

Aus dem Institut für Mikrobiologie und Tierseuchen  
des Fachbereichs Veterinärmedizin  
der Freien Universität Berlin

**Der GimA-Lokus von  
Extraintestinal-pathogenen *E. coli* (ExPEC):  
Reduktive Evolution in Abhängigkeit  
von Habitat und Pathovar**

**Inaugural-Dissertation**  
zur Erlangung des Grades eines  
Doktors der Veterinärmedizin  
an der  
Freien Universität Berlin

vorgelegt von

**Timo Homeier-Bachmann, geb. Homeier**

Tierarzt aus Hannover

**Berlin 2009**

**Journal-Nr.: 3352**

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin  
der Freien Universität Berlin

Dekan: Univ.-Prof. Dr. Leo Brunnberg  
Erster Gutachter: Univ.-Prof. Dr. Lothar H. Wieler  
Zweiter Gutachter: Univ.-Prof. Dr. Klaus Osterrieder  
Dritter Gutachter: Univ.-Prof. Dr. Uwe H. Rösler

*Deskriptoren (nach CAB-Thesaurus):*

*Escherichia coli*, molecular genetics, polymerase chain reaction, DNA hybridization, selection, phylogeny, blood, urine, habitats, phylogeny, evolution, meningitis blood-brain barrier

Tag der Promotion: 08.12.2009

Bibliografische Information der *Deutschen Nationalbibliothek*

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN: 978-3-86664-728-2

**Zugl.: Berlin, Freie Univ., Diss., 2009**

Dissertation, Freie Universität Berlin

**D 188**

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Buches, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Die Wiedergabe von Gebrauchsnamen, Warenbezeichnungen, usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürfen.

This document is protected by copyright law.

No part of this document may be reproduced in any form by any means without prior written authorization of the publisher.

alle Rechte vorbehalten | all rights reserved

© mensch und buch verlag 2009

choriner str. 85 - 10119 berlin

verlag@menschundbuch.de – www.menschundbuch.de

Für Lisa, Bjarne und Mika



<b>1</b>	<b>EINLEITUNG UND FRAGESTELLUNG .....</b>	<b>1</b>
<b>2</b>	<b>LITERATURÜBERSICHT .....</b>	<b>3</b>
2.1	<i>Escherichia coli</i> .....	3
2.2	Kommensalen und Pathogene.....	5
2.3	Extraintestinal-pathogene <i>E. coli</i> (ExPEC).....	6
2.3.1	Uropathogene <i>E. coli</i> (UPEC) .....	7
2.3.2	Neugeborenenmeningitis-assoziierte <i>E. coli</i> (NMEC).....	8
2.3.2.1	Modelle der Blut-Hirn-Schranke.....	9
2.3.3	Aviär pathogene <i>E. coli</i> (APEC) .....	9
2.4	Pathogenitätsinseln.....	10
2.4.1	Allgemein .....	10
2.4.2	GimA (genomic island of meningitic <i>E. coli</i> containing <i>ibeA</i> ) .....	11
2.4.2.1	<i>ibeA</i> .....	12
2.4.2.2	GimA.....	12
2.5	Phylogenetischer Hintergrund.....	14
2.5.1	Multilokus-Enzym-Elektrophorese (MLEE) .....	15
2.5.2	Sequenz-basierte Methoden .....	16
2.5.2.1	Multi-Lokus-Sequenztypisierung (MLST).....	17
2.5.2.2	Möglichkeiten zur weitergehenden Verwendung von MLST-Sequenzdaten ..	18
2.5.2.2.1	Minimum-Spanning-Tree (MSTree) .....	18
2.5.2.2.2	Clonal Frame-Analyse .....	19
2.5.2.2.3	Analyse der Populationsstruktur .....	20
2.6	Single-Nukleotid-Polymorphismen (SNPs) und Selektion .....	21
2.6.1	Pathoadaptation.....	22
2.6.2	Spuren von multiplen horizontalen Gentransfers.....	23
2.6.3	Detektion von positiver Selektion .....	23
<b>3</b>	<b>MATERIAL UND METHODEN .....</b>	<b>26</b>
3.1	Material.....	26
3.1.1	Bakterienstämme .....	26
3.1.2	Oligonukleotide .....	26
3.1.3	Chemikalien, Reaktionskits und Enzyme .....	27
3.1.3.1	Puffer und Lösungen .....	27
3.1.3.1.1	Reagenzien für Polymerase-Ketten-Reaktion (PCR) .....	27
3.1.3.1.2	Hybridisierungslösungen .....	28
3.1.3.1.3	Agarose-Gelelektrophorese .....	29
3.1.3.2	Reaktionskits.....	30
3.1.3.3	Chemikalien .....	30
3.1.4	Software .....	30
3.1.5	Geräte und sonstiges .....	31
3.1.6	Nährmedien und -böden .....	32

<b>3.2</b>	<b>Methoden</b> .....	<b>33</b>
3.2.1	Kultur von <i>E. coli</i> .....	33
3.2.2	DNS-Präparation.....	33
3.2.3	Konzentrationsbestimmung von Nukleinsäuren .....	33
3.2.4	Polymerase-Kettenreaktion (PCR).....	34
3.2.4.1	PCR-Analyse des putativen GimA-Insertionslokus .....	35
3.2.4.1.1	Screening-PCR.....	35
3.2.4.1.2	Long-Range-PCR.....	36
3.2.4.1.3	<i>In silico</i> -Analyse von vollständig sequenzanalysierten <i>E. coli</i> -Stämmen ..37	
3.2.5	Agarose-Gelelektrophorese .....	37
3.2.6	DNS-DNS-Hybridisierung.....	37
3.2.6.1	Dot-Blot.....	37
3.2.6.2	Sondenherstellung mittels PCR.....	38
3.2.6.3	Hybridisierungsvorgang und Visualisierung.....	38
3.2.6.4	Strippen der Membran.....	39
3.2.6.5	Analyse des putativen GimA-Insertionslokus mittels DNS-DNS-Hybridisierung.....	39
3.2.7	Phylogenetische Analyse .....	39
3.2.7.1	Multi-Lokus-Sequenz-Typisierung (MLST) .....	39
3.2.7.2	Minimum-Spanning-Tree (MSTree).....	42
3.2.7.3	Clonal Frame-Analyse.....	42
3.2.7.4	Analyse der Populationsstruktur (Structure-Analyse).....	43
3.2.8	Sequenzanalyse von <i>ibeA</i> und Computergestützte Analyse .....	44
3.2.8.1	Sequenzanalyse von <i>ibeA</i> .....	44
3.2.8.2	Single-Lokus-Sequenz-Typisierung (SLST).....	45
3.2.8.3	Evolutionäre Analyse von <i>ibeA</i> .....	46
3.2.9	Statistik .....	46
<b>4</b>	<b>ERGEBNISSE</b> .....	<b>47</b>
<b>4.1</b>	<b>Nachweis von GimA</b> .....	<b>47</b>
4.1.1	PCR-Analyse des putativen GimA-Insertionslokus.....	47
4.1.1.1	Screening-PCR.....	47
4.1.1.2	Long-Range-PCR.....	48
4.1.1.3	<i>In silico</i> -Analyse von vollständig sequenzanalysierten <i>E. coli</i> -Stämmen.....	50
4.1.2	DNS-DNS-Hybridisierung.....	52
4.1.3	Zusammenfassung der Untersuchung zum Nachweis von GimA.....	54
<b>4.2</b>	<b>Verteilung der Varianten des GimA-Insertionslokus innerhalb der <i>E. coli</i>- Population</b> .....	<b>55</b>
4.2.1	Multi-Lokus-Sequenztypisierung (MLST) .....	55
4.2.1.1	GimA+.....	56
4.2.1.2	GimA-remnant.....	56
4.2.1.3	GimA-.....	57
4.2.1.4	Heterogene STs .....	59
4.2.2	Minimum-Spanning-Tree (MSTree).....	59
4.2.3	Clonal Frame-Analyse.....	61
4.2.4	Structure-Analyse .....	62
4.2.5	Gemeinsame Betrachtung von Clonal Frame- und Structure-Analyse.....	66
4.2.5.1	Verteilung der Varianten des GimA-Insertionslokus.....	67
4.2.5.2	Assoziationen zwischen der Herkunft der Isolate und den Locusvarianten....	68
4.2.5.3	Assoziationen zwischen den Pathovaren und Locusvarianten.....	68

<b>4.3</b>	<b>Sequenzanalyse von <i>ibeA</i></b> .....	<b>70</b>
4.3.1	Sequenzanalyse von <i>ibeA</i> .....	70
4.3.2	Single-Lokus-Sequenztypisierung (SLST).....	71
<b>4.4</b>	<b>Evolutionäre Analyse von <i>ibeA</i></b> .....	<b>79</b>
<b>5.</b>	<b>DISKUSSION</b> .....	<b>81</b>
5.1.	<b>Polymorphismus des GimA-Insertionslokus</b> .....	<b>81</b>
5.2.	<b>Assoziation von GimA mit Phylogruppe B2</b> .....	<b>82</b>
5.2.1.	GimA – eine intrinsische chromosomale Region .....	83
5.2.2.	Entwicklung von <i>ibeA</i> .....	84
5.2.3.	GimA als originärer Bestandteil der Phylogruppe B2 .....	86
5.3.	<b>Entwicklung von GimA</b> .....	<b>89</b>
5.3.1.	Reduktive Evolution .....	90
5.3.2.	„Change of function“ Mechanismus .....	90
5.4.	<b>Variantenbildung und Pathovare</b> .....	<b>91</b>
5.4.1.	APEC und NMEC.....	91
5.4.2.	UPEC.....	93
5.4.3.	Asymptomatische Besiedlung der Harnwege .....	93
5.5.	<b>Weitergehende Charakterisierung des <i>pptE</i>-Fragmentes</b> .....	<b>94</b>
<b>6</b>	<b>ZITIERTE LITERATUR</b> .....	<b>96</b>
<b>7</b>	<b>ZUSAMMENFASSUNG</b> .....	<b>113</b>
<b>8</b>	<b>SUMMARY</b> .....	<b>115</b>
<b>9</b>	<b>ANHANG</b> .....	<b>117</b>
<b>10</b>	<b>PUBLIKATIONSVERZEICHNIS</b> .....	<b>136</b>
<b>11</b>	<b>DANKSAGUNG</b> .....	<b>137</b>
<b>12</b>	<b>SELBSTÄNDIGKEITSERKLÄRUNG</b> .....	<b>139</b>

**Abkürzungsverzeichnis**

A. bidest.	Aqua bidest (zweifach destilliertes Wasser)
Abb.	Abbildung
Acc. No.	Accession Number
APEC	aviär pathogene <i>E. coli</i>
AS	Aminosäure
bp	Basenpaare
BRSS	Blocking Reagenz Stock Solution
bzw.	beziehungsweise
CNF	zytotoxisch-nekrotisierender Faktor
°C	Grad Celsius
d. h.	das heißt
DNS	Desoxyribonukleinsäure
dNTP	Desoxyribonukleotidtriphosphat
<i>E. coli</i>	<i>Escherichia coli</i>
EcoR Sammlung	<i>E. coli</i> Reference Sammlung
EDTA	Ethylendiamintetraacetat
ET	Elektrophoresetyp
ExPEC	extraintestinal-pathogene <i>E. coli</i>
FP	Forward Primer
g	Gramm
GimA	Genomic island of menigitic <i>E. coli</i> containing <i>ibeA</i>
G+C-Gehalt	Guanin- und Cytosin-Gehalt
h	Stunde
H <sub>2</sub> O	Wasser
HBMEC	Human brain microvascular endothelial cells
HCl	Hydrochlorid (Salzsäure)
HGT	Horizontaler Gentransfer
ibe	invasion of brain endothelium
IMT	Institut für Mikrobiologie und Tierseuchen
kb	Kilobasenpaare
KCl	Kaliumchlorid
konz.	konzentriert
l	Liter
Lsg.	Lösung
max.	maximal
µg	Mikrogramm
µl	Mikroliter
µM	Mikromol
mg	Milligramm
MG	Molekulargewicht
MgCl <sub>2</sub>	Magnesiumchlorid
min	Minuten
ml	Milliliter
MLEE	Multilokus Enzym Elektrophorese
MLST	Multi-Lokus-Sequenztypisierung
mM	Millimolar
MSTree	Minimum-Spanning-Tree
Multiplex-PCR	Multiplex-Polymerase-Kettenreaktion
n	number (Anzahl)
NaCl	Natriumchlorid
NaOH	Natriumhydroxid
NMEC	Neugeborenenmeningitis-assoziierte <i>E. coli</i>
OMP	Outer membrane protein
ORF	open reading frame

PCR	Polymerase-Kettenreaktion
pmol	Pikomol
RP	Reverse Primer
s.	siehe
SDS	Sodium (Natrium) Dodecylsulfat
sec	Sekunden
SepEC	Septikämie-assoziierte <i>E. coli</i>
SLST	Single-Lokus-Sequenztypisierung
SNP	Single-Nukleotid-Polymorphismus
s. o.	siehe oben
sog.	sogenannte
spp.	Spezies
SSC	Sodium (Natrium) Chloride Sodium Citrate
ssp.	Subspezies
ST	Sequenztyp
STM	Signature tagged transposon mutagenesis
s. u.	siehe unten
Tab.	Tabelle
TBE	Tris-Borsäure-EDTA
TE	Tris-EDTA
$t_{\text{melt}}$	Schmelztemperatur
U	Units
u. a.	unter anderem
ü. N.	über Nacht
UPEC	uropathogene <i>E coli</i>
V	Volt
VAG	Virulenz-assoziiertes Gen
v. a.	vor allem
XMFA	multi-FASTA Format
z. B.	zum Beispiel
z. T.	zum Teil



## 1 EINLEITUNG UND FRAGESTELLUNG

Die Bakterienspezies *Escherichia coli* zeichnet sich durch einen hohen Grad an Diversität aus. Neben Kommensalen kommen extraintestinal-pathogene (ExPEC) und intestinal-pathogene *E. coli*-Stämme vor (Bergthorsson und Ochman, 1995; Bergthorsson und Ochman, 1998; Binnewies *et al.*, 2006; Jaureguy *et al.*, 2008). ExPEC werden derzeit in vier Pathovaren unterteilt: (i) Neugeborenenmeningitis-assoziierte (NMEC), uropathogene (UPEC), aviär pathogene (APEC) und Septikämie-assoziierte *E. coli* (SepEC) (Gross, 1994; Kaper *et al.*, 2004; Russo und Johnson, 2000).

Durch NMEC ausgelöste Meningitiden bei Neugeborenen gehören weltweit zu den wichtigsten neonatalen Infektionen (Bonacorsi und Bingen, 2005; Johnson und Russo, 2002). Die Pathogenese einer bakteriellen Meningitis läuft in verschiedenen Stufen ab. Zu Beginn erfolgt die Kolonisierung des Wirtes durch den Erreger, danach folgt die Invasion ins Blutgefäßsystem. Dort muss der Erreger den bakteriziden Einflüssen des Serums widerstehen, um schließlich die Blut-Hirn-Schranke überwinden zu können (Kim, 2003; Leib und Tauber, 1999). Der Translokation der Blut-Hirn-Schranke kommt auch bei den NMEC-assoziierten Neugeborenenmeningitiden eine Schlüsselrolle in der Pathogenese zu. Verschiedene Faktoren, die an diesem Schritt beteiligt sind, konnten identifiziert werden. Beispiele sind *asIA* (Hoffman *et al.*, 2000), *ibeA* (Huang *et al.*, 2001b; Huang *et al.*, 1995), *ibeB* (Huang *et al.*, 1999), *yijP* (Wang *et al.*, 1999) und *ompA* (Prasadarao *et al.*, 1996). Lediglich *ibeA* verfügt über kein Homolog bei nicht-pathogenen *E. coli*-Stämmen. Weiterhin konnte für *ibeA* kürzlich eine Beteiligung an der systemischen *E. coli*-Infektion durch APEC nachgewiesen werden (Germon *et al.*, 2005).

Sequenzanalysen ergaben, dass *ibeA* zu einem 20,3 kb großen Gencluster gehört, der aus insgesamt 14 ORFs (open reading frame) besteht. Aufgrund eines im Vergleich zum restlichen Genom (50,8%) deutlich geringeren G+C-Gehaltes (46,2%) wurde diese Region als eine genetische Insel eingestuft und GimA (genetic island of meningitic *E. coli* containing *ibeA*) genannt. Die ORFs sind in vier Operons organisiert und ihre Genprodukte den funktionellen Kategorien Kohlenhydratmetabolismus und Substrattransport zugeordnet (Huang *et al.*, 2001a).

Während die Rolle, die GimA und besonders die des *ibeA*, im Rahmen der Pathogenese von Neugeborenenmeningitiden und aviären systemischen *E. coli*-Infektionen einnimmt, gut charakterisiert ist, ist über die Verteilung von GimA bzw. *ibeA* im phylogenetischen Hintergrund der *E. coli*-Population nur wenig bekannt. Einige Publikationen geben die Anteile *ibeA*-positiver Stämme in den ExPEC-Pathovaren an. Demnach sind NMEC zu 38,5 - 58,9%, UPEC zu 18,2 - 19,2% und APEC zu 14,2 - 26,2% *ibeA*-positiv (Ewers *et al.*, 2007; Germon *et al.*, 2005; Johnson *et al.*, 2008; Moulin-Schouleur *et al.*, 2007).

Es liegen jedoch nur wenige Untersuchungen vor, die das Vorkommen von *ibeA* mit dem phylogenetischen Hintergrund korrelieren. Da *ibeA* i. d. R. in Stämmen auftritt, die zur Phylogruppe B2 gehören, wird angenommen, dass das Gen als ein Marker für diese Phylogruppe dienen könnte, obwohl in verschiedenen Untersuchungen nicht alle B2-Stämme *ibeA*-positiv waren (Ewers *et al.*, 2007; Gordon *et al.*, 2008; Moulin-Schouleur *et al.*, 2007).

Ziel der vorliegenden Arbeit war es, die Verteilung von GimA im phylogenetischen Hintergrund der ExPEC zu bestimmen. Außerdem war die Evolution von GimA Gegenstand eingehender Untersuchungen. Besonderes Gewicht wurde dabei auf die Zusammensetzung von GimA gelegt, d. h. auf die strukturelle Kontinuität. Weiterhin wurden Sequenzanalysen des *ibeA*-Genes phylogenetisch verwandter Stämme vorgenommen mit dem Ziel, zu klären, inwieweit mehrfacher horizontaler Gentransfer an der Entwicklung und Verbreitung der Insel beteiligt ist.

## 2 LITERATURÜBERSICHT

### 2.1 *Escherichia coli*

Die Spezies *Escherichia coli* wurde erstmals 1885 durch Theodor Escherich, damals noch unter der Bezeichnung „*Bacterium coli commune*“, als physiologischer Bewohner der Dickdarmflora beschrieben (Escherich und Bettelheim, 1885). 1958 wurde die Spezies offiziell in *Escherichia coli* umbenannt.

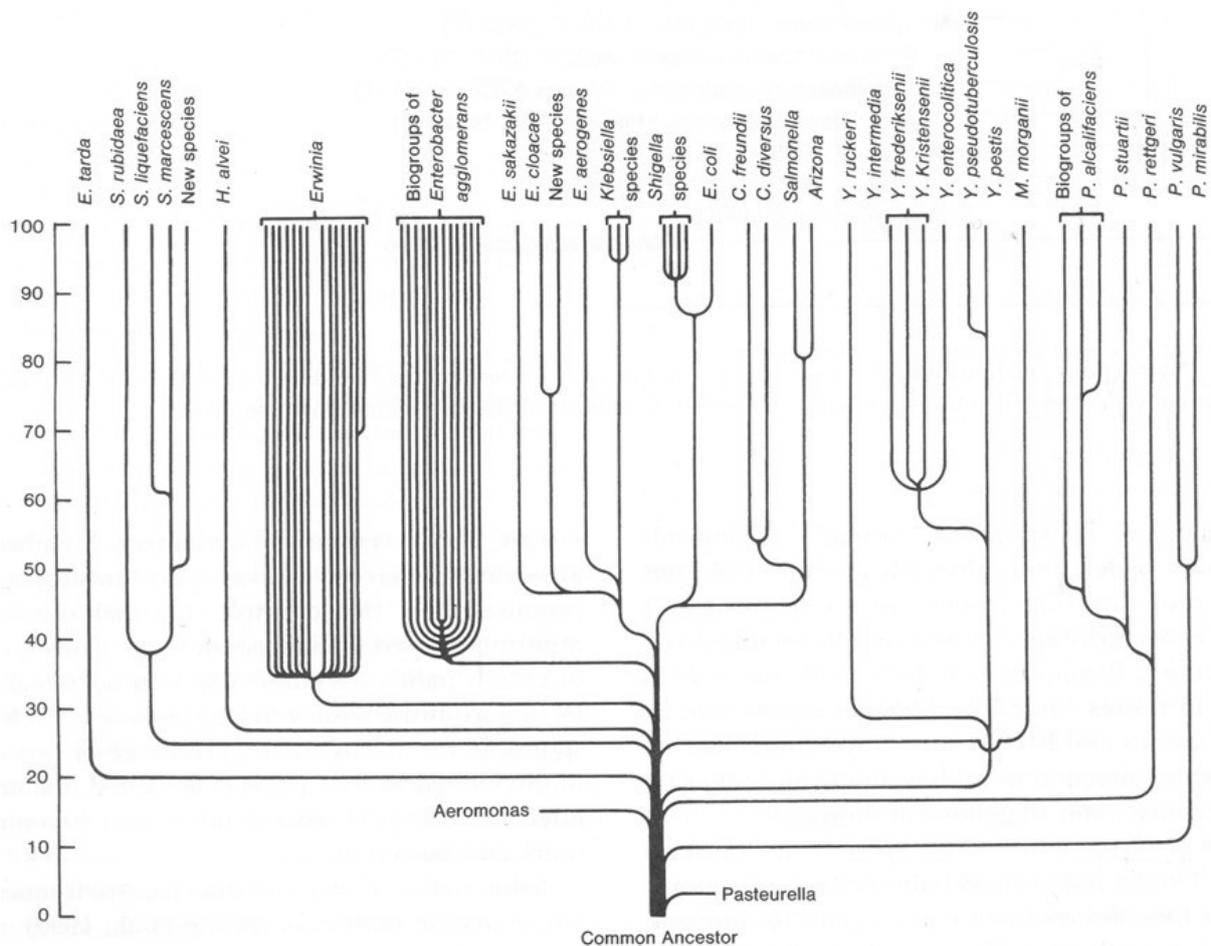


Abb. 1 Divergenzen innerhalb der Familie *Enterobacteriaceae*. Die Ordinate gibt die relative Verwandtschaft in Prozent an. Die vereinfachte Darstellung in der Grafik ermöglicht es die Verwandtschaft von Enterobakterien zu anderen Spezies abzubilden. Es wird angenommen, dass sich alle Organismen von einem gemeinsamen Vorfahren ausgehend entwickelt haben. Der Abstand zum letzten gemeinsamen Vorfahren zweier Organismen drückt ihren relativen Verwandtschaftsgrad aus. Z. B. ist *E. tarda* zu 20 % verwandt mit allen Spezies mit nicht jedoch mit den Gattungen *Aeromonas*, *Proteus*, *Providencia*, *Pasteurella* und *Yersinia* (Brenner und Farmer III, 2005b).

*E. coli* gehört zur Familie der *Enterobacteriaceae* und steht in enger verwandtschaftlicher Beziehung zur Gattung *Shigella*, die ursprünglich einen *E. coli*-Klon darstellt (Lan und

Reeves, 2002), und in naher Verwandtschaft zu den Gattungen *Citrobacter*, *Salmonella* und *Klebsiella*. Etwas größer ist der Abstand zu den anderen Gattungen der *Enterobacteriaceae* (Brenner und Farmer III, 2005a).

*E. coli* ist ein 1,1-1,5 x 2,0-6,0 µm großer Gram-negativer, stäbchenförmiger, fakultativ anaerober Mikroorganismus (Brenner und Farmer III, 2005a). Unter optimalen Bedingungen können sich *E. coli*-Zellen alle 20 Minuten teilen. Diese idealen Voraussetzungen sind allerdings außerhalb der Laborumgebung kaum gegeben. Im üblichen Umfeld werden Verdopplungsraten von 30 Stunden (menschlicher Darm) bis zu mehreren Tagen (wässrige Biotope im Freien) erreicht (Schulze *et al.*, 2006).

Die Serologie als klassisches Differenzierungsverfahren unterscheidet zur Einteilung der unterschiedlichen *E. coli*-Stämme O-Antigene (hitzestabile Bestandteile des Lipopolysaccharid-Komplexes (LPS) der äußeren Membran), K-Antigene (Polysaccharide der Kapsel) und H-Antigene (Geißeln). Bisher sind 181 O-, 80 K- und 56 H-Antigene bekannt. Durch Kombination der einzelnen Antigene ist eine sehr große Zahl an unterschiedlichen Serotypen möglich (Schulze *et al.*, 2006).

Die Entdeckung der Fähigkeit von *E. coli*, genetisches Material zwischen zwei Bakterienzellen auszutauschen (Lederberg und Tatum, 1946), eröffnete die Möglichkeit genexperimenteller Arbeit. Seither ist *E. coli* zum Standardobjekt für Genetiker und Molekularbiologen geworden. Weiterhin dient das Bakterium der therapeutischen und kommerziellen Herstellung von Proteinen (Posfai *et al.*, 2006).

Die große Bedeutung, die *E. coli* beigemessen wird, zeigt sich unter anderem in der großen Zahl vollständig sequenzanalysierter *E. coli*-Genome. Aktuell sind mindestens 30 verschiedene Genome verfügbar. Keine andere Bakteriengattung wurde häufiger vollständig sequenzanalysiert. Dazu gehören unter anderem (Tab. 1):

Tab. 1 Auswahl von vollständig sequenzanalysierten *E. coli*-Chromosomen. Unter den angegebenen Accession-Nummern sind die Sequenzdaten online frei zugänglich, z. B. unter <http://www.ncbi.nlm.nih.gov/sites/entrez>.

<i>E. coli</i>	Referenz	Accession-No
MG1655	(Blattner <i>et al.</i> , 1997)	U00096
W3110	(Riley <i>et al.</i> , 2006)	AP009048
DH10B	(Durfee <i>et al.</i> , 2008)	CP000948
ATCC8739	(Copeland <i>et al.</i> , 2008 unpublished)	CP000946
EDL933	(Perna <i>et al.</i> , 2001)	AE005174
Sakai	(Makino <i>et al.</i> , 1999)	BA000007
EC4115	(Eppinger <i>et al.</i> 2008 unpublished)	CP001164

<b><i>E. coli</i></b>	<b>Referenz</b>	<b>Accession-No</b>
55989	(Genoscope -,C.E.A., 2009 unpublished)	CU928145
E24377A	(Rasko et al., 2007 unpublished)	CP000800
E2348/69	(Iguchi et al., 2009)	FM180568
CFT073	(Welch et al., 2002)	AE014075
UTI89	(Chen et al., 2006)	CP000243
536	(Hochhut et al., 2006)	CP000247
F11	(Rasko <i>et al.</i> , 2008)	AAJU00000000
S88	(Genoscope -,C.E.A., 2009 unpublished)	CU928161
RS218	(Kim et al. 2006) ; <a href="http://www.genome.wisc.edu/sequencing/rs218.htm">http://www.genome.wisc.edu/sequencing/rs218.htm</a>	
APEC O1	(Johnson <i>et al.</i> , 2007)	CP000468
ED1a	(Genoscope -,C.E.A., 2009 unpublished)	CU928162
HS	Rasko et al., 2007 unpublished)	CP000802
IAI1	(Genoscope -,C.E.A., 2009 unpublished)	CU928160
IAI39	(Genoscope -,C.E.A., 2009 unpublished)	CU928164
SE11	(Oshima <i>et al.</i> , 2008)	AP009240
SMS-3-5	(Fricke et al., 2008)	CP000970
UMNO26	(Genoscope -,C.E.A., 2009 unpublished)	CU928163

## 2.2 Kommensalen und Pathogene

*E. coli* kolonisiert den Gastrointestinaltrakt von neonaten Menschen und Tieren innerhalb weniger Stunden nach der Geburt. Beide, Bakterium und Wirt, koexistieren normalerweise zu beiderseitigem Nutzen. Das Bakterium besiedelt vor allem die Muzinschicht des Dickdarms von Menschen, anderen Säugern und Vögeln. Es ist dabei so erfolgreich, dass es der häufigste fakultative Anaerobier der menschlichen intestinalen Darmflora ist (Barnes und Gross, 1997; Gross, 1994; Kaper *et al.*, 2004). Nur selten lösen solche kommensalen *E. coli* Erkrankungen wie Peritonitis aus. Dann liegt meistens eine Schwächung des Immunsystems oder eine Schädigung der intestinalen Barrieren zugrunde (Kaper *et al.*, 2004).

Neben diesen Kommensalen gibt es verschiedene pathogene *E. coli*-Varianten mit spezifischen Virulenzeigenschaften, die es dem Bakterium ermöglichen, neue Nischen zu besiedeln und ein breites Spektrum an Krankheiten auszulösen. Diese Eigenschaften werden durch Virulenzfaktoren vermittelt, die häufig auf mobilen genetischen Elementen (z. B. Pathogenitätsinseln) (s. 2.4) kodiert sind. Wenn diese genetischen Elemente mobilisiert werden, können sie zu neuen Kombinationen von Virulenzfaktoren zusammengesetzt werden. Die Mobilität solcher Elemente ist dabei ebenfalls einer Dynamik unterworfen und kann verloren gehen, so dass mobile Elemente gewissermaßen im Chromosom stranden. Besonders erfolgreiche Kombinationen bleiben über einen langen Zeitraum stabil. Viele Pathovaren sind gekennzeichnet durch solche Kombinationen von Virulenzfaktoren. Beispielsweise bilden diese bei den enterohämorrhagischen (EHEC) oder enteropathogenen *E. coli* (EPEC) die Grundlage für die Zuordnung zu den jeweiligen Pathovaren (Hayashi *et al.*, 2001; Perna *et al.*, 2001). Die Pathovaren sind in der Lage, eine Erkrankung in gesunden Individuen auszulösen (Kaper *et al.*, 2004).

Die Erkrankungen, die durch pathogene *E. coli* ausgelöst werden, können im Allgemeinen je nach ihrer Lokalisation in drei Gruppen eingeteilt werden: zwei lokale, mehr oder weniger auf ein Organsystem begrenzte Erkrankungen (Gastrointestinalerkrankungen und verschiedene Infektionen des Urogenitaltrakts) sowie ein systemisches Krankheitsbild (Sepsis/Meningitis) (Kaper *et al.*, 2004).

Die intestinal-pathogenen *E. coli* gliedern sich in sechs gut charakterisierte Kategorien: i) enteropathogene *E. coli* (EPEC), ii) enterohämorrhagische *E. coli* (EHEC), iii) enterotoxische *E. coli* (ETEC), iv) enteroaggregative *E. coli* (EAEC), v) enteroinvasive *E. coli* (EIEC) und vi) diffus adhärente *E. coli* (DAEC) (Nataro und Kaper, 1998). Demgegenüber stehen extraintestinale Infektionen, die durch uropathogene (UPEC), Neugeborenenmeningitis-assoziierte (NMEC), Sepsis-assoziierte (SepEC) (Russo und Johnson, 2000) sowie aviär pathogene (APEC) *E. coli* (Ewers *et al.*, 2003; Gross, 1994) ausgelöst und insgesamt als extraintestinal-pathogene *E. coli* (ExPEC) angesprochen werden (Kaper *et al.*, 2004).

### **2.3 Extraintestinal-pathogene *E. coli* (ExPEC)**

Untersuchungen an UPEC, SepEC und NMEC ergaben, dass sie sich als Gruppe betrachtet in epidemiologischer und phylogenetischer Hinsicht sowohl von den Kommensalen als auch von den intestinalen Pathogenen unterscheiden. Daher wurden sie als eigene Gruppe von Pathogenen klassifiziert, die extraintestinal-pathogenen *E. coli* (ExPEC) (Russo und Johnson, 2000). So zeichnen sich kommensale *E. coli* durch eine überwiegende Zugehörigkeit zu den phylogenetischen (MLEE) Gruppen A oder B1 (s. 2.5.1) und dem Fehlen der meisten Virulenzfaktoren aus (Herzer *et al.*, 1990; Johnson *et al.*, 2001).

Intestinal-pathogene *E. coli* weisen zwar hinsichtlich der Zuordnung zu den phylogenetischen Gruppen (A, B1 und D) (s. 2.5.1) Überschneidungen mit den Kommensalen auf, aber sie verfügen im Unterschied zu diesen über bestimmte Virulenzfaktoren, die eine Zuordnung zu bestimmten Pathovaren ermöglichen, z. B. *Shiga*-Toxin für EHEC oder Intimin für EPEC (Robins-Browne und Hartland, 2002). Im Gegensatz dazu finden sich ExPEC vor allem in der Phylogruppe B2 und etwas seltener in Gruppe D (s. 2.5.1). Sie verfügen über ein weites Spektrum an Virulenzfaktoren, die mit extraintestinaler Pathogenität assoziiert sind (Johnson *et al.*, 2001; Johnson und Russo, 2002), darunter Adhäsine, Toxine, Kapseln, Siderophore, Proteasen, Invasine und Serumresistenz-Proteine (Johnson, 1991).

ExPEC sind dank der Ausstattung mit Virulenzfaktoren und ggf. unterstützt durch eine Beeinträchtigung des Wirtes (Immunsuppression, Verletzungen) in der Lage, eine Reihe unterschiedlicher extraintestinaler Erkrankungen hervorzurufen. Sie sind die häufigsten Erreger von Infektionen des Urogenitaltraktes sowie von Bakteriämien und Septikämien. Außerdem gehören sie zu den wichtigsten Erregern von Neugeborenenmeningitiden und -septikämien. Zwar sind ExPEC scheinbar nicht in der Lage, intestinale Erkrankungen hervorzurufen (Johnson und Russo, 2002), können aber den Darm dauerhaft besiedeln. In gesunden Wirten können ExPEC die dominante Gruppe innerhalb der *E. coli*-Population ausmachen und folglich das Reservoir für eine Ansteckung bilden (Johnson, 1991). Viele Faktoren, die die Virulenz von ExPEC bestimmen (z. B. Eisenmetabolismus, Adhäsine), stellen dementsprechend auch einen Selektionsvorteil der Stämme in der normalen Umgebung des Darms dar (Le Gall *et al.*, 2007).

### **2.3.1 Uropathogene *E. coli* (UPEC)**

Der Urogenitaltrakt zählt zu den am häufigsten von bakteriellen Infektionen betroffenen Organsystemen. *E. coli* ist an dieser Lokalisation mit Abstand der häufigste Erreger. Verschiedene Untersuchungen zeigten eine Assoziation von verschiedenen Faktoren und dem Auftreten von Zystitis und Pyelonephritis. Dazu zählen Fimbrien (P-Fimbrien, Typ 1- und andere Fimbrien), Aerobactin, Serum-Resistenz-Proteine und Kapselbildung (Johnson, 1991; Manges *et al.*, 2001; Phillips *et al.*, 1988). Außerdem werden Toxine, wie Hämolyysin A und der zytotoxisch-nekrotisierende Faktor (CNF) produziert. Diese Virulenzfaktoren kommen in unterschiedlichen prozentualen Anteilen in den verschiedenen UPEC vor (Johnson und Stell, 2000). Zusätzlich können UPEC Träger von Pathogenitätsinseln sein (Oelschlaeger *et al.*, 2002; Sabaté *et al.*, 2006).

Eine Urogenitalinfektion beginnt i. d. R. mit der Besiedlung des Darmes durch einen uropathogenen Stamm zusätzlich zur kommensalen Flora (Wiles *et al.*, 2008). Solch ein Stamm ist in der Lage, die Periurethralregion zu kolonisieren und von der Harnröhre bis in die Blase aufzusteigen (Kaper *et al.*, 2004). Daneben wird auch die Möglichkeit einer

sexuellen Übertragung diskutiert (Foxman *et al.*, 2002; Johnson *et al.*, 2002). In der Blase wird die Expression von Typ 1-Fimbrien induziert, die die Adhäsion an Zellen des Urothel vermitteln (Connell *et al.*, 1996; Gunther *et al.*, 2001). Die Anheftung triggert Apoptose und Exfoliation des Blasenepithels und spielt somit eine bedeutende Rolle in der Entstehung von Urogenitalinfektionen. Stämme, die bis zum Nierenbecken aufsteigen und eine Pyelonephritis auslösen, exprimieren weniger Typ 1-Fimbrien (Gunther *et al.*, 2001). Eine dadurch weniger feste Bindung an das Blasenepithel ermöglicht eine Loslösung und über die Harnleiter einen weiteren Aufstieg in die Nieren. Eine Anheftung an das Epithel der Nieren erfolgt über P-Fimbrien und kann zusammen mit anderen bakteriellen Produkten (z. B. LPS) eine akute Entzündung auslösen (Eden und Hansson, 1978; Korhonen *et al.*, 1986). Eine Schädigung des Epithels und tieferer Gewebeschichten durch Toxine (Hämolyysin A) kann die Barriere zum Blutstrom schwächen (Trifillis *et al.*, 1994). In der Folge ist es möglich, dass Bakterien das Endothel überwinden und eine Bakteriämie auslösen (Kaper *et al.*, 2004).

### **2.3.2 Neugeborenenmeningitis-assoziierte *E. coli* (NMEC)**

NMEC zählen zu den häufigsten Erregern von durch Gram-negative Bakterien ausgelösten Meningitiden bei Neugeborenen. Die Erkrankung wird mit einer Letalitätsrate von 15-40% angegeben und geht bei den Überlebenden mit schweren neurologischen Defekten einher (Bonacorsi und Bingen, 2005; Dawson *et al.*, 1999; Johnson *et al.*, 2002; Unhanand *et al.*, 1993). Besonders bei niedrigem Geburtsgewicht steigt die Gefahr von durch NMEC bedingten Erkrankungen, während der Anteil von Gram-positiven Erregern an Fällen von Neugeborenen-Meningitis abnimmt (Stoll *et al.*, 2002). Bei den NMEC kommen nur einige wenige Serogruppen vor (O1, O2, O18, O45) (Korhonen *et al.*, 1985; Plainvert *et al.*, 2007). Außerdem sind 80% der Stämme vom Kapseltyp K1 (Kim, 2002).

Grundsätzlich läuft die Entwicklung einer bakteriellen Meningitis in Neugeborenen in vier Stufen ab. Am Beginn steht die (i) Kolonisierung der Eintrittspforte, gefolgt von der (ii) Invasion des Blutgefäßsystems durch die Erreger. Im Anschluss erfolgt die (iii) Anheftung und Invasion der Meningen und schließlich die (iv) Entzündung der Meningen und des Gehirns (De Louvois, 1994). Die Blut-Hirn-Schranke nimmt eine Schlüsselstellung in der Pathogenese ein. Jeder Erreger, der das ZNS über eine hämatogene Streuung erreicht, (blood-borne pathogens) muss diese Barriere überwinden. Die Translokation kann auf mehreren Wegen erfolgen, z. B. durch eine parazelluläre Passage, einen transzellulären Transport oder verborgen in migrierenden Leukozyten (Leib und Tauber, 1999).

Für NMEC stellt sich die Situation wie folgt dar: Auf eine initiale Besiedlung folgend steht am Beginn der Erkrankung eine hämatogene Streuung, wobei das Ausmaß der Bakteriämie die Wahrscheinlichkeit der Entstehung einer Meningitis beeinflusst. Ein Keimgehalt von mehr als  $10^3$  Kolonie-bildenden Einheiten pro Milliliter Blut erhöht die Wahrscheinlichkeit der

Entwicklung einer Meningitis deutlich und scheint ein Schwellenwert zu sein (Dietzman *et al.*, 1974; Kim, 2003). *E. coli* gelangt vom Blut in das zentrale Nervensystem durch Überwindung der Blut-Hirn-Schranke, ohne dabei Schaden an ihren Komponenten zu erzeugen, d. h. auf transzellulärem Weg (Stins *et al.*, 2001)

### **2.3.2.1 Modelle der Blut-Hirn-Schranke**

Wie beschrieben, stellt die Invasion in Zellen des mikrovaskulären Gehirnendotheliums einen elementaren Schritt in der Überwindung der Blut-Hirn-Schranke und damit der Meningitispathogenese dar (Leib und Tauber, 1999). Zwei Modelle stehen für die Untersuchung des Mechanismus dieser Translokation bzw. der daran beteiligten Faktoren zur Verfügung: (i) Ein *in vitro*-System, bestehend aus einem Monolayer von humanen mikrovaskulären Gehirnendothelzellen (HBMEC- human brain microvascular endothelial cells) als Modell der Blut-Hirn-Schranke (Stins *et al.*, 1999). An diesem Modell kann die Fähigkeit von *E. coli*-Stämmen zur Invasion in die Endothelzellen untersucht werden. (ii) Ein *in vivo*-System, bestehend aus neugeborenen Ratten, die peritonäal mit *E. coli* infiziert werden. Bestimmt werden der Keimgehalt des Blutes und der Cerebrospinal-Flüssigkeit, wobei ein Erregernachweis im Liquor als Meningitis gewertet wird (Huang *et al.*, 1999; Kim, 1985; Kim *et al.*, 1992).

### **2.3.3 Aviär pathogene *E. coli* (APEC)**

APEC lösen eine akut verlaufende systemische Erkrankung des Geflügels aus, die von einigen Autoren als Kolibazillose angesprochen wird (Barnes und Gross, 1997; Dho-Moulin und Fairbrother, 1999; Dziva und Stevens, 2008; Gross, 1994; Janßen *et al.*, 2001). Vorwiegend betroffen sind Hühner und Puten, aber auch Enten und anderes Wassergeflügel. Die Erkrankung geht mit einer hohen Morbidität und Mortalität einher und verursacht in der Geflügelindustrie, unabhängig von der Haltungsform (Extensiv- oder Intensivhaltung), erhebliche wirtschaftliche Verluste in Form von verminderter Lebendgewichtszunahme und erhöhter Mortalität (Ewers *et al.*, 2003). Die Erkrankung kann als Primärinfektion auftreten, sowie sekundär auf Infektionen des Respirationstraktes folgen, die beispielsweise durch das Newcastle Disease Virus (NDV), das Infektiöse Bronchitis Virus (IBV) oder *Mycoplasma gallisepticum* verursacht werden (Willinger, 1992).

Eintrittspforte für die Erreger ist der aviäre Respirationstrakt. Der Erreger wird durch Inhalation von Kot-kontaminiertem Staub aufgenommen (Dho-Moulin und Fairbrother, 1999). Zusätzlich werden durch den Kot Futter und Trinkwasser verunreinigt (Ewers *et al.*, 2003). Die Erkrankung tritt gehäuft ab der 4. Lebenswoche auf, was mit einer erhöhten Erregerdichte im Stall, sowie mit dem häufigeren Auftreten von Mykoplasmen- und

Virusinfektionen ab der 3. Lebenswoche in Zusammenhang gebracht werden kann (Willinger, 1992). Eine weitere Häufung von Erkrankungen tritt beim Erreichen der höchsten Legeleistung auf. Stress, z. B. durch eine erhöhte Besatzdichte im Stall, Immunsuppression der Tiere sowie Verletzungen sind weitere prädisponierende Faktoren. Besonders Verletzungen können eine weitere Eintrittspforte darstellen. Die häufigsten O-Gruppen bei APEC sind O1, O2, O18 und O78, sie machen bis zu 61 % der untersuchten Isolate aus (Blanco *et al.*, 1997; Dho-Moulin und Fairbrother, 1999; Ewers *et al.*, 2007).

Eine APEC-Infektion kann zur Ausprägung unterschiedlicher Symptome, wie Dottersackentzündungen, Entzündungen des Respirationstraktes, Eileiterentzündungen, Peritonitis, Zellulitis, Enteritis, Swollen Head Syndrome und neurologischen Ausfallerscheinungen führen. Dabei handelt es sich um lokale Infektionen und solche, die als Folge einer in der Regel tödlich verlaufenden Septikämie auftreten. Die Pathogenese-Mechanismen sind noch weitgehend unklar (Gross, 1994). In den Luftsäcken kommt es nach aerogener Aufnahme des Erregers zu Entzündungen (Barnes und Gross, 1997). In der Folge gelangt der Erreger in die Blutbahn und verursacht so eine systemische Symptomatik, einhergehend mit den genannten entzündlichen Veränderungen an verschiedenen Organen der Tiere. Legehennen können APEC neben der beschriebenen horizontalen Weitergabe auch vertikal entweder über die durch Kot verunreinigte Eierschale oder im Rahmen einer Salpingitis direkt auf das Ei übertragen (Gross, 1994). Aufgrund der Vielzahl der betroffenen Organe und der Verlaufsformen ist es sinnvoll statt Kolibazillose den Begriff systemische *E. coli*-Infektion zu verwenden.

## **2.4 Pathogenitätsinseln**

### **2.4.1 Allgemein**

Umfangreiche Sequenzanalysen der letzten Jahre haben gezeigt, dass das Genom von Prokaryonten prinzipiell aus einem konservierten Kerngenom besteht, in das ein flexibler Anteil eingebettet ist. Die Nukleotidsequenz des sogenannten „core genome“ ist bei *E. coli* durch einen ausgeglichenen Guanin-Cytosin (G+C)-Gehalt gekennzeichnet, wohingegen sich der G+C-Gehalt im sogenannten flexiblen Genpool deutlich vom core genome unterscheidet (Hacker und Kaper, 2000). Der flexible Genpool wird durch horizontalen Gentransfer (HGT) zwischen den Bakterien verbreitet und wird neben Plasmiden und Phagen durch konjugative Transposons, Integrons und genomische Inseln (GEI) vermittelt (Hacker und Carniel, 2001; Kirsch *et al.*, 2004). In Abhängigkeit von den Genen, für die die GEIs kodieren, können sie außer zur Pathogenität (Pathogenitätsinseln) auch zur Symbiose (Symbioseinseln), Resistenz (Resistenzinseln) oder zum Metabolismus (Metabolismusinseln) der Bakterien beitragen (Dobrindt *et al.*, 2004). Pathogenitätsinseln (PAI) stellen demnach

eine Untergruppe von GEIs dar (Hacker *et al.*, 1997) und gelangten in der Regel durch HGT in das Genom (Hacker *et al.*, 2003).

Hacker *et al.* (2003) stellten Kriterien zur Definition von Pathogenitätsinseln auf. Demnach sind PAIs charakterisiert durch (i) das Vorhandensein von Virulenz-assoziierten Genen (VAGs), (ii) eine Größe von üblicherweise mehr als 30 kb, (iii) das Vorkommen in pathogenen und das Fehlen in nicht-pathogenen Stämmen, (iv) einen unterschiedlichen G+C-Gehalt verglichen mit dem des Chromosoms des Wirtsbakteriums (core genome) und (v) die Existenz von flankierenden Mobilitäts- und Insertionselementen, z. B. IS-Elementen, Integrasen, Transposasen, „direct repeats“ und Assoziation mit tRNA-Genen (Hacker *et al.*, 1997; Hacker und Kaper, 2000; Hentschel und Hacker, 2001). Für *E. coli* sind zahlreiche PAIs beschrieben worden (Tab. 2):

Tab. 2 Pathogenitätsinseln bei verschiedenen ExPEC mod. nach (Hacker und Kaper, 2000; Johnson *et al.*, 2007; Rumer *et al.*, 2003) (Auswahl). In den Stämmen 536 und APEC O1 kommen vier verschiedene PAIs unterschiedlicher Größe und assoziiert mit unterschiedlichen tRNA-Genen nebeneinander vor (PAI I-IV), GimA (genomic island of meningitic *E. coli* containing *ibeA*) des Stamms RS218 ist mit keinem tRNA-Gen assoziiert. Die aufgeführten relevanten Gene der PAIs kodieren für einen Hämolyisin-Gen-Cluster (*hly*); einen P-Fimbrien Gen-Cluster (*pap*); einen Gen-Cluster, der den P-Fimbrien ähnelt (*prf*); für das „*E. coli* Attaching and Effacing gene A“ (*eaeA*); für Gene, die für sezernierte Proteine von enteropathogenen *E. coli* (*espAB*), für ein Invasin (*ibeA*), für Siderophore (Yersiniabactin-Operon, *ireA*) und für einen Autotransporter (*vat*).

<i>E. coli</i> -Stamm	Pathovar	Pathogenitätsinsel (PAI)	relevante Gene	Größe (kb)	assoziiertes tRNA-Gen
536	UPEC	PAI I	<i>hly</i>	70	<i>selC</i>
536	UPEC	PAI II	<i>hly, prf</i>	190	<i>leuX</i>
536	UPEC	PAI III	<i>eaeA, espAB</i>	25	<i>thrW</i>
536	UPEC	PAI IV	<i>hly pap</i>	45	<i>asnT</i>
RS218	NMEC	GimA	<i>ibeA</i>	20	-
APEC O1	APEC	PAI <sub>APEC O1</sub> I	<i>ireA, pap</i>	83	<i>phe</i>
APEC O1	APEC	PAI <sub>APEC O1</sub> II	<i>hly</i>	28	<i>asp</i>
APEC O1	APEC	PAI <sub>APEC O1</sub> III	<i>vat</i>	19	<i>thr</i>
APEC O1	APEC	PAI <sub>APEC O1</sub> IV	Yersiniabactin-Operon	87	<i>asn</i>

#### 2.4.2 GimA (genomic island of meningitic *E. coli* containing *ibeA*)

Bevor GimA als genetische Insel erkannt wurde, wurde zunächst einer ihrer Bestandteile (*ibeA* – invasion of brain endothelium A) beschrieben. Um die Chronologie der

Entdeckungen einzuhalten, wird im Folgenden zunächst *ibeA* und im Anschluss die genetische Insel beschrieben.

#### **2.4.2.1 *ibeA***

Huang *et al.* (1995) führten eine Signature-tagged-transposon-Mutagenese (STM) an einem zur Invasion des Zentralen Nervensystems (ZNS) fähigen *E. coli*-Stamm (NMEC-Prototypstamm RS218, O18:K1:H7) durch. Bei der Mutagenese werden zufällig Gene durch Insertion des Transposons ausgeschaltet (Knockout-Mutante) (Hensel *et al.*, 1995; Li *et al.*, 2005). Wie bereits bei der Beschreibung der NMEC erläutert (s. 2.3.2), stellt die Überwindung der Blut-Hirn-Schranke einen wesentlichen Schritt in der Pathogenese dar. Mit Hilfe der zwei Modelle der Blut-Hirn-Schranke [(i) *in vitro* (human brain microvascular endothelial cells – HBMEC) und (ii) *in vivo* (newborn rat)] können die erzeugten Mutanten hinsichtlich einer Attenuierung durch die STM untersucht werden. Eine Mutante mit attenuiertem Phänotyp in beiden Modellen war 10A-23, das entsprechende Gen wurde *ibe10* (invasion of brain endothelium) genannt (Huang *et al.*, 1995). Auf der Oberfläche von HBMEC wurde ein ~45 kDa großes Protein identifiziert, das als Rezeptor für *ibe10* fungiert und Ibe10R genannt wurde. Anti-Ibe10R-Antikörper verhinderten eine Invasion des Wildtypstamms in HBMEC (Prasadarao *et al.*, 1999). Später wurden weitere Oberflächenproteine von HBMEC [Vimentin und PSF (polypyrimidine tract-binding protein (PTB)-associated splicing factor)], die mit *ibe10* interagieren, ausgemacht (Zou *et al.*, 2006; Zou *et al.*, 2007). Weitere Untersuchungen identifizierten *ibe10* als einen Teil eines 1,37 kb großen offenen Leserahmens (ORF - open reading frame), der für ein 456 Aminosäuren großes, IbeA genanntes Protein, mit einem Molekulargewicht von ~50 kDa kodiert. Das IbeA-Protein enthält drei transmembranale Domänen. Wie die Knock-out-Mutante 10A-23 war auch eine isogene Deletionsmutante von *ibeA* (ZD1) *in vitro* und *in vivo* weniger invasiv. Eine Deletion des gesamten Gens führt also ebenfalls zu einem attenuierten Phänotyp, verglichen mit dem Ursprungstamm (Huang *et al.*, 2001b).

38.5-58.9% der NMEC sind *ibeA*-positiv (Ewers *et al.*, 2007; Johnson *et al.*, 2008). Aber auch in APEC kommt *ibeA* mit Häufigkeiten von 14.2-26.2% der Stämme vor (Ewers *et al.*, 2007; Germon *et al.*, 2005). Außerdem konnten Germon *et al.* (2005) zeigen, dass *ibeA* sowohl in APEC eine wesentliche Rolle bei der Invasion in HBMEC als auch in der Pathogenität in einem Hühnerinfektionsmodell spielt.

#### **2.4.2.2 GimA**

2001 wurde veröffentlicht, dass *ibeA* Teil eines 20,3 kb großen Genclusters ist, der sich zwischen den Genen *yjiD* und *yjiE* in der Nähe des *fim*-Operons in der 98 min-Region des *E. coli* K-12 Chromosoms befindet. Das *fim*-Operon kodiert für ein Adhäsin (Typ 1-Fimbrien).

Der GimA-Cluster besteht aus 14 ORFs, die in vier Operons (GimA1: *pptE*, *pgdK*, und *pdaK*; GimA2: *cgrD*, *cgxT*, *cdlD* und *cniT*; GimA3: *gclK*, *ghbD*, *gclR* und *gclA* sowie GimA4: *ibeRAT*) organisiert sind (s. Abb. 2). Hinsichtlich ihrer Funktionen wurden die einzelnen Gene anhand von computergestützten Analysen der kodierenden Sequenzen in die funktionalen Kategorien Kohlenhydratmetabolismus und Substrattransport eingeordnet (Hejnova *et al.*, 2005; Huang *et al.*, 2001a; Zou *et al.*, 2008) (Tab. 3). Die ORFs des GimA1-Operons sind von Bedeutung im Phosphotransferase-System. Dabei handelt es sich um ein aktives Stofftransportsystem, das bisher nur bei Eukaryonten beschrieben wurde (De Crecy-Lagard *et al.*, 1995). Es werden hochenergetische Phosphatreste von Phosphoenolpyruvat (PEP) über mehrere Proteinkinasen (Signaltransduktion) zunächst auf Transportproteine und schließlich auf ein Substrat übertragen, meist Hexosen oder Zuckeralkohole. Das phosphorylierte Substrat kann in der Folge durch Membranproteine in das Zytoplasma eingeschleust werden (De Crecy-Lagard *et al.*, 1995). Die ORFs des GimA2-Operons gehören zum Glycerol-Metabolismus. Die Oxidoreduktasen *cgrD* und *cdlD* katalysieren die Hydrierung von Reduktionsäquivalenten [NADH (Nicotinamid-Adenin-Dinukleotid) bzw. FAD (Flavin-Adenin-Dinukleotid)]. Die Transporter *cgxT* (Glycerol-3-Phosphat) und *cniT* (Carnitin) dienen als Substrat-Shuttle vom Zytosol in das Mitochondrium. Der Glyoxylat-Metabolismus ist für die Funktion des Zitronensäurezyklus von Bedeutung. Da einige Intermediate des Zitronensäurezyklus zur Biosynthese abgezweigt werden, müssen sie über Auffüllreaktionen (anaplerotische Reaktion) ersetzt werden. Einen wichtigen anaplerotischen Reaktionsweg stellt der Glyoxylat-Bypass dar (Holms, 1986). Zu diesem Bypass gehören die vier ORFs des GimA3-Operons. Es wird die Reaktion von Isocitrat und Acetyl-CoA zu Malat und Succinat katalysiert. Dieser Bypass ermöglicht *E. coli* beispielsweise das Wachstum auf Essigsäure. Das letzte GimA-Operon (GimA4) beinhaltet die Gene *ibeR*, *ibeA* und *ibeT*. IbeR ist ein Regulatorprotein, das in die Translokation der Blut-Hirn-Schranke involviert ist und soll außerdem als Regulatorprotein im Pyruvatstoffwechsel für die Stressregulation von *E. coli* in der stationären Wachstumsphase bedeutend sein (Chi *et al.*, 2009). IbeA wurde bereits in Abschnitt 2.4.2.1 beschrieben. Funktionelle Untersuchungen zeigen, dass IbeT als Adhäsion zur Anheftung an das Endothel dient und daher in die Meningitis-Pathogenese eingebunden ist (Zou *et al.*, 2008).

Insgesamt weist die Nukleotidsequenz des GimA-Clusters mit 46,2% im Vergleich zu 50,8% des restlichen Chromosoms einen niedrigeren G+C-Gehalt auf (Huang *et al.*, 2001a). Da solche Unterschiede im G+C-Gehalt ein Charakteristikum für genetische Inseln sind (Hacker *et al.*, 1997), wurde diese Region zu einer genetischen Insel erklärt und GimA (genomic island of meningitic *E. coli* containing *ibeA*) genannt (Huang *et al.*, 2001a).

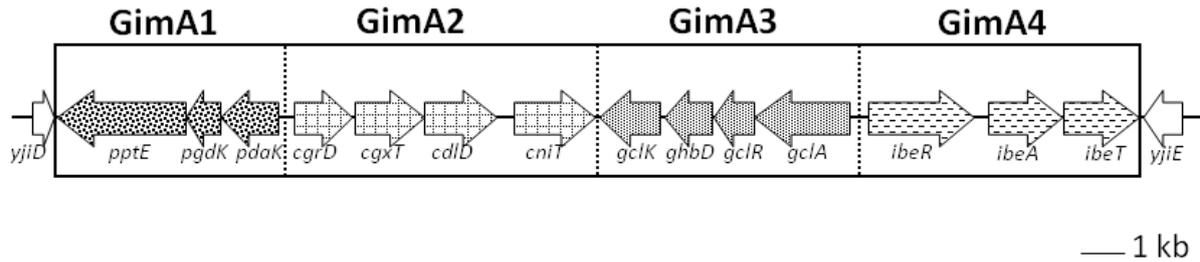


Abb. 2 GimA (genomic island of meningitic *E. coli* containing *ibeA*) der Stamm RS218 (Huang *et al.*, 2001a) und A0 34/86 (Hejnova *et al.*, 2005), bestehend aus vier Operons (GimA1-4) mit 14 ORFs und den flankierenden Genen *yjiD* und *yjiE*.

Tab. 3 Genannotation der ORFs des GimA-Clusters basierend auf der Identifizierung von Homologien mit bekannten Proteinen und deren Einordnung in ein globales Stoffwechselsystem (Hejnova *et al.*, 2005; Huang *et al.*, 2001a; Zou *et al.*, 2008)

Operon	ORF	putative Funktion	biolog. Funktion
GimA1	<i>pptE</i>	Phosphoenolpyruvat-Protein Phosphoryltransferase	Phosphotransferase- System
	<i>pgdK</i>	Dihydroxyaceton-Kinase	
	<i>pdaK</i>	Dihydroxyaceton-Kinase (DHA)	
GimA2	<i>cgrD</i>	Glycerol-Dehydrogenase	Glycerol-Metabolismus
	<i>cgxT</i>	Transporter	
	<i>cdID</i>	Dihydrolipoamid-Dehydrogenase	
	<i>cniT</i>	Carnitin-ransporter (CAIT)	
GimA3	<i>gclK</i>	Glycerat-Kinase	Glyoxylat- Metabolismus
	<i>ghbD</i>	3-Hydroxyisobutyrat- Dehydrogenase	
	<i>gclR</i>	Glycoxylat-Carboligase	
	<i>gclA</i>	Hydroxypyruvat-Isomerase	
GimA4	<i>ibeR</i>	Regulatorprotein	Pyruvat-Metabolismus, Stressregulation
	<i>ibeA</i>	Invasionsprotein	Invasion in HBMEC
	<i>ibeT</i>	Adhäsion	Adhäsion HBMEC

## 2.5 Phylogenetischer Hintergrund

Neben der Unterscheidung in Kommensalen und pathogene Stämme sowie der Serotypisierung können *E. coli*-Isolate in phylogenetische Gruppen eingeteilt werden. Es existieren verschiedene Ansätze zur Untersuchung der Phylogenie innerhalb der *E. coli*-

Population. Die Bildung der Serogruppen ist ein unabhängiger Prozess, folglich spiegelt die Serotypisierung die Phylogenie der *E. coli*-Population nicht wider.

### **2.5.1 Multilokus-Enzym-Elektrophorese (MLEE)**

Goulet und Picard (1989) führten eine Multilokus-Enzym-Elektrophorese (MLEE) mit den Stämmen der *E. coli* Referenz Sammlung (*E. coli* reference collection – EcoR) (Ochman und Selander, 1984) durch. Dabei wurden aufgereinigte Esterasen (Esterasen A, B, C und I) mittels Elektrophorese in ihre Isoenzyme aufgetrennt. Die so entstandenen Bandenmuster der einzelnen Esterasen eines Isolates wurden zu einem Elektrophoresetyp (ET) zusammengefasst. Innerhalb der 72 Stämme der EcoR-Sammlung konnten anhand der Polymorphismen in der Elektrophorese der Esterasen vier phylogenetische Hauptgruppen definiert werden, die als A, B1, B2 und D bezeichnet wurden (Goulet und Picard, 1989; Herzer *et al.*, 1990). Eine weitere potenzielle Phylogruppe (E) spielt eine nur untergeordnete Rolle (Touchon *et al.*, 2009). Mittlerweile liegen PCR-basierte und damit erheblich schnellere und günstigere Methoden zur Feststellung der phylogenetischen Gruppe eines Isolates vor (Clermont *et al.*, 2000). Eine neue Untersuchung kommt jedoch zu dem Ergebnis, dass die Zuordnung zu den (Haupt-) Phylogruppen mittels der PCR-basierten Methode bei einem gewissen Anteil der getesteten Stämme nicht funktioniert. In Abhängigkeit von der Phylogruppe sind die getesteten Isolate in der Größenordnung von fünf Prozent (Phylogruppen B1 und B2) bis 24% (Phylogruppe D) einer falschen Phylogruppe zugeordnet worden (Gordon *et al.*, 2008).

Kommensale *E. coli* konnten überwiegend zu den phylogenetischen Gruppen A und B1 zugeordnet werden (Herzer *et al.*, 1990; Johnson *et al.*, 2001). Verschiedene Untersuchungen an den intestinalen *E. coli* zeigten Assoziationen mit gleich drei phylogenetischen Gruppen (A, B1 und D) (Nataro und Kaper, 1998). Folglich kann nicht die Zugehörigkeit zu einer bestimmten phylogenetischen Gruppe ausschlaggebend für die Definition eines Pathovars sein, sondern das Vorliegen von Pathovar-spezifischen Virulenzgenmustern (s. 2.2 und 2.3). Im Gegensatz zu Kommensalen und intestinal-pathogenen *E. coli* finden sich ExPEC vor allem in Phylogruppe B2 und etwas seltener in Gruppe D (Russo und Johnson, 2000). ExPEC, intestinal-pathogene und kommensale *E. coli* finden sich jedoch auch in anderen Phylogruppen. Die MLEE und die auf ihr basierende Einteilung in phylogenetische Gruppen stellt eine sehr grobe und eher der Übersicht dienende Methode dar. Differenzierte Analysen von z. B. einzelnen Pathovaren sind schwer möglich.

## 2.5.2 Sequenz-basierte Methoden

Obwohl Bakterien sich klonal reproduzieren, wird die Entwicklung ihres Genoms durch eine Reihe von Mechanismen beeinflusst. Unter anderem gehören dazu Punktmutationen, Genom-Rearrangement, Deletion, Duplikation, Transposition sowie homologe und nicht-homologe Rekombination (Achtman, 2004; Feil *et al.*, 1999; Feil *et al.*, 2000; Lawrence und Hendrickson, 2003; Smith *et al.*, 1993). Die Vielfalt evolutionärer Mechanismen, die die Entwicklung von Bakterien beeinflussen, erschwert die Bestimmung der Verwandtschaft von Stämmen. Eine klonale Verwandtschaft kann durch einen Stammbaum dargestellt werden, in dem jedes Blatt einem Stamm und jeder innere Knoten dem letzten gemeinsamen Vorfahren entspricht. Die Abstände zu den einzelnen Verzweigungen drücken den zeitlichen Abstand oder Sequenzunterschied bis zum letzten gemeinsamen Vorfahren mit den anderen von der entsprechenden Verzweigung abgehenden Linien aus. Die Methoden zur Vorhersage solcher Verwandtschaften beruhen meist auf Unterschieden in der DNS-Sequenz. Kämen ausschließlich Punktmutationen vor, könnten phylogenetische Standardmethoden (Neighbour-Joining Algorithmus, Maximum Parsimony Algorithmus) die klonale Verwandtschaft einer Population akkurat darstellen (s. Abb. 3) (Didelot und Falush, 2007).

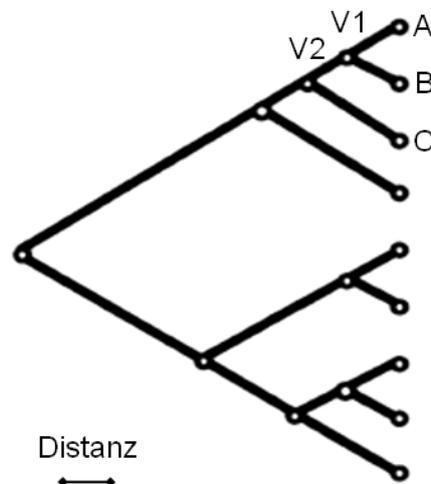


Abb. 3 Beispiel-Dendrogramm für eine Population, die sich ohne Rekombination entwickelt und in zwei Hauptäste gliedert. Die Distanz zwischen A und B entspricht zwei Distanzeinheiten, da A und B jeweils eine Distanzeinheit vom letzten gemeinsamen Vorfahren (V1) entfernt sind. Entsprechend beträgt der Abstand zwischen A und C vier Einheiten, da beide den Vorfahren (V2) gemeinsam haben. mod. nach (Smith *et al.*, 1993)

Homologe Rekombinationen können jedoch in Bakterien der gleichen Art Änderungen von mehreren Nukleotiden gleichzeitig einführen, ohne die Funktionalität der (Ziel-) Gene nachteilig zu beeinflussen, da sie aus einem funktionierenden Spender stammen (Milkman

und Crawford, 1983). Besonders für Gene, die elementare Faktoren im Stoffwechsel kodieren (Housekeeping-Gene), ist dieser Mechanismus von Bedeutung, da andere Veränderungen in ihrer Sequenz, wie Deletion oder Insertion, mit hoher Wahrscheinlichkeit letal sind und daher sehr selten vorkommen (Milkman und Crawford, 1983). Solche homologen Rekombinationen können die Effekte von Punktmutationen deutlich überwiegen und somit zu Berechnung falscher Stammbäume führen (Schierup und Hein, 2000). Die Notwendigkeit, neben Punktmutationen auch Rekombinationseffekte bei der Erzeugung von Stammbäumen zu berücksichtigen, hat u. a. zu der Entwicklung der Multi-Lokus-Sequenztypisierung (MLST) geführt. Die Verwendung von mehreren Housekeeping-Genen, die sich über das gesamte Genom verteilen, reduziert den Einfluss drastisch, den ein einzelnes Rekombinationsereignis hat, da die Wahrscheinlichkeit, dass mehr als ein Locus davon betroffen ist, sehr gering ist (Maiden *et al.*, 1998).

### **2.5.2.1 Multi-Lokus-Sequenztypisierung (MLST)**

Die MLST wurde 1998 von Maiden *et al.* (1998) zur Charakterisierung von Bakterien eingeführt und basiert auf der Annahme, dass es im Genom von Bakterien hoch konservierte Bereiche gibt. Gene in diesen Bereichen werden als Housekeeping-Gene bezeichnet und unterliegen einer nur sehr langsam fortschreitenden Veränderung im Sinne einer Änderung ihrer Nukleotidsequenz. Diese Änderungen treten als Single-Nukleotid-Polymorphismen (SNP) (=Punktmutationen) auf. Für die MLST eines bestimmten Isolates werden mehrere Housekeeping-Gene (=Loci) sequenzanalysiert und die Sequenzen der einzelnen Loci (=Allele) werden webbasiert mit einer Datenbank abgeglichen. Dabei wird jedem Locus ein bereits vorhandener Alleltyp zugeordnet bzw., wenn die entsprechende Sequenz noch nicht hinterlegt ist, ein neuer Alleltyp bestimmt. Die Kombination aus den Alleltypen (=Allelprofil) wird zu einem Sequenztyp (ST). STs, die sich voneinander lediglich in einem Allel unterscheiden, werden zu einem Sequenztyp-Komplex zusammengefasst (Maiden *et al.*, 1998).

Am Beispiel *Neisseria meningitidis* konnte gezeigt werden, dass die partielle Sequenzanalyse von sechs Housekeeping-Genen (oder Teilen davon) in der Lage war, die bekannten, auf der MLEE basierenden Elektrophoresetypen zu diskriminieren (Maiden *et al.*, 1998). Diese Autoren erkannten die Vorteile einer sequenzbasierten Methode in der einfachen Reproduzierbarkeit in jedem Labor und der Möglichkeit, die gewonnenen Daten webbasiert zu hinterlegen und somit ein Werkzeug für globale Epidemiologie zu ermöglichen (Maiden *et al.*, 1998). Weniger als zehn Jahre später waren bereits über 30 MLST Schemata via Internet verfügbar (<http://www.mlst.net/databases/default.asp>) (Maiden, 2006).

Für die MLS-Typisierung von *E. coli* existieren zurzeit zwei unterschiedliche Ansätze, basierend auf der Verwendung unterschiedlicher Housekeeping-Gene. Jaureguy *et al.* (2008) verwenden acht Loki (*dinB*, *icdA*, *pabB*, *polB*, *putB*, *trpA*, *trpB* und *uidA*), wohingegen das von Wirth *et al.* (2006) beschriebene Schema mit sieben Loki (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA* und *recA*) auskommt. Beide Schemata weisen eine Kongruenz mit den auf MLEE basierenden phylogenetischen Gruppen auf (Jaureguy *et al.*, 2008; Wirth *et al.*, 2006) und stimmen in der Zuordnung von in beiden Ansätzen verwendeten öffentlich zugänglichen und vollständig sequenzanalysierten Stämmen überein, d. h. in beiden Typisierungsschemata weisen die Stämme RS218, UTI89 und APEC O1 einen identischen Sequenztyp auf (Jaureguy: ST1 und Wirth: ST95). Das Schema von Wirth *et al.* umfasst zurzeit fast 1400 STs, basierend auf über 2500 Stämmen und liegt der in dieser Arbeit beschriebenen MLS-Typisierung zu Grunde (<http://www.ucc.ie/mlst>).

### **2.5.2.2 Möglichkeiten zur weitergehenden Verwendung von MLST-Sequenzdaten**

#### **2.5.2.2.1 Minimum-Spanning-Tree (MSTree)**

Zur besseren Beurteilung der gewonnenen Daten bietet sich eine grafische Darstellung an. So können Zusammenhänge einfacher erkannt werden. Der MSTree ist eine Methode zur Visualisierung von MLST-Daten. Zugrunde liegen bei der Berechnung eines MSTree die Allelprofile der eingehenden Stämme (Wirth *et al.*, 2006). Mit entsprechender Software ist es möglich, die Allelprofile innerhalb einer MLST-Datenbank als MSTree abzubilden. Dabei werden die Beziehungen unter den einzelnen STs und ST-Komplexen entsprechend der Übereinstimmungen in ihren Allelprofilen als Verbindungslinien unterschiedlicher Dicke und Art ausgedrückt (Wirth *et al.*, 2006).

Beide o. g. MLST-Ansätze führen zu sehr ähnlichen MSTrees. Als Beispiel ist der MSTree von Jaureguy *et al.* (2008) abgebildet (s. Abb. 4)s:

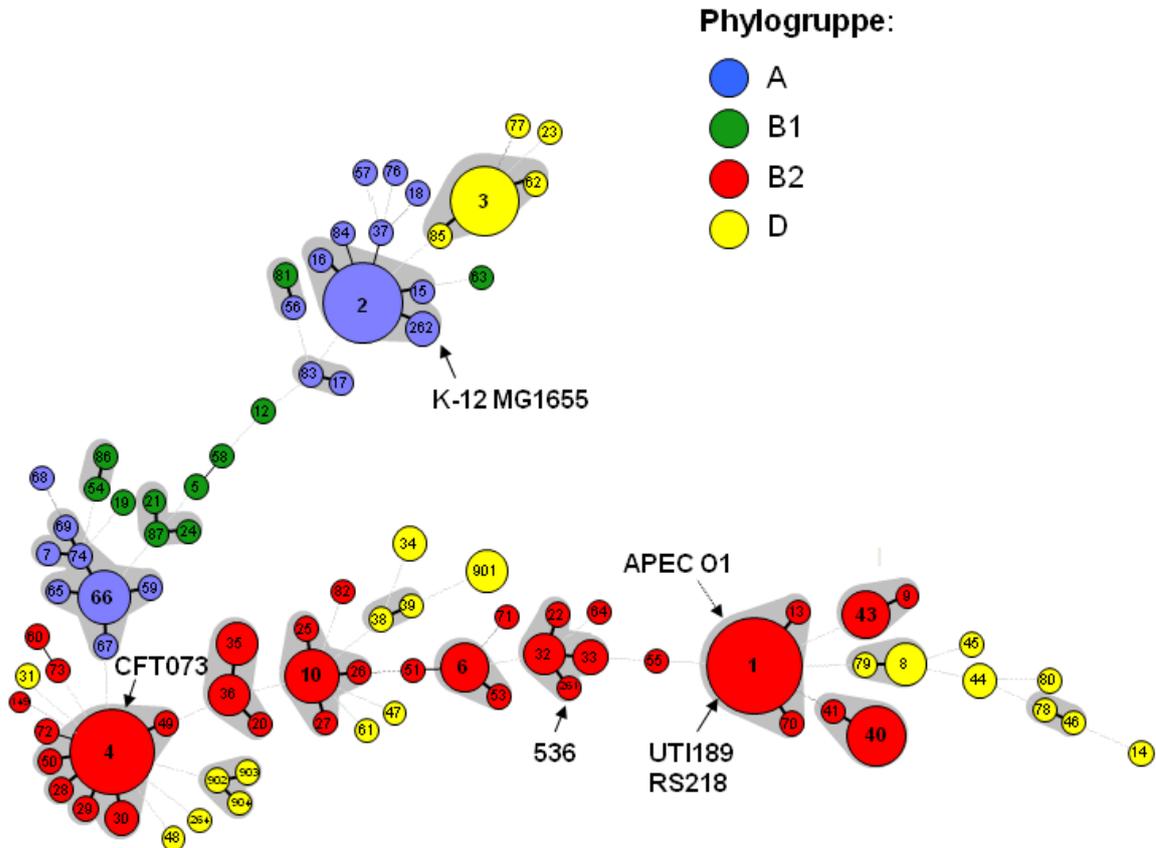


Abb. 4 Minimum-Spanning-Tree-Analyse von 161 *E. coli*-Isolaten. Der MSTree basiert auf den Allelprofilen der Housekeeping-Gene (*dinB*, *icdA*, *pabB*, *polB*, *putP*, *trpA*, *trpB* und *uidA*). Jeder Kreis entspricht einem ST, der im Zentrum des Kreises genannt ist. Die Größe eines Kreises entspricht der Anzahl der enthaltenen Isolate (z. B. enthalten die kleinsten Kreise ein Isolat und der ST4 wird von 17 Isolaten gebildet). Grau unterlegte STs unterscheiden sich nur in einem Allel von ihren direkten Nachbarn und bilden einen ST-Komplex. Die Zugehörigkeit zu einer Phylogruppe wird durch die Farbgebung ausgedrückt. Rote STs gehören zur Phylogruppe B2, gelbe gehören zur Gruppe D, grün steht für die phylogenetische Gruppe B1 und die STs der Gruppe A sind blau gefärbt. Die Zuordnung zu den Phylogruppen erfolgte PCR-basiert (Clermont *et al.*, 2000). Dicke, dünne und unterbrochene Verbindungslinien zwischen den Kreisen entsprechen einem, zwei oder drei Unterschieden im Allelprofil der STs. Zusätzlich ist für einige vollständig sequenzanalytisierte *E. coli*-Stämme die ST-Zugehörigkeit angegeben, die auch in der vorliegenden Arbeit verwendet werden (RS218, UTI89, APEC O1, CFT073, 536 und der *E. coli* K-12-Stamm MG1655). Modifiziert nach Jaureguy *et al.* (2008).

#### 2.5.2.2.2 Clonal Frame-Analyse

Allelbasierte Methoden wie die MLST (s. 2.5.2.1) sind in ihren Anwendungen limitiert. Grundsätzlich gilt, dass Allele, die sich in nur wenigen Nukleotiden unterscheiden, einen höheren Verwandtschaftsgrad aufweisen als solche mit vielen Änderungen. Treten jedoch beispielsweise Rekombinationen mit niedriger Frequenz auf und unterscheidet sich ein Stamm in sechs der sieben Allele in nur einzelnen Nucleotiden von einem zweiten, aber im letzten Locus in vielen Positionen, so würden beide in der MLST als nicht oder nur entfernt

verwandt eingestuft, obwohl sechs Allele eine enge Verwandtschaft deutlich zeigen. Außerdem bleibt der Verwandtschaftsgrad innerhalb eines Allels unberücksichtigt. Wird ein Allel nur teilweise durch homologe Rekombination ausgetauscht, ist in der allelbasierten Methode eine mögliche enge Verwandtschaft des verbliebenen Teils zu einem zweiten Stamm durch den neu eingefügten Teil maskiert. Deshalb sind allelbasierte Methoden eher als Orientierung geeignet als zur detaillierten Datenanalyse (Didelot und Falush, 2007).

Durch Rekombination finden DNS-Fragmente unterschiedlichster Herkunft Eingang in ein Chromosom. Trotz dieser Rekombinationsereignisse bleiben immer Bereiche des Chromosoms unverändert, d. h. sie weisen eine identische Herkunft auf. Diese Bereiche werden Clonal Frame genannt (Milkman und Bridges, 1990). Der Grad der Verwandtschaft zweier Stämme kann somit anhand ihres gemeinsamen Clonal Frame ausgedrückt werden. Eine Clonal Frame-Analyse versucht das Ausmaß des Clonal Frame innerhalb eines jeden Astes eines gegebenen Stammbaumes zu ermitteln, d. h. es wird eine Vorhersage bakterieller Mikroevolution gemacht. Das Ergebnis der Berechnungen kann dann durch entsprechende Software grafisch dargestellt und durch die Bildung von Clustern verdeutlicht werden. Die Methode verwendet dazu die Sequenzdaten, die im Rahmen der MLST generiert wurden (Didelot und Falush, 2007).

Die Clonal Frame-Analyse kann präzise Vorhersagen zur Mikroevolution, die sich in Form von Clusterbildungen darstellen lässt, machen. Die Methode kann jedoch keine Aussage zur Herkunft der durch homologe Rekombination importierten DNS machen (Didelot und Falush, 2007). Eine vollständige Darstellung des Ursprungs sämtlicher DNS in einer Probe ist extrem komplex und selbst für sehr überschaubare, geschlossene und homogene Populationen mit gleichmäßiger Rekombination nur schwer zu berechnen (McVean und Cardin, 2005). Anders als Eukaryonten beherbergen Bakterien häufig DNS-Fragmente anderer Spezies (Dingle *et al.*, 2005) oder sogar Genera (Ochman *et al.*, 2000), und somit ist die Annahme einer geschlossenen und homogenen Population per se unrealistisch.

#### **2.5.2.2.3 Analyse der Populationsstruktur**

Wie bereits weiter oben beschrieben (s. 2.5.2) gehen die Modelle zur Beschreibung der Phylogenie in der Regel von der Annahme aus, dass die Entwicklung der *E. coli*-Population klonal erfolgt und der Einfluss von Rekombination gering ist. Verschiedene Untersuchungen konnten unter Verwendung unterschiedlicher phylogenetischer Standardmethoden (s. 2.5.2) unabhängig voneinander weitgehend übereinstimmende Stammbäume erzeugen, die die in der EcoR-Sammlung definierten Phylogruppen (A, B1, B2 und D) diskriminieren konnten (Feil *et al.*, 2001; Reid *et al.*, 2000).

Die Existenz diskreter Gruppen bedeutet jedoch nicht zwangsläufig eine Klonalität, da die Algorithmen, die den phylogenetischen Modellen zu Grunde liegen, Stammbäume

konstruieren, ungeachtet dessen, ob Individuen Hybriden (aufgrund homologer Rekombination) sind oder nicht (Wirth *et al.*, 2006). Wirth *et al.* (2006) konnten zeigen, dass die (Standard-)Modelle weitgehend ungeeignet sind, die Phylogenie der heutigen *E. coli*-Population darzustellen. In dieser Studie ergab die Anwendung der verschiedenen Modelle unterschiedliche Anordnungen der Isolate innerhalb der Bäume, die darüber hinaus statistisch nicht abzusichern waren. Weiterhin spricht gegen eine ausschließlich klonale Entwicklung, dass in der Studie keine Methode sicher zeigen konnte, welche der Phylogruppen die ursprüngliche ist. Diese Befunde zeigen deutlich, dass homologen Rekombinationen in der Entwicklung der *E. coli*-Population eine wesentliche Rolle zukommen (Wirth *et al.*, 2006).

Für Spezies wie *E. coli*, bei denen homologe Rekombinationen vorkommen, sollten deshalb weitergehende phylogenetische Werkzeuge angewendet werden. Die Structure-Analyse analysiert die Polymorphismen in den sieben MLST-Fragmenten jedes Stammes mit Hilfe einer Methode der Wahrscheinlichkeitsrechnung (Bayesian Methode) mit dem Ziel, Gruppen innerhalb der Testpopulation zu finden (Falush *et al.*, 2003; Pritchard *et al.*, 2000). Wirth *et al.* (2006) konnten zeigen, dass die Gruppen, die in der Structure-Analyse (Software verfügbar unter [http://pritch.bsd.uchicago.edu/software/structure2\\_2.html](http://pritch.bsd.uchicago.edu/software/structure2_2.html)) der untersuchten *E. coli*-Population identifiziert wurden, weitgehend mit den Phylogruppen der MLEE übereinstimmen (Wirth *et al.*, 2006). Ein Teil der Structure-Analyse ist das „linkage model“, das für jeden Nucleotidpolymorphismus eine Wahrscheinlichkeit für die Abstammung aus einer der identifizierten Gruppen berechnet. Aus der Summe dieser einzelnen Wahrscheinlichkeiten kann dann auf die Zugehörigkeit eines Isolates zu einer der Gruppen geschlossen werden. Diese Methode kann ebenfalls die EcoR-Phylogruppen unterscheiden, ordnet aber einen Teil der Isolate Hybridgruppen (AxB1 und ABD) zu und trägt damit der Bedeutung der homologen Rekombination Rechnung (Wirth *et al.*, 2006).

## 2.6 Single-Nukleotid-Polymorphismen (SNPs) und Selektion

SNPs treten in allen Bereichen des *E. coli*-Genoms auf. Die Frequenz ihres Auftretens bestimmt die Geschwindigkeit, mit der die Diversität innerhalb einer Population zunimmt. Sie wird als molekulare Uhr (molecular clock rate) bezeichnet. Für das Genom von *E. coli* wurde eine Frequenz von  $7,6 \times 10^{-10}$  SNPs pro Jahr bestimmt, basierend auf einem Sequenzunterschied der zusammengefügteten Housekeeping-Gene der MLST von 21,3% zwischen *E. coli* und *Salmonella enterica* und der Annahme, dass sich diese beiden Spezies vor ca. 140 Millionen Jahren getrennt haben (Ochman und Wilson, 1987; Wirth *et al.*, 2006).

Das Aufspüren von Verwandtschaft und Ursprung von Organismen mittels SNP-Analyse stellt eine analytische Nutzung der informativen Punktmutationen (=SNPs) dar.

Demgegenüber steht die biologische Bedeutung von SNPs. Meistens bleiben Änderungen in der Nukleotidsequenz folgenlos. Jedoch kommt es vor, dass solche Änderungen innerhalb von kodierenden Bereichen auftreten und dann zu dramatischen Effekten führen. Beim Mensch ist beispielsweise eine Punktmutation im APOE-Gen mit der Entstehung der Alzheimerschen Krankheit assoziiert, wohingegen eine Mutation im Locus CCR5 zu einer Resistenz gegenüber HIV führen kann (Chakravarti, 2001; Stoneking, 2001). Auf *E. coli* bezogen wurde bereits erwähnt, dass Punktmutationen in Housekeeping-Genen letal sein können, sofern das Genprodukt dadurch seine Funktion im Stoffwechsel nicht mehr übernehmen kann (Milkman und Crawford, 1983).

### 2.6.1 Pathoadaptation

Die Punktmutationen (SNPs) bedingen andererseits auch das große Entwicklungspotenzial von *E. coli* und allen anderen Organismen. Sie sind die Triebfeder der Evolution. Gelangt eine Population von ihrer angestammten Nische in ein neues Habitat, so werden nicht alle Gene optimal angepasst sein. Kann sich die Population ausreichend lange in der neuen Umgebung halten, dann werden vorteilhafte Mutationen selektiert. Dies führt zu einer Differenzierung innerhalb der Population. Solche Differenzierungen können der Beginn der Bildung einer neuen Spezies sein (Orr und Smith, 1998). Sie können aber auch zur Bildung von Pathogenen für diese neue Nische aus zuvor wenig oder apathogenen Spezies führen (Sokurenko *et al.*, 1999). Ein Mikroorganismus, der in ein Habitat gelangt, das nicht sein angestammtes ist, wird suboptimal daran angepasst sein. Durch Mutation von bereits existierenden Genen kann eine Adaptation an die neue Umwelt erfolgen, die die Überlebensfähigkeit bzw. die Virulenz im neuen Habitat verbessert. Diese pathoadaptativen Mutationen können zu den beschriebenen Steigerungen der Pathogenität führen (Jores *et al.*, 2006; Sokurenko *et al.*, 1999).

Solch eine Pathoadaptation konnten z. B. Weissman *et al.* für *fimH* zeigen, das für die adhäsive Untereinheit der Typ 1-Fimbrie kodiert,. Es wurden mehrere unterschiedliche Punktmutationen in Stämmen mit gleichem phylogenetischen Hintergrund identifiziert. Derart diversifizierte Stämme wiesen in anschließenden Zellkulturversuchen ein verbessertes Adhäsionsvermögen auf (Weissman *et al.*, 2006). Ein ähnliches Phänomen wurde für das *cadBA*-Operon beschrieben. Die *cad*-Gene kodieren für eine Lysin-Decarboxylase (LDC). Änderungen der Nukleotidsequenzen der Gene können zu einem Verlust der LDC-Aktivität führen. Eine Komplementation mittels eines Plasmids, das ein funktionales *cadBA*-Operon enthält, stellt die Enzymaktivität wieder her. Eine Komplementation der Enzymaktivität in *Shigella spp.* und EIEC (s. 2.2) führt darüber hinaus zu einer Reduktion des Adhäsionsvermögens in Zellkulturversuchen. Ein Defekt des *cadBA*-Operons scheint folglich

ebenfalls eine Pathoadaptation darzustellen, da durch den Funktionsverlust von LDC eine Verbesserung der Adhäsivität erreicht wird (Jores *et al.*, 2006).

### **2.6.2 Spuren von multiplen horizontalen Gentransfers**

Neben der biologischen Funktion als Mechanismus zur Pathoadaptation können SNPs, sofern sie sich innerhalb von mobilen Elementen befinden, Aussagen bezüglich der Frequenz von horizontalem Gentransfer zulassen (Weissman *et al.*, 2006).

So ergab ebenfalls in der Untersuchung von Weissman *et al.* ein Vergleich der Sequenzen des *fimA*-Gens von 43 zum ST95-Komplex gehörenden Stämmen, dass innerhalb dieses Komplexes vier *fimA*-Linien vorkommen. *FimA* kodiert ebenfalls für eine Untereinheit der Typ 1-Fimbrie. Dagegen ergab ein Vergleich von MLST-Sequenzdaten für diese Stämme ein monophyletisches Bild. Das Auftreten von *fimA*-Varianten, die auch als Single-Lokus-Sequenztypen (SLSTs) bezeichnet werden können, innerhalb einer phylogenetischen Einheit zeigt eine mehrfache Insertion von *fimA* bzw. des gesamten *fim*-Operons in den ST95-Komplex. SLSTs basieren auf Unterschieden in der Nukleotidsequenz, die durch das Auftreten von Punktmutationen (SNPs) entstehen. Vergleichbar ist die Frequenz der Varianten mit denen der O- und H-Antigene. Die Antigene kommen in den getesteten Stämmen in jeweils drei Varianten vor und sind ebenfalls durch horizontalen Gentransfer mehrfach und unabhängig in den ST95-Komplex gelangt (Milkman *et al.*, 2003; Nelson und Selander, 1994). Triebfeder dieser antigenetischen Diversifizierung kann z. B. die Selektion durch das Immunsystem des Wirtes sein (Weissman *et al.*, 2006).

Das Auftreten der verschiedenen Antigene innerhalb dieses begrenzten phylogenetischen Bereiches zeigt nicht nur, dass die kodierenden Gene mehrfach in ein *E. coli*-Chromosom inseriert sind, sondern auch den Einfluss, den das Immunsystem des Wirtes durch die Erzeugung von Selektionsdruck auf die genetische Struktur von *E. coli* hat (Reeves, 1992).

### **2.6.3 Detektion von positiver Selektion**

Innerhalb von kodierenden Bereichen des Genoms können SNPs grundsätzlich in zwei Gruppen unterschieden werden. Die einen bleiben folgenlos, d. h. bei der Expression des Gens ändert sich die Aminosäuresequenz nicht und das Genprodukt kann seine Funktion in gleicher Weise wie in einem unveränderten Stamm erfüllen. Dass Mutationen stumm bleiben können, liegt an der Redundanz, mit der für die einzelnen Aminosäuren kodiert wird. Wie die Code-Sonne (s. Abb. 5) zeigt, werden bis auf Tryptophan und Methionin alle Aminosäuren durch mehrere Basen-Triplets kodiert. Eine Änderung der letzten Base im Triplet führt daher in der Mehrzahl der Fälle zu keiner Änderung der Aminosäuresequenz und des gesamten Proteins. Diese Mutationen werden als synonym bezeichnet.

Denen gegenüber stehen die nicht-synonymen Mutationen. Sie führen zu Änderungen der Aminosäureabfolge und ziehen folglich Veränderungen im Protein nach sich. Dies kann zu den oben beschriebenen Folgen führen, wie der Entstehung von Krankheiten, der Bildung von Resistenzen (s. 2.6) und der Pathoadaptation (s. 2.6.1) (Chakravarti, 2001; Stoneking, 2001).

Die Frage, inwieweit es sich bei diesen Mutationen um Zufall oder um gerichtete Entwicklungen im Sinne einer Selektion handelt, kann durch die Berechnung des Verhältnisses der nicht-synonymen ( $D_n$ ) zur synonymen Mutationsrate ( $D_s$ ) beantwortet werden (Hughes und Nei, 1989; Tanaka und Nei, 1989). Die Mutationsraten entsprechen dabei dem Anteil der tatsächlichen an den möglichen (synonymen oder nicht-synonymen) Mutationen.

Liegt keine Selektion vor, so sind  $D_n$  und  $D_s$  etwa gleich groß. Konservierte Bereiche zeichnen sich durch neutrale Selektion aus. Die Zahl nicht-synonymer Mutationen ist deutlich kleiner als die Anzahl synonyme. Folglich ergibt sich eine  $D_n/D_s$ -Ratio  $< 1$  und eine konservierte Aminosäuresequenz. Andersherum weist ein Polymorphismus, der aufgrund von positiver (Darwin-) Selektion entstanden ist, deutlich höhere  $D_n$  als  $D_s$  auf, da die positive Selektion Aminosäureänderungen bevorzugt (Hughes und Nei, 1989; Tanaka und Nei, 1989). Folgende Verhältnisse ergeben sich:

- ◇  $D_n/D_s > 1$  positive Selektion
- ◇  $D_n/D_s \sim 1$  keine Selektion
- ◇  $D_n/D_s < 1$  neutrale Selektion

Mit Hilfe dieser Methode und deren Modifikationen sind zahlreiche Gene unter positiver Selektion identifiziert worden (Nielsen und Yang, 1998; Petersen *et al.*, 2007; Yang und Nielsen, 2000), unter anderem auch das oben beschriebene *fimH* (Weissman *et al.*, 2006).



## 3 MATERIAL UND METHODEN

### 3.1 Material

#### 3.1.1 Bakterienstämme

Insgesamt wurden 410 *E. coli*-Stämme untersucht. Davon waren 338 Feldisolate, und 72 Stämme stammten aus der *E. coli* reference collection (EcoR) (Ochman und Selander, 1984). 98 Stämme waren als APEC (aviär pathogene *E. coli*) klassifiziert und 140 als UPEC (uropathogene *E. coli*), wobei 64 menschlichen und 76 tierischen (Hunde und Katzen) Ursprungs waren. Darüber hinaus wurden noch 25 NMEC (Neugeborenenmeningitis-assoziierte *E. coli*) und 28 Isolate aus humanen Blutproben, die als Septikämie-assoziierte *E. coli* (SepEC) bezeichnet wurden sowie 119 fäkale Isolate klinisch gesunder Menschen (n=86) und Tiere (n=33) einbezogen (Mordhorst *et al.*, 2009; Sankar *et al.*, 2009; Wirth *et al.*, 2006). Die verwendeten Stämme sind im Anhang in Tabelle 27 aufgeführt.

Zusätzlich wurden einige öffentlich verfügbare Sequenzdaten verwendet: Die vollständig sequenzanalysierten Genome der UPEC-Stämme CFT073 (Acc. No AE014075), UTI89 (Acc. No CP000243), F11 (Acc. No AAJU00000000), and 536 (Acc. No CP000247) und des APEC Stamms APEC O1 (Acc. No CP000468) sowie des K12 Stamms MG1655 (Acc. No U00096). Weiterhin wurde die veröffentlichte Sequenz von GimA des NMEC Stamms RS218 (Acc. No AF289032) genutzt.

#### 3.1.2 Oligonukleotide

Die verwendeten Primer wurden von MWG Biotech, Ebersberg synthetisiert. Entsprechend der Angaben des Syntheseberichtes wurde durch Zugabe von Wasser eine Konzentration der Stammlösung auf 100 pmol/µl eingestellt. Die Stammlösungen sowie die daraus hergestellten Gebrauchslösungen wurden bei -20°C aufbewahrt. Die Konzentration der Gebrauchslösungen betrug 10 pmol/µl. Zur Verwendung wurden die Lösungen jeweils auf Eis aufgetaut.

Die in der vorliegenden Arbeit genutzten Primer sind mit ihrer 5'-3'-Sequenz und Schmelztemperatur folgend aufgeführt (s. Tab. 4):

Tab. 4 Verwendete Oligonukleotid-Primer. Die Primer sind entsprechend des Verwendungszwecks aufgeführt. Genannt sind die 5'-3'-Sequenz, die Zielregion und die Schmelztemperatur ( $t_{melt}$ ) jedes Primers. Zusätzlich ist die erwartete Größe des Amplikons angegeben. FP steht für Forward-Primer, RP für Reverse-Primer.

Verwendung für	Bezeichnung	Sequenz 5'-3'	Zielregion	$t_{melt}$	Amplikon-Größe (bp)
Screening-PCR	yjiD-yjiE FP	ttactgtttcaggcatgcc	GimA-Locus	57	variable
	yjiD-yjiE RP	atgaatcccgttgccgaacg		57	
Long-Range-PCR	yjiD-yjiE FP	ttactgtttcaggcatgcc	Long-Range Fragment 1	56	10.549
	gclKGimA3 RP	tgaatgagttacgccagtgtg		56	
	cniTGimA2 FP	ggacactgtactgtgggta	Long-Range Fragment 2	60	10.514
	yjiD-yjiE RP	atgaatcccgttgccgaacg		60	
Synthese der DNS-Sonden für die DNS-DNS-Hybridisierung	pptE FP	atgattgaggtcctgccgc	<i>pptE</i>	57	342
	pptE RP	ttagtgggaggttctgattgc			
	pdaKGimA1 FP	tcggtttgtttgcaccacac	<i>pdaK/cgrD</i>	57	571
	cgrDGimA2 RP	ccggtcacaattcatctgca			
	cniTGimA2 FP	ggacactgtactgtgggta	<i>cniT/gclK</i>	57	475
	gclKGimA3 RP	tgaatgagttacgccagtgtg			
	gclA FP	aatatagcgtgtctccgggtcc	<i>gclA</i>	57	600
	gclA RP	aagcgagtgaattgttgcg			
	ibeA FP	tggaaccgctcgtaataac	<i>ibeA</i>	57	1270
	ibeA RP	tacgccattttgctgtaagcg			
Sequenzanalyse von <i>ibeA</i>	ibeAFr1FP	gtattagcatgatgttgcttg	<i>ibeA</i> -Fragment 1	55	~900
	ibeAFr1RP	attttcttcatccgcccccc			
	ibeAFr2FP	agctaattgttatgacaacc	<i>ibeA</i> -Fragment 2	55	~900
	ibeAFr2RP	aaaaatgatcgggtgaagcg			

### 3.1.3 Chemikalien, Reaktionskits und Enzyme

#### 3.1.3.1 Puffer und Lösungen

##### 3.1.3.1.1 Reagenzien für Polymerase-Ketten-Reaktion (PCR)

Die neben der Taq-Polymerase in der PCR verwendeten Reagenzien waren bereits bei Auslieferung mit der Polymerase als Reddy-Mix (ABgene Germany, Hamburg, Deutschland) gemischt bzw. wurden vom Hersteller der Taq-Polymerase (Rapidozym, Berlin, Deutschland) mitgeliefert (PCR-Puffer und  $MgCl_2$ -Lösung) und wurden entsprechend der

Herstellervorgaben eingesetzt. Weiterhin wurden dNTPs (Rapidozym, Berlin, Deutschland) entsprechend der Herstellerangaben verwendet (nicht im Reddy-Mix):

Tab. 5 PCR-Reagenzien für die Verwendung mit Rapidozym Taq-Polymerase

Reagenz	Zusammensetzung
10xPCR Puffer (im Lieferumfang der Taq-Polymerase enthalten)	100 mM Tris-HCl (pH 8,8) 500 mM KCl 0,1% Triton X-100
1xMgCl <sub>2</sub> Lösung (im Lieferumfang der Taq-Polymerase enthalten)	50 mM MgCl <sub>2</sub>
dATP (Rapidozym, Berlin, Deutschland)	100 mM dATP
dCTP (Rapidozym, Berlin, Deutschland)	100 mM dCTP
dGTP (Rapidozym, Berlin, Deutschland)	100 mM dGTP
dTTP (Rapidozym, Berlin, Deutschland)	100 mM dTTP

### 3.1.3.1.2 Hybridisierungslösungen

Tab. 6 Zusammensetzungen der Lösungen für die DNS-DNS-Hybridisierung

Lösung	Zusammensetzung
Puffer 1 pH 7,5	100 ml 1 m Maleinsäure 150 ml 1 m NaCl ad 1000 ml A. bidest.
Blocking Reagent Stock Solution (BRRS)	10 g in 100 ml Puffer 1 (ohne Tween)
Prähybridisierungslösung	12,5 ml 20x SSC 5 ml BRSS 1 ml 10% N-Laurolylsarcosin 0,1 ml 10% SDS ad 50 ml A. bidest.
Puffer 2	1 ml BRSS 9 ml Puffer 1 (ohne Tween)
Puffer 3 pH 9,5	10 ml 1 m Tris/HCl 10 ml 1 m NaCl ad 100 ml A. bidest.
1 m Tris/HCl pH 9.5	60,55 g in 500 ml A. bidest.
1 m Maleinsäure	58,035 g auf 500 ml A. bidest.
1 m NaCl	58,443 g auf 1000 ml A. bidest.

Lösung	Zusammensetzung
20xSSC pH 7.0	175,32 g NaCl 88,23 g Natriumcitrat ad 1000 ml A. bidest.
5xSSC	43,83 g NaCl 22,06 g Natriumcitrat ad 1000 ml A. bidest.
2xSSC	17,53 g NaCl 8,82 g Natriumcitrat ad 1000 ml A. bidest.
Stripping Puffer	0,2 M NaOH 0,1% SDS

### 3.1.3.1.3 Agarose-Gelelektrophorese

Die Agarose und die Größenmarker für die Gelelektrophorese stammten von Peqlab (Erlangen, Deutschland).

Tab. 7 Reagenzien für die Elektrophorese

Reagenzien	Zusammensetzung
1x TBE	10,8 g Tris base 5,5 g Borsäure 4,0 ml 0.5 M EDTA (pH 8.0) ad 1000 ml A. bidest.
Ladepuffer (10 ml)	10 mg Bromphenolblau 2,0 ml 0,5 M EDTA pH 8,0 0,1 ml 10 mM Tris pH 8,0 2,9 ml A. bidest. 5,0 ml Glycerin
Agarosegel (Universalagarose)	1,0 g Agarose in 100 ml 1x TBE, nach Abkühlung auf 60°C 2µl 1% Ethidiumbromid zufügen
100bp Leiter Größenmarker	162 µl A. bidest 30 µl Ladepuffer 10 µl Größenmarkerkonzentrat (100bp)
1 kb Leiter Größenmarker	200 µl A. bidest. 50 µl Ladepuffer 30 µl Größenmarkerkonzentrat (1kb)

### 3.1.3.2 Reaktionskits

Tab. 8 Verwendete Reaktionskits

Kit	Hersteller
MasterPure™ Genomic DNA Purification Kit	Epicentre, Biozym, Hessisch Oldendorf, Deutschland
PCR DIG Probe Synthesis Kit	Roche Diagnostics GmbH, Mannheim, Deutschland
DIG Luminescent Detection Kit	Boehringer Mannheim GmbH, Mannheim, Deutschland
QIAquick Gel Extraktion Kit	Qiagen, Hilden, Deutschland
Plasmid Purification Midi Kit	Qiagen, Hilden, Deutschland

### 3.1.3.3 Chemikalien

Bezugsfirmen für die in dieser Arbeit verwendeten und in den Tabellen 6 und 7 aufgeführten Chemikalien waren Merck AG, Darmstadt; New England Biolabs GmbH, Schwalbach/Taunus; Oxoid Unipath GmbH, Wesel; Promega GmbH, Mannheim; Roche Diagnostics GmbH, Mannheim Roth GmbH & Co., Karlsruhe,; Serva, Heidelberg sowie Sigma, Deisenhofen.

### 3.1.4 Software

Tab. 9 Verwendete Software

Software	Hersteller/Referenz
Bionumerics V4.6	Applied Maths, Sint-Martens-Latem, Belgien
Kodon V3.6	Applied Maths, Sint-Martens-Latem Belgien
ClonalFrame	(Didelot und Falush, 2007)
Structure 2.2	(Falush <i>et al.</i> , 2003)
MEGA 4.0	(Tamura <i>et al.</i> , 2007)
DnaSP 4.50.3	(Rozas <i>et al.</i> , 2003)

### 3.1.5 Geräte und sonstiges

Tab. 10 verwendete Geräte und sonstiges

<b>Gerät</b>	<b>Typ und Hersteller</b>
Brutschränke	Typ B6060 Heraeus
Elektrophoresekammern	AGS; Hybaid; MWG
Filmkassetten	Kodak BioMax Cassette
Photometer	NanoDrop ND-1000
Hybridisierungsöfen	Hybaid
pH-Meter	Sigma 3K30
Kühlzentrifuge	Knick 766 Calimatic
Pipetten	Eppendorf Research
Röntgenfilme	Kodak X-OMAT AR
Schüttelinkubator	IKA Labortechnik KS 250 basic
Schüttler	Hybaid PS 250
Spannungsgerät	New Brunswick Scientific
Thermocycler	TProfessional, Biometra
Tischzentrifuge	Eppendorf Centrifuge 5415D
Transilluminator	Herolab; E.A.S.Y. 429K; ICU-1; Video-Graphik Printer UP 890 CE
UV-Tisch	Biometra TI 1
Vortexer	IKA MS2 Minishaker
Waage	Julabo; Th. Karow GmbH
Wasserbäder	WTC Binder
Wärmeschrank	Sartorius LA 230S und BP 2100

### 3.1.6 Nährmedien und -böden

Die Nährmedien und -böden zur Kultivierung von *E. coli* wurden entsprechend der Herstellervorgaben (CARL ROTH GmbH + Co. KG, Karlsruhe, Deutschland) hergestellt.

Tab. 11 verwendete Nährmedien und -böden

<b>Nährmedien und -böden</b>	<b>Zusammensetzung</b>
LB (Luria Bertani)-Medium	10 g Trypton 5 g Hefeextrakt 10 g NaCl ad 1000ml A. bidest.
LB-Agar	10 g Trypton 5 g Hefeextrakt 10 g NaCl 17 g Agar ad 1000ml A. bidest.

## 3.2 Methoden

### 3.2.1 Kultur von *E. coli*

Alle Inkubationen erfolgten bei 37°C. Gefrierkulturen (-80°C) wurden auf LB-Agar ausgestrichen und über Nacht inkubiert. Flüssigkulturen in LB-Medium wurden ebenfalls über Nacht im Schüttler bei 200 UPM inkubiert.

### 3.2.2 DNS-Präparation

Chromosomale DNS von *E. coli* wurde mit dem „Master Pure™ Genomic DNA Purification Kit“ entsprechend der Herstellervorgaben isoliert. Dazu wurde 1,0 ml einer über Nacht-Kultur in LB-Medium in ein 1,5 ml Eppendorf-Reaktionsgefäß überführt und die Bakterien per Zentrifugation pelletiert. Der Überstand wurde verworfen und das Pellet in 1,0 ml 1x PBS resuspendiert und erneut abzentrifugiert. Das Pellet wurde in 600 µl Lysis-Puffer 2 aufgenommen, für zehn Minuten bei 80°C inkubiert und für zwei bis drei Minuten auf Eis abgekühlt. Anschließend wurden 3 µl RNaseA (2ng/ml) und eine 40-minütige Inkubation bei 37°C angeschlossen. Zweihundert Mikroliter der Präzipitationslösung fällten den Zelldetritus, der dann bei 16.000 x g für zehn Minuten abgeschieden wurde. Der Überstand wurde weiterverwendet und die gelöste DNS mit 600 µl Isopropanol präzipitiert und nach zehn-minütiger Inkubation bei 16.000 x g für zehn Minuten abzentrifugiert. Nach zweimaligem Waschen mit eiskaltem 70%-igem Ethanol wurde das DNS-Pellet getrocknet und schließlich in 100 µl A. bidest. aufgenommen, ggf. wurde zur Lösungserleichterung max. eine Stunde bei 65°C inkubiert. Anschließend wurde die DNS-Konzentration bestimmt und die Lösung bei -20°C eingefroren.

Eine weitere Möglichkeit zur DNS-Präparation bietet die Hitzelyse der Bakterien. Dazu wurde Koloniematerial in einem 1,5 ml Eppendorf-Reaktionsgefäß in 200 µl A. bidest. resuspendiert, für zehn Minuten bei -10°C eingefroren und anschließend für zehn Minuten aufgeköcht. Der Zelldetritus wurde abzentrifugiert und der Überstand in der PCR und für Hybridisierungen eingesetzt.

### 3.2.3 Konzentrationsbestimmung von Nukleinsäuren

Die Ermittlung der Nukleinsäurekonzentration erfolgte im NanoDrop ND-1000. Zur Messung wird 1 µl Probe auf das Podest aufgetragen und der Arm des Gerätes geschlossen. Die Oberflächenspannung hält die Probe zwischen den Oberflächen des Podestes und des Armes. Das UV-Licht (Wellenlänge 260nm) zur Messung wird mittels Glasfaserleitung zur oberen Oberfläche im Arm, von dort durch die Probe und schließlich auf einen CCD-Chip geleitet. Der Abstand der beiden Oberflächen ist definiert und ermöglicht somit durch Anwendung des Lambert-Beerschen Gesetzes

$$E = \varepsilon \cdot c \cdot d$$

- ◇ E ist die Extinktion
- ◇  $\varepsilon$  ist der molare dekadische Extinktionskoeffizient (=1/50  $\mu\text{l}/\text{ng}$  für dsDNS)
- ◇ c ist die Konzentration der Lösung
- ◇ d ist hier der Abstand der beiden Oberflächen

eine Berechnung der DNS-Konzentration. Ein Messzyklus besteht aus jeweils einer Messung bei 1,0 mm und 0.2 mm Abstand der beiden Oberflächen. Dies ermöglicht einen sehr großen Messbereich (2-3700ng/ $\mu\text{l}$ ), auf eine Verdünnung der Probe kann daher verzichtet werden (Gallagher und Desjardins, 2006).

### 3.2.4 Polymerase-Kettenreaktion (PCR)

Mit der PCR gelingt es, DNS-Fragmente spezifisch zu amplifizieren. Dabei werden Polymerasen aus thermophilen Bakterien zur Synthese der DNS eingesetzt. Diese Polymerasen bewirken die Verlängerung der Nukleotide ausgehend von den verwendeten Oligonukleotid-Primern (Saiki *et al.*, 1988).

Für Standardreaktionen wurde der PCR-Ansatz (Ausubel, 1999) mit der Taq-Polymerase von Rapidozym (Berlin, Deutschland) verwendet. Im 25  $\mu\text{l}$ -Ansatz wurden 2,0  $\mu\text{l}$  Taq-Polymerase (5 Units/ $\mu\text{l}$ ), 2,5  $\mu\text{l}$  PCR-Puffer, 2,0  $\mu\text{l}$  50mM  $\text{MgCl}_2$ , 0,2  $\mu\text{l}$  dNTPs (25mM Ausgangskonzentration) (Rapidozym, Berlin, Deutschland), 1,0  $\mu\text{l}$  je Primer (10pmol/ $\mu\text{l}$ ), 3,0  $\mu\text{l}$  Template DNS (10ng/ $\mu\text{l}$ ) mit Millipore-Wasser auf 25  $\mu\text{l}$  aufgefüllt, eingesetzt. PCR-Puffer und  $\text{MgCl}_2$  waren im Lieferumfang der Polymerase enthalten. Bei Fragmentgrößen über 5 kb (Kilobasenpaare) wurde ein kommerzieller Fertig-Mix (Extensor 2xReddyMix, Applied Biosystems) verwendet. Die Annealing-Temperatur richtet sich nach der Schmelztemperatur ( $t_{\text{melt}}$ ) der Primer/DNS-Komplexe, die durch das Verhältnis der einzelnen Basen zueinander sowie der Länge des Primers bestimmt wird. Die Elongationsdauer richtet sich nach der Größe des zu amplifizierenden DNS-Fragmentes, wobei von der Faustregel von 1000 Nukleotiden pro Minute Syntheseleistung der Polymerase ausgegangen wurde. Als Temperatur für den Elongationsschritt wurde für die Standardreaktion 72°C und für den Fertig-Mix 68°C verwendet. Als Template diente chromosomale DNS.

Nach einem zwei-minütigen Erhitzen bei 94°C (initiale Denaturierung) lief die PCR in der Regel unter den folgenden Bedingungen ab:

1. Hitzedenaturierung (30 Sekunden, 94°C bzw. 92°C bei Verwendung des Fertig-Mixes); Trennung der doppelsträngigen DNS
2. Annealing (30 Sekunden, Temperatur in Abhängigkeit der  $T_{\text{melt}}$  der verwendeten Primer); Primeranlagerung
3. Elongation (1 kb/min, 72°C bzw. 68°C bei Verwendung des Fertig-Mixes); Verlängerung der Primer an ihrem 3'-Ende mit dNTPs

Dieser Zyklus wurde insgesamt 25-mal durchlaufen, bei großen Fragmenten wurden 30 Zyklen angesetzt. Weiterhin verlängerte sich die Elongationszeit zur Erzeugung größerer PCR-Produkte beginnend mit dem elften Zyklus um jeweils zehn Sekunden je Zyklus.

Die entstandenen PCR-Produkte wurden mittels eines analytischen Agarosegels überprüft (s. 3.2.5).

#### **3.2.4.1 PCR-Analyse des putativen GimA-Insertionslokus**

In einem ersten Schritt wurde der beschriebene Insertionslokus von GimA (genomic island of meningitic *E. coli* containing *ibeA*) zwischen den Genen *yjiD* und *yjiE* in allen Stämmen mittels einer Screening-PCR untersucht. In der Screening-PCR war die Elongationszeit mit 60 Sekunden so gewählt, dass bei einer Kapazität der Polymerase von 1000 Nucleotiden pro Minute Fragmente von maximal 1,5 kb erzeugt werden könnten. Bei Isolaten, bei denen in der Screening-PCR kein Produkt erzeugt werden konnte, wurde nach Sicherstellung der Qualität der DNS zum Ausschluss von falsch-negativen Ergebnissen eine Long-Range-PCR durchgeführt. Die Long-Range-PCR ermöglichte die Darstellung des gesamten 20,3 kb umfassenden GimA. Diese sukzessive Vorgehensweise ermöglichte die Darstellung des *yjiD*-*yjiE*-Lokus in allen Isolaten.

##### **3.2.4.1.1 Screening-PCR**

Die Screening-PCR wurde unter den beschriebenen Standardbedingungen (s. 3.2.4) mit Modifikationen entsprechend der verwendeten Primer durchgeführt. Die eingesetzten Primer *yjiD*-*yjiE* FP und *yjiD*-*yjiE* RP weisen jeweils eine Schmelztemperatur von 57°C auf (s. Tab. 4). Die Annealing-Temperatur ist mit 57°C daran ausgerichtet. Die Elongationszeit von einer Minute ermöglicht entsprechend der Synthesekapazität der verwendeten Taq-Polymerase die Erzeugung von Produkten mit einer Maximalgröße von 1000-1500 Basen. Die erwartete Größe des Amplikons, die sich für *E. coli* K-12 (Blattner *et al.*, 1997) ergibt, ist 342 bp (Basenpaare). Das folgende Temperatur-Zeit-Profil wurde verwendet (s. Tab. 12):

Tab. 12 Temperatur-Zeit-Profil für die Screening-PCR. Die Schritte 2.- 4. werden insgesamt 25-mal durchlaufen

Schritt	Temperatur und Dauer
1. Prädenaturierung	94°C für drei Minuten
2. Denaturierung	94°C für 30 Sekunden
3. Annealing	57°C für 30 Sekunden
4. Elongation	72°C für 60 Sekunden
5. Endelongation	72°C für 10 Minuten

### 3.2.4.1.2 Long-Range-PCR

War trotz intakter DNS kein PCR-Produkt in der Screening-PCR darstellbar, so wurde eine Long-Range-PCR zur Prüfung auf Vollständigkeit von GimA durchgeführt. Aufgrund der Größe von GimA von über 20 kb wurde diese Untersuchung in zwei PCR-Ansätze unterteilt. Beide Ansätze verwendeten jeweils einen Primer, der außerhalb von GimA bindet und bereits in der Screening-PCR zur Anwendung kam und einen weiteren Primer, der etwa in der Mitte von GimA ansetzt. Die zentralen Primer waren so gewählt, dass sich die resultierenden PCR-Produkte in der Mitte von GimA überlappen und somit eine vollständige Darstellung von GimA ermöglichen. Es wurde der kommerzielle Fertig-Mix verwendet und die Annealing-Temperaturen entsprechend den verwendeten Primern gewählt (s. Tab. 4). Außerdem wurden entsprechend der Herstellerangaben 30 Zyklen durchlaufen und beginnend mit dem elften Zyklus die Elongationszeit je Zyklus um zehn Sekunden verlängert. Daraus ergab sich das folgende Temperatur-Zeit-Profil (s. Tab. 13):

Tab. 13 Temperatur-Zeit-Profil für die Long-Range-PCR. Die Schritte 2.- 4. werden zehn mal durchlaufen. Anschließend werden die Schritte 5.- 7. insgesamt 20-mal durchlaufen, wobei sich mit jedem Zyklus die Elongationszeit um 10 Sekunden verlängert. Die PCR zur Amplifikation des Fragmentes 1 erfolgt mit einer Annealing-Temperatur von 56°C, die Annealing-Temperatur für das Fragment 2 beträgt 60°C.

Schritt	Temperatur und Dauer
1. Prädenaturierung	92°C für drei Minuten
2. Denaturierung	92°C für 30 Sekunden
3. Annealing	56/60°C für 30 Sekunden
4. Elongation	68°C für acht Minuten
5. Denaturierung	92°C für 30 Sekunden
6. Annealing	56/60°C für 30 Sekunden
7. Elongation	68°C für acht Minuten + 10 Sekunden je Zyklus
8. Endelongation	72°C für 10 Minuten

### **3.2.4.1.3 *In silico*-Analyse von vollständig sequenzanalysierten *E. coli*-Stämmen**

Die frei verfügbaren Sequenzdaten von einigen vollständig sequenzanalysierten *E. coli*-Stämmen wurden analog der PCR-Untersuchungen analysiert. Dazu wurden zunächst die Sequenzen aus der Internetdatenbank (<http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=2&type=0&name=Complete%20Bacteria>) in die Kodon-Software eingespielt. Mithilfe der Suchfunktion konnte die Position der Bindungsstellen für Primer der Screening-PCR ermittelt und der Abstand zwischen beiden Positionen berechnet werden. Da diese *in silico*-Anwendung keinen biochemischen Grenzen unterworfen war, lieferte diese Methode die Ergebnisse der Screening- wie auch der Long-Range-PCR gleichermaßen.

### **3.2.5 Agarose-Gelelektrophorese**

Die Auftrennung von DNS-Molekülen nach ihrer Größe erfolgte mittels Agarose-Gelelektrophorese. Die Agarosekonzentration und die angelegte Spannung richteten sich nach der Größe der aufzutrennenden DNS-Fragmente.

Standardmäßig wurden 1,0 g Agarose in 100 ml 1xTBE durch Kochen gelöst. Nach Abkühlen auf ca. 50°C wurden 2,0 µl 1%-iges Ethidiumbromid zugesetzt, die Lösung in einen abgedichteten Gelträger gegossen und der Gelkamm zur Erzeugung der Taschen für die Proben eingesetzt. Sobald die Agarose erstarrt war, wurde der Kamm entfernt, das Agarosegel in eine Gelkammer mit 1xTBE als Laufpuffer gelegt und pro Tasche mit 8 µl PCR-Produkt bestückt, wobei in der Regel ein 25 µl PCR-Ansatz zuvor mit 3 µl Ladepuffer gemischt wurde. In Abhängigkeit der erwarteten Fragmentgrößen wurden 5 µl eines 100bp oder 1kb Größen-Markers in der ersten Tasche aufgetragen. Zur Erzeugung des elektrischen Felds wurde eine Spannung von 10 Volt pro Zentimeter Gellänge angelegt. Nach Abschluss der Elektrophorese wurde das Gel in einem UV-Betrachter fotografiert.

### **3.2.6 DNS-DNS-Hybridisierung**

Mit Hilfe der DNS-DNS-Hybridisierung lassen sich gleiche oder sehr ähnliche Sequenzen der zu untersuchenden Nukleinsäure und der markierten Sonden nachweisen.

Durch den Hybridisierungsvorgang kommt es zu einem Kontakt zwischen einer DNS-Sonde und der auf der Membran gebundenen denaturierten DNS, so dass die Sonde während der Inkubation an die entsprechend komplementären Bereiche bindet. Es bilden sich sogenannte „Hybrid-Moleküle“, die nach Entfernung der nicht gebundenen DNS-Sonde sichtbar gemacht werden können.

#### **3.2.6.1 Dot-Blot**

Für die Herstellung eines Dot-Blot wurden 3 µl einer 200 ng DNS/µl enthaltenden Lösung bzw. einer Lösung von hitzelysierten Bakterien auf eine positiv geladene Nylonmembran (Boehringer Mannheim GmbH, Mannheim) punktförmig aufgetropft. Nach kurzer Trocknung der Membran

wurde die DNS anschließend durch Inkubation bei 120 °C für 30 Minuten fixiert und bis zur weiteren Verwendung bei Kühlschranktemperatur aufbewahrt.

### **3.2.6.2 Sondenherstellung mittels PCR**

Die DNS-Sonden für die DNS-DNS-Hybridisierung wurden durch den Einbau von Dig-dUTP im Rahmen einer PCR markiert, dazu wurde das „PCR DIG Probe Synthesis Kit“ (Roche Diagnostics GmbH, Mannheim, Deutschland) verwendet und die PCR entsprechend den Empfehlungen des Herstellers und den Erfordernissen der verwendeten Primer durchgeführt. Mittels einer Agarosegelelektrophorese wurde festgestellt, ob die Markierung der Sonden erfolgreich durchgeführt wurde, da das Laufverhalten des markierten Amplifikates gegenüber dem nicht markierten verändert ist. Anschließend wurde die Bande aus dem Agarosegel ausgeschnitten, aufgereinigt und die DNS-Konzentration bestimmt. Bis zur Verwendung wurde die DNS-Sonde bei -20°C aufbewahrt.

### **3.2.6.3 Hybridisierungsvorgang und Visualisierung**

Zunächst wurde die Membran mit 20 ml Prähybridisierungslösung (pro 100 cm<sup>2</sup> Membran) bei 68°C rotierend im Ofen für drei Stunden inkubiert. In der Zwischenzeit wurde die Gensonde vorbereitet, dazu wurde die Sonde aufgetaut und mit 5 ml Prähybridisierungslösung gemischt und eine Endkonzentration von 20-30 ng Sonde pro ml eingestellt. Anschließend wurde das Gemisch bei 100°C für 10 Minuten im Wasserbad gekocht und danach für drei Minuten auf Eis gekühlt. Die eigentliche Hybridisierungsreaktion fand über Nacht bei 68°C rotierend im Ofen statt, dazu wurde die vorbereitete Sonde in den Kolben mit der vorbereiteten Membran gegeben, die Prähybridisierungslösung der drei-stündigen Inkubation wurde zuvor verworfen. Am Folgetag wurde die Membran je zweimal erst für fünf Minuten bei Raumtemperatur mit Waschpuffer 1 und danach für 15 Minuten bei 68°C mit Waschpuffer 2 rotierend im Ofen gewaschen.

Der Nachweis der markierten DNS erfolgte schließlich durch die Bindung eines mit alkalischer Phosphatase gekoppelten Anti-Digoxigenin-Antikörpers an der Dig-Sonde, danach dephosphoryliert die alkalische Phosphatase ein Chemolumineszent-Substrat unter Emission von Licht, das einen Röntgenfilm schwärzt.

Dazu wurde die Membran rotierend eine Minute im Ofen mit Puffer 1+Tween gewaschen und anschließend für 45 Minuten bei RT rotierend im Ofen mit Puffer 2+Tween inkubiert. Im Anschluss wurde für 30 Minuten im Ofen bei RT mit einem Gemisch bestehend aus 5 ml Puffer 2 und 1 µl Anti-Digoxigenin-AP, gefolgt von zwei Waschgängen von je 15 Minuten mit Puffer 1+Tween inkubiert. Vor der Zugabe des CSPD (Chemolumineszent-Substrat) wurde die Membran für zwei Minuten mit Puffer 3 äquilibriert und dann das Gemisch aus 1 ml Puffer 3 und 1 µl CSPD auf die Membran getropft und in einen Klarsichtbeutel eingeschweisst. Die Membranen wurden auf einen Röntgenfilm gelegt und in einer lichtdichten Filmkassette je nach

verwendeter Sonde zwischen 30 Minuten und drei Stunden exponiert. Anhand der Schwärzung der Röntgenfilme verglichen mit Positiv- und Negativkontrolle konnte die Chemilumineszenzreaktion, d.h. die Bindung der Sonde an die, auf der Membran fixierten DNS nachgewiesen werden.

### 3.2.6.4 Strippen der Membran

Um die Membranen mehrfach verwenden zu können, mussten die gebundenen Sonden-Moleküle von der Membran gelöst werden. Dazu wurde die Membran für eine Minute bei RT in 2xSSC, anschließend bei 37°C je 15 Minuten in Stripping-Puffer, der die Dissoziation der Sonden-Moleküle vermittelt und schließlich bei 37°C für zwei Minuten mit 2xSSC gewaschen. Alle Waschschrte wurden zweimal hintereinander durchgeführt. Da bei der Stripping-Prozedur auch chromosomale DNS verloren geht, war die Wiederverwendbarkeit der Membranen auf max. drei Hybridisierungen limitiert.

### 3.2.6.5 Analyse des putativen GimA-Insertionslokus mittels DNS-DNS-Hybridisierung

Die Verifizierung der Ergebnisse aus der PCR-basierten Untersuchung des GimA-Insertionslokus erfolgte mit Hilfe der DNS-DNS-Hybridisierung. Die Sonden für die Hybridisierung wurden mittels PCR generiert, die dafür verwendeten Primer sind in Tab. 4 aufgeführt. Die Elongationszeit der PCR richtete sich nach der Größe der Sonden und variierte zwischen 60 Sekunden für die Sonden *pptE*, *pdaK/cgrD*, *cniT/gclK* und *gclA* und 1,5 Minuten für die *ibeA*-Sonde.

Tab. 14 Temperatur-Zeit-Profil für die Generierung der DNS-Sonden für die DNS-DNS-Hybridisierung. Die Schritte 2.- 4. werden insgesamt 25-mal durchlaufen. Die Elongationszeit variierte bei den einzelnen Sonden. 60 Sekunden für die *pptE*-, *pdaK/cgrD*-, *cniT/gclK*- und *gclA*-Sonde und 90 Sekunden für die *ibeA*-Sonde.

Schritt	Temperatur und Dauer
1. Prädenaturierung	94°C für drei Minuten
2. Denaturierung	94°C für 30 Sekunden
3. Annealing	57°C für 30 Sekunden
4. Elongation	72°C für 60/90 Sekunden
5. Endelongation	72°C für 10 Minuten

## 3.2.7 Phylogenetische Analyse

### 3.2.7.1 Multi-Lokus-Sequenz-Typisierung (MLST)

Bei der Multi-Lokus-Sequenz-Typisierung werden sieben hoch konservierte Housekeeping-Gene (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA* und *recA*) im *E. coli*-Genom mittels einer PCR amplifiziert und dann sequenzanalysiert. Die Sequenzdaten werden über Datenbanken und das

Internet mit anderen Stämmen verglichen. Es werden einzelne Polymorphismen (SNPs – Single-Nukleotid-Polymorphismus) in den einzelnen Genen, d. h. Allele dieser Gene detektiert. Jedes Allel wird aufgrund seiner Nukleotidsequenz einem bestehenden Alleltyp zugeordnet bzw., wenn die Sequenz noch nicht bekannt ist, ein neuer definiert. Das aus den sieben einzelnen Alleltypen zusammengesetzte Allelprofil wird zu einem Sequenztyp (ST) zusammengefasst. STs, die sich in nur einem Allel unterscheiden, bilden einen ST-Komplex. Diese Vergleiche sind die Grundlage zur phylogenetischen Analyse der Isolate untereinander. So ist es möglich, phylogenetische und epidemiologische Untersuchungen an einer großen Zahl an Isolaten durchzuführen (Adiri *et al.*, 2003; Maiden *et al.*, 1998). In der vorliegenden Arbeit wurden die MLST-Daten der 410 untersuchten Stämme verwendet. Einige der STs wurden bereits veröffentlicht (Mordhorst *et al.*, 2009; Sankar *et al.*, 2009; Wirth *et al.*, 2006). Der Großteil wurde jedoch am Institut für Mikrobiologie und Tierseuchen (IMT) der Freien Universität Berlin typisiert und an die MLST-Datenbank (<http://mlst.ucc.ie/mlst>) übermittelt. Insgesamt waren 138 STs vertreten. Die verwendeten Daten umfassten die STs und ggf. ST-Komplexe, zu denen die einzelnen Stämme zugeordnet wurden (s. Tab. 15). Darüber hinaus wurden die Allelprofile, die zu der Eingruppierung in die STs führten (s. Tab. 27 im Anhang), und die Sequenzdaten der einzelnen Allele eines jeden Isolates (als Download verfügbar auf [http://mlst.ucc.ie/mlst/dbs/Ecoli/Downloads\\_HTML](http://mlst.ucc.ie/mlst/dbs/Ecoli/Downloads_HTML)) genutzt.

Tab. 15 STs (n=138) und ST-Komplexe (n=28), zu denen die 410 *E. coli*-Stämme der vorliegenden Untersuchung zugehörig sind. Im Klammern ist die Anzahl der *E. coli*-Stämme angegeben, die dem jeweiligen ST oder ST-Komplex zugeordnet sind.

Sequenztyp-Komplex (Anzahl)	Sequenztyp (Anzahl)
ST10 Komplex (n=48)	ST10 (n=36), ST34 (n=1), ST43 (n=1), ST44 (n=1), ST48 (n=6), ST49 (n=1), ST167 (n=1), ST548 (n=1)
ST11 Komplex (n=1)	ST61 (n=1)
ST12 Komplex (n=20)	ST12 (n=19), ST544 (n=1)
ST14 Komplex (n=5)	ST4 (n=2), ST81 (n=1), ST537 (n=1), ST550 (n=1)
ST23 Komplex (n=34)	ST23 (n=13), ST88 (n=15), ST356 (n=1), ST360 (n=1), ST367 (n=1), ST369 (n=2), ST410 (n=1)
ST31 Komplex (n=4)	ST130 (n=1), ST374 (n=1), ST393 (n=2)
ST38 Komplex (n=5)	ST38 (n=5)
ST46 Komplex (n=2)	ST46 (n=2)
ST59 Komplex (n=2)	ST59 (n=2)
ST69 Komplex (n=3)	ST69 (n=3)
ST73 Komplex (n=43)	ST73 (n=34), ST74 (n=1), ST104 (n=3), ST353 (n=1), ST375 (n=1), ST638 (n=2)
ST86 Komplex (n=2)	ST86 (n=1), ST453 (n=1)

<b>Sequenztyp-Komplex (Anzahl)</b>	<b>Sequenztyp (Anzahl)</b>
ST95 Komplex (n=72)	ST95 (n=56), ST140 (n=6), ST368 (n=1), ST370 (n=1), ST390 (n=4), ST416 (n=1), ST417 (n=1), ST418 (n=1), ST421 (n=1)
ST101 Komplex (n=2)	ST101 (n=1), ST365 (n=1)
ST155 Komplex (n=9)	ST55 (n=1), ST56 (n=3), ST58 (n=2), ST155 (n=3)
ST156 Komplex (n=2)	ST348 (n=2)
ST165 Komplex (n=1)	ST100 (n=1)
ST168 Komplex (n=5)	ST93 (n=4), ST373 (n=1)
ST206 Komplex.(n=1)	ST77 (n=1)
ST346 Komplex (n=1)	ST346 (n=1)
ST350 Komplex (n=5)	ST57 (n=2), ST350 (n=2), ST371 (n=1)
ST398 Komplex (n=1)	ST398 (n=1)
ST399 Komplex (n=2)	ST399 (n=1), ST536 (n=1)
ST405 Komplex (n=3)	ST72 (n=1), ST405 (n=2)
ST469 Komplex (n=2)	ST162 (n=2)
ST522 Komplex (n=1)	ST541 (n=1)
ST538 Komplex (n=1)	ST538 (n=1)
ST568 Komplex (n=3)	ST80 (n=3)
kein Komplex (n=131)	ST45 (n=1), ST47 (n=1), ST50 (n=1), ST51 (n=1), ST52 (n=1), ST53 (n=2), ST54 (n=1), ST60 (n=1), ST62 (n=10), ST63 (n=1), ST64 (n=1), ST65 (n=1), ST66 (n=1), ST67 (n=1), ST68 (n=3), ST70 (n=2), ST71 (n=1), ST75 (n=2), ST76 (n=1), ST78 (n=1), ST79 (n=1), ST82 (n=1), ST83 (n=1), ST84 (n=1), ST85 (n=1), ST87 (n=1), ST89 (n=1), ST92 (n=1), ST115 (n=3), ST117 (n=13), ST118 (n=1), ST126 (n=2), ST127 (n=11), ST131 (n=1), ST135 (n=2), ST141 (n=13), ST144 (n=2), ST216 (n=1), ST297 (n=1), ST347 (n=1), ST351 (n=1), ST352 (n=1), ST355 (n=3), ST357 (n=1), ST358 (n=1), ST359 (n=1), ST361 (n=1), ST362 (n=1), ST363 (n=1), ST372 (n=7), ST409 (n=1), ST428 (n=1), ST533 (n=1), ST539 (n=2), ST540 (n=1), ST542 (n=1), ST543 (n=1), ST545 (n=1), ST547 (n=1), ST549 (n=1), ST625 (n=1), ST646 (n=2), ST681 (n=1), ST925 (n=1), ST929 (n=1), ST937 (n=1), ST959 (n=1), ST961 (n=1)

### 3.2.7.2 Minimum-Spanning-Tree (MSTree)

Der MSTree stellt eine grafische Darstellung dar, die einzelne STs innerhalb eines MLST-Datensatzes in Verbindung zueinander stellt. Die Bionumerics Software (Version 4.6, Applied Maths, Sint Martens-Latem, Belgien) verfügt über einen Algorithmus zur Berechnung von MSTrees aus MLST-Datensätzen. Der MSTree wird durch Anwendung eines Prim's Algorithmus berechnet, der auf jeder Stufe der Konstruktion des Trees auswählt zwischen der Beibehaltung des aktuellen oder alternativen Subtrees. Basis des Netzwerkes ist der ST mit der größten Anzahl an unterschiedlichen Alleltypen. Folgend werden sukzessive weitere STs in die Berechnung mit einbezogen, abgestuft nach der Anzahl an unterschiedlichen Alleltypen und der Größe des ST, d. h. der Anzahl der Isolate in diesem ST.

ST-Komplexe werden erzeugt, wenn sich wenigstens drei STs finden, die sich in maximal einem Allel unterscheiden.

In der Grafik sind dann die quantitativen Beziehungen zwischen den einzelnen STs und ST-Komplexen, d. h. basierend auf der Zahl der gemeinsamen Allele durch Linien dargestellt. Dicke Linien zeigen, dass die verbundenen STs sechs gemeinsame Allele aufweisen. Dünne Linien zeigen fünf und schwarze gestrichelte Linien vier gemeinsame Allele an. Weniger als drei gemeinsame Allele werden durch grau gestrichelte Linien dargestellt.

Da der MSTree nicht auf den STs selbst, sondern auf den Allelprofilen der entsprechenden STs basiert, werden beispielhaft die Allelprofile der STs 95 und 73, als die häufigsten STs in der Untersuchung angeführt (Tab. 16).

Tab. 16 Allelprofile der STs 95 und 73. Angegeben sind die Alleltypen der sieben Housekeeping-Gene der MLST. Die STs 95 und 73 sind die beiden STs, die am häufigsten in der vorliegenden Untersuchung vertreten sind.

Sequenztyp	Allelprofil						
	<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>Mdh</i>	<i>purA</i>	<i>recA</i>
ST95	37	38	19	37	17	11	37
ST73	36	24	9	13	17	11	25

### 3.2.7.3 Clonal Frame-Analyse

Die Software Clonal Frame v1.1 (<http://www2.warwick.ac.uk/fac/sci/statistics/staff/research/didelot/clonalframe/>) dient zur Vorhersage bakterieller Mikroevolution. Sie kann dazu auch Sequenzdaten verwenden, die im Rahmen der MLST generiert wurden. Im Gegensatz zu phylogenetischen Methoden, die bei nicht-rekombinierenden Organismen Anwendung finden, basiert die Software auf einem Modell, das genetische Diversifizierung aufgrund von homologer Rekombination einbezieht. Damit wird eine Bestimmung der klonalen Beziehungen innerhalb einer getesteten Population einer rekombinierenden bakteriellen Spezies, wie beispielsweise *E. coli*, ermöglicht. Weiterhin ist eine Vorhersage des Ortes homologer Rekombinationen

möglich, die die Ursachen für die Auflösung der klonalen Struktur innerhalb einer Population darstellen.

Für die Verarbeitung durch Clonal Frame wurden die Sequenzen der sieben Housekeeping-Gene aus der MLST untereinander abgeglichen, auf eine einheitliche Länge getrimmt und in einem extended multi-FASTA Format (XMFA) zu jeweils sieben Blöcken zusammengefügt. Jeder Block repräsentierte dabei ein Gen.

Ergebnis der Berechnungen ist eine zuverlässige Vorhersage der phylogenetischen Beziehungen innerhalb von bakteriellen Spezies, auch wenn einige der zugrundeliegenden Housekeeping-Gene Rekombinationseffekten unterliegen.

Die Analyse wurde mit den Standardeinstellungen von 50.000 burn-in-Wiederholungen gefolgt von 50.000 MCMC (Markov Chain Monte Carlo Methode)-Durchläufen durchgeführt. Der resultierende „majority-rule consensus tree“ wurde im Anschluss durch die Software MEGA version 4 (Tamura *et al.*, 2007) grafisch dargestellt.

#### **3.2.7.4 Analyse der Populationsstruktur (Structure-Analyse)**

Für Spezies, bei denen homologe Rekombinationen vorkommen, wie bei *E. coli* der Fall, existieren spezielle phylogenetische Werkzeuge, die die Identifizierung von Gruppen und die Zuordnung von einzelnen Stämmen zu diesen Gruppen ermöglichen. Zugrunde liegt eine Analyse der Populationsstruktur. Diese wurde mit Hilfe der Software Structure v2.2 ([http://pritch.bsd.uchicago.edu/software/structure2\\_2.html](http://pritch.bsd.uchicago.edu/software/structure2_2.html)) durchgeführt. Das Programm beruht auf einer Modell-basierten Cluster-Methode (Falush *et al.*, 2003; Falush *et al.*, 2007; Pritchard *et al.*, 2000) und verwendet Genotyp-Daten (hier MLST-Datensätze) von unabhängigen Markern (hier den einzelnen Housekeeping-Genen) zur Bestimmung von Gruppen. Dazu werden die Polymorphismen in den sieben MLST-Fragmenten jedes Stammes mit Hilfe einer Methode der Wahrscheinlichkeitsrechnung (Bayesian Methode - näherungsweise nach der Markov Chain Monte Carlo Methode (MCMC) bestimmt) analysiert, mit dem Ziel, Gruppen innerhalb der Testpopulation zu finden. Basierend auf dem Hardy-Weinberg Gleichgewicht kann die Methode:

1. das Vorliegen einer Populationsstruktur zeigen
2. Gruppen innerhalb einer Population identifizieren
3. Individuen zu den identifizierten Gruppen zuordnen
4. Individuen identifizieren, die Strukturen aus verschiedenen Gruppen in sich vereinigen (diese Individuen werden dann so genannten Hybridgruppen, je nach Herkunft der einzelnen Strukturen, zugeordnet)

Teil der Structure-Analyse ist das „linkage model“, das für jeden Nukleotidpolymorphismus eine Wahrscheinlichkeit ( $q$ ) für die Abstammung aus einer bestimmten Gruppe berechnet. Aus der

Summe dieser einzelnen Wahrscheinlichkeiten kann dann auf Zugehörigkeit eines Isolates zu einer der Gruppen geschlossen werden. Diese Methode kann ebenfalls die EcoR-Phylogruppen unterscheiden, ordnet aber einen Teil der Isolate Hybridgruppen zu und trägt damit der Bedeutung der homologen Rekombination Rechnung (Wirth *et al.*, 2006).

Die Analyse wurde unter der Annahme der Existenz von vier Gruppen innerhalb der Population durchgeführt. Dies entspricht nicht nur den vier Hauptgruppen (A, B1, B2 und D), die mit der MLEE (Multi-Lokus-Enzymelektrophorese) und phylogenetischen Standardmethoden identifiziert wurden (s. 2.5.1.), sondern ebenfalls der Gruppenzahl, die Wirth *et al.* in ihrer Untersuchung als wahrscheinlichste identifiziert haben. Die statistische Sicherheit wurde über eine große Zahl an Wiederholungen erreicht (20.000 Wiederholungen) (Wirth *et al.*, 2006).

Im Rahmen der Analyse wurde die Wahrscheinlichkeit für die Zugehörigkeit eines Isolats zu den einzelnen Gruppen berechnet. Diese Wahrscheinlichkeit ( $q$ ) diente dann der Zuordnung der Isolate zu den Gruppen (A, B1, B2 und D). Der Grenzwert für eine Zuordnung lag bei  $q \geq 0.67$ . Isolate, die keiner Gruppe zugeordnet werden konnten, wurden als Hybridgruppe AxB1 angesprochen, wenn die Werte für  $q$  für die Zugehörigkeit zu A und B1 zusammen  $\geq 0,8$  ist. Alle verbleibenden Isolate wurden zu einer zweiten Hybridgruppe (ABD) zusammengefasst.

### **3.2.8 Sequenzanalyse von *ibeA* und Computergestützte Analyse**

#### **3.2.8.1 Sequenzanalyse von *ibeA***

Die DNS-Sequenzanalysen wurden von der Firma LGC (Berlin, Deutschland) mittels der Kettenabbruchmethode nach F. Sanger durchgeführt. Bei dieser automatisierten Methode dienen synthetische Oligonukleotide als Primer einer *in vitro*-DNS-Synthese. Dabei stellt der zu sequenzierende Strang die Matrize dar. Das synthetisierte DNS-Molekül wird mit einem fluoreszierenden Farbstoff markiert. Zu einem statistisch möglichst gleichmäßigen, basenspezifischen Kettenabbruch kommt es durch Einbau eines 2',3'-Didesoxynukleosidtriphosphates (ddNTPs). Diesen Nukleotiden fehlt die zur weiteren Kettenverlängerung notwendige 3'-Hydroxylgruppe. So werden vier verschiedene Reaktionen benutzt, die jeweils eines der vier ddNTPs (ddATP, ddCTP, ddGTP oder ddTTP) neben den üblichen dNTPs enthalten. Das statistische Auftreten der Kettenabbruch-Reaktionen an der entsprechenden Position der DNS (ddTTP gegenüber A; ddCTP gegenüber G und so weiter) führt zu unterschiedlich großen DNS-Fragmenten. Durch die Größenaufftrennung in einer denaturierenden Polyacrylamid-Gelelektrophorese und Sichtbarmachen der einzelnen Banden durch Fluorographie werden die Positionen des Kettenabbruchs sichtbar gemacht. Das Ablesen dieser Sequenzleitern führt zu den entsprechenden Sequenzinformationen.

Für die Sequenzanalyse von *ibeA* (invasion of brain endothelium A) wurden die in Tab. 4 aufgeführten Sequenzierungs-Primer verwendet.

### 3.2.8.2 Single-Lokus-Sequenz-Typisierung (SLST)

Die Sequenzanalyse eines einzelnen Gen-Lokus funktioniert prinzipiell wie die MLST, mit dem Unterschied, dass der betrachtete Locus weniger konserviert ist als die Housekeeping-Gene der MLST. Mit der SLST eines Genes mit einer höheren Variabilität als die Housekeeping-Gene ist es möglich, Stämme eines gemeinsamen phylogenetischen Hintergrundes zu diskriminieren. Für die SLST wurde das *ibeA*-Gen ausgewählt, da es in die Pathogenese der Neugeborenenmeningitis involviert ist. Wie an den genannten Modellen der Blut-Hirn-Schranke (s. 2.3.2.1) gezeigt werden konnte, ist *ibeA* an der Translokation dieser Barriere beteiligt (*in vitro* und *in vivo*). Folglich findet eine Interaktion mit dem Wirt statt. Diese Interaktion könnte zu einer Immunantwort des Wirtes führen, die wiederum neue Varianten bevorteilt und zu einer größeren Variabilität im Vergleich zu Housekeeping-Genen, im Sinne einer Pathoadaptation (wie von O-Antigenen und dem *fim*-Operon) führt. Weiterhin wurde der genetische Zusammenhang, in dem sich *ibeA* befindet, GimA, als genomische Insel beschrieben. Genomische Inseln sind mobile genetische Elemente und können per horizontalem Gentransfer zwischen Bakterien übertragen werden. Aufgrund dieser Mobilität ist zu vermuten, dass in einem einheitlichen phylogenetischen Hintergrund mehrere Varianten, d. h. Sequenztypen der Insel (SLSTs) auftreten. Eine entsprechende genetische Diversifizierung wurde bereits für O- und H-Antigene (Reeves, 1992) sowie das *fim*-Operon (Weissman *et al.*, 2006) beschrieben.

Für die Sequenzanalyse von *ibeA* wurden insgesamt 34 *E. coli*-Stämme ausgewählt. Sie repräsentieren den MLST-basierten phylogenetischen Hintergrund, in dem *ibeA* vorkommt, d. h. die Phylogruppe B2. Da das *ibeA*-Gen mit 1.341 bp zu groß ist, um es mit nur einem PCR-Ansatz zu amplifizieren und zu sequenzanalysieren, wurde es in zwei überlappende Fragmente geteilt. Beide Fragmente wurden mittels eines Primers, der außerhalb von *ibeA* bindet und einem Primer, der in der Mitte des Gens bindet, analysiert (s. Tab. 4). Mit Ausnahme der Annealing-Temperatur von 55°C und einer Elongationszeit von 60 s werden die Bedingungen wie unter 3.2.4 beschrieben, verwendet (s. Tab. 17). Die Sequenzanalyse wurde von LGC durchgeführt. Die Auswertung der Sequenzen erfolgte mit Hilfe des „assembler module“ der Kodon-Software (Version 3.6 Applied Maths, Sint Martens-Latem, Belgien). Das Alignment der Sequenzen sowie die Berechnung des Dendrogramms nach dem Maximum-Parsimony-Algorithmus wurde ebenfalls mit Kodon durchgeführt. Eine statistische Absicherung des Dendrogramms erfolgte durch eine Bootstrap-Analyse mit 1.000 Wiederholungen (Kodon-Software).

Tab. 17 Temperatur-Zeit-Profil für die PCR zur Generierung der Produkte für die Sequenzanalyse. Die Schritte 2. - 4. werden insgesamt 25-mal durchlaufen

Schritt	Temperatur und Dauer
1. Prädenaturierung	94°C für drei Minuten
2. Denaturierung	94°C für 30 Sekunden
3. Annealing	55°C für 30 Sekunden
4. Elongation	72°C für 60 Sekunden
5. Endelongation	72°C für 10 Minuten

### 3.2.8.3 Evolutionäre Analyse von *ibeA*

Für die evolutionäre Analyse von *ibeA* wurden die *ibeA*-Sequenzen sowie die MLST-Sequenzdaten von 21 Stämmen verwendet, die dem ST95-Komplex zugeordnet sind. Die Sequenzdaten der einzelnen Housekeeping-Gene der MLST wurden für die evolutionäre Analyse für jeden Stamm zu einer einzigen Sequenz „in frame“ zusammengefügt.

Die Detektion von synonymen und nicht-synonymen Mutationen sowie die Berechnung der entsprechenden Mutationsraten  $D_n$  (nicht-synonymen Mutationsrate) und  $D_s$  (synonymen Mutationsrate) wurde anschließend mit der DnaSP 4.50.3-Software (<http://www.ub.es/dnasp/>) (Rozas *et al.*, 2003) durchgeführt. Die  $D_n/D_s$ -Verhältnisse wurden für jeden einzelnen Stamm berechnet und als Durchschnitt der Testpopulation ( $n=23$ ) angegeben.

Zur Bestätigung wurden die so ermittelten  $D_n/D_s$ -Werte mittels einer weiteren Methode zur Detektion positiver Selektion, der SLAC-Analyse (single likelihood ancestor counting) überprüft. Die SLAC-Analyse wurde auf einer webbasierten Plattform zur Detektion von positiver Selektion durchgeführt (<http://www.datamonkey.org>) (Pond und Frost, 2005).

### 3.2.9 Statistik

Die in dieser Arbeit verwendeten statistischen Analysen wurden, sofern sie nicht bereits in die entsprechende Software implementiert sind, mit der SPSS Software (Version 15, SPSS Inc., Chicago, IL, USA) durchgeführt.

## 4 ERGEBNISSE

### 4.1 Nachweis von GimA

Zunächst wurde der GimA (genomic island of meningitic *E. coli* containing *ibeA*)-Insertionslokus aller 410 Isolate mittels einer PCR (Polymerase-Kettenreaktion)-Anwendung untersucht. Anschließend wurden die Ergebnisse aus der PCR-Untersuchung durch eine DNS-DNS-Hybridisierung verifiziert.

#### 4.1.1 PCR-Analyse des putativen GimA-Insertionslokus

Der Bereich zwischen den Genen *yjiD* und *yjiE* wurde in den 410 *E. coli*-Isolate der vorliegenden Untersuchung mittels der zwei PCR-Anwendungen (Screening- und Long Range-PCR) analysiert. Im Stamm RS218 befindet sich GimA zwischen diesen beiden Genen (Huang *et al.*, 2001a). Es konnten drei verschiedene Größenvarianten dieses Lokus identifiziert werden (s. u.).

##### 4.1.1.1 Screening-PCR

In der Screening PCR lassen sich in Abhängigkeit der Reaktionsbedingungen Amplifikate mit einer Maximallänge von 1.000-1.500 bp (Basenpaare) erzeugen. Bei 344 der 410 Isolate konnte ein Produkt erzeugt werden, bei den übrigen 66 Isolaten war nach der Elektrophorese kein Amplifikat in Form einer Bande sichtbar. Diese 66 Isolate wurden anschließend mittels der Long-Range-PCR weiter untersucht.

Bei 264 (64,4%) der positiven Isolate betrug die Amplikongröße 340 bp und entspricht damit der für *E. coli* K-12 berechneten Größe (s. Abb. 7). Die verbleibenden 80 Isolate wiesen Produkte mit 742 bp Größe auf (19,5%) (s. Abb. 8). Abb. 6 zeigt beispielhaft ein Elektrophogramm von PCR-Produkten der Screening-PCR.

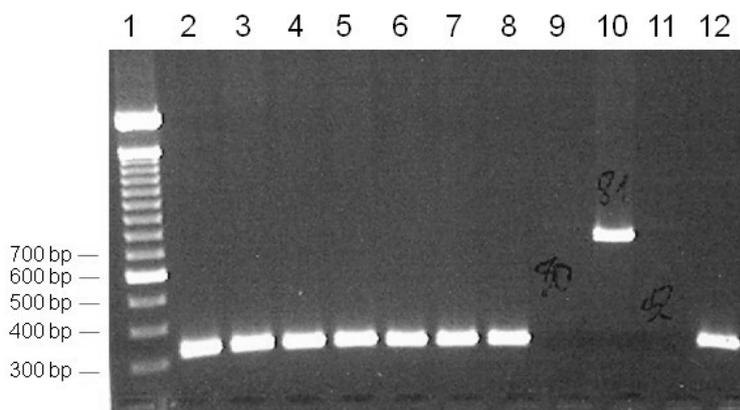


Abb. 6 Elektrophogramm zur Analyse des Bereichs zwischen den Genen *yjiD* und *yjiE* in den *E. coli*-Stämmen (1,5%-iges Agarosegel, 1 h bei 120V). In den Spuren 2-12 sind

PCR-Produkte der Screening-PCR aufgetragen. In Spur 1 ist ein 100bp-Marker aufgetragen. Die Spuren 2-8 sowie die Spur 12 haben eine Bande bei etwa 340 bp. In Spur 10 zeigt sich eine Bande von etwa 740 bp Größe. Die Spuren 9 und 11 haben keine Bande. Diese Proben werden in der Long Range PCR weiteruntersucht.

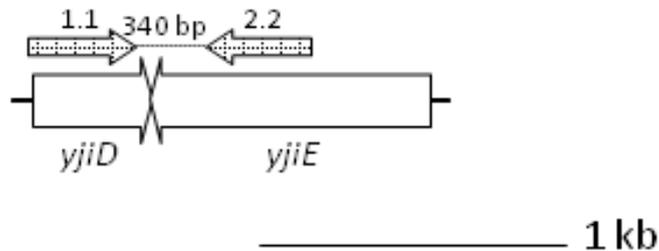


Abb. 7 Bindungsstellen der Primer in der Screening-PCR (Primer-Nummer in der Grafik und Primer-Bezeichnung in Tab. 4: 1.1 = *yjiD*-*yjiE* FP, 2.2 = *yjiD*-*yjiE* RP). Es ergibt sich eine Amplikongröße von 340 bp. Die Abbildung basiert auf der Sequenz von *E. coli* K-12 (Blattner *et al.*, 1997)

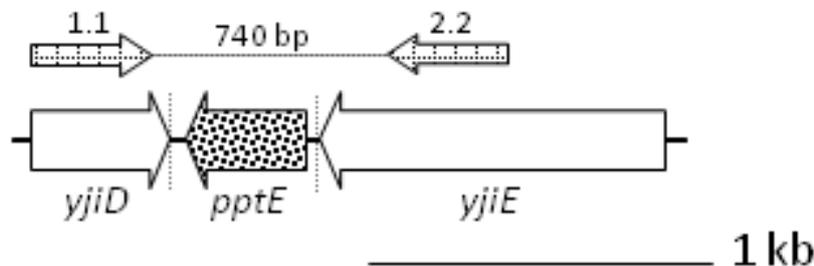


Abb. 8 Bindungsstellen der Primer in der Screening-PCR (Primer-Nummer in der Grafik und Primer-Bezeichnung in Tab. 4: 1.1 = *yjiD*-*yjiE* FP, 2.2 = *yjiD*-*yjiE* RP). Die sich ergebende Größe des Amplikons beträgt 742 bp. Die Abbildung basiert auf den Sequenzen der *E. coli*-Stämme CFT073 (Welch *et al.*, 2002), F11 (Rasko *et al.*, 2008) und 536 (Hochhut *et al.*, 2006)

#### 4.1.1.2 Long-Range-PCR

Die 66 Isolate, mit denen in der Screening-PCR kein Amplifikat zu erzeugen war, wurden mit der Long-Range-PCR weiter untersucht. Basierend auf der Annahme, dass diese Isolate ein vollständiges *GimA* von 20,3 kb (Kilobasenpaare) Größe enthalten, wurden zwei Long-Range-PCRs parallel durchgeführt. Entsprechend der Bindungsstellen der Primer wurden zwei sich überlappende Fragmente von 10,5 kb Größe erzeugt, wenn sich eine vollständige Insel im Insertionslokus befindet (s. Abb. 10). Für alle 66 Isolate (16,1%) waren in beiden Long-Range-

PCRs 10,5 kb große Banden darstellbar. Die Isolate enthalten zwischen den Genen *yjiD* und *yjiE* das vollständige, aus vier Operons (GimA1-4) bestehende GimA. Abb. 9 zeigt beispielhaft ein Elektrophogramm von PCR-Produkten der beiden Long-Range-PCRs.

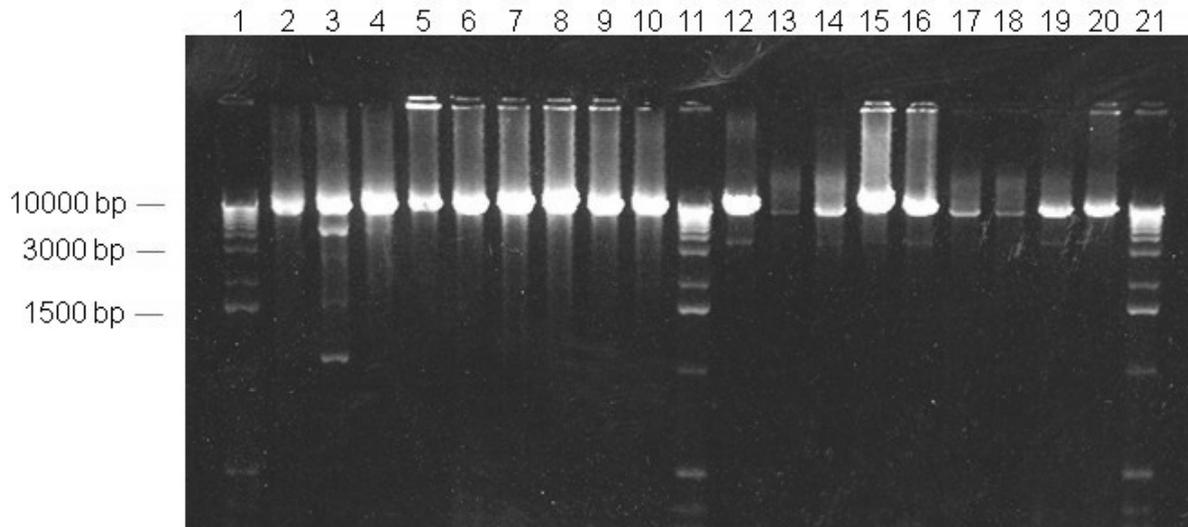


Abb. 9 Elektrophogramm zur Darstellung der Ergebnisse der Analyse des Bereichs zwischen den Genen *yjiD* und *yjiE* (1,0%-iges Agarosegel, 2h bei 120V). In den Spuren 2-10 und 12-20 sind PCR-Produkte der Long-Range-PCR aufgetragen. In Spuren 1, 11 und 21 ist ein 1 kb-Marker aufgetragen. In den Spuren 2-10 sind Produkte der PCR mit den Primern für das Long-Range-Fragment 1 und in den Spuren 12-20 sind Produkte der PCR für das Long-Range-Fragment 2 aufgetragen. Alle Produkte weisen Banden bei etwa 10,5 kb auf.

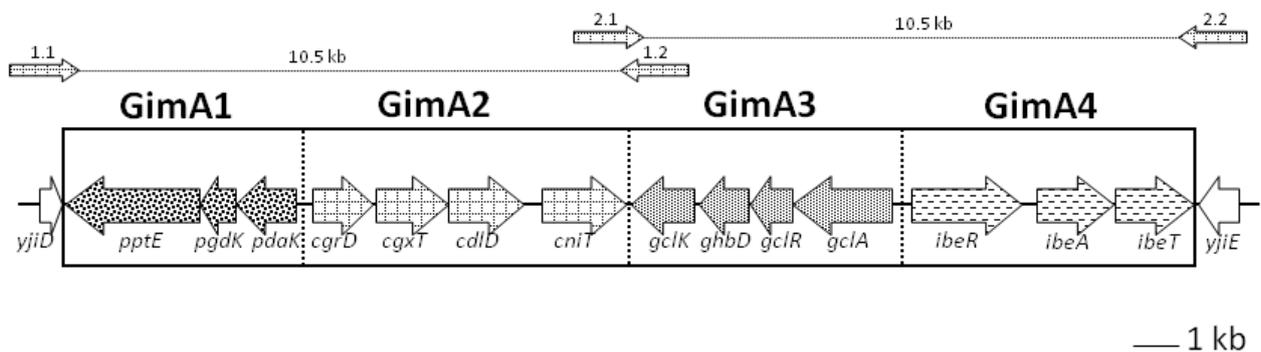


Abb. 10 Bindungsstellen der Primer in der Long Range PCR (Primer-Nummer in der Grafik und Primer-Bezeichnung in Tab. 4: 1.1 = *yjiD*-*yjiE* FP, 1.2 = *gclK*GimA3 RP, 2.1 = *cniT*GimA2 FP, 2.2 = *yjiD*-*yjiE* RP). Es ergeben sich zwei ca. 10,5 kb große Produkte, die Long-Range-Fragmente 1 und 2. Die Abbildung basiert auf den Sequenzen der *E. coli*-Stämme RS218, A0 34/86, UTI89 und APEC O1 (Chen *et al.*, 2006; Hejnova *et al.*, 2005; Huang *et al.*, 2001a; Johnson *et al.*, 2007)

#### 4.1.1.3 *In silico*-Analyse von vollständig sequenzanalysierten *E. coli*-Stämmen

Analog zu den PCR-Untersuchungen wurden die frei verfügbaren Sequenzdaten vollständig sequenzanalysierter *E. coli*-Stämmen untersucht. Es wurden die gleichen (virtuellen) Produktgrößen wie in den beiden PCR-Ansätzen gefunden. Die *E. coli*-Stämme RS218 (NMEC - Neugeborenenmeningitis-assoziierte *E. coli*), APEC O1 (APEC – aviär pathogene *E. coli*) und UTI89 (UPEC – uropathogene *E. coli*) wiesen ein vollständiges GimA auf (s. Abb. 10), während die UPEC-Stämme CFT073, F11 und 536 einen Abstand von 742 bp (s. Abb. 8) und der K-12 Stamm MG1655 einen Abstand von 340 bp zwischen den Bindungsstellen der Primer yjiD-yjiE FP und yjiD-yjiE RP aufwiesen (s. Abb. 7) (s. Tab. 4).

Weiterhin konnte im Rahmen dieser *in silico*-Analyse das 742 bp große Fragment eingehender untersucht werden. Es zeigte sich, dass innerhalb der aus dem *E. coli* K-12 bekannten Sequenz (340 bp) zusätzlich 342 Nucleotide vorkommen. Ein Sequenzvergleich zwischen Stämmen mit vollständigem GimA (APEC O1, RS218 und UTI89) und solchen mit dem 742 bp Amplikon (CFT073, F11 und 536) ergab eine fast 100%-ige Übereinstimmung dieser zusätzlichen Sequenz mit den letzten 342 bp des *pptE*-Gens (insgesamt 2.832 bp) (s. Abb. 12). Das *pptE*-Gen ist das erste Gen im ersten Operon (GimA1) von GimA (s. Abb. 2).

Die drei Vertreter des ST95 (RS218, APEC O1 und UTI89) wiesen sogar eine vollständige (100% Identität) Übereinstimmung in den Sequenzen der *pptE*-Fragmente auf.

Weiterhin war es dank der vollständig sequenzanalysierten Stämme möglich, mit Hilfe der Kodon-Software eine grafische Darstellung der Verhältnisse des Insertionslokus in den drei gefundenen Varianten zu erzeugen (s. Abb. 11).

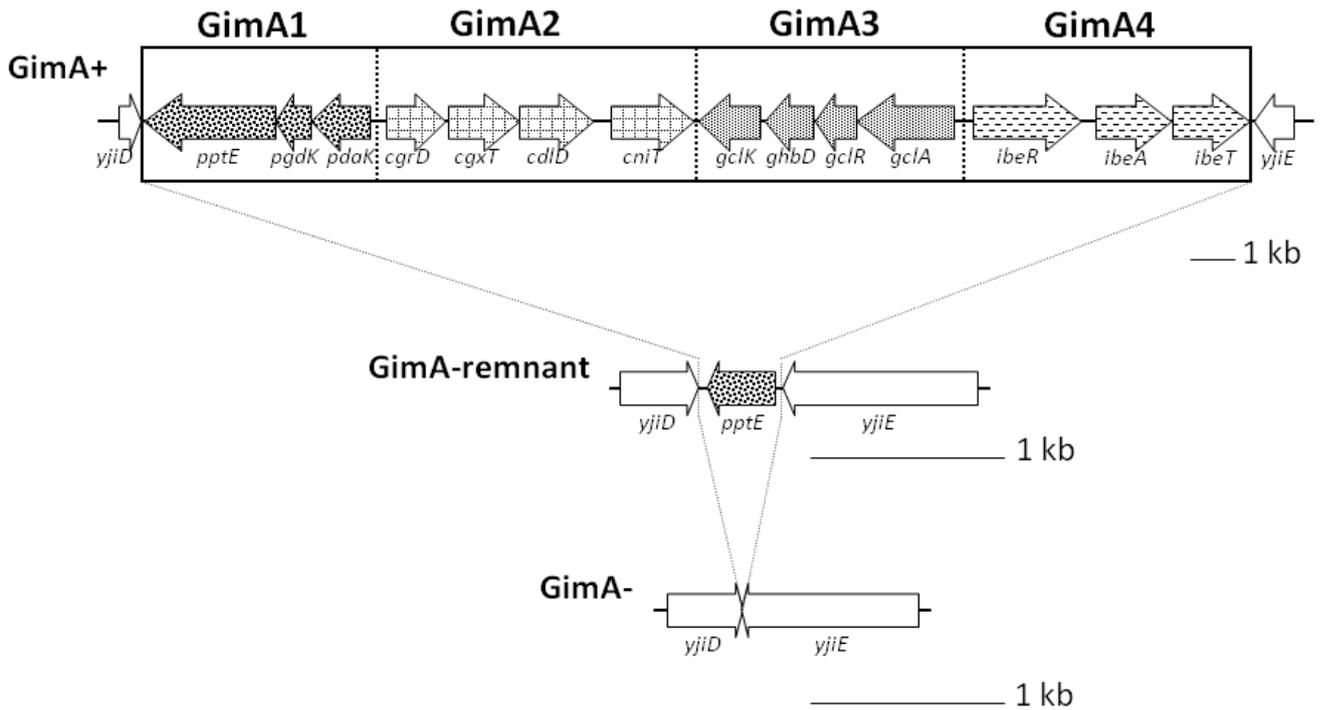


Abb. 11 Drei Varianten des GimA-Insertionslokus. GimA+: enthält das vollständige GimA, bestehend aus den Operons GimA1-4; GimA-remnant: enthält einen 342 bp großen 3' Rest des *pptE*-Gens; sowie GimA-, das keine GimA-Strukturen aufweist. Die Darstellung basiert für GimA+ auf den Sequenzen der *E. coli*-Stämme RS218, A0 34/86, UT189 und APEC O1, für GimA-remnant auf den Stämmen CFT073, F11 und 536, sowie für GimA- auf der Sequenz des *E. coli* K-12-Stamms MG1655. Die Abbildungen der Varianten GimA-remnant und GimA- sind zur besseren Übersicht in einem größeren Maßstab dargestellt als die Variante GimA+.

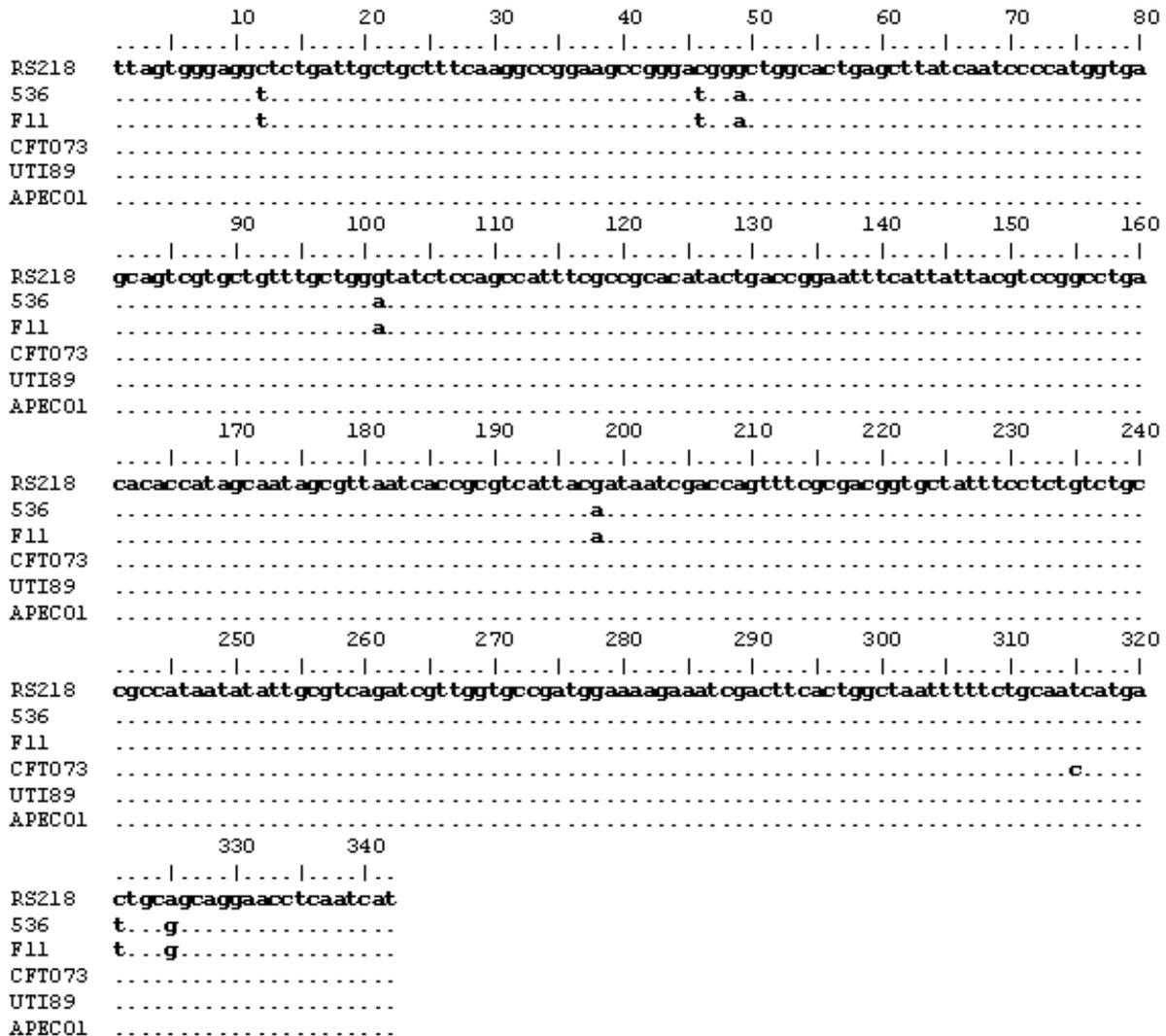


Abb. 12 Sequenzvergleich des 3'-Endes des *pptE*-Gens der Stämme RS218 (ST95) (Huang *et al.*, 2001a), UTI89 (ST95) (Chen *et al.*, 2006) und APEC O1 (ST95) (Johnson *et al.*, 2007) mit den zwischen *yjiD* und *yjiE* befindlichen Sequenzen der Stämme 536 (ST92) (Hochhut *et al.*, 2006), F11 (ST127) (Rasko *et al.*, 2008) und CFT073 (ST73) (Welch *et al.*, 2002). Übereinstimmungen mit der Template-Sequenz RS218 sind durch Punkte und Unterschiede durch Nennung der Base gekennzeichnet.

**4.1.2 DNS-DNS-Hybridisierung**

Zur Verifizierung der mit der PCR (Screening- und Long-Range) erzielten Ergebnisse wurden alle 410 *E. coli*-Isolate zusätzlich mit Hilfe der DNS-DNS-Hybridisierung auf das Vorkommen von GimA (genomic island of meningitic *E. coli* containing *ibeA*)-spezifischen Sequenzen untersucht. Die Hybridisierungsreaktion mit den DNS-Sonden *pdaK/cgrD*, *cniT/gclK*, *gclA*, und *ibeA* war jeweils bei 66 Isolaten positiv. Die *pptE*-Sonde konnte spezifisch in der DNS von 146 Isolaten binden (s. Abb. 15).

Ein Vergleich mit den Ergebnissen der PCR-Untersuchungen ergab eine 100%-ige Übereinstimmung. In den 66 Isolaten, die laut Long-Range-PCR ein vollständiges GimA

enthielten, konnten auch die vier GimA-spezifischen Sonden sowie die *pptE*-Sonde binden. Die *pptE*-Sonde konnte in 80 weiteren Fällen in den DNS-Proben binden. Diese Isolate wiesen in der Screening-PCR das 742 bp Amplikon auf. Mittels der DNS-DNS-Hybridisierung konnte gezeigt werden, dass diese Isolate zusätzlich das Restfragment des *pptE*-Gens enthalten, wie in der *in silico*-Analyse für die Stämme CFT073, F11 und 536 beschrieben.

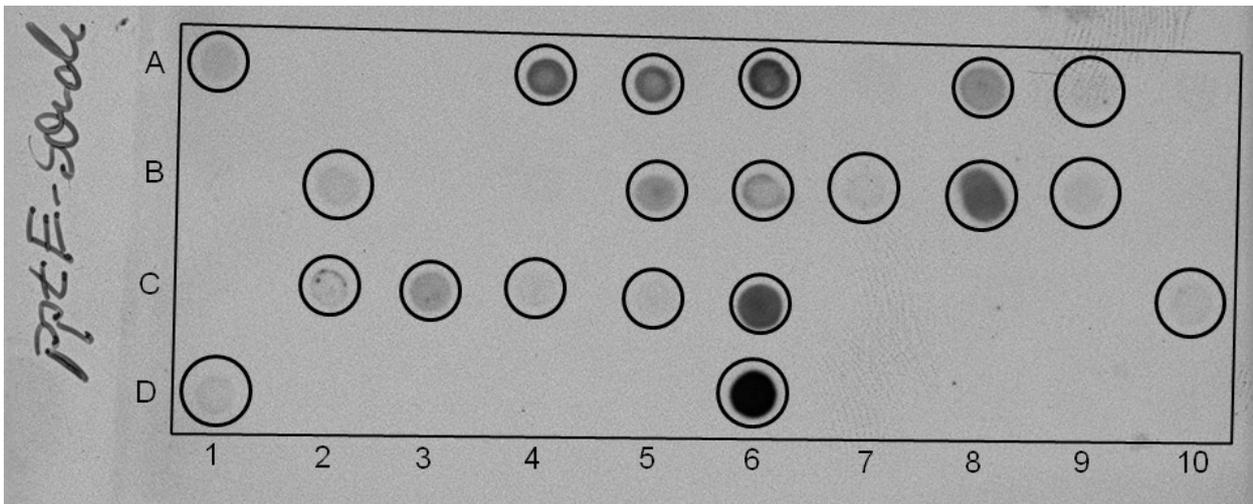


Abb. 13 Chemoluminogramm. DNS-DNS-Hybridisierung mit der *pptE*-Sonde zum Nachweis des *pptE*-Restfragmentes in der aufgetropften DNS der zu untersuchenden Isolate. Die Kreise markieren die Stellen, an denen die Sonde mit der aufgetropften und fest an die Membran gebundenen DNS hybridisieren konnte. Eine Chemolumineszenzreaktion an der markierten *pptE*-Sonde belichtet einen aufgelegten Röntgenfilm. Insgesamt ist die DNS von 35 Isolaten aufgetropft. Die drei oberen Reihen (A-C) enthalten die DNS von zehn Isolaten (1-10). Die unterste, vierte Reihe (D) besteht, links beginnend, aus der aufgetropften DNS von fünf Isolaten (D 1-5) und der fünf Positivkontrollen für die Sonden (Abfolge: *pptE*, *pdaK/cgrD*, *cniT/gclK*, *gclA*, und *ibeA*) (D 6-10). Die hier eingesetzte DNS-Sonde detektiert 19 *pptE*-haltige Isolate und bindet spezifisch an die *pptE*-Positivkontrolle

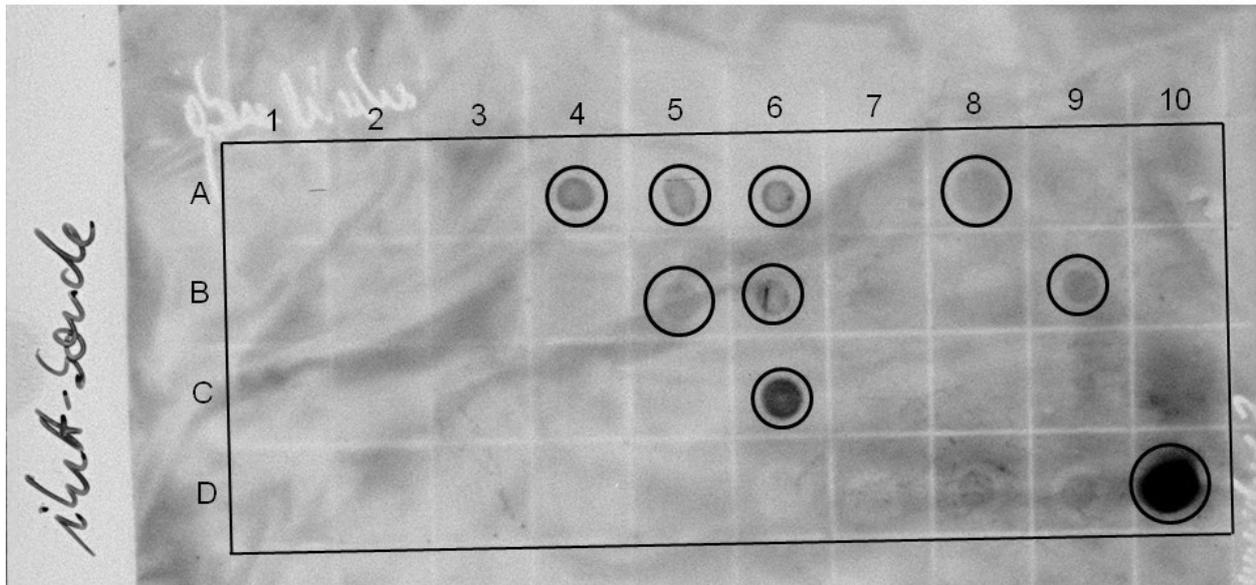


Abb. 14 Chemoluminogramm. DNS-DNS-Hybridisierung mit der *ibeA*-Sonde zum Nachweis des *ibeA*-Gens in der aufgetropften DNS der zu untersuchenden Isolate. Die Kreise markieren die Stellen, an denen die Sonde mit der aufgetropften und fest an die Membran gebundenen DNS hybridisieren konnte. Eine Chemolumineszenzreaktion an der markierten Sonde belichtet einen aufgelegten Röntgenfilm. Insgesamt ist die DNS von 35 Isolaten aufgetropft. Die drei oberen Reihen (A, B und C) enthalten die DNS von zehn Isolaten (1-10). Die unterste, vierte Reihe (D) besteht links beginnend aus der aufgetropften DNS von fünf Isolaten (D 1-5) und der fünf Positivkontrollen für die Sonden (Abfolge: *pptE*, *pdaK/cgrD*, *cniT/gclK*, *gclA*, und *ibeA*) (D 6-10). Aufgetragen sind die gleichen Proben wie in Abb. 13. Es werden jedoch nur acht Isolate als *ibeA*-positiv identifiziert. Die *ibeA*-Positivkontrolle wird spezifisch durch die *ibeA*-Sonde erkannt.

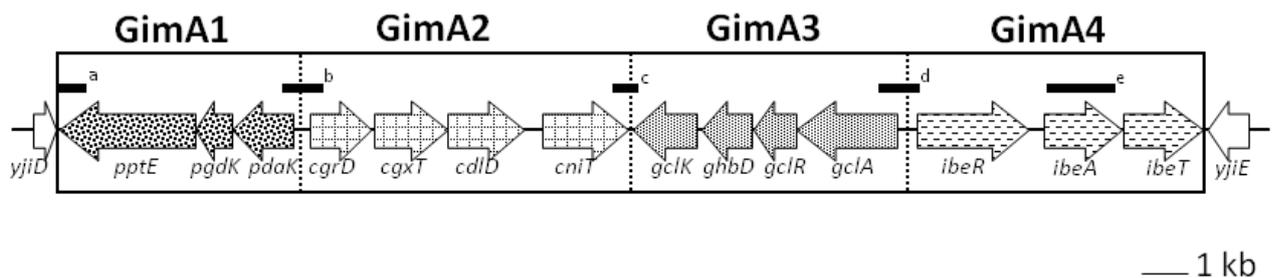


Abb. 15 Bindungsstellen der DNS-Sonden für die DNS-DNS-Hybridisierung (a: *pptE*, b: *pdaK/cgrD*, c: *cniT/gclK*, d: *gclA* und e: *ibeA*) Die Abbildung basiert auf den Sequenzen der *E. coli*-Stämme RS218, A0 34/86, UTI89 und APEC O1 (Chen *et al.*, 2006; Hejnova *et al.*, 2005; Huang *et al.*, 2001a; Johnson *et al.*, 2007)

#### 4.1.3 Zusammenfassung der Untersuchung zum Nachweis von GimA

Aus den Ergebnissen der PCR- und DNS-DNS-Hybridisierungsuntersuchungen ergibt sich folgendes Bild:

Es konnten drei (Größen-) Varianten des GimA Insertionslokus identifiziert werden. In der Folge werden die drei Varianten als GimA<sup>+</sup> für das Vorliegen eines vollständigen GimA (20,3 kb); als

GimA-remnant für den Rest des *pptE*-Gens (742 bp) und als GimA- für die Verhältnisse, wie in *E. coli* K-12, bezeichnet (340 bp). In den Klammern sind jeweils die Abstände zwischen den Bindungsstellen der Primer *yjiD*-*yjiE* FP und *yjiD*-*yjiE* RP angegeben (s. Abb. 11). Für die Variante GimA-remnant ergibt sich folglich im Vergleich mit der Sequenz des *E. coli* K-12-Stamms MG1655, dass 342 zusätzliche Basen enthalten sind.

## 4.2 Verteilung der Varianten des GimA-Insertionslokus innerhalb der *E. coli*-Population

### 4.2.1 Multi-Lokus-Sequenztypisierung (MLST)

Die verwendeten 410 Isolate sind vollständig mittels der MLST-typisiert. Unter Verwendung dieser Daten wurde das Vorkommen der Insertionslokusvarianten in Abhängigkeit von ST (Sequenztyp) und ST-Komplex untersucht.

Da die Mehrheit der STs nur mit einem Isolat in der Untersuchung vertreten waren, konnte eine fundierte Aussage zur Prävalenz der einzelnen Varianten des Insertionslokus in den einzelnen STs nicht gemacht werden. Weiterhin ergab sich für STs, die mehrfach in der Untersuchung vorkommen, ein sehr heterogenes Bild. Neben STs, die nur eine Lokusvariante enthielten [homogene STs: nur GimA-: ST10 (n=36), ST48 (n=6), ST23 (n=13), ST88 (n=15) und ST62 (n=10), nur Gim+: ST372 (n=7)] kommen ebenso viele STs vor, die zwei oder drei Varianten aufweisen [heterogene STs: ST12 (n=19), ST73 (n=34), ST95 (n=56), ST117 (n=13), ST127 (n=11), ST140 (n=6) und ST141 (n=13)]. Bezogen auf die ST-Komplexe, die mehr als fünf Isolate enthielten, ergab sich ein ähnliches Bild. Zwei ST-Komplexen, die lediglich eine Variante aufwiesen [nur GimA-: ST-Komplexe 10 (n=48) und 23 (n=34)] standen drei Komplexe gegenüber, die mehr als eine Variante enthielten [ST-Komplexe 12 (n=20), 73 (n=42) und 95 (n=72)]. Ein Gutteil der in der Untersuchung vertretenen STs (n=68) war keinem ST-Komplex zugeordnet. Für die homogenen STs und ST-Komplexe betragen die Prävalenzen 100%. Signifikante Assoziationen zwischen ST bzw. ST-Komplex und den Varianten des Insertionslokus waren nicht nachweisbar. Die verbleibenden Komplexe repräsentieren nur zwischen einem und fünf Stämme und sind daher für Aussagen zur Prävalenz kaum geeignet.

In den folgenden Tabellen zur Verteilung werden aus Gründen der Übersichtlichkeit lediglich die STs und die ST-Komplexe aufgeführt. Weiterhin werden Angaben zu den Häufigkeiten der STs und ST-Komplexe gemacht. Die zugrundeliegenden Isolate mit den entsprechenden STs und ST-Komplexen sind im Anhang aufgeführt (s. Tab. 27 im Anhang).

Die Verteilung der einzelnen Varianten des Insertionslokus auf die STs und ST-Komplexe stellte sich wie folgt dar.

#### 4.2.1.1 GimA+

Wie in der Untersuchung des Insertionslokus festgestellt, enthielten 66 der insgesamt 410 getesteten Isolate das vollständige GimA (16,1%). Bezogen auf die Daten aus der MLST verteilen sich die GimA+-Isolate auf insgesamt 26 STs. Davon sind 10 STs, die 34 Isolate enthalten, vier ST-Komplexen zugeordnet.

Tab. 18 ST-Komplexe (n=4), in denen Isolate mit einem vollständigen GimA (GimA+) identifiziert werden konnten (n=34). Weiterhin sind die STs (n=10) angegeben, die innerhalb der einzelnen ST-Komplexe die Variante GimA+ enthalten. In Klammern sind die jeweiligen Häufigkeiten angegeben.

Lokusvariante: GimA+	
Sequenztyp-Komplex	enthaltene STs in der Untersuchung
STC95 (n=29)	ST95 (n= 16), ST140 (n= 4), ST370 (n= 1), ST390 (n= 4), ST416 (n= 1), ST417(n= 1), ST418 (n= 1)
STC568 (n=3)	ST80 (n=3)
STC538 (n=1)	ST538 (n=1)
STC14 (n=1)	ST537 (n=1)

Die verbleibenden 32 Isolate setzten sich aus 16 STs zusammen, die bisher keinem ST-Komplex zugeordnet sind.

Tab. 19 STs (n=16), in denen Isolate mit einem vollständigen GimA (GimA+) identifiziert werden konnten (n=32) und die derzeit keinem ST-Komplex zugeordnet sind. In Klammern sind die jeweiligen Häufigkeiten angegeben.

Lokusvariante: GimA+		
STs, die keinem Komplex zugeordnet sind		
ST83 (n=1)	ST352 (n=1)	ST372 (n=7)
ST126 (n=1)	ST355 (n=3)	ST428 (n=1)
ST135 (n=2)	ST357 (n=1)	ST547 (n=1)
ST141 (n=6)	ST358 (n=1)	ST625 (n=1)
ST144 (n=2)	ST363 (n=1)	ST646 (n=2)
ST929 (n=1)		

#### 4.2.1.2 GimA-remnant

Von den 410 getesteten Stämmen wurden insgesamt 80 in der PCR und DNS-DNS-Hybridisierung positiv auf das Restfragment des *pptE*-Gens getestet. Diese, das GimA-remnant tragenden Isolate, verteilen sich auf 16 STs, wovon sieben (n=47 Stämme) zwei ST-Komplexen zugeordnet sind.

Tab. 20 ST-Komplexe (n=2), in denen Isolate mit einem GimA-remnant identifiziert werden konnten (n=47). Weiterhin sind die STs (n=7) angegeben, die innerhalb der beiden

ST-Komplexe die Variante GimA-remnant enthalten. In Klammern sind die jeweiligen Häufigkeiten angegeben.

<b>Lokusvariante: GimA-remnant</b>	
<b>Sequenztyp-Komplex</b>	<b>enthaltene STs in der Untersuchung</b>
STC73 (n=35)	ST73 (n=28), ST74 (n=1), ST104 (n=3), ST353 (n=1), ST375 (n=1), ST638 (n=1)
STC12 (n=12)	ST12 (n=12)

Die restlichen 33 Isolate verteilen sich auf neun STs, die aktuell keinem ST-Komplex zugeordnet sind.

Tab. 21 STs (n=9), in denen Isolate mit einem GimA-remnant identifiziert werden konnten (n=33) und die derzeit keinem ST-Komplex zugeordnet sind. In Klammern sind die jeweiligen Häufigkeiten angegeben.

<b>Lokusvariante: GimA-remnant</b>		
<b>STs, die keinem Komplex zugeordnet sind</b>		
ST76 (n=1)	ST117 (n=12)	ST681 (n=1)
ST82 (n=1)	ST127 (n=9)	ST959 (n=1)
ST92 (n=1)	ST141 (n=6)	ST961 (n=1)

#### 4.2.1.3 GimA-

Im Rahmen der Untersuchung des Insertionslokus von GimA ergab sich für 264 der untersuchten Isolate eine Zusammensetzung dieses Bereiches wie für *E. coli* K-12 bekannt. Eine Verknüpfung mit den MLST-Daten zeigte, dass sich die Isolate mit der Variante GimA- auf insgesamt 106 STs verteilen. Davon sind 58 STs, die 198 Isolate repräsentieren, 26 ST Komplexen zugeordnet.

Tab. 22 ST-Komplexe (n=26), in denen Isolate ohne GimA-assozierte Sequenzen (GimA-) identifiziert werden konnten (n=198). Weiterhin sind die STs (n=58) angegeben, die innerhalb der einzelnen ST-Komplexe die Variante GimA- enthalten. In Klammern sind die jeweiligen Häufigkeiten angegeben.

<b>Lokusvariante: GimA-</b>	
<b>Sequenztyp-Komplex</b>	<b>enthaltene STs in der Untersuchung</b>
STC10 (n=48)	ST10 (n=36), ST34 (n=1), ST43 (n=1), ST44 (n=1), ST48 (n=6), ST49 (n=1), ST 167 (n=1), ST548 (n=1)
STC11 (n=1)	ST61 (n=1)
STC12 (n=8)	ST12 (n=7), ST544 (n=1)
STC14 (n=4)	ST14(n=2), ST81(n=1), ST550(n=1)
STC23 (n=34)	ST23 (n=13), ST88 (n=15), ST356 (n=1), ST360 (n=1), ST367 (n=1), ST369 (n=2), ST410 (n=1)
STC31 (n=4)	ST130 (n=1), ST374 (n=1), ST393 (n=2)

<b>Lokusvariante: GimA-</b>	
<b>Sequenztyp-Komplex</b>	<b>enthaltene STs in der Untersuchung</b>
STC38 (n=5)	ST38 (n=5)
STC46 (n=2)	ST46 (n=2)
STC59 (n=2)	ST59 (n=2)
STC69 (n=3)	ST69 (n=3)
STC73 (n=7)	ST73 (n=6), ST638 (n=1)
STC86 (n=2)	ST86 (n=1), ST453 (n=1)
STC95 (n=43)	ST95 (n=40), ST140 (n=1), ST368 (n=1), ST421 (n=1)
STC101 (n=2)	ST101 (n=1), ST365 (n=1)
STC155 (n=9)	ST55 (n=1), ST56 (n=3), ST58 (n=2), ST155 (n=3)
STC156 (n=2)	ST348 (n=2)
STC165 (n=1)	ST100 (n=1)
STC168 (n=5)	ST93 (n=4), ST373 (n=1)
STC206 (n=1)	ST77 (n=1)
STC349 (n=1)	ST349 (n=1)
STC350 (n=5)	ST57 (n=2), ST350 (n=2), ST371 (n=1)
STC398 (n=1)	ST398 (n=1)
STC399 (n=2)	ST399 (n=1), ST536 (n=1)
STC405 (n=3)	ST72 (n=1), ST405 (n=2)
STC469 (n=2)	ST 162 (n=2)
STC522 (n=1)	S T541 (n=1)

Die verbleibenden 66 Isolate verteilen sich auf 48 STs, die derzeit keinem ST-Komplex zugeordnet sind.

Tab. 23 STs (n=48), in denen Isolate ohne GimA-assoziierte Sequenzen identifiziert werden konnten (n=66) und die derzeit keinem ST-Komplex zugeordnet sind. In Klammern sind die jeweiligen Häufigkeiten angegeben.

<b>Lokusvariante: GimA-</b>		
<b>STs, die keinem Komplex zugeordnet sind</b>		
ST45 (n=1)	ST71 (n=1)	ST297 (n=1)
ST47 (n=1)	ST75 (n=2)	ST347 (n=1)
ST50 (n=1)	ST78 (n=1)	ST351 (n=1)
ST51 (n=1)	ST79 (n=1)	ST359 (n=1)
ST52 (n=1)	ST84 (n=1)	ST361 (n=1)
ST53 (n=2)	ST85 (n=1)	ST362 (n=1)

<b>Lokusvariante: GimA-</b>		
<b>STs, die keinem Komplex zugeordnet sind</b>		
ST54 (n=1)	ST87 (n=1)	ST409 (n=1)
ST60 (n=1)	ST89 (n=1)	ST533 (n=1)
ST62 (n=10)	ST115 (n=3)	ST539 (n=2)
ST63 (n=1)	ST117 (n=1)	ST540 (n=1)
ST64 (n=1)	ST118 (n=1)	ST542 (n=1)
ST65 (n=1)	ST126 (n=1)	ST543 (n=1)
ST66 (n=1)	ST127 (n=2)	ST545 (n=1)
ST67 (n=1)	ST131 (n=1)	ST549 (n=1)
ST68 (n=3)	ST141 (n=1)	ST925 (n=1)
ST70 (n=2)	ST216 (n=1)	ST937 (n=1)

#### 4.2.1.4 Heterogene STs

Wie eingangs beschrieben, traten neben homogenen STs, also solchen, die nur Isolate mit einer einzigen Variante des Insertionslokus enthalten, auch heterogene STs auf. Diese enthielten Isolate mit unterschiedlichen Varianten des Insertionslokus.

Tab. 24 STs (n=9), die Isolate (n=156) enthalten, in denen mehrere unterschiedliche Lokus-Varianten vorkommen. Weiterhin ist angegeben, welche Varianten in den einzelnen STs vertreten sind. Mit Ausnahme des ST141 sind jeweils zwei verschiedene Varianten enthalten. In Klammern sind die jeweiligen Häufigkeiten angegeben.

<b>STs mit zwei bzw. drei Lokus-Varianten</b>	
<b>Sequenztyp</b>	<b>Lokus-Varianten</b>
ST12 (n=19)	GimA- (n=7), GimA-remnant (n=12)
ST73 (n=34)	GimA- (n=6), GimA-remnant (n=28)
ST95 (n=56)	GimA- (n=40), GimA+ (n=16)
ST117 (n=13)	GimA- (n=1), GimA-remnant (n=12)
ST126 (n=2)	GimA- (n=1), GimA+ (n=1)
ST127 (n=11)	GimA- (n=2), GimA-remnant (n=9)
ST140 (n=6)	GimA- (n=1), GimA+ (n=5)
ST638 (n=2)	GimA- (n=1), GimA-remnant (n=1)
ST141 (n=13)	GimA- (n=1), GimA-remnant (n=6), GimA+ (n=6)

#### 4.2.2 Minimum-Spanning-Tree (MSTree)

Mit entsprechender Software (Bionumerics V4.6, Applied Maths, Sint-Martens-Latem, Belgien) können die Daten der MLST-Analyse, die in der obigen Auswertung lediglich in Form der Zuordnung zu einem ST Ausdruck finden, auch zur Erzeugung einer grafischen Darstellung

dienen. Im MSTree dient nicht der ST selbst, sondern das ihn definierende Allelprofil als Grundlage der Berechnung.

Der aus den MLST-Datensätzen der 410 Isolate resultierende MSTree zeigt eine Konzentration der Varianten GimA+ und GimA-remnant auf der rechten Seite des MSTree (s. Abb. 16), wohingegen auf der linken Bildhälfte die Variante GimA- dominiert. Eine Ausnahme bildet der ST117, der, obwohl auf der linken Seite gelegen, neben GimA- (n=1) vor allem GimA-remnant-tragende Isolate enthält (n=12). Die Struktur des MSTree deutet eine Assoziation zwischen ST und Insertionslokusvariante an. Wie bereits erläutert, war jedoch statistisch kein signifikanter Zusammenhang zwischen ST bzw. ST-Komplex und den Insertionslokusvarianten nachweisbar.

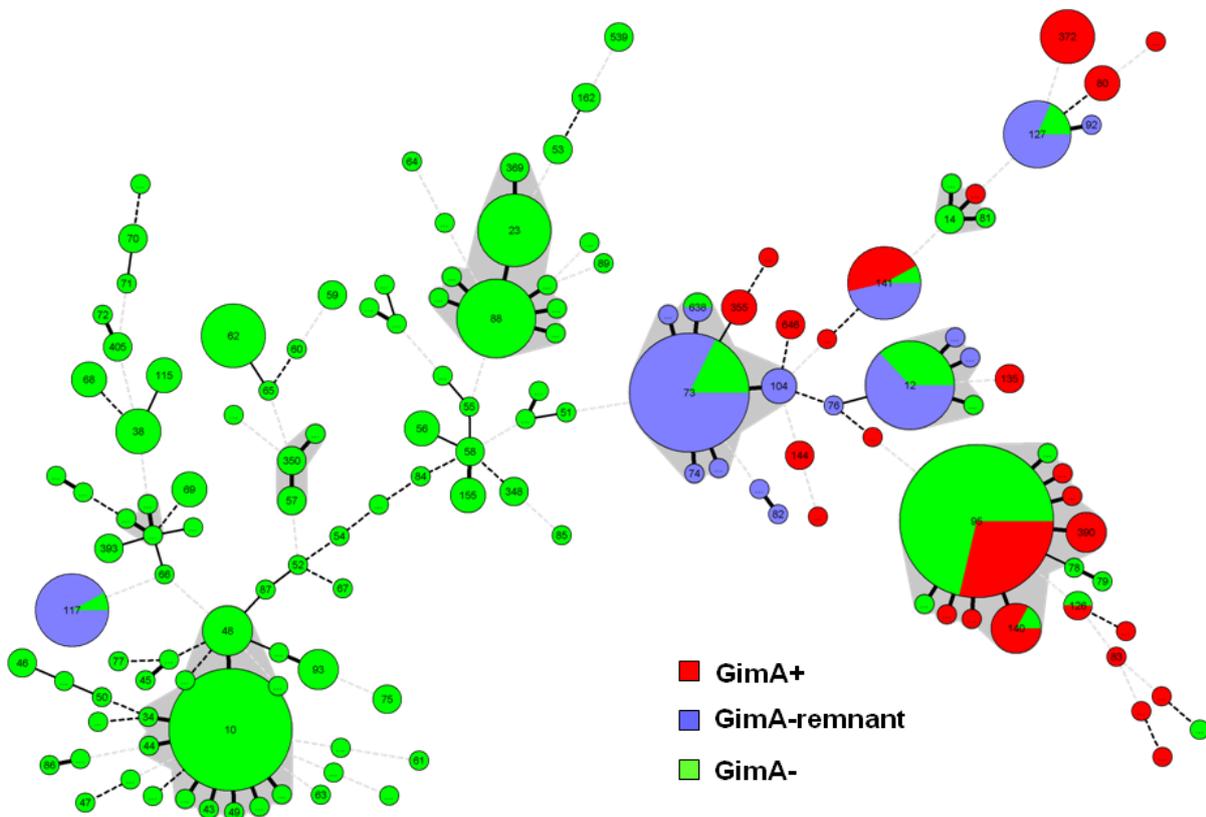


Abb. 16 MSTree basierend auf den MLST-Daten der 410 untersuchten *E. coli*-Stämme. Jeder Kreis entspricht einem ST. Die Anzahl der Stämme in einem ST wird durch die Größe des entsprechenden Kreises ausgedrückt. Die Farben der Kreise geben die Anteile der einzelnen Insertionslokusvarianten an den STs an. ST-Komplexe sind grau unterlegt und bestehen aus wenigstens drei STs, die sich in maximal einem Allel unterscheiden. Weitere quantitative Beziehungen zwischen den STs sind durch unterschiedliche Linien dargestellt. Dicke Linien zeigen, dass die verbundenen STs sechs gemeinsame Allele aufweisen. Dünne Linien zeigen fünf und schwarze gestrichelte Linien vier gemeinsame Allele an. Weniger als drei gemeinsame Allele werden durch grau gestrichelte Linien dargestellt.

### 4.2.3 Clonal Frame-Analyse

Um die Mikroevolution der 410 untersuchten *E. coli*-Stämme detaillierter abzubilden, wurde eine Clonal Frame-Analyse durchgeführt. Als Sequenz-basierte Methode ermöglicht sie die Berücksichtigung von Rekombinationseffekten, die bei den phylogenetischen Standardmethoden enge Beziehungen maskieren könnten. Auch im Vergleich mit der Allel-basierten MLST, die bereits auf Rekombinationseffekte Rücksicht nimmt, reagiert die Clonal Frame-Analyse viel weniger anfällig auf Rekombination und kann die Genealogie der getesteten *E. coli*-Population viel akkurater abbilden. Die Komplexität der notwendigen Berechnungen zeigt sich in den sehr langen Rechenzeiten für diese Analyse.

Für die Analyse mit der Software Clonal Frame wurden die Sequenzen der sieben Housekeeping-Gene aus der MLST, die zuvor in ein extended multi-FASTA Format (XMFA) zusammengefügt wurden, eingesetzt.

Da die Clonal Frame-Software selbst über kein Visualisierungsmodul verfügt, wurde der aus der Analyse resultierende „majority-rule consensus tree“ mit Hilfe der MEGA Software weiterverarbeitet. Der Stammbaum wurde in der Option „Radiation“ erzeugt (Abb. 17).

Es sind vier große Gruppen erkennbar (I-IV). Eine der Gruppen (IV) besteht nur aus dem ST117. Die drei verbleibenden Gruppen können in weitere Untergruppen unterteilt werden. Da eine weitere Unterteilung jedoch ohne Festlegung eines Cut-off nicht sinnvoll ist, werden für die Einteilung der Gruppen die Daten aus der Structure-Analyse verwendet.

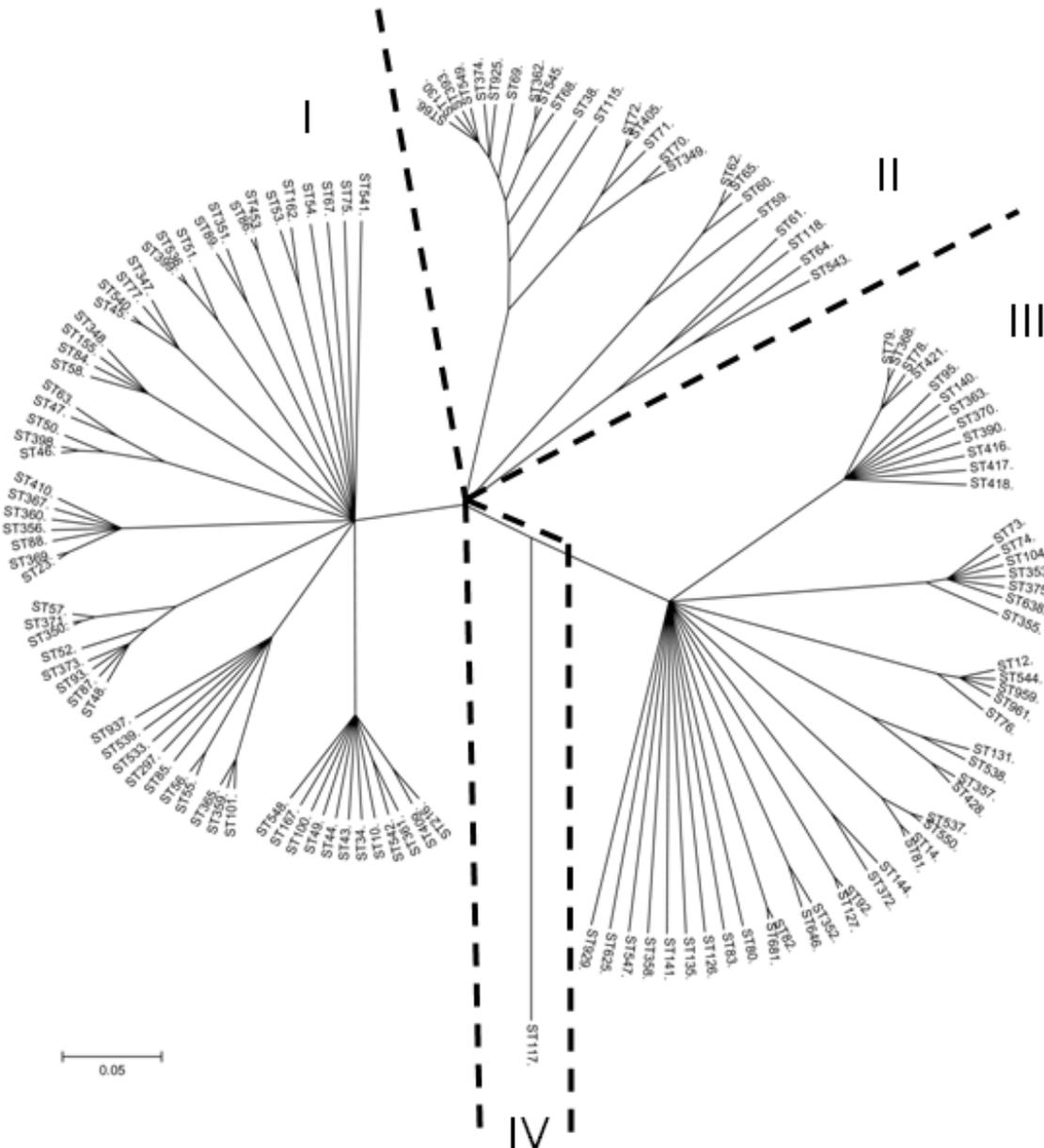


Abb. 17 Visualisierung des sich aus der Clonal Frame-Analyse ergebenden “majority-rule consensus tree” durch die MEGA Software (Version 4) (Tamura *et al.*, 2007). Aufgrund des gemeinsamen Clonal Frame bilden die STs vier große Gruppen (I-IV), wobei die Gruppe IV lediglich aus dem ST117 besteht. Der angegebene Maßstab gibt die genetische Distanz an.

#### 4.2.4 Structure-Analyse

Die Structure-Analyse ist eine Sequenz-basierte Methode, die die Sequenzdaten aus der MLST verwendet. Sie berücksichtigt ähnlich der Clonal Frame-Analyse Rekombinationsereignisse. Anders als bei der Clonal Frame-Analyse ist jedoch nicht die Bestimmung der Beziehungen der einzelnen STs untereinander das Ziel, sondern die Vorhersage von Gruppenbildungen innerhalb einer rekombinierenden Spezies (hier *E. coli*). Das „linkage model“ der Structure-

Software errechnet für jeden Nukleotidpolymorphismus eine Wahrscheinlichkeit ( $q$ ) für die Abstammung aus den einzelnen vorgegebenen Gruppen.

In der vorliegenden Analyse wurden vier Gruppen vorgegeben und analog den EcoR-Phylogruppen benannt. Die letztliche Zuordnung jedes ST erfolgt mittels der aufsummierten Einzelwahrscheinlichkeiten für die Zugehörigkeit zu den vorgegebenen Gruppen jedes Nukleotidpolymorphismus des betreffenden ST. Die Grenzwerte wurden wie folgt festgesetzt: die Zuordnung zu einer Gruppe (A, B1, B2 oder D) erfolgte, wenn  $q \geq 0.67$ ; STs, die keiner Gruppe zugeordnet werden konnten, wurden als Hybridgruppe AxB1 angesprochen, wenn  $q$  für die Zugehörigkeit zu A und B1 zusammen  $\geq 0,8$  war. Alle verbleibenden Isolate wurden zu einer zweiten Hybridgruppe ABD zusammengefasst.

Wie bereits durch Wirth *et al.* (2006) festgestellt, erkannte die Structure-Software aufgrund des Polymorphismus der eingespeisten Sequenzen vier Gruppen, die für die Stämme der EcoR-Sammlung weitgehend mit der Zuordnung in die EcoR-Phylogruppen übereinstimmten. Einige Isolate wurden jedoch in die genannten Hybridgruppen (AxB1 und ABD) eingeordnet (Wirth *et al.*, 2006). In der vorliegenden Untersuchung wurde die Analyse auf weitere STs ausgeweitet. Der Phylogruppe A wurden 14 der 138 STs (10,1%) zugeordnet, B1 13 STs (9,4%), B2 49 STs (35,5%) und D 18 STs (13,0%). Somit sind den ursprünglichen Phylogruppen 94 STs (68,1%) zugeordnet. Auf die Hybridgruppen verteilen sich 44 STs (31,9%), auf AxB1 26 STs (18,8%) und auf ABD 18 STs (13,0%). Die Ergebnisse der Analyse sind in Tab. 25 zusammengestellt.

Tab. 25 Wahrscheinlichkeitswerte ( $q$ ) für die Zugehörigkeit zu den vier festgelegten Gruppen (A, B1, B2 und D) sowie der daraus resultierenden Zuordnung zu einer Phylogruppe (Grenzwert für eine Zuordnung zu einer Gruppe  $q \geq 0.67$ . Isolate, die keiner Gruppe zugeordnet werden konnten, wurden als Hybridgruppe AxB1 angesprochen, wenn die Werte für  $q$  für die Zugehörigkeit zu A und B1 zusammen  $\geq 0,8$  war. Alle verbleibenden Isolate wurden zu einer zweiten Hybridgruppe ABD zusammengefasst).

ST	A	B1	B2	D	Phylogruppe
ST10	0.923	0.026	0.025	0.026	A
ST12	0.027	0.024	0.925	0.024	B2
ST14	0.033	0.134	0.809	0.024	B2
ST23	0.121	0.747	0.066	0.066	B1
ST34	0.830	0.113	0.024	0.032	A
ST38	0.029	0.027	0.032	0.911	D
ST43	0.921	0.028	0.025	0.026	A
ST44	0.814	0.137	0.024	0.025	A
ST45	0.602	0.344	0.024	0.030	AxB1
ST46	0.668	0.281	0.028	0.024	AxB1
ST47	0.602	0.252	0.032	0.114	AxB1
ST48	0.811	0.135	0.027	0.026	A
ST49	0.906	0.044	0.024	0.025	A
ST50	0.704	0.242	0.028	0.027	A

ST	A	B1	B2	D	Phylogruppe
ST51	0.235	0.442	0.295	0.028	ABD
ST52	0.553	0.306	0.060	0.080	AxB1
ST53	0.345	0.605	0.026	0.024	AxB1
ST54	0.334	0.510	0.027	0.129	AxB1
ST55	0.200	0.651	0.072	0.078	AxB1
ST56	0.088	0.624	0.031	0.257	ABD
ST57	0.423	0.307	0.030	0.240	ABD
ST58	0.391	0.437	0.032	0.139	AxB1
ST59	0.084	0.136	0.352	0.428	ABD
ST60	0.165	0.029	0.267	0.538	ABD
ST61	0.145	0.138	0.036	0.681	D
ST62	0.117	0.028	0.324	0.531	ABD
ST63	0.733	0.030	0.027	0.211	A
ST64	0.329	0.113	0.038	0.520	ABD
ST65	0.250	0.031	0.251	0.469	ABD
ST66	0.144	0.151	0.030	0.676	D
ST67	0.323	0.623	0.028	0.025	AxB1
ST68	0.025	0.025	0.026	0.924	D
ST69	0.025	0.028	0.028	0.919	D
ST70	0.026	0.086	0.040	0.849	D
ST71	0.027	0.027	0.034	0.912	D
ST72	0.032	0.026	0.030	0.912	D
ST73	0.028	0.024	0.921	0.026	B2
ST74	0.029	0.027	0.820	0.124	B2
ST75	0.327	0.504	0.086	0.083	AxB1
ST76	0.026	0.025	0.922	0.028	B2
ST77	0.490	0.459	0.025	0.025	AxB1
ST78	0.027	0.029	0.813	0.130	B2
ST79	0.027	0.145	0.801	0.027	B2
ST80	0.114	0.032	0.824	0.030	B2
ST81	0.032	0.261	0.679	0.028	B2
ST82	0.120	0.117	0.735	0.028	B2
ST83	0.031	0.027	0.874	0.068	B2
ST84	0.408	0.396	0.035	0.161	AxB1
ST85	0.029	0.911	0.031	0.030	B1
ST86	0.262	0.421	0.040	0.277	ABD
ST87	0.510	0.336	0.069	0.085	AxB1
ST88	0.128	0.743	0.062	0.067	B1
ST89	0.189	0.666	0.077	0.068	AxB1
ST92	0.022	0.023	0.841	0.113	B2
ST93	0.658	0.177	0.034	0.130	AxB1
ST95	0.028	0.027	0.919	0.026	B2
ST100	0.806	0.070	0.028	0.096	A
ST101	0.208	0.743	0.025	0.023	B1

ST	A	B1	B2	D	Phylogruppe
ST104	0.028	0.026	0.921	0.024	B2
ST115	0.114	0.042	0.103	0.740	D
ST117	0.171	0.017	0.404	0.408	ABD
ST118	0.276	0.124	0.044	0.557	ABD
ST126	0.027	0.024	0.923	0.026	B2
ST127	0.028	0.025	0.828	0.118	B2
ST130	0.025	0.029	0.025	0.921	D
ST131	0.034	0.026	0.877	0.063	B2
ST135	0.026	0.025	0.922	0.027	B2
ST140	0.028	0.028	0.917	0.027	B2
ST141	0.041	0.026	0.906	0.028	B2
ST144	0.028	0.024	0.825	0.122	B2
ST155	0.264	0.563	0.030	0.143	AxB1
ST162	0.423	0.526	0.027	0.024	AxB1
ST167	0.809	0.137	0.028	0.026	A
ST216	0.700	0.134	0.029	0.137	A
ST297	0.252	0.687	0.032	0.029	B1
ST347	0.491	0.458	0.025	0.026	AxB1
ST348	0.265	0.684	0.026	0.025	B1
ST349	0.072	0.118	0.030	0.780	D
ST350	0.293	0.335	0.026	0.346	ABD
ST351	0.098	0.618	0.036	0.247	ABD
ST352	0.116	0.030	0.824	0.030	B2
ST353	0.027	0.139	0.804	0.030	B2
ST355	0.026	0.025	0.924	0.025	B2
ST356	0.127	0.734	0.064	0.076	B1
ST357	0.028	0.026	0.915	0.031	B2
ST358	0.029	0.024	0.918	0.029	B2
ST359	0.311	0.640	0.027	0.022	AxB1
ST360	0.125	0.744	0.065	0.067	B1
ST361	0.621	0.260	0.026	0.093	AxB1
ST362	0.028	0.131	0.026	0.815	D
ST363	0.024	0.027	0.920	0.028	B2
ST365	0.202	0.631	0.142	0.025	AxB1
ST367	0.130	0.737	0.066	0.066	B1
ST368	0.025	0.139	0.810	0.025	B2
ST369	0.126	0.730	0.078	0.066	B1
ST370	0.028	0.027	0.917	0.028	B2
ST371	0.276	0.322	0.078	0.324	ABD
ST372	0.028	0.023	0.824	0.125	B2
ST373	0.677	0.159	0.032	0.132	A
ST374	0.027	0.032	0.028	0.914	D
ST375	0.027	0.025	0.922	0.026	B2
ST390	0.029	0.028	0.916	0.027	B2

ST	A	B1	B2	D	Phylogruppe
ST393	0.025	0.030	0.027	0.917	D
ST398	0.578	0.366	0.028	0.028	AxB1
ST399	0.303	0.550	0.073	0.074	AxB1
ST405	0.030	0.028	0.032	0.910	D
ST409	0.542	0.319	0.049	0.090	AxB1
ST410	0.134	0.731	0.062	0.073	B1
ST416	0.027	0.027	0.919	0.027	B2
ST417	0.028	0.031	0.914	0.027	B2
ST418	0.024	0.026	0.922	0.028	B2
ST421	0.132	0.033	0.810	0.025	B2
ST428	0.030	0.026	0.913	0.030	B2
ST453	0.250	0.425	0.043	0.281	ABD
ST533	0.216	0.725	0.029	0.030	B1
ST536	0.282	0.493	0.045	0.179	ABD
ST537	0.026	0.024	0.923	0.027	B2
ST538	0.029	0.028	0.917	0.025	B2
ST539	0.218	0.718	0.034	0.030	B1
ST540	0.597	0.344	0.026	0.032	AxB1
ST541	0.708	0.240	0.026	0.026	A
ST542	0.549	0.030	0.029	0.392	ABD
ST543	0.323	0.161	0.034	0.482	ABD
ST544	0.029	0.026	0.918	0.027	B2
ST545	0.024	0.130	0.025	0.820	D
ST547	0.030	0.023	0.920	0.027	B2
ST548	0.928	0.025	0.024	0.022	A
ST549	0.026	0.031	0.027	0.917	D
ST550	0.027	0.029	0.918	0.027	B2
ST625	0.068	0.025	0.882	0.024	B2
ST638	0.029	0.025	0.920	0.026	B2
ST646	0.025	0.022	0.929	0.024	B2
ST681	0.125	0.110	0.739	0.026	B2
ST925	0.027	0.030	0.025	0.918	D
ST929	0.031	0.027	0.915	0.027	B2
ST937	0.361	0.463	0.029	0.146	AxB1
ST959	0.027	0.037	0.910	0.026	B2
ST961	0.027	0.026	0.922	0.025	B2

#### 4.2.5 Gemeinsame Betrachtung von Clonal Frame- und Structure-Analyse

Durch die gemeinsame Darstellung der Ergebnisse der Clonal Frame- und der Structure-Analyse zeigte sich eine Kongruenz in der Gruppenbildung. Die drei bereits in der Beschreibung des auf der Clonal Frame-Analyse basierenden Dendrogramms angesprochenen vier großen Gruppen korrelierten mit den in der Structure-Analyse bestimmten Gruppen. Im Clonal Frame-

Dendrogramm bildeten alle Stämme der Phylogruppe B2 eine große Gruppe. Die Stämme mit der Phylogruppe D bildeten zusammen mit einigen Stämmen der Hybridgruppe ABD eine weitere große Gruppe. Die Beziehung der D und der ABD-Stämme zeigt sich auch in den Daten der Structure-Analyse. Die Stämme der ABD-Gruppe wiesen für die Zugehörigkeit zur Phylogruppe D die höchsten Wahrscheinlichkeiten auf ( $q \geq 0,43$ ). Allerdings waren die Wahrscheinlichkeiten nicht ausreichend für eine Eingruppierung in Gruppe D, aber höher als die aller anderen zur Gruppe ABD gehörenden Stämme. In ähnlicher Weise stellte sich die Situation für die dritte große Gruppe dar. In ihr waren Stämme der Phylogruppen A und B1 sowie der Mischgruppe aus beiden (AxB1) versammelt. Daneben kamen einige Stämme der Hybridgruppe ABD vor, die laut Structure-Analyse die höchsten Wahrscheinlichkeiten ( $q$ ) für eine Zugehörigkeit zu den Gruppen A und B1, für die Gruppe D jedoch nur niedrige Werte aufweisen. Die letzte Gruppe besteht ausschließlich aus Vertretern des ST117 und ist daher auch in den Werten der Structure-Analyse übereinstimmend. Die Werte für  $q$  zeigen, dass der ST117 von den vier Phylogruppen die größte Übereinstimmung mit den Gruppen B2 und D aufweist (jeweils  $q \geq 0,4$ ). Abb. 18 verdeutlicht die Kongruenz beider Analyseergebnisse.

#### 4.2.5.1 Verteilung der Varianten des GimA-Insertionslokus

In Abb. 18 ist zu erkennen, dass GimA<sup>+</sup> ausschließlich bei solchen Stämmen vorkommt, deren STs der Phylogruppe B2 zugehören. Ebenso deutlich häufen sich in dieser Phylogruppe STs bzw. Stämme, die das GimA-remnant tragen. Einige weitere GimA-remnant-Isolate finden sich im ST117, der der Hybridphylogruppe ABD zugeordnet ist. Dort weisen zwölf der 13 Isolate die Insertionslokusvariante GimA-remnant auf. Das verbleibende Isolat ist GimA<sup>-</sup>. Die Sonderstellung, die der ST117 aufgrund seiner Alleinstellung innerhalb des Clonal Frame-Diagrammes einnimmt, spiegelt sich auch in dem Vorhandensein des GimA-remnant wider. Mit Ausnahme des ST117 ist die Variante des Insertionslokus GimA-remnant ausschließlich in der Phylogruppe B2 zu finden und weist damit eine Verteilung auf, die der Variante GimA<sup>+</sup> entspricht. Alle anderen Phylogruppen (A, B1 und D) sowie die Hybridgruppe AxB1 sind GimA<sup>-</sup>. Mit Ausnahme des ST117 sind ebenfalls alle Vertreter der Hybridgruppe ABD GimA<sup>-</sup>. In Abb. 18 ist die Verteilung der Insertionslokusvarianten abgebildet.

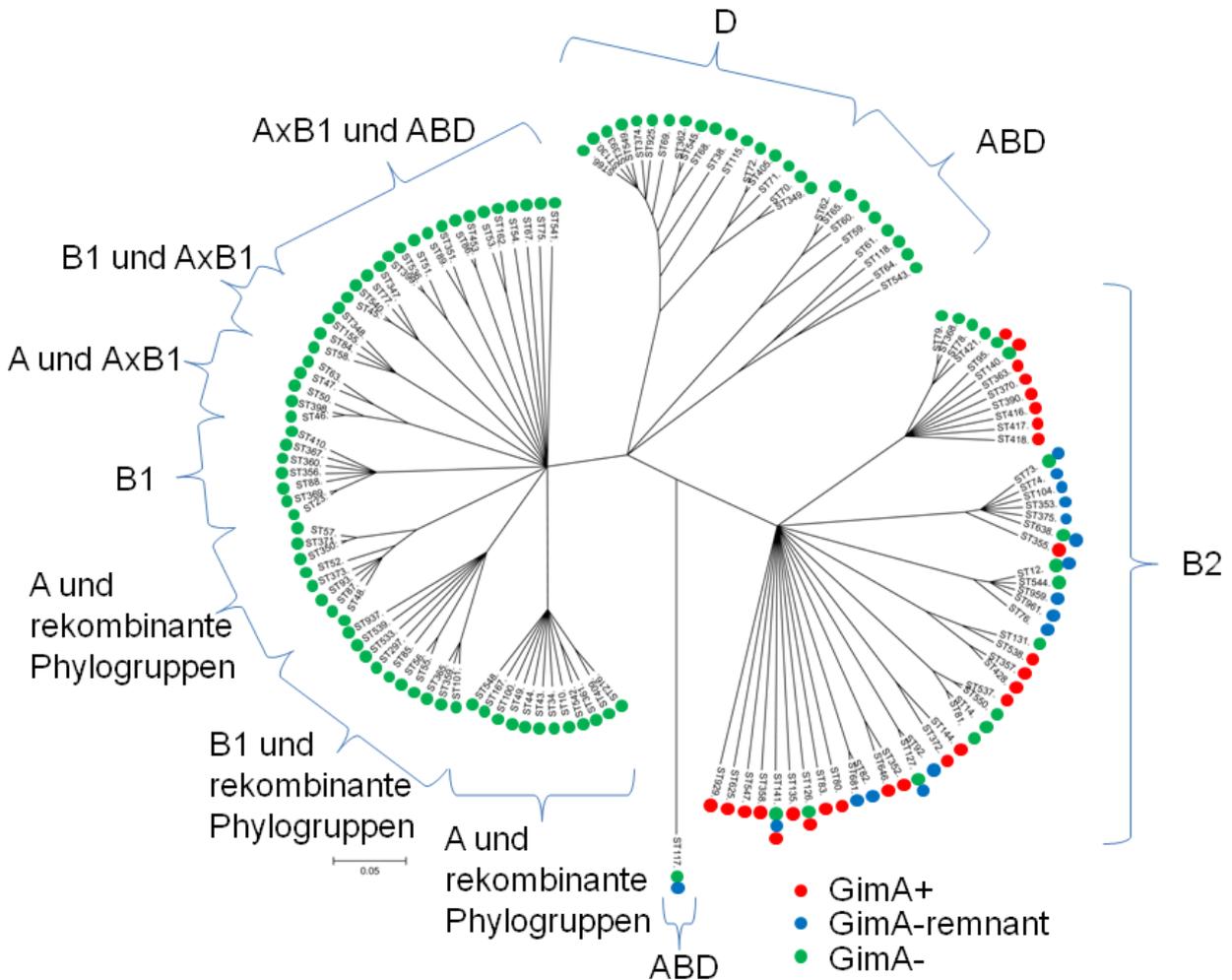


Abb. 18 Clonal Frame-Dendrogramm aus Abb. 17 erweitert durch die Insertionsloкусvarianten, die in den Stämmen in Abhängigkeit vom ST vorkommen (rote Punkte: GimA+; blaue Punkte: GimA-remnant und grüne Punkte: GimA-). Weiterhin sind die Phylogruppen bzw. Hybridgruppen aufgeführt, zu denen die STs basierend auf der Structure-Analyse zugeordnet wurden.

#### 4.2.5.2 Assoziationen zwischen der Herkunft der Isolate und den Lokusvarianten

Statistisch relevante Assoziationen zwischen der Herkunft (Mensch/Tier) und den verschiedenen Varianten des Insertionslokus konnten nicht identifiziert werden (s. Tab. 27 im Anhang).

#### 4.2.5.3 Assoziationen zwischen den Pathovaren und Lokusvarianten

Eine Analyse der Verteilung der Pathovaren und Insertionsloкусvarianten ergab folgende signifikante Assoziationen ( $p \leq 0.05$ , Chi-Quadrat-Test) innerhalb der 410 Stämme der Testpopulation: Die Lokusvariante GimA+ war signifikant assoziiert mit den Pathovaren APEC und NMEC. Demgegenüber war die Variante GimA-remnant signifikant assoziiert mit UPEC-Stämmen. Die Stämme, die aus den Stuhlproben klinisch gesunder Menschen bzw. den Kotproben klinisch gesunder Tiere isoliert wurden, waren signifikant häufiger GimA- (s. Tab. 26).

Tab. 26 Verteilung der Insertionsloкусvarianten und der Pathovaren innerhalb der getesteten 410 ExPEC- und fäkalen *E. coli*-Isolate. In Klammern ist jeweils die Anzahl der Isolate angegeben, die der Phylogruppe B2 bzw. einer der übrigen Phylogruppen zugehörig sind. Signifikante Assoziationen ( $p \leq 0.05$ ; Chi-Quadrat-Test) sind mit Sternchen (\*) gekennzeichnet.

Pathovar	<i>yjiD-yjiE</i> locus pattern			$\Sigma$ (n) [B2/non B2]
	GimA+ (n) [B2/non B2]	GimA-remnant (n) [B2/non B2] <sup>1</sup>	GimA- (n) [B2/non B2]	
APEC	23* (23/0)	13 (3/10)	62 (16/46)	98 (42/56)
NMEC	10* (10/0)	1 (1/0)	14 (7/7)	25 (18/7)
UPEC	22 (22/0)	52* (50/2)	66 (26/40)	140 (98/42)
SepEC	1 (1/0)	5 (5/0)	22 (4/18)	28 (10/18)
Feces <sub>Mensch</sub>	9 (9/0)	6 (6/0)	71* (16/55)	86 (31/55)
Fecal <sub>Tier</sub>	1 (1/0)	3 (3/0)	29* (0/29)	33 (4/29)
$\Sigma$	66 (66/0)	80 (68/12)	264 (69/195)	410 (203/207)

<sup>1</sup> Die zwölf nicht B2-Stämme gehören zum ST117, der der Hybridphylogruppe ABD zugeordnet ist.

Wurden in die Betrachtung nur die Stämme der Phylogruppe B2 einbezogen, ergaben sich die gleichen signifikanten Assoziationen. Jedoch waren die Anteile der einzelnen Varianten des Insertionsloкус bei den einzelnen Pathovaren verändert. Die Anteile der Varianten GimA+ und GimA-remnant stiegen an und demzufolge sank der Anteil an GimA- Stämmen. In Tab. 26 sind in Klammern jeweils die Anzahlen an B2 und nicht-B2-Stämmen angegeben. Entsprechend ergaben sich innerhalb der Phylogruppe B2 die Anteile der einzelnen Lokusvarianten zwischen den Pathovaren wie folgt: Beide, APEC und NMEC wiesen hohe Gehalte an Stämmen mit der Variante GimA+ (54,7% and 55,6%) auf der einen Seite und auf der anderen niedrige Gehalte an GimA-remnant-tragenden Stämmen (7,1% and 5,6%) auf. Das umgekehrte Verhältnis fand sich bei UPEC. Dort standen hohe Anteile an GimA-remnant-tragenden Stämmen (51,0%) weniger Isolaten mit der Variante GimA+ gegenüber (22,4%).

## 4.3 Sequenzanalyse von *ibeA*

### 4.3.1 Sequenzanalyse von *ibeA*

Es wurden insgesamt 34 *E. coli*-Stämme für die Sequenzanalyse von *ibeA* (invasion of brain endothelium) ausgewählt. Die Stämme repräsentieren den phylogenetischen Hintergrund, in dem *ibeA* vorkommt, d. h. 17 verschiedene STs, die alle der Phylogruppe B2 angehören. Aufgrund der Größe des *ibeA*-Gens von 1.341 bp wurde die Sequenzanalyse mittels zweier PCR-Anwendungen vorgenommen, deren Amplifikate sich überlappen (s. Abb. 19). Die PCR-basierte Generierung des Template für die Sequenzanalyse erfolgte im Institut für Mikrobiologie und Tierseuchen an der Freien Universität Berlin. Die Sequenzanalyse selbst wurde extern durch LGC (Berlin, Deutschland) durchgeführt. Das Ergebnis der Sequenzanalyse wurde in Form eines Trace-file bereitgestellt. Ein Trace-file beinhaltet die Messdaten der vier Kanäle (einer für jede Base) und dient u. a. der Qualitätskontrolle der Sequenzanalyse. Die Trace-files wurden mit Hilfe der Kodon Software (Version 3.6 Applied Maths, Sint Martens-Latem, Belgien) weiterbearbeitet. Die Bearbeitung der einzelnen Sequenzen umfasste ein Assembling, d. h. das Zusammenfügen der zwei überlappenden Sequenzen und ein Trimming, d. h. die Entfernung von Sequenzabschnitten, die sich bedingt durch die Bindungsstellen der Primer up- und downstream des *ibeA*-Gens befinden. Am Ende lag für jeden der 34 Stämme eine vollständige *ibeA*-Sequenz (1.341 bp) vor. Die einzelnen Sequenzen sind als Dot-view in Abb. 20 aufgeführt. Dabei sind zur besseren Übersicht lediglich die von der Template-Sequenz des Stamms IMT5155 unterschiedlichen Basen, die SNPs entsprechen, aufgeführt. Alle Positionen, die mit der Template-Sequenz übereinstimmen, sind als Punkt dargestellt.

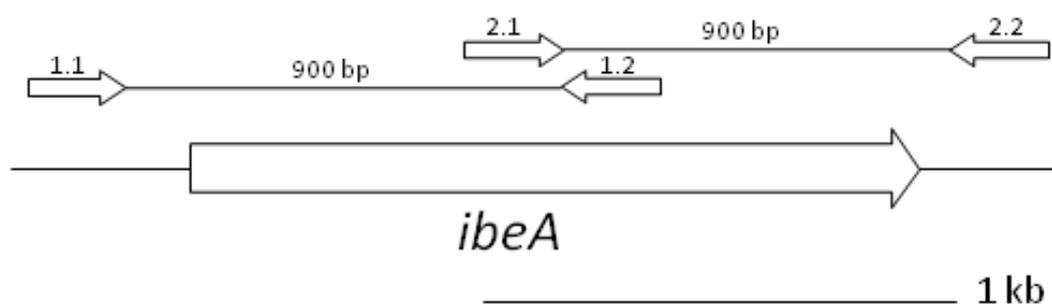


Abb. 19 Bindungsstellen der Primer zur Erzeugung des PCR-Produkts für die Sequenzanalyse von *ibeA* (Primer-Nummer in der Grafik und Primer-Bezeichnung in Tab. 4: 1.1 = *ibeAFr1FP*, 1.2 = *ibeAFr1RP*, 2.1 = *ibeAFr2FP*, 2.2 = *ibeAFr2RP*). Es werden zwei ca. 900 bp große Produkte erzeugt, die *ibeA*-Fragmente 1 und 2. Die Abbildung basiert auf den Sequenzen der *E. coli*-Stämme RS218, A0 34/86, UT189 und APEC O1 (Chen *et al.*, 2006; Hejnova *et al.*, 2005; Huang *et al.*, 2001a; Johnson *et al.*, 2007)

### 4.3.2 Single-Lokus-Sequenztypisierung (SLST)

Im Anschluss an die Bearbeitung der Trace-files wurde ein Alignment durchgeführt, d. h. eine Ausrichtung der Sequenzen untereinander anhand von Übereinstimmungen in der Basenabfolge und mittels des Maximum-Parsimony-Algorithmus ein Dendrogramm berechnet (s. Abb. 21). In dem Dendrogramm sind zwölf unterschiedliche *ibeA*-SLSTs zu erkennen.

Mit Ausnahme des NMEC-Stammes C5 wiesen alle (n=11) Stämme des ST95 eine 100%-ige Übereinstimmung der *ibeA*-Sequenzen auf. Die *ibeA*-Sequenz von C5 unterschied sich von den restlichen ST95 *ibeA*-Sequenzen in zwei Nukleotidpositionen.

Der ST95 stellt den Gründungs-ST und damit namensgebenden ST des ST95-Komplex dar. Zum ST95-Komplex gehören weitere STs, die sich jeweils in maximal einem Allel unterscheiden und folglich nah verwandt sind.

In der vorliegenden Untersuchung waren neben dem ST95 noch weitere Angehörige des ST95-Komplexes vertreten (ST140, ST368, ST370, ST390, ST416, ST 417 und ST418). Allen Vertretern des ST95-Komplexes gemeinsam war eine identische *ibeA*-Sequenz. Die einzige Ausnahme stellte, wie erwähnt, der NMEC-Stamm C5 dar.

Die weiteren STs, die durch mehr als einen Stamm vertreten waren, zeichneten sich ebenfalls durch einheitliche *ibeA*-Sequenzen aus (jeweils zwei Isolate aus den STs 355, 372 und 135).



```

delta 32 .....
delta 54 .....
delta 68 .....
IMT2099 .....
E10094 .....
E422 .....
IMT2112 .....
IMT2121 .....
IMT2248 .....
IMT2295 .....
IMT2492 .....
IMT5119 .....
IMT8998 .....
IHE3036 .....
IMT9300 .....
IMT9650 .....
MT78 .....
RS218 .....
UTI89 .....

IMT5155 .....
248 .....
2201 .....
7011 .....
A459 .....
APEC01 .....
E13155 .....
BK 658 .....
C5 .....
delta 8 .....
delta 18 .....
delta 20 .....
delta 21 .....
delta 22 .....
delta 27 .....
delta 32 .....
delta 54 .....
delta 68 .....
IMT2099 .....
E10094 .....
E422 .....
IMT2112 .....
IMT2121 .....
IMT2248 .....
IMT2295 .....
IMT2492 .....
IMT5119 .....
IMT8998 .....
IHE3036 .....
IMT9300 .....
IMT9650 .....
MT78 .....

```





IMT9650	.....A.....T.....T.....T.....C...
NT78	.....
RS218	.....
UTI89	.....
IMT5155	.....
248	.....
2201	.....T.....G.....
7011	.....T.....G.....
A459	.....
APEC01	.....
B13155	.....
BK 658	.....G.....
C5	.....
delta 8	.....
delta 18	.....
delta 20	.....
delta 21	.....
delta 22	.....
delta 27	.....
delta 32	.....
delta 54	.....
delta 68	.....
IMT2099	.....T.....G.....
E10094	.....T.....G.....
E422	.....C.....G.....A.....
IMT2112	.....T.....T.....G.....
IMT2121	.....C.....G.....A.....
IMT2248	.....C.....G.....A.....C.....
IMT2295	.....T.....G.....
IMT2492	.....T.....G.....
IMT5119	.....C.....G.....A.....C.....
IMT8998	.....T.....G.....
IHE3036	.....
IMT9300	.....T.....G.....
IMT9650	.....T.....G.....
NT78	.....
RS218	.....
UTI89	.....
IMT5155	.....
248	.....
2201	.....
7011	.....
A459	.....
APEC01	.....
B13155	.....
BK 658	.....
C5	.....
delta 8	.....
delta 18	.....

```

delta 20 .....
delta 21 .....
delta 22 .....
delta 27 .....
delta 32 .....
delta 54 .....
delta 68 .....
IMT2099 .....
E10094 .....
E422 .....
IMT2112 .....
IMT2121 .....
IMT2248 .....
IMT2295 .....
IMT2492 .....
IMT5119 .....
IMT8998 .....
IHE3036 .....
IMT9300 .....
IMT9650 .....
MT78 .....
RS218 .....
UT189 .....

1130 1140 1150 1160 1170 1180 1190 1200 1210 1220 1230 1240 1250 126
CGACCTTTTTC AAAATCCCTT TGGGCTTAATGATTC GCGACACAGTCCGACGATATTTGGGCTTACGACATATTCGGGCTGTATATTCGCTGATTCACCTGGCATA
.....
2201 .....
7011 .....
A459 .....
APRC01 .....
B13155 .....
BK 658 .....
C5 .....
delta 8 .....
delta 18 .....
delta 20 .....
delta 21 .....
delta 22 .....
delta 27 .....
delta 32 .....
delta 54 .....
delta 68 .....
IMT2099 .....
E10094 .....
E422 .....
IMT2112 .....
IMT2121 .....
IMT2248 .....
IMT2295 .....
IMT2492 .....
IMT5119 .....
IMT8998 .....

```

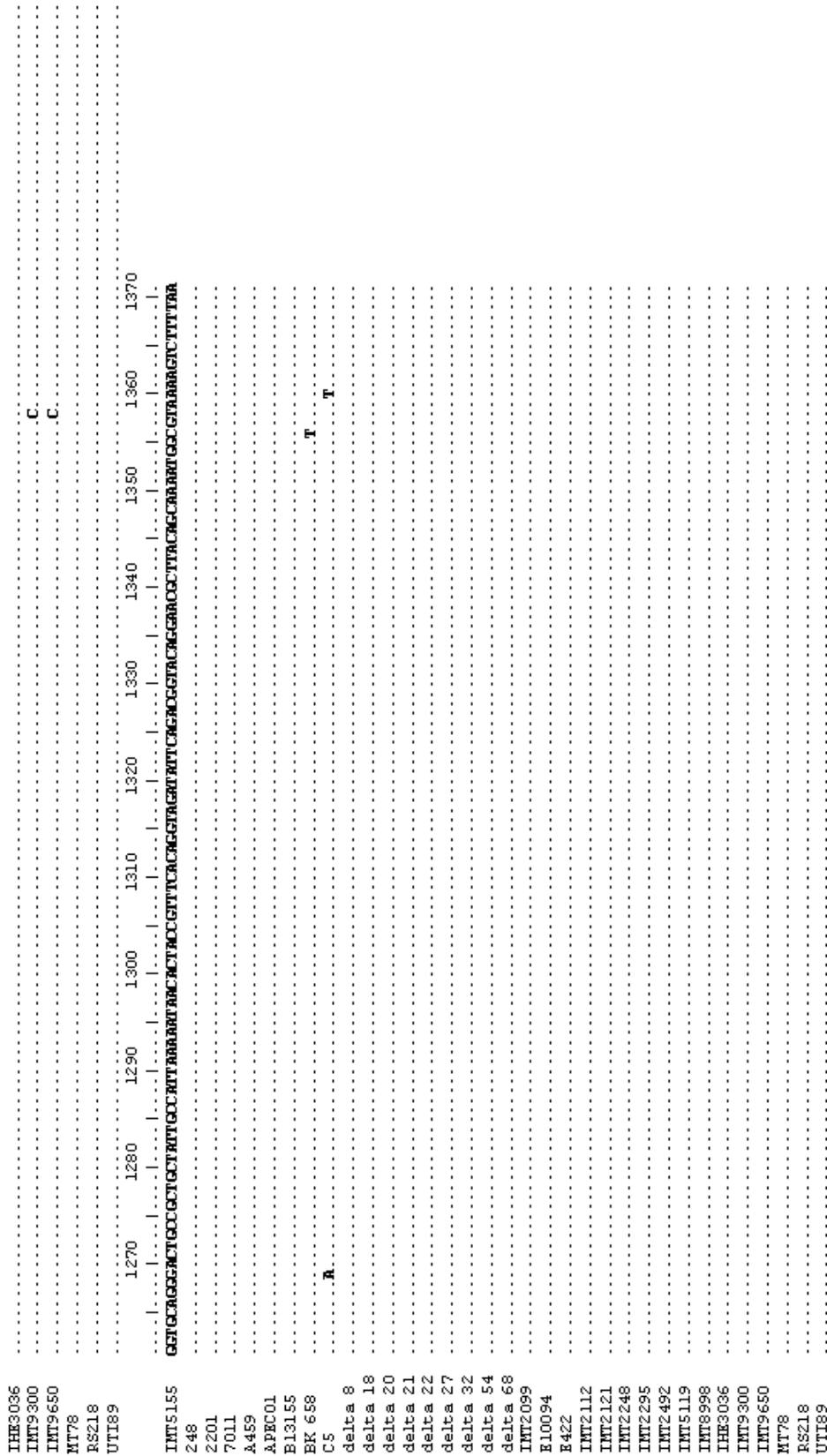


Abb. 20 Vergleich der *ibeA*-Sequenzen der 34 sequenzanalysierten ExPEC-Stämme. Die Stämme repräsentieren den phylogenetischen Hintergrund, in dem *ibeA* auftritt. Übereinstimmungen mit der Template-Sequenz des Stamms IMT5155 sind durch Punkte und Unterschiede durch Nennung der Base gekennzeichnet

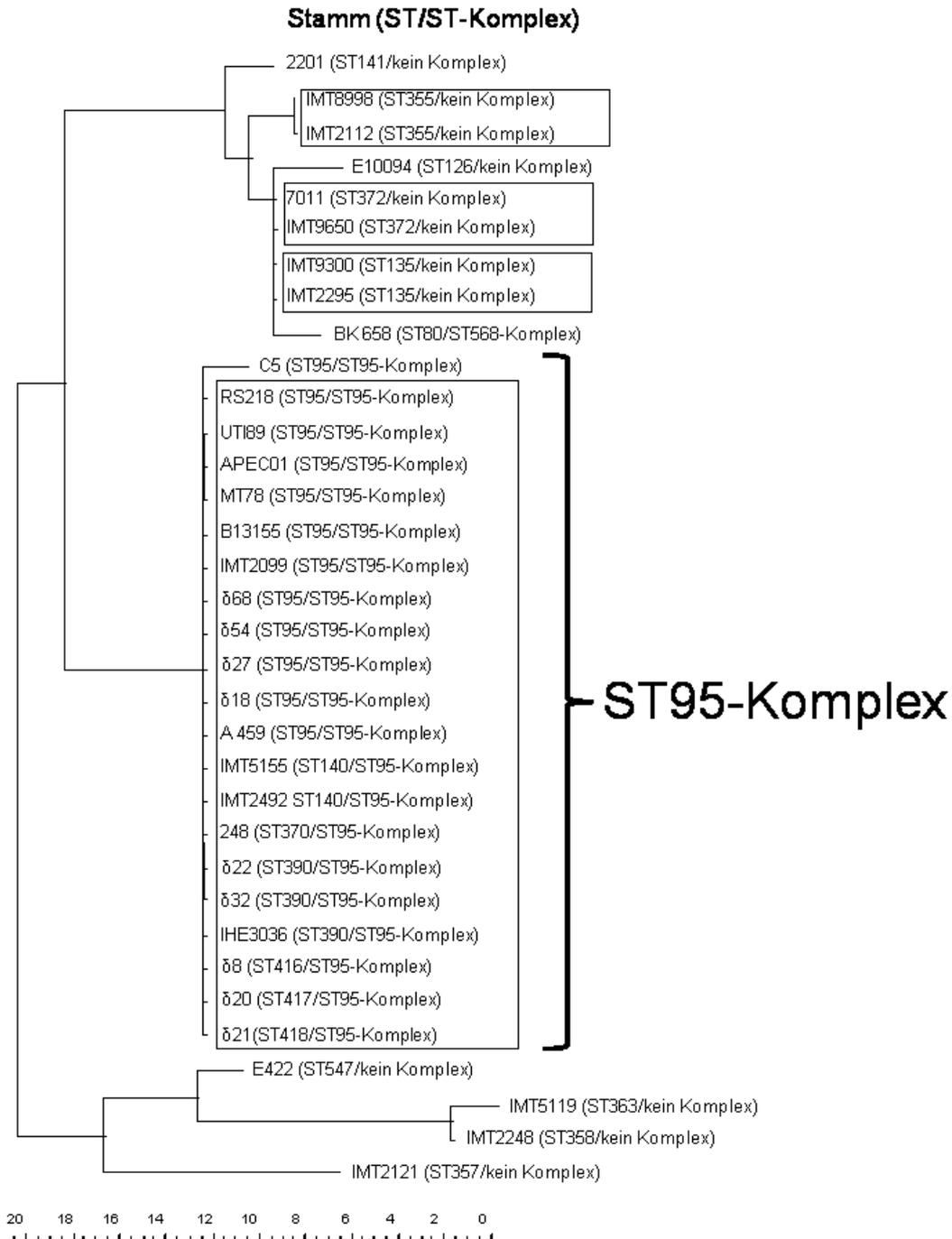


Abb. 21 Maximum-Parsimony-Dendrogramm auf Grundlage der *ibeA*-Sequenzen von 34 *E. coli*-Stämmen, die 17 verschiedenen STs zugehörig sind und den phylogenetischen Hintergrund, in dem *ibeA* vorkommt, repräsentieren. Alle Isolate gehören zur Phylogruppe B2. Es zeigen sich zwölf verschiedene *ibeA*-SLSTs. STs mit identischem phylogenetischen Hintergrund (ST bzw. ST-Komplex) sind eingerahmt. Der Maßstab gibt die genetische Distanz der *ibeA*-Sequenzen an.

#### 4.4 Evolutionäre Analyse von *ibeA*

Es wurden 21 *ibeA*-Sequenzen von Stämmen aus dem ST95-Komplex untersucht. Die Stämme sind in Abb. 21 aufgeführt (großer Rahmen in Abb. 21 sowie NMEC-Stamm C5). Es

wurden die synonymen ( $D_s$ ) und nicht-synonymen Mutationsraten ( $D_n$ ) bestimmt und anschließend das  $D_n/D_s$ -Verhältnis berechnet.

Wie bereits beschrieben, zeigt dieses Verhältnis die Qualität der wirkenden Selektion an (s. 2.6.3):

- ◇  $D_n/D_s > 1 \rightarrow$  positive Selektion
- ◇  $D_n/D_s \sim 1 \rightarrow$  keine Selektion
- ◇  $D_n/D_s < 1 \rightarrow$  neutrale Selektion

In die vorliegende Untersuchung waren, wie erwähnt, sieben Mitglieder des ST95-Komplexes einbezogen (ST95, ST140, ST370, ST390, ST416, ST417 and ST418). Für die zusammengeführten MLST-Sequenzen innerhalb des ST95-Komplexes konnte eine durchschnittliche Anzahl von  $1,43 \pm 0,58$  synonymen Mutationen gefunden werden. Demgegenüber standen  $0,29 \pm 0,45$  nicht-synonyme Mutationen. Die Mutationsraten sind jeweils abhängig von der Anzahl der Nukleotidpositionen, an denen synonyme bzw. nicht-synonyme Mutationen möglich sind, und betragen für  $D_s = 0,0017$  und  $D_n = 0,0001$ . Das  $D_n/D_s$ -Verhältnis der Housekeeping-Gene der MLST betrug 0,0588 und zeigte somit neutrale Selektion an.

Entsprechend wurden die 21 *ibeA*-Sequenzen der Stämme untersucht. Wie im Rahmen der *ibeA*-Sequenzanalyse bereits festgestellt, kommen innerhalb des ST95-Komplexes nur zwei unterschiedliche *ibeA*-SLSTs vor. Einer lediglich im NMEC-Stamm C5 und der andere in den restlichen 20 Stämmen. Die beiden *ibeA*-SLSTs unterscheiden sich in zwei Nukleotidpositionen, wobei eine synonym und die andere nicht-synonym ist. Daraus ergaben sich eine  $D_s$  von 0,0032 und eine  $D_n$  von 0,001. Das resultierende  $D_n/D_s$ -Verhältnis betrug 0,3125 und zeigte ebenfalls neutrale Selektion an.

Eine zur Bestätigung der Ergebnisse durchgeführte SLAC-Analyse (single likelihood ancestor counting) der *ibeA*-Sequenzen ergab ebenfalls das Vorliegen von neutraler Selektion ( $D_n/D_s$ -Verhältnis: 0,1181). Demzufolge konnten auch keine Nukleotidpositionen identifiziert werden, die unter positiver Selektion stehen.

## 5. DISKUSSION

### 5.1. Polymorphismus des GimA-Insertionslokus

Für „blood-borne pathogens“, d. h. solche, deren Verteilung im Wirtsorganismus über das Blut erfolgt, stellt die Translokation der Blut-Hirn-Schranke einen wesentlichen Schritt in der Meningitis-Pathogenese dar (Leib und Tauber, 1999). Auch Meningitiden, die durch *E. coli* hervorgerufen werden, setzen eine solche hämatogene Streuung des Erregers voraus. Folglich ist auch für *E. coli* die Blut-Hirn-Schranke eine wichtige Barriere, die für eine erfolgreiche Etablierung der Erkrankung überwunden werden muss (Dietzman *et al.*, 1974; Kim, 2003). Eine Vielzahl an Untersuchungen mit dem Ziel, Faktoren bzw. Gene zu identifizieren, die an der Translokation dieser Barriere beteiligt sind, wurden am NMEC-Stamm RS218 (O18:H1:H7) (Achtman *et al.*, 1983) durchgeführt (Hoffman *et al.*, 2000; Huang *et al.*, 1999; Huang *et al.*, 2001b; Huang *et al.*, 1995; Prasadarao *et al.*, 1996; Wang *et al.*, 1999; Wang und Kim, 2002). Der Serotyp O18:K1:H7 ist der weltweit meist verbreitete NMEC-Serotyp und wird daher auch als Archetyp bezeichnet (Achtman *et al.*, 1983; Korhonen *et al.*, 1985; Plainvert *et al.*, 2007). Der Stamm RS218 wurde 1974 in Kalifornien bei einem Neugeborenen mit Meningitis isoliert (Achtman *et al.*, 1983) und dient seither als Prototyp für NMEC (Neugeborenenmeningitis-assoziierte *E. coli*).

Für *ibeA* (invasion of brain endothelium) konnte eine Beteiligung an der Überwindung der Blut-Hirn-Schranke, d. h. an der Pathogenese der Neugeborenen-Meningitis, im Zellkultur- und Rattenmodell belegt werden (Huang *et al.*, 2001b; Huang *et al.*, 1995). *IbeA* ist Teil einer genetischen Insel (GimA – genomic island of meningitic *E. coli* containing *ibeA*), die sich aus vier Operons (GimA1-4) mit unterschiedlichen biologischen Funktionen zusammensetzt. Im Stamm RS218 ist GimA im Chromosom zwischen den Genen *yjiD* und *yjiE* lokalisiert (Huang *et al.*, 2001a), weshalb dieser Ort auch als Insertionslokus bezeichnet wird.

Ziel der vorliegenden Arbeit war es, GimA eingehender zu charakterisieren. Besonderes Gewicht sollte dabei auf die Zusammensetzung von GimA, d. h. die strukturelle Kontinuität sowie die Verbreitung der Insel innerhalb einer gut definierten Auswahl an ExPEC (extraintestinal-pathogene *E. coli*) und weiteren *E. coli*-Isolaten von klinisch gesunden Menschen und Tieren gelegt werden.

Die Untersuchung des Insertionslokus ergab überraschender Weise drei verschiedene Größenvarianten:

- ◇ GimA+ → GimA liegt vollständig vor (20,3 kb) (ExPEC: RS218, APECO1 und UTI89)
- ◇ GimA-remnant → nur ein Restfragment von GimA (3'-Ende von *pptE*, dem ersten ORF in GimA1) liegt vor (342 bp) (ExPEC: CFT073, 536 und F11)
- ◇ GimA- → der Insertionslokus ist vollständig frei von GimA (nicht ExPEC: *E. coli* K-12 MG1655)

In Klammern sind jeweils die Stämme angegeben, die die entsprechende Variante tragen und deren Chromosom zum Zeitpunkt der Promotion vollständig sequenzanalysiert vorlag. Demnach findet sich das vollständige GimA sowie das Restfragment GimA-remnant in ExPEC, wohingegen die Variante GimA- in nicht-ExPEC, jedoch wie vorliegende Untersuchungen gezeigt haben, auch in ExPEC vorkommt.

Der Polymorphismus des Insertionslokus hat bisher noch keinen Eingang in die Literatur gefunden.

### Zusammenfassung

Die Untersuchung des Insertionslokus von GimA zwischen den Genen *yjiD* und *yjiE* ergab, dass neben einem vollständig vorhandenen GimA (20,3 kb) (GimA+) und einem Lokus, der keine GimA-assoziierten Strukturen enthält (GimA-), eine dritte Variante vorkommt, die einen 342 bp großen Rest des *pptE*-Gens enthält (GimA-remnant).

## 5.2. Assoziation von GimA mit Phylogruppe B2

Wie bereits erläutert (s. 2.4.1), sind genomische Inseln, zu denen unter anderen auch Pathogenitätsinseln (PAI) zählen, Bereiche im Chromosom von bakteriellen Mikroorganismen, die sich durch eine Reihe von Eigenschaften vom restlichen Chromosom unterscheiden. Diese Eigenschaften bilden Kriterien, die genetische Inseln charakterisieren (Hacker *et al.*, 1997; Hacker und Kaper, 2000; Hentschel und Hacker, 2001).

Genetische Inseln:

- ◇ beinhalten Virulenz-assoziierte Gene (VAGs) und kommen bei vielen Vertretern innerhalb einer Spezies vor, können aber bei nah verwandten Stämmen fehlen; sie fehlen in apathogenen Stämmen und kommen in Pathogenen vor
- ◇ sind relativ groß (üblicherweise  $\geq 30$  kb)
- ◇ unterscheiden sich im G+C-Gehalt (Guanin- und Cytosin-Gehalt) vom restlichen Chromosom
- ◇ sind häufig flankiert von besonderen Sequenzen (direct repeats) und tragen oft Gene, die für genetische Mobilität kodieren (z. B. Phagengene, Insertions-Elemente (IS), Integrasen, Transposasen) und sind in der Regel mit tRNA-Genen assoziiert

### 5.2.1. GimA – eine intrinsische chromosomale Region

Derzeit sind in öffentlich zugänglichen Datenbanken drei vollständige GimA-Sequenzen inklusive flankierendem Bereich verfügbar. In der Umgebung von GimA der drei Stämme UTI89, RS218 und A0 34/86 (O83:K24:H31) (Hejnova *et al.*, 2005), d. h. 10 kb up- und downstream, sind keine der genannten mit Mobilität assoziierten Gene vorhanden. Ebenso fehlen, wie bereits in der Erstbeschreibung publiziert, Hinweise für direct repeats. Die nächstgelegenen tRNA-Gene sind *leuX* (141 kb bzw. 22 kb upstream bei UTI89 bzw. APEC O1) und *leuQ* (44 kb downstream in beiden Stämmen). Eine Assoziation von GimA mit diesen tRNA-Genen kann aufgrund der großen Distanz praktisch ausgeschlossen werden.

Jedoch werden andere der genannten Eigenschaften von genetischen Inseln zumindest teilweise erfüllt. Beispielsweise fehlt GimA in apathogenen Stämmen (Huang *et al.*, 2001a), kommt aber in Pathogenen vor (Bonacorsi *et al.*, 2003). Weiterhin konnte ein geringerer G+C-Gehalt (46,2%) im Vergleich mit dem restlichen Genom von NMEC-Stamm RS218 (50,8%) (Huang *et al.*, 2001a) und MG1655 (50,7%) (Blattner *et al.*, 1997) gefunden werden. Im Vergleich mit anderen PAIs ist der Unterschied im G+C-Gehalt zwischen GimA und restlichem Genom jedoch moderat (Hacker *et al.*, 1997). Zudem existieren beispielsweise im Chromosom des *E. coli* K-12 Stamms MG1655 Bereiche, die, wie GimA, einen geringeren G+C-Gehalt aufweisen, ohne eine genetische Insel zu sein (Blattner *et al.*, 1997).

Die Literaturdaten sowie solche, die aus den verfügbaren Sequenzdaten generiert werden konnten, belegen, dass GimA nur wenige Eigenschaften aufweist, die bei den klassischen PAIs vorkommen. Eine weitergehende Untersuchung erschien daher angezeigt.

Durch die Untersuchung des Insertionslokus konnte zunächst gezeigt werden, dass die Verbreitung von GimA auf Stämme eines bestimmten phylogenetischen Hintergrundes beschränkt ist. Abb. 18 verdeutlicht, dass GimA strikt mit der Phylogruppe B2 assoziiert ist. Solch ein Verteilungsmuster deutet eher auf eine vertikale Verbreitung von GimA hin, d. h. einer Weitergabe der Erbinformationen von einer Generation zur nächsten, als auf horizontalen Gentransfer (HGT). Eine derartige vertikale Verbreitung konnte beispielsweise für das *bgl*-Operon gezeigt werden, dessen Genprodukte für die Aufnahme und Hydrolyse von aryl- $\beta$ -(D)-Glucosiden von Bedeutung sind (Sankar *et al.*, 2009).

Weiterhin ergab die Untersuchung, dass GimA stets zwischen den beiden Genen *yjiD* und *yjiE* liegt. Eine Beschränkung auf nur einen Insertionslokus ist jedoch für genetische Inseln, die via HGT verbreitet werden, unüblich. Beispielsweise kann der „Locus of Enterocyte Effacement“ (LEE) in drei verschiedene tRNA-Gene inserieren (*selC*, *pheU* und *pheV*) (Jores *et al.*, 2001; Jores *et al.*, 2004).

HGT von genetischen Inseln ist in der Regel gekennzeichnet durch das Vorkommen mehrerer Varianten der Insel innerhalb eines MLST-basierten Phylotyps (wie beispielsweise

im ST95-Komplex), die sich in ihrer Nukleotidsequenz und teilweise auch hinsichtlich der Funktionalität der einzelnen Gene unterscheiden (Tartof *et al.*, 2007; Weissman *et al.*, 2006). Das Vorkommen solcher Varianten erklärt sich dadurch, dass die Insel, ebenso wie das Chromosom, einer langsam fortschreitenden Entwicklung aufgrund von Punktmutationen (SNPs - Single-Nukleotid-Polymorphismus) unterliegt. Als Folge der Akkumulation von Mutationen entstehen im Laufe der Zeit unterschiedliche Abkömmlinge der Ur-Insel, die sich in ihrer Nukleotidsequenz unterscheiden. Die Unterschiede in der DNS-Sequenz können im Hinblick auf die Funktionsfähigkeit des betroffenen Gens sowohl folgenlos bleiben, als auch zum Funktionsverlust bzw. Verbesserung der Funktionalität gegenüber des unveränderten Gens führen (Chakravarti, 2001; Stoneking, 2001).

### Zusammenfassung

GimA erfüllt nur wenige Kriterien für PAIs, v. a. fehlen Elemente, die die Mobilität einer genetischen Insel vermitteln. GimA ist strikt mit der Phylogruppe B2 assoziiert, was auf eine vertikale Verbreitung hinweist. Ebenfalls für eine vertikale Verbreitung spricht die konstante Lokalisation zwischen *yjiD* und *yjiE*.

### 5.2.2. Entwicklung von *ibeA*

Die Mobilität einer Insel ermöglicht die Insertion von mehreren unterschiedlichen Inselvarianten in einen Phylotyp (z. B. einen ST-Komplex). Ähnlich der MLST können Varianten einer Insel durch eine SNP-Analyse in Form einer Single-Locus-Sequenz-Typisierung (SLST) von Teilen der Insel (z. B. einem Gen) identifiziert werden. Für verschiedene Loci von genetischen Inseln konnte das Auftreten von unterschiedlichen Varianten in einem Phylotyp bereits gezeigt werden. Für die Adhäsine *fimA* und *fimH* sowie *papG* und *papA*, die in der Pathogenese von Harnwegsinfektionen von Bedeutung sind (s. 2.3.1.), konnte mittels SLST eine Mehrfachinsertion der entsprechenden Insel in den ST95-Komplex identifiziert werden (Weissman *et al.*, 2006). Ähnliches konnte für die High Pathogenicity Island (HPI) der *Enterobacteriaceae* gezeigt werden (Schubert *et al.*, 2009). Die Variabilität der O- und H-Antigene innerhalb eines Phlotyps basiert ebenso auf unabhängigem HGT der genetischen Determinanten (Milkman *et al.*, 2003; Nelson und Selander, 1994).

An eine Insertion kann sich eine Weiterdiversifizierung durch Punktmutationen oder Rekombinationen anschließen. Varianten und die Produkte einer weitergehenden Diversifizierung lassen sich anhand der Anzahl der unterschiedlichen Nukleotide unterscheiden. Eine Diversifizierung umfasst nur wenige SNPs, wohingegen einzelne Varianten einen deutlich größeren Unterschied aufweisen. Für *fimH* und *papG* konnte z. B. gezeigt werden, dass eine Diversifizierung im Sinne einer Pathoadaptation stattgefunden

hat, die sich in der Ausbildung verschiedener Allel-Varianten niederschlägt (*papG* I, II und III). Eine positive Selektion von auftretenden Punktmutationen, die eine verbesserte Anpassung an die Rezeptoren der Wirtszelle von *fimH* und *papG* darstellen, bildet hierfür die Grundlage (Weissman *et al.*, 2006). Nur Mutationen, die zu einer Änderung der Aminosäuresequenz führen, sogenannte nicht-synonyme Mutationen, können Gegenstand der positiven Selektion sein. Demgegenüber stehen die synonymen Mutationen. Das Verhältnis von nicht-synonymen ( $D_n$ ) und synonymen ( $D_s$ ) Mutationsraten gibt Auskunft über die Art der Selektion. Positive Selektion liegt vor, wenn  $D_n/D_s > 1$  ist, d. h. mehr nicht-synonyme als synonyme Mutationen auftreten, wie es der Fall bei *fimH* und *papG* ist.  $D_n/D_s < 1$  zeigt eine neutrale Selektion an (Hughes und Nei, 1989; Tanaka und Nei, 1989).

Für die Untersuchung des *ibeA* wurden Stämme des ST95-Komplexes ausgewählt, da zahlreiche Untersuchungen ergaben, dass dieser ST-Komplex innerhalb der ExPEC-Gruppe von herausragender Bedeutung ist (Johnson *et al.*, 2008; Mora *et al.*, 2009). Eine ebenso herausragende Rolle konnte für *ibeA* gezeigt werden. Es ist, wie beschrieben, im Rahmen der Pathogenese der *E. coli*-bedingten Meningitis der Neugeborenen an der Überwindung der Blut-Hirn-Schranke beteiligt (Huang *et al.*, 2001b; Huang *et al.*, 1995). Darüber hinaus ist es in die Pathogenese der systemischen *E. coli*-Infektion des Geflügels involviert (Germon *et al.*, 2005). Daher wurde *ibeA* als das biologisch interessanteste Gen von GimA für diesen Teil der Untersuchung ausgewählt.

Zur Erfüllung seiner biologischen Funktion als Ligand befindet sich das IbeA-Protein auf der Oberfläche von *E. coli*. Eine Folge davon könnte eine Interaktion des Proteins mit der Wirtsabwehr sein, die einen Selektionsdruck auf diese Struktur erzeugt. Dieser Selektionsdruck könnte in der Folge zu einer forcierten Diversifizierung von *ibeA* im Sinne einer Pathoadaptation führen. Maßstab zur Bewertung des Diversifizierungsgrades ist der Vergleich mit der Diversität der Housekeeping-Gene, die mittels MLST (Multi-Lokus-Sequenztypisierung) identifiziert wurde.

Eine SLST des *ibeA*-Gens des ST95-Komplexes sowie der drei weiteren STs (Sequenztyp), die durch mehr als ein Isolat vertreten waren (STs 135, 355 und 372), ergab, dass diese jeweils eine identische *ibeA*-Sequenz aufweisen (100% Identität). Das Fehlen von Sequenzvariabilität von *ibeA* in diesen STs (im ST95-Komplex mit der Ausnahme des NMEC-Stamms C5) ist vergleichbar mit dem praktisch monophyletischen Erscheinungsbild dieser STs bzw. dieses ST-Komplexes in der MLST. Solch eine Kongruenz zwischen den sieben zusammengeführten Housekeeping-Genen der MLST einerseits und den *ibeA*-Sequenzen andererseits spricht deutlich gegen eine Mehrfachinsertion von *ibeA* in den ST95-Komplex bzw. die anderen STs, wie es bei einer Verbreitung via horizontalem Gentransfer zu erwarten wäre (Weissman *et al.*, 2006). Deshalb sind die einzelnen *ibeA*-Sequenzen strikt mit einem bestimmten phylogenetischen Hintergrund, d. h. ST bzw. ST-

Komplex assoziiert. Anders ausgedrückt: Vertreter eines identischen ST tragen (fast) immer eine identische *ibeA*-Sequenz. Dies zeigt eine analoge Entwicklung des *ibeA*-Gens mit den Housekeeping-Genen der MLST.

Wie beschrieben, treten innerhalb der 21 untersuchten Vertreter des ST95-Komplexes zwei verschiedene *ibeA*-Sequenzen auf. 20 Isolate weisen eine identische *ibeA*-Sequenz auf. Lediglich das Isolat C5 unterscheidet sich davon, allerdings nur in zwei Positionen. Die zwei SNPs führen zu jeweils einer synonymen sowie einer nicht-synonymen Substitution. Eine Bestimmung des Dn/Ds-Verhältnisses mit zwei verschiedenen Methoden ergab eine Ratio deutlich kleiner als eins. Triebfeder für die Entwicklung des *ibeA*-Gens ist folglich keine positive Selektion. Eine Entwicklung im Sinne einer Pathoadaptation kann ausgeschlossen werden. Die Sequenzunterschiede sind somit das Produkt neutraler Selektion.

#### **Zusammenfassung**

Für das *ibeA*-Gen ergeben sich keine Hinweise auf Mehrfachinsertion in den ST95-Komplex noch auf eine Pathoadaptation durch positive Selektion. Das Gen entwickelte sich aufgrund von neutraler Selektion.

#### **5.2.3. GimA als originärer Bestandteil der Phylogruppe B2**

Für die folgenden Betrachtungen wird angenommen, dass *ibeA* stellvertretend für das gesamte GimA steht, da innerhalb einer genetischen Einheit eine einheitliche Entwicklung vorausgesetzt werden kann, zumal die Insel mit 20,3 kb relativ klein ist.

Das Fehlen von (i) einer Mehrzahl der Charakteristika für genetische Inseln und (ii) unterschiedlichen *ibeA*-SLSTs sowie (iii) die zum „core genome“ analoge Phylogenie von *ibeA* lassen GimA als einen originären Bestandteil der Phylogruppe B2 erscheinen. Die geschilderten Ergebnisse sind für eine mobile Insel eher unwahrscheinlich. Offenbar ist GimA als ein Bestandteil des „core genome“ stabil in das *E. coli* Chromosom integriert. Dafür spricht, dass in der direkten Umgebung der Insel keine genetischen Elemente aufzufinden sind, die eine Mobilität vermitteln könnten.

Alternativ besteht die Möglichkeit, dass GimA eine genetische Insel ist, die in einem einmaligen Insertionsereignis in die Phylogruppe B2 gelangt ist. Im Anschluss an diese Insertion erfolgte eine stabile Integration in das Chromosom durch den Verlust der Mobilitätselemente. Neben einer Deletion kann auch eine Inaktivierung dieser Elemente (z. B. durch Punktmutationen) zur Immobilisierung der Insel geführt haben. Dies ist für die PAIs von *Salmonella* und von intestinalen *E. coli* beschrieben (Hentschel und Hacker, 2001). Sofern die Mobilitätselemente nur aus einer kurzen Nukleotidsequenz bestehen, wie bei „direct repeats“, können sogar bereits wenige SNPs das Element so stark maskieren, dass es nicht mehr als Mobilitätselement erkannt wird (Dobrindt *et al.*, 2004). Ein Verlust der

Mobilitätselemente könnte die stabile Integration von GimA erklären, nicht aber die einmalige Insertion in die Phylogruppe B2.

Wie oben bereits erläutert, ergab die Sequenzanalyse von *ibeA* weder Hinweise auf eine auf positiver Selektion basierende Entwicklung (Pathoadaptation), noch auf wiederholte Insertion in den ST95-Komplex oder die STs 135, 355 und 372. Im Gegenteil zeigt sich eine Kongruenz in den *ibeA*-Sequenzdaten und den Housekeeping-Genen aus der MLST-Analyse. Da genetische Inseln i.d.R. durch die enthaltenen Gene einen Selektionsvorteil für den Empfängerstamm bieten, basiert folglich auch ihre Verbreitung auf positiver Selektion (Dobrindt *et al.*, 2004). Solch eine positive Selektion zeigt sich zum einen durch das Auftreten verschiedener Varianten der Insel in Form unterschiedlicher SLSTs, d. h. Mehrfachinsertion (s. 5.2.2), und zum anderen durch das Auftreten von Pathoadaptation innerhalb eines Phylotyps. Beides ist jedoch für die betrachteten *E. coli*-Stämme nicht nachweisbar. Entsprechend den Arbeiten von Weissman *et al.* (2006) kann daraus abgeleitet werden, dass das gesamte GimA originärer Bestandteil der Phylogruppe B2 ist. Ein einmaliger HGT sowie das Fehlen von Pathoadaptation ist bei genetischen Inseln mehr als unwahrscheinlich.

Wie erläutert, deutet die ausschließliche Assoziation mit der phylogenetischen Gruppe B2 (s. Abb. 18) auf eine stabile Integration von GimA hin. In einer großen Studie basierend auf 20 vollständig sequenzanalysierten kommensalischen und pathogenen *E. coli*-Stämmen konstruierten Touchon *et al.* eine Phylogenie der *E. coli*-Population. Durch die Verwendung des gesamten Chromosoms wurden die Unzulänglichkeiten der herkömmlichen phylogenetischen Methoden (sequenzbasiert und nicht-sequenzbasiert) bei der Erzeugung von Stammbäumen umgangen (s. 2.5) (Didelot und Falush, 2007; Wirth *et al.*, 2006). Die errechnete Genealogie zeigt eine erste Verzweigung in der *E. coli*-Population, die auf der einen Seite zu der heute noch existierenden Gruppe B2 und auf der anderen Seite zu einem Vorläufer der verbleibenden Phylogruppen (A, B1 und D) führte (Touchon *et al.*, 2009). Das Ergebnis, dass GimA stets mit der Phylogruppe B2 assoziiert auftritt, unterstützt die Annahme, dass GimA ein originärer Bestandteil der B2-Gruppe ist. Bereits in der Vorläufergruppe der späteren B2-Gruppe war GimA integriert und konnte sich folglich in der gesamten B2-Gruppe verbreiten.

Weiterhin zeigen Touchon *et al.* (2009), dass in der basalen Gruppe innerhalb der *E. coli*-Population, die sich bei dieser ersten Verzweigung bildet und sich zu der heutigen Phylogruppe B2 entwickelt, auch Vorläufer einer weiteren Phylogruppe enthalten sind. Diese entwickeln sich zu einer Untergruppe der Phylogruppe D (Touchon *et al.*, 2009). Die Phylogruppe D ist somit uneinheitlich. Eine Untergruppe ist Teil der basalen Gruppe und eine andere gehört zu den Vorläufern der Phylogruppen A, B1 und D. Diese Instabilität der Phylogruppe D wurde bereits zuvor von anderen Autoren beschrieben (Wirth *et al.*, 2006).

Im Rahmen der Untersuchung des Insertionslokus von GimA waren Reste der Insel, das GimA-remnant, nicht nur in STs zu finden, die der Phylogruppe B2 zugeordnet waren, sondern auch im ST117. Dieser ST ist jedoch nicht der Gruppe D zugeordnet, sondern der Hybridgruppe ABD. Die Hybridgruppe ABD ist aufgrund ihrer Funktion als Sammelbecken für alle STs, die im Rahmen der Structure-Analyse keiner anderen Gruppe zugeordnet werden konnten, sehr heterogen. Der ST117 scheint innerhalb der Gruppe ABD eine Sonderstellung einzunehmen. Er ist deutlich von anderen Vertretern der Gruppe ABD abgegrenzt. Werden jedoch die Einzelwahrscheinlichkeiten ( $q$ ) für die Abstammung aus den Phylogruppen der Structure-Analyse betrachtet, so zeigt sich, dass die höchsten Werte für eine Zugehörigkeit zu B2 und D zu finden sind (jeweils  $q \geq 0,4$ ). Möglicherweise sind die Vorfahren des ST117 ebenfalls Teil der basalen Gruppe gewesen. Dies deutet sich bereits in der Clonal Frame-Analyse an, in der der ST117 sich von der Linie abspaltet, die sich im weiteren Verlauf in die Vertreter der Gruppe B2 aufspaltet (s. Abb. 19). Auch die oben dargelegte Annahme, dass GimA bereits in den Vorläufer der B2-Gruppe integriert war, würde eine Zugehörigkeit des ST117 zur basalen Gruppe unterstützen. Das Vorkommen von GimA-Strukturen im ST117 ließe sich durch die Abstammung von einem gemeinsamen Vorfahren, sprich der basalen Gruppe, mit der B2-Gruppe erklären.

Künftige Untersuchungen sollten daher klären, ob der ST117 Teil der basalen Gruppe der *E. coli*-Phylogenie ist.

Die Entwicklung von GimA wird durch die gleichen Mechanismen bewirkt, die auch für die Evolution des „core genome“ verantwortlich sind, d. h. neutrale Selektion. Das Fehlen sowohl von Anzeichen für Mehrfachintegration als auch von unterschiedlichen SLSTs des *ibeA*-Gens im Sinne einer Pathoadaptation innerhalb einer Phylogruppe ist somit nur konsequent. Solch eine Co-Evolution eines einzelnen Lokus mit dem „core genome“ konnte kürzlich für das *bgl*-Operon gezeigt werden. Ähnlich GimA, soll auch das *bgl*-Operon bereits intrinsischer Bestandteil eines Vorfahren der heutigen *E. coli* sein (Sankar *et al.*, 2009).

Um die Zugehörigkeit zum „core genome“ zu unterstreichen, sollte diese genomische Region nicht als genetische Insel, sondern künftig besser als GimA-Lokus angesprochen werden.

Wie zu Beginn dieses Kapitels angemerkt, basieren die hier vorgestellten Betrachtungen auf der Annahme, dass *ibeA* das gesamte GimA reflektiert. Zur Verifizierung der These, dass *ibeA* als Marker für GimA dienen kann, sind weitere Untersuchungen erforderlich. Es sollten weitere Gene und die Ränder von GimA sequenzanalysiert und entsprechend *ibeA* ausgewertet werden. Weiterhin ist eine Ausweitung der untersuchten Population und ggf. auch die Sequenzanalyse des gesamten GimA sinnvoll.

Die sechs *pptE*-Sequenzen, die im Rahmen der *in silico*-Analyse der online verfügbaren *E. coli*-Sequenzen untersucht wurden (s. 4.1.1.3), deuten jedoch bereits ein *ibeA*-analoges Verhalten von *pptE* bzw. dem Restfragment an. Die drei *pptE*-Restsequenzen der Stämme,

die zum ST95 gehören (APEC O1, RS218 und UTI89) sind zu 100% identisch. Dies unterstützt die These, dass *ibeA* als Marker für GimA dienen kann und folglich auch die oben angeführte Argumentation.

### **Zusammenfassung**

GimA ist originärer Bestandteil der Phylogruppe B2. Die Entwicklung dieser Region erfolgt durch die gleichen Mechanismen, die auch auf das restliche Genom wirken, d. h. neutrale Selektion. Die chromosomale Region, in der GimA lokalisiert ist, sollte daher als GimA-Lokus angesprochen werden. Die Phylogruppe B2 ist aus der basalen Gruppe in der *E. coli*-Population hervorgegangen. Wahrscheinlich ist auch der ST117 Teil der basalen Gruppe.

## **5.3. Entwicklung von GimA**

Das Vorkommen von Größenunterschieden eines Lokus in verschiedenen Stämmen wird als Längenpolymorphismus angesprochen. Solche Polymorphismen kommen sowohl im Chromosom von *E. coli* [z. B. *mutS-rpoS* Region (Culham und Wood, 2000)] als auch in PAIs [z. B. LEE (Locus of enterocyte effacement) (Jores *et al.*, 2001; Jores *et al.*, 2004)] vor. Im Falle des LEE werden verschiedene Szenarien für die Entstehung solcher Polymorphismen diskutiert. Einerseits können sie durch wiederholten Verlust und (Wieder-) Aufnahme von Genen (sog. Patchwork-Modell), andererseits sukzessiv durch Deletion von Genen im Anschluss an eine initiale Aufnahme eines vollständigen (Ur-) LEE (Sandner *et al.*, 2001) entstanden sein.

Auch von anderen Spezies sind Längenpolymorphismen bekannt. Ein Beispiel für Gram-positive Bakterien ist die *staphylococcal cassette chromosome mec* (SCCmec) von *Staphylococcus aureus* (Chongtrakool *et al.*, 2006). Auch bei Archaeobakterien gibt es solche Polymorphismen, beispielsweise bei *Prochlorococcus* (Coleman *et al.*, 2006).

Im Rahmen der vorliegenden Untersuchung konnte solch ein Längenpolymorphismus des GimA-Lokus bei verschiedenen ExPEC und weiteren *E. coli*-Isolaten von klinisch gesunden Menschen und Tieren identifiziert werden.

Auch der Polymorphismus des GimA-Lokus (drei Varianten: GimA+, GimA-remnant und GimA-) kann auf den oben genannten Wegen entstanden sein. GimA-originäre Sequenzen, d. h. das GimA-remnant, können zum einen durch Insertion in das *E. coli*-Chromosom gelangt und zum anderen durch Deletion aus dem Chromosom entfernt worden sein.

Aufgrund der Ergebnislage, dass GimA zu einem bestimmten frühen Zeitpunkt stabil in das *E. coli*-Chromosom integriert wurde, sind Deletionsereignisse und nicht wiederholte Insertionsereignisse der wahrscheinlichere Weg für die Entstehung der drei Lokus-Varianten. Anderenfalls müsste ein Szenario abgelaufen sein, dass unabhängige Insertionen immer der gleichen GimA-Variante in ein und denselben Phylotyp beinhaltet, d. h. so etwas wie eine

phylotypspezifische GimA-Insertion. Das gleiche müsste für das GimA-remnant gelten. Besonders das Fehlen von Mobilität vermittelnden Elementen sowie von Anzeichen für Mehrfachinsertion lassen dieses Szenario mehr als unwahrscheinlich erscheinen.

### 5.3.1. Reduktive Evolution

Die Aufnahme genetischen Materials wird häufig ausgeglichen durch den Verlust von anderen Genen und kann unter Umständen einen Selektionsvorteil bedeuten. So kann eine „en bloc“-Aufnahme mit anschließendem Verlust von Genen als ein Mechanismus der Optimierung des Genoms angesehen werden. Solch eine Optimierung, auch als reduktive Evolution bezeichnet, spiegelt den „lifestyle“ des Stamms wider, d. h. die Interaktion von Erreger, Wirt (der das Habitat für den Erreger bildet) und Umwelt (beeinflusst beide) (Dobrindt *et al.*, 2004). Die Reduktion des Genoms mittels Deletion kommt bei vielen Spezies vor. Das Spektrum reicht von Endosymbionten (beispielsweise *Buchnera amphidicola* und *Blochmannia floridanus*) (Tamames *et al.*, 2007) zu Phyto- (z. B. *Phytoplasma asteris*) (Andersson und Kurland, 1998; Oshima *et al.*, 2004) und Zoopathogenen (beispielsweise *Mycoplasma spp.*) (Meseguer *et al.*, 2003). Als Beispiel sei *Mycobacterium ulcerans* genannt. Dieses humanpathogene Bakterium entwickelte sich durch reduktive Evolution aus seinem apathogenen Vorfahren *Mycobacterium marinum* zu einem an eine Nische adaptierten Spezialisten (Stinear *et al.*, 2007).

### 5.3.2. „Change of function“ Mechanismus

Die Ergebnisse der vorliegenden Arbeit deuten darauf hin, dass eine solche reduktive Evolution bei GimA vorliegt. Das GimA-remnant verbleibt im Chromosom des betreffenden Stamms, während der Rest von GimA in einem Deletionsereignis entfernt wird. Weiterhin kann auch das GimA-remnant deletiert werden, so dass die Variante GimA- entsteht. Neben diesem sukzessiven Verlust besteht die Möglichkeit, dass GimA in einem einzigen Schritt vollständig entfernt wird und die Variante GimA- direkt entsteht. Sokurenko *et al.* (1999) beschreiben einen Mechanismus, der es Bakterien erlaubt, ihre Pathogenität durch Eliminierung von vorhandenen Genen zu steigern. Dieser Mechanismus wird angeschoben, wenn ein Bakterium in ein neues Habitat gelangt. Zur Eliminierung von vorhandenen Genen kommt es, wenn die neue Nische andere phänotypische Leistungen als die bisherige erfordert und ein entsprechend hoher Selektionsdruck in dieser Richtung vorhanden ist (change of function) (Sokurenko *et al.*, 1999). Im Falle von GimA werden in einigen Stämmen der Untersuchung weite Teile der Insel eliminiert, so dass nur ein Fragment des *pptE*-Gens übrigbleibt. Dieser Vorgang tritt v. a. bei Stämmen auf, die aus dem Harnapparat isoliert wurden. Möglicherweise scheinen die GimA-Genprodukte mit Ausnahme des GimA-remnant in den Harnwegen nicht von Bedeutung oder sogar von Nachteil zu sein und gehen daher in einem Optimierungsprozess des Genoms verloren. Die in der vorliegenden

Untersuchung bereits zeigte signifikante Assoziation des GimA-remnant mit UPEC (uropathogene *E. coli*) spricht dafür.

In Abschnitt 5.4 wird auf die biologische Bedeutung der einzelnen GimA-ORFs in den unterschiedlichen Habitaten näher eingegangen.

Interessanterweise ist dieses Phänomen fast ausschließlich mit der Phylogruppe B2 assoziiert. 134 (91,8%) der GimA+-und GimA-remnant-tragenden Stämme gehören dieser Phylogruppe an. Zusätzlich kommen einige Stämme (n=12), die das GimA-remnant tragen, im ST117 (Hybridgruppe ABD) vor.

Solch eine Habitat-abhängige Erosion konnte bereits für das *bgl*-Operon gezeigt werden. Ebenso wie beim GimA-Lokus finden sich vollständige *bgl*-Operons in der Phylogruppe B2 und erodierte Varianten in den anderen Gruppen (A, B1 und D) (Sankar *et al.*, 2009). Wie in der vorliegenden Untersuchung, nehmen diese Autoren Selektionseffekte als Ursache dafür an (Sankar *et al.*, 2009).

Hier zeigt sich, dass weiterer Forschungsbedarf besteht, um zu verifizieren, dass der Polymorphismus im GimA-Lokus durch einen „change of function“-Mechanismus bewirkt wird.

#### **Zusammenfassung**

Eine Optimierung des Genoms extraintestinal-pathogener *E. coli* durch Deletion als Folge der Erschließung neuer Habitate ist wahrscheinlich die Ursache für die Entstehung der Varianten des GimA-Lokus. Zur Klärung sind weitere Untersuchungen erforderlich.

## **5.4. Variantenbildung und Pathovare**

Die Ursachen für die Reduktion von GimA bleiben unklar. Allerdings könnte dies, wie bereits erläutert, eine Anpassung an den „lifestyle“ der einzelnen Stämme darstellen. Der „lifestyle“ wird u. a. bestimmt durch das Habitat, in dem sich der Stamm aufhält und ist somit auch von Wirt und Umwelt beeinflusst.

### **5.4.1. APEC und NMEC**

Einerseits konnte in der vorliegenden Untersuchung gezeigt werden, dass Stämme, die den Pathovaren APEC (aviär pathogene *E. coli*) und NMEC zugeordnet sind, signifikant häufiger die Variante GimA+ tragen. Das kann ein Hinweis auf die pathogenetische Bedeutung sein, die das vollständige GimA für diese Pathovare hat. Für verschiedene Gene von GimA konnte bereits eine Funktion in der Pathogenese der Meningitis der Neugeborenen sowie der systemischen APEC-Infektion des Geflügels gezeigt werden. Neben *ibeA* (Germon *et al.*, 2005; Huang *et al.*, 2001b; Huang *et al.*, 1995) ist dies auch bei den ebenfalls im Operon GimA4 lokalisierten Genen *ibeT* und *ibeR* der Fall (Chi *et al.*, 2009; Zou *et al.*, 2008). *IbeT* ist

für die koordinierte Adhäsion von *E. coli* an die Zellen des Endothels im Rahmen der Meningitis-Pathogenese von Bedeutung (Zou *et al.*, 2008). IbeR ist an der Translokation der Blut-Hirn-Schranke beteiligt, eine Deletion von *ibeR* führt zu einem attenuierten Phänotyp (Chi *et al.*, 2009). Weiterhin scheint IbeR ein Regulatorprotein zu sein, das in die Stressregulation von *E. coli* in der stationären Phase involviert ist. Die stationäre Phase schließt sich an die exponentielle Wachstumsphase von *E. coli* an und ist durch eine hohe Bakteriendichte und das Sistieren des bakteriellen Wachstums gekennzeichnet. Stress für *E. coli* entsteht dabei durch Nährstoffmangel und/oder die Anhäufung von toxischen Stoffwechselprodukten (Nyström, 2004). Vermittelt durch IbeR können Gene des Pyruvat-Stoffwechsels kontrolliert werden. Wie die Gene der GimA-Operons 1-3 ist *ibeR* folglich in den Kohlenhydratstoffwechsel involviert (Chi *et al.*, 2009). Besonders für die Reaktion auf Stress, wie er in der stationären Phase auftritt, ist dieser Metabolismus von großer Bedeutung (Wang *et al.*, 2001). Die drei weiteren Operons von GimA beinhalten drei bzw. vier ORFs (open reading frame) (*pptE*, *pgdK*, *pdaK* = GimA1; *cgrD*, *cgxT*, *cdlD*, *cniT* = GimA2 und *gclK*, *ghbD*, *gclR* *gclA* = GimA3). Eine funktionelle Annotation der ORFs, basierend auf der Bestimmung von Homologien mit bekannten Proteinen, ergab bei den elf ORFs, dass sie wahrscheinlich auch für Enzyme des Kohlenhydratstoffwechsels sowie des Substrattransports kodieren (Hejnova *et al.*, 2005; Huang *et al.*, 2001a). Die Möglichkeit zu einer Stressantwort wie sie in der stationären Phase von Bedeutung ist, könnte auch einen Beitrag für das Überleben im Habitat „Blut“ leisten. Beispielhaft für die Stressoren, die im Blut auf *E. coli* wirken, seien die bakteriziden Eigenschaften des Serums (Roberts und Phillips, 1983) und die sehr niedrige Konzentration von freiem Eisen ( $\text{Fe}^{3+}$ ) im Blut bzw. Serum (Bullen *et al.*, 2000) genannt. Daher können die anderen GimA-Gene mit ihren putativen Funktionen im Kohlenhydratstoffwechsel und Substrattransport ebenso wie *ibeR*, einen möglichen Anteil an der Stressantwort auf die Bedingungen im Habitat „Blut“ haben.

Die Pathovare APEC und NMEC können schwere systemische Erkrankungen verursachen, die letztlich entweder zu einer Eliminierung des Erregers oder zum Tod des Wirtes führen. Dadurch ist eine Übertragung des Erregers von einem Wirt auf den nächsten im Sinne einer hohen Kontagiosität eher unwahrscheinlich. Die kaum vorhandene Verbreitungstendenz des Erregers (Le Gall *et al.*, 2007) könnte die Ursache für das Fehlen eines Selektionsdrucks auf das *ibeA*-Gen sein, obwohl es durch seine Exposition auf der Oberfläche des Erregers für die Wirtsabwehr potenziell zugänglich wäre.

Wie bereits unter 5.3.2 erwähnt, besteht zur Verifizierung des Einflusses der anderen GimA-Gene weiterer Forschungsbedarf. Funktionelle Untersuchungen, beispielsweise mittels knockout-Mutanten der einzelnen Gene könnten einen Beitrag zur Klärung leisten.

### 5.4.2. UPEC

Stämme, die bei Harnwegsinfekten isoliert wurden, sind signifikant unabhängig von ihrer Herkunft (Mensch/Tier) mit dem Auftreten der Variante GimA-remnant assoziiert. Dies könnte auf eine Verschiebung der pathogenetischen Bedeutung der einzelnen Bestandteile von GimA hindeuten. Für uropathogene Stämme könnte das vollständige GimA einschließlich *ibeRAT* an Wichtigkeit verloren haben, da sich das Habitat „Harnapparat“, in dem die GimA-Komponenten ihre Wirkung entfalten, deutlich von den Bedingungen unterscheidet, die im Blut vorliegen. Folglich könnten in einem Optimierungsprozess des Genoms weite Teile von GimA deletiert worden sein.

Die Möglichkeit für eine Stressantwort, wie sie im Habitat „Blut“ von Bedeutung ist, spielt im Harnapparat wahrscheinlich eine untergeordnete Rolle, da es sich beim Harnapparat um ein nährstoffreiches und mit weit weniger Stressoren für *E. coli* behaftetes Habitat handelt (Oelschlaeger und Fünfstück, 2006; Zdziarski *et al.*, 2008). Lediglich das *pptE*-Fragment ist in einer überdurchschnittlich hohen Zahl von UPEC-Stämmen enthalten. Dies deutet darauf hin, dass das GimA-remnant für diese Stämme von Bedeutung ist. Möglicherweise ist diese kurze Nukleotidsequenz in die Pathogenese von Harnwegsinfekten involviert (s.5.5).

### 5.4.3. Asymptomatische Besiedlung der Harnwege

Neben UPEC kann *E. coli* auch als asymptomatischer Besiedler der Harnwege auftreten (asymptomatic bacteriuria – ABU). Das Keimspektrum, das im Rahmen von ABU isoliert werden kann, umfasst *Enterobacteriaceae*, *Pseudomonas aeruginosa*, *Enterococcus* spp., und Streptokokken der Gruppe B. *E. coli* ist die am häufigsten isolierte ABU-Spezies (Colgan *et al.*, 2006). ABU-Stämme zeichnen sich u. a. gegenüber herkömmlichen UPEC durch eine reduzierte Virulenz aus. Einerseits wird die Abnahme der Virulenz durch eine Unfähigkeit zur Expression vorhandener Virulenzgene und andererseits durch Deletion von Virulenzgenen erreicht, d. h. durch reduktive Evolution. ABU-Stämme scheinen somit von virulenten UPEC abzustammen. Mit der Zeit hätte sich demnach in diesen Stämmen ein attenuierter Phänotyp durchgesetzt, der keine Immunantwort des Wirtes mehr provoziert (Wiles *et al.*, 2008; Zdziarski *et al.*, 2008). Eine asymptomatische Besiedlung der Harnwege bedeutet nicht nur die Erschließung eines nährstoffreichen Habitats, sondern auch die Möglichkeit einer Übertragung von einem Wirt zum nächsten (Zdziarski *et al.*, 2008). In diesem Sinne könnte der vollständige Verlust von GimA eine konsequente Fortführung der Mechanismen der Genomoptimierung sein, die zur Entstehung des GimA-remnant geführt hat. Im Falle von ABU könnte auch die Deletion des GimA-remnant einen Selektionsvorteil bedeuten. Wenn eine Beteiligung des Restfragmentes an der Pathogenese von Harnwegsinfektionen vorausgesetzt wird, fügt sich der Verlust dieser Struktur widerspruchlos in das ABU-Konzept der Vermeidung der Provokation einer Immunreaktion des Wirts ein. Auf diese

Weise ist eine mehr oder weniger dauerhafte Besiedlung des Habitats „Harnapparat“ möglich (Zdziarski *et al.*, 2008).

#### **Zusammenfassung**

Aufgrund unterschiedlicher pathogenetischer Bedeutung von einzelnen GimA-Komponenten in den verschiedenen Habitaten führte die Optimierung des Genoms in der Mehrzahl der Fälle in APEC und NMEC zu einer Beibehaltung des vollständigen GimA, da alle GimA-Komponenten einen Beitrag zum Überleben im Habitat „Blut“ bzw. an der Translokation der Blut-Hirn-Schranke leisten. In UPEC ist eine Reduktion von GimA bis auf das GimA-remnant möglich, da im Habitat „Harnapparat“ eine derartige Stressantwort nicht erforderlich ist. Bei asymptomatischen Besiedlern der Harnwege ist sogar eine vollständige GimA-Reduktion von Vorteil, da dies der Vermeidung einer Immunreaktion des Wirts dient, die zur Eliminierung des Erregers führen könnte.

### **5.5. Weitergehende Charakterisierung des *pptE*-Fragmentes**

Die Identifizierung der Variante GimA-remnant des GimA-Lokus stellt ein besonders interessantes Ergebnis der vorliegenden Untersuchung dar: Einerseits, weil sie bisher noch nicht beschrieben worden ist, und andererseits aufgrund der gezeigten Assoziation mit UPEC und der sich daraus ergebenden Möglichkeit, dass diese Struktur in die Pathogenese von Harnwegsinfektionen involviert ist.

Eine *in silico*-Analyse der in öffentlichen Datenbanken hinterlegten UPEC-Stämme CFT073, 536 und F11 zeigte, dass diese Stämme das *pptE*-Restfragment tragen. Die 342 bp große Sequenz beinhaltet ein Start- und ein Stopkodon und wurde daher in einer open reading frame (ORF)-Analyse mittels der Kodon Software als ein *orf* identifiziert und kodiert somit potenziell für die Aminosäuresequenz eines Proteins. Die Stopkodons des Restfragmentes und des originären *pptE* sind identisch. Als Startkodon des Restfragmentes fungiert ein im vollständigen *pptE* für Methionin kodierendes ATG an Position 2340. Zusätzlich konnte mit Hilfe eines webbasierten „Promotor Prediction Tool“ (BPPROM, verfügbar unter <http://linux1.softberry.com>) upstream des Startkodons eine Bindestelle für einen Transkriptionsfaktor gefunden werden.

Die Assoziation mit den uropathogenen Stämmen in der vorliegenden Untersuchung deutet an, dass das resultierende Genprodukt in die Pathogenese von Harnwegsinfektionen involviert sein könnte. Die oben genannte funktionelle Annotation (s. 5.4.1) ergab, dass *pptE* für eine putative Phosphoenolpyruvat-Protein Phosphoryltransferase kodiert und damit, wie die anderen GimA-Gene (mit Ausnahme von *ibeA* und *ibeT*), in den Kohlenhydratstoffwechsel involviert ist.

Inwieweit das *pptE*-Fragment tatsächlich exprimiert wird, bzw. unter welchen Bedingungen dies stattfindet, bleibt allerdings offen. Ebenfalls ungeklärt bleibt die Funktion dieses möglichen Genproduktes. Ob das Restfragment die für das vollständige *pptE* vermutete Funktion im Kohlenhydratstoffwechsel übernehmen kann, sollte Gegenstand weiterer Untersuchungen sein. Um die Rolle dieser Nukleotidsequenz näher zu charakterisieren, bieten sich Expressionsanalysen ebenso wie Untersuchungen in der Zellkultur und ggf. auch im Tierversuch an.

Neben der Expression des GimA-remnant und einer direkten oder indirekten Einwirkung des Genprodukts auf die Pathogenese von Harnwegsinfektionen, besteht die Möglichkeit, dass diese Nukleotidsequenz an der Regulation anderer Gene beteiligt ist, ohne selbst exprimiert zu werden. Beispielsweise könnte sie als Promotor dienen. Diese Möglichkeit sollte in nachfolgende Untersuchungen eingeschlossen werden.

#### **Zusammenfassung**

Das GimA-remnant tritt gehäuft bei uropathogenen Stämmen auf und könnte folglich in die Pathogenese von Harnwegsinfektionen involviert sein. Das *pptE*-Fragment konnte als ORF identifiziert werden. Außerdem existiert eine Bindungsstelle für einen Transkriptionsfaktor. Zur Klärung der Funktion des *pptE*-Fragmentes sind weitere Untersuchungen erforderlich.

## 6 ZITIERTE LITERATUR

- ACHTMAN, M. (2004)  
Population structure of pathogenic bacteria revisited.  
Int J Med Microbiol **294**(2-3): 67-73
- ACHTMAN, M., MERCER, A., KUSECEK, B., POHL, A., HEUZENROEDER, M.,  
AARONSON, W., SUTTON, A. und SILVER, R. P. (1983)  
Six widespread bacterial clones among *Escherichia coli* K1 isolates.  
Infect Immun **39**(1): 315-35
- ADIRI, R. S., GOPHNA, U. und RON, E. Z. (2003)  
Multilocus sequence typing (MLST) of *Escherichia coli* O78 strains.  
FEMS Microbiol Lett **222**(2): 199-203
- ANDERSSON, S. G. und KURLAND, C. G. (1998)  
Reductive evolution of resident genomes.  
Trends Microbiol **6**(7): 263-8
- AUSUBEL, F. M. (1999)  
Short Protocols in molecular biology.  
New York, NY, John Wiley & Sons, Inc.
- BARNES, H. J. und GROSS, W. B. (1997)  
Colibacillosis.  
Diseases of Poultry. H. B. Calnek, CW Beard, LM McDougald, YM Saif. Ames,  
IA, ISU Press: 131-141
- BERGTHORSSON, U. und OCHMAN, H. (1995)  
Heterogeneity of Genome Sizes among Natural Isolates of *Escherichia coli*.  
J Bacteriol **177**(20): 5784-5789
- BERGTHORSSON, U. und OCHMAN, H. (1998)  
Distribution of chromosome length variation in natural isolates of *Escherichia coli*.  
Mol Biol Evol **15**(1): 6-16
- BINNEWIES, T. T., MOTRO, Y., HALLIN, P. F., LUND, O., DUNN, D., LA, T., HAMPSON, D.  
J., BELLGARD, M., WASSENAAR, T. M. und USSERY, D. W. (2006)  
Ten years of bacterial genome sequencing: comparative-genomics-based  
discoveries.  
Funct Integr Genomics(6): 165-185
- BLANCO, J. E., BLANCO, M., MORA, A. und BLANCO, J. (1997)  
Production of toxins (enterotoxins, verotoxins, and necrotoxins) and colicins by  
*Escherichia coli* strains isolated from septicemic and healthy chickens: relationship  
with *in vivo* pathogenicity.  
J Clin Microbiol **35**(11): 2953-7
- BLATTNER, F. R., PLUNKETT, G., III, BLOCH, C. A., PERNA, N. T., BURLAND, V., RILEY,  
M., COLLADO-VIDES, J., GLASNER, J. D., RODE, C. R. K., MAYHEW, G. F.,  
GREGOR, J., DAVIS, N. W., KIRKPATRICK, H. A., GOEDEN, M. A., ROSE, D. J.,  
MAU, B. und SHAO, Y. (1997)  
The Complete Genome Sequence of *Escherichia coli* K-12.  
Science **277**(5331): 1453-1462

- BONACORSI, S. und BINGEN, E. (2005)  
Molecular epidemiology of *Escherichia coli* causing neonatal meningitis.  
Int J Med Microbiol **295**(6-7): 373-381
- BONACORSI, S., CLERMONT, O., HOUDOUIN, V., CORDEVANT, C., BRAHIMI, N.,  
MARECAT, A., TINSLEY, C., NASSIF, X., LANGE, M. und BINGEN, E. (2003)  
Molecular Analysis and Experimental Virulence of French and North American  
*Escherichia coli* Neonatal Meningitis Isolates: Identification of a New Virulent Clone.  
J Infect Dis **187**(12): 1895-1906
- BRENNER, D. J. und FARMER III, J. J. (2005a)  
*Enterobacteriaceae*.  
Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part 2. G. M.  
Garrity. New York, NY, Bergey's Manual Trust: 587-848
- BRENNER, D. J. und FARMER III, J. J. (2005b)  
*Enterobacteriaceae*.  
Bergey's Manual of Systematic Bacteriology, Second Edition, Volume 2, Part 2. G. M.  
Garrity. New York, NY, Bergey's Manual Trust: 590
- BULLEN, J., GRIFFITHS, E., ROGERS, H. und WARD, G. (2000)  
Sepsis: the critical role of iron.  
Microbes Infect **2**(4): 409-415
- CHAKRAVARTI, A. (2001)  
Single nucleotide polymorphisms: . . .to a future of genetic medicine.  
Nature **409**(6822): 822-823
- CHEN, S. L., HUNG, C., XU, J. G., REIGSTAD, C. S., MAGRINI, V., SABO, A., BLASIAR,  
D., BIERI, T., MEYER, R. R., OZERSKY, P., ARMSTRONG, J. R., FULTON, R. S.,  
LATREILLE, J. P., SPIETH, J., HOOTON, T. M., MARDIS, E. R., HULTGREN, S. J.  
und GORDON, J. I. (2006)  
Identification of genes subject to positive selection in uropathogenic strains of  
*Escherichia coli*: A comparative genomics approach.  
Proc Natl Acad Sci U S A **103**(15): 5977-5982
- CHI, F., WANG, Y., GALLAHER, T. K., WU, C. H., JONG, A. und HUANG, S. H. (2009)  
Identification of IbeR as a Stationary-Phase Regulator in Meningitic *Escherichia coli*  
K1 that Carries a Loss-of-Function Mutation in *rpoS*.  
J Biomed Biotechnol **vol. 2009**: Article ID 520283
- CHONGTRAKOOL, P., ITO, T., MA, X. X., KONDO, Y., TRAKULSOMBOON, S.,  
TIENSASITORN, C., JAMKLANG, M., CHAVALIT, T., SONG, J.-H. und  
HIRAMATSU, K. (2006)  
Staphylococcal Cassette Chromosome *mec* (*SCCmec*) Typing of Methicillin-Resistant  
*Staphylococcus aureus* Strains Isolated in 11 Asian Countries: a Proposal for a New  
Nomenclature for *SCCmec* Elements.  
Antimicrob. Agents Chemother. **50**(3): 1001-1012
- CLERMONT, O., BONACORSI, S. und BINGEN, E. (2000)  
Rapid and Simple Determination of the *Escherichia coli* Phylogenetic Group.  
Appl Environ Microbiol **66**(10): 4555-4558

- COLEMAN, M. L., SULLIVAN, M. B., MARTINY, A. C., STEGLICH, C., BARRY, K., DELONG, E. F. und CHISHOLM, S. W. (2006)  
Genomic Islands and the Ecology and Evolution of *Prochlorococcus*.  
*Science* **311**(5768): 1768-1770
- COLGAN, R., NICOLLE, L. E., MCGLONE, A. und HOOTON, T. M. (2006)  
Asymptomatic bacteriuria in adults.  
*Am Fam Physician* **74**(6): 985-990
- CONNELL, I., AGACE, W., KLEMM, P., SCHEMBRI, M., MARILD, S. und SVANBORG, C. (1996)  
Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract.  
*Proc Natl Acad Sci U S A* **93**(18): 9827-9832
- CULHAM, D. E. und WOOD, J. M. (2000)  
An *Escherichia coli* reference collection group B2- and uropathogen-associated polymorphism in the *rpoS-mutS* region of the *E. coli* chromosome.  
*J Bacteriol* **182**(21): 6272-6
- DAWSON, K. G., EMERSON, J. C. und BURNS, J. L. (1999)  
Fifteen years of experience with bacterial meningitis.  
*Pediatr Infect Dis J* **18**(9): 816-822
- DE CRECY-LAGARD, V., BINET, M. und DANCHIN, A. (1995)  
Fructose Phosphotransferase System of *Xanthomonas campestris pv campestris* - Characterization of the *fruB* Gene.  
*Microbiology* **141**: 2253-2260
- DE LOUVOIS, J. (1994)  
Acute bacterial meningitis in the newborn.  
*J Antimicrob Chemother* **34**(suppl\_A): 61-73
- DHO-MOULIN, M. und FAIRBROTHER, J. M. (1999)  
Avian pathogenic *Escherichia coli* (APEC).  
*Vet Res* **30**(2-3): 299-316
- DIDELLOT, X. und FALUSH, D. (2007)  
Inference of bacterial microevolution using multilocus sequence data.  
*Genetics* **175**(3): 1251-1266
- DIETZMAN, D. E., FISCHER, G. W. und SCHOENKNECHT, F. D. (1974)  
Neonatal *Escherichia coli* septicemia-bacterial counts in blood.  
*J Pediatr* **85**(1): 128-30
- DINGLE, K. E., COLLES, F. M., FALUSH, D. und MAIDEN, M. C. J. (2005)  
Sequence Typing and Comparison of Population Biology of *Campylobacter coli* and *Campylobacter jejuni*.  
*J Clin Microbiol* **43**(1): 340-347
- DOBRINDT, U., HOCHHUT, B., HENTSCHEL, U. und HACKER, J. (2004)  
Genomic islands in pathogenic and environmental microorganisms.  
*Nat Rev Microbiol* **2**(5): 414-24

- DURFEE, T., NELSON, R., BALDWIN, S., PLUNKETT, G., III, BURLAND, V., MAU, B., PETROSINO, J. F., QIN, X., MUZNY, D. M., AYELE, M., GIBBS, R. A., CSORGO, B., POSFAI, G., WEINSTOCK, G. M. und BLATTNER, F. R. (2008)  
The Complete Genome Sequence of *Escherichia coli* DH10B: Insights into the Biology of a Laboratory Workhorse.  
J Bacteriol **190**(7): 2597-2606
- DZIVA, F. und STEVENS, M. P. (2008)  
Colibacillosis in poultry: unravelling the molecular basis of virulence of avian pathogenic *Escherichia coli* in their natural hosts.  
Avian Pathol **37**(4): 355 - 366
- EDEN, C. S. und HANSSON, H. A. (1978)  
*Escherichia coli* pili as possible mediators of attachment to human urinary tract epithelial cells.  
Infect Immun **21**(1): 229-237
- ESCHERICH, T. und BETTELHEIM, K. S. (1885)  
The Intestinal Bacteria of the Neonate and Breast-Fed Infant.  
Rev Infect Dis **11**(2): 352-356
- EWERS, C., JANSSEN, T. und WIELER, L. H. (2003)  
Avian pathogenic *Escherichia coli* (APEC).  
Berl Munch Tierarztl Wochenschr **116**(9-10): 381-395
- EWERS, C., LI, G., WILKING, H., KIESSLING, S., ALT, K., ANTAO, E. M., LATURNUS, C., DIEHL, I., GLODDE, S., HOMEIER, T., BOHNKE, U., STEINRUCK, H., PHILIPP, H. C. und WIELER, L. H. (2007)  
Avian pathogenic, uropathogenic, and newborn meningitis-causing *Escherichia coli*: how closely related are they?  
Int J Med Microbiol **297**(3): 163-76
- FALUSH, D., STEPHENS, M. und PRITCHARD, J. K. (2003)  
Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies.  
Genetics **164**(4): 1567-1587
- FALUSH, D., STEPHENS, M. und PRITCHARD, J. K. (2007)  
Inference of population structure using multilocus genotype data: dominant markers and null alleles.  
Mol Ecol Notes **7**(4): 574-578
- FEIL, E. J., HOLMES, E. C., BESSEN, D. E., CHAN, M., DAY, N. P. J., ENRIGHT, M. C., GOLDSTEIN, R., HOOD, D. W., KALIA, A., MOORE, C. E., ZHOU, J. und SPRATT, B. G. (2001)  
Recombination within natural populations of pathogenic bacteria: Short-term empirical estimates and long-term phylogenetic consequences.  
Proc Natl Acad Sci U S A **98**(1): 182-187
- FEIL, E. J., MAIDEN, M. C., ACHTMAN, M. und SPRATT, B. G. (1999)  
The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*.  
Mol Biol Evol **16**(11): 1496-1502

- FEIL, E. J., SMITH, J. M., ENRIGHT, M. C. und SPRATT, B. G. (2000)  
 Estimating Recombinational Parameters in *Streptococcus pneumoniae* From  
 Multilocus Sequence Typing Data.  
 Genetics **154**(4): 1439-1450
- FOXMAN, B., MANNING, S. D., TALLMAN, P., BAUER, R., ZHANG, L., KOOPMAN, J. S.,  
 GILLESPIE, B., SOBEL, J. D. und MARRS, C. F. (2002)  
 Uropathogenic *Escherichia coli* Are More Likely than Commensal *E. coli* to Be Shared  
 between Heterosexual Sex Partners.  
 Am J Epidemiol **156**(12): 1133-1140
- FRICKE, W. F., WRIGHT, M. S., LINDELL, A. H., HARKINS, D. M., BAKER-AUSTIN, C.,  
 RAVEL, J. und STEPANAUSKAS, R. (2008)  
 Insights into the Environmental Resistance Gene Pool from the Genome Sequence of  
 the Multidrug-Resistant Environmental Isolate *Escherichia coli* SMS-3-5.  
 J Bacteriol **190**(20): 6779-6794
- GALLAGHER, S. R. und DESJARDINS, P. R. (2006)  
 Quantitation of DNA and RNA with absorption and fluorescence spectroscopy.  
 Curr Protoc Mol Biol **Appendix 3**: A.3D9-A.3D.10
- GERMON, P., CHEN, Y. H., HE, L., BLANCO, J. E., BREE, A., SCHOULER, C., HUANG, S.  
 H. und MOULIN-SCHOULEUR, M. (2005)  
*ibeA*, a virulence factor of avian pathogenic *Escherichia coli*.  
 Microbiology **151**(Pt 4): 1179-86
- GORDON, D. M., CLERMONT, O., TOLLEY, H. und DENAMUR, E. (2008)  
 Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence  
 typing versus the PCR triplex method.  
 Environ Microbiol **10**(10): 2484-2496
- GOULLET, P. und PICARD, B. (1989)  
 Comparative electrophoretic polymorphism of esterases and other enzymes in  
*Escherichia coli*.  
 J Gen Microbiol **135**(1): 135-143
- GROSS, W. B. (1994)  
 Diseases due to *Escherichia coli* in poultry.  
*Escherichia coli* in domestic animals and humans. C. L. Gyles. Wallingford, UK, CAB  
 International: 237-259
- GUNTHER, N. W., IV, LOCKATELL, V., JOHNSON, D. E. und MOBLEY, H. L. T. (2001)  
 In Vivo Dynamics of Type 1 Fimbria Regulation in Uropathogenic *Escherichia coli*  
 during Experimental Urinary Tract Infection.  
 Infect Immun **69**(5): 2838-2846
- HACKER, J., BLUM-OEHLER, G., MUHLDOERFER, I. und TSCHAPE, H. (1997)  
 Pathogenicity islands of virulent bacteria: structure, function and impact on microbial  
 evolution.  
 Mol Microbiol **23**(6): 1089-97
- HACKER, J. und CARNIEL, E. (2001)  
 Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of  
 the evolution of microbes.  
 EMBO Rep **2**(5): 376-81

- HACKER, J., HENTSCHEL, U. und DOBRINDT, U. (2003)  
 Prokaryotic Chromosomes and Disease.  
 Science **301**(5634): 790-793
- HACKER, J. und KAPER, J. B. (2000)  
 Pathogenicity islands and the evolution of microbes.  
 Annu Rev Microbiol **54**: 641-79
- HAYASHI, T., MAKINO, K., OHNISHI, M., KUROKAWA, K., ISHII, K., YOKOYAMA, K., HAN, C. G., OHTSUBO, E., NAKAYAMA, K., MURATA, T., TANAKA, M., TOBE, T., IIDA, T., TAKAMI, H., HONDA, T., SASAKAWA, C., OGASAWARA, N., YASUNAGA, T., KUHARA, S., SHIBA, T., HATTORI, M. und SHINAGAWA, H. (2001)  
 Complete genome sequence of enterohemorrhagic *Escherichia coli* O157 : H7 and genomic comparison with a laboratory strain K-12.  
 DNA Res **8**(1): 11-22
- HEJNOVA, J., DOBRINDT, U., NEMCOVA, R., RUSNIOK, C., BOMBA, A., FRANGEUL, L., HACKER, J., GLASER, P., SEBO, P. und BUCHRIESER, C. (2005)  
 Characterization of the flexible genome complement of the commensal *Escherichia coli* strain A0 34/86 (O83 : K24 : H31).  
 Microbiology **151**(2): 385-398
- HENSEL, M., SHEA, J. E., GLEESON, C., JONES, M. D., DALTON, E. und HOLDEN, D. W. (1995)  
 Simultaneous identification of bacterial virulence genes by negative selection.  
 Science **269**(5222): 400-3
- HENTSCHEL, U. und HACKER, J. (2001)  
 Pathogenicity islands: the tip of the iceberg.  
 Microbes Infect **3**(7): 545-548
- HERZER, P. J., INOUE, S., INOUE, M. und WHITTAM, T. S. (1990)  
 Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of *Escherichia coli*.  
 J Bacteriol **172**(11): 6175-6181
- HOCHHUT, B., WILDE, C., BALLING, G., MIDDENDORF, B., DOBRINDT, U., BRZUSZKIEWICZ, E., GOTTSCHALK, G., CARNIEL, E. und HACKER, J. (2006)  
 Role of pathogenicity island-associated integrases in the genome plasticity of uropathogenic *Escherichia coli* strain 536.  
 Mol Microbiol **61**(3): 584-595
- HOFFMAN, J. A., BADGER, J. L., ZHANG, Y., HUANG, S. H. und KIM, K. S. (2000)  
*Escherichia coli* K1 *asIA* contributes to invasion of brain microvascular endothelial cells *in vitro* and *in vivo*.  
 Infect Immun **68**(9): 5062-7
- HOLMS, W. H. (1986)  
 The Central Metabolic Pathways of *Escherichia-Coli* - Relationship between Flux and Control at a Branch Point, Efficiency of Conversion to Biomass, and Excretion of Acetate.  
 Curr Top Cell Regul **28**: 69-105

- HUANG, S. H., CHEN, Y. H., FU, Q., STINS, M., WANG, Y., WASS, C. und KIM, K. S. (1999)  
 Identification and characterization of an *Escherichia coli* invasion gene locus, *ibeB*, required for penetration of brain microvascular endothelial cells.  
*Infect Immun* **67**(5): 2103-9
- HUANG, S. H., CHEN, Y. H., KONG, G., CHEN, S. H., BESEMER, J., BORODOVSKY, M. und JONG, A. (2001a)  
 A novel genetic island of meningitic *Escherichia coli* K1 containing the *ibeA* invasion gene (GimA): functional annotation and carbon-source-regulated invasion of human brain microvascular endothelial cells.  
*Funct Integr Genomics* **1**(5): 312-22
- HUANG, S. H., WAN, Z. S., CHEN, Y. H., JONG, A. Y. und KIM, K. S. (2001b)  
 Further characterization of *Escherichia coli* brain microvascular endothelial cell invasion gene *ibeA* by deletion, complementation, and protein expression.  
*J Infect Dis* **183**(7): 1071-8
- HUANG, S. H., WASS, C., FU, Q., PRASADARAO, N. V., STINS, M. und KIM, K. S. (1995)  
*Escherichia coli* invasion of brain microvascular endothelial cells *in vitro* and *in vivo*: molecular cloning and characterization of invasion gene *ibe10*.  
*Infect Immun* **63**(11): 4470-5
- HUGHES, A. L. und NEI, M. (1989)  
 Nucleotide Substitution at Major Histocompatibility Complex Class-II Loci - Evidence for Overdominant Selection.  
*Proc Natl Acad Sci U S A* **86**(3): 958-962
- IGUCHI, A., THOMSON, N. R., OGURA, Y., SAUNDERS, D., OOKA, T., HENDERSON, I. R., HARRIS, D., ASADULGHANI, M., KUROKAWA, K., DEAN, P., KENNY, B., QUAIL, M. A., THURSTON, S., DOUGAN, G., HAYASHI, T., PARKHILL, J. und FRANKEL, G. (2009)  
 Complete Genome Sequence and Comparative Genome Analysis of Enteropathogenic *Escherichia coli* O127:H6 Strain E2348/69.  
*J Bacteriol* **191**(1): 347-354
- JANßEN, T., SCHWARZ, C., PREIKSCHAT, P., VOSS, M., PHILIPP, H. C. und WIELER, L. H. (2001)  
 Virulence-associated genes in avian pathogenic *Escherichia coli* (APEC) isolated from internal organs of poultry having died from colibacillosis.  
*Int J Med Microbiol* **291**(5): 371-378
- JAUREGUY, F., LANDRAUD, L., PASSET, V., DIANCOURT, L., FRAPY, E., GUIGON, G., CARBONNELLE, E., LORTHOLARY, O., CLERMONT, O., DENAMUR, E., PICARD, B., NASSIF, X. und BRISSE, S. (2008)  
 Phylogenetic and genomic diversity of human bacteremic *Escherichia coli* strains.  
*BMC Genomics* **9**: 560
- JOHNSON, J. R. (1991)  
 Virulence Factors in *Escherichia coli* Urinary-Tract Infection.  
*Clin Microbiol Rev* **4**(1): 80-128
- JOHNSON, J. R., DELAVARI, P., KUSKOWSKI, M. und STELL, A. L. (2001)  
 Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*.  
*J Infect Dis* **183**(1): 78-88

- JOHNSON, J. R., OSWALD, E., O'BRYAN, T. T., KUSKOWSKI, M. A. und SPANJAARD, L. (2002)  
Phylogenetic distribution of virulence-associated genes among *Escherichia coli* isolates associated with neonatal bacterial meningitis in the Netherlands.  
J Infect Dis **185**(6): 774-84
- JOHNSON, J. R. und RUSSO, T. A. (2002)  
Extraintestinal pathogenic *Escherichia coli*: "the other bad *E. coli*".  
J Lab Clin Med **139**(3): 155-162
- JOHNSON, J. R. und STELL, A. L. (2000)  
Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise.  
J Infect Dis **181**(1): 261-272
- JOHNSON, T. J., KARIYAWASAM, S., WANNEMUEHLER, Y., MANGIAMELE, P., JOHNSON, S. J., DOETKOTT, C., SKYBERG, J. A., LYNNE, A. M., JOHNSON, J. R. und NOLAN, L. K. (2007)  
The Genome Sequence of Avian Pathogenic *Escherichia coli* Strain O1:K1:H7 Shares Strong Similarities with Human Extraintestinal Pathogenic *E. coli* Genomes.  
J Bacteriol **189**(8): 3228-3236
- JOHNSON, T. J., WANNEMUEHLER, Y., JOHNSON, S. J., STELL, A. L., DOETKOTT, C., JOHNSON, J. R., KIM, K. S., SPANJAARD, L. und NOLAN, L. K. (2008)  
Comparison of Extraintestinal Pathogenic *Escherichia coli* Strains from Human and Avian Sources Reveals a Mixed Subset Representing Potential Zoonotic Pathogens.  
Appl Environ Microbiol **74**(22): 7043-7050
- JORES, J., RUMER, L., KIESSLING, S., KAPER, J. B. und WIELER, L. H. (2001)  
A novel locus of enterocyte effacement (LEE) pathogenicity island inserted at *pheV* in bovine Shiga toxin-producing *Escherichia coli* strain O103:H2.  
FEMS Microbiol Lett **204**(1): 75-9
- JORES, J., RUMER, L. und WIELER, L. H. (2004)  
Impact of the locus of enterocyte effacement pathogenicity island on the evolution of pathogenic *Escherichia coli*.  
Int J Med Microbiol **294**(2-3): 103-113
- JORES, J., TORRES, A. G., WAGNER, S., TUTT, C. B., KAPER, J. B. und WIELER, L. H. (2006)  
Identification and characterization of "pathoadaptive mutations" of the *cadBA* operon in several intestinal *Escherichia coli*.  
Int J Med Microbiol **296**(8): 547-552
- KAPER, J. B., NATARO, J. P. und MOBLEY, H. L. (2004)  
Pathogenic *Escherichia coli*.  
Nat Rev Microbiol **2**(2): 123-140
- KIM, K. S. (1985)  
Comparison of cefotaxime, imipenem-cilastatin, ampicillin-gentamicin, and ampicillin-chloramphenicol in the treatment of experimental *Escherichia coli* bacteremia and meningitis.  
Antimicrob Agents Chemother **28**(3): 433-436

- KIM, K. S. (2002)  
Meningitis-Associated *Escherichia coli*.  
*Escherichia coli: Virulence Mechanisms of a Versatile Pathogen*. D. M.S. Orlando, Florida, USA, Academic Press: 269 - 286
- KIM, K. S. (2003)  
Pathogenesis of bacterial meningitis: from bacteraemia to neuronal injury.  
*Nat Rev Neurosci* **4**(5): 376-385
- KIM, K. S., ITABASHI, H., GEMSKI, P., SADOFF, J., WARREN, R. L. und CROSS, A. S. (1992)  
The K1 capsule is the critical determinant in the development of *Escherichia coli* meningitis in the rat.  
*J Clin Invest* **90**(3): 897-905
- KIRSCH, P., JORES, J. und WIELER, L. H. (2004)  
Plasticity of bacterial genomes: Pathogenicity islands and the Locus of Enterocyte Effacement (LEE).  
*Berl Munch Tierarztl Wochenschr* **117**(3-4): 116-129
- KORHONEN, T. K., VALTONEN, M. V., PARKKINEN, J., VAISANEN-RHEN, V., FINNE, J., ORSKOV, F., ORSKOV, I., SVENSON, S. B. und MAKELA, P. H. (1985)  
Serotypes, hemolysin production, and receptor recognition of *Escherichia coli* strains associated with neonatal sepsis and meningitis.  
*Infect Immun* **48**(2): 486-491
- KORHONEN, T. K., VIRKOLA, R. und HOLTHOFER, H. (1986)  
Localization of binding sites for purified *Escherichia coli* P fimbriae in the human kidney.  
*Infect Immun* **54**(2): 328-332
- LAN, R. und REEVES, P. R. (2002)  
*Escherichia coli* in disguise: molecular origins of Shigella.  
*Microbes Infect* **4**(11): 1125-1132
- LAWRENCE, J. G. und HENDRICKSON, H. (2003)  
Lateral gene transfer: when will adolescence end?  
*Mol Microbiol* **50**(3): 739-749
- LE GALL, T., CLERMONT, O., GOURIOU, S., PICARD, B., NASSIF, X., DENAMUR, E. und TENAILLON, O. (2007)  
Extraintestinal virulence is a coincidental by-product of commensalism in B2 phylogenetic group *Escherichia coli* strains.  
*Mol Biol Evol* **24**(11): 2373-2384
- LEDERBERG, J. und TATUM, E. L. (1946)  
Gene Recombination in *Escherichia coli*.  
*Nature* **158**(4016): 558-558
- LEIB, S. L. und TAUBER, M. G. (1999)  
Pathogenesis of bacterial meningitis.  
*Infect Dis Clin North Am* **13**(3): 527-48, v-vi

- LI, G., LATURNUS, C., EWERS, C. und WIELER, L. H. (2005)  
 Identification of genes required for avian *Escherichia coli* septicemia by signature-tagged mutagenesis.  
*Infect Immun* **73**(5): 2818-2827
- MAIDEN, M. C., BYGRAVES, J. A., FEIL, E., MORELLI, G., RUSSELL, J. E., URWIN, R., ZHANG, Q., ZHOU, J., ZURTH, K., CAUGANT, D. A., FEAVERS, I. M., ACHTMAN, M. und SPRATT, B. G. (1998)  
 Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms.  
*Proc Natl Acad Sci U S A* **95**(6): 3140-5
- MAIDEN, M. C. J. (2006)  
 Multilocus Sequence Typing of Bacteria.  
*Annu Rev Microbiol* **60**(1): 561-588
- MAKINO, K., YOKOYAMA, K., KUBOTA, Y., YUTSUDO, C. H., KIMURA, S., KUROKAWA, K., ISHII, K., HATTORI, M., TATSUNO, I., ABE, H., IIDA, T., YAMAMOTO, K., ONISHI, M., HAYASHI, T., YASUNAGA, T., HONDA, T., SASAKAWA, C. und SHINAGAWA, H. (1999)  
 Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic *Escherichia coli* O157:H7 derived from the Sakai outbreak.  
*Genes Genet Syst* **74**(5): 227-39
- MANGES, A. R., JOHNSON, J. R., FOXMAN, B., O'BRYAN, T. T., FULLERTON, K. E. und RILEY, L. W. (2001)  
 Widespread Distribution of Urinary Tract Infections Caused by a Multidrug-Resistant *Escherichia coli* Clonal Group.  
*N Engl J Med* **345**(14): 1007-1013
- MCVEAN, G. A. T. und CARDIN, N. J. (2005)  
 Approximating the coalescent with recombination.  
*Philos T Roy Soc B* **360**(1459): 1387-1393
- MESEGUER, M. A., ÁLVAREZ, A., REJAS, M. T., SÁNCHEZ, C., PÉREZ-DÍAZ, J. C. und BAQUERO, F. (2003)  
*Mycoplasma pneumoniae*: a reduced-genome intracellular bacterial pathogen.  
*Infect Genet Evol* **3**(1): 47-55
- MILKMAN, R. und BRIDGES, M. M. (1990)  
 Molecular Evolution of the *Escherichia coli* Chromosome: 3. Clonal Frames.  
*Genetics* **126**(3): 505-517
- MILKMAN, R. und CRAWFORD, I. (1983)  
 Clustered third-base substitutions among wild strains of *Escherichia coli*.  
*Science* **221**(4608): 378-380
- MILKMAN, R., JAEGER, E. und MCBRIDE, R. D. (2003)  
 Molecular Evolution of the *Escherichia coli* Chromosome: 6. Two Regions of High Effective Recombination.  
*Genetics* **163**(2): 475-483

- MORA, A., LOPEZ, C., DABHI, G., BLANCO, M., BLANCO, J., ALONSO, M., HERRERA, A., MAMANI, R., BONACORSI, S., MOULIN-SCHOULEUR, M. und BLANCO, J. (2009)  
Extraintestinal pathogenic *Escherichia coli* O1:K1:H7/NM from human and avian origin: detection of clonal groups B2 ST95 and D ST59 with different host distribution. *BMC Microbiol* **9**(1): 132
- MORDHORST, I. L., CLAUS, H., EWERS, C., LAPPANN, M., SCHOEN, C., ELIAS, J., BATZILLA, J., DOBRINDT, U., WIELER, L. H., BERGFELD, A. K., MUHLENHOFF, M. und VOGEL, U. (2009)  
O-acetyltransferase gene *neuO* is segregated according to phylogenetic background and contributes to environmental desiccation resistance in *Escherichia coli* K1. *Environ Microbiol* **Epub ahead of print**
- MOULIN-SCHOULEUR, M., REPERANT, M., LAURENT, S., BREE, A., MIGNON-GRASTEAU, S., GERMON, P., RASSCHAERT, D. und SCHOULER, C. (2007)  
Extraintestinal pathogenic *Escherichia coli* strains of avian and human origin: link between phylogenetic relationships and common virulence patterns. *J Clin Microbiol* **45**(10): 3366-76
- NATARO, J. P. und KAPER, J. B. (1998)  
Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev* **11**(1): 142-201
- NELSON, K. und SELANDER, R. K. (1994)  
Intergeneric Transfer and Recombination of the 6-Phosphogluconate Dehydrogenase Gene (*Gnd*) in Enteric Bacteria. *Proc Natl Acad Sci U S A* **91**(21): 10227-10231
- NIELSEN, R. und YANG, Z. H. (1998)  
Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**(3): 929-936
- NYSTRÖM, T. (2004)  
Stationary-Phase Physiology. *Annu Rev Microbiol* **58**(1): 161-181
- OCHMAN, H., LAWRENCE, J. G. und GROISMAN, E. A. (2000)  
Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**(6784): 299-304
- OCHMAN, H. und SELANDER, R. K. (1984)  
Standard reference strains of *Escherichia coli* from natural populations. *J Bacteriol* **157**(2): 690-693
- OCHMAN, H. und WILSON, A. C. (1987)  
Evolution in Bacteria - Evidence for a Universal Substitution Rate in Cellular Genomes. *J Mol Evol* **26**(1-2): 74-86
- OELSCHLAEGER, T. und FÜNFSTÜCK, R. (2006)  
Recurrent urinary tract infections in women. Virulence of pathogens and host reaction. *Urologe* **45**(4): 412-420

- OELSCHLAEGER, T. A., DOBRINDT, U. und HACKER, J. (2002)  
 Pathogenicity islands of uropathogenic *E. coli* and the evolution of virulence.  
*Int J Antimicrob Agents* **19**(6): 517-521
- ORR, M. R. und SMITH, T. B. (1998)  
 Ecology and speciation.  
*Trends Ecol Evol* **13**(12): 502-506
- OSHIMA, K., KAKIZAWA, S., NISHIGAWA, H., JUNG, H. Y., WEI, W., SUZUKI, S.,  
 ARASHIDA, R., NAKATA, D., MIYATA, S., UGAKI, M. und NAMBA, S. (2004)  
 Reductive evolution suggested from the complete genome sequence of a plant-  
 pathogenic phytoplasma.  
*Nat Genet* **36**(1): 27-9
- OSHIMA, K., TOH, H., OGURA, Y., SASAMOTO, H., MORITA, H., PARK, S., OOKA, T.,  
 IYODA, S., TAYLOR, T. D., HAYASHI, T., ITOH, K. und HATTORI, M. (2008)  
 Complete Genome Sequence and Comparative Analysis of the Wild-type Commensal  
*Escherichia coli* Strain SE11 Isolated from a Healthy Adult.  
*DNA Res* **15**(6): 375-386
- PERNA, N. T., PLUNKETT, G., BURLAND, V., MAU, B., GLASNER, J. D., ROSE, D. J.,  
 MAYHEW, G. F., EVANS, P. S., GREGOR, J., KIRKPATRICK, H. A., POSFAI, G.,  
 HACKETT, J., KLINK, S., BOUTIN, A., SHAO, Y., MILLER, L., GROTEBECK, E. J.,  
 DAVIS, N. W., LIM, A., DIMALANTA, E. T., POTAMOUSIS, K. D., APODACA, J.,  
 ANANTHARAMAN, T. S., LIN, J., YEN, G., SCHWARTZ, D. C., WELCH, R. A. und  
 BLATTNER, F. R. (2001)  
 Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.  
*Nature* **409**(6819): 529-533
- PETERSEN, L., BOLLBACK, J. P., DIMMIC, M., HUBISZ, M. und NIELSEN, R. (2007)  
 Genes under positive selection in *Escherichia coli*.  
*Genome Res* **17**(9): 1336-1343
- PHILLIPS, I., KING, A., ROWE, B., EYKYN, S., GRANSDEN, W. R., FROST, J. und  
 GROSS, R. (1988)  
 Epidemic Multiresistant *Escherichia coli* Infection in west Lambeth Health District.  
*Lancet* **331**(8593): 1038-1041
- PLAINVERT, C., BIDET, P., PEIGNE, C., BARBE, V., MEDIGUE, C., DENAMUR, E.,  
 BINGEN, E. und BONACORSI, S. (2007)  
 A New O-Antigen Gene Cluster Has a Key Role in the Virulence of the *Escherichia*  
*coli* Meningitis Clone O45:K1:H7.  
*J Bacteriol* **189**(23): 8528-8536
- POND, S. L. und FROST, S. D. (2005)  
 Datamonkey: rapid detection of selective pressure on individual sites of codon  
 alignments.  
*Bioinformatics* **21**(10): 2531-3
- POSFAI, G., PLUNKETT, G., III, FEHER, T., FRISCH, D., KEIL, G. M., UMENHOFFER, K.,  
 KOLISNYCHENKO, V., STAHL, B., SHARMA, S. S., DE ARRUDA, M., BURLAND,  
 V., HARCUM, S. W. und BLATTNER, F. R. (2006)  
 Emergent Properties of Reduced-Genome *Escherichia coli*.  
*Science* **312**(5776): 1044-1046

- PRASADARAO, N. V., WASS, C. A., HUANG, S. H. und KIM, K. S. (1999)  
 Identification and characterization of a novel lbe10 binding protein that contributes to *Escherichia coli* invasion of brain microvascular endothelial cells.  
*Infect Immun* **67**(3): 1131-8
- PRASADARAO, N. V., WASS, C. A., WEISER, J. N., STINS, M. F., HUANG, S. H. und KIM, K. S. (1996)  
 Outer membrane protein A of *Escherichia coli* contributes to invasion of brain microvascular endothelial cells.  
*Infect Immun* **64**(1): 146-53
- PRITCHARD, J. K., STEPHENS, M. und DONNELLY, P. (2000)  
 Inference of population structure using multilocus genotype data.  
*Genetics* **155**(2): 945-959
- RASKO, D. A., ROSOVITZ, M. J., MYERS, G. S. A., MONGODIN, E. F., FRICKE, W. F., GAJER, P., CRABTREE, J., SEBAIHIA, M., THOMSON, N. R., CHAUDHURI, R., HENDERSON, I. R., SPERANDIO, V. und RAVEL, J. (2008)  
 The pangenome structure of *Escherichia coli*: Comparative genomic analysis of *E. coli* commensal and pathogenic isolates.  
*J Bacteriol* **190**(20): 6881-6893
- REEVES, P. R. (1992)  
 Variation in O-Antigens, Niche-Specific Selection and Bacterial Populations.  
*Fems Microbiol Lett* **100**(1-3): 509-516
- REID, S. D., HERBELIN, C. J., BUMBAUGH, A. C., SELANDER, R. K. und WHITTAM, T. S. (2000)  
 Parallel evolution of virulence in pathogenic *Escherichia coli*.  
*Nature* **406**(6791): 64-67
- RILEY, M., ABE, T., ARNAUD, M. B., BERLYN, M. K. B., BLATTNER, F. R., CHAUDHURI, R. R., GLASNER, J. D., HORIUCHI, T., KESELER, I. M., KOSUGE, T., MORI, H., PERNA, N. T., PLUNKETT, G., RUDD, K. E., SERRES, M. H., THOMAS, G. H., THOMSON, N. R., WISHART, D. und WANNER, B. L. (2006)  
*Escherichia coli* K-12: a cooperatively developed annotation snapshot - 2005.  
*Nucleic Acids Res* **34**(1): 1-9
- ROBERTS, A. P. und PHILLIPS, R. (1983)  
 The relative importance of the classical and alternative complement pathways in serum bactericidal activity against *Escherichia coli*.  
*J Med Microbiol* **16**(1): 69-74
- ROBINS-BROWNE, R. M. und HARTLAND, E. L. (2002)  
*Escherichia coli* as a cause of diarrhea.  
*J Gastroenterol Hepatol* **17**(4): 467-475
- ROZAS, J., SANCHEZ-DELBARRIO, J. C., MESSEGUER, X. und ROZAS, R. (2003)  
 DnaSP, DNA polymorphism analyses by the coalescent and other methods.  
*Bioinformatics* **19**(18): 2496-7

- RUMER, L., JORES, J., KIRSCH, P., CAVIGNAC, Y., ZEHMKE, K. und WIELER, L. H. (2003)  
 Dissemination of pheU- and pheV-located genomic islands among enteropathogenic (EPEC) and enterohemorrhagic (EHEC) *E. coli* and their possible role in the horizontal transfer of the locus of enterocyte effacement (LEE).  
*Int J Med Microbiol* **292**(7-8): 463-75
- RUSSO, T. A. und JOHNSON, J. R. (2000)  
 Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC.  
*J Infect Dis* **181**(5): 1753-4
- SABATÉ, M., MORENO, E., PÉREZ, T., ANDREU, A. und PRATS, G. (2006)  
 Pathogenicity island markers in commensal and uropathogenic *Escherichia coli* isolates.  
*Clin Microbiol Infect* **12**(9): 880-886
- SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. und ERLICH, H. A. (1988)  
 Primer-Directed Enzymatic Amplification of DNA with a Thermostable DNA-Polymerase.  
*Science* **239**(4839): 487-491
- SANDNER, L., EGUIARTE, L. E., NAVARRO, A., CRAVIOTO, A. und SOUZA, V. (2001)  
 The elements of the focus of enterocyte effacement in human and wild mammal isolates of *Escherichia coli*: evolution by assemblage or disruption?  
*Microbiology-Sgm* **147**: 3149-3158
- SANKAR, T. S., NEELAKANTA, G., SANGAL, V., PLUM, G., ACHTMAN, M. und SCHNETZ, K. (2009)  
 Fate of the H-NS-Repressed *bgl* Operon in Evolution of *Escherichia coli*.  
*PLoS Genet* **5**(3): e1000405
- SCHIERUP, M. H. und HEIN, J. (2000)  
 Consequences of Recombination on Traditional Phylogenetic Analysis.  
*Genetics* **156**(2): 879-891
- SCHUBERT, S., DARLU, P., CLERMONT, O., WIESER, A., MAGISTRO, G., HOFFMANN, C., WEINERT, K., TENAILLON, O., MATIC, I. und DENAMUR, E. (2009)  
 Role of Intraspecies Recombination in the Spread of Pathogenicity Islands within the *Escherichia coli* Species.  
*PLoS Pathog* **5**(1): e1000257
- SCHULZE, J., SCHIEMANN, U. und SONNENBORN, U. (2006)  
 120 Jahre *E. coli*: Bedeutung in Forschung und Medizin.  
 Hagen, Germany, Alfred-Nissle-Gesellschaft
- SMITH, J. M., SMITH, N. H., O'ROURKE, M. und SPRATT, B. G. (1993)  
 How clonal are bacteria?  
*Proc Natl Acad Sci U S A* **90**(10): 4384-4388
- SOKURENKO, E. V., HASTY, D. L. und DYKHUIZEN, D. E. (1999)  
 Pathoadaptive mutations: gene loss and variation in bacterial pathogens.  
*Trends Microbiol* **7**(5): 191-195

- STINEAR, T. P., SEEMANN, T., PIDOT, S., FRIGUI, W., REYSSET, G., GARNIER, T., MEURICE, G., SIMON, D., BOUCHIER, C., MA, L., TICHIT, M., PORTER, J. L., RYAN, J., JOHNSON, P. D., DAVIES, J. K., JENKIN, G. A., SMALL, P. L., JONES, L. M., TEKAIA, F., LAVAL, F., DAFFE, M., PARKHILL, J. und COLE, S. T. (2007)  
 Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer.  
*Genome Res* **17**(2): 192-200
- STINS, M. F., BADGER, J. und SIK KIM, K. (2001)  
 Bacterial invasion and transcytosis in transfected human brain microvascular endothelial cells.  
*Microb Pathog* **30**(1): 19-28
- STINS, M. F., NEMANI, P. V., WASS, C. und KIM, K. S. (1999)  
*Escherichia coli* binding to and invasion of brain microvascular endothelial cells derived from humans and rats of different ages.  
*Infect Immun* **67**(10): 5522-5
- STOLL, B. J., HANSEN, N., FANAROFF, A. A., WRIGHT, L. L., CARLO, W. A., EHRENKRANZ, R. A., LEMONS, J. A., DONOVAN, E. F., STARK, A. R., TYSON, J. E., OH, W., BAUER, C. R., KORONES, S. B., SHANKARAN, S., LAPTOOK, A. R., STEVENSON, D. K., PAPILE, L. A. und POOLE, W. K. (2002)  
 Late-onset sepsis in very low birth weight neonates: the experience of the NICHD Neonatal Research Network.  
*Pediatrics* **110**(2 Pt 1): 285-91
- STONEKING, M. (2001)  
 Single nucleotide polymorphisms: From the evolutionary past. . .  
*Nature* **409**(6822): 821-822
- TAMAMES, J., MOYA, A. und VALENCIA, A. (2007)  
 Modular organization in the reductive evolution of protein-protein interaction networks.  
*Genome Biol* **8**(5): R94
- TAMURA, K., DUDLEY, J., NEI, M. und KUMAR, S. (2007)  
 MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0.  
*Mol Biol Evol* **24**(8): 1596-1599
- TANAKA, T. und NEI, M. (1989)  
 Positive darwinian selection observed at the variable-region genes of immunoglobulins.  
*Mol Biol Evol* **6**(5): 447-59
- TARTOF, S. Y., SOLBERG, O. D. und RILEY, L. W. (2007)  
 Genotypic analyses of uropathogenic *Escherichia coli* based on *fimH* single nucleotide polymorphisms (SNPs).  
*J Med Microbiol* **56**(10): 1363-1369

- TOUCHON, M., HOEDE, C., TENAILLON, O., BARBE, V., BAERISWYL, S., BIDET, P., BINGEN, E., BONACORSI, S., BOUCHIER, C., BOUVET, O., CALTEAU, A., CHIAPELLO, H., CLERMONT, O., CRUVEILLER, S., DANCHIN, A., DIARD, M., DOSSAT, C., EL KAROUI, M., FRAPY, E., GARRY, L., GHIGO, J. M., GILLES, A. M., JOHNSON, J. A., LE BOUGUENEC, C., LESCAT, M., MANGENOT, S., MARTINEZ-JOHANNE, V., MATIC, I., NASSIF, X., OZTAS, S., PETIT, M. A., PICHON, C., ROUY, Z., RUF, C. S., SCHNEIDER, D., TOURRET, J., VACHERIE, B., VALLENET, D., MEDIGUE, C., ROCHA, E. P. C. und DENAMUR, E. (2009)  
Organised Genome Dynamics in the *Escherichia coli* Species Results in Highly Diverse Adaptive Paths.  
PLoS Genet **5**(1): e1000344
- TRIFILLIS, A. L., DONNENBERG, M. S., CUI, X. L., RUSSELL, R. G., UTSALO, S. J., MOBLEY, H. L. T. und WARREN, J. W. (1994)  
Binding to and Killing of Human Renal Epithelial-Cells by Hemolytic P-Fimbriated *Escherichia coli*.  
Kidney Int **46**(4): 1083-1091
- UNHANAND, M., MUSTAFA, M. M., MCCRACKEN, J., G. H. und NELSON, J. D. (1993)  
Gram-negative enteric bacillary meningitis: A twenty-one-year experience.  
J Pediatr **122**(1): 15-21
- WANG, D., DING, X. und RATHER, P. N. (2001)  
Indole Can Act as an Extracellular Signal in *Escherichia coli*.  
J Bacteriol **183**(14): 4210-4216
- WANG, Y., HUANG, S. H., WASS, C. A., STINS, M. F. und KIM, K. S. (1999)  
The gene locus *yijP* contributes to *Escherichia coli* K1 invasion of brain microvascular endothelial cells.  
Infect Immun **67**(9): 4751-6
- WANG, Y. und KIM, K. S. (2002)  
Role of OmpA and IbeB in *Escherichia coli* K1 invasion of brain microvascular endothelial cells *in vitro* and *in vivo*.  
Pediatr Res **51**(5): 559-63
- WEISSMAN, S. J., CHATTOPADHYAY, S., APRIKIAN, P., OBATA-YASUOKA, M., YAROVA-YAROVAYA, Y., STAPLETON, A., BA-THEIN, W., DYKHUIZEN, D., JOHNSON, J. R. und SOKURENKO, E. V. (2006)  
Clonal analysis reveals high rate of structural mutations in fimbrial adhesins of extraintestinal pathogenic *Escherichia coli*.  
Mol Microbiol **59**(3): 975-88
- WELCH, R. A., BURLAND, V., PLUNKETT, G., REDFORD, P., ROESCH, P., RASKO, D., BUCKLES, E. L., LIOU, S.-R., BOUTIN, A., HACKETT, J., STROUD, D., MAYHEW, G. F., ROSE, D. J., ZHOU, S., SCHWARTZ, D. C., PERNA, N. T., MOBLEY, H. L. T., DONNENBERG, M. S. und BLATTNER, F. R. (2002)  
Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*.  
Proc Natl Acad Sci U S A **99**(26): 17020-17024
- WILES, T. J., KULESUS, R. R. und MULVEY, M. A. (2008)  
Origins and virulence mechanisms of uropathogenic *Escherichia coli*.  
Exp Mol Pathol - Special Issue: Molecular Pathology and Molecular Diagnostics **85**(1): 11-19

- WILLINGER, H. (1992)  
Krankheiten des Wirtschaftsgeflügels Band II.  
Enterobacteriaceae-Infektionen. G. u. G. M. Heider. Jena, Gustav Fischer Verlag: 97-110
- WIRTH, T., FALUSH, D., LAN, R., COLLES, F., MENSA, P., WIELER, L. H., KARCH, H., REEVES, P. R., MAIDEN, M. C., OCHMAN, H. und ACHTMAN, M. (2006)  
Sex and virulence in *Escherichia coli*: an evolutionary perspective.  
Mol Microbiol **60**(5): 1136-1151
- YANG, Z. und NIELSEN, R. (2000)  
Estimating Synonymous and Nonsynonymous Substitution Rates Under Realistic Evolutionary Models.  
Mol Biol Evol **17**(1): 32-43
- ZDZIARSKI, J., SVANBORG, C., WULLT, B., HACKER, J. und DOBRINDT, U. (2008)  
Molecular Basis of Commensalism in the Urinary Tract: Low Virulence or Virulence Attenuation?  
Infect Immun **76**(2): 695-703
- ZOU, Y., HE, L. und HUANG, S. H. (2006)  
Identification of a surface protein on human brain microvascular endothelial cells as vimentin interacting with *Escherichia coli* invasion protein IbeA.  
Biochem Biophys Res Commun **351**(3): 625-30
- ZOU, Y., HE, L., WU, C. H., CAO, H., XIE, Z. H., OUYANG, Y., WANG, Y., JONG, A. und HUANG, S. H. (2007)  
PSF is an IbeA-binding protein contributing to meningitic *Escherichia coli* K1 invasion of human brain microvascular endothelial cells.  
Med Microbiol Immunol **196**(3): 135-43
- ZOU, Y. M., HE, L., CHI, F., JONG, A. und HUANG, S. H. (2008)  
Involvement of *Escherichia coli* K1 ibeT in bacterial adhesion that is associated with the entry into human brain microvascular endothelial cells.  
Med Microbiol Immunol **197**(4): 337-344

## 7 ZUSAMMENFASSUNG

### Der GimA-Lokus von Extraintestinal-pathogenen *E. coli* (ExPEC): Reduktive Evolution in Abhängigkeit von Habitat und Pathovar

Untersuchungen an humanen Gehirndothelzellen (HBMEC - brain microvascular endothelial cells) sowie an Tierversuchsmodellen mit neugeborenen Ratten und Hühnern konnten zeigen, dass das *ibeA*-Gen an der Pathogenese von systemischen *E. coli*-Infektionen beteiligt ist. *ibeA* ist in einer 20,3 kb großen chromosomalen Region lokalisiert, die von den Genen *yjiD* und *yjiE* flankiert wird. Diese Region ist als genetische Insel beschrieben und wird GimA (genomic island of meningitic *E. coli* containing *ibeA*) genannt. GimA besteht aus vier Operons (GimA1-4). Die *E. coli*, die die systemischen Infektionen auslösen, gehören den Pathovaren APEC (aviär pathogene) und NMEC (Neugeborenenmeningitis-assoziierte *E. coli*) aus der Gruppe der extraintestinal-pathogenen *E. coli* (ExPEC) an. Beispielsweise ist *ibeA* in die Überwindung der Blut-Hirn-Schranke, einem wesentlichen Schritt in der Pathogenese der Meningitis beim Neugeborenen, involviert.

Der putative Insertionslokus von GimA wurde mittels Long-Range-PCR und DNS-DNS-Hybridisierung mit dem Ziel untersucht, die Phylogenie von GimA aufzudecken. Weiterhin sollte der Inselcharakter von GimA eingehender analysiert werden. Die Untersuchungen wurden an insgesamt 410 *E. coli*-Isolaten durchgeführt. Im Probenumfang waren die Pathovare APEC, NMEC, UPEC (uropathogene) und SepEC (Septikämie-assoziierte *E. coli*) sowie Isolate aus Stuhlproben bzw. Kotproben von klinisch gesunden Menschen und Tieren enthalten. Darüber hinaus waren die 72 *E. coli*-Stämme der EcoR-Sammlung (*E. coli* reference collection) in den 410 Isolaten enthalten.

Es konnte neben einer Variante des Insertionslokus, die ein vollständiges GimA (~20.3 kb) (16.1%) (GimA+) und einer, die keinerlei GimA-assoziierte Strukturen (64.4%) (GimA-) enthält, eine dritte Variante des Lokus identifiziert werden, die ein 342 bp großes Restfragment von GimA aufweist (19.5%) (GimA-remnant).

Das vollständige GimA und das GimA-remnant waren fast ausschließlich mit der Phylogruppe B2 assoziiert. Zusätzlich trat das GimA-remnant bei Stämmen auf, die zum Multi-Lokus-Sequenztyp ST117 gehören und der rekombinanten Phylogruppe ABD zugeordnet sind. Die weiteren Phylogruppen (A, B1, AxB1, ABD (mit Ausnahme des ST117) und D) waren GimA-.

Weiterhin konnten signifikante Assoziationen zwischen dem Vorkommen des vollständigen GimA und den Pathovaren APEC und NMEC gefunden werden. Das GimA-remnant war dagegen signifikant assoziiert mit UPEC sowohl menschlichen als auch tierischen Ursprungs. Ähnlich der Rolle, die *ibeA* in der Neugeborenenmeningitis-Pathogenese spielt,

könnte dies einen möglichen Beitrag dieser Struktur zur Pathogenese von durch UPEC verursachten Harnwegsinfektionen andeuten.

Eine Sequenzanalyse des *ibeA*-Gens ergab eine strikte Assoziation des Gens mit dem phylogenetischen Hintergrund. Dies könnte ein Hinweis darauf sein, dass GimA ein originärer Bestandteil der Phylogruppe B2 ist.

Ferner zeigten die Untersuchungen, dass *ibeA* innerhalb eines definierten phylogenetischen Hintergrundes (ST95-Komplex) keinem positiven Selektionsdruck ausgesetzt ist. Dementsprechend scheint die Evolution von *ibeA* aufgrund von neutraler Selektion zu erfolgen. Obwohl einige Kriterien für genetische Inseln erfüllt werden, sprechen diese Befunde gegen eine Mobilität von GimA und folglich auch gegen die Annahme, GimA sei eine genetische Insel. GimA scheint eher ein ursprünglicher Teil der Phylogruppe B2 zu sein. Um die Zugehörigkeit von GimA zum „core genome“ zu verdeutlichen sollte diese genetische Region besser als GimA-Lokus angesprochen werden

Die Existenz von zwei weiteren Varianten des GimA-Lokus (GimA-remnant und GimA-) deutet auf eine Größenreduktion des Genoms in Form einer reduktiven Evolution hin. Dieser Prozess wird möglicherweise durch Veränderungen der pathogenetischen Bedeutung von einzelnen GimA-Komponenten in verschiedenen Habitaten verursacht. Bei von APEC und NMEC ausgelösten systemischen Infektionen konnten verschiedene Untersuchungen einen Beitrag zur Pathogenese zeigen. Einige GimA-Bestandteile sind für die Translokation der Blut-Hirn-Schranke von Bedeutung (*ibeRAT* – GimA4), während die übrigen in die Stressantwort des Erregers im Habitat „Blut“ involviert zu sein scheinen (GimA1-3).

Der Harnapparat ist im Gegensatz zum Blut ein mit weniger Stressoren für *E. coli* behaftetes Habitat. Entsprechend könnten Gene, die die Überwindung der Blut-Hirn-Schranke vermitteln, an Bedeutung für *E. coli* verloren haben und in einem Prozess der Genomoptimierung deletiert worden sein. Lediglich das GimA-remnant blieb erhalten. Dies deutet an, dass diese Struktur bzw. ihre Erhaltung einen Selektionsvorteil bietet. Das GimA-remnant konnte als ORF identifiziert werden und könnte folglich an der Pathogenese von Harnwegsinfektionen beteiligt sein. Weitere Untersuchungen besonders zur Funktionalität des GimA-remnant sind notwendig.

Die dritte Variante des GimA-Lokus kann auf zweierlei Wegen entstanden sein: (i) GimA ist in einem einzigen Schritt vollständig verloren gegangen oder (ii) im Anschluss an seine Entstehung das GimA-remnant ebenfalls deletiert worden. Im Falle von UPEC könnte letztere Möglichkeit das Resultat eines fortgesetzten Genomoptimierungsprozesses sein. Die resultierenden Stämme vermeiden eine Immunantwort des Wirtes und können somit als asymptomatische Besiedler im Harnapparat verbleiben.

## 8 SUMMARY

### **The GimA locus of extraintestinal pathogenic *E. coli* (ExPEC): reductive evolution correlates with habitat and pathovar**

Studies on human brain microvascular endothelial cells (HBMEC) and on *in vivo* models (newborn rats and chickens) demonstrated that in particular *ibeA* is involved in the pathogenesis of several extraintestinal pathogenic *E. coli* (ExPEC) pathovars, including newborn meningitic (NMEC) and avian pathogenic *E. coli* (APEC). E. g. it is required for translocation of the brain blood barrier, a crucial step in the pathogenesis of newborn meningitis. *IbeA* is located on a 20.3 kbp chromosomal region flanked by *yjiD* and *yjiE*, termed GimA. GimA has also been reported to be a genetic island, consisting of four operons (GimA1-4).

To unravel the phylogeny of GimA and to investigate its island character, the putative insertion locus of GimA was determined via Long Range PCR and DNA-DNA hybridization in 410 *E. coli* isolates, including APEC, NMEC, UPEC (uropathogenic), SepEC (septicemia associated *E. coli*) and human and animal fecal isolates as well as in 72 strains of the *E. coli* reference collection (EcoR).

In addition to a complete GimA (~20.3 kb) (16.1%) (GimA+) and a locus lacking GimA (64.4%) (GimA-) a third pattern containing a 342 bp remnant of GimA could be found (19.5%) (GimA-remnant).

The complete GimA and the GimA-remnant were almost exclusively associated with phylogroup B2. Additionally, the GimA remnant occurs among strains belonging to multilocus sequence type ST117 which is assigned to the recombinant phylogroup ABD. The remaining phylogroups (A, B1, AxB1, ABD (with the exception of ST117), and D) lack any GimA related structure.

Moreover, significant associations between the complete GimA and the pathovars APEC and NMEC on the one hand and between the GimA-remnant and UPEC strains on the other hand could be identified. Alike the known contribution of *ibeA* to the pathogenicity of newborn meningitis, these findings may suggest a possible involvement of this structure in uropathogenicity.

A detailed analysis of the *ibeA* sequences revealed a rigorous association to the phylogenetic background of the strains suggesting GimA originating from phylogroup B2. Moreover, this gene was not under positive selection in strains belonging to the ST95 complex. Thus, the evolution of *ibeA* seems to be driven by neutral selection, arguing against GimA being mobile. Although common criteria for genetic islands are partially fulfilled, GimA rather seems to be an ancestral part of phylogroup B2. In view of this affiliation

to the core genome instead of being an island, it would be appropriate to assign this genomic region the name GimA locus.

The existence of two other patterns reflects a genome reduction in a reductive evolution-like manner. The process of reduction in this genomic region may be driven by a shift in pathogenic importance of the GimA components in certain habitats. For systemic infections caused by APEC and NMEC a complete GimA is required. Some components mediate the entrance to the central nervous system (*ibeRAT* – GimA4), while the others are involved in stress response within the blood stream (GimA1-3). In contrast, the urinary tract is a completely different habitat exhibiting considerably fewer stressors for *E. coli*. Thus, genes allowing the translocation of the brain blood barrier as well as stress response genes may have lost its importance. Consequently, these genes might have been deleted in a genome optimization process and merely a remnant of GimA maintains. This structure may provide selective advantage and therefore maintains in the *E. coli*-chromosome. This GimA-remnant could be identified as an ORF and is highly associated with UPEC. Hence, it may be involved in uropathogenicity. Further investigations are needed to determine the function of the GimA-remnant. The third variant (GimA-) could have emerged from both (i) directly by the loss of the complete GimA in a one step deletion event and (ii) subsequently following the reduction process leading to the GimA-remnant, respectively. The latter one could in the case of UPEC be the result of an ongoing genome optimization process leading to attenuated strains. Such strains avoid to provoke host response and may therefore maintain in the urinary tract as asymptotically colonizers.

## 9 ANHANG

Tab. 27 Übersicht über die in der vorliegenden Arbeit verwendeten *E. coli*-Stämme inkl. dem zugeordneten Pathovar, ST, Alleltyp, ST-Komplex, Phylogruppe, Variante des GimA-Lokus und der Herkunft

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
APEC	IMT664	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
UPEC	IMT1200	353	36	24	9	13	17	11	14	ST73-Komplex	B2	GimA-remnant	Mensch
APEC	IMT1431	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1432	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1433	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1930	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1932	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1935	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1936	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT1937	155	6	4	14	16	24	8	14	ST155-Komplex	AxB1	GimA-	Tier
APEC	IMT1938	155	6	4	14	16	24	8	14	ST155-Komplex	AxB1	GimA-	Tier
APEC	IMT1939	155	6	4	14	16	24	8	14	ST155-Komplex	AxB1	GimA-	Tier
APEC	IMT2087	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2089	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2092	347	6	94	4	86	8	18	2	kein Komplex	AxB1	GimA-	Tier
APEC	IMT2094	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2095	162	9	65	5	1	9	13	6	ST469-Komplex	AxB1	GimA-	Tier
APEC	IMT2097	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2098	115	4	26	39	25	5	31	19	kein Komplex	D	GimA-	Tier
APEC	IMT2099	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT2101	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2102	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT2104	348	6	29	32	16	11	8	14	ST156-Komplex	B1	GimA-	Tier
APEC	IMT2105	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-	Tier
APEC	IMT2106	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT2108	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2111	38	4	26	2	25	5	5	19	ST38-Komplex	D	GimA-	Tier
APEC	IMT2112	355	36	24	10	13	17	10	25	kein Komplex	B2	GimA+	Tier
APEC	IMT2113	101	43	41	15	18	11	7	6	ST101-Komplex	B1	GimA-	Tier
APEC	IMT2120	356	6	97	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2121	357	13	40	13	13	23	25	66	kein Komplex	B2	GimA+	Tier
APEC	IMT2125	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2248	358	13	98	10	13	18	10	25	kein Komplex	B2	GimA+	Tier
APEC	IMT2250	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2251	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2254	126	13	43	19	37	17	25	25	kein Komplex	B2	GimA-	Tier
APEC	IMT2261	349	34	36	39	87	67	16	4	ST349-Komplex	D	GimA-	Tier
APEC	IMT2263	93	6	11	4	10	7	8	6	ST168-Komplex	AxB1	GimA-	Tier
APEC	IMT2264	350	6	31	83	28	1	1	2	ST350-Komplex	ABD	GimA-	Tier
APEC	IMT2265	351	6	95	4	88	7	7	7	kein Komplex	ABD	GimA-	Tier
APEC	IMT2278	140	55	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2283	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2288	140	55	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT2294	115	4	26	39	25	5	31	19	kein Komplex	D	GimA-	Tier
APEC	IMT2295	135	13	39	50	13	16	37	25	kein Komplex	B2	GimA+	Tier
APEC	IMT2297	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT2352	369	6	101	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT2467	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier

Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
APEC	IMT2469	140	55	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT2470	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT2477	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-remnant	Tier
APEC	IMT2489	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT2490	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT2491	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2492	140	55	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT2518	104	13	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
APEC	IMT2532	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT2545	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT4514	352	76	96	19	89	17	1	10	kein Komplex	B2	GimA+	Tier
APEC	IMT4516	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT4517	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT4518	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT4525	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT4529	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT4534	140	55	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT4537	118	31	4	42	44	15	33	17	kein Komplex	ABD	GimA-	Tier
APEC	IMT4541	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT5110	100	10	27	5	10	12	9	2	ST165-Komplex	A	GimA-	Tier
APEC	IMT5119	363	13	38	84	13	17	64	26	kein Komplex	B2	GimA+	Tier
APEC	IMT5132	365	43	41	15	18	11	11	6	ST101-Komplex	AxB1	GimA-	Tier
APEC	IMT5144	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-remnant	Tier
APEC	IMT5155	140	55	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT5215	93	6	11	4	10	7	8	6	ST168-Komplex	AxB1	GimA-	Tier
APEC	IMT5494	359	43	41	15	90	11	8	6	kein Komplex	AxB1	GimA-	Tier
UPEC	IMT6008	367	6	4	12	1	20	65	7	ST23-Komplex	B1	GimA-	Tier
UPEC	536	92	40	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Mensch
UPEC	J96	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Mensch
UPEC	764	14	14	14	10	14	17	7	10	ST14-Komplex	B2	GimA-	Mensch
UPEC	AD110	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
UPEC	RZ422	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Mensch
UPEC	RZ479	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Mensch
UPEC	U-7	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Mensch
UPEC	C97	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	C90	70	34	36	28	25	28	16	4	kein Komplex	D	GimA-	Mensch
UPEC	C89	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Mensch
UPEC	C70	14	14	14	10	14	17	7	10	ST14-Komplex	B2	GimA-	Mensch
UPEC	ABV84	360	6	4	12	1	20	7	7	ST23-Komplex	B1	GimA-	Mensch
APEC	2164	368	37	38	19	37	17	13	26	ST95-Komplex	B2	GimA-	Tier
APEC	2201	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA+	Tier
APEC	2446	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
UPEC	CFT073	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
APEC	IMT8994	355	36	24	10	13	17	10	25	kein Komplex	B2	GimA+	Tier
APEC	IMT8998	355	36	24	10	13	17	10	25	kein Komplex	B2	GimA+	Tier
APEC	IMT9003	369	6	101	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
UPEC	B6977-1	131	53	40	47	13	36	28	29	kein Komplex	B2	GimA-	Mensch
UPEC	B6985-2	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
UPEC	B6993	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-remnant	Mensch
UPEC	B6998	361	10	99	5	91	8	7	2	kein Komplex	AxB1	GimA-	Mensch
UPEC	B7001	362	62	100	17	31	5	5	4	kein Komplex	D	GimA-	Mensch
UPEC	IMT9096	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT9179	374	18	22	2	6	5	5	4	ST31-Komplex	D	GimA-	Tier
UPEC	IMT9206	375	36	24	9	13	17	11	68	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT9227	68	33	26	2	31	5	16	19	kein Komplex	D	GimA-	Tier
APEC	IMT9232	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT9238	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier

Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
APEC	IMT9241	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
APEC	IMT9242	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT9250	57	6	31	5	28	1	1	2	ST350-Komplex	ABD	GimA-	Tier
NMEC	IHE 3036	390	37	38	86	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
NMEC	IHE 3080	390	37	38	86	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
NMEC	BK658	80	13	24	19	14	23	1	10	ST568-Komplex	B2	GimA+	Mensch
NMEC	B10363	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
NMEC	B13155	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
NMEC	VE239	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
NMEC	VE1140	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
NMEC	E817	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-	Mensch
NMEC	E247	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Mensch
UPEC	IMT9280	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	IMT9289	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
APEC	MT78	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	V-Z	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	248	370	37	102	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	2363	135	13	39	50	13	16	37	25	kein Komplex	B2	GimA+	Tier
UPEC	IMT9402	104	13	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
APEC	IMT9581	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT9582	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT9583	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT9584	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Tier
APEC	IMT9586	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
APEC	IMT9588	371	6	31	83	28	1	1	67	ST350-Komplex	ABD	GimA-	Tier
UPEC	IMT9650	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
APEC	IMT9698	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Tier
APEC	IMT9706	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
APEC	IMT9712	373	6	11	4	10	7	8	2	ST168-Komplex	A	GimA-	Tier
APEC	IMT9713	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
Feces (Tier)	IMT10666	58	6	4	4	16	24	8	14	ST155-Komplex	AxB1	GimA-	Tier
Feces (Mensch)	δ1	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	δ8	416	37	110	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
Feces (Mensch)	δ9	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ11	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ15	421	37	38	19	37	17	8	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	δ18	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
UPEC	δ20	417	37	38	89	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
UPEC	δ21	418	37	38	19	37	72	11	26	ST95-Komplex	B2	GimA+	Mensch
UPEC	δ22	390	37	38	86	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
SepEC	δ27	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
SepEC	δ29	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ30	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ32	390	37	38	86	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
Feces (Mensch)	δ34	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ37	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
SepEC	δ39	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ54	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
Feces (Mensch)	δ58	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	δ62	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	δ64	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ67	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	δ68	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
Feces (Mensch)	E10097	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E10077	162	9	65	5	1	9	13	6	ST469-Komplex	AxB1	GimA-	Mensch
Feces (Mensch)	E10096	167	10	11	4	8	8	13	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E10099	34	10	11	4	1	8	8	2	ST10-Komplex	A	GimA-	Mensch

Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
Feces (Mensch)	E10100	405	35	37	29	25	4	5	73	ST405-Komplex	D	GimA-	Mensch
Feces (Mensch)	E10084	38	4	26	2	25	5	5	19	ST38-Komplex	D	GimA-	Mensch
Feces (Mensch)	E10082	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E10095	549	21	22	20	6	5	91	4	kein Komplex	D	GimA-	Mensch
Feces (Mensch)	E10085	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E10098	69	21	35	27	6	5	5	4	ST69-Komplex	D	GimA-	Mensch
Feces (Mensch)	E476	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E164	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E167	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E291	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E264	399	6	4	1	95	69	8	20	ST399-Komplex	AxB1	GimA-	Mensch
Feces (Mensch)	E7370	538	13	40	19	13	36	28	30	ST538-Komplex	B2	GimA+	Mensch
Feces (Mensch)	E424	38	4	26	2	25	5	5	19	ST38-Komplex	D	GimA-	Mensch
Feces (Mensch)	E452	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
Feces (Mensch)	E179	350	6	31	83	28	1	1	2	ST350-Komplex	ABD	GimA-	Mensch
Feces (Mensch)	E176	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
Feces (Mensch)	E175	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
SepEC	i484	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
SepEC	F1	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
SepEC	F287	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
SepEC	F385	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
SepEC	F557	23	6	4	12	1	20	13	7	ST23-Komplex	B1	GimA-	Mensch
SepEC	F560	544	13	137	9	13	16	10	9	ST12-Komplex	B2	GimA-	Mensch
SepEC	F569	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Mensch
SepEC	F645	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
SepEC	F742	539	6	19	57	18	9	13	6	kein Komplex	B1	GimA-	Mensch
SepEC	F775	540	6	7	57	1	8	8	2	kein Komplex	AxB1	GimA-	Mensch
SepEC	F785	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
SepEC	F905	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
SepEC	F911	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Mensch
SepEC	F1251	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
SepEC	St4723	297	6	65	32	26	9	8	2	kein Komplex	B1	GimA-	Mensch
SepEC	St5119	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-remnant	Mensch
SepEC	St5679	38	4	26	2	25	5	5	19	ST38-Komplex	D	GimA-	Mensch
SepEC	V9261	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Mensch
SepEC	V9343	216	10	11	57	8	7	18	6	kein Komplex	A	GimA-	Mensch
SepEC	V10744	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Mensch
SepEC	W7483	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Mensch
SepEC	W7716	545	62	100	111	31	5	5	4	kein Komplex	D	GimA-	Mensch
SepEC	W8987	542	112	11	5	12	8	8	86	kein Komplex	ABD	GimA-	Mensch
SepEC	W9763	46	8	7	1	8	8	8	6	ST46-Komplex	AxB1	GimA-	Mensch
SepEC	W9887	48	6	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	U2183	453	99	6	33	33	24	8	7	ST86-Komplex	ABD	GimA-	Mensch
UPEC	U2366	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	U2388	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
UPEC	U2873	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Mensch
UPEC	U3104	533	6	4	5	18	11	8	14	kein Komplex	B1	GimA-	Mensch
UPEC	U3145	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
UPEC	U3372	409	10	11	1	1	71	8	6	kein Komplex	AxB1	GimA-	Mensch
UPEC	U3407	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	U3292	130	18	22	20	6	5	5	4	ST31-Komplex	D	GimA-	Mensch
UPEC	U3362	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
UPEC	U3454	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
UPEC	U3633	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	U3622	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Mensch
UPEC	U4191	93	6	11	4	10	7	8	6	ST168-Komplex	AxB1	GimA-	Mensch
UPEC	U4252	48	6	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch

Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
UPEC	U4418	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	U4437	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-	Mensch
UPEC	U4417	398	64	7	1	1	8	8	6	ST398-Komplex	AxB1	GimA-	Mensch
UPEC	U4409	393	18	106	17	6	5	5	4	ST31-Komplex	D	GimA-	Mensch
UPEC	U5107	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	U5033	93	6	11	4	10	7	8	6	ST168-Komplex	AxB1	GimA-	Mensch
UPEC	U5070	69	21	35	27	6	5	5	4	ST69-Komplex	D	GimA-	Mensch
Feces (Mensch)	E10093	405	35	37	29	25	4	5	73	ST405-Komplex	D	GimA-	Mensch
Feces (Mensch)	E10089	543	83	136	110	117	80	1	2	kein Komplex	ABD	GimA-	Mensch
Feces (Mensch)	E10092	536	6	107	1	95	69	8	20	ST399-Komplex	ABD	GimA-	Mensch
Feces (Mensch)	E10086	541	111	23	109	8	8	8	2	ST522-Komplex	A	GimA-	Mensch
Feces (Mensch)	E10094	126	13	43	19	37	17	25	25	kein Komplex	B2	GimA+	Mensch
Feces (Mensch)	E10090	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E10091	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-remnant	Mensch
Feces (Mensch)	E472	68	33	26	2	31	5	16	19	kein Komplex	D	GimA-	Mensch
Feces (Mensch)	E475	537	14	14	10	14	17	11	10	ST14-Komplex	B2	GimA+	Mensch
Feces (Mensch)	E460	38	4	26	2	25	5	5	19	ST38-Komplex	D	GimA-	Mensch
Feces (Mensch)	E165	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	E166	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E292	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E345	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E182	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Mensch
Feces (Mensch)	E422	547	13	52	10	119	17	37	25	kein Komplex	B2	GimA+	Mensch
Feces (Mensch)	E444	548	10	11	4	8	81	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E478	428	96	40	13	100	23	28	66	kein Komplex	B2	GimA+	Mensch
Feces (Mensch)	E457	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	E177	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	E178	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
Feces (Mensch)	E180	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	E173	393	18	106	17	6	5	5	4	ST31-Komplex	D	GimA-	Mensch
Feces (Mensch)	E174	348	6	29	32	16	11	8	14	ST156-Komplex	B1	GimA-	Mensch
Feces (Mensch)	E464	550	14	14	10	14	17	92	10	ST14-Komplex	B2	GimA-	Mensch
Feces (Mensch)	E471	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
Feces (Mensch)	E467	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Mensch
UPEC	IMT14965	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT4966	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
NMEC	RS218	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
UPEC	IMT14967	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14968	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14969	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT15014	117	20	45	41	43	5	32	2	kein Komplex	ABD	GimA-remnant	Tier
UPEC	IMT15015	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
NMEC	RS176	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
NMEC	C183-77	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
NMEC	IHE3055	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
NMEC	A459	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
NMEC	RS179	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
UPEC	14970	104	13	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	15016	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA+	Tier
NMEC	C5	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
NMEC	EC10	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
NMEC	A90	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
NMEC	RS167	144	13	43	9	36	30	44	25	kein Komplex	B2	GimA+	Mensch
NMEC	RS168	59	27	32	24	29	26	19	22	ST59-Komplex	ABD	GimA-	Mensch
NMEC	S88	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
NMEC	S95	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Mensch
NMEC	E253	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch

Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
NMEC	E334	144	13	43	9	36	30	44	25	kein Komplex	B2	GimA+	Mensch
UPEC	IMT15017	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
UPEC	15019	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT15021	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Tier
UPEC	IMT15022	115	4	26	39	25	5	31	19	kein Komplex	D	GimA-	Tier
UPEC	IMT15024	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT15025	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT15028	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Tier
UPEC	IMT14972	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
UPEC	IMT15029	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
UPEC	IMT15030	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
UPEC	IMT15031	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT15033	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14973	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14974	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14955	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Tier
UPEC	IMT14975	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT14976	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14977	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14978	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT14979	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Tier
UPEC	IMT14980	638	76	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Tier
UPEC	IMT14981	638	76	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14983	539	6	19	57	18	9	13	6	kein Komplex	B1	GimA-	Tier
UPEC	IMT14984	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
UPEC	IMT14958	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14959	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14985	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Tier
UPEC	IMT14986	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA+	Tier
UPEC	IMT14987	75	6	6	5	10	20	23	6	kein Komplex	AxB1	GimA-	Tier
UPEC	IMT14989	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA+	Tier
UPEC	IMT14990	625	13	43	13	14	30	94	92	kein Komplex	B2	GimA+	Tier
UPEC	IMT14991	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Tier
UPEC	IMT14961	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Tier
UPEC	IMT14993	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT14994	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
UPEC	IMT14995	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14963	646	13	96	121	13	17	11	10	kein Komplex	B2	GimA+	Tier
UPEC	IMT14964	646	13	96	121	13	17	11	10	kein Komplex	B2	GimA+	Tier
UPEC	IMT14997	372	88	103	19	36	23	44	26	kein Komplex	B2	GimA+	Tier
UPEC	IMT14998	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT14999	681	38	39	30	13	17	25	28	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT15000	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA-	Tier
UPEC	IMT15001	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA+	Tier
UPEC	IMT15002	410	6	4	12	1	20	18	7	ST23-Komplex	B1	GimA-	Tier
UPEC	IMT15003	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT15004	127	13	14	19	36	23	11	10	kein Komplex	B2	GimA-remnant	Tier
NMEC	IHE3034	95	37	38	19	37	17	11	26	ST95-Komplex	B2	GimA+	Mensch
UPEC	IMT15006	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
UPEC	IMT15007	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA-	Tier
UPEC	IMT15008	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
UPEC	IMT15009	80	13	24	19	14	23	1	10	ST568-Komplex	B2	GimA+	Tier
UPEC	IMT15010	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Tier
UPEC	IMT15011	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
UPEC	IMT15013	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Tier
UPEC	IMT14953	925	18	22	111	6	5	5	4	kein Komplex	D	GimA-	Tier
UPEC	IMT14954	141	13	52	10	14	17	25	17	kein Komplex	B2	GimA+	Tier

## Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
UPEC	IMT14982	959	13	13	9	13	16	10	119	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT14988	929	13	43	49	37	17	44	96	kein Komplex	B2	GimA+	Tier
UPEC	IMT15005	961	13	13	9	13	16	126	9	kein Komplex	B2	GimA-remnant	Tier
UPEC	IMT15020	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
UPEC	IMT15027	937	6	29	5	26	24	8	6	kein Komplex	AxB1	GimA-	Tier
Feces (Mensch)	RM74A	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	STM1	49	25	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Tier)	W1R1(a)	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
Feces (Mensch)	RM39A	63	30	11	5	8	7	21	2	kein Komplex	A	GimA-	Mensch
Feces (Mensch)	RM60A	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	RM66C	77	6	7	5	1	8	18	27	ST206-Komplex	AxB1	GimA-	Mensch
Feces (Tier)	RM73C	87	6	11	14	8	20	8	2	kein Komplex	AxB1	GimA-	Tier
Feces (Mensch)	RM77C	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	FN98	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	ANI	43	24	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
UPEC	C97	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	FN59	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	FN10	44	10	11	4	8	8	8	7	ST10-Komplex	A	GimA-	Mensch
UPEC	P62	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Mensch
Feces (Mensch)	FN3	45	6	7	23	1	8	8	2	kein Komplex	AxB1	GimA-	Mensch
Feces (Tier)	RM191F	46	8	7	1	8	8	8	6	ST46-Komplex	AxB1	GimA-	Tier
Feces (Tier)	RM200Q	47	8	11	5	1	7	18	6	kein Komplex	AxB1	GimA-	Tier
Feces (Tier)	RM210F	48	6	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
Feces (Tier)	RM210J	48	6	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
Feces (Tier)	RM213J	48	6	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
Feces (Tier)	RM213K	48	6	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
Feces (Tier)	RM215C	50	26	7	4	1	8	8	6	kein Komplex	A	GimA-	Tier
Feces (Tier)	RM183E	51	6	4	1	13	17	8	20	kein Komplex	ABD	GimA-	Tier
Feces (Mensch)	FN33	52	6	11	5	1	20	8	2	kein Komplex	AxB1	GimA-	Mensch
Feces (Tier)	MS1	10	10	11	4	8	8	8	2	ST10-Komplex	A	GimA-	Tier
Feces (Mensch)	LL	53	9	29	12	1	11	13	6	kein Komplex	AxB1	GimA-	Mensch
Feces (Tier)	RM24J	53	9	29	12	1	11	13	6	kein Komplex	AxB1	GimA-	Tier
Feces (Mensch)	RM52B	54	6	30	5	1	24	8	21	kein Komplex	AxB1	GimA-	Mensch
Feces (Tier)	RM3A	55	6	4	4	18	20	8	14	ST155-Komplex	AxB1	GimA-	Tier
Feces (Tier)	RM10A	56	6	4	4	18	24	5	14	ST155-Komplex	ABD	GimA-	Tier
Feces (Tier)	RM12	57	6	31	5	28	1	1	2	ST350-Komplex	ABD	GimA-	Tier
Feces (Tier)	RM28	56	6	4	4	18	24	5	14	ST155-Komplex	ABD	GimA-	Tier
Feces (Tier)	RM56C	56	6	4	4	18	24	5	14	ST155-Komplex	ABD	GimA-	Tier
Feces (Tier)	WIR2(a)	58	6	4	4	16	24	8	14	ST155-Komplex	AxB1	GimA-	Tier
Feces (Mensch)	RM42B	59	27	32	24	29	26	19	22	ST59-Komplex	ABD	GimA-	Mensch
Feces (Mensch)	RM77B	60	28	33	24	29	26	20	2	kein Komplex	ABD	GimA-	Mensch
Feces (Tier)	RM44B	61	29	12	8	12	15	2	23	ST11-Komplex	D	GimA-	Tier
Feces (Mensch)	RM75A	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
Feces (Mensch)	FN104	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
UPEC	P60	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
Feces (Mensch)	T44	62	28	33	25	29	7	11	24	kein Komplex	ABD	GimA-	Mensch
Feces (Mensch)	DAR1	64	31	34	26	30	27	22	17	kein Komplex	ABD	GimA-	Mensch
Feces (Mensch)	FN36	65	28	33	25	29	1	11	2	kein Komplex	ABD	GimA-	Mensch
Feces (Tier)	RM1891	66	6	22	20	6	5	5	2	kein Komplex	D	GimA-	Tier
Feces (Tier)	RM201C	67	6	19	5	1	9	8	7	kein Komplex	AxB1	GimA-	Tier
Feces (Tier)	RM202F	68	33	26	2	31	5	16	19	kein Komplex	D	GimA-	Tier
Feces (Tier)	RM211C	69	21	35	27	6	5	5	4	ST69-Komplex	D	GimA-	Tier
Feces (Mensch)	C90/28512	70	34	36	28	25	28	16	4	kein Komplex	D	GimA-	Mensch
Feces (Mensch)	FN90	71	34	37	28	25	4	16	4	kein Komplex	D	GimA-	Mensch
UPEC	FN90P97	72	35	37	29	25	4	5	18	ST405-Komplex	D	GimA-	Mensch
Feces (Mensch)	DD	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
Feces (Tier)	RM73A	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier

Anhang

Pathovar	Bezeichnung	ST	Alleltyp							ST-Komplex	Phylogruppe	GimA-Lokus-Variante	Herkunft
			<i>adk</i>	<i>fumC</i>	<i>gyrB</i>	<i>icd</i>	<i>mdh</i>	<i>purA</i>	<i>recA</i>				
Feces (Mensch)	RM33B	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-	Mensch
Feces (Mensch)	RM64A	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Mensch
UPEC	FN4	74	36	24	9	13	17	5	25	ST73-Komplex	B2	GimA-remnant	Mensch
Feces (Mensch)	P106	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-	Mensch
Feces (Tier)	RM71B	73	36	24	9	13	17	11	25	ST73-Komplex	B2	GimA-remnant	Tier
Feces (Tier)	RM185S	75	6	6	5	10	20	23	6	kein Komplex	AxB1	GimA-	Tier
Feces (Mensch)	SIL8	76	13	13	9	13	17	10	26	kein Komplex	B2	GimA-remnant	Mensch
UPEC	C89	12	13	13	9	13	16	10	9	ST12-Komplex	B2	GimA-remnant	Mensch
Feces (Mensch)	FN83	78	37	38	19	32	17	24	26	kein Komplex	B2	GimA-	Mensch
UPEC	P69	79	37	38	19	32	17	7	26	kein Komplex	B2	GimA-	Mensch
Feces (Mensch)	FN21	80	13	24	19	14	23	1	10	ST568-Komplex	B2	GimA+	Mensch
UPEC	C70	81	14	14	10	14	17	7	21	ST14-Komplex	B2	GimA-	Mensch
Feces (Tier)	RM202I	82	38	39	30	13	29	25	28	kein Komplex	B2	GimA-remnant	Tier
Feces (Tier)	RM209I	83	13	38	19	13	30	25	29	kein Komplex	B2	GimA+	Tier
Feces (Tier)	RM217T	84	6	29	31	16	24	8	2	kein Komplex	AxB1	GimA-	Tier
Feces (Tier)	RM224H	85	6	6	32	26	11	26	7	kein Komplex	B1	GimA-	Tier
Feces (Tier)	RM45EM	86	9	6	33	33	24	8	7	ST86-Komplex	ABD	GimA-	Tier
Feces (Tier)	RM70B	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Tier
UPEC	ABV84	88	6	4	12	1	20	12	7	ST23-Komplex	B1	GimA-	Mensch
UPEC	P68	89	39	7	4	34	20	7	7	kein Komplex	AxB1	GimA-	Mensch

Tab. 28 Übersicht der Ergebnisse der Untersuchungen des GimA-Lokus mittels PCR und DNS-DNS-Hybridisierung sowie der daraus resultierenden GimA-Lokus-Variante der 410 *E. coli*-Stämme (n. a. = nicht angewendet; + = spezifische DNS-Sequenz mittels Sonde erkannt; - = spezifische DNS-Sequenz mittels Sonde nicht erkannt)

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cniT/gclK</i>	<i>ibeA</i>	
IMT664	340bp	n.a.	-	-	-	-	-	GimA-
IMT1200	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT1431	340bp	n.a.	-	-	-	-	-	GimA-
IMT1432	340bp	n.a.	-	-	-	-	-	GimA-
IMT1433	340bp	n.a.	-	-	-	-	-	GimA-
IMT1930	340bp	n.a.	-	-	-	-	-	GimA-
IMT1932	340bp	n.a.	-	-	-	-	-	GimA-
IMT1935	340bp	n.a.	-	-	-	-	-	GimA-
IMT1936	340bp	n.a.	-	-	-	-	-	GimA-
IMT1937	340bp	n.a.	-	-	-	-	-	GimA-
IMT1938	340bp	n.a.	-	-	-	-	-	GimA-
IMT1939	340bp	n.a.	-	-	-	-	-	GimA-
IMT2087	340bp	n.a.	-	-	-	-	-	GimA-
IMT2089	340bp	n.a.	-	-	-	-	-	GimA-
IMT2092	340bp	n.a.	-	-	-	-	-	GimA-
IMT2094	340bp	n.a.	-	-	-	-	-	GimA-
IMT2095	340bp	n.a.	-	-	-	-	-	GimA-
IMT2097	340bp	n.a.	-	-	-	-	-	GimA-
IMT2098	340bp	n.a.	-	-	-	-	-	GimA-
IMT2099	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2101	340bp	n.a.	-	-	-	-	-	GimA-
IMT2102	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT2104	340bp	n.a.	-	-	-	-	-	GimA-
IMT2105	340bp	n.a.	-	-	-	-	-	GimA-
IMT2106	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT2108	340bp	n.a.	-	-	-	-	-	GimA-
IMT2111	340bp	n.a.	-	-	-	-	-	GimA-
IMT2112	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2113	340bp	n.a.	-	-	-	-	-	GimA-
IMT2120	340bp	n.a.	-	-	-	-	-	GimA-
IMT2121	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2125	340bp	n.a.	-	-	-	-	-	GimA-
IMT2248	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2250	340bp	n.a.	-	-	-	-	-	GimA-
IMT2251	340bp	n.a.	-	-	-	-	-	GimA-
IMT2254	340bp	n.a.	-	-	-	-	-	GimA-
IMT2261	340bp	n.a.	-	-	-	-	-	GimA-
IMT2263	340bp	n.a.	-	-	-	-	-	GimA-
IMT2264	340bp	n.a.	-	-	-	-	-	GimA-
IMT2265	340bp	n.a.	-	-	-	-	-	GimA-
IMT2278	340bp	n.a.	-	-	-	-	-	GimA-
IMT2283	340bp	n.a.	-	-	-	-	-	GimA-
IMT2288	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2294	340bp	n.a.	-	-	-	-	-	GimA-
IMT2295	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2297	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2352	340bp	n.a.	-	-	-	-	-	GimA-

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cnIT/gclK</i>	<i>ibeA</i>	
IMT2467	340bp	n.a.	-	-	-	-	-	GimA-
IMT2469	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2470	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2477	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT2489	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT2490	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT2491	340bp	n.a.	-	-	-	-	-	GimA-
IMT2492	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT2518	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT2532	340bp	n.a.	-	-	-	-	-	GimA-
IMT2545	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT4514	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT4516	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT4517	340bp	n.a.	-	-	-	-	-	GimA-
IMT4518	340bp	n.a.	-	-	-	-	-	GimA-
IMT4525	340bp	n.a.	-	-	-	-	-	GimA-
IMT4529	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT4534	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT4537	340bp	n.a.	-	-	-	-	-	GimA-
IMT4541	340bp	n.a.	-	-	-	-	-	GimA-
IMT5110	340bp	n.a.	-	-	-	-	-	GimA-
IMT5119	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT5132	340bp	n.a.	-	-	-	-	-	GimA-
IMT5144	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT5155	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT5215	340bp	n.a.	-	-	-	-	-	GimA-
IMT5494	340bp	n.a.	-	-	-	-	-	GimA-
IMT6008	340bp	n.a.	-	-	-	-	-	GimA-
536	740bp	n.a.	+	-	-	-	-	GimA-remnant
J96	340bp	n.a.	-	-	-	-	-	GimA-
764	340bp	n.a.	-	-	-	-	-	GimA-
AD110	740bp	n.a.	+	-	-	-	-	GimA-remnant
RZ422	740bp	n.a.	+	-	-	-	-	GimA-remnant
RZ479	740bp	n.a.	+	-	-	-	-	GimA-remnant
U-7	740bp	n.a.	+	-	-	-	-	GimA-remnant
C97	340bp	n.a.	-	-	-	-	-	GimA-
C90	340bp	n.a.	-	-	-	-	-	GimA-
C89	740bp	n.a.	+	-	-	-	-	GimA-remnant
C70	340bp	n.a.	-	-	-	-	-	GimA-
ABV84	340bp	n.a.	-	-	-	-	-	GimA-
2164	340bp	n.a.	-	-	-	-	-	GimA-
2201	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
2446	340bp	n.a.	-	-	-	-	-	GimA-
CFT073	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT8994	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT8998	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT9003	340bp	n.a.	-	-	-	-	-	GimA-
B6977-1	340bp	n.a.	-	-	-	-	-	GimA-
B6985-2	740bp	n.a.	+	-	-	-	-	GimA-remnant
B6993	740bp	n.a.	+	-	-	-	-	GimA-remnant
B6998	340bp	n.a.	-	-	-	-	-	GimA-

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cniT/gclK</i>	<i>ibeA</i>	
B7001	340bp	n.a.	-	-	-	-	-	GimA-
IMT9096	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9179	340bp	n.a.	-	-	-	-	-	GimA-
IMT9206	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9227	340bp	n.a.	-	-	-	-	-	GimA-
IMT9232	340bp	n.a.	-	-	-	-	-	GimA-
IMT9238	340bp	n.a.	-	-	-	-	-	GimA-
IMT9241	340bp	n.a.	-	-	-	-	-	GimA-
IMT9242	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9250	340bp	n.a.	-	-	-	-	-	GimA-
IHE3036	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IHE3080	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
BK658	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
B10363	340bp	n.a.	-	-	-	-	-	GimA-
B13155	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
VE239	340bp	n.a.	-	-	-	-	-	GimA-
VE1140	340bp	n.a.	-	-	-	-	-	GimA-
E817	340bp	n.a.	-	-	-	-	-	GimA-
E247	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9280	340bp	n.a.	-	-	-	-	-	GimA-
IMT9289	740bp	n.a.	+	-	-	-	-	GimA-remnant
MT78	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
V-Z	740bp	n.a.	+	-	-	-	-	GimA-remnant
248	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
2363	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT9402	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9581	340bp	n.a.	-	-	-	-	-	GimA-
IMT9582	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9583	340bp	n.a.	-	-	-	-	-	GimA-
IMT9584	340bp	n.a.	-	-	-	-	-	GimA-
IMT9586	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT9588	340bp	n.a.	-	-	-	-	-	GimA-
IMT9650	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT9698	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT9706	340bp	n.a.	-	-	-	-	-	GimA-
IMT9712	340bp	n.a.	-	-	-	-	-	GimA-
IMT9713	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT10666	340bp	n.a.	-	-	-	-	-	GimA-
δ1	340bp	n.a.	-	-	-	-	-	GimA-
δ8	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ9	340bp	n.a.	-	-	-	-	-	GimA-
δ11	340bp	n.a.	-	-	-	-	-	GimA-
δ15	340bp	n.a.	-	-	-	-	-	GimA-
δ18	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ20	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ21	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ22	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ27	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ29	340bp	n.a.	-	-	-	-	-	GimA-
δ30	340bp	n.a.	-	-	-	-	-	GimA-
δ32	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cnIT/gclK</i>	<i>ibeA</i>	
δ34	340bp	n.a.	-	-	-	-	-	GimA-
δ37	340bp	n.a.	-	-	-	-	-	GimA-
δ39	340bp	n.a.	-	-	-	-	-	GimA-
δ54	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
δ58	340bp	n.a.	-	-	-	-	-	GimA-
δ62	340bp	n.a.	-	-	-	-	-	GimA-
δ64	340bp	n.a.	-	-	-	-	-	GimA-
δ67	340bp	n.a.	-	-	-	-	-	GimA-
δ68	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E 10097	340bp	n.a.	-	-	-	-	-	GimA-
E10077	340bp	n.a.	-	-	-	-	-	GimA-
E10096	340bp	n.a.	-	-	-	-	-	GimA-
E10099	340bp	n.a.	-	-	-	-	-	GimA-
E10100	340bp	n.a.	-	-	-	-	-	GimA-
E10084	340bp	n.a.	-	-	-	-	-	GimA-
E10082	340bp	n.a.	-	-	-	-	-	GimA-
E10095	340bp	n.a.	-	-	-	-	-	GimA-
E10085	340bp	n.a.	-	-	-	-	-	GimA-
E10098	340bp	n.a.	-	-	-	-	-	GimA-
E476	340bp	n.a.	-	-	-	-	-	GimA-
E164	340bp	n.a.	-	-	-	-	-	GimA-
E167	340bp	n.a.	-	-	-	-	-	GimA-
E291	340bp	n.a.	-	-	-	-	-	GimA-
E264	340bp	n.a.	-	-	-	-	-	GimA-
E7370	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E424	340bp	n.a.	-	-	-	-	-	GimA-
E452	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E179	340bp	n.a.	-	-	-	-	-	GimA-
E176	740bp	n.a.	+	-	-	-	-	GimA-remnant
E175	340bp	n.a.	-	-	-	-	-	GimA-
i484	740bp	n.a.	+	-	-	-	-	GimA-remnant
F1	740bp	n.a.	+	-	-	-	-	GimA-remnant
F287	340bp	n.a.	-	-	-	-	-	GimA-
F385	740bp	n.a.	+	-	-	-	-	GimA-remnant
F557	340bp	n.a.	-	-	-	-	-	GimA-
F560	340bp	n.a.	-	-	-	-	-	GimA-
F569	340bp	n.a.	-	-	-	-	-	GimA-
F645	340bp	n.a.	-	-	-	-	-	GimA-
F742	340bp	n.a.	-	-	-	-	-	GimA-
F775	340bp	n.a.	-	-	-	-	-	GimA-
F785	340bp	n.a.	-	-	-	-	-	GimA-
F905	340bp	n.a.	-	-	-	-	-	GimA-
F911	740bp	n.a.	+	-	-	-	-	GimA-remnant
F1251	340bp	n.a.	-	-	-	-	-	GimA-
St4723	340bp	n.a.	-	-	-	-	-	GimA-
St5119	740bp	n.a.	+	-	-	-	-	GimA-remnant
St5679	340bp	n.a.	-	-	-	-	-	GimA-
V9261	340bp	n.a.	-	-	-	-	-	GimA-
V9343	340bp	n.a.	-	-	-	-	-	GimA-
V10744	340bp	n.a.	-	-	-	-	-	GimA-
W7483	340bp	n.a.	-	-	-	-	-	GimA-

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cnIT/gclK</i>	<i>ibeA</i>	
W7716	340bp	n.a.	-	-	-	-	-	GimA-
W8987	340bp	n.a.	-	-	-	-	-	GimA-
W9763	340bp	n.a.	-	-	-	-	-	GimA-
W9887	340bp	n.a.	-	-	-	-	-	GimA-
U2183	340bp	n.a.	-	-	-	-	-	GimA-
U2366	340bp	n.a.	-	-	-	-	-	GimA-
U2388	740bp	n.a.	+	-	-	-	-	GimA-remnant
U2873	740bp	n.a.	+	-	-	-	-	GimA-remnant
U3104	340bp	n.a.	-	-	-	-	-	GimA-
U3145	740bp	n.a.	+	-	-	-	-	GimA-remnant
U3372	340bp	n.a.	-	-	-	-	-	GimA-
U3407	340bp	n.a.	-	-	-	-	-	GimA-
U3292	340bp	n.a.	-	-	-	-	-	GimA-
U3362	740bp	n.a.	+	-	-	-	-	GimA-remnant
U3454	340bp	n.a.	-	-	-	-	-	GimA-
U3633	340bp	n.a.	-	-	-	-	-	GimA-
U3622	340bp	n.a.	-	-	-	-	-	GimA-
U4191	340bp	n.a.	-	-	-	-	-	GimA-
U4252	340bp	n.a.	-	-	-	-	-	GimA-
U4418	340bp	n.a.	-	-	-	-	-	GimA-
U4437	340bp	n.a.	-	-	-	-	-	GimA-
U4417	340bp	n.a.	-	-	-	-	-	GimA-
U4409	340bp	n.a.	-	-	-	-	-	GimA-
U5107	340bp	n.a.	-	-	-	-	-	GimA-
U5033	340bp	n.a.	-	-	-	-	-	GimA-
U5070	340bp	n.a.	-	-	-	-	-	GimA-
E10093	340bp	n.a.	-	-	-	-	-	GimA-
E10089	340bp	n.a.	-	-	-	-	-	GimA-
E10092	340bp	n.a.	-	-	-	-	-	GimA-
E10086	340bp	n.a.	-	-	-	-	-	GimA-
E10094	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E10090	340bp	n.a.	-	-	-	-	-	GimA-
E10091	740bp	n.a.	+	-	-	-	-	GimA-remnant
E472	340bp	n.a.	-	-	-	-	-	GimA-
E475	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E460	340bp	n.a.	-	-	-	-	-	GimA-
E165	340bp	n.a.	-	-	-	-	-	GimA-
E166	340bp	n.a.	-	-	-	-	-	GimA-
E292	340bp	n.a.	-	-	-	-	-	GimA-
E345	340bp	n.a.	-	-	-	-	-	GimA-
E182	340bp	n.a.	-	-	-	-	-	GimA-
E422	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E444	340bp	n.a.	-	-	-	-	-	GimA-
E478	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
E457	340bp	n.a.	-	-	-	-	-	GimA-
E177	340bp	n.a.	-	-	-	-	-	GimA-
E178	340bp	n.a.	-	-	-	-	-	GimA-
E180	340bp	n.a.	-	-	-	-	-	GimA-
E173	340bp	n.a.	-	-	-	-	-	GimA-
E174	340bp	n.a.	-	-	-	-	-	GimA-
E464	340bp	n.a.	-	-	-	-	-	GimA-

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cniT/gclK</i>	<i>ibeA</i>	
E471	740bp	n.a.	+	-	-	-	-	GimA-remnant
E467	340bp	n.a.	-	-	-	-	-	GimA-
IMT14965	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14966	740bp	n.a.	+	-	-	-	-	GimA-remnant
RS218	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14967	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14968	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14969	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15014	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15015	740bp	n.a.	+	-	-	-	-	GimA-remnant
RS176	340bp	n.a.	-	-	-	-	-	GimA-
C183-77	340bp	n.a.	-	-	-	-	-	GimA-
IHE3055	340bp	n.a.	-	-	-	-	-	GimA-
BE1-2928	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
RS179	340bp	n.a.	-	-	-	-	-	GimA-
IMT14970	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15016	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
C5	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
EC10	340bp	n.a.	-	-	-	-	-	GimA-
A90	340bp	n.a.	-	-	-	-	-	GimA-
RS167	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
RS168	340bp	n.a.	-	-	-	-	-	GimA-
S88	340bp	n.a.	-	-	-	-	-	GimA-
S95	340bp	n.a.	-	-	-	-	-	GimA-
E253	340bp	n.a.	-	-	-	-	-	GimA-
E334	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT15017	340bp	n.a.	-	-	-	-	-	GimA-
IMT15019	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15021	340bp	n.a.	-	-	-	-	-	GimA-
IMT15022	340bp	n.a.	-	-	-	-	-	GimA-
IMT15024	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15025	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15028	340bp	n.a.	-	-	-	-	-	GimA-
IMT14972	340bp	n.a.	-	-	-	-	-	GimA-
IMT15029	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT15030	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT15031	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15033	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14973	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14974	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14955	340bp	n.a.	-	-	-	-	-	GimA-
IMT14975	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14976	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14977	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14978	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14979	340bp	n.a.	-	-	-	-	-	GimA-
IMT14980	340bp	n.a.	-	-	-	-	-	GimA-
IMT14981	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14983	340bp	n.a.	-	-	-	-	-	GimA-
IMT14984	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14958	740bp	n.a.	+	-	-	-	-	GimA-remnant

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cnIT/gclK</i>	<i>ibeA</i>	
IMT14959	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14985	340bp	n.a.	-	-	-	-	-	GimA-
IMT14986	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14987	340bp	n.a.	-	-	-	-	-	GimA-
IMT14989	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14990	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14991	340bp	n.a.	-	-	-	-	-	GimA-
IMT14961	340bp	n.a.	-	-	-	-	-	GimA-
IMT14993	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14994	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14995	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14963	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14964	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14997	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14998	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14999	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15000	340bp	n.a.	-	-	-	-	-	GimA-
IMT15001	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT5002	340bp	n.a.	-	-	-	-	-	GimA-
IMT15003	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15004	740bp	n.a.	+	-	-	-	-	GimA-remnant
IHE3034	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT15006	340bp	n.a.	-	-	-	-	-	GimA-
IMT15007	340bp	n.a.	-	-	-	-	-	GimA-
IMT15008	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15009	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT15010	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15011	340bp	n.a.	-	-	-	-	-	GimA-
IMT15013	340bp	n.a.	-	-	-	-	-	GimA-
IMT14953	340bp	n.a.	-	-	-	-	-	GimA-
IMT14954	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT14982	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT14988	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
IMT15005	740bp	n.a.	+	-	-	-	-	GimA-remnant
IMT15020	340bp	n.a.	-	-	-	-	-	GimA-
IMT15027	340bp	n.a.	-	-	-	-	-	GimA-
RM74A	340bp	n.a.	-	-	-	-	-	GimA-
STM1	340bp	n.a.	-	-	-	-	-	GimA-
W1R1(a)	340bp	n.a.	-	-	-	-	-	GimA-
RM39A	340bp	n.a.	-	-	-	-	-	GimA-
RM60A	340bp	n.a.	-	-	-	-	-	GimA-
RM66C	340bp	n.a.	-	-	-	-	-	GimA-
RM73C	340bp	n.a.	-	-	-	-	-	GimA-
RM77C	340bp	n.a.	-	-	-	-	-	GimA-
FN98	340bp	n.a.	-	-	-	-	-	GimA-
ANI	340bp	n.a.	-	-	-	-	-	GimA-
C97	340bp	n.a.	-	-	-	-	-	GimA-
FN59	340bp	n.a.	-	-	-	-	-	GimA-
FN10	340bp	n.a.	-	-	-	-	-	GimA-
P62	340bp	n.a.	-	-	-	-	-	GimA-
FN3	340bp	n.a.	-	-	-	-	-	GimA-

Anhang

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cniT/gclK</i>	<i>ibeA</i>	
RM191F	340bp	n.a.	-	-	-	-	-	GimA-
RM200Q	340bp	n.a.	-	-	-	-	-	GimA-
RM210F	340bp	n.a.	-	-	-	-	-	GimA-
RM210J	340bp	n.a.	-	-	-	-	-	GimA-
RM213J	340bp	n.a.	-	-	-	-	-	GimA-
RM213K	340bp	n.a.	-	-	-	-	-	GimA-
RM215C	340bp	n.a.	-	-	-	-	-	GimA-
RM183E	340bp	n.a.	-	-	-	-	-	GimA-
FN33	340bp	n.a.	-	-	-	-	-	GimA-
MS1	340bp	n.a.	-	-	-	-	-	GimA-
LL	340bp	n.a.	-	-	-	-	-	GimA-
RM24J	340bp	n.a.	-	-	-	-	-	GimA-
RM52B	340bp	n.a.	-	-	-	-	-	GimA-
RM3A	340bp	n.a.	-	-	-	-	-	GimA-
RM10A	340bp	n.a.	-	-	-	-	-	GimA-
RM12	340bp	n.a.	-	-	-	-	-	GimA-
RM28	340bp	n.a.	-	-	-	-	-	GimA-
RM56C	340bp	n.a.	-	-	-	-	-	GimA-
WIR2(a)	340bp	n.a.	-	-	-	-	-	GimA-
RM42B	340bp	n.a.	-	-	-	-	-	GimA-
RM77B	340bp	n.a.	-	-	-	-	-	GimA-
RM44B	340bp	n.a.	-	-	-	-	-	GimA-
RM75A	340bp	n.a.	-	-	-	-	-	GimA-
FN104	340bp	n.a.	-	-	-	-	-	GimA-
P60	340bp	n.a.	-	-	-	-	-	GimA-
T44	340bp	n.a.	-	-	-	-	-	GimA-
DAR1	340bp	n.a.	-	-	-	-	-	GimA-
FN36	340bp	n.a.	-	-	-	-	-	GimA-
RM1891	340bp	n.a.	-	-	-	-	-	GimA-
RM201C	340bp	n.a.	-	-	-	-	-	GimA-
RM202F	340bp	n.a.	-	-	-	-	-	GimA-
RM211C	340bp	n.a.	-	-	-	-	-	GimA-
C90/28512	340bp	n.a.	-	-	-	-	-	GimA-
FN90	340bp	n.a.	-	-	-	-	-	GimA-
FN90P97	340bp	n.a.	-	-	-	-	-	GimA-
DD	740bp	n.a.	+	-	-	-	-	GimA-remnant
RM73A	740bp	n.a.	+	-	-	-	-	GimA-remnant
RM33B	340bp	n.a.	-	-	-	-	-	GimA-
RM64A	740bp	n.a.	+	-	-	-	-	GimA-remnant
FN4	740bp	n.a.	+	-	-	-	-	GimA-remnant
P106	340bp	n.a.	-	-	-	-	-	GimA-
RM71B	740bp	n.a.	+	-	-	-	-	GimA-remnant
RM185S	340bp	n.a.	-	-	-	-	-	GimA-
SIL8	740bp	n.a.	+	-	-	-	-	GimA-remnant
C89	740bp	n.a.	+	-	-	-	-	GimA-remnant
FN83	340bp	n.a.	-	-	-	-	-	GimA-
P69	340bp	n.a.	-	-	-	-	-	GimA-
FN21	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+
C70	340bp	n.a.	-	-	-	-	-	GimA-
RM202I	740bp	n.a.	+	-	-	-	-	GimA-remnant
RM209I	k.A.	10,5kb/10,5kb	+	+	+	+	+	GimA+

Bezeichnung	PCR		Hybridisierung					GimA-Lokus-Variante
	Screening	Long Range	<i>pptE</i>	<i>gclA</i>	<i>pdaK/cgrD</i>	<i>cniT/gclK</i>	<i>ibeA</i>	
RM217T	340bp	n.a.	-	-	-	-	-	GimA-
RM224H	340bp	n.a.	-	-	-	-	-	GimA-
RM45EM	340bp	n.a.	-	-	-	-	-	GimA-
RM70B	340bp	n.a.	-	-	-	-	-	GimA-
ABV84	340bp	n.a.	-	-	-	-	-	GimA-
P68	340bp	n.a.	-	-	-	-	-	GimA-

Tab. 29 Verwendete *E. coli*-Stämme der EcoR-Sammlung (Ochman und Selander, 1984) mit ST, ST-Komplex, Variante des GimA-Lokus und Phylogruppe

EcoR Nr.	Bezeichnung	ST	ST-Komplex	GimA-Lokus-Variante	Phylogruppe
ECOR-01	RM74A	10	ST10 -Komplex	GimA-	A
ECOR-02	STM1	49	ST10 -Komplex	GimA-	A
ECOR-03	W1R1(a)	10	ST10 -Komplex	GimA-	A
ECOR-04	RM39A	63	kein Komplex	GimA-	A
ECOR-05	RM60A	10	ST10 -Komplex	GimA-	A
ECOR-06	RM66C	77	ST206 -Komplex	GimA-	AxB1
ECOR-07	RM73C	87	kein Komplex	GimA-	AxB1
ECOR-08	RM77C	10	ST10 -Komplex	GimA-	A
ECOR-09	FN98	10	ST10 -Komplex	GimA-	A
ECOR-10	ANI	43	ST10 -Komplex	GimA-	A
ECOR-11	C97	10	ST10 -Komplex	GimA-	A
ECOR-12	FN59	10	ST10 -Komplex	GimA-	A
ECOR-13	FN10	44	ST10 -Komplex	GimA-	A
ECOR-14	P62	10	ST10 -Komplex	GimA-	A
ECOR-15	FN3	45	kein Komplex	GimA-	AxB1
ECOR-16	RM191F	46	ST46 -Komplex	GimA-	AxB1
ECOR-17	RM200Q	47	kein Komplex	GimA-	AxB1
ECOR-18	RM210F	48	ST10 -Komplex	GimA-	A
ECOR-19	RM210J	48	ST10 -Komplex	GimA-	A
ECOR-20	RM213J	48	ST10 -Komplex	GimA-	A
ECOR-21	RM213K	48	ST10 -Komplex	GimA-	A
ECOR-22	RM215C	50	kein Komplex	GimA-	A
ECOR-23	RM183E	51	kein Komplex	GimA-	ABD
ECOR-24	FN33	52	kein Komplex	GimA-	AxB1
ECOR-25	MS1	10	ST10 -Komplex	GimA-	A
ECOR-26	LL	53	kein Komplex	GimA-	AxB1
ECOR-27	RM24J	53	kein Komplex	GimA-	AxB1
ECOR-28	RM52B	54	kein Komplex	GimA-	AxB1
ECOR-29	RM3A	55	ST155 -Komplex	GimA-	AxB1
ECOR-30	RM10A	56	ST155 -Komplex	GimA-	ABD
ECOR-31	RM12	57	ST350 -Komplex	GimA-	ABD
ECOR-32	RM28	56	ST155 -Komplex	GimA-	ABD
ECOR-33	RM56C	56	ST155 -Komplex	GimA-	ABD
ECOR-34	WIR2(a)	58	ST155 -Komplex	GimA-	AxB1
ECOR-35	RM42B	59	ST59 -Komplex	GimA-	ABD
ECOR-36	RM77B	60	kein Komplex	GimA-	ABD
ECOR-37	RM44B	61	ST11 -Komplex	GimA-	D
ECOR-38	RM75A	62	kein Komplex	GimA-	ABD
ECOR-39	FN104	62	kein Komplex	GimA-	ABD
ECOR-40	P60	62	kein Komplex	GimA-	ABD
ECOR-41	T44	62	kein Komplex	GimA-	ABD
ECOR-42	DAR1	64	kein Komplex	GimA-	ABD
ECOR-43	FN36	65	kein Komplex	GimA-	ABD

## Anhang

<b>EcoR Nr.</b>	<b>Bezeichnung</b>	<b>ST</b>	<b>ST-Komplex</b>	<b>GimA-Lokus-Variante</b>	<b>Phylogruppe</b>
ECOR-44	RM1891	66	kein Komplex	GimA-	D
ECOR-45	RM201C	67	kein Komplex	GimA-	AxB1
ECOR-46	RM202F	68	kein Komplex	GimA-	D
ECOR-47	RM211C	69	ST69 -Komplex	GimA-	D
ECOR-48	C90/28512	70	kein Komplex	GimA-	D
ECOR-49	FN90	71	kein Komplex	GimA-	D
ECOR-50	FN90P97	72	ST405 -Komplex	GimA-	D
ECOR-51	DD	73	ST73 -Komplex	GimA-remnant	B2
ECOR-52	RM73A	73	ST73 -Komplex	GimA-remnant	B2
ECOR-53	RM33B	12	ST12 -Komplex	GimA-	B2
ECOR-54	RM64A	73	ST73 -Komplex	GimA-remnant	B2
ECOR-55	FN4	74	ST73 -Komplex	GimA-remnant	B2
ECOR-56	P106	73	ST73 -Komplex	GimA-	B2
ECOR-57	RM71B	73	ST73 -Komplex	GimA-remnant	B2
ECOR-58	RM185S	75	kein Komplex	GimA-	AxB1
ECOR-59	SIL8	76	kein Komplex	GimA-remnant	B2
ECOR-60	C89	12	ST12 -Komplex	GimA-remnant	B2
ECOR-61	FN83	78	kein Komplex	GimA-	B2
ECOR-62	P69	79	kein Komplex	GimA-	B2
ECOR-63	FN21	80	ST568 -Komplex	GimA+	B2
ECOR-64	C70	81	ST14 -Komplex	GimA-	B2
ECOR-65	RM202I	82	kein Komplex	GimA-remnant	B2
ECOR-66	RM209I	83	kein Komplex	GimA+	B2
ECOR-67	RM217T	84	kein Komplex	GimA-	AxB1
ECOR-68	RM224H	85	kein Komplex	GimA-	B1
ECOR-69	RM45EM	86	ST86 -Komplex	GimA-	ABD
ECOR-70	RM70B	88	ST23 -Komplex	GimA-	B1
ECOR-71	ABV84	88	ST23 -Komplex	GimA-	B1
ECOR-72	P68	89	kein Komplex	GimA-	B1

## 10 PUBLIKATIONSVERZEICHNIS

Timo Homeier, Lothar H. Wieler, Christa Ewers

Island or no island: phylogeny of GimA involved in pathogenesis of newborn meningitis caused by *E. coli*

VAAM Jahrestagung März 2009

Timo Homeier, Lothar H. Wieler, Christa Ewers

Charakterisierung von GimA, einer putativen genetischen Insel bei Neugeborenenmeningitis *E. coli*, in extraintestinal pathogenen *E. coli*  
Fachgruppentagung der DVG Juni 2008

Timo Homeier

Funktionelle und phylogenetische Analyse von *E. coli* (K1)-Isolaten human- und veterinärmedizinischen Ursprungs

2. Doktoranden-Symposium am Fachbereich Veterinärmedizin der FU Berlin Juni 2007

## 11 DANKSAGUNG

Mein herzlicher Dank gilt Herrn Prof. Dr. L. H. Wieler für die Überlassung des Promotionsthemas, die ausgezeichneten Arbeitsmöglichkeiten am Institut für Mikrobiologie und Tierseuchen der FU Berlin sowie seine stets offene Tür. Besonders danken möchte ich Herrn Wieler für die Möglichkeit meinen Sohn Bjarne zeitweise im Institut zu betreuen.

Ebenso herzlich möchte ich Frau Dr. C. Ewers für die fachliche Betreuung der Arbeit, zahlreiche hilfreiche Anregungen und die stetige Diskussionsbereitschaft danken.

Ich bin Herrn T. Semmler sehr dankbar für die Unterstützung bei allen Computer-basierten Auswertungen und Analysen.

Allen Mitarbeitern des Instituts für Mikrobiologie und Tierseuchen danke ich für die freundliche Aufnahme, die Unterstützung bei der praktischen Durchführung, die Einarbeitung in diverse mikro- und molekularbiologische Techniken und besonders Frau A. Bethe und Frau E. M. Antao die gelegentliche Betreuung meines Sohnes Bjarne.

Für die Überlassung von *E. coli*-Stämmen danke ich Herrn Dr. S. Bonacorsi (Service de Microbiologie, Laboratoire associé au Centre Nationale Référence de *Escherichia coli*, Hôpital Robert Debré (APHP), Frankreich), Herrn PD Dr. U. Dobrindt (Institut für Molekulare Infektionsbiologie, Universität Würzburg), Herrn Prof. Dr. K. S. Kim (Division of Pediatric Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore, USA), Frau Prof. Dr. K. Schnetz (Institut für Genetik, Universität Köln) und Herrn Prof. Dr. U. Vogel (Institut für Hygiene und Mikrobiologie, Universität Würzburg).

Herrn Prof. Dr. G. Amtsberg (Institut für Mikrobiologie, Stiftung Tierärztliche Hochschule Hannover) danke ich dafür, dass er mich schon frühzeitig für das Fach Mikrobiologie begeistert hat.

Besonders danken möchte ich Herrn Prof. Dr. U. Truyen (Institut für Tierhygiene und Öffentliches Veterinärwesen, Universität Leipzig), der mir die zeitnahe Abfassung dieser Arbeit ermöglicht hat.

Für die finanzielle und ideelle Förderung während der Doktorarbeit danke ich der Studienstiftung des deutschen Volkes.

Allen Freunden und Kommilitonen danke ich für die seelische Unterstützung in der Zeit des Studiums und der Promotion.

Ich danke meinen Eltern für die finanzielle Unterstützung und Motivation während des Studiums und der Doktorarbeit.

Am Ende möchte ich der Person danken, die am Gelingen der Doktorarbeit (und des Studiums) den größten Anteil hat: meiner Frau. Du hast nicht nur alle Tiefpunkte mitdurchlebt und mich wieder aufgerichtet, sondern warst auch immer wertvolle Ratgeberin. Dafür und für vieles anderes liebe ich Dich. Dafür, dass sie mit einem Lächeln alle Sorgen vergessen machen können, danke ich unseren Söhnen Bjarne und Mika.

## **12 SELBSTÄNDIGKEITSERKLÄRUNG**

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Hilfsmittel und Literaturquellen angefertigt habe.

Berlin, November 2009

Timo Homeier-Bachmann