratio were kept on the microarray due to the multiple-probe concept (Tab. 3.10). In total, 61 of 164 probes were not suitable for the identification of bacteria from the intestinal tract and were discarded during the verification process. This also included the *Shigella* probes, which could not be replaced by other probes. Additionally, ten probes were modified or newly designed to solve problems of specificity and sensitivity towards non-targets and the target, respectively (Tab. 3.8). In Tab. 3.2, all 174 designed probes are listed. The sequences of the final probe set are marked with an 'x'.

In case of *Listeria monocytogenes*, it was possible to reach specificity towards *L. innocua* by shortening the 27 nucleotides long probe for two nucleotides at both, the 3' and 5' end (Fig. 3.13). The 27 nucleotides long probe had two mismatches with the *L. innocua*-sequence at position 5 and 19 of the sense strand, which were still present in the 23 nt-probe at position 3 and 17. The relative hybridization signals of the shorter sense and antisense probes towards *L. monocytogenes* were reduced by 20% and 18.5%. However, the relative intensity of the cross-hybridization with *L. innocua* of the sense and antisense probe was reduced by 57% and 83%. The shorter antisense probe displayed only 1.5 RI after hybridization with 200 ng amplified target DNA from *L. innocua*, which corresponded to 393 and 228 absolute FU in two hybridization experiments. Compared to *L. monocytogenes* it made for 1.6% of the correct positive hybridization signals. The new antisense probe now allowed reliable

detection of *L. monocytogenes*. The sense probe still showed a cross-hybridization of 21.8 RI and was not useful for the array.

Tab. 3.10: Cross-hybridizations of probes of the final probe set above 2% MM/PM ratio, which were accepted due to the multiple probe concept, and the mismatches and insertions between the probe and the respective area in the target DNA of the hybridized species. The table does not include multi-targeting probes.

| Probe | Hybridized species | Nb. of mismatches ^{a)} | Nb. of ins./del. ^{b)} | Signal ratio MM/PM ^{c)} |
|---------------------|---|---------------------------------|--------------------------------|----------------------------------|
| EHEC23S_1164_18 | <i>E. coli</i> | 3 | 1 ins. | 0.07 |
| V.chol23S_1379_20 | <i>V. fischeri</i> , <i>V. parahaemolyticus</i> | 2, 2 | 0, 0 | n.d. |
| V.chol23S_1580_21/i | <i>V. fischeri</i> , <i>V. parahaemolyticus</i> | 2, 2 | 0, 0 | n.d. |
| Camp23S_2174_22 | <i>M. av. subsp. paratuberculosis</i> | 10 | 0 | 0.027 |
| Camp23S_2174_22 | <i>B. ureolyticus</i> | no sequence | data | 0.024 |
| List23S_2175_20 | <i>M. av. subsp. paratuberculosis</i> | 11 | 0 | 0.056 |
| Entc23S_175_19/i | <i>L. lactis</i> | 1 | 0 | 0.24/0.07 |
| E.faec23S_1172_20i | EHEC | 10 | 0 | 0.025 |
| E.faec23S_1172_20i | <i>V. fischeri</i> | 10 | 0 | 0.031 |
| E.faec23S_1172_20i | <i>E. faecium</i> | 2 | 0 | 0.043 |
| E.faec23S_284_25 | <i>A. bestiarum</i> (sample 24) | 7 | 4 del. | 0.10 |
| Camp23S_932_18 | <i>C. haemolyticum</i> | 1 | 0 | 0.10 |
| Camp23S_932_18i | <i>B. ureolyticus</i> | no sequence | data | 0.17 |
| Y.ent23S_1478_18/i | <i>Y. pseudotuberculosis</i> | 1 | 0 | 0.89/1.28 |
| Y.ent23S_1501_20 | <i>Y. pseudotuberculosis</i> | 2 | 0 | 0.54 |
| E.coli23S_271_18B/i | EHEC | 0 | 0 | 1.71/1.92 |
| E.coli23S_271_18B/i | <i>S. enteritidis</i> | 2 | 0 | 0.49/0.19 |
| E.coli23S_271_18B/i | <i>S. typhimurium</i> | 2 | 0 | 0.33/0.12 |
| Atop16S_1020_18 | <i>E. faecium</i> | 10 | 0 | 0.06 |
| B.frag23S_152_20 | <i>B. bifidum</i> | no sequence | data | 0.14 |

^{a)} between probe sequence and hybridized species in the same region

^{b)} number of insertions or deletions in the target sequence towards the probe sequence

^{c)} hybridization signal of mismatching target vs. hybridization signal of perfectly matching target; '/' separates values of sense and antisense probe

n.d. – not determined, but the RI of cross-hybridization was above 2.0

The *Campylobacter jejuni* probe of the first probe set for this species was unspecific towards the reference DNA of *Campylobacter coli* and *Campylobacter lari*. *Campylobacter coli* even showed higher hybridization signals with these probes than *C. jejuni*. By sequencing, it was found that the *C. coli* reference strain had no mismatches towards the *C. jejuni*-probes, although the *C. coli* sequences in the database had seven mismatches and two insertions in this area. The reference strain of *C. lari* had one mismatch and one insertion resulting in five mismatches in total. On the bases of the new sequences two new probes, C.jej23S_387_21/i, were developed, which can distinguish between both species (Fig. 3.13). Only the antisense probe of position 394 still hybridized with *C. jejuni* and *C. lari* with an intensity of 27% and 9% of the perfect match. The relative fluorescence intensity was 15.2 and 5.0 RI.

The genus probes for clostridia displayed cross-hybridization signals with *E. faecalis*, *E. faecium*, *L. lactis*, *R. intestinalis*, and *B. fragilis*. A redesign of the probes by cutting two nucleotides at the 3' end reduced the cross-hybridizations, but did not fully exclude false-negatives (Tab. 3.8). The MM/PM ratio of the *E. faecalis* cross-hybridization with sense and antisense probe was reduced from 10% and 3% to 2% and 0.

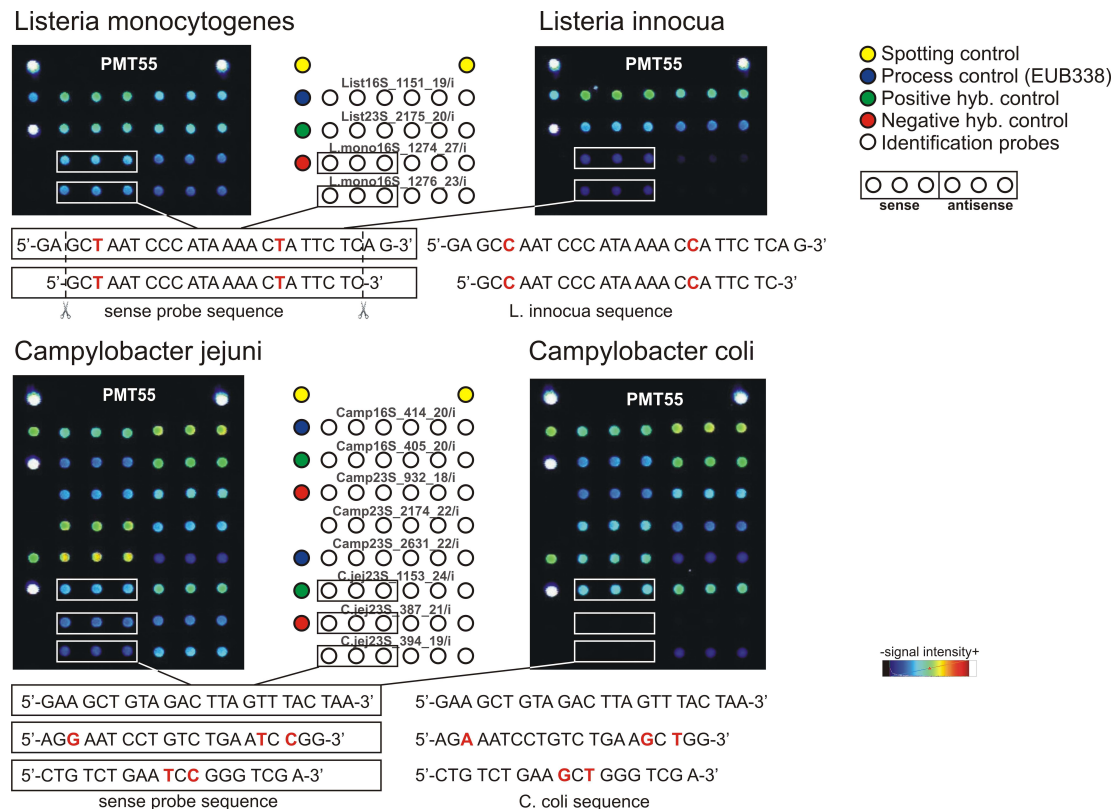


Fig. 3.13: Examples for the redesign of an unspecific *L. monocytogenes*-probe and the design of new probes to replace an unspecific *C. jejuni*-probe. The sequences of the sense probes (frames) are opposed to the sequences of the cross-hybridizing bacteria. Mismatches are marked in red. The antisense probe sequences are not shown. The hybridization result of the correct positive (left) and the cross-hybridizing organism (right) are set in contrast with each other.

The *E. faecium* cross- hybridization with probes Clos23S_1038_24/i was reduced from 22% and 7% to 8.5% and 2% and the *L. lactis* cross-hybridization from 22% and 22% to 6% and 4%. *B. fragilis* displayed a MM/PM ratio with the sense and antisense probe of 1% and 21% before shortening and 0.4% and 10% after shortening. In case of *R. intestinalis*, it seemed from the average values that the MM/PM ratio had not changed or even increased slightly. For the sense and antisense probe, a change of the ratio from 45% and 1% to 45% and 2% was observed by shortening. However, the average value of the relative hybridization signals for the 24 nt long probes was calculated from four hybridizations, while the signals for the 22 nt long probes originated from one experiment only. In the ladder, the RI signals were considerably higher in comparison to the previous hybridizations. This resulted in a bias. If only the results from hybridizations with array V6 were considered including both, the short and the long clostridia probes, a clear reduction of the cross-hybridization was observed. The MM/PM ratio of the *R. intestinalis* hybridization with the clostridia sense and antisense probe was clearly reduced from 90% and 3% to 45% and 2% by probe shortening. Due to the remaining cross-hybridizations the shortened probes were also excluded from the microarray.

Some probes were accounted multi-targeting probes. The probes E.coli23S_271_18B/i hybridized with *E. coli*, EHEC, *S. enteritidis*, and *S. typhimurium*. The probes Y.ent23S_1478_18/i, which hybridized with *Y. enterocolitica* and *Y. pseudotuberculosis*, and V.chol23S_1379_20 and V.chol23S_1580_21/i hybridizing with *V. fischeri* and *V. parahaemolyticus* are likewise regarded as group-probes. The EHEC23S_1164_18 sense and antisense probes hybridized with EHEC and *S. enteritidis*, but not with *S. typhimurium*. Due to the hybridization ability of the EHEC-probe towards *Salmonella* spp., it is not possible to detect EHEC in a co-infection with *Salmonella*.

The specific probes for *Aeromonas hydrophila* A.hyd23S_1170_18/i displayed no hybridization with the reference DNA, which was identified as *A. bestiarum* by sequencing. The 23S sequence contained three mismatches between probe and target. Before unambiguous identification of this species, a specific probe for DNA isolate 24 was designed based on the partially finished sequencing data. A.hyd23S_241_18/i was designed to detect the reference isolate 24. This probe displayed strong cross-hybridization with other species (Tab. 3.8), but was finally excluded from the array, because it actually targets *A. bestiarum* and not *A. hydrophila*.

The final probe set comprised 108 identification probes, which are shown in Tab. 3.11 illustrating the array performance with the tested reference species. The spotted final array V6 (Fig. 2.1) comprised all probes after redesign but also the excluded probes. A virtual array, array V6f (Fig. 2.2), shows all probes of the final probe set, but was not spotted like this.

Tab. 3.11: Final probe set of the diagnostic Gastroenteritis-microarray according to layout V6f. (green) correct positives, (light red) false positives between 1 and 5% RI, (orange) false positives between 5 and 10% RI, (dark red) false positives above 10% RI, (grey) no hybridization performed, (*) hybridization result in conflict with clinical identification or tube labelling.

| Internal number Nb. hyb. exp. à 2 arrays ¹⁾ Probe | 10 | 15 | 16 | 6A + 6B | 1 | 33 | 13 | 18 | 24* | 25 | 26 | 7 | 8 | 9 | 11 | 31 | 5 | 32 | 12 | 2 | 22 | 23* | 4 | 3 | 20 | 21 | 14 | 17A* | 17B | 29 | 30 | 19 | | |
|--|---------------|----------|------------|---------------|---------------|-----------|---------------|-------------|--------------|-------------|-------------|--------------|--------------|------------|-----------------|---------------|---------------|-------------|---------------|-----------------|------------|------------|--------------|---------------|---------------|-------------|-----------------|-------------|---------------|--------------|-------------|------------|-----|-----|
| | P.shig 2/1 | MAP 1 | M.tub 1 | L.mono 3/2 | L.inno 2/1 | EHEC 1 | S.ente 3/1 | S.typh 1 | A.best. 1 | V.para 1 | V.fish 1 | C.jej 2/1 | C.col 2/1 | C.lar 1 | C.diff 3/2/1 | C.haem 2/1 | Y.ente 2/1 | Y.pseu 1 | E.coli 2/1 | R.inte 3/2/1 | A.min 1 | A.min 1 | A.vag 2/1 | V.parv 2/1 | B.frag 2/1 | B.ureo 1 | E.faec 3/2/1 | E.faec 1 | E.faem 2/1 | L.lac 2/1 | L.delb 1 | B.bif 1 | | |
| P.shig16S_468_20 | 215,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,8 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| P.shig16S_468_20i | 162,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 1,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| P.shig23S_1475_19 | 43,7 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| P.shig23S_1475_19i | 81,0 | 0,0 | 0,3 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,3 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Mycb16S_987_20 | 0,0 | 111,0 | 72,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Mycb16S_987_20i | 0,0 | 218,3 | 202,0 | 2,7 | 2,6 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 2,0 | 0,0 | 0,0 | 0,2 | 1,4 | 0,8 | 0,1 | 0,3 | 0,0 | 0,0 | |
| Mycb23S_1872_20 | 0,0 | 60,5 | 58,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Mycb23S_1872_20i | 0,0 | 130,2 | 182,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | |
| M.avC23S_650_17 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| M.avC23S_650_17i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| M.avP23S_650_17 | 0,0 | 91,4 | 0,3 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 1,2 | 0,0 | 0,0 | 0,0 | |
| M.avP23S_650_17i | 0,0 | 41,2 | 0,7 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,6 | 0,0 | 0,0 | 0,0 | 0,0 | |
| List23S_2175_20 | 0,0 | 6,7 | 0,0 | 154,9 | 118,4 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| List23S_2175_20i | 0,0 | 0,5 | 0,0 | 154,0 | 89,9 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| L.mono16S_1276_23i | 0,0 | 0,0 | 0,0 | 75,3 | 1,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,4 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| EHEC/Salm23S_1164_18 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 34,9 | 27,7 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 2,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| EHEC/Salm23S_1164_18i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 15,4 | 13,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Salm23S_274_20 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 44,3 | 43,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,7 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Salm23S_1704_17 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 23,8 | 34,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Salm23S_1704_17i | 0,7 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 38,5 | 37,3 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,7 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Aero23S_1402_20 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 75,7 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Aero23S_1402_20i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 85,5 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| A.hyd23S_1170_18 | 0,0 | 0,0 | 0,0 | 0,3 | 0,2 | 0,0 | 0,0 | 0,0 | 0,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| A.hyd23S_1170_18i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 1,4 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Vibri16S_570_20i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 170,8 | 121,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.chol23S_1207_21 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,3 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| V.chol23S_133_22 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.chol23S_133_22i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.chol23S_1379_20 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 30,6 | 19,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.chol23S_1379_20i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,1 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.chol23S_1580_21 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 10,8 | 8,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.chol23S_1580_21i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 46,7 | 28,6 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.parRel16S_63_19 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 102,0 | 11,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.parRel16S_63_19i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 260,9 | 36,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.par23S_339_19 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 89,8 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.par23S_339_19i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 47,8 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.par23S_1212_19 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 100,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| V.par23S_1212_19i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 92,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 |
| Camp23S_2174_22 | 0,0 | 2,9 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 118,4 | 112,0 | 166,5 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 2,5 | 0,2 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | |
| Camp23S_2174_22i | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 0,0 | 41,2 | 47,0 | 43,9 | 0,0 | 0,0 | | | | | | | | | | | | | | | | | | |

3.10 Limit of detection of the DNA microarray

The limit of detection (LOD) of the array was determined in dependence from the grade of PCR product digestion before hybridization and from the influence of background DNA of a faecal sample (chap. 2.2.10). The detection limit for *E. coli* without background DNA was determined as 10^3 genome equivalents when digested with 0.4 or 4 mU DNase/ng DNA (Fig. 3.14). Digested with 4 mU DNase I per ng DNA, *E. coli* DNA showed the highest hybridization fluorescence signals with three out of four probes. The signals were, compared to undigested DNA, 3 to 35 times higher. However, in case of probe Eco23S_271_18Bi the highest hybridization signal was observed after target digestion with 0.4 mU DNase I/ng. For this probe, the hybridization response decreased from 20,998 to 4,301 FU when the DNA was digested with 4 instead of 0.4 mU DNase I/ng when using 10^5 genome equivalents in PCR.

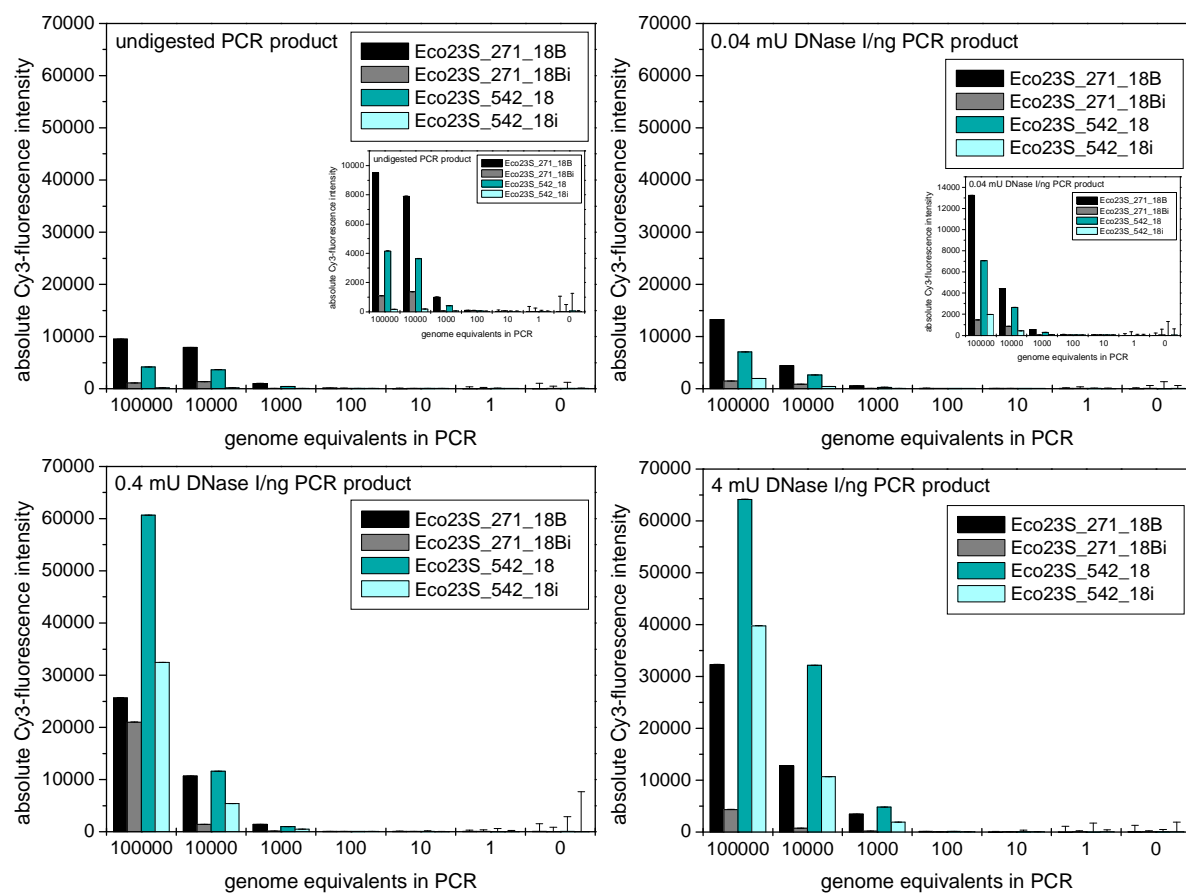


Fig. 3.14: Hybridization of *E. coli* PCR product amplified from different amounts of template DNA ($0-10^5$ genome equivalents) after digestion with three different amounts of DNase I (0.04, 0.4, and 4 mU/ng) and without prior digestion.

Applying high concentrations of DNA (10^5 genome equivalents) to PCR and hybridization led to stronger hybridization signals with increasing grade of digestion up to a turning point, as it was seen in probe Eco23S_271_18Bi. At the detection limit (10^3 genome equivalents in PCR), a digestion prior to hybridization with 0.04 mU/ng only, led to lower hybridization signals than with undigested DNA (Fig. 3.15). However, a higher grade of digestion resulted in higher hybridization responses. If 10^2 genome equivalents were applied to PCR and hybridization, the fluorescence signals were in the range of the average response of the negative hybridization control (NHC) with 60 FU (Fig. 3.15).

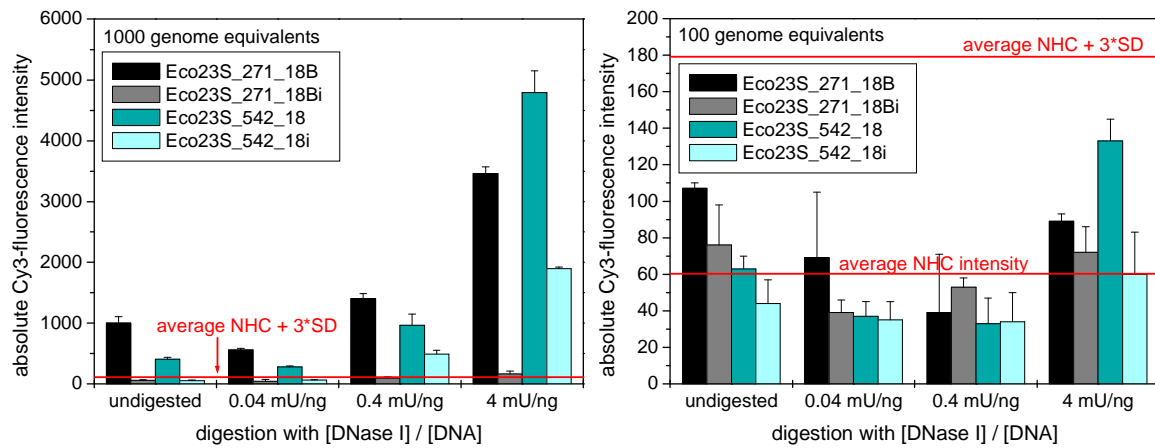


Fig. 3.15: Hybridization of *E. coli* PCR product amplified from 10^3 (left) or 10^2 (right) genome equivalents of template DNA after digestion with three different amounts of DNase I (0.04, 0.4, and 4 mU/ng) and without prior digestion. Red lines indicate the average fluorescence values of the negative hybridization control (NHC) and the NHC plus three times its standard deviation (3^*SD) above all arrays.

For *C. jejuni* the sensitivity was determined in the background of *Campylobacter*-free faecal DNA. A detection limit of 10^3 genome equivalents was determined (Fig. 3.16) as it was for *E. coli* without background DNA. The average signal intensity of the NHC was 162 FU, while the responses of the specific probes were 1,644, 608, 1,249, and 994 using 10^3 genome equivalents in PCR. The excess of background DNA accounted for 2,770 fold in case of 10^3 genome equivalents.

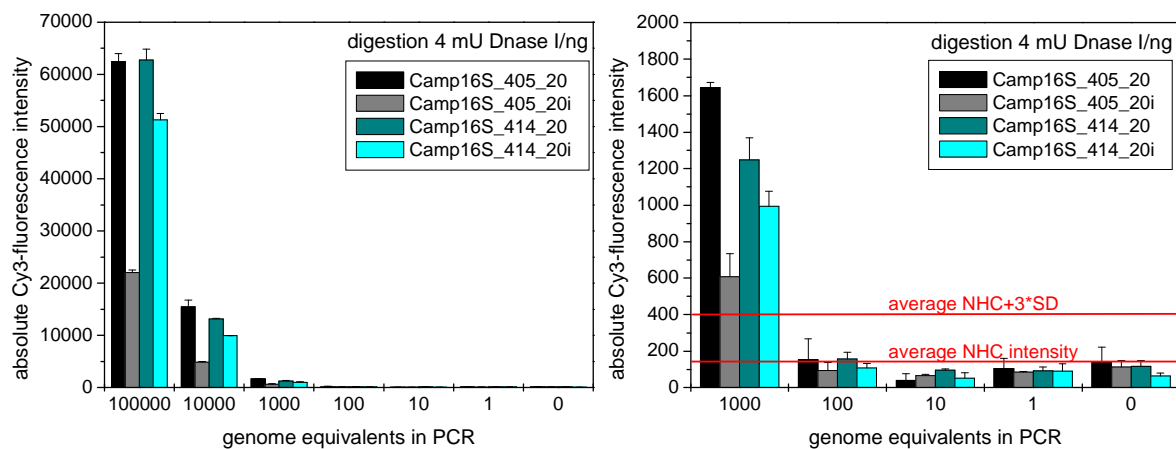


Fig. 3.16: (left) Hybridization of *Campylobacter jejuni* PCR product amplified from different amounts of template DNA ($0-10^5$ genome equivalents) in the excess of background DNA after digestion with 4 mU DNase I/ng DNA. (right) Magnification of the values between 0 and 1,000 genome equivalents of the left bar diagram. Red lines indicate the average fluorescence values of the negative hybridization control (NHC) and the NHC plus three times its standard deviation (3^*SD) above all arrays.

The two experiments proved a detection limit of the present array of 10^3 genome equivalents, which was also valid for bacteria in front of a background of faecal DNA and depended on the grade of digestion of the PCR products before hybridization.

3.11 Application of the identification microarray

3.11.1 Clinical faecal isolates from gastroenteritis patients and healthy individuals

In order to investigate the array performance with real samples, 58 clinical faecal isolates from patients with gastroenteritic symptoms and six samples from healthy volunteers were analyzed. The age distribution of patients is shown in Tab. 3.12. Using specific real-time PCR and stool culture 17 samples were previously tested positive for *Salmonella* spp., eight for *Clostridium* spp. (toxin A and B), four were identified as *Campylobacter* infection, and one as *Yersinia* infection by our clinical partners. Twenty-seven samples were negative for the tested bacterial pathogens. The age distribution with respect to the clinically confirmed pathogen is depicted in Tab. 3.13.

Tab. 3.12: Age distribution of hospitalized individuals with Gastroenteritis and healthy volunteers.

| Age group | Number of samples | Mean age \pm SD (years) | Age range (years) |
|---------------------|-------------------|---------------------------|-------------------|
| 4-17 years | 21 | 8 \pm 3.4 | 4-16 |
| 18-60 years | 18 | 38 \pm 12.6 | 18-55 |
| 61-86 years | 18 | 75 \pm 7.0 | 61-86 |
| Healthy individuals | 6 | n.d. | 30-40 |

Tab. 3.13: Age distribution with respect to the clinically confirmed pathogen.

| Pathogen | Nb. of samples | Mean age \pm SD (years) | Age range (years) | Age group distribution 4-17 / 18-60 / 61-86 years (nb.) |
|------------------------------|----------------|---------------------------|-------------------|---|
| <i>Clostridium difficile</i> | 8 | 62 \pm 24.7 | 8-78 | 1(12.5%) / 1(12.5%) / 6(75%) |
| <i>Salmonella</i> spp. | 17 | 18 \pm 15.4 | 5-55 | 12(71%) / 5 (29%) / 0 |
| <i>Campylobacter</i> spp. | 4 | 44 \pm 19.3 | 18-63 | 0 / 3(75%) / 1(25%) |
| <i>Yersinia</i> spp. | 1 | 47 | | 0 / 1(100%) / 0 |
| unknown | 27 | 44 \pm 31.4 | 4-86 | 8(30%) / 8(30%) / 11(41%) |

The ribosomal genes of each sample were amplified in five separate fragments H, J, DE, F and G (chap. 2.2.4.3, 2.2.5) and the purified PCR products were mixed in equimolar amounts resulting in 1 μ g PCR product per sample. The fragment mix was digested with 0.4 mU DNase I/ng (chap. 2.2.6) and 500 ng digested DNA were then hybridized to each of two arrays with layout V6 (chap. 2.2.1.2, 2.2.7). The Cy3 incorporation rates varied for each PCR product between 60 and 150. In case of sample S29, S37, S48, and S52, the whole PCR products of all fragments were mixed due to low amplification efficiency and the DNA was hybridized without prior digestion. In sample S29, the H, DE, and F fragments were not visible on the agarose gel. Fragments J and G showed only weak bands. For S37, all fragments could be detected by gel electrophoresis, but the bands were very faint. For sample S48, the DE fragment was not detected and all other fragments showed weak bands. Fragment DE of sample S52 was not detected on the agarose gel, while all other fragments displayed weaker bands than other amplified samples. By using the whole PCR product, 364, 474, 561, and 397 ng of sample S29, S37, S48, and S52, respectively, were hybridized to each array. However, the values might be inaccurate due to the limitation of accuracy of the spectrophotometer at low DNA concentrations. Samples S9, S15, S20, S23, S27, S39, S43, S50, S54, S57, S58, and S59 were hybridized a second time using a mixture of the

whole PCR products. In all cases >1 µg digested DNA was hybridized to each array. The amplification products of samples S60-S64 from healthy volunteers were completely hybridized only. Besides for S63, which had low amplification efficiency, more than 1 µg DNase I-digested DNA was hybridized to each array. For S63 only 440 ng PCR product could be hybridized per array. Data processing followed protocol 2.2.9, accepting only probe signals >NHC+3SD with a standard deviation <50% as positive hybridization result. A genus or species was regarded as identified, if more than 50% of the respective probes showed a positive signal. Additionally, the identification of a species required a positive result for the genus, if appropriate probes were included in the array.

The observed standard variation of the hybridization signals from six spots hybridized in two separate arrays (each array contained three spots) (Fig. 3.17) was in 90% of cases ≤20%. More than half of the probes, 63%, displayed a standard variation ≤10%. Only 1% of the hybridized probes had signal standard variation >50%. The mean standard deviation of signals from replicate spots was 11%.

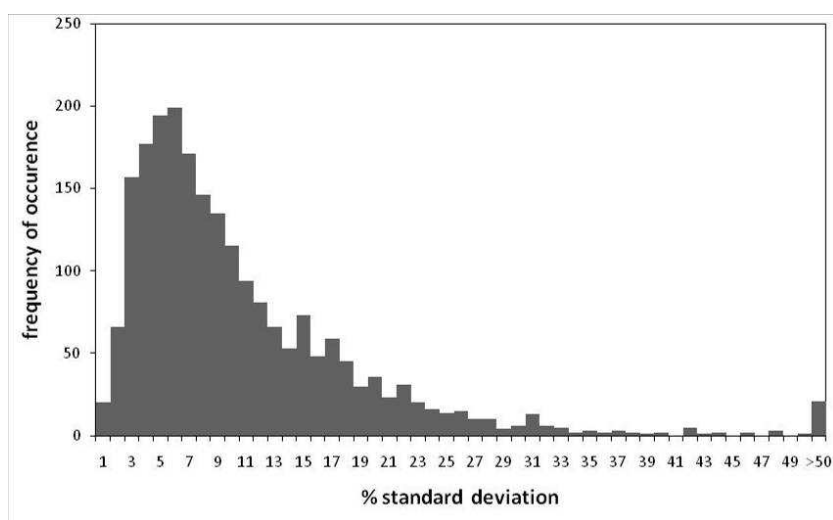


Fig. 3.17: Level distribution of standard deviation of the net signal intensities of 2188 probes comprising each three spots from two hybridized arrays (n=6) in 78 hybridization experiments.

Microarray analysis of the clinical samples revealed individual community patterns regarding the natural intestinal flora (Tab. 3.14). The predominant species were *Bacteroides* spp., *B. fragilis*, *F. prausnitzii*, and *Veillonella* spp., which were detected in 45, 35, 38, and 41 out of 58 patients samples, respectively. These species were also dominant in the healthy individuals. *Roseburia* spp. and *R. intestinalis* were found in 22 and 15 diseased individuals. *Enterococcus* spp. and *E. faecalis* were found in 30 and 6 samples and *E. coli* in 24 patients' samples. *Bifidobacterium bifidum* was detected in six samples only, whereas *Lactococcus lactis* was identified in 14 samples of gastroenteritis patients. *Atopobium* spp. was underrepresented and detected only in two specimens but *A. minutum* was not identified. *Eubacterium bifforme* and *L. acidophilus* were found in four and one faecal DNA isolates from gastroenteritis patients. In samples S16, S51 and S55, nearly no background flora was detected. The six healthy individuals displayed a comparably diverse background flora with *Bacteroides* spp. and *B. fragilis*, *F. prausnitzii* and *Veillonella* spp. in five, two, five, and four samples. *Roseburia* spp. and *R. intestinalis* were found in two samples, *E. coli* in three. *Lactococcus lactis* was identified in three individuals, *Enterococcus* spp. and *E. faecalis* in two and one person, respectively. Fig. 3.22 shows the detected bacterial profile of one healthy adult.

Considering all isolates from hospitalized patients, 77% were analyzed in agreement with the clinical result. Samples from healthy individuals and samples that could not be amplified were not included. If only those samples were analyzed, where the etiologic pathogen had been detected by a combination of culture and specific real-time PCR assays, the percentage of correct results according to clinical routine diagnostic was 67%. Four samples, S29, S37, S48, and S52, could not be amplified and therefore not be analyzed with the

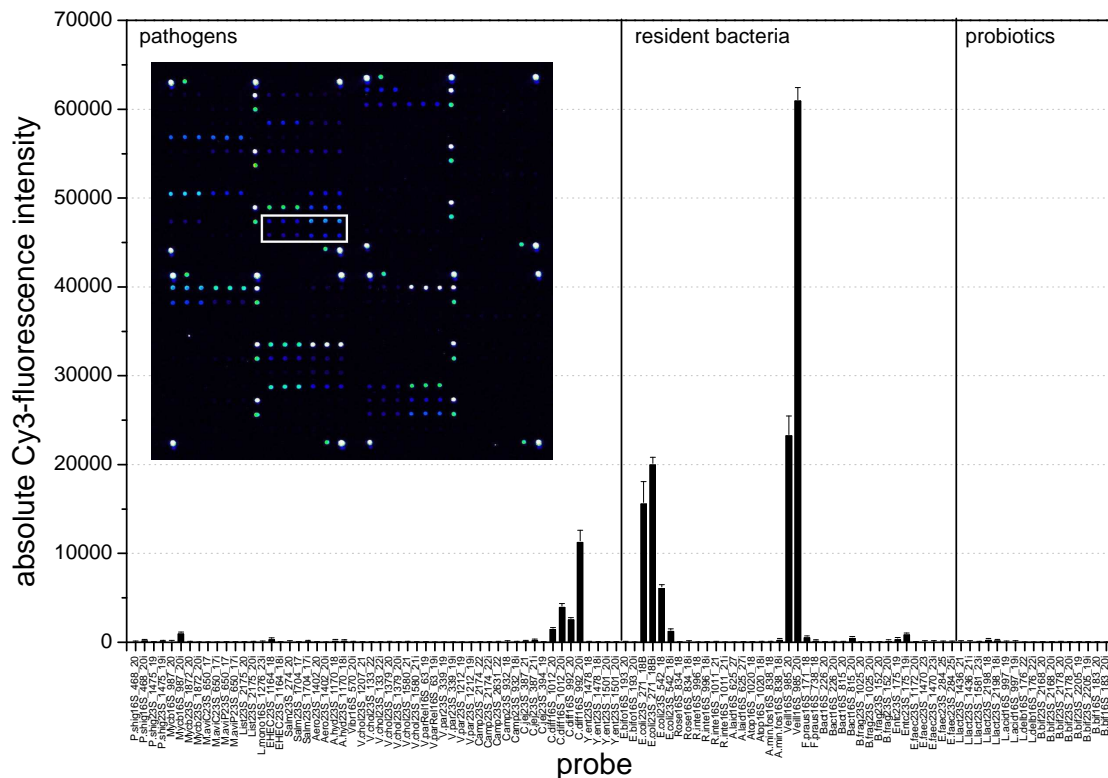


Fig. 3.18: Infection with *Clostridium difficile*: Hybridization of amplified faecal DNA (sample S14) from a *C. difficile* infected patient with the Gastroenteritis-Chip V6. Quantification according to layout V6f (PMT 60). (white frame) *C. difficile*-specific probes (V6f).

In five from eight *Clostridium* toxin-positive samples, the bacterium was correctly detected. Fig. 3.18 shows the hybridization result of sample S14, which was positive for *C. difficile* detected by four specific probes. From the resident microbiota *E. coli*, *Veillonella* spp., *F. prausnitzii*, and *Enterococcus* spp. were detected. *Campylobacter jejuni* was identified in three out of four samples (example in Fig. 3.19), which were supposed by real-time PCR to contain this pathogen. In 12 from 18 *Salmonella*-positive samples, the pathogen was also identified by microarray. Two *Salmonella*-positive samples, S15 and S39, showed a weak but positive result only after hybridization of the whole PCR product onto two arrays (1-1.5 µg/array). In Fig. 3.21, the hybridization pattern of a faecal sample from a patient with *Salmonella*-infection can be seen. From the resident intestinal community only *Bacteroides*, *E. coli*, and *Veillonella* spp. were weakly detected. The *Vibrio*-genus probe displayed a weak signal, as well, but the species-specific probes did not respond. The *Salmonella*-infection was clearly detected by three specific probes and four multi-targeting probes (EHEC23S_1164_18/i and Ecol23S_271_18B/i). The *Yersinia*-infection in one specimen was not detected by the microarray. The probe set for *Mycobacterium* spp. was positive in nine faecal specimens and for *Vibrio* spp. in six samples.

Five samples also showed wrong positive results for pathogens, which were not detected in the clinical routine. These were *C. difficile* (S61, S41, S45), *Campylobacter* ssp. (S32), and *Salmonella* ssp. (S50). In case of *C. jejuni*, this indicated a double infection with *Clostridium difficile* and *Campylobacter jejuni* (Fig. 3.20). False positive hybridization signals were also found for the *Vibrio* genus probe and the *Mycobacterium* genus probes.

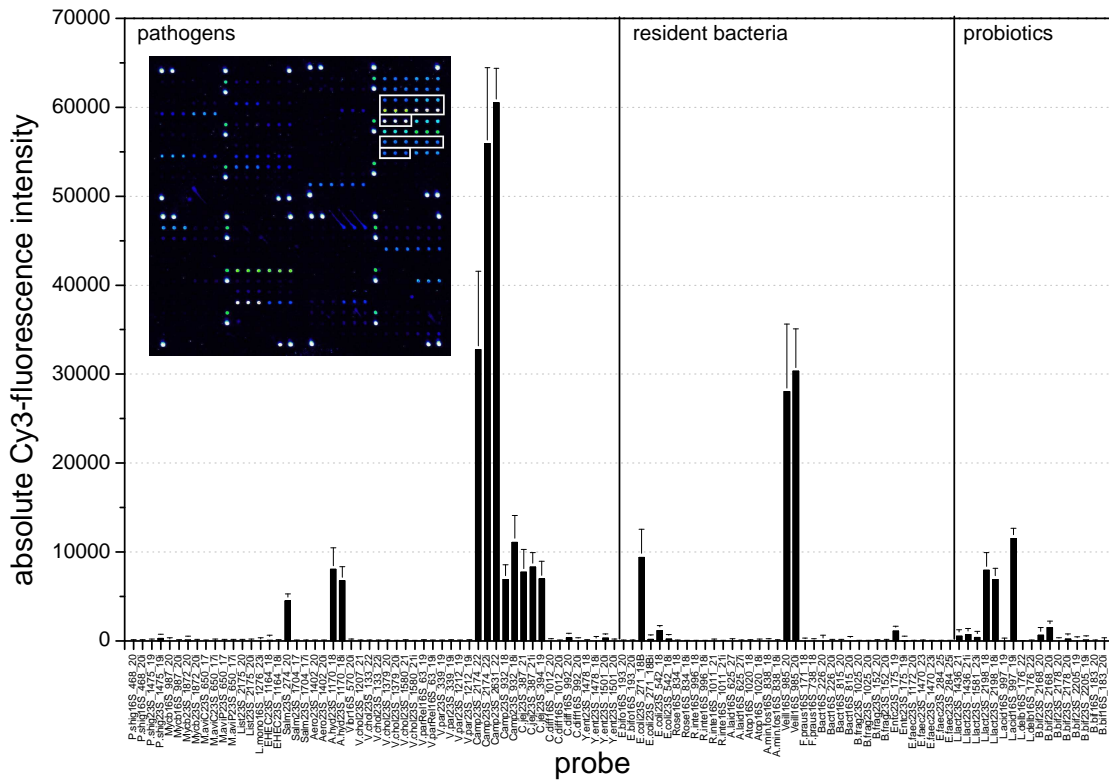


Fig. 3.19: Infection with *Campylobacter jejuni*: Hybridization of amplified faecal DNA (sample S11) from a *Campylobacter*-infected patient with the Gastroenteritis-Chip V6. Quantification according to layout V6f (PMT 60). (white frame) *Campylobacter*-specific probes (V6f).

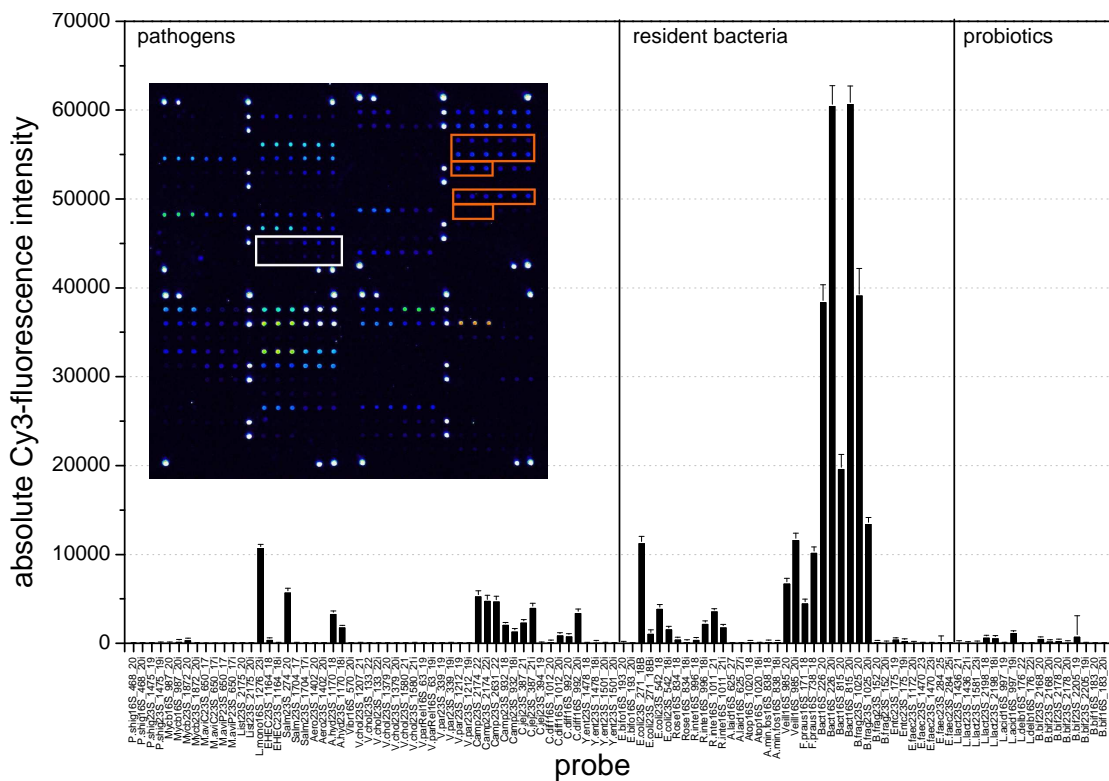


Fig. 3.20: Potential double-infection with *Clostridium difficile* and *Campylobacter jejuni*: Hybridization of amplified faecal DNA (sample S32) from a *C. difficile*-infected patient with the Gastroenteritis-Chip V6. Quantification according to layout V6f (PMT 60). (white frame) *C. difficile*-specific probes (V6f); (orange frame) *C. jejuni*-specific probes (V6f).

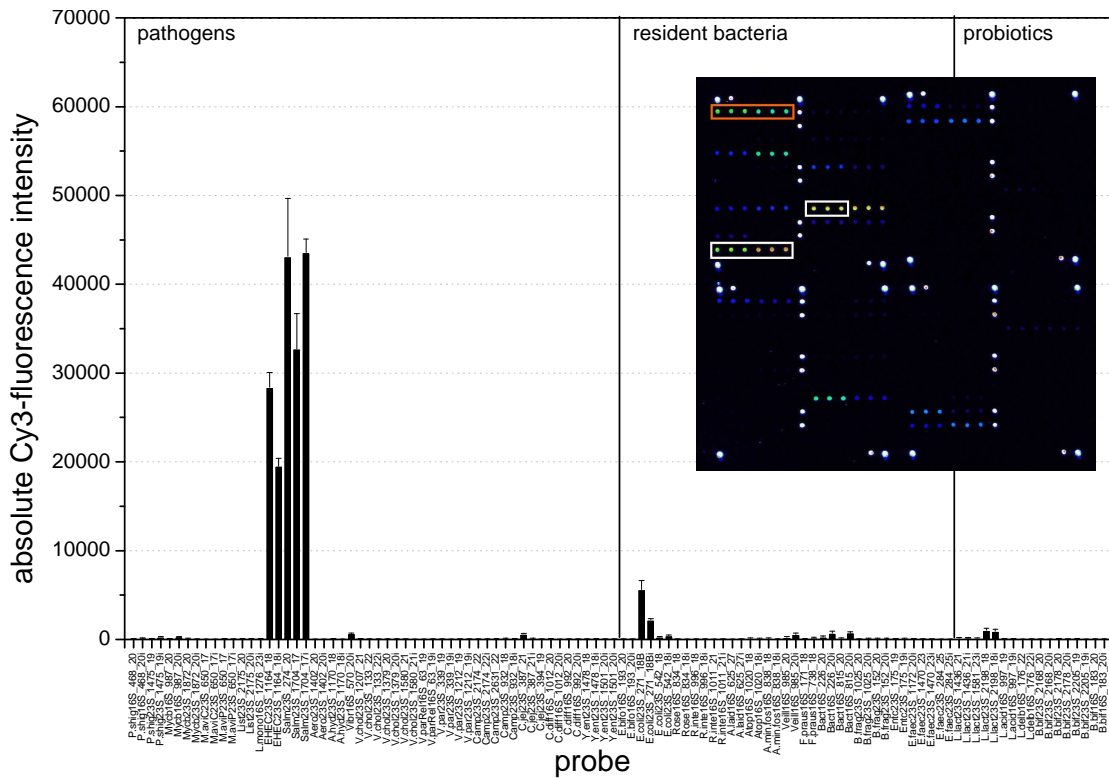


Fig. 3.21: Infection with *Salmonella* spp.: Hybridization of amplified faecal DNA (sample S18) from a *Salmonella* infected patient with the Gastroenteritis-Chip V6. Quantification according to layout V6f (PMT 60). (white frame) *Salmonella*-specific probes (V6f); (orange frame) EHEC- and *Salmonella*-specific probes (V6f).

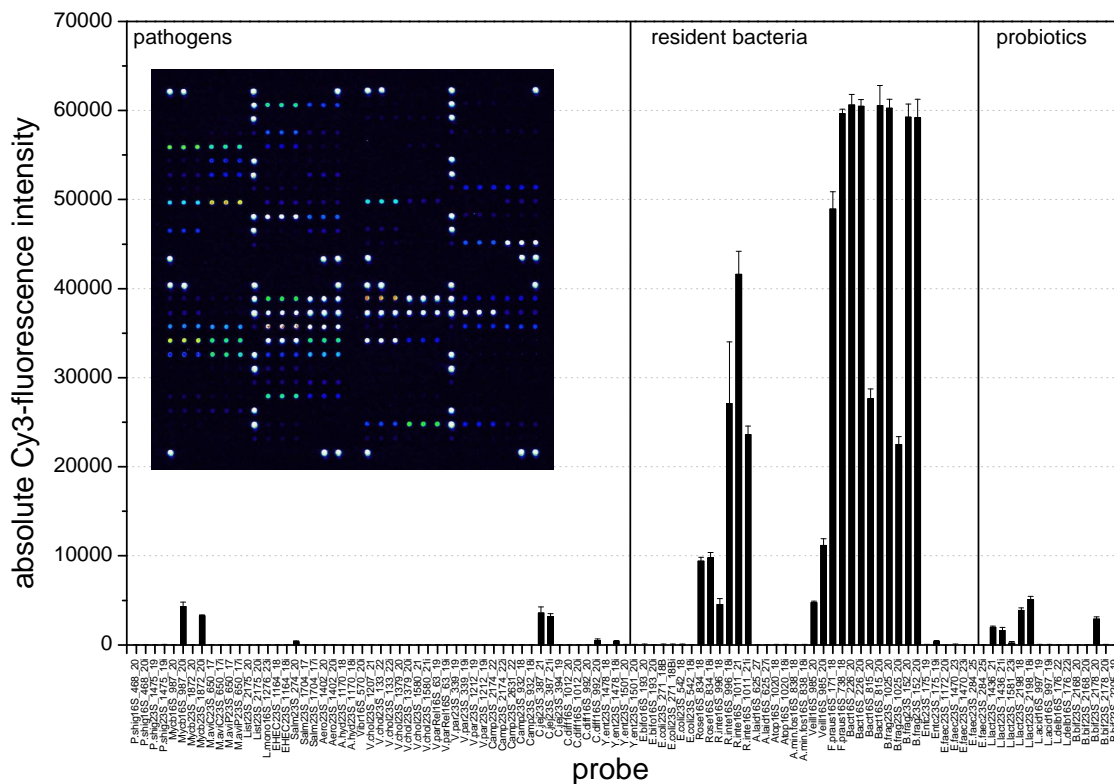


Fig. 3.22: Healthy adult: Hybridization of amplified faecal DNA (sample 59) from a healthy adult with the Gastroenteritis-Chip V6. Quantification according to layout V6f (PMT 60).

PCA and PLSA did not reveal significant differences in the microarray-detected intestinal flora between hospitalized patients suffering from gastroenteritis and healthy individuals (Fig. 3.23). In PCA the two principle components, which allowed best discrimination of both groups, explained only low amount of variation (PC7=4.9%; PC4=11.4%) and did not allow separation of both groups. In PLSA both groups clustered together as well. The classification error of cross-validation was 10.3% for a one- or two-component PLS model. Using a one-component PLS model at a significance of $p=0.05$, *B. fragilis* was identified as the most important variable for group discrimination. However, one-way ANOVA did not confirm this result. The presence of *B. fragilis* in both groups was not significantly different ($p=0.147$).

Separating the samples into three groups comprising (a) hospitalized individuals with diagnosed pathogen, (b) hospitalized individuals with undiagnosed pathogen, and (c) healthy individuals did not improve separation of the groups (Fig. 3.24). The correct classification rates for the PLS models were 60.3% and 62.1% for one or two components, respectively. Using either the one- or two-component model, the data were also not separated according to the grouping.

Thus, the intestinal background microbiota in this sample cohort was not significantly different between patients and the small group of healthy people. Nevertheless, each individual contained an individual intestinal microbiota.

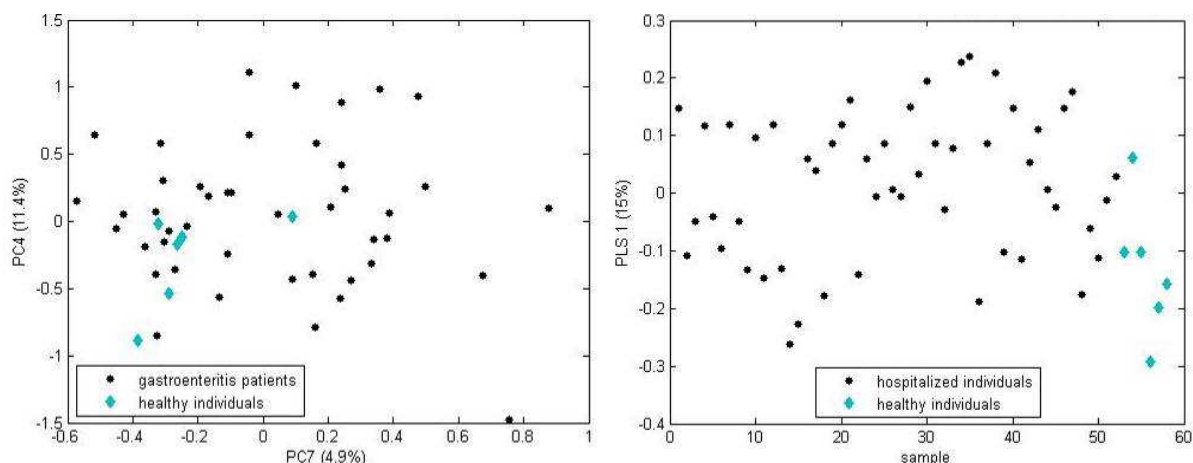
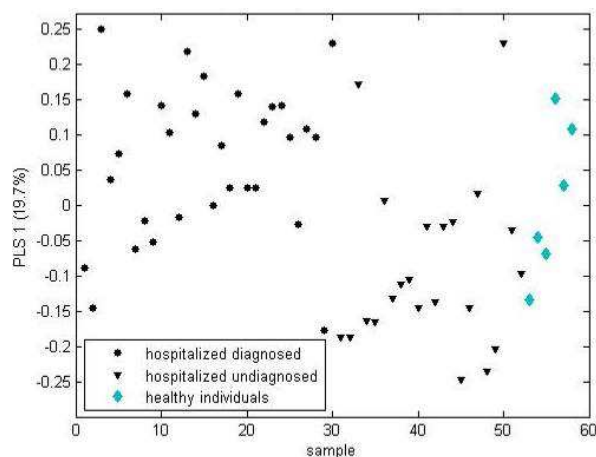


Fig. 3.23: (left) Principle component analysis and (right) partial least squares analysis with two-group classification of the results from microarray hybridization of faecal samples from hospitalized individuals suffering from gastroenteritis and healthy persons.

Fig. 3.24: Partial least squares analysis with three-group classification of the results from microarray hybridization of faecal samples from hospitalized individuals suffering from gastroenteritis and healthy persons.



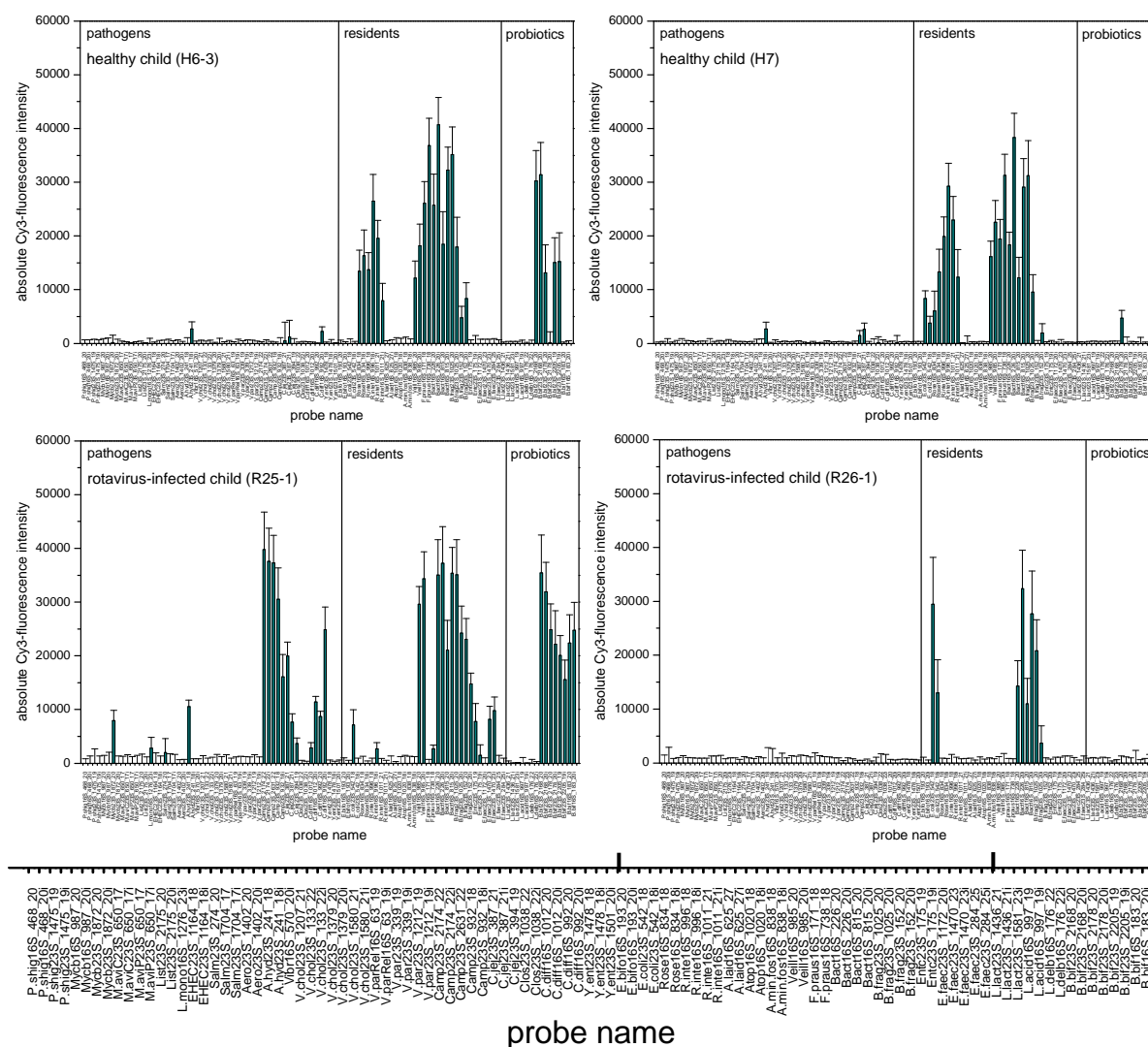


Fig. 3.25: Bacterial profiles of the intestinal flora of two healthy (H6-3, H7) and two infected (R25-1, R26-1) children measured by the identification array V7. Signals are average values of six spots from two hybridizations corrected for the average signal intensity of the negative hybridization control plus three times its standard deviation. (bottom) Magnification of the x-axis of the four graphs.

Multivariate analysis (chap. 2.2.13) was used to investigate whether the data regarding the intestinal background flora obtained by microarray analysis were sufficient to separate healthy from ill children. Two methods, principle component analysis (PCA) and partial least squares (PLS) analysis, were applied.

In the PCA score plot (Fig. 3.26), the infected and healthy children assembled in two separate groups. The principle components PC1 and PC3 explained considerable variance of the data (PC1=36.3%; PC3=13.8%). In the PLS plot (Fig. 3.26) the separation of both groups was even more obvious. According to leave-one-out cross-validation (LOOCV) 90% of samples were correctly classified if one PLS component was used for the PLS model. A PLS model with two components yielded 95% accuracy. Using the two-component PLS model at a significance of $p=0.05$, four microarray targets were identified as the most important in discriminating healthy from infected children: *Roseburia* spp., *E. coli*, *F. prausnitzii*, and *Atopobium* spp. (PLS regression plot, Fig. 3.27). Positive regression coefficients correlated with the diseased state, while negative ones correlated with the healthy state. *Roseburia* was detected in seven out of ten healthy children but in none of the ten infected children. *Fusobacterium prausnitzii* was detected in all healthy individuals but only in one infected child. *Escherichia coli* was found in five infected and four healthy children

and *Atopobium* in two infected but no healthy child. One-way ANOVA (Tab. 3.16), as an alternative method, showed that the difference of *Roseburia* spp. ($p=2.3 \cdot 10^{-4}$) and *F. prausnitzii* ($p=4.4 \cdot 10^{-8}$) between the two groups was significant ($p < 0.05$), whereas the difference of *E. coli* ($p=0.673$) and *Atopobium* ($p=0.151$) was not. One-way ANOVA found also *R. intestinalis* ($p=1.7 \cdot 10^{-3}$) to be significantly different between both groups of children, whereas all others were not.

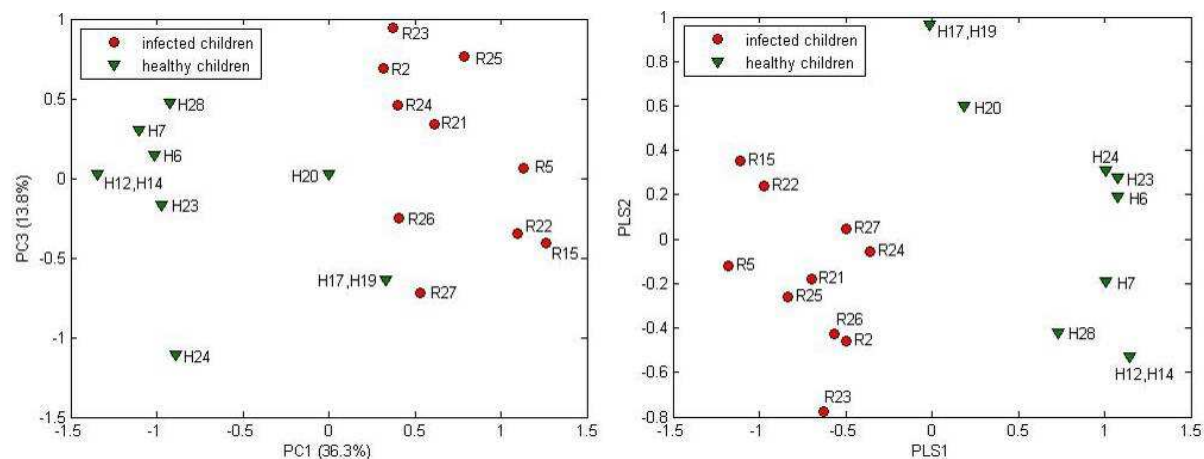


Fig. 3.26: Principle component analysis (left) and partial least squares analysis (right) of the results from microarray hybridization of faecal samples from healthy and rotavirus-infected children.

Tab. 3.16: One-way ANOVA calculated p-values of the microarray results for the intestinal background flora giving the significance of differentiation between the two groups of healthy and rotavirus-infected children. The ANOVA p-values of species identified as significant variables by PLSA are given in bold.

| Variable | Microarray target | p-value |
|----------|--------------------------|--|
| 1 | <i>Bacteroides</i> spp. | 0.151 |
| 2 | <i>B. fragilis</i> | 0.196 |
| 3 | <i>Enterococcus</i> spp. | 0.331 |
| 4 | <i>E. faecalis</i> | n.d. |
| 5 | <i>Roseburia</i> spp. | $2.31 \cdot 10^{-4}$ |
| 6 | <i>R. intestinalis</i> | $1.7 \cdot 10^{-3}$ |
| 7 | <i>E. coli</i> | 0.673 |
| 8 | <i>F. prausnitzii</i> | $4.4 \cdot 10^{-8}$ |
| 9 | <i>Veillonella</i> spp. | 0.355 |
| 10 | <i>B. bifidum</i> | n.d. |
| 11 | <i>L. lactis</i> | 0.331 |
| 12 | <i>L. delbrueckii</i> | n.d. |
| 13 | <i>L. acidophilus</i> | n.d. |
| 14 | <i>E. bifforme</i> | n.d. |
| 15 | <i>Atopobium</i> spp. | 0.151 |
| 16 | <i>A. minutum</i> | n.d. |

n.d. – not determined (due to mathematical reasons, e.g. only zeros in both groups)

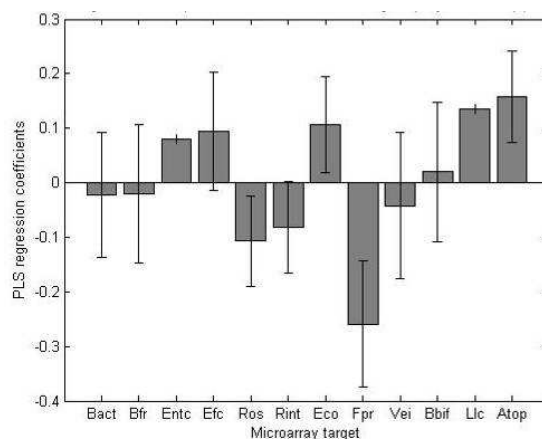


Fig. 3.27: PLS regression coefficients plot of twelve microarray-detected species or genera when two PLS components were used ($p=0.05$). The bar indicates the uncertainty limits.

Based on the data from *Roseburia* spp., *E. coli*, *F. prausnitzii*, and *Atopobium* spp. only, a new PLS model was established. The cross-validated correct classification rate was then 95% for a one-component model and 90% for a two-component model. However, when *Roseburia* spp. and *F. prausnitzii* were used only, both models yielded the maximum

accuracy of 95%. Including *R. intestinalis* into the establishment of the PLS model (*Roseburia* spp., *R. intestinalis*, and *F. prausnitzii*), which was suggested by one-way ANOVA, yielded a correct classification rate of only 80% and 95% for a one- and two-component model, respectively.

Thus, the background microbiota of rotavirus-infected children was clearly influenced by the infection and some species were identified, which mostly contributed to distinguish between diseased and healthy individuals.

3.11.3 Establishment of human and pig intestinal flora in piglets

Two groups of piglets with human-associated (HFA) and pig-associated (PFA) flora were investigated. Both animal groups were meant to investigate the establishment of the human gut flora in pigs in comparison with the pig intestinal flora, but a diarrheal infection occurred in both groups, which was more severe in the group with inoculated pig flora. Two of the pig flora-associated piglets died on day 10 and 21 after birth. In Tab. 3.17 the course of disease during the time of sampling is illustrated.

Tab. 3.17: Diarrheal symptoms of the piglets during the sampling. P-pig flora-associated animals H-human flora-associated animals.

| Age of the piglets (days) | P1 | P2 | P3 | P4 | P5 | P6 | P7 | P8 | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 |
|---------------------------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| 5 (sampling) | x | x | √ | x | x | x | x | x | x | x | x | x | x | x | x | x |
| 6 | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
| 7 (sampling) | x | x | √ | √ | x | √ | √ | x | x | √ | x | x | x | x | x | x |
| 8 | x | x | x | x | x | √ | √ | x | x | x | x | x | x | x | x | √ |
| 9 | x | x | x | x | √ | √ | √ | √ | √ | √ | x | x | x | x | x | x |
| 10 | √† | x | √ | x | x | √ | x | √ | √ | √ | x | x | x | x | x | x |
| 11 | | x | √ | x | x | √ | x | √ | √ | x | x | x | x | x | x | x |
| 12 | | x | √ | x | x | √ | x | x | √ | x | x | x | x | x | x | x |
| 13 | | x | √ | √ | x | √ | x | √ | x | x | x | x | x | x | x | x |
| 14 (sampling) | | x | √ | √ | x | √ | √ | x | x | x | x | x | x | x | x | x |
| 15 | | x | √ | x | x | √ | √ | x | x | x | x | x | x | x | x | x |
| 16 | | √ | √ | √ | √ | √ | √ | √ | | | | | | | | |
| 17 | | x | ℃ | x | x | x | x | x | | | | | | | | |
| 18 | | x | x | x | x | x | x | x | | | | | | | | |
| 19 | | x | x | x | x | x | x | x | | | | | | | | |
| 20 | | x | x | x | x | x | x | x | | | | | | | | |
| 21 (sampling) | | x | x | † | x | x | x | x | | | | | | | | |
| 22 | | x | √ | | x | √ | x | x | x | x | x | x | x | √ | x | x |
| 23 | | x | √ | | x | √ | x | x | x | x | x | x | x | √ | x | x |
| 24 | | x | √ | | x | x | x | x | | | | | | | | |

x - no symptoms; √ - diarrhea; ℃ - fever; † - died

Samples taken on day 14 and 21 after birth were investigated by microarray and DGGE (day 14 only). For the microarray experiments the array layout V7 (Fig. 2.3) was used. The samples were prepared as described in chapter 2.2.3.3. The DNA was amplified and Biotin-labelled (chap. 2.2.4.4, 2.2.5) and the PCR products were digested before hybridization with 0.04 mU DNase I per ng DNA (chap. 2.2.6) due to higher specific activity of this enzyme. The hybridization (chap. 2.2.7) was detected using the silver deposition method (chap. 2.2.8, 2.2.9). In Tab. 3.18 and Tab. 3.19, the results are illustrated for both points in time. Every piglet had its own individual flora established after fourteen days, which was not identical with the donor flora. *Bacteroides* was detected in every sample, *B. fragilis* only in some. *Bacteroides fragilis* was found neither in the conventionally raised piglets (CV-piglets) nor in

the pig flora donor, but the human donor held this species above the detection limit. The recipients partly contained it and partly did not. *Escherichia coli* was found in the human donor and nearly all its recipients except for H8. In the pig donor it was not detected but in all its recipients. From the CV-piglets, only one had *E. coli* in detectable amounts in its gut flora. The genus *Enterococcus* was not detected at all. The probes for *E. faecalis* were negative for the donors and the CV-piglets, but positive in all recipients except for P5. However, *E. faecalis* could be regarded as detected only, if the genera probes were positive as well. *Roseburia* and *R. intestinalis* were detected in the human donor only. *Bifidobacterium bifidum* was present in the human donor and one of its recipients. The CV-piglets exclusively contained *Veillonella*. *Lactobacillus delbrueckii* was found in six out of seven pig flora recipients but in no other sample. *Fusobacterium prausnitzii* was detected in the CV-piglets and both donors but had established only in five pig flora-associated piglets. Regarding the pathogens, the array detected *C. jejuni* in some samples but due to missing signals from the genus-specific probes, this was assumed a cross-hybridization. Most interesting were the two samples from the HFA-piglets, which were positive for enterohemorrhagic *E. coli* (EHEC).

Seven days after, the bacterial profiles had partly changed. The overall impression was a more diverse flora in the HFA piglets and a reduced flora in the PFA piglets. This was in accordance with the observed course of the disease, which is tabulated above (Tab. 3.17). The bacterial profile of the HFA piglets was now relatively similar to the donor profile. In case of the PFA piglets, the detected donor profile was not very complex, thus it made a comparison of similarity relatively useless. *Bacteroides fragilis* had disappeared in all PFA piglets, and in pigs P3 and P4 even the genus was not detected anymore. In the HFA piglets, *B. fragilis* was not found in pig H4 anymore but was now present in piglets H3 and H8. The genus *Enterococcus* was now detected in one sample, H6. The positive signals for *E. faecalis* had reduced in the PFA group. Regarding *Roseburia* spp. and *R. intestinalis*, nothing had changed. However, *E. coli* was less frequently detected in the PFA group, while now all HFA piglets contained this species. Except for pig H4, *F. prausnitzii* had now established in all HFA piglets, where it was not found before. This was the case for the PFA piglets P5 and P8 as well. On the other hand, in pigs P3 and P4 it was found seven days ago but not on day 21. *Bifidobacterium bifidum* had also established in the HFA group. Only piglet H5 was missing this species. In all other piglets, it was not detected. *Lactobacillus delbrueckii*, which was frequently found in the pig flora recipients at day fourteen, was missing in pigs P3, P4, and P6 seven days later. *Campylobacter jejuni* was still detected in three HFA piglets but again without a positive signal from the genus-specific probe. This was the same for one HFA piglet, which displayed a positive signal for *Aeromonas hydrophila*. The pathogenic EHEC was now identified in piglets H1 and H6 but not in H2 and H4, like seven days earlier.

The faecal samples from HFA and PFA piglets of day 14 were analysed by denaturing gradient gel electrophoresis (chap. 2.2.12) regarding all bacteria (V3 region), to compare the microarray results in terms of the detected diversity with a second method. However, species were not identified with DGGE. Identification of species from DGGE gels would require extraction of the fragments, cloning and subsequent sequencing. The V3 region is a conserved region of the 16S ribosomal gene spanning the nucleotides 442-492 (*E. coli* numbering) (Rajilic-Stojanovic 2009). The DGGE gel of the V3 region is shown in Fig. 3.28. For all samples, a diverse fragment profile was found. A higher fragment homology was visible within the HFA group compared to the PFA group. Moreover, the human donor V3 profile was more similar to its recipients than the pig donor profile to the PFA piglets. The pig donor had a more similar structure to the conventionally raised pig. One dominant band was common to both groups but was faint in H7 and the pig donor.

Tab. 3.18: Identified pathogens and background flora in piglets with pig-associated or human-associated intestinal flora suffering from diarrhea and conventionally raised piglets 14 days after birth. The donors of the intestinal flora are also listed. The species were identified after PCR amplification of the DNA, hybridization to the microarray and acquisition of the silver deposition (green squares).

| piglets | | Background flora and probiotics | | | | | | | | | | | | | | Pathogens | | | | | | | | | | | | | | | | | | |
|---------|------------------|---------------------------------|-------------|--------------|-------------|-----------|-----------------|---------|----------------|-------------|------------|--------------------|---------------------|---------------------|-----------------|-----------|------------|-------------|--------------|---------------|-----------|----------|-------------------|------------|----------|------------------|---------------|------------------|-------------------|------|-----------|---------------|--|--|
| | | Bacteroides | B. fragilis | Enterococcus | E. faecalis | Roseburia | R. intestinalis | E. coli | F. prausnitzii | Veillonella | B. bifidum | Lactococcus lactis | Lactob. delbrueckii | Lactob. acidophilus | Eubact. biforme | Atopobium | A. minutum | Clostridium | C. difficile | Campylobacter | C. jejuni | Yersinia | Y. enterocolitica | Salmonella | Listeria | L. monocytogenes | Mycobacterium | M. avium complex | M. avium paratub. | EHEC | Aeromonas | A. hydrophila | | |
| CV1 | conventional pig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CV2 | conventional pig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| PD | pig donor | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P2d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P3d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P4d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P5d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P6d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P7d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P8d14 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| HD | human donor | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H1d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H2d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H3d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H4d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H5d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H6d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H7d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H8d14 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Tab. 3.19: Identified pathogens and background flora in piglets with pig-associated or human-associated intestinal flora suffering from diarrhea and conventionally raised piglets 21 days after birth. The donors of the intestinal flora are also listed. The species were identified after PCR amplification of the DNA, hybridization to the microarray and acquisition of the silver deposition (green squares).

| piglets | | Background flora and probiotics | | | | | | | | | | | | | | Pathogens | | | | | | | | | | | | | | | | | | |
|---------|------------------|---------------------------------|-------------|--------------|-------------|-----------|-----------------|---------|----------------|-------------|------------|--------------------|---------------------|---------------------|-----------------|-----------|------------|-------------|--------------|---------------|-----------|----------|-------------------|------------|----------|------------------|---------------|------------------|-------------------|------|-----------|---------------|--|--|
| | | Bacteroides | B. fragilis | Enterococcus | E. faecalis | Roseburia | R. intestinalis | E. coli | F. prausnitzii | Veillonella | B. bifidum | Lactococcus lactis | Lactob. delbrueckii | Lactob. acidophilus | Eubact. biforme | Atopobium | A. minutum | Clostridium | C. difficile | Campylobacter | C. jejuni | Yersinia | Y. enterocolitica | Salmonella | Listeria | L. monocytogenes | Mycobacterium | M. avium complex | M. avium paratub. | EHEC | Aeromonas | A. hydrophila | | |
| CV1 | conventional pig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| CV2 | conventional pig | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| PD | pig donor | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P2d21 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P3d21 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P4d21 | died | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P5d21 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P6d21 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P7d21 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| P8d21 | PFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| HD | human donor | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H1d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H2d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H3d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H4d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H5d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H6d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H7d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| H8d21 | HFA piglet | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

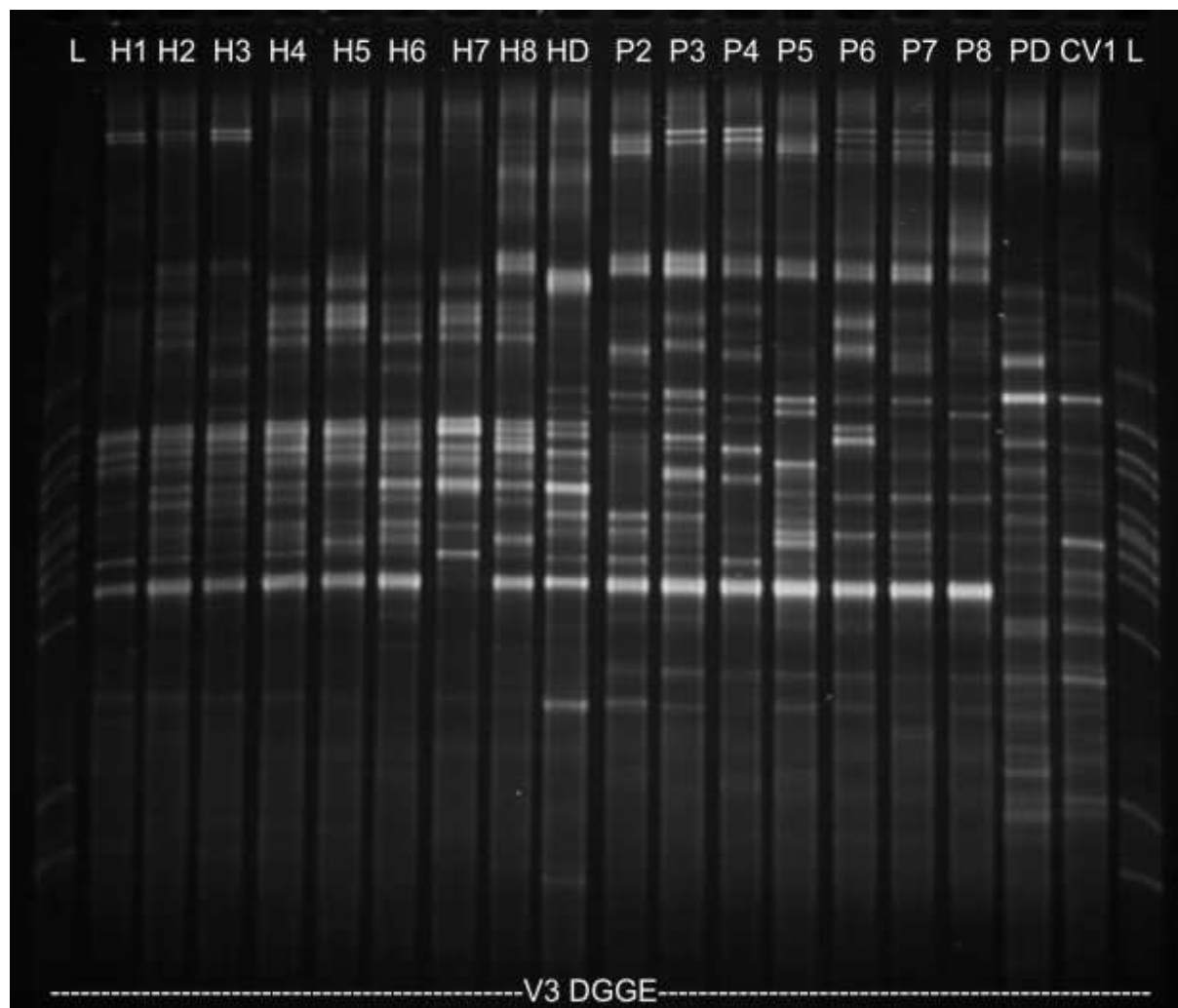


Fig. 3.28: Denaturing gradient gel electrophoresis of the V3 fragment of intestinal bacteria from human (H) and pig (P) flora-associated piglets, human and pig flora donor (HD, PD) and conventionally raised pig (CV) at day 14 after birth in a polyacrylamide gel with an urea gradient from 27% to 52%. L – V3 ladder

Multivariate analysis (chap. 2.2.13) was applied to the microarray data of the intestinal flora to investigate whether the bacterial composition of conventionally raised piglets, PFA piglets, and HFA piglets differ significantly and whether the microbiota of the donors had established in the recipients. The samples were first assigned to two groups comprising CV, PD, and PFA in one group ($y=0$) and HD and HFA in another group ($y=1$). A second analysis was performed under the assumption of three independent groups comprising the CV piglets in one group ($y=0$), the PD and PFA piglets in the second group ($y=1$), and the HD and HFA piglets in the third group ($y=2$).

Fourteen days after birth of the piglets, the PCA score plot (Fig. 3.29) of faecal microbiota did not clearly cluster in two or three separate groups. In the two-group analysis the principle components PC2 and PC7 discriminated the groups best, but explained only little variance of the data (PC2=26.9%; PC7=1.9%). The human donor was located close to the pig donor and the CV piglets. Piglet P5 was located in the HFA group being most similar to H1.

In the PLS plot (Fig. 3.30) the separation of both groups (CV+PD+PFA vs. HD+HFA) was more obvious. According to leave-one-out cross-validation (LOOCV) 84% of samples were correctly classified if one PLS component was used for the PLS model. A PLS model with two components yielded 90% accuracy. Using the one-component PLS model at a significance of $p=0.05$, four microarray targets were identified as the most important in discriminating healthy from infected children: *Fusobacterium prausnitzii*, *Veillonella* spp.,

Bifidobacterium bifidum, and *Lactobacillus delbrueckii* (Fig. 3.31). Positive regression coefficients correlated with the human microbiota, while negative ones correlated with the pig microbiota. *Fusobacterium prausnitzii* was detected in the CV piglets, the pig donor, in five out of seven PFA piglets, and in the human donor but in none of the HFA piglets. *Veillonella* spp. was detected only in the CV piglets. *B. bifidum* was found only in the human donor and one HFA piglet. *L. delbrueckii* could be identified in six out of seven PFA piglets. One-way ANOVA (Tab. 3.20) showed that the difference of *L. delbrueckii* ($p=2.9 \times 10^{-3}$) and *F. prausnitzii* ($p=1.1 \times 10^{-3}$) between the two groups was significant ($p < 0.05$), whereas the difference of *Veillonella* spp. ($p=0.174$), *B. bifidum* ($p=0.128$), and all other variables was not.

Remodelling the PLS model with only *F. prausnitzii* and *L. delbrueckii* resulted in a one-component model with 89% accuracy and a two-component model with 84% accuracy. Using *F. prausnitzii*, *L. delbrueckii*, *Veillonella* spp., and *B. bifidum* for the model did not improve accuracy. The two-component model provided even lower correct classification.

A two-component PLS model allowed also separation of the CV piglets from HFA and PFA piglets with their donors (three-group analysis) with 79% accuracy. The pig donor and PFA piglets were located between the CV piglets and the HFA piglets with their donor. A one-component model would reach 73% correct classification. Since one group, the CV piglets, comprised only two samples, the significant variables were not calculated using PLSA. Another method, fold cross-validation, would be more useful for determination of the significant variables, but could not be performed with the current program. In one-way ANOVA, *F. prausnitzii* ($p=4.2 \times 10^{-3}$), *L. delbrueckii* ($p=3.2 \times 10^{-4}$), and *E. faecalis* ($p=3.1 \times 10^{-2}$) were identified as the significant variables.

Twenty-one days after birth of the piglets, the PCA score plot (Fig. 3.29) of faecal microbiota now assembled in two separate groups. In the two-group analysis, the principle components PC1 and PC3 clearly discriminated the HFA piglets and their donor from the PFA piglets, the pig donor, and the CV piglets. Moreover, 63.6% of variance of the data was explained by these two principle components (PC1=48.6%; PC3=14.8%).

The PLS analysis (Fig. 3.30) showed that a one-component PLS model was able to separate the HD and HFA piglets from all other samples, but separation of the CV piglets in the three-group analysis was not observed. They clearly clustered with the PFA group and the pig donor. The one-component model in the two-group analysis reached 100% correct classification. Five significant variables were identified (Fig. 3.31), which contributed most to group separation: *Bacteroides fragilis*, *Enterococcus faecalis*, *Escherichia coli*, *Bifidobacterium bifidum*, and *Lactobacillus delbrueckii*. Positive regression coefficients correlated with the human microbiota, while negative ones correlated with the pig microbiota. *Bacteroides fragilis* and *Bifidobacterium bifidum* were found in the human donor and seven out of eight HFA piglets. *Enterococcus faecalis* was identified in all HFA piglets and two PFA piglets. *Escherichia coli* was detected in the human donor, all HFA piglets, one CV piglet, and three PFA piglets. In contrast, *L. delbrueckii* was present only in three out of six PFA piglets. In one-way ANOVA (Tab. 3.20), this result was supported for all variables besides for *L. delbrueckii* ($p=0.063$).

A re-calculated PLS model based on the five identified significant variables and using one PLS component reached 100% correct classification as well. However, this was also the case for a model based on the four significant variables found by one-way ANOVA.

For three-group comparison, one-way ANOVA found *B. fragilis*, *B. bifidum*, and *E. faecalis* as significant variables.

Thus, both intestinal microbiotas had established in the recipients after inoculation and were similar to the donors. The human flora had stabilized a little faster than the pig flora, which correlated with the severity of disease symptoms in the individuals of the trial.

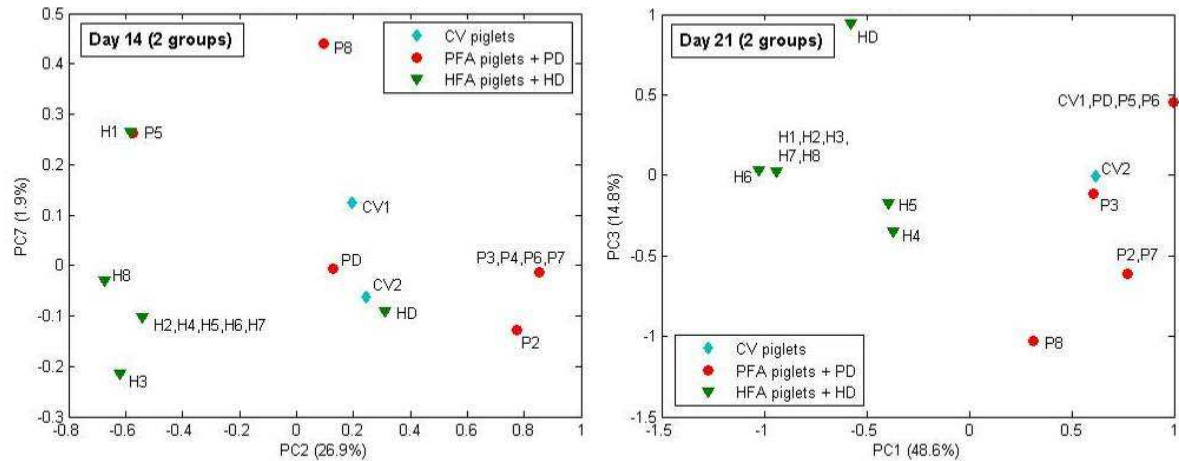


Fig. 3.29: Principle component analysis (two-group classification) of the results from microarray hybridization of faecal samples from conventionally raised (CV) piglets, pig flora-associated (PFA) piglets, the donor of pig microbiota (PD), human flora-associated (HFA) piglets, and the donor of human microbiota (HD) taken at day 14 (left) and day 21 (right) after birth.

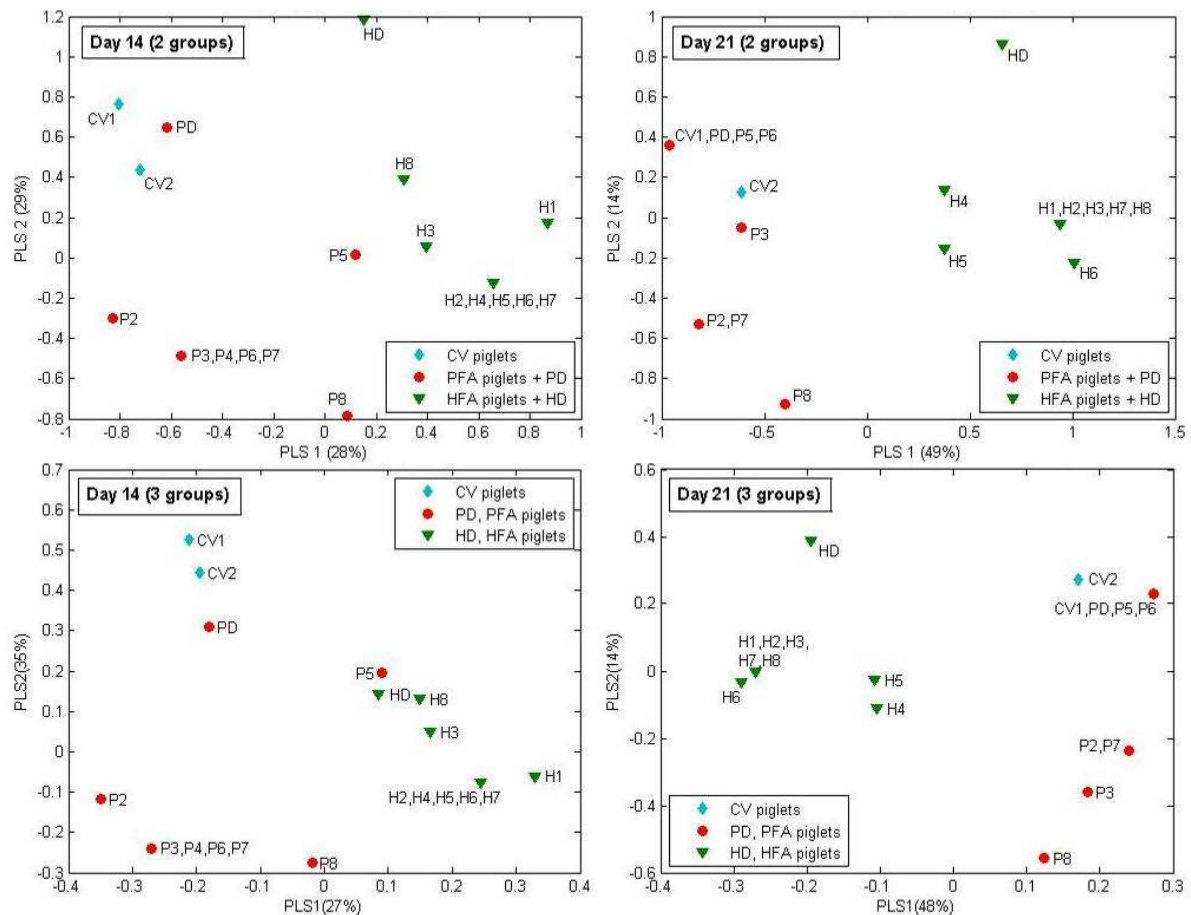


Fig. 3.30: Partial least squares analysis of the results from microarray hybridization of faecal samples from conventionally raised piglets (CV), pig flora-associated (PFA) piglets, the donor of pig microbiota (PD), human flora-associated (HFA) piglets, and the donor of human microbiota (HD) taken at day 14 (left) and day 21 (right) after birth. (top) two-group classification, (bottom) three-group classification.

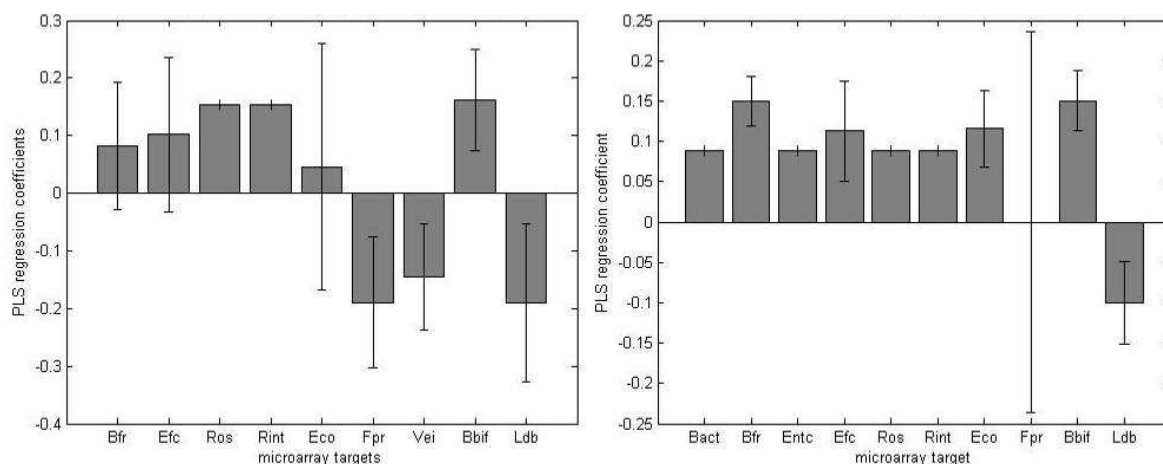


Fig. 3.31: PLS regression coefficients plots of microarray-detected species or genera in the PFA/HFA trial at (left) day 14 and (right) day 21 when one PLS component was used ($p=0.05$). The bar indicates the uncertainty limits.

Tab. 3.20: One-way ANOVA calculated p-values of the microarray results for the intestinal background flora giving the significance of differentiation between the two or three groups of conventionally raised piglets (CV), and/or pig flora-associated (PFA) piglets + the donor of pig microbiota (PD), and human flora-associated (HFA) piglets + the donor of human microbiota (HD). The ANOVA p-values of species identified as significant variables by PLSA are given in bold.

| Variable | Microarray target | Day 14 | | Day 21 | |
|----------|--------------------------|--|----------------------|--|----------------------|
| | | p (2 groups) | p (3 groups) | p (2 groups) | p (3 groups) |
| 1 | <i>Bacteroides</i> spp. | n.d. | n.d. | 0.332 | 0.483 |
| 2 | <i>B. fragilis</i> | 0.233 | 0.131 | 5.5×10^{-7} | 5.7×10^{-6} |
| 3 | <i>Enterococcus</i> spp. | n.d. | n.d. | 0.332 | 0.635 |
| 4 | <i>E. faecalis</i> | 0.171 | 3.1×10^{-2} | 2.3×10^{-3} | 7.6×10^{-3} |
| 5 | <i>Roseburia</i> spp. | 0.305 | 0.601 | 0.332 | 0.635 |
| 6 | <i>R. intestinalis</i> | 0.305 | 0.601 | 0.332 | 0.635 |
| 7 | <i>E. coli</i> | 0.620 | 0.416 | 6.0×10^{-3} | 0.026 |
| 8 | <i>F. prausnitzii</i> | 1.1×10^{-3} | 4.2×10^{-3} | n.d. | 0.874 |
| 9 | <i>Veillonella</i> spp. | 0.174 | n.d. | n.d. | n.d. |
| 10 | <i>B. bifidum</i> | 0.128 | 0.326 | 5.5×10^{-7} | 5.7×10^{-6} |
| 11 | <i>L. lactis</i> | n.d. | n.d. | n.d. | n.d. |
| 12 | <i>L. delbrueckii</i> | 2.9×10^{-3} | 3.2×10^{-4} | 0.063 | 0.059 |
| 13 | <i>L. acidophilus</i> | n.d. | n.d. | n.d. | n.d. |
| 14 | <i>E. bifforme</i> | n.d. | n.d. | n.d. | n.d. |
| 15 | <i>Atopobium</i> spp. | n.d. | n.d. | n.d. | n.d. |
| 16 | <i>A. minutum</i> | n.d. | n.d. | n.d. | n.d. |

n.d. – not determined (due to mathematical reasons, e.g. only zeros in groups or groups identical)

3.12 DNA target quantification with DNA microarrays

Preliminary experiments were carried out by Barbara Hörmann as part of her student research project (Hörmann 2006). The influence of a probe-bound quencher (BHQ2) on target-coupled Cyanine 3 was investigated. In brief, a microarray containing a series of spots with a constant amount of a probe mixture containing the positive hybridization control (PHC) and a 3'-black hole quencher-labelled positive hybridization control (3'-BHQ2-PHC) in a variable ratio of 20 μM + 0 μM , 18 μM + 2 μM , 16 μM + 4 μM up to 0 μM + 20 μM was spotted on epoxy-coated glass slides (Nexterion) (Fig. 2.4 A), covalently immobilized and the surface was blocked. Increasing amounts of the Cy3-labelled complementary strand of the PHC (Cy3-Comp) of 0, 0.5, 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 fmol were hybridized to the slides. When the spots without quencher (20 μM PHC) were analyzed for their fluorescence intensity, the Cy3 signal reached saturation at around 60 fmol target applied to the array surface, which corresponds to $\sim 50,000$ FU (Fig. 3.32). This saturation was probably due to the limitation of the linear range of the scanner and not the limitation of available binding partners. The scanner had a linear range between $\sim 10,000$ and 55,000 FU. When the hybridization signals of all spots with increasing quencher concentration were analyzed after hybridization of 50 fmol Cy3-coupled target at PMT40, a rapid decrease in the hybridization signal was observed (Fig. 3.32). With 10% quencher-coupled probe in the spot the signal decreased by 60%. With 100% quencher-coupled probe in the spot, the signal still accounted for 2-3% of the hybridization response measured in the spots without quencher. This was found also for the other target concentrations as far as the signal could be detected in the linear range of the scanner (data not shown). A plot of the signal intensities versus the applied target concentration for each quencher concentration in the spot seemed to show a linear dependence, if the signals did not reach saturation nor kept below 10,000 FU (not shown). However, the linear fit revealed regression coefficients between 0.58 and 0.93, if at least four data points were between 10,000 and 55,000 FU to be used for regression. Nevertheless, it was supposed, that this was due to variability in the array spotting and hybridization process. Between 10 and 100 fmol target concentration, it was possible to detect the different amounts of DNA at one PMT in the linear range of the scanner using different spots with increasing amount of quencher in the spot.

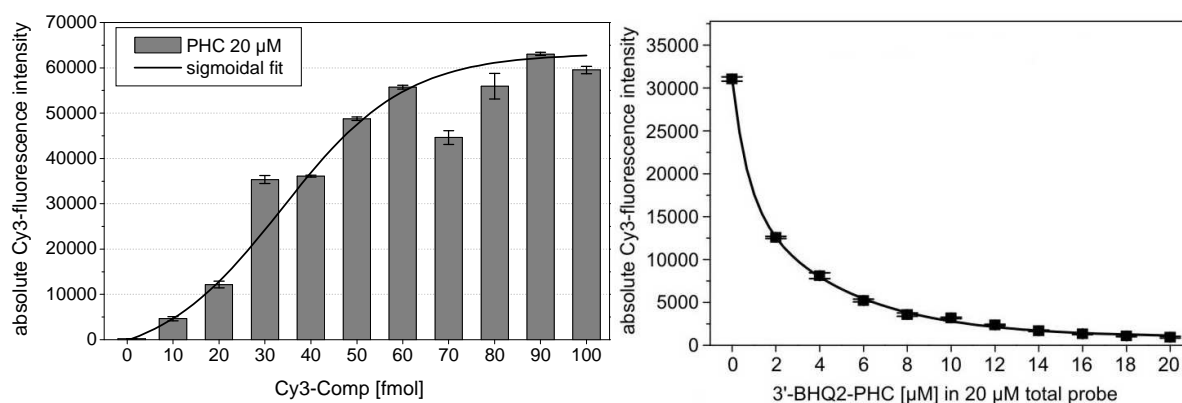


Fig. 3.32: (left) Hybridization of increasing amounts of Cy3-Comp to the complementary probe PHC. Probe PHC was spotted with a concentration of 20 μM . The signals were scanned at PMT43. Sigmoidal fit of the hybridization signals using the Boltzmann model. (right) Hybridization of 50 fmol Cy3-Comp to the complementary probe PHC containing variable amounts of 3'-BHQ2 quencher (PMT 40). [from (Hörmann 2006)]

To investigate whether the Cy3 signal decrease in spots with higher quencher concentration was due to a real quenching effect or only due to steric hindrance of the hybridization, another experiment with the quencher coupled to the target and the fluorophore coupled to the probe was performed by Hörmann. While the fluorophore was bound to the 3'end of the

probe, which was surface immobilized via the 5' end, the quencher was 3'- or 5'-coupled to the target to analyze the influence of the fluorophore-quencher distance.

In brief, 20 μM 3'-Cy3-labelled positive hybridization control (Cy3-PHC) was spotted to the array surface, which was then hybridized with 0, 0.5, 1, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 fmol 3'-BHQ2-Comp or 5'-BHQ2-Comp. With increasing amount of quencher-target in hybridization, the signals of the spots decreased (Fig. 3.33). The spots in the negative control (0 fmol target) had an average signal intensity of $38,412 \pm 2,915$ FU. The signal after hybridization with 100 fmol 3'-BHQ2-Comp was $23,249 \pm 11,578$ FU and with 100 fmol 5'-BHQ2-Comp $18,118 \pm 5,999$ FU, which accounts for 61% and 47% of the maximum signal. The very high standard deviation resulted from inter-array variations on one slide, which was probably a scanner problem. If only one array comprising ten spots was analyzed, the result was $37,537 \pm 1,554$ FU for the negative control and $19,375 \pm 875$ FU or $24,112 \pm 552$ FU for 3'-BHQ2-Comp or 5'-BHQ2-Comp, respectively (data not shown). This accounts for 52% and 64% of the maximum signal. However, a linear fit of the Cy3-signal decrease upon hybridization of increasing amounts of target with coupled quencher had a regression coefficient of 0.71 and 0.87, respectively, if all arrays were analyzed despite the high standard deviation. If only one array was quantified, the result was 0.79 and 0.83, respectively. For both targets, the linear dependence seemed to be likely. The calculated signal after hybridization with 100 fmol 3'-BHQ2-Comp or 5'-BHQ2-Comp by using the linear fit equation accounted for 47% and 52% of the maximum signal. A difference in the quenching efficiency was not found between the two different targets. The quenching of the Cy3 immobilized in the spot was not complete upon hybridization of 100 fmol target.

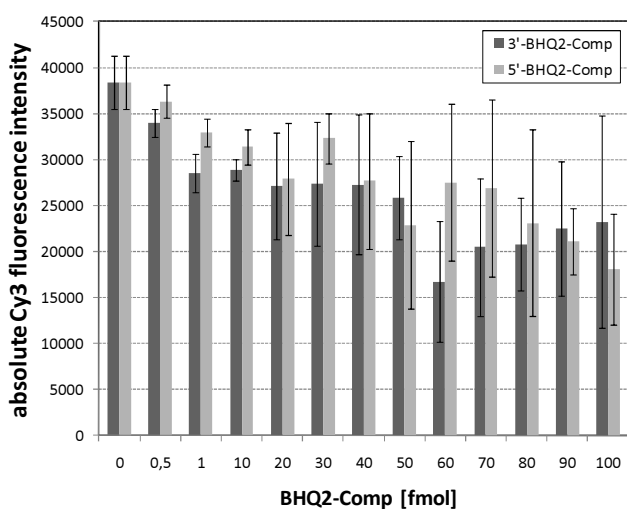


Fig. 3.33: Quenching effect of the target-bound BHQ2-quencher (5'-BHQ2-Comp or 3'-BHQ2-Comp) on the PHC probe-immobilized Cy3 upon hybridization on a microarray [from (Hörmann 2006)].

3.12.1 Hybridization of PCR product to *E. coli*-specific probes with a quencher

The different amounts of *E. coli*-DNA corresponding to $2 \cdot 10^6$, 10^6 , 10^5 , 10^4 , 10^3 , 10^2 , 10, 1, and 0 genome equivalents were successfully amplified and Cy3-labelled by PCR (chap. 2.2.14.1). The highest amount of $2 \cdot 10^6$ genome equivalents corresponded to the maximum amount of DNA applied in PCR, which was 10 ng. The PCR products were detectable by agarose gel electrophoresis in all samples down to 10^4 genome equivalents (Fig. 3.34). This amount of template DNA corresponds to 50 pg genomic DNA (*E. coli* 5.085 fg/genome). The DNA and Cy3 concentration could not be measured in all PCR products, as it was below the detection limit (Tab. 3.21). Although, six samples always contained the same amount of template DNA in PCR the resulting PCR product concentration was different. The average DNA concentration for $2 \cdot 10^6$ genome equivalents as template was 25.3 ± 5.4 ng/ μL , for 10^6 genome equivalents 18.0 ± 6.5 ng/ μL , and for 10^5 genome equivalents 8.8 ± 0.4 ng/ μL . This is a standard deviation of 21%, 36%, and 4.8%, respectively. Below 10^5 genome equivalents

in PCR, the PCR products were not spectrophotometrically quantifiable. Cyanine 3 was also not detectable in the samples with template less than 10^5 genome equivalents and even not in all samples with more template DNA. The average incorporation rate was 257 ± 48 for $2 \cdot 10^6$ genome equivalents, and 313 ± 179 for 10^6 . In the samples containing 10^5 genome equivalents as template DNA, only one had a detectable Cy3 concentration with a resulting incorporation rate of 276.

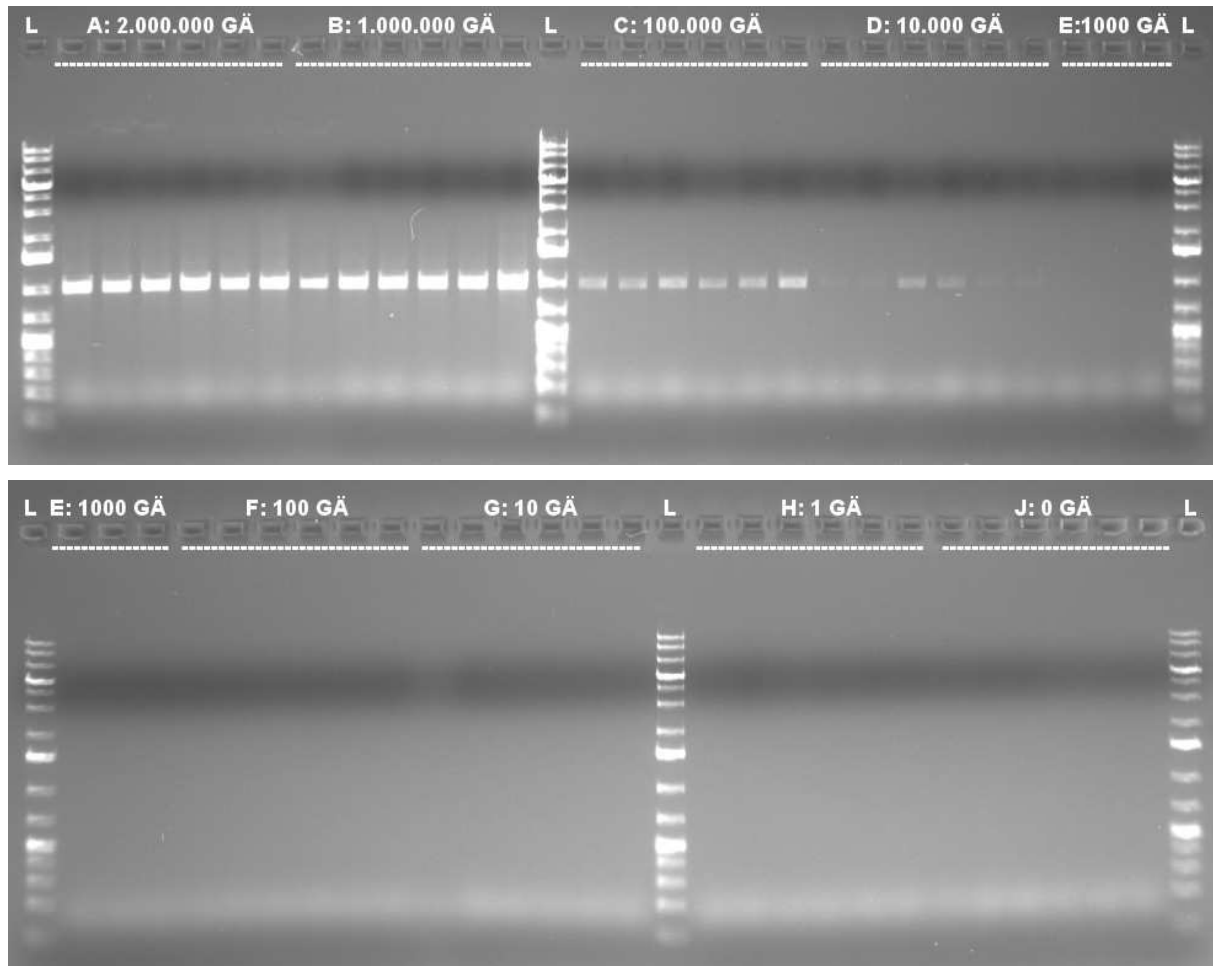


Fig. 3.34: Agarose gel analysis of the PCR products using one primer pair (114V/1084R) to amplify the DE fragment (23S gene) from different amounts of *E. coli* DNA ranging from $2 \cdot 10^6$ to 0 genome equivalents (GÄ). L - ladder.

Tab. 3.21: DNA concentration and Cy3 incorporation rate of the amplified *E. coli* DNA.

| samples | [DNA] ng/ μ L | | | | | | Cy3 incorporation NT/F | | | | | |
|-----------|-------------------|------|------|------|------|------|------------------------|-----|------|------|------|------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 1 | 2 | 3 | 4 | 5 | 6 |
| 2,000,000 | 23.1 | 28.5 | 24.8 | 17.3 | 24.5 | 33.6 | 237 | 216 | 251 | n.d. | 247 | 339 |
| 1,000,000 | 13.2 | 19.1 | 10.3 | 19.1 | 17.2 | 29.1 | n.d. | 579 | n.d. | 193 | 261 | 220 |
| 100,000 | 8.6 | 9.1 | 9.5 | 8.3 | 8.9 | 8.6 | n.d. | 276 | n.d. | n.d. | n.d. | n.d. |

(n.d.) not determined

The PCR products were then hybridized to microarrays after or without prior DNase I digestion (chap. 2.2.14.1). In both cases, the sense probes displayed decreasing fluorescence signals upon increasing quencher concentration in the spots, but the signals did not reduce to zero, which was visible when the arrays were scanned at higher enhancing factors (PMT) (Fig. 3.35, Fig. 3.36). At PMT 60, 65 and 70 some signals were saturated by means of scanner properties. On the contrary, the hybridization response of the antisense probes decreased to zero upon increasing quencher amount until 50% quencher-probe in the spots. Only, the probe Ecoli23S_542_18i did not hybridize with undigested DNA, if no quencher was in the spot, but after digestion, this probe also displayed a strong hybridization signal. Unexpectedly, the antisense probes showed an increasing fluorescence signal upon hybridization, if the spotted solution contained more than 12 μM probe with coupled quencher in the 20 μM spotting solution. The absolute fluorescence signals of the Ecoli23S_542_18/i probes were higher when digested PCR product was hybridized, while for the Ecoli23S_271_18B/i probes it was contradictory.

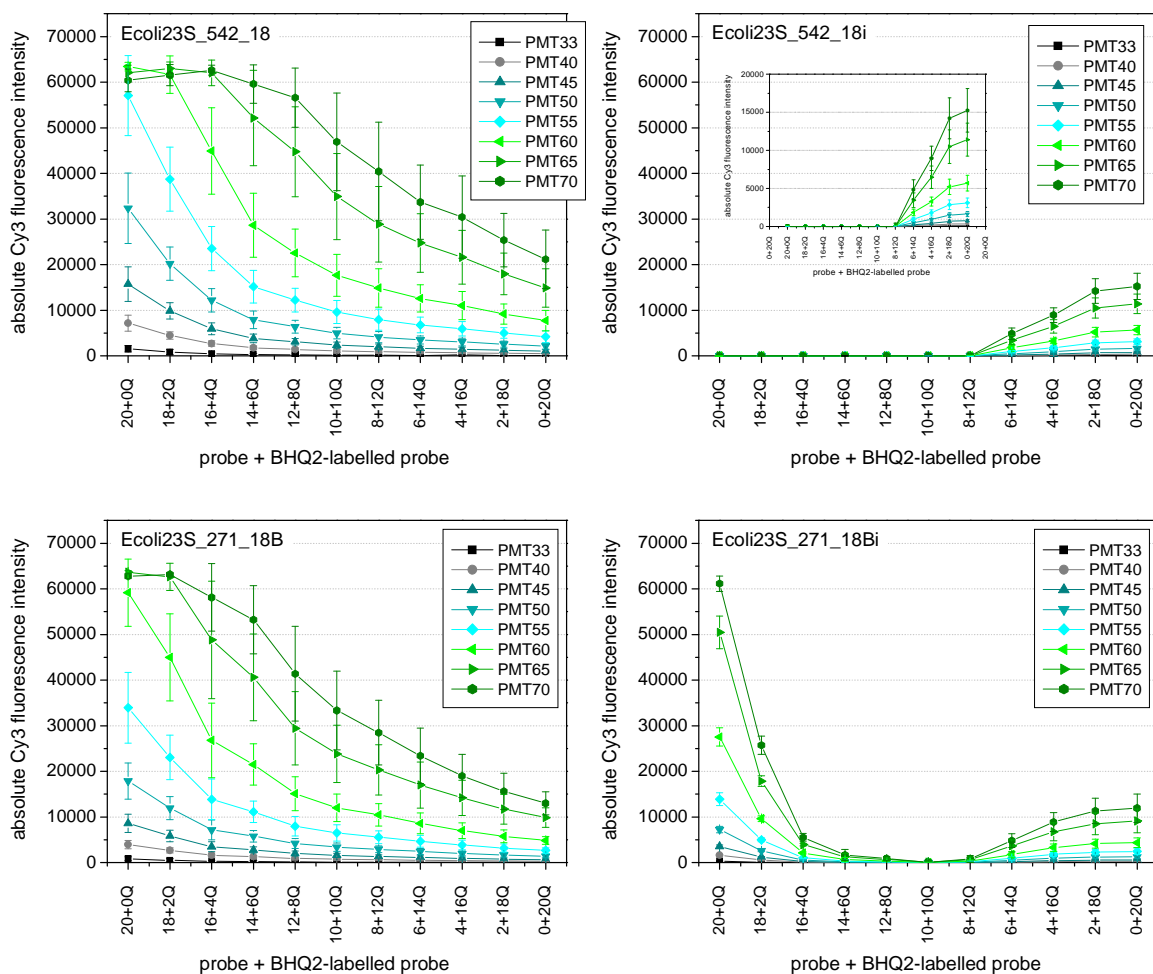


Fig. 3.35: Hybridization response of undigested PCR product derived from 2×10^6 genome equivalents of *E. coli* with spots containing a complementary probe and different amounts of black hole quencher (BHQ2) at different enhancing steps by photo multiplier tube (PMT). Each plot represents the results for one probe sequence.

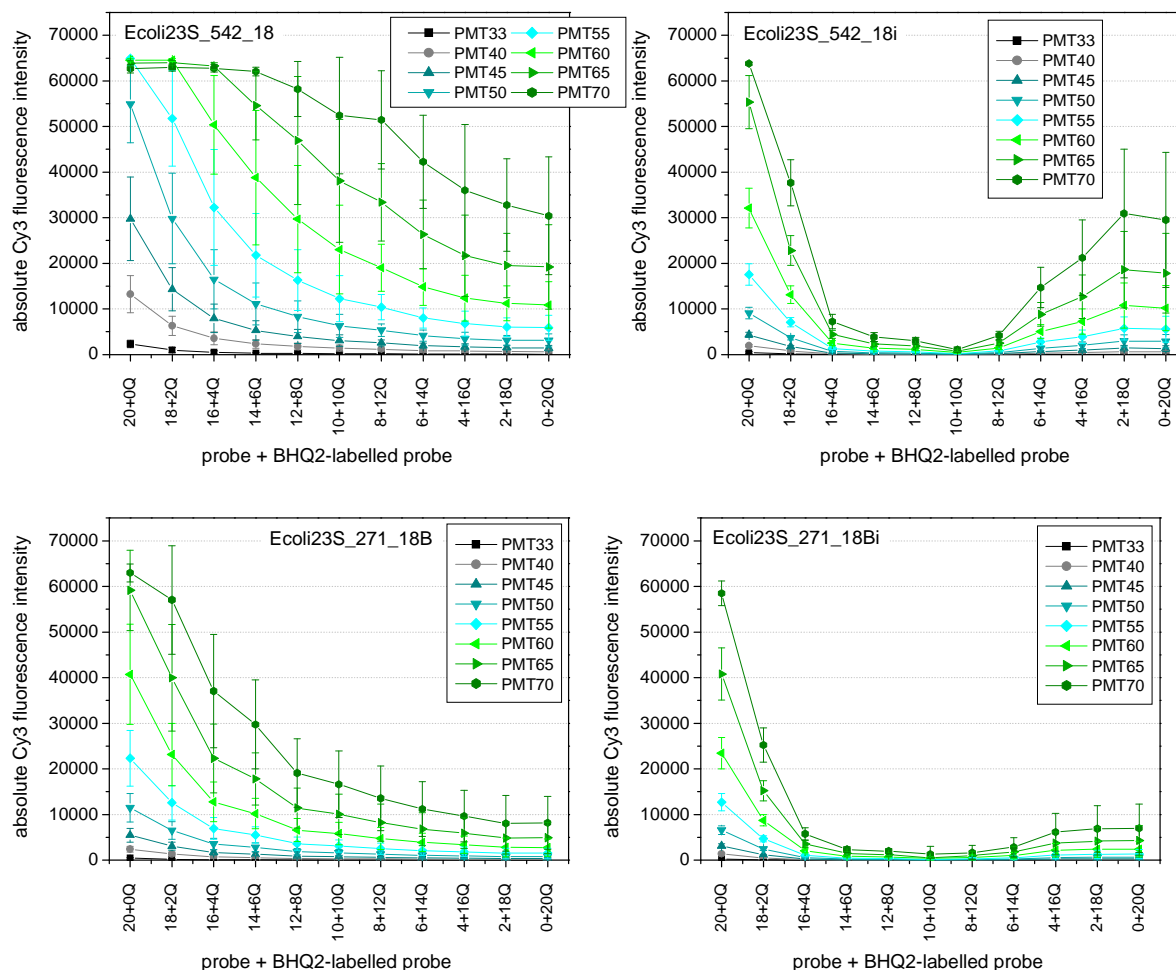


Fig. 3.36: Hybridization response of DNase I-digested PCR product derived from $2 \cdot 10^6$ genome equivalents of *E. coli* with spots containing a complementary probe and different amounts of black hole quencher (BHQ2) at different enhancing steps by photo multiplier tube (PMT). Each plot represents the results for one probe sequence.

Fig. 3.37 shows the hybridization result of all different amounts of genome equivalents applied to PCR with the complementary probes containing increasing amounts of quencher at PMT 65. The PCR product was not digested. Depending on the amount of applied DNA, the signal height changes. Lower amounts of DNA result in less hybridization signal intensities. This was the same for digested DNA shown in Fig. 3.38. As described for Fig. 3.35 and Fig. 3.36 the fluorescence signals decrease upon increasing quencher concentration in the spots, except for the antisense probes containing more than 50% quencher-probe in the spots.

In Fig. 3.39 the fluorescence signals of the spots containing different amounts of quencher are plotted against the amount of *E. coli*-DNA used in PCR to be hybridized to the array after DNase I digestion. The arrays were scanned at PMT 65. The detection limit in this experiment was 10^4 genome equivalents applied to PCR. However, the Cy3-signals for 10^4 genome equivalents at PMT 65 with the spots containing no quencher (20+0Q) was $5,766 \pm 3,244$ for probe Ecoli23S_542_18, $1,707 \pm 1,233$ for probe Ecoli23S_542_18i, $2,658 \pm 1691$ for probe Ecoli23S_271_18B, and $4,116 \pm 2,468$ for probe Ecoli23S_271_18Bi. The high standard deviation for these fluorescence signals has its seeds in the proximate spotting control, which was slightly overdosed and outshining the neighbouring spots by increasing the background signal, which is subtracted from the raw spot intensity. The absolute signals suggest that about 1/3 to 1/5 of DNA would have been detectable as well. That means, $2 \cdot 10^3$ to $3.3 \cdot 10^3$ genome equivalents would have been detectable, although 10^3 genome equivalents displayed to low signals to be identified as *E. coli*.

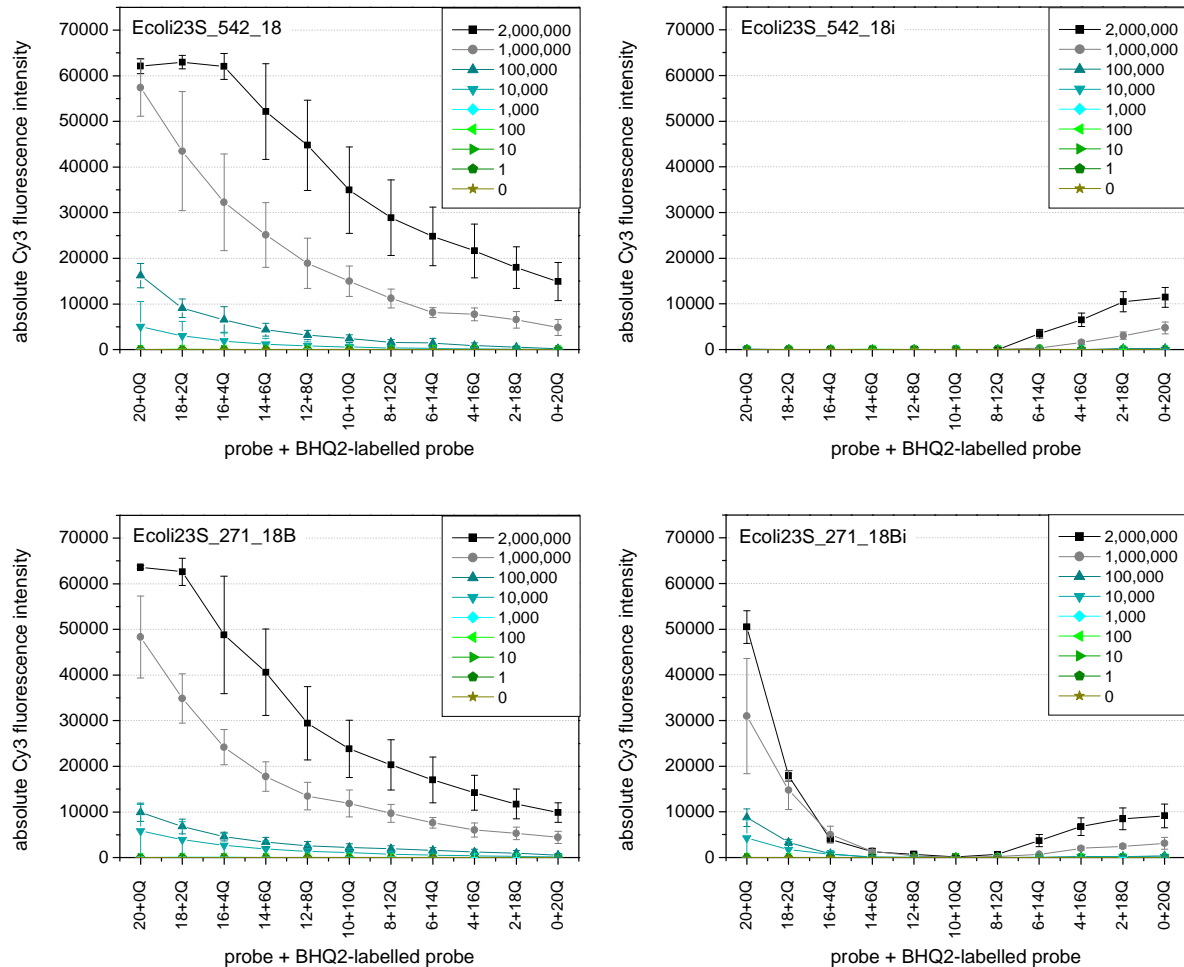


Fig. 3.37: Hybridization response of undigested PCR product derived from different amounts of genome equivalents of *E. coli* with spots containing a complementary probe and different amounts of black hole quencher (BHQ2) at PMT 65. Each plot represents the results for one probe sequence.

In this experiments it was shown, that a spot-immobilized quencher can fully quench the hybridization signal of a labeled, hybridizing probe. This effect was independent from DNase I digestion of the PCR product before hybridization. However, a second effect was observed, which resulted in returning signals at higher quencher concentrations in the spot indicating an unspecific binding of DNA target to the quencher. This effect also overlaid the quenching in some probes.

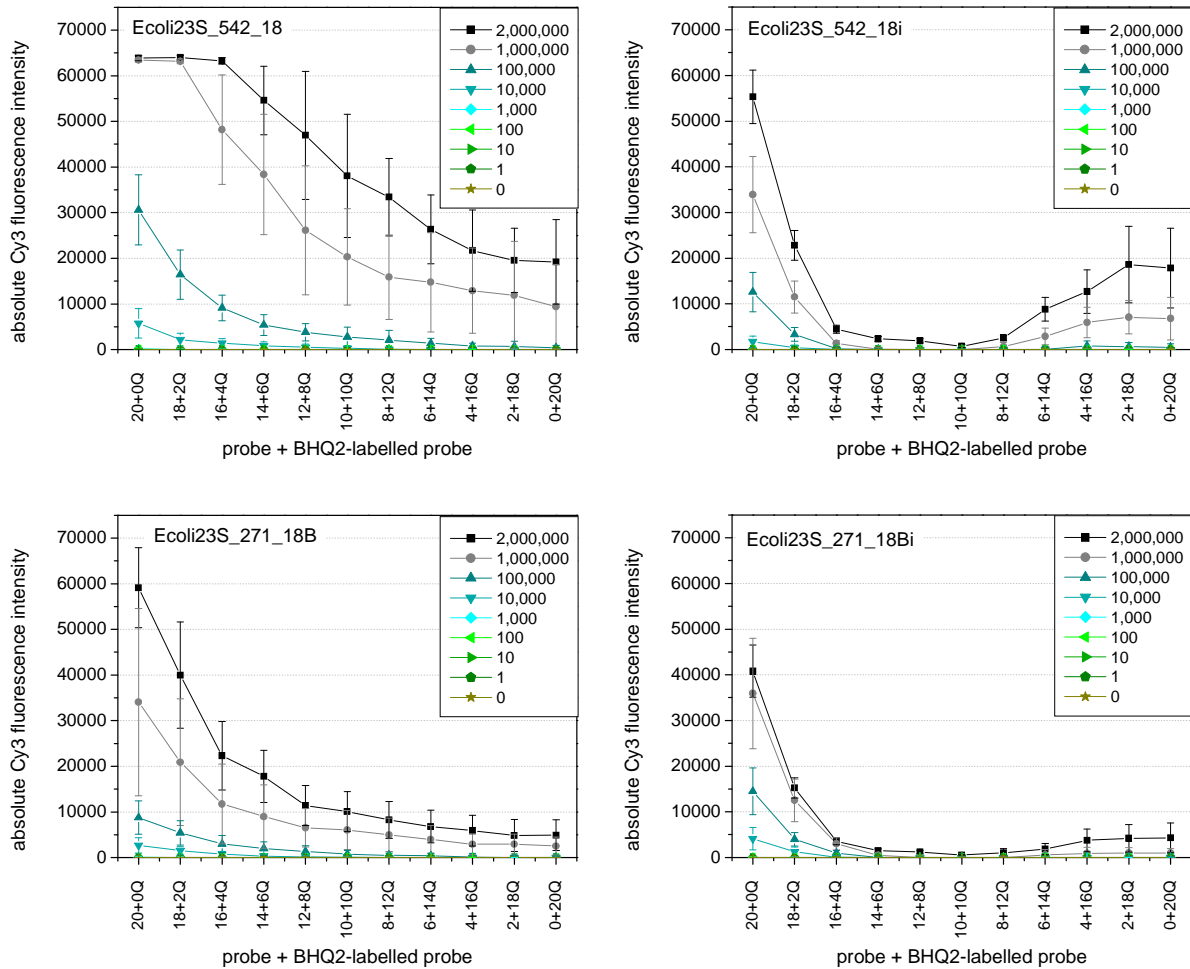


Fig. 3.38: Hybridization response of DNase I-digested PCR product derived from different amounts of genome equivalents of *E. coli* with spots containing a complementary probe and different amounts of black hole quencher (BHQ2) at PMT 65 (enhancing step by photo multiplier tube). Each plot represents the results for one probe sequence.

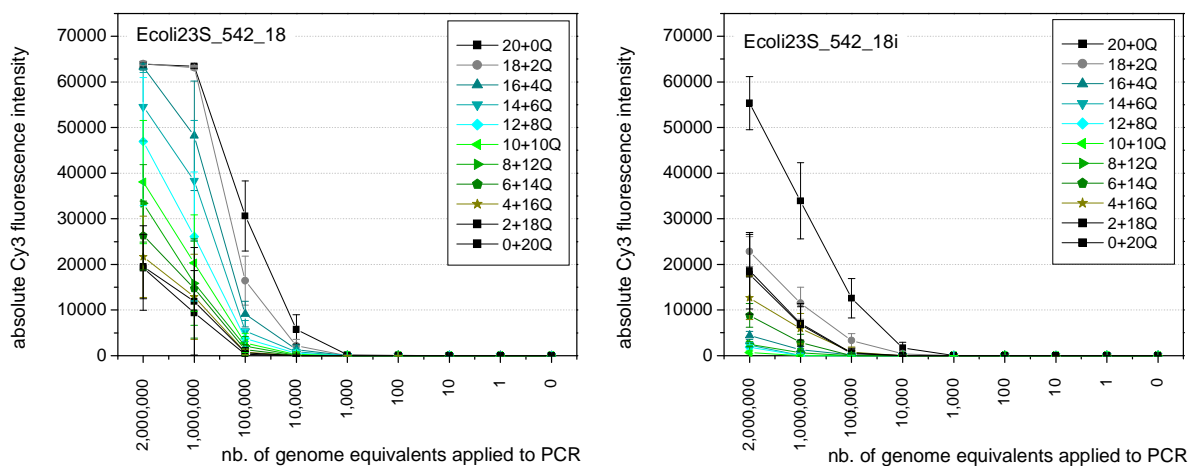


Fig. 3.39: Hybridization response of DNase I-digested PCR product derived from different amounts of genome equivalents of *E. coli* with spots containing a complementary probe and different amounts of black hole quencher (BHQ2) at PMT 65 (enhancing step by photo multiplier tube). Each plot represents the results for one probe sequence.

4 Discussion

Infections with bacterial, intestinal pathogens are an international medical problem, which cause high expenses in the industrialized countries and are a leading cause of childhood morbidity and mortality in developing countries (WHO 2009). The identification of the causative agent is of major concern regarding the choice of therapy, the clinical hygiene management, the patient's confidence in the clinician, and acceptance of and satisfaction with the medical treatment. Currently, it is mainly based on phenotypic classification using stool culture and a set of morphological, physiological, and serological tests (chap. 1.4.1). This time-consuming, complex procedure does not meet the requirements of an evidence-based therapy and a cost effective medical care system. Three problems had to be solved in this work, to improve the situation: (I) the time required for pathogen identification had to be reduced, (II) the coverage of possible pathogens, which is tested for, had to be increased, and (III) the pathogen should be quantified over several orders of magnitude. The second contributes to solve the first problem of required diagnostic time. It reduces the likelihood that a sample is negative for all analyzed pathogens, which would require a new sample to be taken and investigated for other pathogens. The aim of this work was to develop a genotypic method to identify the most relevant gastroenteritic pathogens in parallel and short time. Therefore, a microarray was developed that is able to detect these intestinal, bacterial pathogens. In addition, the most predominant intestinal residents and some lactic acid bacteria (LAB) that can serve as probiotics are detectable with the microarray. In case of therapy with probiotics, this can help the clinician to follow the course of therapy and the recovery of the intestinal flora. The array was validated with clinical samples and the sensitivity was determined. To solve the problem of reliable quantification of intestinal bacteria and pathogens with a concentration range over several orders of magnitude on microarrays, a system using a quencher was evaluated.

4.1 Genotyping of intestinal pathogens

Genotyping methods can fulfil the needs of medical facilities for fast and easy pathogen identification. They have several advantages towards phenotyping methods.

(I) The main advantage of this approach is the abolition of the necessity to cultivate the pathogen before identification. Generally, DNA is directly isolated from a faecal sample and is introduced to the further assay process. This solves several problems regarding cultivation of pathogens. Intestinal pathogens require anaerobic cultivation conditions and some, such as *Vibrio* species and *Yersinia enterocolitica*, are difficult to culture necessitating special media or a lasting cold enrichment. Moreover, the sample transport and storage strongly influences the vitality of the bacteria and thus the cultivability. Wrong sample handling and demanding bacteria increase the risk of failed pathogen diagnosis. Ethelberg et al. analyzed data of 620,000 faecal samples of a central laboratory in Denmark from 1995 to 2003 for the significance of the number of submitted samples and patient-related factors for faecal bacterial diagnostics (Ethelberg 2007). They found that only 15.3% of the sample-series, that is samples taken from one patient on the same day, were positive for any pathogen. The analysis of more than one sample from one day improved the sensitivity of faecal culture by 20% for each additional sample. This illustrates the relatively high failure rate of classical faecal analysis using culturing technique. The study also showed that the probability of culturing bacteria from samples decreased within seven days by 3-7%. This means, in case of a failed diagnosis for a common pathogen, the investigation of a newly taken sample few days later will have less success.

(II) Furthermore, genotyping methods allow pathogen detection in shorter time, which is a consequence of the abolition of the cultivation, and can cover a broad spectrum of potential

pathogens in one assay. By this, these methods will reduce the need for more than one sample. Even if a second sample is needed, it can be taken on the same day.

(III) Moreover, co-infections with a second pathogen can be detected more likely, if a sample is analyzed for a broad bacterial spectrum at once.

(IV) Another advantage of genotyping methods is the possibility to enlarge the target spectrum without changing any laboratory routine and training of the personnel. However, the implementation of new targets to the system is not always easy, but can be realized in relatively short time.

(V) The sensitivity is one of the most crucial factors of a pathogen detection system. Ethelberg et al. demonstrated in their study that in case of persistent diarrhea a positive diagnosis was about three-fold less likely relative to the category “standard analysis” (Ethelberg 2007). It was supposed that this is due to low numbers of pathogens in the faeces. This illustrates the need for a highly sensitive method and shows, that classical laboratory stool diagnostics cannot always provide the required sensitivity. Genotyping methods have proven to be able to detect pathogens with high sensitivity. Sensitivity in terms of microarrays is discussed in detail in chapter 4.2.6.

(VI) Compared to phenotypic tests the results of genotypic pathogen identification are less subjective through computational data analysis instead of visual assessment of phenotypic assays. However, this can also be a disadvantage. Experienced laboratory personnel can react on ambiguous or unusual results of morphological and physiological investigations, which is not possible in case of yes-or-no answers from *in-silico* data analysis.

Although genotyping methods have many advantages towards cultivation-based phenotypic pathogen identification, they also suffer from some drawbacks, which are often linked to the absence of the pathogen cultivation.

(I) The main challenge for cultivation-independent pathogen genotyping in faecal samples is the non-sterile material containing huge amounts of background DNA. Especially, in case of low pathogen concentrations strategies for up-concentration of the target are required, as the enrichment by pathogen cultivation omits. Regarding stool samples commercial kits are available, which combine DNA extraction and enrichment (QIAamp® DNA stool mini kit, Qiagen, Hilden/GER; InviMag® Stool DNA Mini Kit or PSP® Spin Stool DNA Plus Kit, Invitex, Berlin/GER; CRYSTAL Stool DNA Lysis & Isolation Kit, BioLabProducts, Gödenstorf/GER; ACCUPREP Stool DNA extraction kit, Bioneer, Alameda CA/USA; EurX GeneMatrix Stool DNA Extraction Kit, Roboklon, Berlin/GER; E.Z.N.A.® Stool DNA kit, Omega Bio-Tek, Norcross GA/USA), but selective enrichment of the pathogen is not achieved. Special kits for separation of bacterial and eukaryotic DNA (LOOXTER, SIRS-Lab, Jena/GER) can reduce the human DNA background but not the background from the intestinal flora. Enrichment of bacterial cells may be achieved using antibodies (Voitoux 2002; Uyttendaele 2000), bacteriophage proteins (Kretzer 2007) or glycan binding lectins (Payne 1992; Patchett 1991). A PCR amplification step before detection is a standard procedure to enrich the target and is applied as alone-standing genotyping method, as well.

(II) Another problem, which is directly linked to the missing cultivation, is the inability of genotyping methods to distinguish between live and dead cells or free DNA. Enrichment strategies with surface binding molecules could at least separate lysed from non-lysed cells. However, in case of acute gastroenteritis a differentiation between live and dead pathogens is not of therapeutic consequence.

(III) A major limitation of genotyping methods is the restriction to the detection of known pathogens and the intolerance towards changes in the DNA sequence. However, in most clinical applications the spectrum of potential pathogens is also limited and well known. Moreover, classical culture-based methods also have low chance to detect unusual pathogens. Additionally, genotyping assays can provide redundant information by e.g. including several primers or probes for one target bacterium to make the assay robust towards changes in the DNA sequences.

(IV) The quantification of the pathogen is important information in the assessment of an intestinal infection. Cultivation-based techniques allow doubtless quantification by colony counting. Genotyping methods require sophisticated calibration to generate reliable quantitative information. This is caused by limited linearity between the target amount and the generated signal and different behaviour of the same amount of distinct pathogens in the assay. This is, depending on the method and concrete target, a consequence of the limitation in the linear signal acquisition, the variable dye incorporation rates, the different gene copy numbers of discrete pathogens, and individual sequence differences resulting in different binding behaviour or amplification efficiency.

(V) At present, only few clinical standards exist for the application of genotyping methods in clinical routine. Clinical studies have to prove the efficiency, accuracy, and reproducibility of these methods.

(VI) Currently, genotyping methods are still more expensive than classical phenotyping. Efforts are made regarding the development of miniaturized systems and the use of cheaper materials to reduce the cost. Nevertheless, the saved expenses due to faster diagnostic information and the advantages for the patient already justify the application of genotyping methods.

4.2 Design and performance of the Gastroenteritis-Chip

The developed microarray is a step forward towards the development of diagnostic tools to complement the classical identification routine and fill the gap of uncertain identifications of pathogens. The species coverage of the developed diagnostic array comprises the most relevant bacterial pathogens of gastroenteritis, which are described in literature (Janka 2001) and are regularly detected in the clinical routine diagnostics, as communicated by our clinical partners. Additionally, it provides further information on the constitution of the dominant commensal bacterial community in the patient's GI tract, as it includes probes for *Bacteroides* spp., *Enterococcus* spp., *Fusobacterium prausnitzii*, *Veillonella* spp., *E. coli*, *Roseburia* spp., *Atopobium* spp., *Acholeplasma laidlawii*, *B. bifidum* and *L. lactis*. Probes detecting potentially probiotic species like *L. acidophilus*, *L. delbrueckii*, *L. lactis*, and *B. bifidum* allow monitoring of administration of these bacteria in the course of therapy. Information on the background flora and probiotic species may allow basic studies on the effect of probiotics in antimicrobial therapy and a comparative investigation of the influence of antibiotics and probiotics on the commensal flora. Probiotic bacteria are an interesting alternative to antibiotic therapy. Strains *E. coli* Nissle 1917, *Lactobacillus rhamnosus* GG, and a mix of *L. delbrueckii* var. *Bulgaricus*, *Streptococcus thermophilus*, *L. acidophilus*, and *B. bifidum* have proven to reduce the stool frequency and duration of diarrhea in infants and toddlers (Henker 2007; Canani 2007). Positive effects of probiotics were also described for inflammatory bowel diseases such as pouchitis, ulcerative colitis, and Crohn's disease (Hormannspenger 2010). Restrictively it has to be mentioned that probiotic action is strain dependent. Although the array detects potential probiotics only on the species level and not on strain level, it allows at least relative quantification of these bacteria, even if the probiotic strains are not specifically detected. This already enables monitoring of probiotics during administration. Several mechanisms of probiotic action are discussed and investigated, including modulation of the gut microbiota composition, production of antibacterial substances, competitive exclusion of pathogenic bacteria, induction of defensin production, repair initiation of the intestinal barrier function after damage, and modulation of host immune function (Wohlgemuth 2010).

By combining pathogen detection and monitoring of the commensal microbiota, the present array distinguishes from previously published identification chips. Jin et al., Kostic et al., and Mao et al., for example, focused on the detection of intestinal pathogens (Jin 2006; Kostic 2007; Mao 2008). Jin et al. detected 15 intestinal pathogens on species or genus level using

a set of 20 specific probes, which had the 16S and 23S ribosomal genes as targets. Kostic et al. detected 11 pathogens with a set of 35 probes against the *gyrB* gene. Targeting the 16S rRNA gene and two virulence genes with 21 probes, Mao et al. identified 14 species or genera of intestinal pathogens. Nevertheless, regarding alternative therapy with probiotics information on the background flora will give valuable information for the clinician to assess the therapeutic success. You et al. included eight pathogens in their microarray, but used specific virulence genes for identification (You 2008). This approach had the advantage of better specificity but the disadvantage that for each virulence gene a separate primer set was needed. By this, an enlargement of the target spectrum is more difficult to realize and it is not taken advantage of the multiplexing capacity of the microarray detection. Li et al. identified 15 pathogenic *Shigella* and *E. coli* serotypes targeting O-serotype-specific genes (Li 2006). This array gives valuable information for epidemiology but does not cover enough pathogens to serve as independent diagnostic tool. In contrast, Wang et al. detected only intestinal, resident bacteria using the 16S rRNA gene as target (Wang 2004a), while Lehner et al. focused on *Enterococcus* species and reached discrimination by using 41 nested 16S or 23S rRNA gene-targeted probes (Lehner 2005). Boesten et al. discriminated bifidobacteria to the species level by using a microarray with a 16S-based hierarchical probe cluster (Boesten 2009). Some recent publications presented microarrays, which cover the intestinal diversity by identifying phylogenetic clusters and were used to study the intestinal ecosystem. Rajilic-Stojanovic et al. analyzed 129 bacterial, phylogenetic groups in young and elderly individuals using 4,800 tiling oligonucleotide probes targeting two hypervariable regions of the SSU rRNA sequences (Rajilic-Stojanovic 2009). However, species identification was not shown in this work. Paliy et al. also used a phylogenetic cluster array to identify bacterial groups in the intestine (Paliy 2009). This approach is very useful for a comparative investigation of the intestinal ecosystem under certain aspects. For clinical applications, a small array is preferable, due to the necessity of clear yes-or-no answers and the cost of large microarrays. Small, highly specific probe sets allow identification of species instead of phylogenetic groups. For the present array developed in this work, the clinical applicability had the priority. Restricted information on the background flora and probiotics may help the clinician to follow the course of therapy without producing too much information that needs to be handled by the medical personnel and is not necessary for the medical care.

4.2.1 Choice of the target gene

Two types of targets have been frequently applied for identification of microbes with microarrays, that is amplified genes and rRNA (Pozhitkov 2007). Specific genes or whole-genome DNA were used for identification purposes of distinct microbial groups, which in some cases promise higher specificity and separation to the strain level compared to rRNA genes. The *gyrB* gene, for example, was used for intestinal pathogens (Kostic 2007; Kakinuma 2003), antigenic determinants, and virulence factors for *Campylobacter* spp. (Volokhov 2003), *Salmonella* spp., *Shigella* spp. (Chizhikov 2001), *E. coli* strains (Chizhikov 2001; Bekal 2003), *Listeria* spp. (Volokhov 2002), and various species (You 2008) or the internal transcribed spacer (ITS) region for pathogenic fungi (Leinberger 2005) and mycobacteria (Park 2005). Lactic acid bacteria were also detected using full bacterial genomes as probes (Bae 2005). For pathogen and probiotics detection, this could have been a promising approach, but this concept would have required either a very complex multiplex-PCR, restriction to few pathogens, or a whole genome amplification for not losing the multiplexing capacity of the microarray to the prior enrichment step. Using rRNA as target a higher sensitivity may be achieved, due to the higher copy number in the cell, and the fast degradation after cell death may allow differentiation between live and dead cells. However, the low stability of this molecule makes it not a favoured target for clinical applications and reliable quantification is even a greater challenge, because the extend of unspecific hybridization in natural samples is not known and difficult to determine (Pozhitkov 2007). Ribosomal RNA was applied for pathogen detection in RT real-time PCR (Amar 2005) and on microarrays (Anthony 2005), and for community analysis in dot blot hybridization and temperature gradient gel electrophoresis (TGGE) (Felske 1996), on microarrays (Adamczyk

2003; Hoshino 2007; El Fantroussi 2003), and in fluorescence in-situ hybridization (FISH) (Amann 1995; Fallani 2006; Lay 2005).

In this study, the 16S and 23S rRNA genes were chosen for species detection. These genes are commonly used as targets for species identification (Desantis 2007; Mao 2008; Jin 2006) because of their high sequence availability covering large bacterial communities, their variability allowing specific detection, and their multi-copy-existence in many species increasing the sensitivity of an assay towards them.

The high sequence coverage of the rRNA genes allowed prediction of potential cross-hybridizations *in-silico* by checking the probes against a huge database of sequences of these genes. The used 16S and 23S databases contained ~41,000 and ~7,000 sequences. The current sequence databases released by the ARB Silva project contain ~460,000 and ~17,000 high quality sequences (>1200 nt) (www.arb-silva.de). Despite the comprehensive 16S database, it was necessary to use the 23S gene as well, because this gene shows more variation between species of medical importance (Anthony 2000). The probe design based on a comprehensive database enlarges reliability of the array (Loy 2006) and is especially important for microarrays detecting bacteria in complex ecosystems, as a full verification of the array performance with all members of this community is impossible. This was the case for the developed Gastroenteritis-Chip, which had to detect pathogenic and non-pathogenic bacteria specifically in front of the intestinal genomic background. Additionally, the sequence redundancy within the database enables to design probes based on more than one sequence. This is of advantage, because sequence variations or sequencing mistakes can be considered during probe design. Additionally, the universality of the ribosomal genes reduces the complexity of the amplification procedure and allows a straightforward extension of the species coverage of the DNA chip. Only in case of *C. difficile*, it might be discussed that the organism itself, as a natural inhabitant of the gut, is not pathogenic and therefore, the toxin should better be detected. In this study, however, only three wrong-positive samples with *C. difficile* were detected. In most cases, the natural *C. difficile* concentration was below the detection limit of the assay. It is conceivable to think about a follow-up test for the toxin after positive *C. difficile* results or to include this target into the array. Furthermore, *Shigella* spp. could not be detected by the ribosomal RNA genes. To include these pathogens, detection of the toxin genes would be necessary.

4.2.2 Target amplification

Target amplification and labelling is one of the crucial steps in microarray technology. It is quite common to perform a PCR prior to hybridization to enhance sensitivity and to incorporate a detection label. However, despite the sensitivity and specificity associated with PCR amplification, the inherent bias and throughput of this approach often constrain the principle benefits of downstream microarray application (Vora 2004; Suzuki 1996). If one target is used for the identification of several species, universal primers flanking the specific regions have to be found to perform a consensus PCR. The use of different primers for different species is impractical, as this would nullify the advantage of parallel species detection with the microarray. Primers that amplify the ribosomal genes were previously published (Wang 1996; Anthony 2000), but the design of new primers or their modification was necessary to match the specific choice of target bacteria and the PCR conditions and to cover all designed probes. The current microarray comprised probes, which were spread over both ribosomal genes. Initially, it was aimed to amplify both target genes from isolated DNA as one amplicon. This was not successful due to the significant fragmentation of the template DNA after isolation, similar to what was described previously (Jin 2006). The isolation procedure, which was used in the clinical laboratory, seemed to be too harsh to obtain high quality DNA. In order to adjust to the clinical situation, five PCR primer pairs were developed which amplified shorter fragments. This strategy allowed the application of the microarray relatively independent from the DNA extraction method and template DNA degradation. In contrast, Jin et al. limited themselves to amplify only two short fragments of 500 and 640 nt length from the 16S and 23S genes (Jin 2006). This also restricts the

possibility to find specific probes and the chance to enlarge the array for further target bacteria. To enhance the signal intensity on the array, Jin et al. performed an asymmetric PCR to generate more single-stranded DNA (Jin 2006). This approach was not applicable for the Gastroenteritis-Chip, because it contained sense and antisense probes. Moreover, the reached sensitivity of the published array was similar or even lower than the sensitivity of the Gastroenteritis-Chip. Jin et al. found a detection limit of 10^3 CFU/mL for single species and in a multiplex PCR with several bacteria a limit of 10^5 CFU/mL for some species. The sensitivity of the Gastroenteritis-Chip is discussed in detail in chapter 4.2.6.

A multiplex-PCR would be advantageous for the assay to reduce cost and complexity. However, the use of several primer pairs in one reaction decreases sensitivity and increases the chance that unrelated primers produce spurious products (Anthony 2000). This is most likely in case of adjacent targets with overlapping primers, as for the Gastroenteritis-Chip. In general, multiplex PCR protocols were preferentially used for the amplification of short, non-adjacent targets (Jin 2006; Mao 2008; Volokhov 2002). Less efficient amplification of some fragments in multiplex PCR was also observed in this work, e.g. for the long DE fragment. In some cases, the DE fragment was not amplified at all. The adjacent fragment F was often weakly amplified or could not be distinguished from fragment H by Lab-on-a-chip capillary gel electrophoresis. An inefficient amplification of DE and F in the multiplex PCR could have been caused by the high dimer stability of the DE reverse primer and the F forward primer, which overlapped by 12 nucleotides. Due to the narrow conserved regions in the ribosomal genes, it was impossible to design primers without sequence overlap. Even if the amount of produced PCR product was enough for detection of the bacteria from pure culture after multiplex-PCR, in some cases the amplification efficiency might be too low regarding real faecal samples, where DNA from many different species competes for the primers. Additionally, some unexpected products were found, due to efficient amplification by unrelated primers. With regard to the desired sensitivity towards intestinal pathogens, it was decided to waive the multiplex-PCR. The achieved broad information from a microarray analysis justifies the procedure of five separate PCR reactions. In comparison, in real-time PCR five PCR reactions are required to detect five targets, which could be five different pathogens or five gene segments belonging to one pathogen. If further reduction of the necessary PCR reactions is required, it was conceivable to check for the possibility to multiplex fragments H and F and fragments J and G to reduce the number of PCR reaction to three. First tests with *E. coli* were promising (data not shown).

One of the major problems in PCR amplification is the bias in amplification efficiency between different targets, which is introduced during simultaneous amplification of specific sequences from complex template mixtures (Polz 1998; Suzuki 1996). This bias was also observed in this work. The amplification efficiency varied by a factor of five between different fragments and species. This was of no consequence for species identification, but if quantification of the pathogen was requested, alternative amplification strategies were necessary. To circumvent the problems related to multiplexing of PCR reactions and to allow probe design towards any gene without changing the amplification procedure, whole genome amplification with random primers can be adopted. Although, random amplification approaches, like random primed, isothermal Klenow fragment-based, and ϕ 29 polymerase-based amplification, have shown to be three orders of magnitude less sensitive than multiplex PCR, the use of Klenow-plus-Klenow and ϕ 29 polymerase-plus-Klenow tandem amplification provided increased sensitivity (Vora 2004). Additionally, the amplification bias among the five genetic loci tested was clearly reduced. Nevertheless, whole genome amplification also increases the risk of unspecific hybridization and the sensitivity may be decreased for such a complex material as faecal samples are, because the target genes will not be specifically amplified.

4.2.3 Steric and kinetic effects on the microarray surface

The hybridization of a target to its specific probe is influenced by several steric and kinetic effects. The most important parameters are (I) the distance to the array surface, (II) the

probe density, (III) the direction of probe immobilization, (IV) the secondary structure of the target, (V) the sequence composition of the target, (VI) the target labelling, and (VII) the Langmuir adsorption coefficient.

The distance of the probe from the array support strongly influences hybridization. Negative steric effects of the microarray surface on the hybridization signal were previously shown (Shchepinov 1997). Guo et al. examined the influence of different spacer length between 0 and 15 nucleotides on the hybridization efficiency of immobilized probes and found the highest signals for the probes with a 15 dT-spacer (Guo 1994). The optimum for bound probes might be higher and it strongly depends from the array support. A hydrophobic surface, such as polypropylene, is solvent repellent and hinders hybridization. The best hybridization efficiency can be expected from probes in solution. However, a compromise between optimal hybridization efficiency and expense in probe synthesis has to be made. In this work, a spacer length of 11 nucleotides was successfully used. An alternative solution could be chemical linkers as described by Shchepinov et al. (Shchepinov 1997).

The probe density can negatively influence the hybridization signal by either being too low or too high. It influences the accessibility of probes and hybridization kinetics by defining the maximal absorbable amount of target. This is expressed in the Langmuir adsorption isotherm describing the kinetics of an adsorption process on a surface, which is influenced by the intrinsic adsorption coefficient, the maximal absorbable amount of target, and the concentration of the target in solution. The surface density of the probes can be controlled by the probe concentration in the spotting solution, as there is a linear relation between both parameters (Guo 1994). One oligonucleotide occupies around 6-12 Å² of the chip surface (Maskos 1992). Guo et al. found an optimal surface density of 500 Å²/molecule ($2 \cdot 10^{13}$ molecules/cm²) by applying an oligonucleotide solution of 5 mM concentration (Guo 1994). However, for a longer fragment to be hybridized the optimal surface density was significantly lower. It can be assumed that optimal density has to be determined for each array system separately. Other investigations resulted in $\sim 10^{11}$ molecules/cm² (Michel 2007) or $\sim 10^{12}$ molecules/cm² (Peterson 2002). The main parameter determining the upper limit of the surface density is not steric interference but electrostatic Coulomb repulsion of the DNA strands (Vainrub 2002). The probe density is also a key parameter for correct quantification on microarrays, as described by Michel et al. (Michel 2007). Non-charged peptide nucleic acids provide an interesting alternative to lessen the unfavourable electrostatic interactions (Shakeel 2006), but would also increase assay cost. Only in electronic-based array systems the charge properties of DNA are positively used (Sosnowski 2002; Zhang 2005b). In this work, the probes were diluted and spotted with a concentration of 20 μM. Under the ideal assumption that the spotting solution is put down to the array surface as a hemisphere and that all probes are immobilized, the probe density was $4.8 \cdot 10^{13}$ molecules/cm². This is in agreement with previously published optimal probe densities, if one assumes that not all probes are immobilized to the chip surface.

Due to the lower stability of A:T versus G:C pairs, the hybridization efficiency is also influenced by the target sequence. One study suggests that not only the overall base composition is important but also the sequence (Maskos 1993). Sequence effects are expected, as it is known that base stacking interactions, which depend on nearest neighbours, significantly affect duplex stability (Southern 1999).

Peplies et al. showed that the probe immobilization via 3' and 5' end resulted in different hybridization efficiencies (Peplies 2003). It was supposed that this is due to steric interference of the non-hybridizing overhang of the target strand with the array support.

The previously mentioned effects may play a role in an observation made in this work. Each probe on the microarray was spotted as sense and antisense version resulting in different signal intensities. This observation remains a phenomenon to be investigated. In most cases, these differences were highest when undigested DNA strands were hybridized. The hybridization may have a direction starting from one end depending on the base composition of the probe. Hybridization might then be sterically hindered by the close microarray surface

mainly for one probe of each pair due to their immobilisation direction. As the observed effect could not be correlated with either sense or antisense probe, it might have its seeds in the probe-binding site in the target strand and the base composition of each probe strand. In case of hybridization of non-digested PCR products, a 5'-immobilized sense probe sharing sequence identity with a 5' end and an antisense probe sharing sequence identity with a 3' end of the target sense sequence should theoretically display higher signals than their opposites, as the overhang of the hybridizing target strand should not interfere with the microarray surface. In contrast, Peytavi et al. found an inverse correlation between hybridization signal intensity and the length of the 5' overhanging end of the captured strand (Peytavi 2005). The length of the 3' overhang had no influence on the signal intensity. They assumed that in case of long solution-directed overhangs kinetic effects and re-association of the PCR product's complementary strand could lead to destabilization of the capture probe/DNA target duplex. The theory of Peytavi et al. was slightly supported in this work, although clear observations were not made in this regard. A strong correlation of the relative signal response (RI) of probes depending on their position in the target strand was not found for the 19 probes, which hybridized with the four undigested PCR products of *E. coli*, *R. intestinalis*, *P. shigelloides*, and *C. coli*. Nevertheless, for few probes a signal increase was seen if they hybridized to the 5' end of the target, which resulted in a surface-directed DNA strand overhang. The direct comparison of corresponding sense and antisense probes also disproved the assumption that a surface-directed overhang negatively influences hybridization. With increasing distance of the 5' immobilized sense probe to the target middle and location closer to the 3' end (binding the 5' end of the target), the signal ratio of sense vs. antisense probe also increased. This means the surface-directed target overhang resulted in higher hybridization signals than the solution-directed overhang. Probes in the target middle displayed nearly equal signals. However, if the sense probe was closer to the 5' end and the corresponding antisense probe closer to the 3' end, the signal ratio of antisense vs. sense did not continuously increase. This result shows, that a simple theoretical consideration of the 3' and 5' overhanging ends of the target cannot fully explain hybridization behaviour. Moreover, it was observed that the target digestion partly equalized differences between response intensities of sense and antisense probes belonging to one probe pair. This could be due to a reduction of secondary structures of the target strands. Therefore, it is supposed that secondary structures within the target strands, which affect accessibility of the binding motif, and the individual base composition, have comparably big influence on the hybridization behaviour as the relative position of the probe.

Although DNA targets are less problematic than RNA targets in terms of accessibility, secondary structures and duplex formation can hinder hybridization, and thus reduction of the sequence complexity is preferable (Southern 1999). In a recent study of genome-scale probe design, up to a third of all probe binding sites were affected by secondary structures, which made target regions inaccessible for hybridization (Ratushna 2005). By comparing the hybridization efficiency of amplified DNA from four bacterial reference strains, it was shown that most fluorescence signals increased after DNase I-digestion. The average signal intensity was 1.5 times higher after digestion. According to this result, all PCR products were digested prior to hybridization. Digestion influences all target-probe pairs differently depending on the intrinsic thermodynamic parameters of each target strand. Two contradictory effects have to be considered, that is secondary structures are reduced when target strands become shorter and by this hybridization is facilitated but the number of fluorophores per hybridizing DNA strand will be reduced as well. Additionally, it was observed that target fragmentation can support false-positive signals in some cases (Peplies 2003). Therefore, the degree of digestion had to be optimized to find a compromise between these effects, which allows a sensitive and specific detection of all species in parallel. The optimal digestion degree was reached with 0.4 mU DNase I per ng DNA for five minutes at room temperature. However, this parameter can be different in case of a more or less active enzyme, as it was seen in the experiments performed in Shanghai. As an alternative to DNA digestion before hybridization, Nick translation was used to incorporate the label and by this disrupt secondary structures in the target (Lane 2004). However, Nick translation should

have a comparable effect to label incorporation during PCR, as the target length remains constant. In this work, an additional improvement of the signal intensities was observed after target digestion. This effect was not investigated by Lane. The direct fluorophore incorporation in combination with DNase I digestion, as it was conducted in this work, takes less time than a PCR with subsequent nick translation. The only advantage of the ladder approach could be an increased amplification efficiency of the target DNA. Instead of shortening the target strands, also helper oligonucleotides were applied to open inaccessible rRNA (Fuchs 2001) and rRNA gene (Peplies 2003) target strands. However, in some cases they reduced the discrimination power, which is an important factor for pathogen identification from complex faecal samples.

4.2.4 Probe design

One important aspect of probe design is the necessity to find probes, which not only detect a pathogen highly specific but also perform well under the same hybridization conditions. Parameters, like probe length, GC-content, and the associated melting temperature, influence the hybridization and have to be adjusted for all probes (Bodrossy 2003). The melting temperature of the probes should be around 5-10°C higher than the hybridization temperature and can be influenced by the probe length. Longer probes have proven to give higher signal intensities (Letowski 2004). Therefore, it is not always advisable to shorten a probe, if the GC content and the melting temperature are very high. Due to the necessity to discriminate pathogens from a huge bacterial background, the probe position cannot be chosen freely. Moreover, it is advantageous to distribute mismatches over the entire length of the probe or at least to place single mismatches in the middle to reach best discrimination (Letowski 2004). Thus, there is little scope left to adapt all probes to one hybridization condition. Secondary structures of the probes have to be considered as well. In this regard, self-dimers of probes are less likely than hairpin structures, because the former require close distance of the probes. All these parameters and restrictions make the probe design a very elaborate process where permanent compromises have to be made. In this work, the distribution of the mismatches had priority towards all other parameters to ensure a satisfying discrimination.

In the final probe set, the average GC content was 48.9% and the average melting temperature was 53.6°C (OligoAnalyzer), which was in accordance with the previously defined values. For these parameters the study of Letowski et al. suggested a hybridization temperature between 47 and 53°C relying on the GC content or alternatively 8-13°C below the melting temperature (Letowski 2004). By choosing 48°C as hybridization temperature, the first criterion was met. A hybridization temperature of 8-13°C below the melting point is often linked to loss in specificity. Increasing the hybridization temperature further may slightly improve specificity but can reduce sensitivity of the array. In the diagnostic application, a false negative is more problematic than a wrong positive and the validation of the Gastroenteritis-Array revealed that the sensitivity was the critical point. Therefore, the hybridization temperature was not changed to increase specificity.

4.2.5 Probe verification and redesign

Specificity is, besides sensitivity, the most important feature of a diagnostic tool and thus much effort is made in array development to guaranty an excellent performance. Cultivation techniques used in classical pathogen detection deliver relatively low specificity and, therefore, are combined with methods of higher specificity, like physiological or serological investigation. Genotyping methods have the ambition to generate specificity in one step only. In terms of designing specific probes, the complexity of the biological sample plays an important role. Hybridization signals of one spot may be a mixture of specific hybridization of the perfect match and unspecific hybridization of a mismatch (Zhang 2005a). The *in-silico* assessment of probe specificity is limited to *a priori* known sequences. However, the complete mixture of sequences that need to be discriminated is not known for the intestinal

tract. The comprehensive ARB databases of ribosomal DNA were a good basis for probe design. Nevertheless, cross-hybridizations were observed in the array validation. The first probe set showed an unsatisfying performance with the 32 reference strains. From the 164 probes 61 (37%) were not suitable for specific identification of bacteria.

The specificity of probes can be enhanced by different ways, but often variations work in the opposite direction regarding sensitivity. Therefore, a compromise has to be made to match the applications purpose. Long oligomers of 50 nucleotides have a threshold of differentiation at 75-87% sequence similarity (Kane 2000). High sensitivity is bought dearly by loss of specificity. It was shown that decreasing probe length from 30- to 25-oligomers improves specificity (Religio 2002). Especially, in complex mixtures shorter oligomers (25-30mers) were superior towards long oligomers (60mers), although sensitivity was decreased. A critical length that allows random hybridization limits shortening of the probes. In this work, it was also limited by the necessity to discriminate many different species. A certain length had to be chosen to include mismatching nucleotides against all non-target species and to place them not at the probe ends. Nevertheless, the shortening concept was successfully applied in one case of non-specificity. Shortening of one probe that detected *L. monocytogenes* improved specificity towards *L. innocua*. At least the antisense probe was then specific for *L. monocytogenes*. Further length reduction was limited by the distribution of the mismatches.

The amount and relative position of mismatches in the targeted region strongly influences specificity (Letowski 2004). However, the different behaviour of sense and antisense probes in terms of cross-hybridization supported the conclusion that these parameters do not solely influence, whether a probe can hybridize or not. Nevertheless, in some cases the mismatch positions could explain cross-hybridizations of probes.

Specificity can also be increased by higher hybridization temperatures but from a certain temperature on (depending on the probes and the system) it can also decrease (Religio 2002). This approach is often linked to loss in sensitivity and was not pursued further (see also chap. 4.2.4). Other strategies imply major changes to the array system and are not favoured when developing a diagnostic tool, which should be as simple as possible. By introducing an enzymatic step into the system, the specificity can also be improved. Gerry et al. applied a combination of ligase detection reaction (LDR) and universal ZipCodes for the identification of point mutations (Gerry 1999), which was also used for the detection of bacteria (Busti 2002). Kostic et al. used single nucleotide extension labelling to improve the specificity of their microarray detecting low-abundance pathogens via *gyrB* gene (Kostic 2007). Competitor probes, that is unlabelled perfectly matching probes, could be added to the hybridization solution to reduce the binding efficiency of cross-hybridizing targets as it was shown for FISH probes (Lay 2005; Manz 1992). However, this approach is attended by lower sensitivity for the correct target, which also competes with the unlabelled DNA strand for the probe binding.

In most cases, the redundant probe set allowed exclusion of the unspecific probes. Few probes were defined as multi-targeting probes, namely that were the EHEC probes, two *E. coli* probes, and three *Y. enterocolitica* probes. Ten probes were redesigned or newly developed to replace unspecific ones. The probes of the final probe set on the microarray displayed cross-hybridizations below 12% signal intensity of the process control EUB338, which was below 1,000 FU at PMT 65. By applying a cut-off of the average observed signal from the negative control plus three times its standard deviation, in most cases these signals were reduced nearly to zero. By this, a microarray was developed, which is robust and can specifically detect bacteria in front of a huge bacterial DNA background.

Jin et al. used the previously determined average signal intensity of cross-hybridizing species and the blank control plus two standard deviations as a fixed cut-off for each individual probe with values between 1,600 and 4,500 arbitrary units (Jin 2006). However, in the background of a complex faecal community signals from several cross-hybridizing species might sum-up and lead to false-positive signals, especially when they are in excess towards the pathogen.

Herein, it was approached to minimize cross-hybridizations as far as possible and to ensure correct identification by using multiple probe sets. Remaining cross-hybridizations occurred only with single probes of a probe set. By accepting positive signals only, if more than 50% of the probes in the respective probe set displayed fluorescence and the identification of a species only if the genus-specific probe was positive, these cross-hybridizations could be considered uncritical. Moreover, these minor cross-hybridizations occurred when 200 ng of PCR product from a pure culture isolate were hybridized. In case of a faecal isolate 500-2,000 ng PCR product were hybridized. A single species has to account for 10-40% of the faecal sample DNA to gain the same signal intensities. The multiple probe concept has been used successfully in other publications (Wang 2004a; Liu 2001). In contrast, Mao et al. just relied on one probe for each pathogen, which makes the array less robust towards changes in the target sequence and potential cross-hybridizations (Mao 2008).

4.2.6 Limit of detection of the Gastroenteritis-Chip

The detection limit of the microarray is of main relevance in diagnostic application. In classical pathogen detection, a high sensitivity is reached by prior pathogen enrichment using cultivation technique. Herein, the enrichment by cultivation was replaced by DNA target enrichment using PCR amplification. This, together with DNA extraction from the sample, is the most crucial step in terms of assay sensitivity. This was also the main reason to refrain from performing a multiplex PCR (see also chap. 4.2.2).

The developed microarray was tested for its detection limit towards bacterial DNA and its sensitivity for a pathogen in presence of faecal background DNA. It was reported that the bacterial background DNA significantly decreases the signal intensities from a targeted PCR product (Denef 2003). Without background DNA, *E. coli* was detected down to 1,000 genome equivalents applied to PCR reaction. In front of a $2.7 \cdot 10^3$ fold excess of background DNA also 1,000 genome equivalents of *C. jejuni* were detectable. This amount corresponds to 0.04% of total DNA applied to the PCR or an estimated 0.08% relative abundance of the target bacterium (4 fg average DNA content/cell and 1.8 fg DNA/*C. jejuni* cell). However, one has to take into account that this determination did not include DNA extraction. The sensitivity for real faecal samples may be lower, not due to limitations of the array, but due to limitations of the DNA extraction method. Nevertheless, even if a DNA recovery rate of only 10% is assumed, this result is in agreement with other reported detection limits, which are between 1 to 5% relative abundance of one species spiked into a bacterial community. Denef et al. found a detection limit of 0.54% for a multi copy gene when spiking human genomic DNA into mouse DNA (Denef 2003). Recently, a microarray for intestinal pathogens was published, that relocates the specific hybridization event from the solid surface to the solution by using probe-complementary oligonucleotides for a hybridization-dependent, specific end labelling reaction in solution. For this array a sensitivity of 0.1% relative abundance was reported (Kostic 2007). Wang et al. determined a detection limit of 10^5 CFU/g down to 10 CFU/g for some species obtained from direct cell PCR (Wang 2004a) and Jin et al. found a limit of 10^3 CFU/mL for one pathogen and 10^5 CFU/mL for six simultaneously detected pathogens in front of a faecal background (Jin 2006). With a PCR-free genome-probing microarray for lactic acid bacteria Bae et al. detected 2.5 ng genomic DNA in the presence of background DNA accounting for 0.25% of the microbial composition and around 10^5 to 10^6 cells (Bae 2005). Depending from the *rrn*-operon copy number and the amplification efficiency the sensitivity might be slightly different for other target species. It was also reported, that the recovery efficiency of Gram-positive and Gram-negative bacteria is different with 1-20% compared to 80-100% using the DNA easy kit from Invitrogen (Wang 2004a). This will affect low abundant Gram-positive species below 10^5 CFU/g faecal sample, which might be not detected if the recovery rate is only 1%.

Currently, the sensitivity of culture and real-time PCR is still higher than that of microarrays. The detection limit of conventional culture is 1 CFU and combined with appropriate filter steps it can be lower than 1 CFU/mL, as determined for *Legionella pneumophila* in water samples (Villari 1998). However, this method requires 5-14 days, specific culture media for

each expected pathogen and well-trained microbiologists. For example, the pathogenic *Mycobacterium avium*-species belong to the slow growers, which need one to four weeks for cultivation (Wallace 1997). The sensitivity of real-time PCR lies between 1-1,000 CFU depending on matrix and extraction methods. For the detection of *Campylobacter* species in food a real-time PCR assay was developed, which had a detection limit of 12 genome equivalents as determined with a dilution series of the target DNA (Sails 2003). However, to detect *Campylobacter* species from food samples an enrichment culture was necessary and the sensitivity determination did not include a DNA background. A detection limit for *E. coli* O157:H7 of 5×10^2 cells \cdot mL $^{-1}$ in a buffer solution was determined by real-time PCR combined with immunomagnetic separation. The sensitivity for this pathogen was 1.3×10^4 cells \cdot mL $^{-1}$ when spiked into ground beef (Fu 2005). Ott et al. found a detection limit for certain members of the GI tract community by real-time PCR and dilution experiments of 10 to 10^3 cells. The mean recovery rate in spiking experiments into biopsy or faecal specimens was 79% (Ott 2004b). *Bifidobacterium bifidum*, *C. difficile* and *E. coli* were detected down to 5×10^3 , 2×10^3 , and 4×10^3 CFU/g faeces (Penders 2005). The quantitative linear detection is limited to 10^6 cells/g faeces (Matsuki 2004). The main disadvantage of real-time PCR is the limited multiplexing capacity.

The microarray technique has the capability to lower the detection limit to the level of real-time PCR by using more sensitive labels, array surfaces with higher probe capacity (Benters 2002), pre-hybridization steps in solution (Hashimoto 2006), reduction of the reaction volumes by microfluidic systems, and RNA as the target for detection.

Labels that are more sensitive are fluorescence dyes with higher quantum yield, enzymatic labels that include signal amplification by the enzymatic reaction, and silver deposition. Due to instrumental aspects, silver deposition was used in this work during the research stay in Shanghai, but it was not investigated, whether a higher sensitivity was reached with this label. Vora et al. compared different labelling strategies for microarrays with respect to sensitivity, replicability, and cost (Vora 2008). They found three methods that were capable of detecting *Vibrio cholerae* toxin genes present in a single copy, namely Alexa Fluor 647 aminoallyl-dNTP indirect labelling (MolecularProbes/Invitrogen), modified Cy3-Tyramide Signal Amplification Biotin System (PerkinElmer), and gold resonance light scattering nanoparticles (Genicon/Invitrogen), and two methods that had a detection sensitivity of 10 copies, namely Cy5-anti-biotin-antibodies plus Cy5-anti-mouse-antibodies and NeutrAvidin/Cy5-dual labelled cowpea mosaic virus particles. Unfortunately, the most sensitive methods also exhibited the highest experiment-to-experiment variability. Vora et al. concluded that this was due to the more complex detection protocols, which increased the opportunity for experimental error. Additionally, these methods were also the most costly ones. An interesting way to enhance on-chip fluorescence signals is the use of fluorophore-labelled dendrimers as it was published by Striebel et al. (Striebel 2004). They observed a signal-amplifying effect up to the factor of 30 after incorporation of the dendrimers into the PCR product by coupling a dendrimer to the downstream primer and performance of an asymmetric PCR.

This principle of branched molecules was also transferred to the array surface. An array surface with three-dimensional structure was described by Benters et al. (Benters 2002). Dendrimers containing 64 primary amino groups were immobilized on the array surface and functionalized for oligomer binding. The dendrimer surfaces showed dramatically increased hybridization capacity and therefore sensitivity. The arrays could also be regenerated several times. The epoxy-coated glass slides of the company Eppendorf, which were used in the present work, also had a 3D structure. The amount of functional groups was not known. During this work, the production of the Eppendorf glass slides was stopped and they were replaced by Nexterion glass slides without loss in quality. Although, the epoxy-coated Nexterion glass slides did not provide a 3D structure, they performed comparably well.

The relocation of the specific detection reaction into the solution keeping only the detection of the signal on the array surface could enhance specificity and sensitivity. Hashimoto et al. used discriminating primers and an adjacent, labelled, common primer to perform a ligase

detection reaction (LDR) in solution upon binding of both primers to the target. The LDR products containing a specific ZipCode were hybridized to an array, where the fluorescence signals were read out (Hashimoto 2006). Kostic et al. showed that single nucleotide extension labelling in solution upon primer binding to the target could increase sensitivity of the array (Kostic 2007). However, both publications did not show that hybridization to the array surface requires less stringent conditions, which was beneficial for the sensitivity.

Using RNA as target could improve sensitivity, because RNA is produced by the cell in huge amounts during the exponential phase. In contrast, the ribosomal genes are present in 1 to 15 copies (rrnDB: <http://ribosome.mmg.msu.edu/rrndb/index.php>) per bacterium only (Acinas 2004). As previously mentioned, the instability of RNA makes it a difficult target for clinical routine diagnostics. As all DNA isolates from the hospital showed a high degree of fragmentation, there is no information about the RNA content. Although, the Qiagen kit does not include an RNase-step but claims to yield pure DNA, the samples were supposed to contain mainly DNA and only minor amounts of rRNA.

Additionally, pre-enrichment of the pathogen with selective binding strategies could improve sensitivity. Kretzer et al. applied cell wall-binding domains of bacteriophage endolysins for the selective immobilization and separation of *L. monocytogenes*, *Bacillus cereus*, and *Clostridium perfringens* (Kretzer 2007). Specific antibodies were used to bind *E. coli* O157 selectively on magnetic beads (Voitoux 2002). However, the present Gastroenteritis-Chip would require a less specific enrichment strategy covering all bacteria that should be detected by the microarray.

4.3 Application of the Gastroenteritis-Chip

The primary application of the developed microarray was the detection of intestinal pathogens that cause gastroenteritis. Additionally, the chip should provide information on the resident intestinal flora and potentially probiotic bacteria. The Gastroenteritis-Chip was tested with clinical samples of patients and healthy individuals to compare its performance and value as diagnostic tool with current clinical methods. However, this comparison is only possible in terms of pathogenic bacteria, which for the clinical laboratories routinely test, but not the resident and potentially probiotic bacteria. The amount of available samples was not as high as desired, because a direct clinical access was not given in our institute. Furthermore, the Gastroenteritis-Chip was used to investigate faecal samples from children and a trial with piglets with not array-detectable or unknown pathogen.

Multivariate analysis was applied to investigate, whether the intestinal flora was significantly changed by infection with bacterial or viral pathogens and whether the human and pig flora had established in the inoculated piglets. Three independent methods of multivariate analysis were applied, that is to say principle component analysis (PCA), partial least squares analysis (PLSA), and one-way analysis of variance (ANOVA).

PCA and PLSA are popular methods to analyze complex biological data (Franklin 1999; Wang 2004b; Zhang 2009b) and have been already applied for the analysis of microarray data. However, according to literature survey this application was restricted to expression microarrays only (Fernandez-Perez 2010; Kim 2007; Jonnalagadda 2008; de Haan 2009). Community changes upon intestinal infections were also studied using multivariate analysis for DGGE and clone library data (Feng 2010; Zhang 2009b). Here, the first application of PCA and PLSA for the analysis of microarray data regarding the intestinal community composition was shown.

PCA analyzes data without previous knowledge about classification of samples. The calculated principle components are sorted according to their ability to predict group affiliation. This has the advantage of being independent from the investigators intention. If a classification of samples according to biological characteristics can be seen in PCA and the principle components explain considerable amount of variation in the data, the predicted

difference is probably significant. However, separation of data by PCA is sometimes not very clear. PLSA is superior towards PCA because it solves both, classification and regression problems (Zhang 2009b; Rudi 2004; Idborg 2005). In PLSA, the data are analyzed under the knowledge of sample classification. The results in this work support this observation, as groups were separated slightly better by PLSA than PCA.

ANOVA in its simplest form is a statistical test to compare whether the means of several data sets are significantly different. It was published that in finding the important variables for group differentiation, PLS outperforms classical ANOVA (Zhang 2009b). This was not supported by this work. Here, PLSA generally picked more variables than necessary for a model with highest accuracy. One-way ANOVA usually found only the necessary variables. Only in one case, it selected one additional variable, which reduced PLS model accuracy. This observation suggests combining both methods to make reliable predictions.

4.3.1 Clinical samples from gastroenteritis patients and healthy volunteers

The identification of pathogens is the most important task of the developed microarray. The analysis of clinical faecal isolates with the final array allowed the correct identification of 67% of the samples with a clinically confirmed bacterial pathogen. Unfortunately, one from four (25%) *Campylobacter*-, three from eight (38%) *Clostridium* toxin-, one (100%) *Yersinia*-, and six from eighteen (33%) *Salmonella*-positive samples were not correctly identified. This was not satisfying and might have several reasons. The present array had a detection limit of 10^3 genome equivalents after DNA isolation. The infectious dose can be less than 10^3 cells (Tab. 1.2). When comparing the Ct values from real-time PCR and the detection by microarray, it can be noted, that mainly samples with high Ct values (low pathogen numbers) were not identified by the microarray. Perhaps, the pathogen number was lower than the detection limit of the microarray. In case of *Salmonella* spp. with one exception infections with Ct<28.1 were identified but samples with higher Ct values were not. The *Campylobacter*-positive sample, which was not identified, had a Ct value of 29.41 in real-time PCR, whereas the correctly determined specimens had Ct values of 29.44, 24.89 and 19.66. For *C. difficile*, the correlation was not given, but this might be related to the different targets, that is the toxin genes in real-time PCR and the 16S rRNA gene on the microarray.

Additionally, the DNA extraction and sample storage may have influence on the successful identification of intestinal pathogens. Previous attempts to amplify fragments longer than 1,000 bp from the isolated DNA failed due to highly degraded DNA. In this regard, especially low abundant species would be more affected in terms of percentages. The storage of the isolated DNA at -20°C accounted for several weeks and transport to the laboratory for 1-2 days. This might also be the reason for the weak or impossible amplification of five faecal DNA isolates. In contrast to the microarray approach, the detection by real-time PCR was based on amplification of shorter DNA fragments (~100 bp).

Missing some pathogens resulted in a relatively low clinical sensitivity. The sensitivity of microarrays for intestinal pathogens in real samples is a general problem. Jin et al. missed 11% of the pathogens in clinical samples with their array, which had a comparable detection limit (Jin 2006). However, another publication with the same detection limit stated 100% sensitivity but the array had only 95% specificity (Mao 2008). Other authors of microarrays for pathogen detection did not validate their arrays with clinical samples (Kostic 2007; Li 2006).

Few samples displayed positive results for pathogens, which were not detected in the clinical routine. This lowered the clinical specificity to 89%, but as the clinical routine diagnostics was not corrected for potential failure, one has to judge this result carefully. Regarding sample S32, it is very likely that after positive identification of *C. difficile*, no further investigation for other potential pathogens was performed and a real co-infection with *C. jejuni* was present in this sample. In S50, where *Salmonella* ssp. was detected, the real situation remained unclear. One healthy individual, who was a staff member of the hospital, and two hospitalized persons also showed a false positive hybridization result for *C. difficile*, which is in low

amounts a normal inhabitant of the gut flora. In hospitals the occurrence of spores is much higher than elsewhere (Bartlett 2002), which might have resulted in a detectable colonization with *C. difficile* not secreting any toxin. False positive results, when referring to clinical diagnosis, were also found by other studies (Mao 2008). In principle, false positive results are a less critical problem than false negatives, but for an alone-standing diagnostic tool, they are not acceptable.

The false positive results for the *Vibrio* genus-specific probe in some faecal DNA isolates were supposed to be a product of undetected cross-reactivity of this probe with other intestinal bacteria. A *Vibrio*-infection was not detected in the hospital in either of the samples and the array gave no positive signals for the usual *Vibrio*-species, which cause intestinal infections. *In-silico* and during verification of the probes, no relevant cross-reactivity was found for the *Vibrio*-probe, but as there is only one *Vibrio*-probe on the chip, a positive result for this species is not as reliable as for all other species. A pathogen identification of a *Vibrio* spp. should therefore always require a positive signal of the respective species-specific probe sets. The positive results for the *Mycobacterium* genus-specific probe in some specimens were also not accompanied by signals for the *M. avium* species. *In-silico* no full matches with other intestinal organisms were found, but the 16S-probe had a 16nt match with *Oryza sativa* chromosome 5 and the 23S-probe a 16nt match with some *E. coli* sequences. This could have resulted in positive signals for this genus and illustrates the necessity of a multiple probe-concept as well as the importance to check for cross-reactivity with food components, although it is expected that most food components have been degraded through intestinal tract passage.

Analyzing the age distribution of patients with respect to the clinically confirmed pathogen revealed differences, but the significance had to be validated with higher sample numbers. *C. difficile* was more prevalent in the age group >60 years (75%), while *Salmonella* spp. was found mainly in the younger patients of 4-17 years (71%). *Campylobacter* spp. was detected mainly in the middle ages of 18-60 years (75%). In case of *C. difficile*, this could be attributed to the higher hospitalization rate of elderly people, who are then exposed to this nosocomial pathogen.

The probes for the non-pathogenic, intestinal community on the Gastroenteritis-Chip detect dominant members of the intestinal community. Nonetheless, the observed individual species pattern was highly variable in all specimens and the most abundant bacterial species were generally detected more often than the less abundant ones. The most frequently detected species were *Bacteroides* spp., *B. fragilis*, *Enterococcus*, *Veillonella*, *F. prausnitzii*, *Roseburia* spp., and *E. coli*. It was published that in healthy individuals *F. prausnitzii* and related species account for $5.3 \pm 3\%$ to $16.5 \pm 7\%$ (Suau 2001) and *R. intestinalis* and related species for around 7% (Aminov 2006) of faecal bacteria. *F. prausnitzii* (clostridial cluster IV) and *R. intestinalis* (cluster XIV) immensely contribute to the intestinal butyrate formation (Barcenilla 2000; Duncan 2002). The *Bacteroides* spp. account for about 20% of the faecal flora (Franks 1998) and *Veillonella* spp. for about 0.1% (Harmsen 2002). The bifidobacteria proportion was determined to be 3% of faecal bacteria by Franks et al. According to their investigation this genus is highly variable compared to other bacterial groups (Franks 1998). As most faecal DNA samples in this work were derived from ill individuals, it cannot be expected to detect all members of a healthy intestinal flora. The natural flora is negatively influenced by both the entered pathogen and the antibiotic therapy, if yet started. However, a diseased state due to an infection might have different influence on highly variable or dominant groups of bacteria and on stable or less abundant species. Furthermore, *Lactococcus lactis* was identified, which is a normal inhabitant of the gut flora, a dairy starter culture and used as probiotic bacterium, was identified as well. According to literature, *Enterobacteriaceae*, including *E. coli*, account for about 0.1-0.2% and the *Enterococcus/Lactobacillus* group for 0.01% of faecal bacteria (Harmsen 2002). Although *Atopobium* spp. makes up for around 5% of faecal bacteria (Matsuki 2004), this genera was detected only in two samples of the patients in our study. *Eubacterium bifforme* and *L. acidophilus* were both found in three faecal DNA isolates.

Multivariate analysis revealed no significant difference in the composition of the resident intestinal flora detected by microarray between healthy and hospitalized individuals. It can be regarded as problematic that the group of healthy individuals was very small compared to the group of gastroenteritis patients. Moreover, the group of gastroenteritis patients was inhomogeneous with respect to age and clinical diagnose. Each around 1/3 of the hospitalized persons belonged to one of the following age groups: 4-17 years, 18-60 years, and 61-86 years. Additionally, no pathogen was clinically confirmed in 22 samples (27 including not amplified samples) and four different bacterial pathogens were identified in 30 samples. Although detailed data about the patient's background were not available, it can be supposed that the gastroenteritis had multiple reasons including also viral infections and perhaps inflammatory bowel diseases or irritable bowel syndrome. The influence on the intestinal background flora might be very different. However, performing PCA and PLSA with the samples of individuals with bacterial gastroenteritis only and healthy subjects did not improve group discrimination.

However, the main reason for unsuccessful discrimination of gastroenteritis patients from healthy individuals might have been the insufficient information depth of the microarray with respect to the intestinal microbiota. It is possible that intestinal infections affect at first the low abundant species, which were not covered by the present microarray. Nevertheless, PLSA identified *B. fragilis* as a species, which was mainly influenced by the health status of the host. This result was not supported by one-way ANOVA, but *B. fragilis* was previously detected with higher abundance in infected guts (Myers 1987; Sack 1994) and was also suggested a biomarker species for gut infection (Zhang 2009b).

The hybridization results of real samples were additionally analyzed with respect to the observed standard deviation between replicate spots of probe sets. This analysis proved a robust and reproducible array performance with standard deviations below 20% in 90% of cases and an average standard deviation of 11%. This was a satisfying result for a not fully automated system. A comparison with other arrays is hardly possible, as these data are seldom published. For some arrays correlation coefficients of technical replicates between 0.5 and 0.95 were published (Draghici 2006).

4.3.2 Children with rotavirus infection and healthy individuals

The faecal flora of rotavirus-infected children and healthy individuals was investigated by the Gastroenteritis-Chip using the silver deposition method for detection and a new type of array support. The change in the array support from Eppendorf epoxy-coated 3D slides to Nexterion epoxy-coated slides was a consequence of the stop of production of the former ones. After intensive optimization of the spotting conditions (not described in this work), the same array quality could be reached. Despite the non-3D surface of the new slides, no loss in signal intensity was observed. The switch to the silver deposition detection method was an adaption to the instrumental conditions in the laboratory at Shanghai University. By this, the successful implementation of an alternative detection method was shown, which allows using the Gastroenteritis-Chip without expensive and large instrumentation. Nevertheless, the silver deposition method introduces many additional handling steps into the identification procedure, which is undesirable for clinical applications. The miniaturization and automation of the detection procedure may produce relief in this regard (see chap. 4.5).

The samples from infected and healthy children supplied by Shanghai University were not supposed to contain intestinal bacterial pathogens related to gastroenteritis. The infected children suffered from rotavirus gastroenteritis. Therefore, it was aimed to investigate the influence of the infection on the resident intestinal flora, which was possible due to the availability of faecal samples from healthy children.

Unexpectedly, bacterial pathogens were found in three samples of the infected children. The detected *C. difficile* in three samples might have been a bacterial overgrowth because of the virus attack. *Clostridium difficile* is a normal inhabitant of the human gut flora but is usually suppressed by other bacteria. When the bacterial equilibrium in the gut is disturbed,

C. difficile can quickly proliferate. In contrast, the *C. jejuni* infection in one sample was perhaps nutritionally acquired. In this case, the virus infection might have been the secondary one. This result suggests that implementation of probes for intestinal viruses in the Gastroenteritis-Chip may be useful, because bacterial and viral infections can occur together and the clinical application of separate tests may lead to undiagnosed infections. Co-infections with multiple pathogens have a relative high prevalence, as it was shown by Chen et al. in a study with hospitalized children suffering from acute non-bloody, non-mucoid diarrhea. In 303 children, they found 22.8% of polymicrobial infections, including 17.2% multiple viral and 5.6% viral and bacterial co-infections (Chen 2009). Implementation of further targets, which cannot be detected on basis of the ribosomal genes, would require a modified amplification strategy.

The comparison of the residential flora of the healthy and rotavirus-infected children revealed major differences in the diversity. Dominant groups and species, like *Bacteroides*, *F. prausnitzii*, and *R. intestinalis*, partly of fully disappeared upon infection, which might have been a reaction on the infection but not the antibiotic therapy, which had not yet started. The principle component analysis of the microarray data revealed, that by the composition of the resident flora the children's faecal samples could be assigned to either the healthy or the infected group. Partial least squares analysis supported this result and the two-component model had a very high accuracy (95%). This was in contrast to the PCA and PLSA of the resident flora of the gastroenteritis patients and healthy individuals from the hospital of Giessen, where a significant separation of both groups was not observed. In this investigation 2/3 of samples were derived from adults. This could explain the different results. Perhaps the infantile intestinal flora is more prone to changes evoked by intestinal infections, because it is still under development and less stable. PLS analysis identified four microarray targets that were most important in discriminating healthy from infected children: *Roseburia* spp., *F. prausnitzii*, *Atopobium* spp., and *E. coli*. For *Roseburia* spp. and *Fusobacterium prausnitzii* this was partly confirmed by one-way ANOVA, and *R. intestinalis* was also identified by this method as significantly changed depending on the health status. Remodelling the PLS model with *Roseburia* spp. and *Fusobacterium prausnitzii* only, showed that these two species were sufficient to establish a model with highest accuracy. These two genera belong to the *Firmicutes* and are dominant members of the intestinal flora as described in chapter 4.3.1. In contrast to the adult intestinal microbiota, the infantile flora is not formed in-depth, which is why the dominant members might be more affected by intestinal infections.

The results from this trial supplement a published work by Zhang et al., where nearly the same samples were analyzed for their faecal composition of *Bacteroides* spp. using a clone library approach (Zhang 2009b). Samples R1, R2, and R4-R9 corresponded to samples R2-R25 in this work. R26 and R27 were not part of the investigation by Zhang et al. but one additional R-sample. Samples H1-H4 and H7-H12 corresponded to H6-H28 in this work. Samples H5 and H6 of the publication by Zhang et al. were not part of this work. Zhang et al. could also distinguish both groups based on the *Bacteroides* spp. composition. Three species, *B. vulgates*, *B. fragilis*, and *B. stercoris*, were identified to be significantly changed depending on the health status. One sample, R4 that was equivalent to R15 in this thesis, could not be clearly assigned to the infected group by PLSA. In contrast, PLSA allowed a clear separation of all samples by the microarray-detected resident microbiota.

The results of this application indicate that the present array could also serve as a tool to detect unusual changes in the infantile intestine, which can be an indicator for undetected diseases. In this regard however, the array would require much more validation with clinical samples from children and perhaps a broader spectrum of intestinal residents should be detectable on the species level as well.

4.3.3 HFA and PFA piglets

The microarray was also applied to investigate two groups of piglets, which suffered from an unknown intestinal pathogen after inoculation of human or pig microbiota. The main aim was to identify the cause of the disease, but it was also expected that the array and DGGE data would provide information on the establishment of the intestinal flora in the gastro-intestinal tract. This was most interesting with respect to the human intestinal flora, which had to establish in a foreign organism.

Regarding the identification of the pathogen, which was responsible for the diarrhea and which mainly the PFA group suffered from, the information derived from the microarray investigation was contradictory to the observed course of disease. The only pathogen, which was unambiguously identified, was enteropathogenic *E. coli* in the human flora-associated pigs H2, H4, and H1, H6 after fourteen and twenty one days, respectively. However, in the pig flora associated pigs no pathogen was identified, although these pigs showed the more severe course of disease. Pigs H1 and H2 displayed symptoms before day fourteen, while H6 suffered from diarrhea after day twenty-one. Piglet H4 did not show any symptoms. Two explanations are conceivable; the pathogen was not detected in the other samples due to lacking sensitivity or a second pathogen was causing the disease, which could not be detected by the present microarray. EHEC has a low infective dose of only 10 cells (FDA 2010).

Regarding the bacterial diversity, the results indicated that the inner-species and inter-species transplantation of gut microbiota allows establishment of a donor-like intestinal flora. This was already shown by ERIC-PCR and PCR-TGGE fingerprinting (Pang 2007) and provides a system for research on gut ecology in human metabolism, nutrition and drug discovery. The pig animal model is in aspects of anatomy and physiology more similar to the human than rodent animal models. In this trial, the inoculated human flora had established much faster in the piglets than the inoculated pig flora did. This originated most likely from the infection, which affected the PFA group much more than the HFA group.

This was supported by the DGGE pattern. Fourteen days after birth, the PFA piglets had still very diverse DGGE patterns, which, moreover, differed from the donor profile. In contrast, the HFA profiles were relatively homogenous and similar to the donor.

PCA and PLSA also revealed that the intestinal human flora significantly differs from the pig flora. Using PCA, the full separation occurred not yet 14 days after birth but it was clearly visible at day 21. At day 21, the conventionally raised piglets clustered together with the PFA piglets and the pig flora donor, indicating that the pig flora had established in the recipients and was similar to the conventional intestinal pig flora. By PLSA, the group separation was already observed at day 14 after birth, while 21 days after birth it was more obvious. Several species and genera were identified, which contributed to differentiation between the human and pig flora after inoculation in the piglets. Fourteen days after birth, those were *F. prausnitzii* and *L. delbrueckii*, which were typically detected in the pig flora but not in the human flora. Seven days later, *F. prausnitzii* had also established in most human flora-associated piglets and *L. delbrueckii* had partly disappeared in the pig microbiota. Now, the species that distinguished the both microbiota were *B. fragilis*, *E. faecalis*, *E. coli*, and *B. bifidum*. *B. fragilis* and *B. bifidum* were unique for the human flora and the two other species were mainly found there, as well. However, these results do not represent marker species to tell between human and pig flora, because this trial was affected by a gastroenteritic infection. Additionally, *E. faecalis* was detected in nearly all cases without positive genus-probes and can actually not be regarded as present.

From this application of the microarray the conclusion can be drawn, that this diagnostic tool can also be a valuable tool in research to follow the establishment of microbiota in recipients of foreign intestinal flora and, moreover, in infantile GI tracts. In terms of information depth, however, this array would require supplementation because it was not planned and designed for this purpose.

4.4 DNA target quantification with DNA microarrays

For the practical application of microarrays in clinical diagnostic, it is not only important to identify pathogens but also to quantify them. The main challenges are the equal distribution of the sample on the array surface, the bias, which is introduced by DNA isolation and amplification (Suzuki 1996), and the sensitivity, which is still lower than for real-time PCR. Donhauser et al. tried to minimize the PCR bias effect by stopping the PCR in the middle of the logarithmical phase (Donhauser 2009). However, using that method one has to accept a loss in sensitivity. The determined detection limit was 1.1×10^5 copies/mL water sample. Inhibiting substances can also lead to inaccurate target quantification and internal amplification controls have been developed to alleviate this problem (van Doorn 2009). Nevertheless, it can be assumed that this problem is minimized for faecal diagnostics, because stool extraction kits contain inhibitor blockers. Reliable quantification is also strongly influenced by spatial variations on the microarray surface. Regional biases can occur through the spotting process and different temperature, liquid flow rate, washing efficiency, or target diffusion rate in different chip areas (Reimers 2005). A general limitation is the fact, that there is only a limited linear range between the hybridization rate and the initial target concentration in solution (Michel 2007). A direct correlation of signal intensity with target amount requires heavy validation of each probe with its target, but is not assignable from one probe to another due to differences in binding efficiency. DeSantis et al. quantified SSU amplicons by using a HybScore, which is the average signal intensity of the PM minus MM intensity differences of PM/MM probe pairs in a given probe set. They found a correlation between spiked-in DNA concentrations and the HybScore signal intensities from the probe sets, which seem to be more independent from non-uniform probe melting temperatures and secondary structures compared to single probe approaches. Addressing the problems of DNA extraction and PCR by testing different extraction protocols, they found significant differences in community structure, which are not eliminated by an internal control concept (Desantis 2005). A control concept for uneven probe distribution and target hybridization was published by Cho and Tiedje (2002), who co-hybridized a Cy5-labelled reference with the Cy3-labelled sample DNA and calculated relative signals from both. Differences in labelling efficiency were corrected by an internal standard. The observed linear dependency between relative fluorescence intensity and relative abundance of the target gene in a genomic background resulted in a regression equation and coefficients, which can be used for calculation of the target gene amount from the signal intensity under the given experimental setup. They avoided a bias through uneven amplification by labelling the target DNA without PCR, but found a detection limit of 10 pg DNA only, which corresponds to about 10^7 to 2×10^7 gene equivalents and is not sufficient for complex ecosystems (Cho 2002). Palmer et al. developed a microarray comprising 10,462 SSU probes for the monitoring of complex microbial populations on several taxonomic levels. They did not focus on a certain habitat but tried to cover a huge spectrum of species. Quantification was achieved by co-hybridization of a differently labelled reference DNA pool of known composition and evaluation of relative signals for the respective probe set for each species (Palmer 2006). Satokari and colleagues published an alternative surface-free assay that combines characteristics of microarray and real-time PCR techniques. The identification was based on differently labelled probes in solution. Additionally, the multiplexing capacity was enlarged by analysing not only the fluorophore dye but also the probe length by capillary electrophoresis. Quantitative analysis was limited to dominating groups of bacteria, as fluorescence signals were normalized to a bacterial universal probe (Satokari 2005).

Although, one advantage of microarrays is the relatively wide dynamic range of modern platforms (Leparc 2009), the wide concentration range of 10 to 10^6 cells that can cause an infection (Lamps 2007) is still a central problem for the quantification of intestinal pathogens on most arrays. Low pathogen numbers will result in weak signals on the microarray, which have to be enhanced by technical facilities like photo multiplier tube. On the other hand, the hybridization response of high pathogen numbers will reach saturation due to limitation of the

linear detection range of the scanner. A parallel quantification of all species requires new strategies that force hybridization response of very abundant species to non-saturated intensities without losing low signals from rare species. Therefore, it was aimed to design a microarray that allows linear detection of a wide concentration range of bacteria using only one enhancement factor for the acquisition of all signals. For this purpose, a FRET-based detection system was developed. A Cy3-quencher was immobilized in the spot by coupling to the specific probes. The amount of probe with bound quencher was varied between 0 and 20 μM , while the overall concentration of spotted probe was constant at 20 μM . Previous experiments showed that the surface-bound *black hole quencher* BHQ2 was able to reduce the Cy3 signal upon hybridization of the target by 97 to 98% (Glazer 1997a). This *black hole quencher* has an absorption range of 550-650 nm, while Cy3 emits at 565 nm. Further experiments revealed that the signal reduction was at least partly due to fluorophore quenching and not only because of steric hindrance of the target-probe binding. However, the signal of spot-immobilized Cyanine 3 was quenched only down to 47-64% by target-coupled quencher, independent from the position of the quencher in the target and towards Cy3 after hybridization. This could be explained by an excess of the immobilized probe compared to the locally available target, a partial, but not complete, hindrance of the hybridization, or an inefficient quenching. The distance of both molecules was 0 or 18 bp due to the length of probe and target. It was reported that distances of 4 to 16 bp result in clearly visible quenching effects (Glazer 1997b). Therefore, the quenching might have been not hundred percent effective. The effective proportion between probe and target can be only assumed by mathematical approximation. If the probe solution is spotted to the surface in the form of a hemisphere with 120 μM diameter, the applied volume accounts for 0.452 nL. This volume contains 9 fmol of probe, if the concentration is 20 μM . Even if only 10% of the applied probe is immobilized to the surface, it still accounts for 0.9 fmol. On the contrary, the locally available target solution is more difficult to estimate. Using the known area of the whole hybridization chamber and a single spot and the volume of the applied hybridization solution in the chamber, the volume of hybridization solution above one spot can be calculated as 3.3 nL. The amount of PCR product applied to the chamber was at maximum 100 fmol. The volume above the spot area contained then 0.0047 fmol of target, which is about 0.5% of 0.9 fmol probe molecules. Even if the volume, which the molecules can pass by diffusion during hybridization, is one hundred times bigger, the probe is still in excess of the target. This would explain the incomplete quenching. However, this calculation is based on many assumptions and can only give an idea of the proportions of the hybridization partners.

The experimental setup of Hörmann was transferred to the detection of *E. coli* by four complementary probes. The main difference was the target length, the position, and the amount of fluorophore within the DNA strand. The *E. coli* target strand was derived from the beginning of the 23S rRNA gene and had a length of about 1,000 nt. Each target strand had between three to five fluorophores incorporated at random places. In case of hybridization of undigested target, some or all fluorophores may be out of the effective quenching distance to the quencher at the probe. In case of digested target, statistically less fluorophores are in the remaining target, but the distance may be still too far to allow quenching and additionally the hybridization may be more effective due to the shorter fragments. Moreover, the maximum amount of applied DNA was about 250 fmol, which is 2.5 times more than in Hörmanns experiments. Therefore, a less effective quenching was expected compared to the experiments by Hörmann. In fact, an effective signal reduction was observed for all probes independent from digestion. In case of the antisense probes the hybridization signals were reduced to zero upon 50% probe with coupled quencher in the spot, but a signal increase was observed, if the quencher concentration in the spots was increased beyond that. These results suggest an unspecific binding of the target to higher amounts of quencher in the spot in combination with ineffective quenching. In case of the sense probes, this effect was overlaid by the remaining fluorescence signal. There was found no information on these effects in literature. Therefore, more than 50% of probes in the spots should not be coupled to a quencher. This is backed by the fact that the signal reduction was not linear with the

amount of immobilized quencher but rather exponential. Nevertheless, for each group of spots with a certain amount of quencher a calibration curve could be plotted, which is linear for a distinct range of target concentration. Two different spots, one without quencher and one with 8 μM probe-coupled quencher, will be sufficient to detect 10^4 up to $2 \cdot 10^6$ genome equivalents in the linear range of the scanner. Lower germ numbers, however, approach the detection limit, which is in the lower non-linear range of the scanner. It is not likely that much higher germ numbers will be applied to PCR, as $2 \cdot 10^6$ genome equivalents of *E. coli* account for 10 ng of DNA, which was the maximum amount used in PCR. Moreover, in real samples one species is only part of the whole.

The here developed microarray setup partly overcomes the technical limitation of linear fluorescence signal acquisition for different amounts of bacteria. This problem could not be solved for low germ numbers. A main disadvantage of this setup is the fact, that each probe requires an own calibration and for multi-specific probes it has to be shown, that one calibration curve can be used for different targets with variable mixtures of bacteria. Additionally, internal standards for variations in amplification and labelling efficiency, digestion degree and hybridization conditions are necessary to quantify bacteria.

A recent publication presented single-molecule detection on microarrays, which could significantly improve the exact quantification of low bacterial numbers. Muresan et al. used a wavelet-based method to detect single bright features in each subimage and developed an algorithm based on spatial statistics to separate true signals in the spot area from unspecifically bound material outside the spot (Muresan 2010). Single molecule detection based on rolling circle amplification (RCA) allowed pathogen detection over a dynamic range of seven orders of magnitude and could solve the quantification problem, if transferred to the array surface (Jarvius 2006). Real-time detection of hybridization on microarrays is another promising approach to correctly quantify analytes with very large concentration differences, as it was shown by Hassibi et al. (Hassibi 2009). The combination of the structured microarray surface with real-time detection of on-chip amplification combines advantages of both techniques and allows implementation in miniaturized devices. Solid-phase PCR with immobilized primers was shown for various clinically relevant pathogens (Mitterer 2004), but was not combined with real-time detection. The principle of an on-chip helicase dependent amplification for pathogen detection has been shown as well (Andresen 2009). This isothermal amplification strategy omits the unfavourable heating unit, which would be counterproductive for a point-of-care diagnostic tool. Liu et al. showed the applicability of non-equilibrium melting curve analysis on arrays to discriminate between perfect match and mismatch hybridization (Liu 2001). More recently, they also published a TaqMan probe array for quantitative detection of DNA targets. TaqMan probes were immobilized via 3'-amino-modified poly(T)₂₀-spacer with the reporter dye following the spacer and the quencher at the 5'-end on a glutaraldehyde-coated glass slide. After hybridization of the target, a complementary strand is synthesized during one PCR cycle, while the quencher is released from the probe by exonuclease-mediated cleavage leaving the reporter dye attached to the slide surface emitting its fluorescence. Ct-values showed linear dependency from target amount and the detection limit for genomic DNA was about 10^4 copies (Liu 2006). Real-time PCR approaches have been already transferred to the miniaturized, microfluidic format (Ramalingam 2010; Lee 2010), but still lack the full multiplexing capacity of microarrays.

4.5 Market relevance and aspects of future development

The demand for fast and reliable pathogen detection methods in clinical routine is increasing, as they allow evidence-based therapy of patients. Genotyping has been described as appropriate way to improve assays in terms of analytical speed, information depth, multiplexing capacity, handling efficiency, and cost. Microarrays provide many of these features and have therefore already found their way into the diagnostic market (Tab. 1.1). Although most diagnostic arrays are in the development or validation stage, some are

already CE marked or FDA approved. With respect to this work, three marketed identification DNA chips have to be mentioned. The 'Dr. Food Kit' (Dr. Chip, Taiwan) detects the nine most common food pathogens, including *E. coli*, *Y. enterocolitica*, *L. monocytogenes*, *Salmonella* spp., *Shigella* spp., and *Vibrio* spp., while the 'Check&Trace Salmonella' Kit (CheckPoints, Holland) detects 98 *Salmonella* serotypes. The currently released 'G-MAP infant' array (GeneticAnalysis AS, Norway) analyzes the gut microbiota with respect to necrotizing enterocolitis. These examples demonstrate the high expectations in the market relevance of microarrays for pathogen detection. Molecular diagnostics represents one of the fastest growing segments of the in-vitro diagnostics (IVD) industry. The market is expected to reach \$5.42 billion by 2012 with an estimated annual growth rate of 11% (Park 2009) and it was appraised that the greatest growth will come in three sectors: diabetes, nucleic-acid based infectious disease diagnostics, and cell-based cancer diagnostics (Petrik 2006). Major factors of IVD market growth are miniaturization, assay sensitivity, cost containment, and automation (Petrik 2006).

The present microarray for the detection of intestinal, bacterial pathogens in combination with resident bacteria and probiotics has an interesting potential for developments in the clinical routine. It could be a tool for initial diagnosis and for the monitoring of the therapeutic success in gastroenteritis patients in combination with the recovering of the natural microbiota. Moreover, improved detection of intestinal pathogens extends the knowledge about local outbreaks, which are often unrecognized (DuPont 2009), and thus can lead to new risk prevention strategies. However, for independent application of the array in clinical routine further improvements regarding the sensitivity are desperately required. Furthermore, it should be of interest to combine the species identification with the antibiotic resistance detection on the chip for therapy optimization and epidemiology. Currently, a spread of vancomycin-resistant enterococci (VRE) is observed (Bryant 2007) and intestinal bacteria are supposed to serve as reservoirs for resistance determinants (Ammor 2007; Lofmark 2006). An increasing health problem are ESBL-producing *Enterobacteriaceae* which are able to distribute their antibiotic resistance determinants quickly among other Gram-negatives (Mshana 2009). Additionally, the microarray could also serve as identification tool in food (Liu-Stratton 2004; Lauri 2009) or water contamination (Lemarchand 2004; Straub 2003), as this is the source for most gastroenteritic pathogens. Faecal pollution of water could be detected by the probes for the commensal microbiota. However, in regards of food the DNA extraction from the generally complex matrices is the challenging and limiting step, which must not be underestimated, whereas the low pathogen concentrations in polluted water require a high sensitivity of the analytical method or enrichment strategies.

The major task for further development is to make genotyping assays, including microarrays, an alone-standing diagnostic tool, which complement culturing techniques with respect to sensitivity and accuracy. From a cost estimation of clinical diagnostic assays for intestinal pathogens it was concluded that a combination of rapid genotyping assays with routine culture is unlikely to be cost-effective. By evaluation of published sources, manufacturer's information, and discussion with laboratory staff the cost for testing one sample for *Campylobacter* spp., *Salmonella* spp., and *E. coli* were calculated to be £18.85 (22,62 €) with PCR, £15.66 (18,79 €) with immunological tests, and £15.01 (18,01 €) by culture methods (Abubakar 2007). This supports recent strategies to design fully integrated, cheap, and fast point-of-care (POC) diagnostic tools for pathogen sensing, which are of major interest in consideration of increasing cost for national health care systems and the demand to relocate the diagnostic process from the laboratory to the clinician and, if possible, to the patient. In 2005, the POC market controlled roughly 35% of the total diagnostics product market and it will probably maintain 41% by 2011 (Park 2008). Therefore, integration of assay steps into integrated devices is now the key requirement for further improvement of microarray applicability (Weigl 2008; Schulze 2009). This will reduce handling steps, minimize the contamination risk, speed up the diagnostic assay, reduce sample and reagent volumes, and by this make the array a cheap, disposable, diagnostic tool. The possibility of fast implementation of microarrays as the detection platform into microfluidic devices is a definite advantage to other established pathogen detection principles, such as Luminex technology.

The features, which have to be included into a POC tool, are pre-analytics (enrichment, cell pulping, separation of the target molecule, transport, target amplification) and detection. The two most important aspects of POC tool development are comparability to central laboratory test results and suitability for use by an untrained person (Petrik 2006). Array integration has been shown from a technical point of view (Schuler 2009) as well as for pathogen and SNP detection (Liu 2004; Lenigk 2002; Chen 2008; Karsunke 2009), and low abundant point mutations related to cancer (Xu 2007; Hashimoto 2006). Other microfluidic devices integrate detection methods like real-time PCR, LDR (ligase detection reaction), single hybridization events, immunoassays, Raman spectroscopy, or combinations of them (Sista 2008; Berdat 2008; Kwakye 2003; Cheng 2007; Easley 2006; de la Rosa 2007). A fully integrated system for detection of enteric pathogens including *Shigella* spp., EHEC, *C. jejuni*, and *Salmonella* spp. on a Disposable Enterics Card (DEC) has been developed by the non-profit Program for Appropriate Technology in Health (PATH). The system comprises immunocapturing, multiplexed nucleic acid amplification, and visual on-chip detection by the naked eye integrated into the Micronics MagnaFlow device (Weigl 2006; Ramachandran 2006). For further review of microfluidic platforms see Lui et al. (Lui 2009), Haeberle et al. (Haeberle 2007), and Mairhofer et al. (Mairhofer 2009).

Biophotonics, as a new emerging field using interactions between biological items and photons, provides a new set of tools with the capability to fit perfectly the requirements of miniaturized and multiplexed diagnostic applications (Huser 2008; West 2003; Foulter 2005). Nanoparticles, such as quantum dots, may replace conventional molecular tags due to their higher quantum efficiencies, the broad spectral emission range that can be excited with a single laser at 405 nm, greater chemical stability, and stability against photo bleaching (Karlin-Neumann 2007; Han 2001). Silver plasmon resonant particles, which scatter light based on their surface plasmon resonance and which can be tuned by varying their size, shape, the coupling between particles, and their core-shell ratio, are another interesting alternative to fluorophores (Oldenburg 2002; Francois 2003).

A remaining challenge is to enable reliable quantification on microarrays in a miniaturized format. Strategies to achieve this aim were discussed previously (chap. 4.4). One promising approach concerning miniaturization is the combination of microarray and real-time PCR technology (Pemov 2005; Khodakov 2008), as this approach could fully maintain the multiplexing capacity of microarrays while obtaining the sensitivity of real-time PCR.

It can be anticipated that the final design of a microarray will require a compromise between multiplexing capacity, specificity, sensitivity, throughput, and cost.

4.6 Conclusion

The central aim of this work was the development of a diagnostic DNA microarray for the detection of intestinal pathogens related to gastroenteritis. By developing this array, the two main limitations of classical clinical diagnostics could be solved. At first, the time required for pathogen identification compared to classical, culture-based methods was markedly reduced from several days to few hours. Secondly, the coverage of detectable pathogens was increased by the ability of the microarray to analyse many species in parallel. This could improve standard laboratory diagnostics. The new concept of detecting also resident and potentially probiotic bacteria may improve therapy management by generating further information of diagnostic value. For other applications than disease diagnostic, an increased analytical depth could be interesting. Additionally, successful application of a FRET-based system for pathogen quantification over several orders of magnitude was shown, addressing a problem of main relevance for clinical and research applications especially in multiplexing approaches.

The clinical applicability of the microarray was promising for pathogen detection. However, improvement in terms of sensitivity is still needed to compete with standard culture

technique. Additionally, this study demonstrated a new approach to assess the health status of patients during infection and therapy by analyzing the bacterial intestinal communities with the Gastroenteritis-Chip. This approach may be most interesting in case of diseased children and further research on this topic is of interest for health management.

The Gastroenteritis-Chip, with its covered pathogen spectra in combination with natural intestinal microbiota, is a clear improvement to previous developments for gastroenteritis diagnostic, but could be further improved with respect to the intestinal diversity and the growing number of pathogenic species. The main advantage of microarray technology is the possibility to realize such coverage enhancement with relatively low effort.

5 Literature

- Abubakar, I., Irvine, L., Aldus, C. M., Wyatt, G. M., Fordham, R., Schelenz, S., Shepstone, L., Howe, A., Peck, M., and Hunter, P. R. (2007) A systematic review of the clinical, public health and cost-effectiveness of rapid diagnostic tests for the detection and identification of bacterial intestinal pathogens in faeces and food. *Health Technology Assessment* 11[36], 1-+.
- Acinas, S. G., Marcelino, L. A., Klepac-Ceraj, V., and Polz, M. F. (2004) Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *Journal of Bacteriology* 186[9], 2629-2635.
- Adamczyk, J., Hesselsoe, M., Iversen, N., Horn, M., Lehner, A., Nielsen, P. H., Schloter, M., Roslev, P., and Wagner, M. (2003) The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. *Applied and Environmental Microbiology* 69[11], 6875-6887.
- Ajjampur, S. S. R., Rajendran, P., Ramani, S., Banerjee, I., Monica, B., Sankaran, P., Rosario, V., Arumugam, R., Sarkar, R., Ward, H., and Kang, G. (2008) Closing the diarrhoea diagnostic gap in Indian children by the application of molecular techniques. *Journal of Medical Microbiology* 57[11], 1364-1368.
- Almeida, C., Azevedo, N. F., Fernandes, R. M., Keevil, C. W., and Vieira, M. J. (2010) Fluorescence in situ hybridization method using a peptide nucleic acid probe for identification of *Salmonella* spp. in a broad spectrum of samples. *Applied and Environmental Microbiology* 76[13], 4476-4485.
- Amann, R. I., Binder, B. J., Olson, R. J., Chisholm, S. W., Devereux, R., and Stahl, D. A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Applied and Environmental Microbiology* 56[6], 1919-1925.
- Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995) Phylogenetic Identification and In-Situ Detection of Individual Microbial-Cells Without Cultivation. *Microbiological Reviews* 59[1], 143-169.
- Amar, C. F., East, C. L., Grant, K. A., Gray, J., Iturriza-Gomara, M., Maclure, E. A., and McLauchlin, J. (2005) Detection of viral, bacterial, and parasitological RNA or DNA of nine intestinal pathogens in fecal samples archived as part of the english infectious intestinal disease study: assessment of the stability of target nucleic acid. *Diagn.Mol.Pathol.* 14[2], 90-96.
- Aminov, R. I., Walker, A. W., Duncan, S. H., Harmsen, H. J., Welling, G. W., and Flint, H. J. (2006) Molecular diversity, cultivation, and improved detection by fluorescent in situ hybridization of a dominant group of human gut bacteria related to *Roseburia* spp. or *Eubacterium rectale*. *Appl.Environ.Microbiol* 72[9], 6371-6376.
- Ammor, M. S., Belen, F. A., and Mayo, B. (2007) Antibiotic resistance in non-enterococcal lactic acid bacteria and bifidobacteria. *Food Microbiol.* 24[6], 559-570.
- Andresen, D., Nickisch-Rosenegk, M., and Bier, F. F. (2009) Helicase dependent OnChip-amplification and its use in multiplex pathogen detection. *Clinica Chimica Acta* 403[1-2], 244-248.
- Anthony, R. M., Brown, T. J., and French, G. L. (2000) Rapid Diagnosis of Bacteremia by Universal Amplification of 23S Ribosomal DNA Followed by Hybridization to an Oligonucleotide Array. *Journal of Clinical Microbiology* 38[2], 781-788.
- Anthony, R. M., Schuitema, A. R. J., Oskam, L., and Klatser, P. R. (2005) Direct detection of *Staphylococcus aureus* mRNA using a flow through microarray. *Journal of Microbiological Methods* 60[1], 47-54.
- Armougom, F., Henry, M., Vialettes, B., Raccach, D., and Raoult, D. (2009) Monitoring bacterial community of human gut microbiota reveals an increase in *Lactobacillus* in obese patients and *Methanogens* in anorexic patients. *PLoS One* 4[9], e7125.
- Bae, J. W., Rhee, S. K., Park, J. R., Chung, W. H., Nam, Y. D., Lee, I., Kim, H., and Park, Y. H. (2005) Development and evaluation of genome-probing microarrays for monitoring lactic acid bacteria. *Applied and Environmental Microbiology* 71[12], 8825-8835.
- Baldi, F., Bianco, M. A., Nardone, G., Pilotto, A., and Zamparo, E. (2009) Focus on acute diarrhoeal disease. *World J.Gastroenterol.* 15[27], 3341-3348.
- Barcenilla, A., Pryde, S. E., Martin, J. C., Duncan, S. H., Stewart, C. S., Henderson, C., and Flint, H. J. (2000) Phylogenetic Relationships of Butyrate-Producing Bacteria from the Human Gut. *Applied and Environmental Microbiology* 66[4], 1654-1661.
- Barl, T., Dobrindt, U., Yu, X., Katcoff, D. J., Sompolinsky, D., Bonacorsi, S., Hacker, J., and Bachmann, T. T. (2008) Genotyping DNA chip for the simultaneous assessment of antibiotic resistance and pathogenic potential of extraintestinal pathogenic *Escherichia coli*. *Int.J.Antimicrob.Agents* 32[3], 272-277.
- Bartlett, J. G. (2002) Clinical practice. Antibiotic-associated diarrhea. *N.Engl.J.Med.* 346[5], 334-339.
- Bekal, S., Brousseau, R., Masson, L., Prefontaine, G., Fairbrother, J., and Harel, J. (2003) Rapid identification of *Escherichia coli* pathotypes by virulence gene detection with DNA microarrays. *Journal of Clinical Microbiology* 41[5], 2113-2125.
- Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., and Sayers, E. W. (2010) GenBank. *Nucleic Acids Res.* 38[Database issue], D46-D51.
- Benters, R., Niemeyer, C. M., Drutschmann, D., Blohm, D., and Wöhrle, D. (2002) DNA microarrays with PAMAM dendritic linker systems. *Nucleic Acids Research* 30[2].
- Berdar, D., Martin Rodriguez, A. C., Herrera, F., and Gijss, M. A. (2008) Label-free detection of DNA with interdigitated micro-electrodes in a fluidic cell. *Lab Chip* 8[2], 302-308.
- Bertucci, F., Finetti, P., Cervera, N., and Birnbaum, D. (2008) Prognostic classification of breast cancer and gene expression profiling. *M S-Medecine Sciences* 24[6-7], 599-606.

- Bezirtzoglou, E., Maipa, V., Chotoura, N., Apazidou, E., Tsiotsias, A., Voidarou, C., Kostakis, D., and Alexopoulos, A. (2006) Occurrence of Bifidobacterium in the intestine of newborns by fluorescence in situ hybridization. *Comp Immunol.Microbiol Infect.Dis.* 29[5-6], 345-352.
- BfR (2010) Erreger von Zoonosen in Deutschland im Jahr 2008: Mitteilungen der Länder zu Lebensmitteln, Tieren, Futtermitteln und Umweltproben. Report, Bundesinstitut für Risikobewertung, BfR-Wissenschaft 06/2010
- Bibiloni, R., Mangold, M., Madsen, K. L., Fedorak, R. N., and Tannock, G. W. (2006) The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients. *Journal of Medical Microbiology* 55[8], 1141-1149.
- Bik, E. M., Eckburg, P. B., Gill, S. R., Nelson, K. E., Purdom, E. A., Francois, F., Perez-Perez, G., Blaser, M. J., and Relman, D. A. (2006) Molecular analysis of the bacterial microbiota in the human stomach. *PNAS of the United States of America* 103[3], 732-737.
- Blessmann, J., Buss, H., Nu, P. A., Dinh, B. T., Ngo, Q. T., Van, A. L., Alla, M. D., Jackson, T. F., Ravdin, J. I., and Tannich, E. (2002) Real-time PCR for detection and differentiation of *Entamoeba histolytica* and *Entamoeba dispar* in fecal samples. *Journal of Clinical Microbiology* 40[12], 4413-4417.
- Bodrossy, L. (2003) Diagnostic oligonucleotide microarrays for microbiology. [2], 43-92. New York, Kluwer Academic Publishers.
- Boesten, R. J., Schuren, F. H., and de Vos, W. M. (2009) A Bifidobacterium mixed-species microarray for high resolution discrimination between intestinal bifidobacteria. *J Microbiol.Methods* 76[3], 269-277.
- Bourlioux, P., Koletzko, B., Guarner, F., and Braesco, V. (2003) The intestine and its microflora are partners for the protection of the host: report on the Danone Symposium "The Intelligent Intestine," held in Paris, June 14, 2002. *Am J Clin.Nutr.* 78[4], 675-683.
- Braasch, D. A. and Corey, D. R. (2001) Locked nucleic acid (LNA): fine-tuning the recognition of DNA and RNA. *Chem.Biol.* 8[1], 1-7.
- Braasch, D. A., Nulf, C. J., and Corey, D. R. (2002) Synthesis and purification of peptide nucleic acids. *Curr.Protoc.Nucleic Acid Chem.* Chapter 4, Unit.
- Bryant, S. and Wilbeck, J. (2007) Vancomycin-resistant *Enterococcus* in critical care areas. *Critical Care Nursing Clinics of North America* 19[1], 69-75.
- Busti, E., Bordoni, R., Castiglioni, B., Monciardini, P., Sosio, M., Donadio, S., Consolandi, C., Rossi, B. L., Battaglia, C., and De Bellis, G. (2002) Bacterial discrimination by means of a universal array approach mediated by LDR (ligase detection reaction). *BMC Microbiol* 2, ArtNb-27.
- Call, D. R., Kang, M. S., Daniels, J., and Besser, T. E. (2006) Assessing genetic diversity in plasmids from *Escherichia coli* and *Salmonella enterica* using a mixed-plasmid microarray. *Journal of Applied Microbiology* 100[1], 15-28.
- Canani, R. B., Cirillo, P., Terrin, G., Cesarano, L., Spagnuolo, M. I., De Vincenzo, A., Albano, F., Passariello, A., De Marco, G., Manguso, F., and Guarino, A. (2007) Probiotics for treatment of acute diarrhoea in children: randomised clinical trial of five different preparations. *British Medical Journal* 335[7615], 340-342.
- Candela, M., Consolandi, C., Severgnini, M., Biagi, E., Castiglioni, B., Vitali, B., De Bellis, G., and Brigidi, P. (2010) High taxonomic level fingerprint of the human intestinal microbiota by Ligase Detection Reaction - Universal Array approach. *BMC Microbiol.* 10[1], 116.
- Cannon, G. A., Carr, M. J., Yandle, Z., Schaffer, K., Kidney, R., Hosny, G., Doyle, A., Ryan, J., Gunson, R., Collins, T., Carman, W. F., Connell, J., and Hall, W. W. (2010) A low density oligonucleotide microarray for the detection of viral and atypical bacterial respiratory pathogens. *J.Virol.Methods* 163[1], 17-24.
- Castiglioni, B., Rizzi, E., Frosini, A., Sivonen, K., Rajaniemi, P., Rantala, A., Mugnai, M. A., Ventura, S., Wilmotte, A., Boute, C., Grubisic, S., Balthasart, P., Consolandi, C., Bordoni, R., Mezzelani, A., Battaglia, C., and De Bellis, G. (2004) Development of a universal microarray based on the ligation detection reaction and 16S rRNA gene polymorphism to target diversity of cyanobacteria. *Applied and Environmental Microbiology* 70[12], 7161-7172.
- Cebra, J. J. (1999) Influences of microbiota on intestinal immune system development. *Am J Clin.Nutr.* 69[5], 1046S-1051S.
- Chan, S. S., Ng, K. C., Lyon, D. J., Cheung, W. L., Cheng, A. F., and Rainer, T. H. (2003) Acute bacterial gastroenteritis: a study of adult patients with positive stool cultures treated in the emergency department. *Emerg.Med.J* 20[4], 335-338.
- Chan, W. C., Maxwell, D. J., Gao, X., Bailey, R. E., Han, M., and Nie, S. (2002) Luminescent quantum dots for multiplexed biological detection and imaging. *Curr.Opin.Biotechnol.* 13[1], 40-46.
- Chen, H., Wang, L., and Li, P. C. (2008) Nucleic acid microarrays created in the double-spiral format on a circular microfluidic disk. *Lab Chip* 8[5], 826-829.
- Chen, S. Y., Tsai, C. N., Chao, H. C., Lai, M. W., Lin, T. Y., Ko, T. Y., and Chiu, C. H. (2009) Acute gastroenteritis caused by multiple enteric pathogens in children. *Epidemiol.Infect.* 137[7], 932-935.
- Chen, W., Li, D., Paulus, B., Wilson, I., and Chadwick, V. S. (2000) Detection of *Listeria monocytogenes* by polymerase chain reaction in intestinal mucosal biopsies from patients with inflammatory bowel disease and controls. *J.Gastroenterol.Hepatol.* 15[10], 1145-1150.
- Cheng, I. F., Chang, H. C., Hou, D., and Chang, H. C. (2007) An integrated dielectrophoretic chip for continuous bioparticle filtering, focusing, sorting, trapping, and detecting. *Biomicrofluidics.* 1[2], 21503.
- Chizhikov, V., Rasooly, A., Chumakov, K., and Levy, D. D. (2001) Microarray Analysis of Microbial Virulence Factors. *Applied and Environmental Microbiology* 67[7], 3258-3263.
- Cho, J. C. and Tiedje, J. M. (2002) Quantitative Detection of Microbial Genes by Using DNA Microarrays. *Applied and Environmental Microbiology* 68[3], 1425-1430.

- Claesson, M. J., O'Sullivan, O., Wang, Q., Nikkila, J., Marchesi, J. R., Smidt, H., de Vos, W. M., Ross, R. P., and O'Toole, P. W. (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE*. 4[8], e6669.
- Clement, B. G., Kehl, L. E., Debord, K. L., and Kitts, C. L. (1998) Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities. *Journal of Microbiological Methods* 31[3], 135-142.
- Corr, S. C., Gahan, C. G. M., and Hill, C. (2007) Impact of selected *Lactobacillus* and *Bifidobacterium* species on *Listeria monocytogenes* infection and the mucosal immune response. *Fems Immunology and Medical Microbiology* 50[3], 380-388.
- D'Souza, A. L. (2007) Ageing and the gut. *Postgrad.Med.J.* 83[975], 44-53.
- de Haan, J. R., Bauerschmidt, S., van Schaik, R. C., Piek, E., Buydens, L. M. C., and Wehrens, R. (2009) Robust ANOVA for microarray data. *Chemometrics and Intelligent Laboratory Systems* 98[1], 38-44.
- de la Rosa, C., Prakash, R., Tilley, P. A., Fox, J. D., and Kaler, K. V. (2007) Integrated microfluidic systems for sample preparation and detection of respiratory pathogen *Bordetella pertussis*. *Conf.Proc.IEEE Eng Med.Biol.Soc.* 2007, 6303-6306.
- de Oliveira, M. A., Ribeiro, E. G. A., Bergamini, A. M. M., and De Martinis, E. C. P. (2010) Quantification of *Listeria monocytogenes* in minimally processed leafy vegetables using a combined method based on enrichment and 16S rRNA real-time PCR. *Food Microbiology* 27[1], 19-23.
- De Preter, V., Vanhoutte, T., Huys, G., Swings, J., De Vuyst, L., Rutgeerts, P., and Verbeke, K. (2007) Effects of *Lactobacillus casei* Shirota, *Bifidobacterium breve*, and oligofructose-enriched inulin on colonic nitrogen-protein metabolism in healthy humans. *Am J Physiol Gastrointest.Liver Physiol* 292[1], G358-G368.
- Delgado, S., Suarez, A., and Mayo, B. (2006) Identification of dominant bacteria in feces and colonic mucosa from healthy Spanish adults by culturing and by 16S rDNA sequence analysis. *Digestive Diseases and Sciences* 51[4], 744-751.
- Denef, V. J., Park, J., Rodrigues, J. L. M., Tsoi, T. V., Hashsham, S. A., and Tiedje, J. M. (2003) Validation of a more sensitive method for using spotted oligonucleotide DNA microarrays for functional genomics studies on bacterial communities. *Environmental Microbiology* 5[10], 933-943.
- Desantis, T. Z., Brodie, E. L., Moberg, J. P., Zubieta, I. X., Piceno, Y. M., and Andersen, G. L. (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. *Microb.Ecol.* 53[3], 371-383.
- Desantis, T. Z., Stone, C. E., Murray, S. R., Moberg, J. P., and Andersen, G. L. (2005) Rapid quantification and taxonomic classification of environmental DNA from both prokaryotic and eukaryotic origins using a microarray. *FEMS Microbiol.Lett.* 245[2], 271-278.
- Dethlefsen, L., Huse, S., Sogin, M. L., and Relman, D. A. (2008) The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol.* 6[11], e280.
- Dinoto, A., Suksomcheep, A., Ishizuka, S., Kimura, H., Hanada, S., Kamagata, Y., Asano, K., Tomita, F., and Yokota, A. (2006) Modulation of rat cecal microbiota by administration of raffinose and encapsulated *Bifidobacterium breve*. *Appl.Environ.Microbiol* 72[1], 784-792.
- Dolezel, J., Bartos, J., Voglmayr, H., and Greilhuber, J. (2003) Nuclear DNA content and genome size of trout and human. *Cytometry Part A* 51A[2], 127-128.
- Domann, E., Hong, G., Imirzalioglu, C., Turschner, S., Kuhle, J., Watzel, C., Hain, T., Hossain, H., and Chakraborty, T. (2003) Culture-independent identification of pathogenic bacteria and polymicrobial infections in the genitourinary tract of renal transplant recipients. *Journal of Clinical Microbiology* 41[12], 5500-5510.
- Donhauser, S. C., Niessner, R., and Seidel, M. (2009) Quantification of *E. coli* DNA on a Flow-through Chemiluminescence Microarray Readout System after PCR Amplification. *Analytical Sciences* 25[5], 669-674.
- Donskey, C. J. (2006) Antibiotic regimens and intestinal colonization with antibiotic-resistant gram-negative bacilli. *Clin.Infect.Dis.* 43 Suppl 2, S62-S69.
- Dore, J., Sghir, A., Hannequart-Gramet, G., Corthier, G., and Pochart, P. (1998) Design and evaluation of a 16S rRNA-targeted oligonucleotide probe for specific detection and quantitation of human faecal *Bacteroides* populations. *Systematic and Applied Microbiology* 21[1], 65-71.
- Draghici, S., Khatri, P., Eklund, A. C., and Szallasi, Z. (2006) Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet.* 22[2], 101-109.
- Du Plessis, E. M. and Dicks, L. M. (1995) Evaluation of random amplified polymorphic DNA (RAPD)-PCR as a method to differentiate *Lactobacillus acidophilus*, *Lactobacillus crispatus*, *Lactobacillus amylovorus*, *Lactobacillus gallinarum*, *Lactobacillus gasseri*, and *Lactobacillus johnsonii*. *Curr.Microbiol* 31[2], 114-118.
- Dufva, M. (2009a) Fabrication of DNA microarray. *Methods Mol.Biol.* 529, 63-79.
- Dufva, M. (2009b) Introduction to microarray technology. *Methods Mol.Biol.* 529, 1-22.
- Dunbar, S. A., Vander Zee, C. A., Oliver, K. G., Kareem, K. L., and Jacobson, J. W. (2003) Quantitative, multiplexed detection of bacterial pathogens: DNA and protein applications of the Luminex LabMAP system. *J.Microbiol.Methods* 53[2], 245-252.
- Duncan, S. H., Barcenilla, A., Stewart, C. S., Pryde, S. E., and Flint, H. J. (2002) Acetate utilization and butyryl coenzyme A (CoA): acetate-CoA transferase in butyrate-producing bacteria from the human large intestine. *Applied and Environmental Microbiology* 68[10], 5186-5190.
- DuPont, H. L. (2007) The growing threat of foodborne bacterial enteropathogens of animal origin. *Clin.Infect.Dis.* 45[10], 1353-1361.
- DuPont, H. L. (2009) Clinical practice. Bacterial diarrhea. *N.Engl.J.Med.* 361[16], 1560-1569.

- Easley, C. J., Karlinsky, J. M., Bienvenue, J. M., Legendre, L. A., Roper, M. G., Feldman, S. H., Hughes, M. A., Hewlett, E. L., Merkel, T. J., Ferrance, J. P., and Landers, J. P. (2006) A fully integrated microfluidic genetic analysis system with sample-in-answer-out capability. *Proc.Natl.Acad.Sci.U.S.A* 103[51], 19272-19277.
- El Fantroussi, S., Urakawa, H., Bernhard, A. E., Kelly, J. J., Noble, P. A., Smidt, H., Yershov, G. M., and Stahl, D. A. (2003) Direct profiling of environmental microbial populations by thermal dissociation analysis of native rRNAs hybridized to oligonucleotide microarrays. *Appl.Environ.Microbiol* 69[4], 2377-2382.
- Eller, C., Crabill, M. R., and Bryant, M. P. (1971) Anaerobic roll tube media for nonselective enumeration and isolation of bacteria in human feces. *Appl.Microbiol.* 22[4], 522-529.
- Espineira, M., Atanassova, M., Vieites, J. M., and Santaclara, F. J. (2010) Validation of a method for the detection of five species, serogroups, biotypes and virulence factors of *Vibrio* by multiplex PCR in fish and seafood. *Food Microbiology* 27[1], 122-131.
- Ethelberg, S., Olsen, K. E. P., Gerner-Smidt, P., and Molbak, K. (2007) The significance of the number of submitted samples and patient-related factors for faecal bacterial diagnostics. *Clinical Microbiology and Infection* 13[11], 1095-1099.
- Fairchild, A. S., Smith, J. L., Idris, U., Lu, J., Sanchez, S., Purvis, L. B., Hofacre, C., and Lee, M. D. (2005) Effects of Orally Administered Tetracycline on the Intestinal Community Structure of Chickens and on tet Determinant Carriage by Commensal Bacteria and *Campylobacter jejuni*. *Applied and Environmental Microbiology* 71[10], 5865-5872.
- Fallani, M., Rigottier-Gois, L., Aguilera, M., Bridonneau, C., Collignon, A., Edwards, C. A., Corthier, G., and Dore, J. (2006) *Clostridium difficile* and *Clostridium perfringens* species detected in infant faecal microbiota using 16S rRNA targeted probes. *J Microbiol Methods* 67[1], 150-161.
- Farfan, M. J., Garay, T. A., Prado, C. A., Filliol, I., Ulloa, M. T., and Toro, C. S. (2010) A new multiplex PCR for differential identification of *Shigella flexneri* and *Shigella sonnei* and detection of *Shigella* virulence determinants. *Epidemiol.Infect.* 138[4], 525-533.
- Farthing, M. (2002) *Infectious Diarrhoea*. [35], 409-423. Martin Duntiz Ltd, a member of the Taylor & Francis group.
- Favier, C. F., de Vos, W. M., and Akkermans, A. D. (2003) Development of bacterial and bifidobacterial communities in feces of newborn babies. *Anaerobe* 9[5], 219-229.
- FDA. (2010) *Bad Bug Book*.
<http://www.fda.gov/Food/FoodSafety/FoodborneIllness/FoodborneIllnessFoodbornePathogensNaturalToxins/BadBugBook/ucm071284.htm> Food and Drug Administration. Food and Drug Administration.
- Felske, A., Engelen, B., Nubel, U., and Backhaus, H. (1996) Direct ribosome isolation from soil to extract bacterial rRNA for community analysis. *Applied and Environmental Microbiology* 62[11], 4162-4167.
- Feng, Y. N., Gong, J., Yu, H., Jin, Y. P., Zhu, J., and Han, Y. M. (2010) Identification of changes in the composition of ileal bacterial microbiota of broiler chickens infected with *Clostridium perfringens*. *Veterinary Microbiology* 140[1-2], 116-121.
- Fernandez-Perez, L., Novoa, J., Stahlberg, N., Santana-Farre, R., Boronat, M., Marrero, D., Henriquez-Hernandez, L., Norstedt, G., and Flores-Morales, A. (2010) The effect of in vivo growth hormone treatment on blood gene expression in adults with growth hormone deficiency reveals potential biomarkers to monitor growth hormone therapy. *Clinical Endocrinology* 72[6], 800-806.
- Fitzgerald, C., Collins, M., van Duyn, S., Mikoleit, M., Brown, T., and Fields, P. (2007) Multiplex, bead-based suspension array for molecular determination of common *Salmonella* serogroups. *Journal of Clinical Microbiology* 45[10], 3323-3334.
- Fons, M., Gomez, A., and Karjalainen, T. (2000) Mechanisms of Colonisation and Colonisation Resistance of the Digestive Tract Part 2: Bacteria/Bacteria Interactions. *Microbial Ecology in Health and Disease* 12[1 supp 2], 240-246.
- Fordtran, J. S. (2006) Colitis due to *Clostridium difficile* toxins: underdiagnosed, highly virulent, and nosocomial. *Proc.(Bayl.Univ Med.Cent.)* 19[1], 3-12.
- Foultier, B., Moreno-Hagelsieb, L., Flandre, D., and Remacle, J. (2005) Comparison of DNA detection methods using nanoparticles and silver enhancement. *IEE.Proc.Nanobiotechnol.* 152[1], 3-12.
- Francois, P., Bento, M., Vaudaux, P., and Schrenzel, J. (2003) Comparison of fluorescence and resonance light scattering for highly sensitive microarray detection of bacterial pathogens. *J.Microbiol.Methods* 55[3], 755-762.
- Franklin, R. B., Taylor, D. R., and Mills, A. L. (1999) Characterization of microbial communities using randomly amplified polymorphic DNA (RAPD). *J.Microbiol.Methods* 35[3], 225-235.
- Franks, A. H., Harmsen, H. J., Raangs, G. C., Jansen, G. J., Schut, F., and Welling, G. W. (1998) Variations of bacterial populations in human feces measured by fluorescent in situ hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. *Applied and Environmental Microbiology* 64[9], 3336-3345.
- Fritz, J., Cooper, E. B., Gaudet, S., Sorger, P. K., and Manalis, S. R. (2002) Electronic detection of DNA by its intrinsic molecular charge. *Proc.Natl.Acad.Sci.U.S.A* 99[22], 14142-14146.
- Fu, Z., Rogelj, S., and Kieft, T. L. (2005) Rapid detection of *Escherichia coli* O157:H7 by immunomagnetic separation and real-time PCR. *Int.J.Food Microbiol.* 99[1], 47-57.
- Fuchs, B. M., Syutsubo, K., Ludwig, W., and Amann, R. (2001) In situ accessibility of *Escherichia coli* 23S rRNA to fluorescently labeled oligonucleotide probes. *Applied and Environmental Microbiology* 67[2], 961-968.
- Fugett, E. B., Schoonmaker-Bopp, D., Dumas, N. B., Corby, J., and Wiedmann, M. (2007) Pulsed field gel electrophoresis (PFGE) analysis of temporally matched *Listeria monocytogenes* isolates from human clinical cases, foods, ruminant farms, and urban and natural environments reveals source associated as well as widely distributed PFGE types. *J Clin.Microbiol.*

- Garcia, P., Valenzuela, N., Rodriguez, M. V., Leon, E., and Fernandez, H. (2009) Antimicrobial susceptibility of *Campylobacter jejuni* isolates from stool cultures in Santiago, Chile. *Revista Chilena de Infectologia* 26[6], 511-514.
- Gascon, J. (2006) Epidemiology, etiology and pathophysiology of traveler's diarrhea. *Digestion* 73 Suppl 1, 102-108.
- Gerry, N. P., Witowski, N. E., Day, J., Hammer, R. P., Barany, G., and Barany, F. (1999) Universal DNA microarray method for multiplex detection of low abundance point mutations. *Journal of Molecular Biology* 292[2], 251-262.
- Gielen, J., Erhard, M., Kallow, W., Kronke, M., and Krut, O. (2007) Rapid pathogen identification by MALDI-TOF mass spectrometry/SARAMIS database in clinical microbiological routine diagnostics. *International Journal of Antimicrobial Agents* 29, S314.
- Glazer, A. N. and Mathies, R. A. (1997b) Energy-transfer fluorescent reagents for DNA analyses. *Curr.Opin.Biotechnol.* 8[1], 94-102.
- Glazer, A. N. and Mathies, R. A. (1997a) Energy-transfer fluorescent reagents for DNA analyses. *Curr.Opin.Biotechnol.* 8[1], 94-102.
- Gleesen, A. S., Grarup, C., Dargis, R., Andresen, K., Christensen, J. J., and Kemp, M. (2008) PCR and DNA sequencing in establishing the aetiology of bacterial infections in children. *Apms* 116[9], 811-815.
- Goldenberg, O., Herrmann, S., Adam, T., Marjoram, G., Hong, G., Gobel, U. B., and Graf, B. (2005) Use of denaturing high-performance liquid chromatography for rapid detection and identification of seven *Candida* species. *Journal of Clinical Microbiology* 43[12], 5912-5915.
- Goldenberg, O., Herrmann, S., Marjoram, G., Noyer-Weidner, M., Hong, G., Bereswill, S., and Gobel, U. B. (2007) Molecular monitoring of the intestinal flora by denaturing high performance liquid chromatography. *J Microbiol Methods* 68[1], 94-105.
- Guan, L. L., Hagen, K. E., Tannock, G. W., Korver, D. R., Fasenko, G. M., and Allison, G. E. (2003) Detection and identification of *Lactobacillus* species in crops of broilers of different ages by using PCR-denaturing gradient gel electrophoresis and amplified ribosomal DNA restriction analysis. *Applied and Environmental Microbiology* 69[11], 6750-6757.
- Guarner, F. (2006) Enteric flora in health and disease. *Digestion* 73, 5-12.
- Guisseppi-Elie, A. and Lingerfelt, L. (2005) Impedimetric detection of DNA hybridization: Towards near-patient DNA diagnostics. *Immobilisation of Dna on Chips I* 260, 161-186.
- Guo, Z., Guilfoyle, R. A., Thiel, A. J., Wang, R. F., and Smith, L. M. (1994) Direct Fluorescence Analysis of Genetic Polymorphisms by Hybridization with Oligonucleotide Arrays on Glass Supports. *Nucleic Acids Research* 22[24], 5456-5465.
- Haarman, M. and Knol, J. (2005) Quantitative Real-Time PCR Assays To Identify and Quantify Fecal *Bifidobacterium* Species in Infants Receiving a Prebiotic Infant Formula. *Applied and Environmental Microbiology* 71[5], 2318-2324.
- Haeberle, S. and Zengerle, R. (2007) Microfluidic platforms for lab-on-a-chip applications. *Lab Chip* 7[9], 1094-1110.
- Hajdukiewicz, J., Boland, S., Kavanagh, P., and Leech, D. (2010) An enzyme-amplified amperometric DNA hybridisation assay using DNA immobilised in a carboxymethylated dextran film anchored to a graphite surface. *Biosens.Bioelectron.* 25[5], 1037-1042.
- Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001) PAST: Paleontological Statistics Software Package for Education and Data Analysis. *Palaeontologia Electronica* 4[1].
- Han, M., Gao, X., Su, J. Z., and Nie, S. (2001) Quantum-dot-tagged microbeads for multiplexed optical coding of biomolecules. *Nat.Biotechnol.* 19[7], 631-635.
- Hang, T. C. and Guisseppi-Elie, A. (2004) Frequency dependent and surface characterization of DNA immobilization and hybridization. *Biosensors & Bioelectronics* 19[11], 1537-1548.
- Harmsen, H. J. M., Raangs, G. C., He, T., Degener, J. E., and Welling, G. W. (2002) Extensive set of 16S rRNA-based probes for detection of bacteria in human feces. *Applied and Environmental Microbiology* 68[6], 2982-2990.
- Harmsen, H. J. M., Wildeboer-Veloo, A. C. M., Grijpstra, J., Knol, J., Degener, J. E., and Welling, G. W. (2000) Development of 16S rRNA-Based Probes for the *Coriobacterium* Group and the *Atopobium* Cluster and Their Application for Enumeration of *Coriobacteriaceae* in Human Feces from Volunteers of Different Age Groups. *Applied and Environmental Microbiology* 66[10], 4523-4527.
- Hashimoto, M., Barany, F., and Soper, S. A. (2006) Polymerase chain reaction/ligase detection reaction/hybridization assays using flow-through microfluidic devices for the detection of low-abundant DNA point mutations. *Biosens.Bioelectron.* 21[10], 1915-1923.
- Hassibi, A., Vikalo, H., Riechmann, J. L., and Hassibi, B. (2009) Real-time DNA microarray analysis. *Nucleic Acids Research* 37[20].
- Hayashi, H., Sakamoto, M., and Benno, Y. (2002) Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. *Microbiol Immunol.* 46[8], 535-548.
- Hayashi, H., Sakamoto, M., Kitahara, M., and Benno, Y. (2006) Diversity of the *Clostridium coccoides* group in human fecal microbiota as determined by 16S rRNA gene library. *FEMS Microbiol Lett.* 257[2], 202-207.
- Hayashi, H., Takahashi, R., Nishi, T., Sakamoto, M., and Benno, Y. (2005) Molecular analysis of jejunal, ileal, caecal and recto-sigmoidal human colonic microbiota using 16S rRNA gene libraries and terminal restriction fragment length polymorphism. *Journal of Medical Microbiology* 54[Pt 11], 1093-1101.
- Heid, C. A., Stevens, J., Livak, K. J., and Williams, P. M. (1996) Real time quantitative PCR. *Genome Research* 6[10], 986-994.

- Heilig, H. G. H. J., Zoetendal, E. G., Vaughan, E. E., Marteau, P., Akkermans, A. D. L., and de Vos, W. M. (2002) Molecular Diversity of *Lactobacillus* spp. and Other Lactic Acid Bacteria in the Human Intestine as Determined by Specific Amplification of 16S Ribosomal DNA. *Applied and Environmental Microbiology* 68[1], 114-123.
- Henker, J., Laass, M., Blokhin, B. M., Bolbot, Y. K., Maydannik, V. G., Elze, M., Wolff, C., and Schulze, J. (2007) The probiotic *Escherichia coli* strain Nissle 1917 (EcN) stops acute diarrhoea in infants and toddlers. *European Journal of Pediatrics* 166[4], 311-318.
- Hill, M. J. and Drasar, B. S. (1975) The Normal Colonic Bacterial-Flora. *Gut* 16[4], 318-323.
- Hold, G. L., Schwierz, A., Aminov, R. I., Blaut, M., and Flint, H. J. (2003) Oligonucleotide Probes That Detect Quantitatively Significant Groups of Butyrate-Producing Bacteria in Human Feces. *Applied and Environmental Microbiology* 69[7], 4320-4324.
- Holland, P. M., Abramson, R. D., Watson, R., and Gelfand, D. H. (1991) Detection of Specific Polymerase Chain-Reaction Product by Utilizing the 5'-]3' Exonuclease Activity of *Thermus-Aquaticus* Dna-Polymerase. *Proceedings of the National Academy of Sciences of the United States of America* 88[16], 7276-7280.
- Hooper, L. V. and Macpherson, A. J. (2010) Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nature Reviews Immunology* 10[3], 159-169.
- Hooper, L. V., Wong, M. H., Thelin, A., Hansson, L., Falk, P. G., and Gordon, J. I. (2001) Molecular Analysis of Commensal Host-Microbial Relationships in the Intestine. *Science* 291[5505], 881-884.
- Hopkins, M. J., Sharp, R., and Macfarlane, G. T. (2001) Age and disease related changes in intestinal bacterial populations assessed by cell culture, 16S rRNA abundance, and community cellular fatty acid profiles. *Gut* 48[2], 198-205.
- Hörmann, B. (2006) Grundlegende Studien zur Entwicklung eines Quantifizierungs-Systems auf DNA-Microarrays (Studienarbeit). 1-66. Universität Stuttgart, Institut für Technische Biochemie.
- Hormansperger, G. and Haller, D. (2010) Molecular crosstalk of probiotic bacteria with the intestinal immune system: Clinical relevance in the context of inflammatory bowel disease. *International Journal of Medical Microbiology* 300[1], 63-73.
- Hoshiko, M. (1994) Laboratory diagnosis of infectious diarrhea. *Pediatr. Ann.* 23[10], 570-574.
- Hoshino, T., Furukawa, K., Tsuneda, S., and Inamori, Y. (2007) RNA microarray for estimating relative abundance of 16S rRNA in microbial communities. *J. Microbiol. Methods* 69[2], 406-410.
- Hou, X. L., Cao, Q. Y., Jia, H. Y., and Chen, Z. (2008) Pyrosequencing analysis of the *gyrB* gene to differentiate bacteria responsible for diarrheal diseases. *Eur. J. Clin. Microbiol. Infect. Dis.* 27[7], 587-596.
- Hruskova, V. and Kaclikova, E. (2009) Rapid and sensitive detection of pathogenic *Yersinia enterocolitica* strains in food using selective enrichment and real-time PCR. *Journal of Food and Nutrition Research* 48[2], 100-108.
- Hsu, C. F., Tsai, T. Y., and Pan, T. M. (2005) Use of the Duplex TaqMan PCR System for Detection of Shiga-Like Toxin-Producing *Escherichia coli* O157. *Journal of Clinical Microbiology* 43[6], 2668-2673.
- Huang, A., Li, J. W., Shen, Z. Q., Wang, X. W., and Jin, M. (2006) High-throughput identification of clinical pathogenic fungi by hybridization to an oligonucleotide microarray. *Journal of Clinical Microbiology* 44[9], 3299-3305.
- Huijsdens, X. W., Linskens, R. K., Mak, M. T., Meuwissen, S. G. M., Vandenbroucke-Grauls, C. M. J. E., and Savelkoul, P. H. M. (2002) Quantification of bacteria adherent to gastrointestinal mucosa by real-time PCR. *Journal of Clinical Microbiology* 40[12], 4423-4427.
- Huijsdens, X. W., Linskens, R. K., Taspinar, H., Meuwissen, S. G. M., Vandenbroucke-Grauls, C. M. J. E., and Savelkoul, P. H. M. (2003) *Listeria monocytogenes* and inflammatory bowel disease - Detection of *Listeria* species in intestinal mucosal biopsies by real-time PCR. *Scandinavian Journal of Gastroenterology* 38[3], 332-333.
- Hurtle, W., Shoemaker, D., Henchal, E., and Norwood, D. (2002) Denaturing HPLC for identifying bacteria. *Biotechniques* 33[2], 386-1.
- Huser, T. (2008) Nano-biophotonics: new tools for chemical nano-analytics. *Curr. Opin. Chem. Biol.* 12[5], 497-504.
- Idborg, H., Zamani, L., Edlund, P. O., Schuppe-Koistinen, I., and Jacobsson, S. P. (2005) Metabolic fingerprinting of rat urine by LC/MS Part 2. Data pretreatment methods for handling of complex data. *Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences* 828[1-2], 14-20.
- Iizuka, M., Konno, S., Itou, H., Chihara, J., Toyoshima, I., Horie, Y., Sasaki, K., Sato, A., Shindo, K., and Watanabe, S. (2004) Novel evidence suggesting *Clostridium difficile* is present in human gut microbiota more frequently than previously suspected. *Microbiol. Immunol.* 48[11], 889-892.
- Ince, J. and McNally, A. (2009) Development of rapid, automated diagnostics for infectious disease: advances and challenges. *Expert Rev. Med. Devices* 6[6], 641-651.
- Ishkanian, A. S., Malloff, C. A., Watson, S. K., deLeeuw, R. J., Chi, B., Coe, B. P., Snijders, A., Albertson, D. G., Pinkel, D., Marra, M. A., Ling, V., MacAulay, C., and Lam, W. L. (2004) A tiling resolution DNA microarray with complete coverage of the human genome. *Nature Genetics* 36[3], 299-303.
- Isola, N. R., Allman, S. L., Golovlev, V. V., and Chen, C. H. (2001) MALDI-TOF mass spectrometric method for detection of hybridized DNA oligomers. *Anal. Chem.* 73[9], 2126-2131.
- Jaaskelainen, A. J. and Maunula, L. (2006) Applicability of microarray technique for the detection of noro- and astroviruses. *Journal of Virological Methods* 136[1-2], 210-216.
- Janka, A., Friedrich, A. W., and Karch, H. (2001) [What should be used for salmonella and company? Antibiotic therapy for bacterial diarrhea]. *Pharm. Unserer Zeit* 30[5], 406-410.
- Jarvis, J., Melin, J., Goransson, J., Stenberg, J., Fredriksson, S., Gonzalez-Rey, C., Bertilsson, S., and Nilsson, M. (2006) Digital quantification using amplified single-molecule detection. *Nature Methods* 3[9], 725-727.

- Jernberg, C., Sullivan, A., Edlund, C., and Jansson, J. K. (2005) Monitoring of Antibiotic-Induced Alterations in the Human Intestinal Microflora and Detection of Probiotic Strains by Use of Terminal Restriction Fragment Length Polymorphism. *Applied and Environmental Microbiology* 71[1], 501-506.
- Jin, D. Z., Wen, S. Y., Chen, S. H., Lin, F., and Wang, S. Q. (2006) Detection and identification of intestinal pathogens in clinical specimens using DNA microarrays. *Mol. Cell Probes* 20, 337-347.
- Jonnalagadda, S. and Srinivasan, R. (2008) Principal components analysis based methodology to identify differentially expressed genes in time-course microarray data. *BMC Bioinformatics* 9.
- Justesen, U. S., Skov, M. N., Knudsen, E., Holt, H. M., Sogaard, P., and Justesen, T. (2010) 16S rRNA Gene Sequencing in Routine Identification of Anaerobic Bacteria Isolated from Blood Cultures. *Journal of Clinical Microbiology* .
- Kajander, K., Myllyluoma, E., Rajilic-Stojanovic, M., Kyronpalo, S., Rasmussen, M., Jarvenpaa, S., Zoetendal, E. G., de Vos, W. M., Vapaatalo, H., and Korpela, R. (2008) Clinical trial: multispecies probiotic supplementation alleviates the symptoms of irritable bowel syndrome and stabilizes intestinal microbiota. *Aliment. Pharmacol. Ther.* 27[1], 48-57.
- Kakinuma, K., Fukushima, M., and Kawaguchi, R. (2003) Detection and identification of *Escherichia coli*, *Shigella*, and *Salmonella* by microarrays using the *gyrB* gene. *Biotechnol. Bioeng.* 83[6], 721-728.
- Kanazawa, Y., Ishikawa, T., Shimizu, K., and Inaba, S. (2007) Enterohemorrhagic *Escherichia coli* outbreaks in nursery and primary schools. *Jpn. J. Infect. Dis.* 60[5], 326-327.
- Kane, M. D., Jatkoa, T. A., Stumpf, C. R., Lu, J., Thomas, J. D., and Madore, S. J. (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Research* 28[22], 4552-4557.
- Kararli, T. T. (1995) Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. *Biopharm. Drug Dispos.* 16[5], 351-380.
- Karlin-Neumann, G., Sedova, M., Falkowski, M., Wang, Z., Lin, S., and Jain, M. (2007) Application of quantum dots to multicolor microarray experiments: four-color genotyping. *Methods Mol. Biol.* 374, 239-251.
- Karsunke, X. Y. Z., Niessner, R., and Seidel, M. (2009) Development of a multichannel flow-through chemiluminescence microarray chip for parallel calibration and detection of pathogenic bacteria. *Analytical and Bioanalytical Chemistry* 395[6], 1623-1630.
- Ke, D., Picard, F. J., Martineau, F., Menard, C., Roy, P. H., Ouellette, M., and Bergeron, M. G. (1999) Development of a PCR assay for rapid detection of enterococci. *J Clin. Microbiol.* 37[11], 3497-3503.
- Kelly, C. P. and LaMont, J. T. (2008) *Clostridium difficile* - More difficult than ever. *New England Journal of Medicine* 359[18].
- Keramas, G., Bang, D. D., Lund, M., Madsen, M., Bunkenborg, H., Telleman, P., and Christensen, C. B. V. (2004) Use of culture, PCR analysis, and DNA microarrays for detection of *Campylobacter jejuni* and *Campylobacter coli* from chicken feces. *Journal of Clinical Microbiology* 42[9], 3985-3991.
- Khodakov, D. A., Zakharova, N. V., Gryadunov, D. A., Filatov, F. P., Zasedatelev, A. S., and Mikhailovich, V. M. (2008) An oligonucleotide microarray for multiplex real-time PCR identification of HIV-1/HBV and HCV. *Biotechniques* 44[2], 241-+.
- Kim, B. C. and Gu, M. B. (2007) Discrimination of toxic impacts of various chemicals using chemical-gene expression profiling of *Escherichia coli* DNA microarray. *Process Biochemistry* 42[3], 392-400.
- Kirk, M. D., Veitch, M. G., and Hall, G. V. (2010) Gastroenteritis and food-borne disease in elderly people living in long-term care. *Clin. Infect. Dis.* 50[3], 397-404.
- Kohara, T., Hayashi, S., Hamatsuka, J., Nishitani, Y., Masuda, Y., and Osawa, R. (2006) Intra-specific composition and succession of *Bifidobacterium longum* in human feces. *Syst Appl. Microbiol.* .
- Kolbert, C. P. and Persing, D. H. (1999) Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. *Current Opinion in Microbiology* 2[3], 299-305.
- Koletzko, S. and Osterrieder, S. (2009) Acute infectious diarrhea in children. *Dtsch. Arztebl. Int.* 106[33], 539-547.
- Kostic, T., Weilharter, A., Rubino, S., Delogu, G., Uzzau, S., Rudi, K., Sessitsch, A., and Bodrossy, L. (2007) A microbial diagnostic microarray technique for the sensitive detection and identification of pathogenic bacteria in a background of nonpathogens. *Analytical Biochemistry* 360[2], 244-254.
- Kretzer, J. W., Lehmann, R., Schmelcher, M., Banz, M., Kim, K. P., Korn, C., and Loessner, M. J. (2007) Use of high-affinity cell wall-binding domains of bacteriophage endolysins for immobilization and separation of bacterial cells. *Appl. Environ. Microbiol.* 73[6], 1992-2000.
- Krizova, J., Spanova, A., and Rittich, B. (2006) Evaluation of amplified ribosomal DNA restriction analysis (ARDRA) and species-specific PCR for identification of *Bifidobacterium* species. *Syst Appl. Microbiol.* 29[1], 36-44.
- Kuijper, E. J., Coignard, B., and Tull, P. (2006) Emergence of *Clostridium difficile*-associated disease in North America and Europe. *Clinical Microbiology and Infection* 12, 2-18.
- Kwakye, S. and Baemner, A. (2003) A microfluidic biosensor based on nucleic acid sequence recognition. *Anal. Bioanal. Chem.* 376[7], 1062-1068.
- Lamhoujeb, S., Charest, H., Fliss, I., Ngazoa, S., and Jean, J. (2009) Real-time molecular beacon NASBA for rapid and sensitive detection of norovirus GII in clinical samples. *Can. J. Microbiol.* 55[12], 1375-1380.
- Lamps, L. W. (2007) Infective disorders of the gastrointestinal tract. *Histopathology* 50[1], 55-63.
- Lan, P. T. N., Hayashi, H., Sakamoto, M., and Benno, Y. (2002) Phylogenetic analysis of cecal microbiota in chicken by the use of 16S rDNA clone libraries. *Microbiology and Immunology* 46[6], 371-382.
- Lane, S., Evermann, J., Loge, F., and Call, D. R. (2004) Amplicon secondary structure prevents target hybridization to oligonucleotide microarrays. *Biosens. Bioelectron.* 20[4], 728-735.
- Langendijk, P. S., Schut, F., Jansen, G. J., Raangs, G. C., Kamphuis, G. R., Wilkinson, M. H. F., and Welling, G. W. (1995) Quantitative Fluorescence In-Situ Hybridization of *Bifidobacterium* Spp with Genus-Specific 16S

- Ribosomal-RNA-Targeted Probes and Its Application in Fecal Samples. *Applied and Environmental Microbiology* 61[8], 3069-3075.
- Lauri, A. and Mariani, P. O. (2009) Potentials and limitations of molecular diagnostic methods in food safety. *Genes and Nutrition* 4[1], 1-12.
- Lay, C., Sutren, M., Rochet, V., Saunier, K., Dore, J., and Rigottier-Gois, L. (2005) Design and validation of 16S rRNA probes to enumerate members of the *Clostridium leptum* subgroup in human faecal microbiota. *Environmental Microbiology* 7[7], 933-946.
- Lee, D., Chen, P. J., and Lee, G. B. (2010) The evolution of real-time PCR machines to real-time PCR chips. *Biosensors & Bioelectronics* 25[7], 1820-1824.
- Lee, L. G., Connell, C. R., and Bloch, W. (1993) Allelic Discrimination by Nick-Translation PCR with Fluorogenic Probes. *Nucleic Acids Research* 21[16], 3761-3766.
- Lehner, A., Loy, A., Behr, T., Gaenge, H., Ludwig, W., Wagner, M., and Schleifer, K. H. (2005) Oligonucleotide microarray for identification of *Enterococcus* species. *FEMS Microbiol.Lett.* 246[1], 133-142.
- Leinberger, D. M., Grimm, V., Rubtsova, M., Weile, J., Schroppe, K., Wichelhaus, T. A., Knabbe, C., Schmid, R. D., and Bachmann, T. T. (2009) Integrated Detection of Extended Spectrum Beta-Lactam Resistance by DNA Microarray-Based Genotyping of TEM, SHV, and CTX-M genes. *Journal of Clinical Microbiology* .
- Leinberger, D. M., Schumacher, U., Autenrieth, I. B., and Bachmann, T. T. (2005) Development of a DNA microarray for detection and identification of fungal pathogens involved in invasive mycoses. *J Clin.Microbiol* 43[10], 4943-4953.
- Lemarchand, K., Masson, L., and Brousseau, R. (2004) Molecular biology and DNA microarray technology for microbial quality monitoring of water. *Critical Reviews in Microbiology* 30[3], 145-172.
- Lenigk, R., Liu, R. H., Athavale, M., Chen, Z., Ganser, D., Yang, J., Rauch, C., Liu, Y., Chan, B., Yu, H., Ray, M., Marrero, R., and Grodzinski, P. (2002) Plastic biochannel hybridization devices: a new concept for microfluidic DNA arrays. *Anal.Biochem.* 311[1], 40-49.
- Leparc, G. G., Tuchler, T., Striedner, G., Bayer, K., Sykacek, P., Hofacker, I. L., and Kreil, D. P. (2009) Model-based probe set optimization for high-performance microarrays. *Nucleic Acids Research* 37[3].
- Letowski, J., Brousseau, R., and Masson, L. (2004) Designing better probes: effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide microarrays. *Journal of Microbiological Methods* 57[2], 269-278.
- Ley, R. E., Backhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D., and Gordon, J. I. (2005) Obesity alters gut microbial ecology. *Proc.Natl.Acad.Sci.U.S A* 102[31], 11070-11075.
- Li, Y., Liu, D., Cao, B., Han, W., Liu, Y., Liu, F., Guo, X., Bastin, D. A., Feng, L., and Wang, L. (2006) Development of a serotype-specific DNA microarray for identification of some *Shigella* and pathogenic *Escherichia coli* strains. *Journal of Clinical Microbiology* 44[12], 4376-4383.
- Limor, J. R., Lal, A. A., and Xiao, L. (2002) Detection and differentiation of *Cryptosporidium* parasites that are pathogenic for humans by real-time PCR. *Journal of Clinical Microbiology* 40[7], 2335-2338.
- Lin, B. C., Blaney, K. M., Malanoski, A. P., Ligler, A. G., Schnur, J. M., Metzgar, D., Russell, K. L., and Stenger, D. A. (2007) Using a resequencing microarray as a multiple respiratory pathogen detection assay. *Journal of Clinical Microbiology* 45[2], 443-452.
- Lin, B., Wang, Z., Vora, G. J., Thornton, J. A., Schnur, J. M., Thach, D. C., Blaney, K. M., Ligler, A. G., Malanoski, A. P., Santiago, J., Walter, E. A., Agan, B. K., Metzgar, D., Seto, D., Daum, L. T., Kruzelock, R., Rowley, R. K., Hanson, E. H., Tibbetts, C., and Stenger, D. A. (2006) Broad-spectrum respiratory tract pathogen identification using resequencing DNA microarrays. *Genome Research* 16[4], 527-535.
- Lin, J. C. (2010) Protein microarrays for cancer diagnostics and therapy. *Med.Princ.Pract.* 19[4], 247-254.
- Lin, L. L., Zhou, J. Z., Zhang, Y., and Lin, Z. H. (2008) High sensitive electrochemical detection of sequence-specific DNA using low current voltammetry. *Electroanalysis* 20[16], 1798-1804.
- Lind, K., Anders Ståhlberg, Neven Zoric, and Mikael Kubista. (2006) Combining sequence-specific probes and DNA binding dyes in real-time PCR for specific nucleic acid quantification and melting curve analysis. *Biotechniques* 40, 315-319.
- Liu, H., Wang, H., Shi, Z., Wang, H., Yang, C., Silke, S., Tan, W., and Lu, Z. (2006) TaqMan probe array for quantitative detection of DNA targets. *Nucleic Acids Research* 34[1], e4.
- Liu, R. H., Yang, J., Lenigk, R., Bonanno, J., and Grodzinski, P. (2004) Self-contained, fully integrated biochip for sample preparation, polymerase chain reaction amplification, and DNA microarray detection. *Anal.Chem.* 76[7], 1824-1831.
- Liu, W. T., Mirzabekov, A. D., and Stahl, D. A. (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. *Environ.Microbiol.* 3[10], 619-629.
- Liu-Stratton, Y. L., Roy, S., and Sen, C. K. (2004) DNA microarray technology in nutraceutical and food safety. *Toxicology Letters* 150[1], 29-42.
- Lofmark, S., Jernberg, C., Jansson, J. K., and Edlund, C. (2006) Clindamycin-induced enrichment and long-term persistence of resistant *Bacteroides* spp. and resistance genes. *J Antimicrob.Chemother.* 58, 1160-1167.
- Loy, A. and Bodrossy, L. (2006) Highly parallel microbial diagnostics using oligonucleotide microarrays. *Clin.Chim.Acta* 363[1-2], 106-119.
- Loy, A., Maixner, F., Wagner, M., and Horn, M. (2007) probeBase - an online resource for rRNA-targeted oligonucleotide probes: new features 2007. *Nucleic Acids Research* 35[suppl_1], D800-D804.
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A.,

- Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., and Schleifer, K. H. (2004) ARB: a software environment for sequence data. *Nucleic Acids Research* 32[4], 1363-1371.
- Lui, C., Cady, N. C., and Batt, C. A. (2009) Nucleic Acid-based Detection of Bacterial Pathogens Using Integrated Microfluidic Platform Systems. *Sensors* 9[5], 3713-3744.
- Luna, R. A., Fasciano, L. R., Jones, S. C., Boyanton, B. L., Jr., Ton, T. T., and Versalovic, J. (2007) DNA pyrosequencing-based bacterial pathogen identification in a pediatric hospital setting. *Journal of Clinical Microbiology* 45[9], 2985-2992.
- Lyon, E. and Wittwer, C. T. (2009) LightCycler technology in molecular diagnostics. *J.Mol.Diagn.* 11[2], 93-101.
- Mafu, A. A., Pitre, M., and Sirois, S. (2009) Real-Time PCR as a Tool for Detection of Pathogenic Bacteria on Contaminated Food Contact Surfaces by Using a Single Enrichment Medium. *Journal of Food Protection* 72[6], 1310-1314.
- Mairhofer, J., Roppert, K., and Ertl, P. (2009) Microfluidic Systems for Pathogen Sensing: A Review. *Sensors* 9[6], 4804-4823.
- Majtan, T., Majtanova, L., Timko, J., and Majtan, V. (2007) Oligonucleotide microarray for molecular characterization and genotyping of *Salmonella* spp. strains. *J Antimicrob.Chemother.* 60[5], 937-946.
- Malinen, E., Kassinen, A., Rinttila, T., and Palva, A. (2003) Comparison of real-time PCR with SYBR Green I or 5'-nuclease assays and dot-blot hybridization with rDNA-targeted oligonucleotide probes in quantification of selected faecal bacteria. *Microbiology* 149[Pt 1], 269-277.
- Malinen, E., Rinttila, T., Kajander, K., Mattu, J., Kassinen, A., Krogius, L., Saarela, M., Korpela, R., and Palva, A. (2005) Analysis of the fecal microbiota of irritable bowel syndrome patients and healthy controls with real-time PCR. *Am.J.Gastroenterol.* 100[2], 373-382.
- Mandrell, R. E., Harden, L. A., Bates, A., Miller, W. G., Haddon, W. F., and Fagerquist, C. K. (2005) Speciation of *Campylobacter coli*, *C. jejuni*, *C. helveticus*, *C. lari*, *C. sputorum*, and *C. upsaliensis* by Matrix-Assisted Laser Desorption Ionization-Time of Flight Mass Spectrometry. *Applied and Environmental Microbiology* 71[10], 6292-6307.
- Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K. H. (1996) Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology-Uk* 142, 1097-1106.
- Manz, W., Amann, R., Ludwig, W., Wagner, M., and Schleifer, K. H. (1992) Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria - Problems and Solutions. *Systematic and Applied Microbiology* 15[4], 593-600.
- Mao, X., Yang, L., Su, X. L., and Li, Y. (2006) A nanoparticle amplification based quartz crystal microbalance DNA sensor for detection of *Escherichia coli* O157:H7. *Biosens.Bioelectron.* 21[7], 1178-1185.
- Mao, Z., Zheng, H., Wang, X., Lin, S., Sun, Y., and Jiang, B. (2008) DNA microarray for direct identification of bacterial pathogens in human stool samples. *Digestion* 78[2-3], 131-138.
- Marteau, P., Pochart, P., Dore, J., Bera-Maillet, C., Bernalier, A., and Corthier, G. (2001) Comparative study of bacterial groups within the human cecal and fecal microbiota. *Applied and Environmental Microbiology* 67[10], 4939-4942.
- Maruo, T., Sakamoto, M., Toda, T., and Benno, Y. (2006) Monitoring the cell number of *Lactococcus lactis* subsp *cremoris* FC in human feces by real-time PCR with strain-specific primers designed using the RAPD technique. *International Journal of Food Microbiology* 110[1], 69-76.
- Maskos, U. and Southern, E. M. (1992) Oligonucleotide Hybridizations on Glass Supports - A Novel Linker for Oligonucleotide Synthesis and Hybridization Properties of Oligonucleotides Synthesized Insitu. *Nucleic Acids Research* 20[7], 1679-1684.
- Maskos, U. and Southern, E. M. (1993) A study of oligonucleotide reassociation using large arrays of oligonucleotides synthesised on a glass support. *Nucleic Acids Research* 21[20], 4663-9.
- Matsuda, K., Tsuji, H., Asahara, T., Matsumoto, K., Takada, T., and Nomoto, K. (2009) Establishment of an analytical system for the human fecal microbiota, based on reverse transcription-quantitative PCR targeting of multicopy rRNA molecules. *Applied and Environmental Microbiology* 75[7], 1961-1969.
- Matsuki, T., Watanabe, K., Fujimoto, J., Takada, T., and Tanaka, R. (2004) Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human feces. *Applied and Environmental Microbiology* 70[12], 7220-7228.
- McFarland, L. V. (2008) Update on the changing epidemiology of *Clostridium difficile*-associated disease. *Nat.Clin.Pract.Gastroenterol.Hepatol.* 5[1], 40-48.
- Michel, W., Mai, T., Naiser, T., and Ott, A. (2007) Optical study of DNA surface hybridization reveals DNA surface density as a key parameter for microarray hybridization kinetics. *Biophysical Journal* 92[3], 999-1004.
- Michelland, R. J., Monteils, V., Zened, A., Combes, S., Cauquil, L., Gidenne, T., Hamelin, J., and Fortun-Lamothe, L. (2009) Spatial and temporal variations of the bacterial community in the bovine digestive tract. *Journal of Applied Microbiology* 107[5], 1642-1650.
- Mikhailovich, V., Gryadunov, D., Kolchinsky, A., Makarov, A. A., and Zasedatelev, A. (2008) DNA microarrays in the clinic: infectious diseases. *Bioessays* 30[7], 673-682.
- Mitterer, G., Huber, M., Leidinger, E., Kirisits, C., Lubitz, W., Mueller, M. W., and Schmidt, W. M. (2004) Microarray-based identification of bacteria in clinical samples by solid-phase PCR amplification of 23S ribosomal DNA sequences. *Journal of Clinical Microbiology* 42[3], 1048-1057.
- Mohan, R., Namsolleck, P., Lawson, P. A., Osterhoff, M., Collins, M. D., Alpert, C. A., and Blaut, M. (2006) *Clostridium asparagiforme* sp. nov., isolated from a human faecal sample. *Syst.Appl.Microbiol* 29[4], 292-299.

- Moore, W. E. and Holdeman, L. V. (1974) Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl.Microbiol.* 27[5], 961-979.
- Moreau, M. C. and Gaboriau-Routhiau, V. r. (2001) Influence of Resident Intestinal Microflora on the Development and Functions of the Gut-Associated Lymphoid Tissue. *Microbial Ecology in Health and Disease* 13[2], 65-86.
- Moreira, J. L., Mota, R. M., Horta, M. F., Teixeira, S. M., Neumann, E., Nicoli, J. R., and Nunes, A. C. (2005) Identification to the species level of *Lactobacillus* isolated in probiotic prospecting studies of human, animal or food origin by 16S-23S rRNA restriction profiling. *BMC Microbiol* 5[1], 15.
- Mshana, S. E., Imirzalioglu, C., Hossain, H., Hain, T., Domann, E., and Chakraborty, T. (2009) Conjugative IncFI plasmids carrying CTX-M-15 among *Escherichia coli* ESBL producing isolates at a University hospital in Germany. *BMC Infect.Dis.* 9, ArtNb-97.
- Much, P., Pichler, J., Kasper, S. S., and Allerberger, F. (2009) Foodborne outbreaks, Austria 2007. *Wien.Klin.Wochenschr.* 121[3-4], 77-85.
- Muresan, L., Jacak, J., Klement, E. P., Hesse, J., and Schutz, G. J. (2010) Microarray Analysis at Single-Molecule Resolution. *Ieee Transactions on Nanobioscience* 9[1], 51-58.
- Muyzer, G., Dewaal, E. C., and Uitterlinden, A. G. (1993) Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16S Ribosomal-Rna. *Applied and Environmental Microbiology* 59[3], 695-700.
- Myers, K. M., Gaba, J., and Al Khalidi, S. F. (2006) Molecular identification of *Yersinia enterocolitica* isolated from pasteurized whole milk using DNA microarray chip hybridization. *Mol Cell Probes* 20[2], 71-80.
- Myers, L. L., Shoop, D. S., Stackhouse, L. L., Newman, F. S., Flaherty, R. J., Letson, G. W., and Sack, R. B. (1987) Isolation of Enterotoxigenic *Bacteroides-Fragilis* from Humans with Diarrhea. *Journal of Clinical Microbiology* 25[12], 2330-2333.
- NARMS (2009) National antimicrobial resistance monitoring system for enteric bacteria (NARMS): Executive Report 2006. Report
- Ng, C. T., Gilchrist, C. A., Lane, A., Roy, S., Haque, R., and Houghton, E. R. (2005) Multiplex real-time PCR assay using Scorpion probes and DNA capture for genotype-specific detection of *Giardia lamblia* on fecal samples. *J Clin.Microbiol.* 43[3], 1256-1260.
- Nuding, S., Fellermann, K., Wehkamp, J., and Stange, E. F. (2007) Reduced mucosal antimicrobial activity in Crohn's disease of the colon. *Gut* .
- Oldenburg, S. J., Genick, C. C., Clark, K. A., and Schultz, D. A. (2002) Base pair mismatch recognition using plasmon resonant particle labels. *Analytical Biochemistry* 309[1], 109-116.
- Omiccioli, E., Amagliani, G., Brandi, G., and Magnani, M. (2009) A new platform for Real-Time PCR detection of *Salmonella* spp., *Listeria monocytogenes* and *Escherichia coli* O157 in milk. *Food Microbiology* 26[6], 615-622.
- Ott, S. J., Musfeldt, M., Timmis, K. N., Hampe, J., Wenderoth, D. F., and Schreiber, S. (2004a) In vitro alterations of intestinal bacterial microbiota in fecal samples during storage. *Diagnostic Microbiology and Infectious Disease* 50[4], 237-245.
- Ott, S. J., Musfeldt, M., Ullmann, U., Hampe, J., and Schreiber, S. (2004b) Quantification of intestinal bacterial populations by real-time PCR with a universal primer set and minor groove binder probes: a global approach to the enteric flora. *Journal of Clinical Microbiology* 42[6], 2566-2572.
- Ozutsumi, Y., Hayashi, H., Sakamoto, M., Itabashi, H., and Benno, Y. (2005) Culture-independent analysis of fecal microbiota in cattle. *Biosci.Biotechnol.Biochem.* 69[9], 1793-1797.
- Palacios, G., Quan, P. L., Jabado, O. J., Conlan, S., Hirschberg, D. L., Liu, Y., Zhai, J., Renwick, N., Hui, J., Hegyi, H., Grolla, A., Strong, J. E., Towner, J. S., Geisbert, T. W., Jahrling, P. B., Buchen-Osmond, C., Ellerbrok, H., Sanchez-Seco, M. P., Lussier, Y., Formenty, P., Nichol, M. S., Feldmann, H., Briese, T., and Lipkin, W. I. (2007) Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg.Infect.Dis.* 13[1], 73-81.
- Paliy, O., Kenche, H., Abernathy, F., and Michail, S. (2009) High-throughput quantitative analysis of the human intestinal microbiota with a phylogenetic microarray. *Applied and Environmental Microbiology* 75[11], 3572-3579.
- Palmer, C., Bik, E. M., Digiulio, D. B., Relman, D. A., and Brown, P. O. (2007) Development of the Human Infant Intestinal Microbiota. *PLoS.Biol.* 5[7], e177.
- Palmer, C., Bik, E. M., Eisen, M. B., Eckburg, P. B., Sana, T. R., Wolber, P. K., Relman, D. A., and Brown, P. O. (2006) Rapid quantitative profiling of complex microbial populations. *Nucleic Acids Res.* 34[1], e5.
- Pang, X., Hua, X., Yang, Q., Ding, D., Che, C., Cui, L., Jia, W., Bucheli, P., and Zhao, L. (2007) Inter-species transplantation of gut microbiota from human to pigs. *ISME.J* 1[2], 156-162.
- Pang, X., Ding, D., Wei, G., Zhang, M., Wang, L., and Zhao, L. (2005) Molecular profiling of *Bacteroides* spp. in human feces by PCR-temperature gradient gel electrophoresis. *Journal of Microbiological Methods* 61[3], 413-417.
- Park, H., Jang, H., Song, E., Chang, C. L., Lee, M., Jeong, S., Park, J., Kang, B., and Kim, C. (2005) Detection and genotyping of *Mycobacterium* species from clinical isolates and specimens by oligonucleotide array. *Journal of Clinical Microbiology* 43[4], 1782-1788.
- Park, R. (2008) Strong growth projected for POC products. *IVD Technology* (January).
- Park, R. (2009) Continually moving forward. *IVD Technology* (November).
- Park, S. J., Taton, T. A., and Mirkin, C. A. (2002) Array-based electrical detection of DNA with nanoparticle probes. *Science* 295[5559], 1503-1506.

- Patchett, R. A., Kelly, A. F., and Kroll, R. G. (1991) The Adsorption of Bacteria to Immobilized Lectins. *Journal of Applied Bacteriology* 71[3], 277-284.
- Payne, M. J., Campbell, S., Patchett, R. A., and Kroll, R. G. (1992) The Use of Immobilized Lectins in the Separation of *Staphylococcus-Aureus*, *Escherichia-Coli*, *Listeria* and *Salmonella* Spp from Pure Cultures and Foods. *Journal of Applied Bacteriology* 73[1], 41-52.
- Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series* 6 2[11], 559-572.
- Pemov, A., Modi, H., Chandler, D. P., and Bavykin, S. (2005) DNA analysis with multiplex microarray-enhanced PCR. *Nucleic Acids Research* 33[2], e11.
- Penders, J., Vink, C., Driessen, C., London, N., Thijs, C., and Stobberingh, E. E. (2005) Quantification of *Bifidobacterium* spp., *Escherichia coli* and *Clostridium difficile* in faecal samples of breast-fed and formula-fed infants by real-time PCR. *Fems Microbiology Letters* 243[1], 141-147.
- Peplies, J., Glockner, F. O., and Amann, R. (2003) Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes. *Applied and Environmental Microbiology* 69[3], 1397-1407.
- Peterson, A. W., Wolf, L. K., and Georgiadis, R. M. (2002) Hybridization of mismatched or partially matched DNA at surfaces. *Journal of the American Chemical Society* 124[49], 14601-14607.
- Petrik, J. (2006) Diagnostic applications of microarrays. *Transfusion Medicine* 16[4], 233-247.
- Petrosino, J. F., Highlander, S., Luna, R. A., Gibbs, R. A., and Versalovic, J. (2009) Metagenomic pyrosequencing and microbial identification. *Clin.Chem.* 55[5], 856-866.
- Peytavi, R., Liu-Ying, T., Raymond, F. R., Boissinot, K., Bissonnette, L., Boissinot, M., Picard, F. J., Huletsky, A., Ouellette, M., and Bergeron, M. G. (2005) Correlation between microarray DNA hybridization efficiency and the position of short capture probe on the target nucleic acid. *Biotechniques* 39[1], 89-96.
- Piliarik, M., Parova, L., and Homola, J. (2009) High-throughput SPR sensor for food safety. *Biosens.Bioelectron.* 24[5], 1399-1404.
- Polz, M. F. and Cavanaugh, C. M. (1998) Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology* 64[10], 3724-3730.
- Pozhitkov, A. E., Bailey, K. D., and Noble, P. A. (2007) Development of a statistically robust quantification method for microorganisms in mixtures using oligonucleotide microarrays. *J Microbiol Methods* 70[2], 292-300.
- Rajilic-Stojanovic, M., Heilig, H. G., Molenaar, D., Kajander, K., Surakka, A., Smidt, H., and de Vos, W. M. (2009) Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ.Microbiol.* 11[7], 1736-1751.
- Ramachandran, S., Weigl, B. H., Gerdes, J., Tarr, P., Yager, P., Dillman, L., Peck, R., Kokoris, M., Nabavi, M., Battrell, F., and Hoeckstra, D. (2006) Dry-reagent storage for disposable lab-on-card diagnosis of enteric pathogens.
- Ramalingam, N., Rui, Z., Liu, H. B., Dai, C. C., Kaushik, R., Ratnahratika, B., and Gong, H. Q. (2010) Real-time PCR-based microfluidic array chip for simultaneous detection of multiple waterborne pathogens. *Sensors and Actuators B-Chemical* 145[1], 543-552.
- Ratushna, V. G., Weller, J. W., and Gibas, C. J. (2005) Secondary structure in the target as a confounding factor in synthetic oligomer microarray design. *BMC Genomics* 6.
- Reimers, M. and Weinstein, J. N. (2005) Quality assessment of microarrays: visualization of spatial artifacts and quantitation of regional biases. *BMC Bioinformatics* 6, 166.
- Religio, A., Schwager, C., Richter, A., Ansorge, W., and Valcarcel, J. (2002) Optimization of oligonucleotide-based DNA microarrays. *Nucleic Acids Research* 30[11], e-51.
- Requena, T., Burton, J., Matsuki, T., Munro, K., Simon, M. A., Tanaka, R., Watanabe, K., and Tannock, G. W. (2002) Identification, Detection, and Enumeration of Human *Bifidobacterium* Species by PCR Targeting the Transaldolase Gene. *Applied and Environmental Microbiology* 68[5], 2420-2427.
- Rigottier-Gois, L., Le Bourhis, A. G., Gramet, G., Rochet, V., and Dore, J. (2003a) Fluorescent hybridisation combined with flow cytometry and hybridisation of total RNA to analyse the composition of microbial communities in human faeces using 16S rRNA probes. *Fems Microbiology Ecology* 43[2], 237-245.
- Rigottier-Gois, L., Rochet, V., Garrec, N., Suau, A., and Dore, J. (2003b) Enumeration of *Bacteroides* species in human faeces by fluorescent in situ hybridisation combined with flow cytometry using 16S rRNA probes. *Systematic and Applied Microbiology* 26[1], 110-118.
- Ringel, A. F., Jameson, G. L., and Foster, E. S. (1995) Diarrhea in the intensive care patient. *Crit Care Clin.* 11[2], 465-477.
- Robert-Koch-Institut (2007a) RKI: Jahresstatistik 2005/2006. Report
- Robert-Koch-Institut. (2007b) *Salmonella* Enteritidis – aktuelle Bedeutung. *Epidemiologisches Bulletin* 3, 17-19.
- Robert-Koch-Institut (2009) Gruppenerkrankung in einem Alten- und Pflegeheim durch *Salmonella* Enteritidis (Lysotyp 3a/17). Report
- Roche Applied Science. (2010) LightCycler Kits. https://www.roche-applied-science.com/sis/rtqcr/lightcycler/index.jsp?id=ict_030000 .
- Rodriguez-Lazaro, D., Lloyd, J., Herrewegh, A., Ikonopoulou, J., D'Agostino, M., Pla, M., and Cook, N. (2004) A molecular beacon-based real-time NASBA assay for detection of *Mycobacterium avium* subsp *paratuberculosis* in water and milk. *Fems Microbiology Letters* 237[1], 119-126.
- Ronaghi, M. and Elahi, E. (2002) Pyrosequencing for microbial typing. *J.Chromatogr.B Analyt.Technol.Biomed.Life Sci.* 782[1-2], 67-72.

- Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M., and Nyren, P. (1996) Real-time DNA sequencing using detection of pyrophosphate release. *Anal.Biochem.* 242[1], 84-89.
- Ronaghi, M., Uhlen, M., and Nyren, P. (1998) A sequencing method based on real-time pyrophosphate. *Science* 281[5375], 363, 365.
- Round, J. L., O'Connell, R. M., and Mazmanian, S. K. (2010) Coordination of tolerogenic immune responses by the commensal microbiota. *Journal of Autoimmunity* 34[3], J220-J225.
- Rudi, K., Maugesten, T., Hannevik, S. E., and Nissen, H. (2004) Explorative multivariate analyses of 16S rRNA gene data from microbial communities in modified-atmosphere-packed salmon and coalfish. *Applied and Environmental Microbiology* 70[8], 5010-5018.
- Sack, R. B., Albert, M. J., Alam, K., Neogi, P. K., and Akbar, M. S. (1994) Isolation of enterotoxigenic *Bacteroides fragilis* from Bangladeshi children with diarrhea: a controlled study. *Journal of Clinical Microbiology* 32[4], 960-963.
- Sails, A. D., Fox, A. J., Bolton, F. J., Wareing, D. R., and Greenway, D. L. (2003) A real-time PCR assay for the detection of *Campylobacter jejuni* in foods after enrichment culture. *Appl.Environ.Microbiol* 69[3], 1383-1390.
- Sails, A. D., Fox, A. J., Bolton, F. J., Wareing, D. R., Greenway, D. L., and Borrow, R. (2001) Development of a PCR ELISA assay for the identification of *Campylobacter jejuni* and *Campylobacter coli*. *Mol.Cell Probes* 15[5], 291-300.
- Satokari, R. M., Kataja, K., and Soderlund, H. (2005) Multiplexed quantification of bacterial 16S rRNA by solution hybridization with oligonucleotide probes and affinity capture. *Microbial Ecology* 50[1], 120-127.
- Satokari, R. M., Vaughan, E. E., Akkermans, A. D. L., Saarela, M., and de Vos, W. M. (2001) Bifidobacterial diversity in human feces detected by genus-specific PCR and denaturing gradient gel electrophoresis. *Applied and Environmental Microbiology* 67[2], 504-513.
- Savage, D. C. (1977) *Microbial Ecology of Gastrointestinal-Tract*. *Annual Review of Microbiology* 31, 107-133.
- Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270[5235], 467-470.
- Schramm, A., Fuchs, B. M., Nielsen, J. L., Tonolla, M., and Stahl, D. A. (2002) Fluorescence in situ hybridization of 16S rRNA gene clones (Clone-FISH) for probe validation and screening of clone libraries. *Environmental Microbiology* 4[11], 713-720.
- Schuler, T., Kretschmer, R., Jessing, S., Urban, M., Fritzsche, W., Moller, R., and Popp, J. (2009) A disposable and cost efficient microfluidic device for the rapid chip-based electrical detection of DNA. *Biosens.Bioelectron.* 25[1], 15-21.
- Schulze, H., Giraud, G., Crain, J., and Bachmann, T. T. (2009) Multiplexed optical pathogen detection with lab-on-a-chip devices. *J.Biophotonics.* 2[4], 199-211.
- Schumann, A., Nutten, S., Donnicola, D., Comelli, E. M., Mansourian, R., Cherbut, C., Corthesy-Theulaz, I., and Garcia-Rodenas, C. (2005) Neonatal antibiotic treatment alters gastrointestinal tract developmental gene expression and intestinal barrier transcriptome. *Physiol Genomics* 23[2], 235-245.
- Schwartz, A., Le Blay, G., and Blaut, M. (2000) Quantification of Different *Eubacterium* spp. in Human Fecal Samples with Species-Specific 16S rRNA-Targeted Oligonucleotide Probes. *Applied and Environmental Microbiology* 66[1], 375-382.
- Scott, L., McGee, P., Minihan, D., Sheridan, J. J., Earley, B., and Leonard, N. (2006) The characterisation of *E. coli* O157:H7 isolates from cattle faeces and feedlot environment using PFGE. *Vet.Microbiol* 114[3-4], 331-336.
- Senok, A., Yousif, A., Mazi, W., Sharaf, E., Bindayna, K., Elnima, E. A., and Botta, G. (2007) Pattern of antibiotic susceptibility in *Campylobacter jejuni* isolates of human and poultry origin. *Japanese Journal of Infectious Diseases* 60[1], 1-4.
- Septhri, S., Kotlowski, R., Bernstein, C. N., and Krause, D. O. (2007) Microbial diversity of inflamed and noninflamed gut biopsy tissues in inflammatory bowel disease. *Inflamm.Bowel Dis.*
- Sghir, A., Gramet, G., Suau, A., Rochet, V., Pochart, P., and Dore, J. (2000) Quantification of Bacterial Groups within Human Fecal Flora by Oligonucleotide Probe Hybridization. *Applied and Environmental Microbiology* 66[5], 2263-2266.
- Shah, N., DuPont, H. L., and Ramsey, D. J. (2009) Global Etiology of Travelers' Diarrhea: Systematic Review from 1973 to the Present. *American journal of tropical medicine and hygiene* 80[4], 609-614.
- Shakeel, S., Karim, S., and Ali, A. (2006) Peptide nucleic acid (PNA) - a review. *Journal of Chemical Technology and Biotechnology* 81[6], 892-899.
- Shchepinov, M. S., Case-Green, S. C., and Southern, E. M. (1997) Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. *Nucleic Acids Research* 25[6], 1155-61.
- Shen, J., Zhang, B., Wei, G., Pang, X., Wei, H., Li, M., Zhang, Y., Jia, W., and Zhao, L. (2006) Molecular profiling of the *Clostridium leptum* subgroup in human fecal microflora by PCR-denaturing gradient gel electrophoresis and clone library analysis. *Appl.Environ.Microbiol* 72[8], 5232-5238.
- Shi, L., Perkins, R. G., Fang, H., and Tong, W. (2008) Reproducible and reliable microarray results through quality control: good laboratory proficiency and appropriate data analysis practices are essential. *Curr.Opin.Biotechnol.* 19[1], 10-18.
- Shima, T., Fukushima, K., Setoyama, H., Imaoka, A., Matsumoto, S., Hara, T., Suda, K., and Umesaki, Y. (2008) Differential effects of two probiotic strains with different bacteriological properties on intestinal gene expression, with special reference to indigenous bacteria. *FEMS Immunol.Med Microbiol* 52[1], 69-77.
- Sista, R., Hua, Z., Thwar, P., Sudarsan, A., Srinivasan, V., Eckhardt, A., Pollack, M., and Pamula, V. (2008) Development of a digital microfluidic platform for point of care testing. *Lab Chip* 8[12], 2091-2104.

- Skanseng, B., Kaldhusdal, M., and Rudi, K. (2006) Comparison of chicken gut colonisation by the pathogens *Campylobacter jejuni* and *Clostridium perfringens* by real-time quantitative PCR. *Molecular and Cellular Probes* 20[5], 269-279.
- Sosnowski, R., Heller, M. J., To, E., Forster, A. H., and Radtkey, R. (2002) Active microelectronic array system for DNA hybridization, genotyping and pharmacogenomic applications. *Psychiatric Genetics* 12[4], 181-192.
- Southern, E., Mir, K., and Shchepinov, M. (1999) Molecular interactions on microarrays. *Nature Genetics* 21[1 Suppl], 5-9.
- Stecher, B., Chaffron, S., Kappeli, R., Hapfelmeier, S., Friedrich, S., Weber, T. C., Kirundi, J., Suar, M., McCoy, K. D., von Mering, C., Macpherson, A. J., and Hardt, W. D. (2010) Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. *PLoS Pathog.* 6[1], e1000711.
- Straub, T. M. and Chandler, D. P. (2003) Towards a unified system for detecting waterborne pathogens. *J.Microbiol.Methods* 53[2], 185-197.
- Striebel, H. M., Birch-Hirschfeld, E., Egerer, R., Foldes-Papp, Z., Titz, G. P., and Stelzner, A. (2004) Enhancing sensitivity of human herpes virus diagnosis with DNA microarrays using dendrimers. *Experimental and Molecular Pathology* 77[2], 89-97.
- Suau, A., Rochet, V., Sghir, A., Gramet, G., Brewaeys, S., Sutren, M., Rigottier-Gois, L., and Dore, J. (2001) *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. *Systematic and Applied Microbiology* 24[1], 139-145.
- Suau, A., Bonnet, R., Sutren, M., Godon, J. J., Gibson, G. R., Collins, M. D., and Dore, J. (1999) Direct Analysis of Genes Encoding 16S rRNA from Complex Communities Reveals Many Novel Molecular Species within the Human Gut. *Applied and Environmental Microbiology* 65[11], 4799-4807.
- Suzuki, M. T. and Giovannoni, S. J. (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and Environmental Microbiology* 62[2], 625-630.
- Tanaka, S., Kobayashi, T., Songjinda, P., Tateyama, A., Tsubouchi, M., Kiyohara, C., Shirakawa, T., Sonomoto, K., and Nakayama, J. (2009) Influence of antibiotic exposure in the early postnatal period on the development of intestinal microbiota. *FEMS Immunol.Med.Microbiol.* 56[1], 80-87.
- Taton, T. A., Mirkin, C. A., and Letsinger, R. L. (2000) Scanometric DNA array detection with nanoparticle probes. *Science* 289[5485], 1757-1760.
- Thiel, R. and Blaut, M. (2005) An improved method for the automated enumeration of fluorescently labelled bacteria in human faeces. *J.Microbiol.Methods* 61[3], 369-379.
- Threlfall, E. J., Day, M., de Pinna, E., Lewis, H., and Lawrence, J. (2006) Drug-resistant enteric fever in the UK. *Lancet* 367[9522], 1576.
- Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R., and Gordon, J. I. (2007) The human microbiome project. *Nature* 449[7164], 804-810.
- Tynkkynen, S., Satokari, R., Saarela, M., Mattila-Sandholm, T., and Saxelin, M. (1999) Comparison of ribotyping, randomly amplified polymorphic DNA analysis, and pulsed-field gel electrophoresis in typing of *Lactobacillus rhamnosus* and *L. casei* strains. *Appl.Environ.Microbiol* 65[9], 3908-3914.
- Uyttendaele, M., Van, H., I, and Debevere, J. (2000) The use of immuno-magnetic separation (IMS) as a tool in a sample preparation method for direct detection of *L. monocytogenes* in cheese. *Int.J.Food Microbiol.* 54[3], 205-212.
- Vainrub, A. and Pettitt, B. M. (2002) Coulomb blockage of hybridization in two-dimensional DNA arrays. *Physical Review e* 66[4].
- Vali, L., Hamouda, A., Pearce, M. C., Knight, H. I., Evans, J., and Amyes, S. G. (2007) Detection of genetic diversity by pulsed-field gel electrophoresis among *Escherichia coli* O157 isolated from bovine faecal samples by immunomagnetic separation technique. *Lett.Appl.Microbiol* 44[1], 19-23.
- van Doorn, R., Klerks, M. M., Gent-Pelzer, M. P., Speksnijder, A. G., Kowalchuk, G. A., and Schoen, C. D. (2009) Accurate quantification of microorganisms in PCR-inhibiting environmental DNA extracts by a novel internal amplification control approach using Biotrove OpenArrays. *Applied and Environmental Microbiology* 75[22], 7253-7260.
- Vanhoutte, T., De Preter, V., De Brandt, E., Verbeke, K., Swings, J., and Huys, G. (2006) Molecular monitoring of the fecal microbiota of healthy human subjects during administration of lactulose and *Saccharomyces boulardii*. *Applied and Environmental Microbiology* 72[9], 5990-5997.
- Ventura, M., Casas, I. A., Morelli, L., and Callegari, M. L. (2000) Rapid amplified ribosomal DNA restriction analysis (ARDRA) identification of *Lactobacillus* spp. isolated from fecal and vaginal samples. *Syst Appl.Microbiol* 23[4], 504-509.
- Ventura, M., Elli, M., Reniero, R., and Zink, R. (2001) Molecular microbial analysis of *Bifidobacterium* isolates from different environments by the species-specific amplified ribosomal DNA restriction analysis (ARDRA). *FEMS Microbiol Ecol.* 36[2-3], 113-121.
- Verweij, J. J., Blange, R. A., Templeton, K., Schinkel, J., Brienen, E. A. T., Van Rooyen, M. A. A., Van Lieshout, L., and Polderman, A. M. (2004) Simultaneous detection of *Entamoeba histolytica*, *Giardia lamblia*, and *Cryptosporidium parvum* in fecal samples by using multiplex real-time PCR. *Journal of Clinical Microbiology* 42[3], 1220-1223.
- Vianna, M. E., Horz, H. P., Gomes, B. P., and Conrads, G. (2005) Microarrays complement culture methods for identification of bacteria in endodontic infections. *Oral Microbiol.Immunol.* 20[4], 253-258.
- Villari, P., Motti, E., Farullo, C., and Torre, I. (1998) Comparison of conventional culture and PCR methods for the detection of *Legionella pneumophila* in water. *Lett.Appl.Microbiol.* 27[2], 106-110.

- Vincent, D., Roy, D., Mondou, F., and Dery, C. (1998) Characterization of bifidobacteria by random DNA amplification. *Int J Food Microbiol* 43[3], 185-193.
- Voetsch, A. C., Van Gilder, T. J., Angulo, F. J., Farley, M. M., Shallow, S., Marcus, R., Cieslak, P. R., Deneen, V. C., and Tauxe, R. V. (2004) FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. *Clinical Infectious Diseases* 38, S127-S134.
- Voitoux, E., Lafarge, V., Collette, C., and Lombard, B. (2002) Applicability of the draft standard method for the detection of *Escherichia coli* O157 in dairy products. *Int.J.Food Microbiol.* 77[3], 213-221.
- Volokhov, D., Chizhikov, V., Chumakov, K., and Rasooly, A. (2003) Microarray-based identification of thermophilic *Campylobacter jejuni*, *C.-coli*, *C.-lari*, and *C.-upsaliensis*. *Journal of Clinical Microbiology* 41[9], 4071-4080.
- Volokhov, D., Rasooly, A., Chumakov, K., and Chizhikov, V. (2002) Identification of *Listeria* species by microarray-based assay. *Journal of Clinical Microbiology* 40[12], 4720-4728.
- Vora, G. J., Meador, C. E., Anderson, G. P., and Taitt, C. R. (2008) Comparison of detection and signal amplification methods for DNA microarrays. *Mol.Cell Probes* 22[5-6], 294-300.
- Vora, G. J., Meador, C. E., Stenger, D. A., and Andreadis, J. D. (2004) Nucleic Acid Amplification Strategies for DNA Microarray-Based Pathogen Detection. *Applied and Environmental Microbiology* 70[5], 3047-3054.
- Wache, Y. J., Valat, C., Postollec, G., Bougeard, S., Burel, C., Oswald, I. P., and Fravallo, P. (2009) Impact of Deoxynivalenol on the Intestinal Microflora of Pigs. *International Journal of Molecular Sciences* 10[1], 1-17.
- Wallace, R. J., Glassroth, J., Griffith, D. E., Olivier, K. N., Cook, J. L., and Gordin, F. (1997) Diagnosis and treatment of disease caused by nontuberculous mycobacteria. *American Journal of Respiratory and Critical Care Medicine* 156[2], S1-S25.
- Walter, J., Hertel, C., Tannock, G. W., Lis, C. M., Munro, K., and Hammes, W. P. (2001) Detection of *Lactobacillus*, *Pediococcus*, *Leuconostoc*, and *Weissella* Species in Human Feces by Using Group-Specific PCR Primers and Denaturing Gradient Gel Electrophoresis. *Applied and Environmental Microbiology* 67[6], 2578-2585.
- Wang, R. F., Cao, W. W., and Cerniglia, C. E. (1996) Phylogenetic analysis of *Fusobacterium prausnitzii* based upon the 16S rRNA gene sequence and PCR confirmation. *Int.J.Syst.Bacteriol.* 46[1], 341-343.
- Wang, R. F., Beggs, M. L., Erickson, B. D., and Cerniglia, C. E. (2004a) DNA microarray analysis of predominant human intestinal bacteria in fecal samples. *Molecular and Cellular Probes* 18[4], 223-234.
- Wang, Y., Holmes, E., Nicholson, J. K., Cloarec, O., Chollet, J., Tanner, M., Singer, B. H., and Utzinger, J. (2004b) Metabonomic investigations in mice infected with *Schistosoma mansoni*: an approach for biomarker identification. *Proc.Natl.Acad.Sci.U.S A* 101[34], 12676-12681.
- Wang, Z., Vora, G. J., and Stenger, D. A. (2004c) Detection and genotyping of *Entamoeba histolytica*, *Entamoeba dispar*, *Giardia lamblia*, and *Cryptosporidium parvum* by oligonucleotide microarray. *Journal of Clinical Microbiology* 42[7], 3262-3271.
- Wang, Z., Orlandi, P. A., and Stenger, D. A. (2005) Simultaneous detection of four human pathogenic microsporidian species from clinical samples by oligonucleotide microarray. *Journal of Clinical Microbiology* 43[8], 4121-4128.
- Waterman, S. R. and Small, P. L. (1998) Acid-sensitive enteric pathogens are protected from killing under extremely acidic conditions of pH 2.5 when they are inoculated onto certain solid food sources. *Applied and Environmental Microbiology* 64[10], 3882-3886.
- Weigl, B., Domingo, G., LaBarre, P., and Gerlach, J. (2008) Towards non- and minimally instrumented, microfluidics-based diagnostic devices. *Lab on A Chip* 8[12], 1999-2014.
- Weigl, B. H., Gerdes, J., Tarr, P., Yager, P., Dillman, L., Peck, R., Ramachandran, S., Lemba, S., Kokoris, M., Nabavi, M., Battrell, F., Hoekstra, D., Klein, E. J., and Denno, D. M. (2006) Fully integrated multiplexed lab-on-a-card assay for enteric pathogens - art. no. 611202. 6112. *Microfluidics, BioMEMS, and Medical Microsystems IV, Proceedings of the SPIE.*
- West, J. L. and Halas, N. J. (2003) Engineered nanomaterials for biophotonics applications: improving sensing, imaging, and therapeutics. *Annu.Rev.Biomed.Eng* 5, 285-292.
- Westad, F. and Martens, H. (2000) Variable selection in near infrared spectroscopy based on significance testing in partial least squares regression. *Journal of Near Infrared Spectroscopy* 8[2], 117-124.
- WHO (2009) World Health Statistics 2009. Report
- Wilkes, T., Laux, H., and Foy, C. A. (2007) Microarray data quality - Review of current developments. *Omics-A Journal of Integrative Biology* 11[1], 1-13.
- Wilson, K. H. and Blichington, R. B. (1996) Human colonic biota studied by ribosomal DNA sequence analysis. *Applied and Environmental Microbiology* 62[7], 2273-2278.
- Wilson, W. J., Strout, C. L., Desantis, T. Z., Stilwell, J. L., Carrano, A. V., and Andersen, G. L. (2002) Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. *Molecular and Cellular Probes* 16[2], 119-127.
- Wittwer, C. T., Herrmann, M. G., Moss, A. A., and Rasmussen, R. P. (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. *Biotechniques* 22[1], 130-138.
- Wohlgemuth, S., Loh, G., and Blaut, M. (2010) Recent developments and perspectives in the investigation of probiotic effects. *International Journal of Medical Microbiology* 300[1], 3-10.
- Wold, S. (1978) Cross-Validatory Estimation of Number of Components in Factor and Principal Components Models. *Technometrics* 20[4], 397-405.
- Wold, S., Sjostrom, M., and Eriksson, L. (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* 58, 109-130.
- Woodmansey, E. J. (2007) Intestinal bacteria and ageing. *J.Appl.Microbiol.* 102[5], 1178-1186.

- Wu, X. Y., Walker, M. J., Hornitzky, M., and Chin, J. (2006) Development of a group-specific PCR combined with ARDRA for the identification of *Bacillus* species of environmental significance. *J Microbiol Methods* 64[1], 107-119.
- Xu, F., Datta, P., Wang, H., Gurung, S., Hashimoto, M., Wei, S., Goettert, J., McCarley, R. L., and Soper, S. A. (2007) Polymer microfluidic chips with integrated waveguides for reading microarrays. *Anal.Chem.* 79[23], 9007-9013.
- Yao, G. and Tan, W. H. (2004) Molecular-beacon-based array for sensitive DNA analysis. *Analytical Biochemistry* 331[2], 216-223.
- You, Y., Fu, C., Zeng, X., Fang, D., Yan, X., Sun, B., Xiao, D., and Zhang, J. (2008) A novel DNA microarray for rapid diagnosis of enteropathogenic bacteria in stool specimens of patients with diarrhea. *J.Microbiol.Methods* 75[3], 566-571.
- Yu, X., Susa, M., Weile, J., Knabbe, C., Schmid, R. D., and Bachmann, T. T. (2007) Rapid and sensitive detection of fluoroquinolone-resistant *Escherichia coli* from urine samples using a genotyping DNA microarray. *Int J Med Microbiol* 297[6], 417-429.
- Zeng, S. Q., Halkosalo, A., Salminen, M., Szakal, E. D., Puustinen, L., and Vesikari, T. (2008) One-step quantitative RT-PCR for the detection of rotavirus in acute gastroenteritis. *J.Virol.Methods* 153[2], 238-240.
- Zengler, K., Toledo, G., Rappe, M., Elkins, J., Mathur, E. J., Short, J. M., and Keller, M. (2002) Cultivating the uncultured. *Proceedings of the National Academy of Sciences of the United States of America* 99[24], 15681-15686.
- Zhang, H., Dibaise, J. K., Zuccolo, A., Kudrna, D., Braidotti, M., Yu, Y., Parameswaran, P., Crowell, M. D., Wing, R., Rittmann, B. E., and Krajmalnik-Brown, R. (2009a) Human gut microbiota in obesity and after gastric bypass. *Proc.Natl.Acad.Sci.U.S A* 106[7], 2365-2370.
- Zhang, M., Zhang, M., Zhang, C., Du, H., Wei, G., Pang, X., Zhou, H., Liu, B., and Zhao, L. (2009b) Pattern extraction of structural responses of gut microbiota to rotavirus infection via multivariate statistical analysis of clone library data. *FEMS Microbiol.Ecol.* 70[2], 21-29.
- Zhang, Y., Hammer, D. A., and Graves, D. J. (2005a) Competitive hybridization kinetics reveals unexpected behavior patterns. *Biophysical Journal* 89[5], 2950-2959.
- Zhang, Z. X. and Li, M. Q. (2005b) Electrostatic microcantilever array biosensor and its application in DNA detection. *Progress in Biochemistry and Biophysics* 32[4], 314-317.
- Zheng, H., Sun, Y., Lin, S., Mao, Z., and Jiang, B. (2008) *Yersinia enterocolitica* infection in diarrheal patients. *Eur.J.Clin.Microbiol.Infect.Dis.* 27[8], 741-752.
- Zoetendal, E. G., Vaughan, E. E., and de Vos, W. M. (2006) A microbial world within us. *Molecular Microbiology* 59[6], 1639-1650.
- Zoetendal, E. G., Akkermans, A. D., and de Vos, W. M. (1998) Temperature Gradient Gel Electrophoresis Analysis of 16S rRNA from Human Fecal Samples Reveals Stable and Host-Specific Communities of Active Bacteria. *Applied and Environmental Microbiology* 64[10], 3854-3859.

Acknowledgment

At first, I would like to thank Prof. Dr. Rolf D. Schmid for the confidence in entrusting me this interesting scientific topic and for the opportunity to earn a doctorate at the Institute for Technical Biochemistry at the University of Stuttgart. The excellent technical equipment at the institute, his scientific advice, and his motivational support gave the basis of this work. I am deeply grateful for giving me the opportunity for a research stay in Shanghai/ China, which has broaden my mind and gave new input to my scientific work.

I would especially like to thank my supervisor PD Dr. Till T. Bachmann, who's enthusiasm for microarrays and clinical biotechnology created a working environment full of valuable and inspiring scientific discussions. His supervision of my work, the scientific advice, and the support in publishing were invaluable.

A special thanks goes also to all the present and former lab members, with whom I had the opportunity to work. The excellent working atmosphere, valuable scientific discussions, and mutual assistance made my time at the institute enjoyable and fruitful. I would like to mention especially Susi, Dirk, Timo, Harald, Beate, Holger, Thomas, Luam, Victoria, Marco, Michaela, Johanna, Barbara S., Xiaolei, and Verena, who shared not only the lab and scientific experience with me but also enjoyable leisure experiences.

I would also like to express my special gratitude to Barbara Hörmann, who unimpeachably and conscientiously assisted me in the laboratory. Parts of her '*Studienarbeit*' on '*Grundlegende Studien zur Entwicklung eines Quantifizierungssystems auf DNA Microarrays*' were the basis of my work.

My sincere thanks to PD Dr. Eugen Domann and Prof. Dr. Trinad Chakraborty, as our project partners, who supported me with information on clinically relevant bacteria and reference material for evaluation of the chip. In this regard, I am also much obliged to Kirsten-Susann Bommersheim and Isabell Trur for their efforts in preparing the clinical samples. Thanks also for the possibility to observe the clinical routine as a guest visitor for two days.

Many thanks go to Prof. Liping Zhao, head of School of Life Sciences at Shanghai Jiao Tong University, for giving me the opportunity of a scientific research stay at his institute and the excellent working facilities. I also thank Dr. Xiaojun Zhang for the prearrangement of my stay in Shanghai and his untiring support during the whole five month. My special thanks to Jing Zhu, who directly worked together with me, assisted my experiments, and supported me in daily life in a foreign country. I also owe all colleagues in Shanghai a debt of gratitude for the kind reception and the inspiring and friendly working atmosphere.

I also thank Dr. Wolfgang Ludwig from the Institute of Microbiology at the Technical University of Munich for his support with ARB and the time he spent on answering my questions.

I am indebted to the German Federal Ministry of Research BMBF (Bundesministerium für Bildung und Forschung) for funding of this project under the PathoGenoMik program and the Federal State of Baden-Württemberg (Landesgraduiertenförderung) for supporting this work by a scholarship. I also acknowledge the Eppendorf AG for providing this work with laboratory equipment and scientific advice from the industrial point of view and for the support in Shanghai/China. My sincere thanks also to our GenoMik project partners for their inspiring discussions and advice during our meetings.

Furthermore, I would like to thank my parents, my parents-in-law, and my husband Andrej for their everlasting support, patience, and motivation. Frank, you did me a big favour in reading my paper. Andi, thanks for all. Ella, you partly grounded my perfectionism.

Abbreviations and international units

| | |
|----------|--|
| °C | Degree Celsius |
| µL | Microlitre |
| µM | Micromolar |
| ANOVA | Analysis of variance |
| APS | ammonium persulfate |
| BHQ | Black hole quencher |
| BLAST | Basic local alignment and search tool |
| bp | Base pairs |
| CDC | Centre for Disease Control |
| CE-IVD | Consumer safety label: <i>Conformité Européenne</i> „European conformity“ – In-vitro Diagnostics |
| CFU | Colony forming units |
| Comp | Complementary |
| Ct | Cycle threshold |
| CV | Conventionally raised |
| Cy3 | Cyanine 3 |
| dATP | Deoxyadenosine triphosphate |
| dCTP | Deoxycytidine triphosphate |
| dd | double distilled |
| DGGE | denaturing gel electrophoresis |
| dGTP | Deoxyguanosine triphosphate |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| DSMZ | Deutsche Stammsammlung von Mikroorganismen und Zellkulturen (German collection of microorganisms and cell cultures) |
| dTTP | Deoxythymidine triphosphate |
| dUTP | Deoxyuridine triphosphate |
| EDTA | Ethylendiamin-tetraacetat |
| EGTA | Ethylene glycol tetraacetic acid |
| EHEC | Enterohemorrhagic <i>Escherichia coli</i> |
| FAO | Food and Agriculture Organization |
| fg | Femtogramm |
| FISH | Fluorescence in situ hybridization |
| fm | Femtomole |
| FRET | Fluorescence resonance energy transfer |
| FU | Fluorescence units |
| GI tract | Gastro-intestinal tract |
| HD | Human donor |
| HFA | Human flora associated |
| HPLC | High performance liquid chromatography |
| i | inverse (antisense) |
| IBS | Irritable bowel syndrome |
| Inc. | Incorporation |
| ITB | Institute for Technical Biochemistry, University of Stuttgart |
| ITS | Internal transcribed spacer |
| kb | Kilobases |
| L | Ladder |
| LAB | Lactic acid bacteria |
| LOD | Limit of detection |
| LOOVC | Leave-one-out cross-validation |
| LSU | Large subunit |
| M | Molar |
| min | Minute(s) |

| | |
|------------|---|
| mL | Millilitre |
| mM | Millimolar |
| MM | Mismatch |
| mU | Milliunit(s) |
| NC | Negative control |
| n.d. | not determined |
| ng | Nanogramm |
| NHC | Negative hybridization control |
| nm | Nanometre |
| nt | Nucleotides |
| NT/F | Number of nucleotides/number of incorporated fluorescent dyes |
| OTU | Operational taxonomic unit |
| p/p-value | Measure of significance |
| PC | Process control |
| PCA | Principle component analysis |
| PCR | Polymerase chain reaction |
| PD | Pig donor |
| PFA | Pig flora associated |
| PHC | Positive hybridization control |
| PLSA | Partial least squares analysis |
| PM | Perfect match |
| PMT | Photo multiplier tube |
| POC | Point-of-care |
| RFLP | restriction fragment length polymorphism |
| RI | Relative intensity |
| RNA | Ribonucleic acid |
| rpm | Revolutions per minute |
| rRNA | Ribosomal ribonucleic acid |
| s | Second(s) |
| S | Svedberg |
| SD | Standard deviation |
| SDS | Sodium dodecyl sulfate |
| SNP | Single nucleotide polymorphism |
| SPC | Spotting control |
| spp. | Species plural |
| SSC | Saline-sodium citrate buffer |
| SSP | Saline-sodium phosphate buffer |
| SSPE | Saline-sodium phosphate EDTA |
| SSU | Small subunit |
| TAE | Tris-acetate-EDTA |
| <i>Taq</i> | <i>Thermophilus aquaticus</i> |
| TEMED | Tetramethylethylenediamine |
| Tox | Toxin |
| tRNA | Transfer ribonucleic acid |
| U | Unit(s) |
| V | Volt(s) |
| WHO | World Health Organization |
| wt/vol | Weight per volume |

Curriculum vitae

Personal Information:

Name: Kristina Hänel (geb. Knösche)
 Date and place of birth: 05.07.1978, Potsdam
 Nationality: German
 Family status: married, 1 daughter

Education:

1985-1991 Polytechnische Oberschule (POS 33) in Potsdam
 1991-1998 Helmholtz-Gymnasium Potsdam
 1998 A-level exams (general qualification for university entrance)

University and postgraduate education:

1998-2004 **Diploma study of Biochemistry** at the University of Potsdam
 Degree: Diplom-Biochemikerin, predicate: „magna cum laude”

07-08/2001 **Practical training course in Kiev/Ukraine** at Palladin-Institute for Biochemistry, „Investigation of cell interactions with fibrin and its fragments”

2004 **Diploma thesis** at University of Potsdam, Institute of Analytical Biochemistry, chair of Professor Dr. Frieder W. Scheller, „Characterisation of the binding behaviour of several choline esterases to the cocaine derivative benzoylecgonine-1,8-diamino-3,4-dioxaoctane”

04-08/2007 **Scientific research stay in Shanghai/China** at School of Life Sciences at Shanghai Jiao Tong University, chair of Prof. Dr. Liping Zhao

2004-2010 **Doctoral thesis** at the University of Stuttgart, Institute of Technical Biochemistry, chair of Professor Dr. Rolf D. Schmid, funded by a scholarship of the Federal State of Baden-Württemberg, “Design and application of a DNA microarray for the identification of intestinal pathogens during gastroenteritis and monitoring of the resident intestinal microbiota”

03/2008-09/2008 **Leadership of „Biosystems Technology Group“** at Institute of Technical Biochemistry, University of Stuttgart

Awards:

2004-2008 **Scholarship** (3 years) of the Federal State of Baden-Württemberg (Landesgraduierföderung)

Publications:

Knösche,K., Halamek,J., Makower,A., Fournier,D., Scheller,F.W., (2004) "Molecular recognition of cocaine by acetylcholinesterases for affinity purification and bio-sensing", Biosens. Bioelectron. 20(2), 153-160.

Halamek,J., Makower,A., Knösche,K., Skladal,P., Scheller,F.W., (2005) "Piezoelectric affinity sensors for cocaine and cholinesterase inhibitors", Talanta 65(2), 337-342.

Hänel,K., Hain,T., Chakraborty,T., Schmid,R.D., Domann,E., Bachmann,T.T., (2010) "Detection of intestinal bacterial pathogens and monitoring of the resident microbiota in patients with gastroenteritis by using a diagnostic microarray", Int. J. Medical Microbiology, in revision