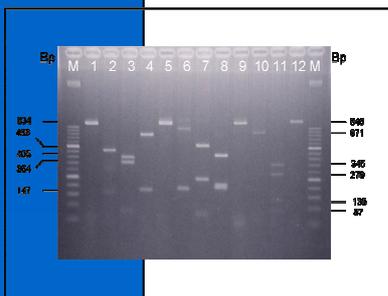
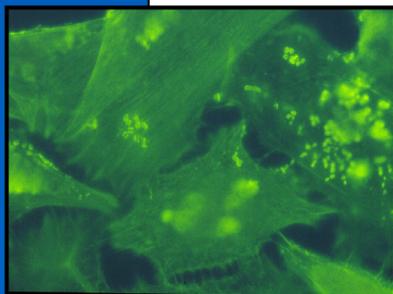
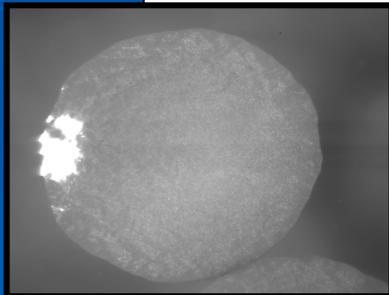


Untersuchungen zur Typisierung des *eae*-Gens enterohämorrhagischer und enteropathogener *Escherichia coli*



INAUGURAL-DISSERTATION
zur Erlangung des Doktorgrades beim
Fachbereich Veterinärmedizin
der Justus-Liebig-Universität Gießen

Eingereicht von

HOLGER SCHÖNENBRÜCHER

Gießen 2006

Aus dem Institut für Tierärztliche Nahrungsmittelkunde
der Justus-Liebig-Universität Gießen

Betreuer: Prof. Dr. M. Bülte

**Untersuchungen zur Typisierung des *eae*-Gens
enterohämorrhagischer und enteropathogener
*Escherichia coli***

INAUGURAL-DISSERTATION
zur Erlangung des Doktorgrades
beim Fachbereich Veterinärmedizin
der Justus-Liebig-Universität Gießen

Eingereicht von

HOLGER SCHÖNENBRÜCHER

Tierarzt aus Siegburg

Gießen 2006

Mit Genehmigung des Fachbereichs Veterinärmedizin
der Justus-Liebig-Universität Gießen

Dekan: Prof. Dr. M. Reinacher

1. Berichterstatter: Prof. Dr. M. Bülte

2. Berichterstatter: Prof. Dr. Dr. habil G. Baljer

Tag der mündlichen Prüfung: 12.07.2006

Teile dieser Arbeit wurden bereits veröffentlicht

Schönenbrücher H., A. Abdulmawjood und M. Bülte (2005)

Typisierung des Intimingens von *E. coli*-Stämmen mit einem PCR-RFLP- Verfahren

46. Arbeitstagung des „Arbeitsgebietes Lebensmittelhygiene“ der Deutsche Veterinärmedizinische Gesellschaft (DVG) vom 27. bis 30.09.2005, Garmisch-Partenkirchen, Proceed. S. 586-592

Schönenbrücher H., A. Abdulmawjood und M. Bülte (2004)

Typisierung des Intimingens bei Attaching and Effacing *Escherichia coli*.

EHEC-Workshop 2004, Wildbad Kreuth, 22.-24.07.2004

Schönenbrücher H., A. Abdulmawjood und M. Bülte (2003)

Typisierung des Intimingens bei Attaching and Effacing *Escherichia coli* und Bereitstellung neuer Gensequenzen.

44. Arbeitstagung des „Arbeitsgebiets Lebensmittelhygiene“, Verlag der Deutsche Veterinär- medizinische Gesellschaft e. V., 35392 Gießen, Proceed. S. 532-537

Schönenbrücher H., A. Abdulmawjood und M. Bülte (2003)

Typisierung des Intimingens bei Attaching and Effacing *Escherichia coli*:

Bereitstellung neuer Gensequenzen.

5. Fachsymposium Lebensmittelmikrobiologie, Vereinigung für allgemeine und angewandte Mikrobiologie (VAAM), Fachgruppe Lebensmittelmikrobiologie in Zusammenarbeit mit der Fachgruppe der DGHM, Kloster Seeon vom 08.-10.05.2003

Schönenbrücher H., A. Abdulmawjood und M. Bülte (2002):

Typisierung des Intimingens enterohämorrhagischer und enteropathogener *Escherichia coli*.

43. Arbeitstagung des „Arbeitsgebiets Lebensmittelhygiene“, Verlag der Deutsche Veterinärmedizinische Gesellschaft e. V., 35392 Gießen, Proceed. S. 568-571

MEINEN ELTERN UND MEINER FRAU

INHALTSVERZEICHNIS

1	Ausgangssituation und Fragestellung	1
2	Literaturübersicht	2
2.1	<i>Escherichia coli</i> (<i>E. coli</i>)	2
2.1.1	Taxonomie und Bedeutung	2
2.2	Darmpathogene <i>E. coli</i>	4
2.2.1	Enteropathogene <i>E. coli</i> (EPEC)	7
2.2.2	Enterotoxinogene <i>E. coli</i> (ETEC)	8
2.2.3	Enteroinvasive <i>E. coli</i> (EIEC)	9
2.2.4	Verotoxin-bildende <i>E. coli</i> (VTEC)/enterohämorrhagische <i>E. coli</i> (EHEC)	10
2.2.5	Enteroaggregative <i>E. coli</i> (EAaggEC)	17
2.2.6	Diffus-adhärente <i>E. coli</i> (DAEC)	18
2.2.7	Nekrotoxische <i>E. coli</i> (NTEC)	18
2.2.8	„Cytotolethal distending toxin“-produzierende <i>E. coli</i> (CLDTEC)	19
2.2.9	Zell-ablösende <i>E. coli</i> , Diarrhoe-assoziierte hämolytische <i>E. coli</i> (DHEC)	20
2.3	Beschreibung des <i>eae</i> -Gens	20
2.3.1	Erstbeschreibung und Charakterisierung	20
2.3.2	Sequenzvergleich, Nomenklatur und diagnostischer Nachweis	22
2.3.3	Intimin-vermittelte Anheftung	24
2.3.4	Bindung an die Wirtszelle ohne Intimin	26
2.4	Ableitung molekularer Phylogenie	26
2.4.1	Terminologie	26
2.4.2	Genetische Distanzen	28
2.4.3	Modelle der Sequenzevolution	29
2.5	Sequenzvergleich	32
2.6	Ableitung populationsgenetischer Parameter	34
2.6.1	Charakterisierung des Sequenzvergleichs	34

2.6.2	Ermittlung des Selektionsmechanismus	35
2.6.3	Genetische Rekombination	35
2.7	Phylogenetische Bäume	40
2.7.1	Distanz- und Charakter-basierte Verfahren	40
2.7.2	Split-Zerlegung	43
2.8	Bioinformatik	46
3	Eigene Untersuchungen	47
3.1	Gliederung der Arbeiten	47
3.1.1	Vorversuche	47
3.1.2	Hauptversuche	48
3.2	Material und Methoden	51
3.2.1	Eingesetzte Referenzstämme	51
3.2.2	Untersuchte Prüf-/Wildstämme	53
3.2.3	Material und Laborgeräte	54
3.2.4	Chemikalien und Reagenzien	56
3.2.5	Herstellung von Puffern und Gebrauchslösungen	58
3.2.6	Revitalisierung und Anzuchtung der Bakterienkulturen	59
3.3	Polymerasekettenreaktion (PCR)	60
3.3.1	Extraktion genomischer DNA aus Bakterienkulturen	60
3.3.2	Extraktion genomischer DNA aus Flüssigmedien	60
3.3.3	Eigenschaften der eingesetzten Polymerase-Typen	62
3.3.4	Darstellung der PCR-Amplifikate mittels Agarose-Gelelektrophorese	64
3.3.5	Ethidiumbromidfärbung und Fotodokumentation	64
3.3.6	Primer-Systeme	65
3.3.7	Primer-System zum Nachweis des <i>eae</i> -Gens	65
3.3.8	Primer-Systeme zur Intimintypisierung	68
3.3.9	Primer-Systeme zur Sequenzierung des <i>eae</i> -Gens	75
3.3.10	Thermocycler-Profile	75
3.4	Restriktionsfragmentlängenpolymorphismus (RFLP)	77
3.4.1	Auswahl und Eigenschaften der Restriktionsendonukleasen	77

3.4.2	Reaktionsansatz und gelelektrophoretische Auswertung	82
3.5	Sequenzanalyse	84
3.5.1	Verwendete Computerprogramme	85
3.5.2	Sequenzierung	91
3.5.3	Verwendete Referenzsequenzen	91
3.5.4	Sequenzvergleich der Intiminvarianten	92
3.5.5	Ableitung molekulargenetischer Phylogenie	93
3.5.6	Ableitung populationsgenetischer Parameter	97
3.6	Zelladhäsion	102
4	Ergebnisse	104
4.1	Nachweis und Typisierung des <i>eae</i> -Gens	104
4.1.1	Amplifikation des konservierten 5'-Bereichs	104
4.1.2	Sequenzanalyse zur Typisierung anhand des hypervariablen 3'-Bereichs	107
4.2	Etablierung und Optimierung der PCR-Typisierungsverfahren	116
4.2.1	Exemplarische Darstellung der PCR-Ergebnisse	116
4.3	Etablierung einer universellen RFLP-Typisierungsmethode	123
4.3.1	Analyse der verwendeten Oligonukleotidprimerkombinationen	123
4.3.2	Exemplarische Darstellung der PCR-Ergebnisse	126
4.4	Ergebnisse der Intimintypisierung einbezogener <i>E. coli</i> -Stämme	133
4.4.1	Klassifizierung <i>eae</i> -positiver <i>E. coli</i> -Stämme	133
4.4.2	Serovaren der <i>eae</i> -positiven <i>E. coli</i> -Stämme	133
4.4.3	Typisierung der Intimingene	137
4.5	Sequenzierung des <i>eae</i> -Gens	137
4.5.1	Analyse der Rohdaten	137
4.6	Ableitung molekulargenetischer Phylogenie	141
4.6.1	Dendrogramme	141
4.6.2	Netzwerkdarstellung	148
4.7	Ableitung populationsgenetischer Parameter	154
4.7.1	Sequenzvergleich der Intiminuntereinheiten	154

4.7.2	Charakterisierung des Sequenzvergleichs der Referenzsequenzen	159
4.7.3	Darstellung des Selektionsmechanismus	161
4.7.4	Genetische Rekombination	168
4.8	Darstellung der zellulären Interaktion	172
4.8.1	Visualisierung der unspezifischen, Fimbrien-vermittelten Adhäsion	172
4.8.2	Visualisierung der spezifischen, Intimin-vermittelten Adhäsion	173
5	Diskussion	175
5.1	Typisierung der <i>eae</i> -Gene	175
5.2	Sequenzierung des <i>eae</i> -Gens	180
5.3	Phylogenetische Analysen	180
5.3.1	Charakterisierung des Alignments	181
5.3.2	Strukturelle Aspekte	183
5.3.3	Darstellung der Baumdiagramme	184
5.3.4	Bewertung der Nomenklatur	186
5.3.5	Netzwerkdarstellung	188
5.3.6	Selektion und genetische Rekombination	190
5.4	Bewertung der verwendeten Computerprogramme	195
5.5	Bedeutung der Variationen innerhalb der Intimingene	197
6	Schlussfolgerungen	202
7	Zusammenfassung	203
8	Summary	207
9	Anhang	208
10	Literaturverzeichnis	226

VERZEICHNIS RELEVANTER BZW. HÄUFIG VERWENDETER ABKÜRZUNGEN UND EINHEITEN

↓ ↑	Schnittstelle einer Restriktionsendonuklease
%	Prozent
=	gleich
+	positiv, plus
-	negativ, minus
±	plus minus
>	größer als
≥	größer oder gleich
<	kleiner als
≤	kleiner oder gleich
®	geschütztes Warenzeichen
™	geschütztes Warenzeichen (engl.: trademark)
§	Paragraph
°C	Grad Celsius
μl	Mikroliter
μM	mikromolar
A	Adenin
α	griech.: Alpha (Minuskel)
A/E	engl.: attaching and effacing
AEEC	engl.: attaching and effacing <i>Escherichia coli</i>
ALTS	Arbeitskreis der auf dem Gebiet der Lebensmittelhygiene und der vom Tier stammenden Lebensmittel tätigen Sachverständigen
Aqua bidest.	Aqua bidestillata
Aqua dest.	Aqua destillata
ATCC	American Type Culture Collection, Manassas, USA
ATP	Adenosintriphosphat

aw	engl.: activity of water (Wasseraktivität)
β	griech.: Beta (Minuskel)
BHI	Hirn-Herz-Bouillon (engl.: brain-heart-infusion broth)
Bp	Basenpaar(e)
C	Cytosin
ca.	circa
CaCl ₂	Calciumchlorid
CDC	Centers for Disease Control and Prevention, Atlanta, USA
CDT	engl.: CDLT cytolethal distending toxin
CHO	engl.: Chinese hamster ovary [Zellkulturlinie]
CLDTEC	engl.: cytolethal distending toxin-producing <i>Escherichia coli</i>
cm	Zentimeter
CTP	Cytidintriphosphat
d. h.	das heißt
DAEC	diffus-adhärente <i>Escherichia coli</i>
δ	griech.: Delta (Minuskel)
DIN	Deutsches Institut für Normung, Berlin
DNA	Desoxyribonukleinsäure (engl.: deoxyribonucleic acid)
<i>E.</i>	<i>Escherichia</i>
dNTP	Desoxynukleotidtriphosphat
eae-Gen	<i>E. coli</i> attaching- and effacing-Gen
EAggEC	engl.: enteroaggregative <i>Escherichia coli</i>
EAST1	EAggEC heat stable enterotoxin
ECD-Agar	<i>Escherichia coli</i> -Direkt-Agar
EDTA	Ethylendiamintetraacetat
EHEC	enterohämorrhagische <i>Escherichia coli</i>
Ehly [<i>hly</i> _{EHEC}]	Enterohämolysin [kursiv: Genbezeichnung]
EIA	Enzymimmunoassay
EIEC	enteroinvasive <i>Escherichia coli</i>
engl.	englisch
EPEC	enteropathogene <i>Escherichia coli</i>

ε	griech.: Epsilon (Minuskel)
EspP [<i>espP</i>]	extrazelluläre Serin-Protease [kursiv: Genbezeichnung]
et al.	und andere (lat.: et alii)
Eta	Gesamtzahl der Mutationen innerhalb eines Alignments
ETEC	enterotoxische <i>Escherichia coli</i>
EU	Europäische Union
Fa.	Firma
g	Gramm
γ	griech.: Gamma (Minuskel)
G	Guanin
G+C-Gehalt	Guanin- und Cytosingehalt
Gb ₃	Globotriosylceramid
Gb ₄	Globotetraosylceramid
ggf.	gegebenenfalls
GmbH	Gesellschaft mit beschränkter Haftung
GTP	Guanosintriphosphat
h	Stunde(n)
H-Antigen	Geißel-Antigen (H: von Hauch)
HC	hämorrhagische Colitis
HCl	Salzsäure
HeLa-Zellen	Zellkulturlinie aus Zervixkarzinomzellen der Patientin Henrietta Lacks
HEp2-Zellen	Mit HeLa-Zellen kontaminierte Zelllinie, die ursprünglich auf eine aus einem humanem Larynxkarzinom gewonnene Linie zurückgeführt wurde
HPI	engl.: high-pathogenicity island
Hrsg.	Herausgeber
http	engl.:hypertext transfer protocol
HUS	hämolytisch-urämisches Syndrom
ι	griech.: Iota (Minuskel)
i. A.	im Allgemeinen
i. d. R.	in der Regel

IfSG	Infektionsschutzgesetz
κ	griech.: Kappa (Minuskel)
<i>K.</i>	<i>Klebsiella</i>
Ka	Verhältnis der Anzahl nonsynonymer Substitutionen zu den nonsynonymen Alignmentpositionen (entspricht p_N)
KatP (<i>katP</i>)	Plasmid-kodierte Katalase-Peroxidase [kursiv: Genbezeichnung]
kb	Kilobase(n)
KbE	Kolonie-bildende Einheiten
kDa	Kilo Dalton
k_{max}	Lokalisation innerhalb eines Alignments mit der höchsten Wahrscheinlichkeit der Rekombination
Ks	Verhältnis der Anzahl synonyme Substitutionen zu den synonymen Alignmentpositionen (entspricht p_s)
λ	griech.: Lambda (Minuskel)
l	Liter
LEE	engl.: locus of enterocyte effacement
LFGB	Lebensmittel- und Futtermittelgesetzbuch
μ	griech.: My (Minuskel)
M	molar
max.	maximal
Mb	Megabase(n)
MDa	Mega Dalton
MgCl ₂	Magnesiumchlorid
min	Minute(n)
mind.	mindestens
ml	Milliliter
mm	Millimeter
mM	millimolar
MUG	4-Methylumbelliferyl- β -D-Glucuronid
v	griech.: Ny (Minuskel)
n	Anzahl

NaCl	Natriumchlorid
NaClO ₄	Natriumperchlorat
NCTC	National Collection of Type Cultures, Central Public Health Laboratories, Colindale, London, Vereinigtes Königreich
nm	Nanometer
Nr.	Nummer
NTEC	nekrotoxische <i>Escherichia coli</i>
NTP	Desoxynukleotidtriphosphat
N-WASP	engl.: neural Wiskott–Aldrich-syndrome protein
o	griech.: Omikron (Minuskel)
o. e.	oben erwähnt
o. g.	oben genannt
O-Antigen	somatisches Antigen
ORF	offene Leserahmen (engl.: open reading frame)
π	griech.: Pi (Minuskel)
PC-Agar	Plate Count-Agar
PCR	Polymerase-Kettenreaktion (engl.: polymerase chain reaction)
PFGE	Pulsfeldgelelektrophorese
<i>pheU</i> , <i>pheV</i>	Phenylalanin-tRNA-Gen
p_N	Verhältnis der Anzahl nonsynonymer Substitutionen zu den nonsynonymen Alignmentpositionen (entspricht K_a)
p_s	Verhältnis der Anzahl synonyme Substitutionen zu den synonymen Alignmentpositionen (entspricht K_s)
ρ	griech.: Rho (Minuskel)
<i>rfbE</i>	O-Antigen-Gen
RFLP	Restriktions-Fragment-Längen-Polymorphismus
RKI	Robert Koch-Institut, Berlin
rpm	rotations per minute
rRNA	ribosomale Ribonukleinsäure (engl.: ribonucleic acid)
<i>Rsa</i> I	Restriktionsenzym aus <i>Rhodopseudomonas sphaeroides</i>
σ	griech.: Sigma (Minuskel)

s	Sekunde(n)
s.	siehe
S	Anzahl variierender Alignmentpositionen
Saa [saa]	engl.: STEC autoagglutination adhesin [kursiv: Genbezeichnung]
SCVPH	Scientific Committee on Veterinary Measures relating to Public Health der Europäischen Kommission
se/C	Selenocystein-tRNA-Gen
SF	Sorbit-fermentierend
SLT	Shiga-like-Toxin
sp./spp.	Spezies (Singular/Plural)
SPM (X)	Anzahl der Alignmentpositionen mit mindestens X identischen Nukleotiden, die zu keiner Änderung der Aminosäurekodierung führen (engl. Singelton Polymorphic Sites)
ssp.	Subspezies
sog.	so genannte
STEC	Shiga-Toxin-bildende <i>Escherichia coli</i>
Stx [stx]	Shiga-Toxin [kursiv: Genbezeichnung]
ε	griech.: Eta (Minuskel)
θ	griech.: Theta (Minuskel)
T	Thymidin
TBE-Puffer	Tris-(hydroxymethyl)aminomethan-Borat-EDTA-Puffer
Tir [tir]	engl.: translocated intimin receptor [kursiv: Genbezeichnung]
tRNA	transferierende Ribonukleinsäure
TSB	Trypton-Soja-Bouillon
TTP	Thymidintriphosphat; thrombotisch-thrombozytopenische Purpura
U	Unit(s)
u. a.	unter anderem

UPGMA	engl.: unweighted-pair-group matching by arithmetic averages-algorithm
UV	Ultra-violett
V	Volt
v. a.	vor allem
VT [vtx]	Verotoxin [kursiv: Genbezeichnung]
VTEC	Verotoxin-bildende <i>Escherichia coli</i> /verotoxinogene <i>Escherichia coli</i>
WHO	World Health Organization, Genf, Schweiz
www	World Wide Web
ξ	griech.: Xi (Minuskel)
YOPIS	engl.: young, old, pregnant, immunocompromised segments of the public
ζ	griech.: Zeta (Minuskel)
z. B.	zum Beispiel
z. Zt.	zur Zeit

TABELLENVERZEICHNIS

Tabelle 2.1.1-1:	Nomenklatur und Einteilung pathogener <i>E. coli</i>	5
Tabelle 2.2.4-1:	Prominente VT 1-Varianten (vereinheitlichte Nomenklatur)	13
Tabelle 2.2.4-2:	Prominente VT 2-Varianten (vereinheitlichte Nomenklatur)	14
Tabelle 2.6.3-1:	Charakteristika verschiedener Programme zur Überprüfung der genetischen Rekombination nach MARTIN und RYBICKI (2000)	39
Tabelle 2.7.1-1:	Einteilung der Rekonstruktionsmethoden zur phylogenetischen Analyse	40
Tabelle 3.2.1-1:	Etablierung der PCR-Systeme nach ADU-BOBIE et al. (1998) und OSWALD et al. (2000) unter Berücksichtigung verwendete Positiv- und Negativkontrollstämmen (...)	52
Tabelle 3.2.2-1:	Zusammensetzung der untersuchten Prüfstämme (n=317 [%])	53
Tabelle 3.3.3-1:	Eigenschaften der verwendeten Polymerasetypen	63
Tabelle 3.3.7-1:	Charakteristika der Primer zur Detektion des <i>eae</i> -Gens nach KARCH et al. (1993)	66
Tabelle 3.3.8-1:	Charakteristika der Primer zur Typisierung des <i>eae</i> -Gens nach ADU-BOBIE et al. (1998a) und der in dieser Arbeit entwickelten Oligonukleotide zur Erfassung von epsilon-Intimin	69
Tabelle 3.3.8-2:	Charakteristika der Primer zur Typisierung des <i>eae</i> -Gens nach OSWALD et al. (2000)	71
Tabelle 3.3.8-3:	Charakteristika der Primer zur Typisierung des <i>eae</i> -Gens nach RAMACHANDRAN et al. (2003)	73
Tabelle 3.3.10-1:	Thermocycler-Profile zur Intimintypisierung und Sequenzierung	76
Tabelle 3.4.1-1:	Nomenklatur, Schnittstellen und Klassifikation der eingesetzten Restriktionsendonukleasen für die PCR-Produktüberprüfung der modifizierten PCR nach ADU-BOBIE et al. (1998a) und der PCR nach RAMACHANDRAN et al. (2003)	78
Tabelle 3.4.1-2:	Produktüberprüfung (n=5) der PCR nach ABU-BOBIE et al. (1998a): Eingesetzte Restriktionsenzyme, Anzahl der Schnittstellen und resultierende Fragmentgrößen	79

Tabelle 3.4.1-3:	Produktüberprüfung (n=19) der PCR nach RAMACHAN-DRAN et al. (2003): Generierte Amplifikatgrößen und erwartete Größe der Restriktionsfragmente in Abhängigkeit vom ausgewählten Restriktionsenzym	80
Tabelle 3.4.1-4:	Differenzierung der Intiminuntereinheiten (n=19) anhand der erwarteten Fragmentgrößen (+/- mit dem Restriktionsenzym differenzierbar bzw. nicht differenzierbar)	81
Tabelle 3.4.2-1:	Reaktionsansatz für den restriktionsenzymatischen Verdau	83
Tabelle 3.5.1-1:	Computerprogramme und Datenbanken unter Berücksichtigung des Verwendungszweckes	86
Tabelle 3.5.5-1:	Parameter zur Berechnung der p-Distanz zur Erstellung des phylogenetischen Baumes unter Verwendung der modifizierten Methodik nach NEI und GOJOBORI (1986, [p-Distanz])	94
Tabelle 3.5.5-2:	Parameter zur Erstellung des phylogenetischen Baumes nach TAMURA und NEI (1993)	95
Tabelle 3.5.6-1:	Abgeleitete populationsgenetische Parameter	101
Tabelle 4.4.2-1:	Zuordnung unterschiedlicher Serovare (n=56) zu den Intimintypen	134
Tabelle 4.4.2-2:	Anteil der Intiminvarianten (n[%]) an den 317 untersuchten Prüfstämmen	135
Tabelle 4.4.2-3:	Verteilung (n) der Intiminuntereinheiten, gruppiert nach Habitaten	136
Tabelle 4.5.1-1:	Tabellarische Übersicht der Charakteristika der für die Sequenzanalyse verwendeten zeta (ζ)-Intimingene	139
Tabelle 4.5.1-2:	Tabellarische Übersicht der prozentualen Sequenzdistanzen (Übereinstimmung, Unterschiede) der zeta (ζ)-Intiminvarianten (n=7) erstellt mit MegAlign™ (Fa. DNastar)	141
Tabelle 4.7.2-1:	Verteilung der polymorphen Alignmentpositionen unter Berücksichtigung des Genabschnittes	160
Tabelle 4.7.3-1:	Variation der Verhältnisse von p _S und pN [s] in den Genabschnitten und funktionellen Domänen (Einteilung nach LUO et al., 2000)	164
Tabelle 4.7.4-1:	Mögliche rekombinante Regionen innerhalb des zeta (ζ)-Intimingens mit p>0,05	170

Tabelle 9.1.1-1:	Für die phylogenetische Analyse und die Berechnung populationsgenetischer Parameter eingesetzte Referenzstämme	209
Tabelle 9.1.1-2:	Sequenz-Übereinstimmungs-Matrix, erstellt anhand des CLUSTALW 1.6-Sequenzvergleichs	212
Tabelle 9.1.1-3:	Berechnung des wahrscheinlichsten Baumes mit DNA-ML V.3.6a3 (enthalten in PAML)	216
Tabelle 9.1.2-1:	Daten der Neighbor-Net-Darstellung nach BRYANT und MOULTON (2004) für den gesamten Intimingenbereich	219
Tabelle 9.1.3-1:	3'-Bereich: Berechnung von „ p_{NR} “ für den Parameter „Polarität“	220
Tabelle 9.1.3-2:	3'-Bereich: Berechnung von „ p_{NC} “ für den Parameter „Polarität“	221
Tabelle 9.1.3-3:	3'-Bereich: Berechnung von „ p_{NR} “ für den Parameter „Ladung“	222
Tabelle 9.1.3-4:	3'-Bereich: Berechnung von „ p_{NC} “ für den Parameter „Ladung“	223
Tabelle 9.1.3-5:	Serovarenverteilung unter Berücksichtigung der nachgewiesenen Intiminvarianten	224

ABBILDUNGSVERZEICHNIS

Abbildung 2.4.1-1:	Phylogenetische Terminologie am Beispiel eines ungewurzelten Baumes in Anlehnung an BRINKMANN (2005)	27
Abbildung 2.4.2-1:	Grundprinzipien möglicher Nukleotidsubstitutionen	29
Abbildung 2.7.2-1:	Schematische Darstellung der Split-Zerlegung nach HUSON und BRYANT (2006)	45
Abbildung 4.1.1-1:	Beispielhafte Darstellung der Amplifikation der konservierten Region des <i>eae</i> -Gens mit den Oligonukleotid-Primern „SK1/ SK2“ (...)	105
Abbildung 4.1.1-2:	Lokalisation der Oligonukleotid-Primer „SK1/ SK2“ (...)	106
Abbildung 4.1.2-1:	Lokalisation der spezifischen Vorwärts-Primer „Int-A, Int-B, Int-Y und Int-D“ nach ADU-BOBIE et al. (1998a) innerhalb des Sequenzvergleichs	109
Abbildung 4.1.2-2:	Lokalisation des Universalrückwärts-Primers „Int-Ru“ nach ADU-BOBIE et al. (1998a) innerhalb des Sequenzvergleichs	110
Abbildung 4.1.2-3:	Lokalisation des abgeleiteten Vorwärts-Primers „epsilonF“ nach SCHÖNENBRÜCHER et al. (2002) innerhalb des Sequenzvergleichs (...)	111
Abbildung 4.1.2-4:	Lokalisation des abgeleiteten Rückwärts-Primers „epsilonR“ nach SCHÖNENBRÜCHER et al. (2002) innerhalb des Sequenzvergleichs (...)	112
Abbildung 4.1.2-5:	Lokalisation des Rückwärts-Primers „LP4“ zum Nachweis von β -Intimin nach OSWALD et al. (2000) innerhalb des Sequenzvergleichs	113
Abbildung 4.1.2-6:	Lokalisation des Rückwärts-Primers „LP5“ zum Nachweis von ε -Intimin nach OSWALD et al. (2000) innerhalb des Sequenzvergleichs (...)	114
Abbildung 4.1.2-7:	Lokalisation der Rückwärts-Primer „LP2“ und „LP3“ zum Nachweis von α - und γ -Intimin nach OSWALD et al. (2000) und der Primer „EaeVR“ (Universalrückwärts-Primer, alle Untereinheiten), „EaeZetaVR“ (Rückwärts-Primer für ζ -Intimin), „EaelotaVR“ (Rückwärts-Primer für ι -Intimin nach RAMACHANDRAN et al. [2003]) innerhalb des Sequenzvergleichs	115

- Abbildung 4.2.1-1:** Exemplarische Darstellung der Amplifikation von α -Intimin mit der Primer-Kombination „Int-Ru/Int-A“ nach ADU-BOBIE et al. (1998a) 117
- Abbildung 4.2.1-2:** Exemplarische Darstellung der Amplifikation von β -Intimin mit der Primer-Kombination „Int-Ru/Int-B“ nach ADU-BOBIE et al. (1998a) und des restriktionsenzymatischen Verdau mit *Taq I* und *Fok I* 118
- Abbildung 4.2.1-3:** Exemplarische Darstellung der Amplifikation von γ -Intimin mit der Primer-Kombination „Int-Ru/Int-Y“ nach ADU-BOBIE et al. (1998a) 119
- Abbildung 4.2.1-4:** Restriktionsenzymatischer Verdau von γ -Intimin mit *Taq I* und *Fok I* 120
- Abbildung 4.2.1-5:** Exemplarische Darstellung der Amplifikation von ε -Intimin mit der Primer-Kombination „epsilonF/epsilonR“ 121
- Abbildung 4.2.1-6:** Restriktionsenzymatischer Verdau von ε -Intimin generiert mit der Primer-Kombination „epsilonF/epsilonR“ 122
- Abbildung 4.3.1-1:** Lokalisation des Universalvorwärts-Primers „EaeVR“ nach RAMACHANDRAN et al. (2003) innerhalb des Alignments der bislang beschriebenen Intiminvarianten alpha (α) bis sigma (σ) 125
- Abbildung 4.3.2-1:** Exemplarische Darstellung der universellen Amplifikation verschiedener Intimingene mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003) 127
- Abbildung 4.3.2-2:** Exemplarische Darstellung der Amplifikation von ζ -Intimin mit der Primer-Kombination „EaeVF/EaeZetaR“ nach RAMACHANDRAN et al. (2003) 128
- Abbildung 4.3.2-3:** Exemplarische Darstellung des restriktionsenzymatischen Verdau von β -Intimin nach Amplifikation mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003) 129
- Abbildung 4.3.2-4:** Exemplarische Darstellung des restriktionsenzymatischen Verdau von γ -Intimin nach Amplifikation mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003) 130

-
- Abbildung 4.3.2-5:** Exemplarische Darstellung des restriktionsenzymatischen Verdaus von ϵ -Intimin nach Amplifikation mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003) 131
- Abbildung 4.3.2-6:** Exemplarische Darstellung des restriktionsenzymatischen Verdaus von θ , β - und ζ -Intimin nach Amplifikation mit der Primer-Kombination „EaeVF/EaeVR“ und „EaeVF/EaeZetaR“ nach RAMACHANDRAN et al. (2003) 132
- Abbildung 4.5.1-1:** Exemplarische Darstellung der schrittweisen Amplifikation des *eae*-Gens der untersuchten Isolate P 146 und P 256 140
- Abbildung 4.6.1-1:** Linearisierter, ungewurzelter Baum der Intiminvarianten unter Verwendung des modifizierten Modells nach NEI und GOJOBORI (p-Distanz) in der Neighbor-Joining-Darstellung. Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent 143
- Abbildung 4.6.1-2:** Interior Branch Test-Auswertung unter Verwendung der p-Distanz, berechnet anhand des modifizierten Modells nach NEI und GOJOBORI (1986) in der Neighbor-Joining-Darstellung 144
- Abbildung 4.6.1-3:** Linearisierter, ungewurzelter Baum der Intiminvarianten unter Verwendung des Modells nach TAMURA und NEI (1998) in der Neighbor-Joining-Darstellung 145
- Abbildung 4.6.1-4:** Interior Branch Test-Auswertung unter Verwendung des Modells nach TAMURA und NEI (1998) in der Neighbor-Joining-Darstellung 146
- Abbildung 4.6.1-5:** Möglicher unverzweigter „Maximum Likelihood“-Baum der Intiminvarianten 147
- Abbildung 4.6.2-1:** Neighbor-Net-Darstellung des phylogenetischen Netzwerkes unter Berücksichtigung der gesamten kodierenden Intimingensequenz anhand des Modells nach BRYANT und MOULTON (2004) 149
- Abbildung 4.6.2-2:** Neighbor-Net-Darstellung des phylogenetischen Netzwerkes der konservierten 5'-Intimingensequenzen anhand des Modells nach BRYANT und MOULTON (2004) 150

Abbildung 4.6.2-3: Neighbor-Net-Darstellung des phylogenetischen Netzwerkes der hypervariablen 3'-Intimingengebieten anhand des Modells nach BRYANT und MOULTON (2004)	151
Abbildung 4.6.2-4: Darstellung des phylogenetischen Netzwerkes der konservierten 5'-Intimingebiete anhand der Split-Zerlegung nach BANDELT und DRESS (1992) sowie HUSON (1998)	152
Abbildung 4.6.2-5: Darstellung des phylogenetischen Netzwerkes unter Berücksichtigung der gesamten Intimingensequenz anhand der Split-Zerlegung nach BANDELT und DRESS (1992) sowie HUSON (1998)	153
Abbildung 4.7.1-1: Sequenzvergleich der Proteinsequenzen	156
Abbildung 4.7.1-2: Visualisierung polymorpher Nukleotidpositionen innerhalb des Sequenzvergleichs	158
Abbildung 4.7.2-1: Nukleotidfrequenzen [%] innerhalb der kodierenden Sequenzregionen	1
Abbildung 4.7.3-1: Darstellung der Anzahl der Transitionen und Transversionen in Abhängigkeit von der Distanz im 5'-Sequenzbereich	161
Abbildung 4.7.3-2: Darstellung der Anzahl der Transitionen und Transversionen in Abhängigkeit von der Distanz im 3'-Sequenzbereich	161
Abbildung 4.7.3-3: Proportion der synonymen- (K_s -) und nonsynonymen- (K_a -) Substitutionen des zeta (ζ)-Intimings im Vergleich mit den alpha (α)- und nu (ν)-Intimingenen	165
Abbildung 4.7.3-4: Proportion der synonymen- (K_s -) und nonsynonymen- (K_a -) Substitutionen des zeta (ζ)-Intimings im Vergleich zum omikron (\omicron)-Intimin	166
Abbildung 4.7.3-5: Relation der radikalen, nonsynonymen Substitutionen (p_{NR}) gegen die konservativen, nonsynonymen Substitutionen (p_{NC})	167
Abbildung 4.7.4-1: Manuelle Überprüfung wahrscheinlicher rekombinanter Regionen ($p > 0,01$) im paarweisen Sequenzvergleich mit ζ -Intimin	171

- Abbildung 4.8.1-1:** Unspezifische, Fimbrien-vermittelte Adhäsion von EPEC O127:H6 an humanen Mundschleimhautzellen (Fluoreszenzmikroskop, Vergrößerung: 400fach) 172
- Abbildung 4.8.2-1:** Lokalisierte Adhäsion des EPEC E2348/69 O127:H6 an HEp2-Zellen. (links: Vergrößerung: 400fach; rechts: Teilausschnitt, Vergrößerung: 600fach) 173
- Abbildung 4.8.2-2:** Darstellung der Stressfasern und der Aktinakkumulation 174
- Abbildung 4.8.2-3:** Kontrolle: Unbeimpfte HEp2-Zellen 174

1 AUSGANGSSITUATION UND FRAGESTELLUNG

Attaching und Effacing (A/E)-Läsionen sind das pathognomische Kennzeichen (engl. „hallmark“) der Enteropathogenen Escherichia coli (EPEC) und der Verotoxin-bildenden Escherichia coli (VTEC) bzw. der Shiga toxin-bildenden E. coli (STEC). Zunächst nur im Zellkulturmodell darstellbar, wurden die beteiligten bakteriellen Virulenz- sowie Adhäsionsfaktoren mittlerweile auch molekularbiologisch gut charakterisiert. So konnte gezeigt werden, dass das auf dem „Locus of Enterocyte Effacement“ (LEE) verankerte „E. coli attaching and effacing“ (*eae*)-Gen für den Anheftungsvorgang, einer innigen Zell- zu Zellverbindung (engl. „intimate adhearance“), verantwortlich sowie für die Ausprägung der A/E-Läsionen essenziell ist. Das für das Protein Intimin kodierende *eae*-Gen stellt dabei neben den Vero-/Shigatoxingenen und dem Enterohämolsingen einen bedeutenden Virulenzfaktor zur Charakterisierung des Gefährdungspotentials der *E. coli* dar. Die taxonomische Bedeutung bei der Klassifikation der VTEC/STEC bzw. der Enterohämorrhagischen E. coli (EHEC) wird kontrovers diskutiert. Der Intimintypisierung auf Basis des hypervariablen 3'-Genabschnittes mittels der von MULLIS und FALOOKA (1987) allgemein beschriebenen Technik der Polymerasekettenreaktion (engl. „Polymerase Chain Reaction“, [PCR]) kommt dabei eine wesentlich diagnostische und epidemiologische Bedeutung zu. So konnten bislang einschließlich der Untereinheiten mehr als 20 verschiedene Intimintypen differenziert werden (α bis σ).

Vor diesem Hintergrund sollte in der eigenen Arbeit mit der Etablierung eines PCR-Verfahrens unter Verwendung der Restriktionsfragment-Längenpolymorphismus (RFLP) basierten Differenzierung ein Beitrag zur methodischen Weiterentwicklung und Verteilung der Intiminvarianten bei 317 *E. coli*-Stämmen aus unterschiedlichen Habitaten geleistet werden. Über die phylogenetische Analyse und die Charakterisierung der Genstruktur sollte darüber hinaus eine Bewertung der möglichen evolutiven Entwicklung vorgenommen werden.

2 LITERATURÜBERSICHT

Der Schwerpunkt der vorgestellten Arbeit behandelt Untersuchungen zum *Escherichia coli* attaching and effacing (eae)-Gen. Die dargestellte Literaturübersicht folgt diesem Prinzip.

2.1 *ESCHERICHIA COLI* (*E. COLI*)

2.1.1 TAXONOMIE UND BEDEUTUNG

Zur Familie der Enterobacteriaceae gehörend sind *Escherichia coli* (*E. coli*) Gram-negative, fakultativ anaerobe, sporenlose, Oxidase-negative, Katalase-positive, stäbchenförmige Bakterien. Ihre Größe beträgt ca. 1,1 - 1,5 µm x 2,0 - 6,0 µm. Überwiegend tritt Kapselbildung auf. Durch peritriche Begeißelung können bewegliche von unbeweglichen Stämmen differenziert werden. Als mesophile Mikroorganismen liegt ihr Temperaturoptimum bei 37°C. Unter diesen optimalen Bedingungen beträgt die Generationszeit etwa 20 min (HOLT et al., 1994). *E. coli* sind zu aerobem sowie fakultativ anaerobem Wachstum befähigt.

Der Dedikationsname geht auf den Pädiater Dr. Theodor Escherich zurück, der solche Stämme 1885 aus Säuglingsstuhl isolierte und diese als *Bacterium coli commune* beschrieb. Neben *E. coli* als dem wichtigsten Vertreter werden unter dem Genus *Escherichia* im „Bergey's Manual of Determinative Bacteriology“ (9. Auflage, 1984) die Spezies *E. blattae*, *E. fergusonii*, *E. hermanii* und *E. vulneris* biochemisch differenziert (HOLT et al., 1994). HUYS et al. (2003) beschrieben mit *E. albertii* eine weitere Spezies. ABBOTT et al. (2003) und HYMA et al. (2005) gelangen die biochemische und evolutive Charakterisierung dieser neuen Spezies.

E. coli werden entsprechend ihren Oberflächen (O)-, Kapsel (K)- und Geißel (H: von Hauch)- Antigene in verschiedene Serovaren unterteilt (ØRSKOV und ØRSKOV, 1984). Die Entwicklung der Serotypisierung ging 1944 von F. Kauffmann aus (KAUFFMANN, 1966). Damit wurde die Untergliederung in Serogruppen (nur O-Antigen) und/oder Serovaren (O- und H-Antigen) ermöglicht.

Sowohl beim Menschen als auch bei den Säugetieren (Ausnahmen: Meerschweinchen, Chinchilla) dient der Dickdarm als natürliches Habitat. Allerdings liegen die Keimzahlgehalte mit 10^4 bis 10^9 KbE/g im 1%-Bereich. Damit gehört *E. coli* zu der Begleitflora (ROLLE und MAYR, 1993). Als mögliche Ursache für die als sehr erfolgreich charakterisierte Eubiose der *E. coli* wird die Fähigkeit zu einer effizienteren Gluconsäureverwertung angesehen (SWEENEY et al., 1996).

Hinsichtlich ihrer klinischen Bedeutung ist zu berücksichtigen, dass sowohl fakultativ pathogene als auch obligat pathogene Stämme unterschieden werden müssen. Letztgenannte verfügen über ein breites, mobilisierbares Spektrum an Virulenzfaktoren (HAHN et al., 1999 und KAPER et al., 2004). Diese *E. coli*-Pathotypen können Harnwegsinfektionen, Sepsis, Hirnhautentzündungen, enterale sowie Durchfallerkrankungen und Appendizitiden, Peritonitiden und postoperative Wundinfektionen verursachen. Zusätzlich zu den uropathogenen *E. coli* (UPEC), den nephropathogenen (NPEC) *E. coli* und septisch-pathogenen (SPEC) *E. coli* (NATARO und KAPER, 1998; KAPER et al., 2004) sind sechs gut charakterisierte intestinale Pathovaren beschrieben worden, die nachfolgend erläutert werden. Damit zählen *E. coli*-Pathotypen mit zu den häufigsten Erregern bakterieller Infektionen des Menschen.

In der Veterinärmedizin spielen *E. coli* als Infektionserreger bei Septikämien, enterotoxischen Enteropathien, Dysenterien sowie der Ödemkrankheit der Absatzferkel eine Rolle.

Bei Adulten ist auf Mastitiden, Harnwegsinfektionen, Pneumonien, Pyometren, Wundinfektionen sowie die Coligranulomatose des Geflügels zu verweisen (ROLLE und MAYR, 1993). An den „avian pathogenic *E. coli* (APEC)“ erkranken primär Puten (EWERS et al., 2003).

Für die Lebensmittelhygiene ist der Darmkommensale *E. coli* als Indikatororganismus für fäkale Verunreinigungen innerhalb der Prozessführung eingestuft worden (REUTER, 1996). Eine detaillierte Auflistung des Umfangs der Untersuchungen auf *E. coli* in Abhängigkeit von der untersuchten Lebensmittelgruppe im Sinne von Prozesshygiene- und Lebensmittelsicherheitskriterien beinhaltet die VERORDNUNG (EG) Nr. 2073/2005 der Kommission vom 15. November 2005 über mikrobiologische Kriterien für Lebensmittel. Aufgrund ihres potenziellen gesundheitlichen Gefahrenpotenzials werden *E. coli* darüber hinaus auch den Indexorganismen zugerechnet (REUTER, 1996).

2.2 DARMPATHOGENE *E. COLI*

Zusätzlich zu den *E. coli*-Pathotypen, die extraintestinale Erkrankungen verursachen und daher auch als ExpEC (RUSSO und JOHNSON, 2000) bezeichnet werden, sind die nach dem Infektionsschutzgesetz (IfSG; ANONYMOUS, 2000) meldepflichtigen darmpathogenen *E. coli* von Bedeutung. Unter ihnen können Stämme der EPEC-, VTEC-/EHEC- und ETEC-Gruppe auch beim Tier klinische Erkrankungen hervorrufen (NATARO und KAPER, 1998; RKI, 1996). Eine Übersicht enthält **Tabelle 2.1.1-1** in Anlehnung an die Zusammenstellung nach GOLL (2005).

Table 2.1.1-1: Nomenklatur und Einteilung pathogener *E. coli*

Akronym	Bezeichnung	Unterteilung	Virulenzprofil	Erkrankung	
				Mensch	Tier
EPEC	Enteropathogene <i>E. coli</i>	class I class II	Haftungsfähigkeit (LEE, EAF-Plasmid ²⁾ LEE	Säuglingsdiarrhoe	Jungtierdiarrhoe
ETEC	Enterotoxische <i>E. coli</i>	LT-Stämme ³⁾ ST-Stämme ⁴⁾	Enterotoxine, Kolonisationsfaktoren	Reisediarrhoe (Cholera-ähnlich)	Enterotoxische Enteropathie
EIEC	Enteroinvasive <i>E. coli</i>		Invasivität, Enterotoxine	Reisediarrhoe (Ruhr-ähnlich)	
VTEC/STEC	Verotoxinogene <i>E. coli</i> /Shiga-Toxin- bildende <i>E. coli</i>		Verotoxine/ Shiga- Toxine		Kälberdiarrhoe; Ödemkrankheit der Absatzferkel (nur VT 2e- Stämme ⁵⁾)
EHEC	Enterohämorrhagische <i>E. coli</i>		Verotoxine/ Shiga- Toxine, Haftungsfähigkeit (LEE), Enterohämolysin	Diarrhoe, HC ⁶⁾ , HUS ⁷⁾ , TTP ⁸⁾	

1) locus of enterocyte effacement

2) EPEC adherence factor

3) hitzelabiles Enterotoxin

4) hitzestabiles Enterotoxin

6) hämorrhagische Colitis

7) hämolytisch-urämisches Syndrom

8) thrombotisch-thrombozytopenische Purpura

Fortsetzung Tabelle 2.1.1.1-1: Nomenklatur und Einteilung pathogener *E. coli*

Akronym	Bezeichnung	Unterteilung	Virulenzprofil	Erkrankung	
				Mensch	Tier
EAggEC/ EAEC	Enteroaggregative <i>E. coli</i>		Enterotoxin (EAST1 ¹⁾) Haftungsfähigkeit	Persistierende Diarrhoe, Reisediarrhoe	Jungtierdiarrhoe
DAEC	Diffus-adhärente <i>E. coli</i>		Adhäsion	Infantile Diarrhoe	Enterotoxische Enteropathie
NTEC	Nekrototoxische <i>E. coli</i>		CNF 1 ²⁾ CNF 2	Diarrhoe, extraintestinale Erkrankungen	Kälberdiarrhoe, -septikämie
CLDTEC	“Cytotolethal distending” toxin”- produzierende <i>E. coli</i>		Cytotolethal distending toxin (CDT I-V ³⁾)	Infantile Diarrhoe	
CDHEC/ DHEC	Diarrhoe- assoziierte hämolytische <i>E. coli</i>		α-Hämolysin	Infantile Diarrhoe	

¹⁾ EAggEC heat stable enterotoxin

²⁾ cytotoxic necrotizing factor

³⁾ cytolethal distending toxin

2.2.1 ENTEROPATHOGENE *E. COLI* (EPEC)

BRAY beschrieb 1945 eine Gruppe serologisch unterscheidbarer *E. coli*, die bei durchfallerkrankten Kindern, nicht jedoch bei gesunden Kindern auftraten. Den EPEC wird der histopathologische Effekt des „attaching and effacing (A/E)“ zugeordnet (CANTEY und HOSTERMANN, 1979). Dieser ergibt sich aus der Anheftung (engl. „attaching“) des EPEC-Stammes und durch den lokalen Verlust der Mikrovilli (engl. „effacing“) sowie einer massiven Anreicherung von Zytoskelett-Proteinen in Form von polymerisiertem Aktin unterhalb der bakteriellen Anheftungsstelle (MOON et al., 1983; KNUTTON et al., 1989). Per Definition fehlt EPEC das Verotoxinbildungsvermögen. Anhand dieser Unterscheidung erfolgt die Abgrenzung zu den VTEC/STEC, die ebenfalls zur Bildung von A/E-Läsionen fähig sind. Eine genaue Beschreibung des dreistufigen Anheftungsprozesses wird im Abschnitt „Virulenzfaktoren“ unter der Beschreibung der Anheftungsmechanismen bei EPEC und VTEC/EHEC vorgenommen. Die „class I“-EPEC verfügen über das EAF-„EPEC adherence factor“-Plasmid, das das Gen der „bundle forming pili“ (BFP) beherbergt (NATARO und KAPER, 1998). Atypischen „class II“-EPEC fehlt das EAF-Plasmid. Die weiteren Virulenzfaktoren sind auf dem „Locus of Enterocyte Effacement“ (LEE) verankert (McDANIEL et al., 1995). Hierzu zählen neben Intimin der translozierte Intiminrezeptor Tir (engl. „translocated intimin receptor“) und weitere Komponenten des Typ III-Sekretionsapparates zuzüglich weiterer Proteine.

Als Erreger der Säuglingsdiarrhoe kommt EPEC nur noch in den Entwicklungsländern eine wesentliche Bedeutung zu. Neben dem akut verlaufenden Erscheinungsbild mit meist wässrigem Durchfall, Erbrechen und geringgradigem Fieber, treten auch protrahierte Formen auf (NATARO und KAPER, 1998). Den anhand der im Rahmen des IfSG gemeldeten Nachweisen ist zu entnehmen, dass 72,5% der durch *E. coli* verursachten intestinalen Erkrankungen auf EPEC zurückzuführen sind. Zu den EPEC gehören Stämme der Serogruppen O26, O44, O55, O86, O111, O114, O119, O125 und O147

(NATARO et al., 1985). In Deutschland dominieren die Serovaren O26:H11 und O86:H34 (RKI, 1996).

EPEC wurden bislang auch bei Durchfallerkrankungen von Kaninchen, Hühnern, Hunden, Katzen sowie Kälbern, Ferkeln und Lämmern isoliert.

2.2.2 ENTEROTOXINOGENE *E. COLI* (ETEC)

Die enterotoxischen *E. coli* werden durch die Eigenschaft der Enterotoxin-Bildung charakterisiert. Die Enterotoxine werden in die hitzlabilen (LT) und die hitzestabilen (ST) Toxine unterteilt (LEVINE, 1987). GREEN et al. (1983) differenzieren in der ersten Gruppe LT I und LT II. Beide Holotoxine (griech. „holos“ = ganz) bestehen aus einer A-Untereinheit und weiterhin 4 identischen B-Untereinheiten. LT I (überwiegend Humanisolate) und LT II (überwiegend Isolate aus Tierbeständen) weisen sowohl in ihrer Struktur (80% Übereinstimmung auf der Proteinsequenzebene) als auch bezüglich ihrer Wirkungsweise große Ähnlichkeit mit dem Cholera-Toxin des *Vibrio cholerae* auf (O'BRIEN und HOLMES, 1987). Aus der Aktivierung einer Reaktionskette über die Adenylatzyklase resultiert die Aktivierung der Chloridkanäle der Epithelzellen. Das Nettoergebnis der Phosphorylierung der CFTR (engl. „cystic fibrosis transmembrane conductance regulator“) ist eine gesteigerte Chlorid-ionensekretion mit der Folge einer sekretorischen Diarrhoe (SEARS und KAPER, 1996; NATARO und KAPER, 1998). Unterstützend wirkt dabei die Stimulation der Prostaglandinsynthese und des enteralen Nervensystems (SEARS und KAPER, 1996).

Bei den hitzestabilen (ST) Toxinen handelt es sich um einfache Peptidtoxine, die in zwei in der Größe und im Pathomechanismus differierende Klassen (STa und STb) unterteilt werden. STa aktiviert über die Guanylatzyklase die oben beschriebene Hypersekretion. STb trägt über die Erhöhung der cytosolischen Calciumionenkonzentration, die Freisetzung von Prostaglandin E₂ und Serotonin zu einer erhöhten Ionensekretion bei (DUBREUIL, 1997). STa tritt bei ETEC-Stämmen des Menschen, der Ferkel und Kälber auf. STb wird bei

humanen Erkrankungen und sporadisch bei Kälbern und Hühnern nachgewiesen (DUBREUIL, 1997).

Die Adhäsion der ETEC-Zellen an das Dünndarmepithel erfolgt durch Fimbrien, die gleichzeitig die Spezies-Spezifität der ETEC begründen. Zunächst als K-Antigene bezeichnet, werden sie jetzt zusammen mit den Fimbrien der uropathogenen *E. coli* als F-Fimbrien bezeichnet (ØRSKOV und ØRSKOV, 1984). NATARO und KAPER (1998) beschreiben sie als vorwiegend Plasmid kodiert.

ETEC machen den Hauptanteil bei der „Colibazillose“ als sog. enterotoxische Enteropathie (Colidiarrhoe, Coliruhr) der neugeborenen Ferkel, Lämmer und Kälber aus (ROLLE und MAYR, 1993; NAGGY und FEKETE, 2005). So werden bei Ferkeln und Kälbern überwiegend F4-(K88-), F5-(K99-), F6-(987P-), F17-(Fy/Att25-), F41- und F18-Fimbrien nachgewiesen (NAGY und FEKETE, 1999).

ETEC gelten beim Menschen als Hauptverursacher der sog. Reisediarrhoe, die mit plötzlich einsetzendem, Cholera-ähnlichem Durchfall einhergeht. Fieber und Erbrechen werden selten beobachtet. Kontaminierte Lebensmittel einschließlich Trinkwasser lassen überwiegend Kleinkinder bzw. Reisende erkranken. (NATARO und KAPER, 1998).

2.2.3 ENTEROINVASIVE *E. COLI* (EIEC)

EIEC sind biochemisch, patho- und phylogenetisch eng mit *Shigella* spp. verwandt. Die Invasion der Bakterienzelle in das Kolonepithel erfolgt über den auf dem 140 MDa großen Plasmid plnV kodierten TypIII-Sekretionsapparat und das äußere Membranprotein IcsA (EGILE et al., 1998). Verschiedene Sekretionsproteine (IpaA bis IpaD) führen zu Neuarrangements des Zytoskelettes und zur Apoptose (TRAN VAN NHIEU et al., 2000). NATARO et al. (1985) und NATARO und KAPER (1998) verweisen auf die mögliche Existenz von Enterotoxinen.

Das klinische Erscheinungsbild umfasst wässrige Diarrhoe, teilweise mit Blut- und Schleimbeimengungen. Mit 0,8% der 2003 nach IfSG gemeldeten Fälle weist Deutschland die für die Industrieländer typische geringe Inzidenz auf (RKI, 2004). Die Serovaren O124:H-, O148:H- und O164:H- dominieren das Infektionsgeschehen in Deutschland (RKI, 1996). Als Vektoren werden Lebensmittel und Trinkwasser sowie „person-to-person“-Übertragungen beschrieben (NATARO und KAPER, 1998).

2.2.4 VEROTOXIN-BILDENDE *E. COLI* (VTEC)/ ENTEROHÄMORRHAGISCHE *E. COLI* (EHEC)

Insbesondere in den Industrienationen kommt der VTEC- bzw. EHEC-Gruppe eine besondere Bedeutung zu (RKI, 1996; BÜLTE und HECKÖTTER, 1997; KARCH et al., 2005). Im Jahr 2004 wurden 927 EHEC-Erkrankungen (außer HUS) und 52 HUS-Fälle sowie 5.558 Erkrankungen durch „sonstige darm-pathogene *E. coli*“ nach IfSG gemeldet (RKI, 2005). In dieser Statistik werden EHEC gesondert von sonstigen darm-pathogenen *E. coli* erfasst.

Die Bezeichnung „VTEC“ geht auf KONOWALCHUK et al. (1977) zurück, die solche *E. coli* aus verschiedenen Probenmatrizes isolieren konnten. Diese verfügten überein bislang unbekanntes Zytotoxinbildungs-vermögen. Aufgrund des zytotoxischen Effektes auf Verozellen (Zelllinie aus Nierenzellen grüner Meerkatzen) bezeichneten sie es als Verotoxin (VT). Die synonym verwendete Bezeichnung „Shiga-like Toxin (SLT)“ geht auf die Arbeiten O'BRIENS et al. (1997) zurück, die bei EPEC ein dem Shiga-Toxin von *Shigella dysenteriae* Typ 1 hochähnliches Toxin nachwiesen. Unter den einbezogenen Stämmen befand sich auch ein durch KONOWALCHUK et al. (1977) charakterisierter Stamm. O'BRIEN und LA VECK (1983) wiesen später nach, dass es sich bei VT und SLT um das gleiche Toxin handelt. Der von CALDERWOOD et al. (1996) eingebrachte Vorschlag einer Vereinheitlichung der Nomenklatur fand keine Zustimmung. Die Bezeichnungen VTEC und STEC werden daher synonym gebraucht.

Das Robert Koch-Institut schlägt die synonyme Verwendung der drei Begriffe EHEC/VTEC und STEC vor. Das IfSG enthält daher nur den Begriff „EHEC“ (ANONYMOUS, 2000). Eine gegensätzliche Meinung vertritt BÜLTE, der die Bezeichnung EHEC nur bei solchen VTEC-Stämmen empfiehlt, die bei entsprechend erkrankten Menschen isoliert worden sind. Ansonsten ist der Begriff „potenzielle EHEC“ zu verwenden (BÜLTE, 2000 und 2001). Das „Scientific Committee on Veterinary Measures Relating to Public Health“ (SCVPH) verwendet die Bezeichnung „human pathogenic VTEC (HP-VTEC)“ und empfiehlt die Einführung einer einheitlichen Terminologie (SCVPH, 2003). LEVINE et al. (1987) definierten solche Stämme als EHEC, die das klinische Erscheinungsbild der HC oder des HUS verursachen, Verotoxine exprimieren, A/E-Läsionen hervorrufen und über ein 60 MDa-Plasmid verfügen.

E. coli O157-Stämme stellen die dominanten, aber nicht einzigen Erreger des HUS dar (KARCH et al., 2005). Das Verotoxinbildungsvermögen konnte mittlerweile bei über 250 unterschiedlichen Serovaren nachgewiesen werden, von denen etwa 100 Stämme Erkrankungen beim Menschen verursachten (BETTELHEIM, 2005). Die klinische Bedeutung der non-O157-VTEC ist regional unterschiedlich. Besonders hohe Nachweisraten werden in Nordamerika verzeichnet (WHO, 1997; WHO, 1998; LOCKING et al., 2003). In Deutschland entfielen im Jahr 2003 78% der nach IfSG gemeldeten EHEC-Fälle (ohne HUS) sowie 12% der HUS-Meldungen, denen Angaben zur Serogruppe beigefügt waren, auf non-O157-EHEC (RKI, 2004). Das VT/STx wird allgemein als Hauptvirulenzfaktor angesehen. Dabei wird grundsätzlich zwischen VT1, VT2 und Varianten dieser Toxintypen unterschieden. So lässt sich die VT1-Gruppe in VT1, VT1c (ZHANG et al., 2002a; FRIEDRICH et al., 2003) und VT1d (BÜRK et al., 2003) untergliedern. In der VT2-Gruppe werden VT2, VT2c (SCHMITT et al., 1991), VT2c2 (JELACIC et al., 2003), VT2d (MELTON-CELSA et al., 1996), VT2e (WEINSTEIN et al., 1988), VT2f (SCHMIDT et al., 2000) und VT2g (LEUNG et al., 2003) unterschieden.

Eine Charakterisierung des Verotoxinbildungsvermögens, der Verteilung und Bedeutung von O157-Stämmen wurde von BÜLTE und HECKÖTTER (1997), QUINTERO BOTERO (2003) und GOLL (2005) vorgenommen (**Tabellen 2.2.4-1 und 2.2.4.-2**).

Table 2.2.4-1: Prominente VT 1-Varianten (vereinheitlichte Nomenklatur), wiedergegeben nach GOLL (2005)

Toxin-Variante	Referenzstamm	Nukleotidsequenzhomologie zur vtx 1 A ¹⁾	Nukleotidsequenzhomologie zur vtx 1 B ¹⁾	Quelle	Bemerkungen
VT 1 ²⁾	H19, C600 (933J)	100%	100%	SCOTLAND et al., 1985 STROCKBINE et al., 1986	
VT 1c	3115/97, 4756/98, 295/00		95%	ZHANG et al., 2002a	Identisch mit VT 1-OX3 nach PATON et al., 1995
VT 1d	MHI813		91%	BÜRCK et al., 2003	Aminosäuresequenz identisch mit VT 1v51 und VT 1v52 nach OHMURA-HOSHINO et al., 2003
VT 1v51	92-1251,	k. A.	k. A.	OHMURA-	Aminosäuresequenz
VT 1v52	92-1252	[94% AS] ³⁾	[94% AS] ³⁾	HOSHINO et al., 2003	identisch mit VT 1v51 und VT 1v52 nach BÜRCK et al., 2003
VT1-Variante mit IS1203v(1)	2000-182	k. A.	k. A.	SUZUKI et al., 2004	IS1203-ähnliches Element in A-Untereinheit

¹⁾ Sequenzen, die für die A- bzw. B-Untereinheit kodieren

²⁾ Verotoxin (VT)-Nomenklatur einheitlich angewandt; ursprüngliche Publikation teilweise als Shiga-Toxin (Stx) oder Shiga-like-Toxin (SLT)

³⁾ k. A.: keine Angabe; % AS: Aminosäuresequenz-Homologie

Table 2.2.4-2: Prominente VT 2-Varianten (vereinheitlichte Nomenklatur), wiedergegeben nach GOLL (2005)

Toxin-Variante	Referenzstamm	Nukleotidsequenzhomologie zur vtx 2 A ¹⁾	B ¹⁾	Quelle	Bemerkungen
VT 2 ²⁾	C600 (933W)	100,0%	100,0%	STROCKBINE et al., 1986	
VT 2vh	B2F1	98,6% ³⁾	95,5% ³⁾	ITO et al., 1990	Reklassifizierung zu VT 2c nach WHO, 1991; Umbenennung in VT 2d (VT 2d1, VT 2d2) nach MELTON-CELSA et al., 1998
VT 2vh-a					
VT 2vh-b					
VT 2vh-c	7279	99%	95%	MEYER et al., 1992	
VT 2va	H.I.8	69,5%	78,1%	GANNON et al., 1990	Vorschlag von SCHMIDT et al., 2000; Reklassifizierung in VT 2f
VT 2c	E32511	99,7%	95,2%	SCHMITT et al., 1991	Miteinbeziehung von VT 2vh-a und VT 2vh-b gemäß WHO, 1991
VT 2e	S1191	94%	79%	WEINSTEIN et al., 1988	ursprünglich VT 2v

¹⁾ Sequenzen, die für die A- bzw. B-Untereinheit kodieren

²⁾ Verotoxin (VT)-Nomenklatur einheitlich angewandt; ursprüngliche Publikation teilweise als Shiga-Toxin (Stx) oder Shiga-like-Toxin (SLT)

³⁾ bezogen auf VT 2vh-a

Fortsetzung Tabelle 2.2.4-2: Prominente VT 2-Varianten (vereinheitlichte Nomenklatur), wiedergegeben nach GOLL (2005)

Toxin-Variante	Referenzstamm	Nukleotidsequenzhomologie zur vt _x 2 A ¹⁾	B ¹⁾	Quelle	Bemerkungen
VT 2d ²⁾ (VT 2d1, VT 2d2)	B2F1	[siehe bei VT 2vh]		MELTON-CELSA et al., 1998	Nicht identisch mit VT 2d nach PIERARD et al., 1998; identisch mit VT 2vh (VT 2vh-a, VT 2vh-b) nach ITO et al., 1990
VT 2d (VT 2d-Ount, VT 2d-O111, VT 2d-OX3a)	EH250	94,9% ³⁾	86,3% ³⁾	PIERARD et al., 1998	Nicht identisch mit VT 2d nach MELTON-CELSA et al., 1998; Miteinbeziehung von VT 2-OX3a und VT 2-O111 nach PATON et al., 1992 und PATON et al., 1993
VT 2f	T4/97	63,4%	75,4% ⁴⁾	SCHMIDT et al., 2000	Miteinbeziehung von VT 2a
VT 2g	7v	k. A. ⁵⁾ [63,0-94,9%] ⁶⁾	k. A. [76,7-90,7%] ⁶⁾	LEUNG et al., 2003	
VT 2-NV206	NV206	k. A. [94,5-99,0%] ⁶⁾	k. A. [81,5-96,0%] ⁶⁾	BERTIN et al., 2001	

1) Sequenzen, die für die A- bzw. B-Untereinheit kodieren

2) Verotoxin (VT)-Nomenklatur einheitlich angewandt; ursprüngliche Publikation teilweise als Shiga-Toxin (Stx) oder Shiga-lik-Toxin (SLT) nach FRIEDRICH et al., 2002

3) nach FRIEDRICH et al., 2002; Angabe nach SCHMIDT et al., 2000: 57,4%

4) keine Angabe; 6) bezogen auf VT 2 und VT 2-Varianten

Durch *E. coli* O157-Stämme verursachte EHEC-Erkrankungen beim Menschen wurden erstmalig 1983 durch RILEY et al. (2003) in den USA beschrieben. Als Krankheitsbild dominierte die hämorrhagische Colitis (HC). KARMALI et al. (1983) beschrieben sporadisch auftretende Infektionen mit Entwicklung des hämolytisch-urämischen Syndroms (HUS).

Bei Säuglingen und Kleinkindern (engl. „young“), alten (engl. „old“), schwangeren (engl. „pregnant“) und immunsupprimierten (engl. „immunocompromised“) Bevölkerungsanteilen (engl. „segments of the public“) („YOPIIS-Gruppe“) können VTEC-Erkrankungen, die ansonsten wegen ihres milden Verlaufs nicht erfasst werden, lebensbedrohlich kompliziert werden (RKI und BGVV, 2001). So treten im Durchschnitt drei bis neun Tage nach der Infektion kolikartige Darmkrämpfe, anschließend wässrige Diarrhoe, gelegentlich in Begleitung von leichtem Fieber und Erbrechen auf. Dabei entwickelt sich bei 10 bis 20% der Fälle eine hämorrhagische Colitis mit starken Abdominalschmerzen, blutigem Stuhl zuzüglich Fieber (RKI und BGVV, 2001). NATARO und KAPER (1998) beschreiben die Ausheilung der Infektion nach etwa sieben Tagen. Bei etwa 1-5% der Fälle kommt es zu extraintestinalen Komplikationen wie HUS oder thrombotisch-thrombozytopenische Purpura (TTP). Es werden „komplettes HUS“ mit hämolytischer Anämie, Thrombozytopenie und akutem Nierenversagen und das „enteropathische HUS“ im Anschluss an ein Durchfallgeschehen unterschieden (GASSER et al., 1955; SIEGLER, 1995; GERBER et al., 2002). Während HUS überwiegend bei Kleinkindern im Alter von unter fünf Jahren auftritt, sind durch die TTP eher adulte Patienten betroffen. Neben der Schwere der Erkrankungsverläufe kommt eine erhebliche Beeinträchtigung der Lebensqualität hinzu. 60% der HUS-Patienten werden dauerhaft dialysepflichtig. Die Mortalitätsrate beträgt etwa 5% (ELLIOT et al., 2001; GERBER et al., 2002).

Dem Menschen kommt als symptomlosen VTEC-Ausscheider eine nicht abschließend zu bewertende Bedeutung zu. LUDWIG et al. (2002) wiesen im Umfeld von HUS-Erkrankten symptomlose Ausscheider nach. Landwirtschaftliches bzw. in der fleischverarbeitenden Industrie tätiges

Personal kommt als VTEC-/STEC-Träger in Betracht (GAREIS et al., 2000; STEPHAN et al., 2000 und SILVESTRO et al., 2004).

Ein weiterer Indikator für die gesundheitspolitische Relevanz der VTEC-Infektionen sind die ökonomischen Aufwendungen für die Medikation und Hospitalisation der Patienten. ABE et al. (2002) verzeichnen expandierende Kosten erst durch die aufwändige Behandlung komplizierter Verläufe. ROBERTS et al. (2000) konnten anhand eines 1994 in Schottland erfolgten *E. coli* O157:H7-Ausbruchs Kosten zwischen 20.000 (TTP-Fall) und etwa 60.000 (HUS-Fall) britischen Pfund ermitteln.

2.2.5 ENTEROAGGREGATIVE *E. COLI* (EAggEC)

Die EAggEC bzw. EAEC treten laut ITOH et al. (1997) zunehmend häufiger als Infektionserreger des Menschen in Erscheinung. In Entwicklungsländern besitzen sie eine besondere Bedeutung. Als Erreger persistierender wässriger Diarrhoeen mit Schleimbeimengungen erlangten sie besondere Bedeutung (FANG et al., 1995). Nach NATARO et al. (1987) handelt es sich um LT- und ST-Toxin negative Stämme, die an HEp2-Zellen (Zelllinie, die auf eine aus einem humanem Larynxkarzinom gewonnene Linie zurückgeführt wird) ein typisches aggregatives Adhärenzmuster zeigen. Etwa die Hälfte der Stämme verfügt über das Potenzial zur Bildung des sogenannten EAST 1 (EAggEC heat stable enterotoxin 1), das auch bei EPEC-, ETEC- und VTEC-Stämmen detektiert werden kann (SAVARINO et al., 1996; PAIVA DE SOUSA und DUBREUIL, 2001). Die Adhäsine AAF/I und AAF/II (aggregative adherence fimbriae) werden nur bei einem geringen Teil der Stämme identifiziert und sind plasmidverankert (NATARO et al., 1998; PIVA et al., 2003). PABST et al. (2003) identifizierten sie als Verursacher der Reisediarrhoe. In Deutschland entfielen von den 4.705 im Jahre 2003 nach IfSG gemeldeten *E. coli*-Infektionen (ohne EHEC) 0,2% auf diese Gruppe (RKI, 2004).

2.2.6 DIFFUS-ADHÄRENTE *E. COLI* (DAEC)

DAEC werden über ihr charakteristisches, diffuses Anheftungsmuster an HEp-2-Zellen definiert. Sie werden als Verursacher wässriger Durchfallerkrankungen bei Kindern, überwiegend älter als 12 Monate, beschrieben (NATARO und KAPER, 1998 und SCALETSKY et al., 2002). Detaillierte Beschreibungen des Pathomechanismus, der ein „Aufrollen“ des DAEC innerhalb langer Wirtszellfortsätze beinhaltet, fehlen bislang. Die Beteiligung von Fimbrien wird bei mehr als 75% der Isolate beschrieben (BILGE et al., 1989). In Deutschland wurden 0,3% der nach IfSG gemeldeten *E. coli*-Fälle dieser Erregergruppe zugeordnet (RKI, 2004).

2.2.7 NEKROTOXISCHE *E. COLI* (NTEC)

NTEC zeichnen sich durch die Bildung eines Zytonekrose-Faktors (engl. „cytotoxic necrotizing factor“ [CNF]) aus (CAPRIOLI et al., 1983). DE RYCKE et al. (1999) untergliederten sie anhand der Veränderungen bei HeLa-Zellkulturen in CNF 1 (NTEC 1)- und CNF 2 (NTEC 2)-Stämme. Aufgrund der bei einigen Stämmen nachgewiesenen Hämolyse (Hly, FALBO et al., 1992) und des CDT (engl. „cytolethal distending toxins“) kommt ihnen nach Ansicht dieser Autoren ein hohes pathogenes Potenzial zu. CNF 1-Stämme werden sowohl beim Menschen als auch bei Haustieren, CNF 2-Stämme bisher ausschließliche bei Wiederkäuern nachgewiesen. BLANCO et al. (1992) gewannen Isolate von septikämischen bzw. durchfallerkrankten Kälbern; TOTH et al. (2000) von durchfallerkrankten Saugferkeln. ORDEN et al. (2002) gelang der Nachweis auch bei klinisch gesunden Tieren. Beim Menschen treten NTEC-Infektionen überwiegend extraintestinal auf. TAVECHIO et al. (2004) konnten allerdings bei Kindern mit und ohne Diarrhoe CNF 1-positive Stämme isolieren. In Lebensmitteln gelang ein Nachweis bisher nur selten (QUINTO und CEPEDA, 1996).

2.2.8 „CYTOLETHAL DISTENDING TOXIN“-PRODUZIERENDE *E. COLI* (CLDTEC)

„Cytotoxic distending toxins“ (CDT) gehören zu einer Gruppe genotoxischer Proteine, die durch eine Blockade des eukaryotischen Zellzyklus im G1- oder G2-Stadium zum nichtapoptotischen Zelltod führen (COPE et al., 1997; CORTES-BRATTI et al., 2001). Bei *E. coli* tritt die Blockade im G2/M-Stadium auf (BIELASZEWSKA et al., 2005). Die CDT sind aus den drei Untereinheiten CDTA bis CDTC zusammengesetzt. Dabei bildet CDTB die enzymatisch aktive Komponente als DNase I (DE RYCKE und OSWALD, 2001), während CDTA und CDTC am Transport von CDTB beteiligt sind (LARA-TEJERO und GALAN, 2000). Ihr Vorkommen ist neben *E. coli* auch bei *Campylobacter* spp., *Shigella* spp., *Salmonella enterica* serovar Typhi, *Haemophilus ducreyi*, *Actinobacillus actinomycetemcomitans* und *Helicobacter* spp. beschrieben (DE RYCKE und OSWALD, 2001). Bislang wurden bei *E. coli* fünf CDT-Varianten beschrieben, die im Falle von CDT I, CDT II und CDT V chromosomal verankert sind (SCOTT und KAPER, 1994; PICKET et al., 1994; JANKA et al., 2003). PERES et al. (1997) identifizierten CDT III auf dem pVlr-Plasmid. CDT V konnte dabei sowohl bei Sorbit-fermentierenden (SF) STEC O157:H⁻ und in geringeren Anteilen bei O157:H7 nachgewiesen werden. Zusätzlich konnten auch die *eae*-negativen, humanpathogenen Serovaren O91:H21 und O113:H21 als CDT V-Produzenten ermittelt werden (BIELASZEWSKA et al., 2004). CDT wird zusätzlich auch bei uropathogenen, nekrotoxischen und enteropathogenen Stämmen nachgewiesen (CLARK et al., 2002).

2.2.9 ZELL-ABLÖSENDE *E. COLI*, DIARRHOE-ASSOZIIERTE HÄMOLYTISCHE *E. COLI* (DHEC)

GUNZBURG et al. (1993) beschrieben diese Pathogruppe anhand eines als typisch eingestuften Zell-Ablösungs (engl. „cell detaching“)-Phänomens in der Zellkultur und dem Vorkommen von α -Hämolysin (MARQUES et al., 1995).

Die Stämme werden daher im englischen Sprachgebrauch als „cell-detaching, *E. coli*“ bzw. „diarrhoe-associated hemolytic *E. coli*“ bezeichnet. Sie weisen weitere Virulenzfaktoren auf, so z. B. das ST-Toxin der enterotoxinogenen *E. coli* (ABDUCH FABREGA et al., 2002), was eine eindeutige Gruppenzuordnung erschwert.

2.3 BESCHREIBUNG DES *EAE*-GENS

Das bei vielen EPEC und VTEC/STEC vorkommende *eae*-Gen kodiert für das äußere Membranprotein (engl. „outer membrane protein [omp]“). Intimin ist als Adhäsionsfaktor mittlerweile gut beschrieben worden. Ihm kommt sowohl *in vitro* (AGIN et al., 1996; BEEBAKHEE et al., 1992; DYTOC et al., 1993; LOUIE et al., 1993; McDANIEL et al., 1995; SCHOOLNIK, 1993; YU und KAPER, 1992) als auch *in vivo* eine wichtige Rolle in der Pathogenese zu (DONNENBERG et al., 1993A, 1993B, 1993C; TZIPORI et al., 1995).

2.3.1 ERSTBESCHREIBUNG UND CHARAKTERISIERUNG

JERSE et al. gelang es im Jahr 1990 ein chromosomales Gen bei EPEC, RDEC und VTEC zu identifizieren und seine Beteiligung an der Entstehung der A/E-Läsionen (STALEY et al., 1969) aufzuzeigen. Über Mutagenese des identifizierten Gens unter Verwendung des Transposons *TnphA* konnten sie die

charakteristischen Veränderungen bei zuvor A/E-negativen EPEC induzieren. Das identifizierte Gen bestand aus einem offenen Leserahmen (engl. „open reading frame“ [orf]), dem sie die Bezeichnung *eae*-Gen zuwiesen. Anhand der Aminosäuresequenz leiteten sie ein Molekulargewicht von 102 kDa (engl. „kilo Dalton“) ab. Der genetische Nachweis des neuen Anheftungsfaktors aus bakterieller DNS erfolgte zunächst über eine DNA-Hybridisierungssonde. 1991 ermittelten JERSE und KAPER die tatsächliche Proteingröße mit 94 kDa bei dem EPEC-Stamm E2348/69. Den Unterschied zur theoretisch anhand des 2817 Basenpaare umfassenden Gens publizierten Größe, begründeten sie mit posttranslationalen Veränderungen bzw. mangelhafter Auflösungsstärke ihres SDS-Page- (engl. „sodium dodecylsulfate polyacrylamide gel electrophoresis) Verfahrens. DONNEBERG und KAPER (1991) bestätigten die Daten der Arbeitsgruppe um JERSE, zusätzlich erweitert durch die Charakterisierung der erfolgreichen Deletion ihrer Mutante mittels spezifischer PCR-Primer. Die Aminosäuresequenz verfügt über einen hohen Homologiegrad des N-Terminus (5'-Ende des Gens) mit dem daher auch als *eae*-Genhomolog bezeichneten Äquivalent von *Citrobacter freundii* (SCHAUER und FALKOW, 1993a; 1993b). Homologe Gene existieren ebenfalls bei *Yersinia enterocolitica* und *Yersinia pseudotuberculosis* (ISBERG et al., 1987) sowie *Hafnia alvei* (FRANKEL et al., 1994). Hinsichtlich der funktionellen Auslegung wurden bereits über Invasin bekannte Fakten, wie etwa die Rolle der Disulfidbrücke, der KNUTTON et al. (1987) zunächst noch eigenes Bindungspotenzial zuwiesen, übertragen. KELLY et al. (1999) ermittelten die Struktur des hypervariablen Intiminabschnittes Int280 (Intiminabschnitt der hypervariablen Region, der 280 Aminosäuren umfasst) über das Nuklear-Magnetresonanz-Verfahren. Diese Region umfasst drei Domänen, zwei Immunglobulin-ähnliche und eine Domäne, die dem „C-Typ-Lektin“ (150 Aminosäuren) ähnelt und für die Bindung an Tir über Zucker (Karbhydrate) verantwortlich sein soll. LUO et al. (2000) ermittelten eine gute Übereinstimmung zwischen den Sekundärstrukturelementen und den Domänengrenzen, die von KELLY et al. (1999) festgelegt wurden. Sie charakterisierten Intimin als strukturell hochähnlich zum Invasin, einem äußeren Membranprotein der Shigellen. Invasin verfügt über fünf Domänen und

ebenfalls einen Karbohydratanker (HAMBURGER et al., 1999, vgl. **Kapitel 5.3.2**).

Das Intimigen ist auf dem LEE (engl. „Locus of Enterocyte Effacement“ verankert (McDANIEL et al., 1995, 1997). Dabei handelt es sich um eine aquirierte Pathogenitätsinsel der chromosomalen DNA. Aufgrund des mit etwa 41% geringeren G+C-Gehaltes, im Vergleich zu dem des bakteriellen Genoms mit etwa 50%, gehen die Autoren von einer Fremdaufnahme aus. PERNA et al. (1998) sequenzierten den LEE des EHEC-Stammes EDL 933 O157:H7.

Im Vergleich zu der durch ELLIOT et al. (1998) erarbeiteten Sequenz des EPEC E2348/69 O127:H6 LEE, verfügt er mit circa 44 kBp im Gegensatz zu den circa 36 kBp LEE des EPEC über den größeren Umfang. SANDNER et al. (2001) konnten zeigen, dass der LEE eine dynamische genetische Einheit darstellt, die zur Aufnahme und Abgabe von genetischer Information befähigt ist. RUMER et al. (2003) identifizierten mit *pheU/pheV* als Ergänzung zu *se/C* zwei weitere Verankerungszonen. Hierbei handelt es sich um transfer (t)RNA-Gene. Beide LEE beinhalten 41 offene Leseraster, die für einen TypIII-Sekretionsapparat, das Intimin und den translozierten Intiminrezeptor Tir kodieren. Der LEE des O157 verfügt zusätzlich über 13 offene Leseraster eines Prophagen.

2.3.2 SEQUENZVERGLEICH, NOMENKLATUR UND DIAGNOSTISCHER NACHWEIS

BEEBAKHEE et al. (1992) lieferten die *eae*-Gensequenz des EHEC-Stammes EDL 933 O157:H7, bei dem das genannte Gen von den zuvor genannten Autoren mehrfach nachgewiesen worden war. LOUIE et al. (1993) sequenzierten den Stamm CL8 O157:H7. Die genetische Übereinstimmung betrug für die ersten 2200 Bp 97% bzw. 59% für den 3'-Bereich. Sie fanden ein 97 kDa großes *eae*-Genprodukt bei *E. coli* O157:H7, das auf der Oberfläche des Bakteriums exprimiert wird. Aufgrund der Sequenzunterschiede wurde eine serovarspezifische Zuordnung angenommen. Das phasenweise auch als *eaeA*-

Gen bezeichnete *eae*-Gen wurde daher vielfach mit einem die Bezeichnung des O-Antigens tragenden Suffix gekennzeichnet (LOUIE et al., 1993). Die bereits erkannten Sequenzunterschiede im hypervariablen Bereich wurden zunächst für die Etablierung diagnostischer Nachweisverfahren, als evolutive Marker und - immer in Analogie zu Invasin - als mögliche bindungsvermittelnde Regionen angesehen. LOUIE et al. (1994) zeigten dabei, dass O55:H7- und O157:H7-Stämme das *eae*-Gen besitzen.

Neben der Verwendung unterschiedlicher Gensonden und unterschiedlicher PCR-Verfahren kam den Zellkulturtests besondere Bedeutung zu. KNUTTON et al. (1989) stellten den „Fluorescent Actin Staining“ (FAS)-Test als diagnostisches Hilfsmittel zur Charakterisierung A/E-Läsionen verursachender EPEC und VTEC vor. In Verbindung mit der zunächst verfügbaren Sonde für das EAF-Plasmid evaluierten SHARIFF et al. (1993) den FAS-Test, den sie als zuverlässiger und spezifischer im Vergleich zur Sonde einstufen. BEUTIN et al. (2003) setzten das Verfahren zur Charakterisierung von EPEC und VTEC, die aus dem Stuhl deutscher und australischer Kinder gewonnen wurden und über α -, β -, γ - oder ζ -Intimin verfügten, ein.

Für die PCR-gestützte Detektion des *eae*-Gens wurden eine Vielzahl unterschiedlicher Oligonukleotid-Primer beschrieben. Häufige Verwendung findet dabei das Primer-Paar „SK1/SK2“ (KARCH et al., 1993). Obwohl KARCH et al. (1993) als Erstbeschreibende genannt werden, findet sich in der genannten Arbeit lediglich die auch in den eigenen Untersuchungen verwendete Sequenz für „SK1“. Die Angabe der heute verwendeten Primer-Sequenzen wie auch die Bezeichnung als „SK1/SK2“ wurde von SCHMIDT et al. (1994) publiziert. Im Folgenden wird die gängige Zitierweise verwendet. WIELER et al. (2000) verwenden die Primer-Bezeichnung „SK1 (VT)/SK2 (VT)“ zur Amplifikation des *vtx 1/stx 1*-Gens.

2.3.3 INTIMIN-VERMITTELTE ANHEFTUNG

DONNENBERG und KAPER (1992) sowie KNUTTON (2003) charakterisierten die Bildung der A/E-Läsionen als vierstufigen Prozess. In der ersten Phase kommt es zur Exprimierung des „bundle forming pilus“ (BfP). In der zweiten Phase wird die innige Verbindung über das Sekretionsprotein EspA und Intimin eingeleitet. Über ein TypIII-Sekretionsystem (TTSS) werden der Rezeptor Tir und weitere Sekretionsfaktoren (engl. „E. coli secreted proteins“ [ESP]) exprimiert. Tir wird phosphoryliert und die Umstrukturierung des Zytoskelettes beginnt. Das TTSS und EspA bilden eine physische Verbindung in Form einer „molekularen Spritze“ (FIVAZ und VAN DER GOOT, 1999), durch die EspB und EspD-Proteine als Poren bildende Proteine in die Zellmembran eingeschleust werden können (KNUTTON et al., 1998; SHAW et al., 2001). WAINWRIGHT und KAPER (1998) sowie ABE et al. (1999) belegten die Notwendigkeit des CesD/CesT als (engl. „chaperone for secreted *E. coli* proteins“) sog. Gouvernantenproteine zur Sekretion von EspB und EspD. An der Umstrukturierung des Zytoskelettes sind die intrazellulären Bestandteile Vinculin, Cortactin, Talin und α -Aktinin beteiligt (CANTARELLI et al., 2000a, 2000b; FREEMAN et al., 2000; GOOSNEY et al., 2000). Die drei letztgenannten Komponenten interagieren direkt mit Tir. Dabei handelt es sich um das EspE (DEIBEL et al., 1998), das zuvor durch ROSENSHINE et al. (1992) als Hitzeschockprotein (hsp90) bezeichnet wurde. MUZA-MOONS et al. (2003) konnten zeigen, dass EPEC die Wirtszellpolarität verändern können. Dabei wandern insbesondere β 1-Integrine an die Apikalseite der Zelle, um Intimin zu binden. SINCLAIR und O'BRIEN (2002) konnten für das γ -Intimin von *E. coli* O157 das Intimin unabhängige Bindungsvermögen an Nucleolin zeigen. Bei Versuchen an α - und β -Intimin tragenden Stämmen schienen Tir und Nucleolin um freie Bindungsstellen zu konkurrieren. War die Bindung an Tir erfolgt, wurde die Bindung an Nucleolin blockiert. Unterschiede zwischen EPEC und EHEC bei der Formung der Signalkaskade bestehen für EPEC in der

Bindung von Tir an das Wirtszell-eigene Signalprotein Nck, das über das N-WASP (engl. „neural Wiskott-Aldrich-sndrome protein“) an das Arp 2/3 (engl. „actin-related protein 2/3“) bindet (KALMAN et al., 1999). Der Tir der EHEC benötigt Nck nicht. In der dritten Phase kommt es zur Abtrennung der EspA-Filamente von der Zelloberfläche. Intimin bindet an Tir und die zellulären Veränderungen werden erkennbar. EPEC sind dabei durch die „localized adhearance“ charakterisiert. Neben den genannten Sekretionsfaktoren sind eine Vielzahl weiterer Faktoren beteiligt. Dazu gehören EspF als an der Apoptose beteiligter Faktor, EspG, EspH und Map als mitochondrientoxischer Faktor. CAMPELLONE und LEONG (2003) konnten zeigen, dass EPEC und EHEC Tir aufgrund der Bindung des EPEC Tir an Nck und dessen Tyrosinphosphorylierung nicht vollständig austauschbar sind.

DE VINNEY et al. (2001) zeigten in Mutationsversuchen, dass EPEC Tir die vollständige Anheftung bei EHEC nicht vermitteln kann und *vice versa*. GOOSNEY et al. (2000) sehen die Tyrosinveränderung als notwendig zur Bindung der Wirtszelladapterproteine Grb2 und CrkIII an. BATCHELOR et al. (2000) ermittelten die strukturelle Basis für die Erkennung des Tir durch Intimin. Sie liegt innerhalb des Int190 (beinhaltet D3 und D4 nach LUO et al., 2000) und lässt sich sogar bis auf ein 55 Aminosäuren großes Motiv eingrenzen (KENNY et al., 1997). Dabei sollen weniger als fünf Aminosäuren des Tir mit Intimin interagieren (**Kapitel 5.3.2**).

2.3.4 BINDUNG AN DIE WIRTSZELLE OHNE INTIMIN

Bereits die ersten vergleichenden Untersuchungen mittels FAS-Test und DNA-Sondenhybridisierung ergaben, dass das Intimin zwar für die Ausprägung der A/E-Läsionen notwendig, aber nicht allein verantwortlich ist. Aufmerksamkeit erregten LEE-negative Stämme, die trotzdem mit schwerwiegenden Krankheitsverläufen assoziiert waren (FRIEDRICH et al., 2003; JENKINS et al., 2003). PATON et al. (2001) beschrieben das „STEC autoagglutination adhesin“ (Saa) erstmals bei einem HUS-Ausbruchsstamm des Serovars O113:H21. Alle Saa-positiven Isolate besitzen zusätzlich das Enterohämolysin- (FRIEDRICH et al., 2003) und häufig *vtx 1c*-Gen (KUMAR et al., 2004b; ZWEIFEL et al., 2004).

2.4 Ableitung molekularer Phylogenie

2.4.1 TERMINOLOGIE

Die Phylogenetik (griech. „phylon“ = Stamm, Geschlecht; genesis= Entstehung) oder Kladistik (griech. „klados“ = Zweig) beschreibt die evolutive Entwicklungsgeschichte der jeweils betrachteten, taxonomischen Merkmale oder Einheiten. Die phylogenetische Analyse verwendet zur Untersuchung und Darstellung der Rangfolge evolutiver Beziehungen überwiegend Baumdiagramme. Die Nomenklatur der Darstellung variiert je nach Quellenangabe und wird daher kurz zusammengefasst (**Abbildung 2.4.1-1**). Ein Stammbaumdiagramm (engl. „tree diagram“) stellt ein anhand eines mathematischen Modells ermitteltes Ergebnis dar. Als sog. abgeleiteter Baum hängt die Übereinstimmung mit dem wahren Baum neben der Anzahl und der Qualität der untersuchten Merkmale (hier: Nukleotid- sowie Aminosäuresequenzen der *eae*-Gene) von der Eignung des gewählten Berechnungsmodells ab.

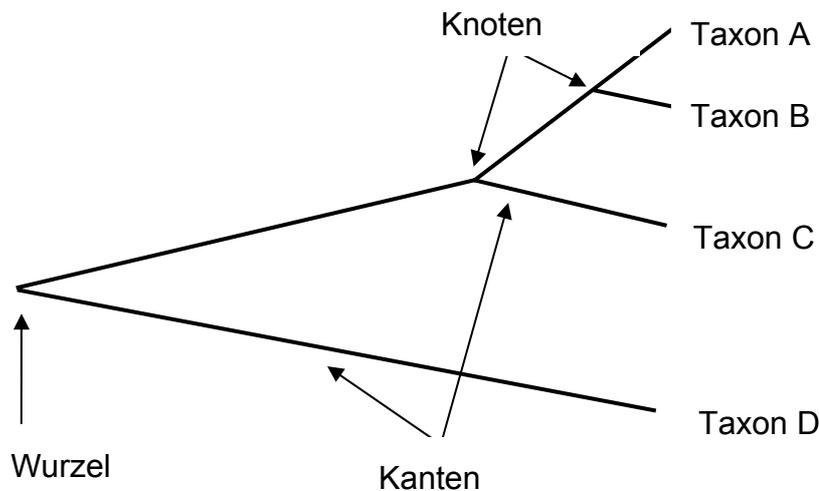


Abbildung 2.4.1-1: Phylogenetische Terminologie am Beispiel eines ungewurzelten Baumes in Anlehnung an BRINKMANN (2005)

Als Taxon wird das betrachtete phylogenetische Merkmal bzw. die Einheit (z. B. Sequenz 1 für „Intimin α “), mehrere Taxa (z. B. alle „Intimin α “-Sequenzen) werden als Gruppe (engl. = „clade“) bezeichnet. Ein Baumdiagramm besteht aus Knoten (engl. = „nodes“), die durch Kanten (engl. = „edges“) bzw. Äste (engl. = „branches“) verbunden sind. Knoten symbolisieren Differenzierungsereignisse. Terminale Knoten (engl. = „terminal nodes“ oder „leaves“) bzw. terminale Taxa repräsentieren die Sequenzen, für die entsprechende Daten vorhanden sind. Da sie Grundlage der Berechnungen sind, werden sie häufig auch als „Operational Taxonomic Units“ (OTU) geführt. Die vorgelagerten internen Knoten (engl. = „internal nodes“) stehen für hypothetische Vorfahren. Der Vorläufer aller Sequenzen bildet die Wurzel, d. h. die Kante oder den Ast, die/der vom letzten gemeinsamen Vorfahren aller terminalen Knoten ausgeht. Ist ein gemeinsamer Vorfahre nicht bekannt oder kann dieser nicht ermittelt werden, ergibt sich eine „ungewurzelte Darstellung“ (BRINKMANN, 2005). Alternativ können Außengruppen (z.B. das phylogenetisch weit entfernte *Citrobacter rodentium eae*-Genhomolog) zur

Wurzelung des Baumdiagramms verwendet werden. Dendrogramme (griech. „dendros“ = Baum) werden überwiegend als dichotome (griech. „dicha“ = getrennt und „tome“ = Schnitt) Diagramme dargestellt, d. h. jeder Separation folgen zwei divergierende Populationen. Folgende Arten von Baumdarstellungen werden unterschieden: Bei dem sog. Kladogramm (griech. „klados“ = Ast) ist die dargestellte Länge der Kanten nicht von Bedeutung. Es findet Anwendung bei der Split-Zerlegung. Beim Phylogramm entspricht die Länge der Kanten der genetischen Veränderung des betrachteten Parameters. Es wird z. B. bei der phylogenetischen Analyse der Intiminvarianten eingesetzt. Der sog. ultrametrische Baum gibt mit der Kantenlänge die evolutive Zeitdauer wieder. Beim sog. kodierten Baum erfolgt die Anordnung der Taxa zueinander in Form eines Klammerterminus, z. B. bei der Verarbeitung berechneter Bäume durch unterschiedliche Computerprogramme (BRINKMANN, 2005).

2.4.2 GENETISCHE DISTANZEN

Die molekulare Phylogenetik setzt voraus, dass die evolutive mit der genetischen Distanz zweier Taxa korreliert. Die Kalkulation der Nukleotidsubstitutionen zwischen Adenin (A), Cytosin (C), Guanin (G) und Thymin (T) dient als Grundlage zur Ermittlung phylogenetischer Distanzen. Die entwickelten Berechnungsmodelle basieren auf den in **Abbildung 2.4.2-1** dargestellten Grundprinzipien. Als Transitionen (α) werden Wechsel von Purinen oder Pyrimidinen, als Transversionen (β) werden Wechsel von Purinen gegen Pyrimidine und umgekehrt bezeichnet.

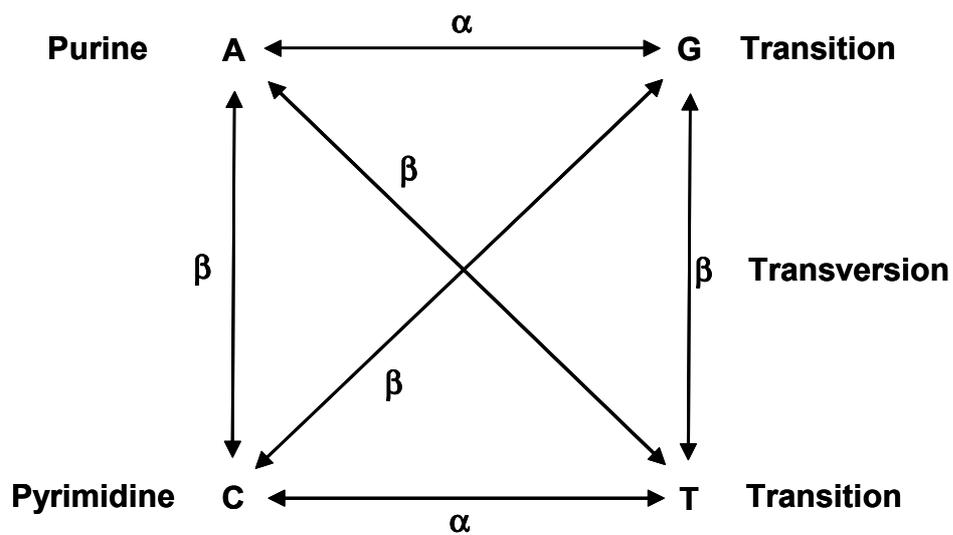


Abbildung 2.4.2-1: Grundprinzipien möglicher Nukleotidsubstitutionen

Unter Berücksichtigung der Auswirkung der Substitution auf die Aminosäuresequenz werden Nukleotidsubstitutionen, die zu keiner Veränderung der Aminosäuresequenz führen als synonyme (engl. synonymous) oder stille (engl. „silent“) Veränderungen benannt. Im Gegensatz dazu werden Mutationen als nonsynonym (engl. „nonsynonoumus“) klassifiziert, aus denen eine Veränderung der Aminosäureabfolge resultiert. Sinnlose (engl. „nonsense“) Veränderungen führen zu Stopcodons (GOJOBORI et al., 1982).

2.4.3 MODELLE DER SEQUENZEVOLUTION

Das Verhältnis (engl. = proportion) „ p “ der Nukleotidpositionen, in denen sich zwei Sequenzen unterscheiden dient als einfache Bezugsgröße zur Beschreibung der Divergenz. Das Verhältnis „ p “ wird über die folgende Gleichung ermittelt: $p = n_d/n$, mit n_d = Anzahl unterschiedlicher Nukleotide, n = Gesamtzahl überprüfter Nukleotide.

Synonym dazu wird der erhaltene Wert als „ p -Distanz“ für Nukleotidsequenzen bezeichnet (GOJOBORI et al., 1982). Das Verhältnis „ p “ wird dabei für synonyme Substitutionen als „ p_S “ und nonsynonyme Substitutionen als „ p_N “ dargestellt. Die absolute Anzahl synonymmer bzw. nonsynonymer Substitutionen wird mit „ d_S “ und „ d_N “ bezeichnet.

Auf Basis der „ p -Distanz“ wurden unterschiedliche Modelle zur Darstellung der Sequenzevolution entwickelt. Ihr wesentlicher Unterschied liegt in der Gewichtung der Häufigkeit des Vorkommens der Transitionen und Transversionen. YANG (1997a, 1997b), YANG und YODER (1999) sowie HUGHES und FRIEDMAN (2000) konnten zeigen, dass eine einfache Bestimmung des Verhältnisses der Transitionen zu den Transversionen (engl. „transition/transversion ratio“) „ R “ über die Gleichung $R = P/Q$ mit P = Gesamtzahl der Transitionen und Q = Gesamtzahl der Transversionen zu fehlerhaften Ergebnissen führen kann. Daher wurde für das eigene Vorgehen „ R “ unter ausschließlicher Berücksichtigung der vierfach veränderten (engl. „4-fold-degenerated-sites“) Positionen ermittelt. Darin sind ausschließlich synonyme Positionen enthalten.

Für die eigenen Berechnungen (siehe **Kap. 3.5.6**) wurde das modifizierte Modell nach NEI und GOJOBORI (NEI und GOJOBORI, 1986 und ZHANG et al., 1998) und das Verfahren nach TAMURA und NEI (1993) angewandt. Sie ordnen sich in den nachfolgend erläuterten kontinuierlichen Entwicklungsprozess ein. Hier finden sich sowohl Distanz- als auch auf der Berechnung der größten Wahrscheinlichkeit (engl. = „maximum-likelihood“ [ML])-basierende Formeln.

2.4.3.1 JUKES-CANTOR (JC)-MODELL

Unter der Voraussetzung, dass die Nukleotidsubstitutionen über die Zeit konstant bleiben und die Basenzusammensetzung im Gleichgewicht liegt, entwickelten JUKES und CANTOR (1969) ein einfaches Modell. Die Wahrscheinlichkeit für Transitionen und Transversionen wurde gleich gesetzt.

Anders ausgedrückt: Der Einfluss der Selektion auf einzelne Genabschnitte wurde nicht berücksichtigt. Die Sequenzevolution wird als mechanischer Zufallsprozess behandelt.

2.4.3.2 HASEGAWA, KISHINO UND YANO (HKY85)-MODELL

KIMURA (1980) konnte zeigen, dass Transitionen häufiger vorkommen als Transversionen. Im daraufhin erarbeiteten Kimuras 2-Parameter (K2P)-Modell wird dieser Beobachtung Rechnung getragen, indem die Transitionsrate (α) und die Transversionsrate (β) je Position zu einer totalen Substitutionsrate von $\lambda = \alpha + 2\beta$ zusammengefasst wurden. Damit ergaben sich für ein Nukleotid drei Substitutionsmöglichkeiten, bestehend aus einer Transition und zwei Transversionen. FELSENSTEIN (1981) berücksichtigte zusätzlich die Variation der Basenzusammensetzung. So ist z. B. der GC-Gehalt des LEE höher als der des gesamten *E. coli*-Genoms. Die unterschiedliche Basenzusammensetzung wird in FELSENSTEINS (F81)-Modell als Näherung (π_i) in Form der durchschnittlichen Frequenz der Base i über die gesamte Sequenz berücksichtigt. Eine Kombination der Modelle gelang HASEGAWA et al. (1985). Das HKY85-Modell gestattet unterschiedliche Transitions- und Transversionsraten, kombiniert mit variabler Basenfrequenz.

2.4.3.3 TAMURA UND NEI (TRN)-MODELL

Das HKY85-Modell repräsentiert einen ML-Ansatz. TAMURA und NEI (1993) erweiterten es um einen Parameter für Transitionen zwischen Purinen (α_1 , A \leftrightarrow G) und Transitionen zwischen Pyrimidinen (α_2 , T \leftrightarrow C) und etablierten das TrN-Modell. Die Untergliederung der Transitionsrate (α) war notwendig geworden, nachdem bei der Untersuchung der humanen D-loop-Region Abweichungen um den Faktor zwei ermittelt wurden.

Zusammenfassend sind alle genannten Modelle auf die Grundlage von JUKES und CANTOR (1969) zurückzuführen. Mit Zunahme der berücksichtigten Parameter steigt allerdings die Leistungsanforderung an das Computersystem.

Alle dargestellten Vorgehensweisen gehen von bestimmten Annahmen über die Wahrscheinlichkeit einer Substitution aus. Dabei wird die Art des Nukleotids, seine relative Häufigkeit etc. beachtet. Alle Modelle gehen jedoch von einer für alle Nukleotidpositionen gleichen Mutationswahrscheinlichkeit aus. Berücksichtigt man hier die für das *eae*-Gen mehrfach beschriebene Hypervariabilität im 3'-Sequenzbereich wird deutlich, dass die beschriebene mathematische Vereinfachung zu einer Fehleinschätzung der genetischen Divergenz führen kann. In einem solchen Fall sind die jeweiligen funktionellen Regionen des Gens, wie in der vorliegenden Arbeit durchgeführt, abschnittsweise zu analysieren (ZHANG und GU, 1998). YANG (1994) erarbeiteten auf Basis des ML-Ansatzes ein Verfahren zur Kalkulation der Variation in Form der Gamma (γ)-Verteilung. Die Berechnung der γ -Verteilung wurde ebenfalls für die „Distanz-basierten“-Methoden übernommen (ZHANG und GU, 1998).

2.5 SEQUENZVERGLEICH

Grundlage der Darstellung phylogenetischer Relationen ist der gleichzeitige Vergleich (engl. = alignment“) mehrerer Sequenzen (engl. = „multiple sequence-alignment“). Hierfür erarbeiteten TAYLOR (1987) sowie FENG und DOOLITTLE (1987) einen heuristischen (griech. heuristein = finden, entdecken), progressiven Algorithmus. Dabei handelt es sich um ein Verfahren, dass die Berechnungszeit für große Datenmengen deutlich reduziert. Anhand eines paarweisen Sequenzvergleichs wird zunächst eine Datenmatrix mit den Abweichungswerten der einzelnen Sequenzpaare erstellt. Diese dienen als Basis für die Erstellung eines Führungsbaums (engl. „guided tree“). Die fortschreitende Alignierung folgt dem als Berechnungsgrundlage dienenden

Führungsbaum und wird daher als progressive Alignierung bezeichnet. Die heuristische Ausrichtung des Modells setzt dabei grundsätzlich voraus, dass alle betrachteten homologen Sequenzen in gegenseitiger evolutiver Beziehung stehen. Die daraus abgeleitete mathematische Kalkulation reduziert die erforderliche Rechenzeit, nimmt aber in Kauf, dass - im Gegensatz zur dynamischen Programmierung - das optimale Alignment nicht ermittelt werden kann. Der Algorithmus wurde von THOMPSON et al. (1994) modifiziert und in die weit verbreitete und frei verfügbare (engl. „freeware“) Anwendung „ClustalW“ implementiert. Die Modifikationen umfassen neben einer veränderten Gewichtung der einzelnen Sequenzen innerhalb des paarweisen Sequenzvergleichs zur Vermeidung falscher Ergebnisse durch die Überbewertung sehr ähnlicher Sequenzen auch die Kalkulation des Führungsbaumes. Die Berechnung erfolgt im Gegensatz zu anderen Ansätzen anhand der Neighbor-Joining-Methode (SAITOU und NEI, 1987), die eine bessere Bewertung ungleicher Evolutionsraten innerhalb der Sequenzen und eine exaktere Kalkulation der phylogenetischen Distanzen erlaubt. Ferner wurde die zusätzliche Einfügung von Lücken (engl. „gaps“) zur Kompensation unterschiedlicher Sequenzlängen oder Insertionen optimiert. Eine Lücke pro 20 betrachteter Positionen gilt für die Gewährleistung plausibler Daten als Faustregel. Die Parameter „Gap opening penalties“ (GOP) und „Gap extension penalty“ (GEP) dienen als Berechnungsgrundlage für den Einbau neuer Lücken bzw. die Verlängerung von Lücken während des Vergleichs zweier Sequenzen. Hohe GOP-Werte repräsentieren hohe Kosten für das Einfügen von Lücken und reduzieren die Frequenz der Lücken. Entsprechend führen hohe GEP-Werte zur Bevorzugung kurzer Lücken.

Die Programme BioEdit v. 7.03 (HALL, 1999), MEGA3 (KUMAR et al., 2004a) und MegAlign V. 5.0 (DNaStar, Fa. Lasergene) verwenden die ClustalW-Methodik. BioEdit und MEGA3 ermöglichen zusätzlich die Optimierung der Kalkulation über die Angabe der GEP- und GOP-Werte.

2.6 ABLEITUNG POPULATIONSGENETISCHER PARAMETER

2.6.1 CHARAKTERISIERUNG DES SEQUENZVERGLEICHS

Für die Ermittlung der Qualität und das Erzielen eines hohen Informationsgehaltes eines Alignments sind nach BARTON (2005) und BRINKMANN (2005) das Erreichen einer optimalen Alignmentqualität und die Auswahl eines geeigneten Algorithmus bei gleichzeitiger Verwendung einer hohen Anzahl repräsentativer, qualitativ hochwertiger Referenzsequenzen entscheidend. Die manuelle Editierung, d. h. insbesondere das Überprüfen und ggf. Korrigieren der durch die Software eingefügten Lücken sichert gegen die Überbewertung eng verwandter Sequenzen bzw. die Unterbewertung stärker divergierender Gene ab.

Zur Charakterisierung der Alignmentdaten gibt die Anzahl berücksichtigter Alignmentpositionen abzüglich der Lücken Auskunft über die generelle Aussagekraft des Alignments. Die Anzahl variierender Positionen „S“ und die Gesamtzahl der Mutationen „Eta“ werden nach ROZAS und ROZAS (1995) und ROZAS et al. (2003) für die abschnittsweise Charakterisierung funktioneller Bereiche und der Guanin-/Cytosin-Gehalt als Hinweis auf den möglichen evolutiven Ursprung im Vergleich zum Gesamtgenom eingesetzt. Sequenzpolymorphismen innerhalb des Alignments lassen sich hinsichtlich ihrer einheitlichen Identifizierung als „Singleton Polymorphic Sites (SPM)“ oder „Parsimony Informative Sites (PIPM)“ klassifizieren. Unter dem erstgenannten Begriff werden alle Positionen innerhalb der Alignierung zusammengefasst, die zu keiner Veränderung der Aminosäuresequenz führen. Letztere umfassen die in einer Änderung der Aminosäurekodierung resultierenden Substitutionen. Unterschieden werden Alignmentpositionen mit zwei, drei oder vier identischen Nukleotiden. Für eine Übersichtsdarstellung der ermittelten populationsgenetischen Parameter sei auf die **Tabelle 3.5.6-1:** Abgeleitete populationsgenetische Parameter verwiesen.

2.6.2 ERMITTLUNG DES SELEKTIONSMECHANISMUS

Die anhand der Differenzen „ d_S “ und d_N “ bzw. der Proportionen „ p_S “ und p_N “ ermittelten (**siehe Kap. 2.3.3**) Selektionsmechanismen unterteilen sich in:

Positive Selektion mit „ d_N “ signifikant größer *als* „ d_S “,“

Negative Selektion mit „ d_S “ signifikant größer *als* „ d_N “ und die

Neutrale Selektion mit „ d_N “ = „ d_S “.

Die statistische Absicherung erfolgt üblicherweise über Z-Tests (FISHER, 1973). Eine detailliertere Bewertung im Hinblick auf den Einfluss der Veränderung innerhalb des Proteins kann anhand der Parameter Ladung und Polarität vorgenommen werden (HUGHES et al., 1990). So können Aussagen über die funktionelle Divergenz von Proteinen getroffen werden (KNUDSEN et al., 2003). So untersuchten ENDO et al. (1996) mehr als 3500 homologe Proteinsequenzen auf positive Selektion. HUGHES und FRIEDMAN (2005) lieferten grundlegende Arbeiten anhand von 4133 proteinkodierenden Sequenzen von *Saccharomyces cerevisiae* und *Saccharomyces paradoxus*. YANG und NIELSEN (2000) erweiterten die hier verwendeten Berechnungsmodelle um einen ML-Ansatz, den sie an mitochondrialen Genen validierten.

2.6.3 GENETISCHE REKOMBINATION

Unter **Rekombination** wird die Bildung neuer Genkombinationen aus genetisch verschiedenen Genomen verstanden. Die Integration der neu geordneten Anteile kann dabei an spezifischen Stellen, in anderen Situationen zufällig erfolgen.

Mit Blick auf das Gesamtgenom der Bakterien können spezifische *rec* (engl. „recombinant“)-Gene identifiziert werden, die den Rekombinationsprozess steuern. Nach POSADA und CRANDALL (2001a, 2001b) beeinflusst die Rekombination die Evolution auf verschiedenen Ebenen, die insbesondere zu einer signifikanten Beeinflussung humanpathogener Mikroorganismen führen:

- Beeinflussung der genetischen Selektion
- Formung der Architektur von Genomen
- Unterbrechung des Verbindungs-Ungleichgewichtes (engl. „linkage disequilibrium“)

Auf phylogenetischer Ebene unterscheidet WHITTAM (1995) drei Rekombinationsereignisse:

1. Die **assortative Rekombination**, bei der bereits existierende Allele neu kombiniert werden und neue chromosomale Genotypen entstehen.
2. Die **intragenetische Rekombination**, die den Austausch kleiner Genbereiche beschreibt, der Mosaik-Allele bzw. -Gene hervorruft und zu neuen Genotypen führt.
3. Die **additive Rekombination**, bei der es sich um die Integration neuer, bislang nicht vorhandener Gene handelt. Damit entsteht ein zusammengesetzter Genotyp, der die Summe der rekombinierten Merkmale aufweist.

Bezogen auf die Betrachtung eines einzelnen Gens oder Merkmals können hohe Rekombinationsraten zu Problemen bei der Ermittlung des tatsächlichen Verwandtschaftsgrades führen. Das bedeutet, dass die Darstellung der einzig möglichen Entwicklung fehlerhaft oder ungenau werden kann. Für den Anwender bestand bislang die Schwierigkeit, aus einer Vielzahl von statistischen Methoden das geeignete Verfahren auszuwählen. Allgemein lassen sich die Algorithmen wie unten aufgeführt einordnen. Zusätzliche Unterschiede bestehen bei der Berechnung der statistischen Signifikanz der Ergebnisse:

1. Distanz-basierte Methoden:

Die auch als Distanz-Methoden beschriebenen Algorithmen führen eine abschnittsweise Untersuchung auf Inversionen der genetischen Distanzen durch. Gekoppelt an die statistische Auswertung, kann auf die Berechnung des kompletten phylogenetischen Stammbaumes verzichtet werden. Sie finden selten Verwendung (MARTIN und RYBICKI, 2000; MARTIN et al., 2005).

2. Phylogenetische Verfahren

Dabei handelt es sich um die am häufigsten eingesetzten Modelle. Von einem Rekombinationsereignis wird dabei in Abhängigkeit vom Auftreten möglicher diskordanter phylogenetischer Bäume ausgegangen. Die beste Übereinstimmung mit der tatsächlichen Phylogenie kann dabei z. B. in Form eines „Bootstrap-Baumes“ erreicht werden. Hierzu zählen das eingesetzte „Bootscanning-Verfahren“ (SALMINEN et al., 1995) und das „Recombination Detection Programme (RDP)“ (MARTIN und RYBICKI, 2000; MARTIN et al., 2005)

3. Kompatibilitäts-Methoden

Diese Algorithmen untersuchen die teilweise phylogenetische Inkongruenz pro Position innerhalb des Alignments. Eine komplexe Kalkulation des phylogenetischen Baumes wird nicht benötigt.

4. Methoden, die nach Zonierungen von Nukleotidsubstitutionen suchen

Die alignierten Sequenzen werden nach signifikanten Clustern der Substitutionen durchsucht. Alternativ kann auch die Übereinstimmung der Veränderungen mit einem statistischen Modell überprüft werden. Die Algorithmen „Geneconv“ (PADIDAM et al., 1999), der „Maximum Chi-Quadrat-Test“ (SMITH, 1992; SMITH, 1999) und der „Maximum Mismatch Chi-Quadrat-Test“ (Chimaera, POSADA und CRANDALL, 2001b) arbeiten nach diesem Prinzip.

Das Leistungsspektrum der eingesetzten Algorithmen unterscheidet sich nach der Möglichkeit, ausschließlich mögliche rekombinante Genabschnitte zu identifizieren bzw. die Genauigkeit der Analyse durch die Festlegung von Anfangs- und Endpunkten (engl. „breakpoints“) und gesamten Bereichen bzw. mögliche evolutive Vorläufersequenzen (engl. „parent-daughter-identification“) zu erhöhen. Die nachfolgende Übersicht (**Tabelle 2.6.3-1**) charakterisiert die Verfahren gemäß des nach MARTIN und RYBICKI (2000) ermittelten Leistungsvermögens.

Tabelle 2.6.3-1: Charakteristika verschiedener Programme zur Überprüfung der genetischen Rekombination nach MARTIN und RYBICKI (2000)

Methode	Identifikation von Rekombinationen	Ermittlung von Endpunkten	Ermittlung rekombinanter Regionen	P-Wert Kalkulation	Referenz
Originale RDP-Methode	+	+	+	Binominalverteilung	MARTIN und RYBICKI, 2000
Geneconv	+	+	+	Permutation	PADIDAM et al., 1999
Bootscan	+	+	+	Boostrapping, Binominalverteilung, Chi-Quadrat-Test	SALMINEN et al., 1995
Maximum Chi-Quadrat	+	+	+/-	Chi-Quadrat-Test und Permutation	SMITH, 1992, 1999
Chimaera	+	+	+	Chi-Quadrat-Test und Permutation	POSADA und CRANDALL, 2001b
Sister Scan	+	+	+	Permutation und Z-Test	GIBBS et al., 2000

+ möglich,

+/- nur indirekt über Bruchpunkt-Ermittlung möglich

2.7 PHYLOGENETISCHE BÄUME

2.7.1 DISTANZ- UND CHARAKTER-BASIERTE VERFAHREN

Es werden zahlreiche Methoden zur Wiedergabe der phylogenetischen Relation in Form von Bäumen oder Phylogrammen im Internet zur Verfügung gestellt und in der Literatur diskutiert (SWOFFORD et al.; 1996; LI, 1997). Grundsätzlich lassen sich die Verfahren anhand der verwendeten Datentypen und der Rekonstruktionsmethode gemäß der **Tabelle 2.7.1-1** einteilen. Ein Gruppierender (engl. „clustering“)-Algorithmus“ konstruiert auf Grundlage eines Optimalkriteriums sukzessive einen Baum aus den Daten. Eine Suchstrategie ermittelt viele, unter Umständen alle möglichen Bäume und sucht anhand eines Optimalitätskriteriums den oder die optimalen Darstellung(en) heraus.

Tabelle 2.7.1-1: Einteilung der Rekonstruktionsmethoden zur phylogenetischen Analyse

Rekonstruktions- methode	Datentyp	
	Distanzen	Charakter
Gruppierender- Algorithmus	UPGMA	
	Neighbor Joining	
Suchstrategie	Minimum	Maximum Parsimony
	Evolution	Maximum Likelihood

Distanz-basierte Methoden/Distanz-Methoden:

Sie verwenden die Summe der Unterschiede, d. h. die Distanz zwischen zwei alignierten Sequenzen zur Ableitung des Baumes. Dabei wird ein additiver Baum berechnet. Der Distanzwert zweier Sequenzen setzt sich aus der Summe der Astlänge zweier Taxa zusammen (SWOFFORD et al. 1996).

Ein wesentlicher Vorteil der Distanz-basierten Methoden liegt in der geringeren benötigten Rechenleistung. Ihr wesentlicher Nachteil besteht zunächst in der Annahme einer konstanten Evolution. Erfolgt die Sequenzvariation nicht über die gesamte Zeit konstant oder werden bestimmte Mutationen bevorzugt, müssen diese Verhältnisänderungen durch eine geeignete Datenmatrix korrigiert werden.

Zu den einfachsten Methoden zählt UPGMA (engl. „unweighted pair group method using arithmetic averages“) nach SOKAL und MICHENER (1958) sowie SNEATH und SOKAL (1973). Exakte Ergebnisse können dabei nur bei konstanter Evolutionsrate erwartet werden. Die Korrektur für diese Fehlerquelle ist mangelhaft (TAKEZAKI und NEI, 1996).

Der ME- (engl. „minimum evolution“-)Ansatz (EDWARDS und CAVALLI-SFORZA, 1963) verfolgt die „Such-Strategie“ und ermittelt alle möglichen Lagebeziehungen. Als optimaler Baum wird das Ergebnis mit der kleinsten Summe der Astlängen ausgegeben. Dieses Vorgehen ist sehr aufwendig (RZHETSKY und NEI, 1992a und 1992b; SWOFFORD et al., 1996; FELSENSTEIN, 1997). Dabei werden nicht alle Distanzwerte einer Datenmatrix berechnet, sondern interne und externe Knoten definiert. Ihre korrekte Wiedergabe wird anschließend durch die Analyse der Astlänge zwischen beiden Punkten überprüft.

Das 1987 von SAITOU und NEI vorgestellte NJ- (engl. „neighbor joining“-Verfahren) stellt eine Vereinfachung des ME-Ansatzes dar. Dabei wird zunächst ein sternförmiger Baum (engl. „star tree“) erstellt.

Die Gruppierung der eng verwandten Sequenzen zu sog. Nachbarn erfolgt dabei schrittweise. Vereinfacht dargestellt, wird die benötigte Rechenleistung durch Zusammenfassung verschiedener Nachbarn bis zum nächsten Knoten im Baum erzielt. Der so reduzierte Knoten kann als ein Nachbar für weitere Berechnungen betrachtet werden.

Basierend auf diesen grundlegenden Modellen existieren zahlreiche Weiterentwicklungen oder Modifikationen. Ihr wesentlicher Bestandteil ist dabei die Korrektur für ungleiche Mutationsraten, bedingt durch Rekombination oder Selektion.

Das von TAMURA und NEI (1993) vorgestellte Modell beinhaltet neben der Korrektur ungleicher Substitutionsraten zusätzlich die Optimierung verschiedener Transitionsraten der Purine und Transversionenraten der Pyrimidine. Die Korrektur unterschiedlicher Evolutionsraten der einzelnen Alignmentpositionen kann über den gamma-Parameter erfolgen (ZHANG und GU, 1998). Das modifizierte Verfahren nach NEI und GOJOBORI (1986) und ZHANG et al. (1998) korrigiert die im ursprünglichen Modell enthaltene Überbewertung von p_S im Gegensatz zu p_N . Die Optimierung erfolgt über Einbeziehung des Parameters „R“, der das Verhältnis von Transitionen zu Transversionen wiedergibt.

Charakter-basierte Methoden (engl. „character state“) analysieren das Sequenzalignment und bewerten neben der Anzahl auch die Art der Veränderungen. Die Berechnungsgrundlagen sind aufwendig und zeitintensiv. Im Idealfall werden alle möglichen Bäume konstruiert und anschließend überprüft. MP (engl. „maximum parsimony“) verfolgt das Prinzip der maximalen Sparsamkeit (DAYHOFF et al., 1978). Abschließend wird der Evolutionspfad ausgewählt, der die geringsten Veränderungen beinhaltet. Für die Kalkulation werden nur die variablen und gleichzeitig informativen Positionen des Alignments, d. h. mindestens zwei verschiedene Nukleotide, die mindestens zweimal im Alignment vorhanden sind, herangezogen (SWOFFORD et al., 1996).

Ihr wesentlicher Schwachpunkt liegt in der falschen Wiedergabe paralleler Substitutionen (sog. Homoplasien, FELSENSTEIN, 1978).

ML-gestützte Analysen suchen nach dem optimalen evolutiven Modell zur Erklärung eines Stammbaumes (FELSENSTEIN, 1981). Anschließend wird der Baum, der mit der höchsten Wahrscheinlichkeit die Daten widerspiegelt, ausgewählt. Der Algorithmus wurde in verschiedene Programme implementiert, so z. B. „PAUP“ (FELSENSTEIN, 1997), „PAML“ (YANG, 1997a) oder „HyPhy“ (KOSAKOVSKY-POND et al., 2005). Die Unterschiede liegen in Abwandlungen der Suchstrategie und dem benötigten Vorwissen des Anwenders. Neben der Bereitstellung eines exakten Alignments wird dabei gleichzeitig ein bereits erstellter Baum zur Überprüfung der Wahrscheinlichkeit anhand eines vorgegebenen Modells der Sequenzevolution (**Kap. 2.4.3**) konstruiert.

2.7.2 SPLIT-ZERLEGUNG

Evolutive Daten beinhalten häufig eine große Anzahl mitunter die Auswertung erschwerende Signale. Die Darstellung eines einheitlichen Baumes wird dabei durch Homoplasien bzw. genetische Rekombination nicht unterstützt. Zur Darstellung des Einflusses dieser Faktoren entwickelten BANDELT und DRESS (1992) die **Split-Zerlegung** (engl. „split decomposition“). Unter einem „Split“ (engl. „split“ = Aufteilung) wird die Aufteilung zweier Taxa in zwei nicht-leere Untergruppen verstanden (HUSON und BRYANT, 2006). Im Gegensatz zu den oben genannten „MP“- oder „ML“-Verfahren, die den optimalen Baum durch Vergleich verschiedener Parameter bestimmen, verwendet die Split-Zerlegung einen transformatorischen Ansatz. Mathematisch erfolgt eine kanonische Zerlegung in eine Summe „schwach kompatibler Splits“. Diese werden anschließend als „Split-Graph“ wiedergegeben.

Hier entspricht jede Linie bzw. jeder Ast oder jede Kante einem Split, d. h., dass Anfangs- und Endpunkt der Linie die durch den „Split“ getrennten Taxa darstellen. Für ideale Daten ergibt sich ein Baum. Alternative Phylogenien werden als netzwerkartige Strukturen wiedergegeben (FITCH, 1997).

Unterschiedliches Evolutionsgeschehen wird durch parallele Äste verdeutlicht. Die Darstellung entspricht den bereits bekannten Kladogrammen, d. h., dass die Astlängen zwischen den Taxa der ermittelten phylogenetischen Distanz entsprechen. Der hier verwendete Algorithmus nach BRYANT und MOULTON (2004) arbeitet nach dem Distanz-basierten Prinzip, die die häufiger verwendete zweite Klasse der Split-Algorithmen umfasst. Eine exemplarische Darstellung mit Erläuterung der Nomenklatur ist der **Abbildung 2.7.2-1** zu entnehmen. Die ebenfalls verwendete Berechnung nach BANDELT und DRESS (1992) gehört zur ersten Klasse der weniger häufig eingesetzten Charakter-basierten Methoden (POSADA und CRANDALL, 2001a; BRYANT und MOULTON, 2004).

Die Split-Verfahren finden breite Verwendung in der bakteriellen und viralen Phylogenetik. So setzten KOTETISHVILI et al. (2002) das Verfahren zur Analytik des „Multilocus Sequene Typing (MLST)“ bei Salmonellen ein, REVAZISHVILI et al. (2004) verglichen die Pulsfeldgelektrophorese- (PFGE-) und MLTS-Profile von *Listeria monocytogenes*. Exemplarisch sei auf die Untersuchungen humaner Lentiviren von SALEMI et al. (2003) verwiesen.

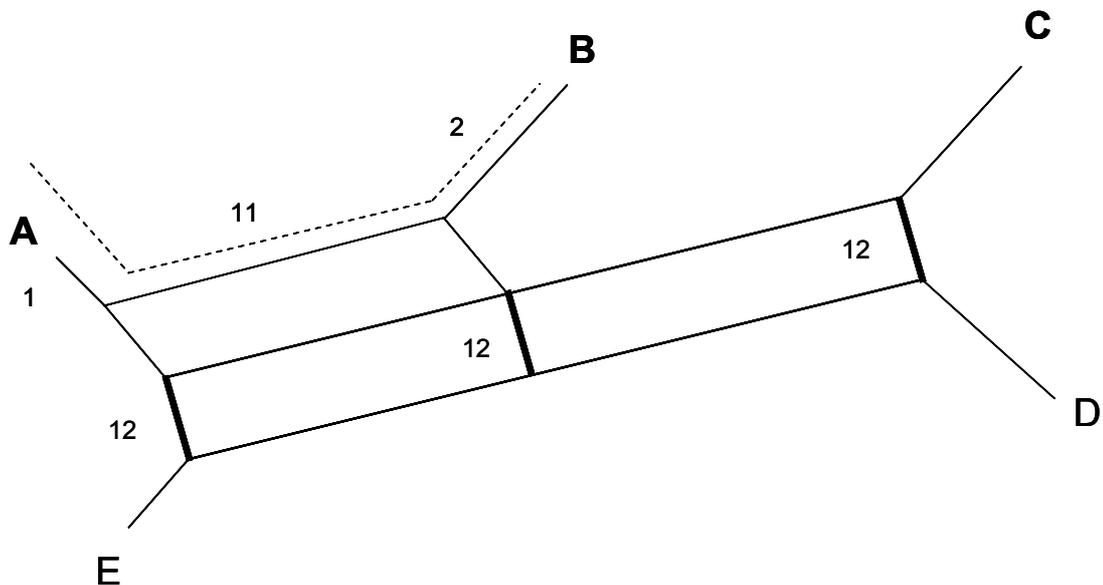


Abbildung 2.7.2-1: Schematische Darstellung der Split-Zerlegung nach HUSON und BRYANT (2006)

Legende:

----- Phylogenetische Distanz zwischen den Taxa „A“ und „B“

———— Parallele Kanten, die zu einem Split gehören

Werte neben den Kanten: Gewichtung der Splits; dies erfolgte für die eigenen Untersuchungen durch Angabe der Bootstrap-Werte [%].

Fettgedruckte Taxa: Gruppe der Taxa, die durch den Split mit der Gewichtung „12“ (3 Kanten) von den anderen getrennt werden

2.8 BIOINFORMATIK

Eine Übersicht der eingesetzten Computerprogramme ist der **Tabelle 3.5.1-1** unter Angabe der verwendeten Optionen, der Bezugsquelle bzw. Internetadresse und der Referenz zu entnehmen. Dabei wurde auf Anwendungen, die direkt über das Internet (engl. „online-applications“) und als eigenständige, auf lokalen Computersystemen installierbare Programme (engl. „stand alone-applications“) zurückgegriffen. Bei den über das Internet verwendbaren Anwendungen, wie z. B. BANKIT, BLAST, NUCLEOTIDE, STRUCTURE und LOCUS LINK, handelt es sich um Bestandteile ständig weiter entwickelter Gendatenbanken des National Center for Biotechnology Information (NCBI), Bethesda, USA (www.ncbi.nlm.nih.gov). Bei Comprehensive Microbial Resource (CMR) wird unter dem Kürzel TIGR (www.tigr.org/CMR2) eine weitere kostenlose Genomdatenbank mit frühzeitigem Zugriff auf Genomsequenzierungsdaten etabliert (PETERSON et al., 2001). MEGA 3 (KUMAR et al., 2004a) und BioEdit V7 (HALL, 1999) sind Programmpakete, die gegenüber dem kommerziell verfügbaren Programmpaket DNASTAR (Inc. Wisconsin, USA) die Verwendung unterschiedlicher Dateiformate ermöglichen. Im erstgenannten Fall liegt ihr Schwerpunkt dabei auf phylogenetischen Anwendungen. YANG (1997a) entwickelte mit PAML 3.14 eine Sammlung verschiedener Programmdateien. Eine Bewertung des praktischen Nutzwertes, insbesondere der graphischen Anwendung Splits Tree (HUSON, 1998) und der Berechnungsprogramme RDP V. 2.0 (MARTIN und RYBICKI, 2000), HON-NEW (ZHANG, 2000) und SC3 (HUGHES et al., 1990 und HUGHES und FRIEDMANN (2000) ist der Diskussion (**Kap. 5.3.5, Kap. 5.3.6 und Kap. 5.4**) zu entnehmen.

3 EIGENE UNTERSUCHUNGEN

Bedeutung und Funktion des *E. coli* attaching and effacing (*eae*)-Gens und seines Translationsproduktes Intimin für den spezifischen Anheftungsmechanismus der Mikroorganismen und das Vermögen zur Auslösung von Erkrankungen bei Mensch und Tier sind hinreichend beschrieben worden. Dabei konnte gezeigt werden, dass das Intimin über konservierte und hypervariable Regionen, letztere lokalisiert in der für die zelluläre Interaktion verantwortlichen Region des Proteins, verfügt. Neben der Bedeutung der daraus resultierenden Proteinvariationen für die Lokalisation von *E. coli* in unterschiedlichen Darmabschnitten, sind die phylogenetische Zuordnung der Sequenzheterologien mit besonderer Berücksichtigung der Herkunft (Habitat) einbezogener Stämme Bestandteil aktueller Untersuchungen. Die genotypische Typisierung erfolgt mittels der Polymerasekettenreaktion.

3.1 Gliederung der Arbeiten

3.1.1 VORVERSUCHE

Ziel der Vorversuche war die Etablierung eines konventionellen PCR-Systems zur Typisierung des Intimins bei *E. coli*-Stämmen. Dazu wurde zunächst die Eignung der in den Arbeiten von ADU-BOBIE et al. (1998a) sowie OSWALD et al. (2000) beschriebenen Oligonukleotid-Primer-Systeme überprüft. Der vorgeschaltete Nachweis des *eae*-Gens erfolgte mittels der Primer-Kombinationen „SK1“ und „SK2“ nach KARCH et al. (1993). Eine Übersicht über die verwendeten Oligonukleotid-Primersequenzen, -positionen und Amplifikatgrößen wird im Folgenden gegeben. Abschließend wurden mit dem als routinetauglich klassifizierten PCR-System 100 *E. coli*-Stämme verschiedener Serovaren und Habitate typisiert. Bei der Auswahl der Stämme wurde auf *E. coli*-Isolate mit aus der Literatur als bekannt vorausgesetzten

Intiminvarianten zurückgegriffen, um falsch-positive Reaktionen ausschließen zu können.

3.1.2 HAUPTVERSUCHE

Die Hauptversuche umfassten fünf Teilbereiche. Sie sind im Folgenden vorab charakterisiert:

Teilbereich 1: Intimintypisierung von *E. coli*-Stämmen

Basierend auf dem optimierten PCR-Nachweis nach ADU-BOBIE et al. (1998a) wurden in der vorliegenden Arbeit 317 *E. coli*-Stämme untersucht. Die Isolate sind Bestandteil der institutseigenen Stammsammlung. 157 *E. coli*-Serovare wurden bereits im Rahmen des Projektes „Entwicklung quantitativer VTEC/STEC-Bestimmungsmethoden für Lebensmittel“ (Az.: 422-7030-56/48), gefördert durch das damalige Bundesministerium für Gesundheit und Soziale Sicherung, hinsichtlich der Virulenzfaktoren Shigatoxin und Enterohämolysin charakterisiert. Diese Daten wurden zur abschließenden Erstellung des Virulenzprofils herangezogen.

Die Überprüfung der erhaltenen PCR-Amplifikate wurde mittels Restriktionsfragmentlängenpolymorphismus (RFLP) vorgenommen. Mit den verwendeten Restriktionsendonukleasen *Taq* I (Produkt-Nr. 404128, Fa. Roche Diagnostics) und *Fok* I (Produkt-Nr. 1 004 816, Fa. Roche Diagnostics) konnte gleichzeitig das Vorhandensein anderer, bislang nicht bekannter Intimintypen bzw. möglicher Intiminuntereinheiten überprüft werden.

Teilbereich 2: Sequenzierung des *eae*-Gens zweier *E. coli*-Stämme

Die verotoxinbildenden Isolate P 146 O156:H⁻ und P 256 O86:H⁻ waren mit dem verwendeten PCR-System keiner der fünf Intiminvarianten zuzuordnen. Über schrittweise Amplifikation und Sequenzierung wurde der Intimintyp der beiden Stämme ermittelt.

Teilbereich 3: Phylogenetische Analyse

Die phylogenetische Relation der beschriebenen Intiminvarianten bildete die Grundlage für eine Klassifikation der verschiedenen Genausprägungen als eigenständige Untereinheiten oder Subtypen. Die angeschlossene Strukturanalyse charakterisierte konservierte und hypervariable Sequenzabschnitte und ermöglichte eine Bewertung der für die genetische Rekombination relevanten Selektionsmechanismen.

Für die Sequenzanalysen und die Erstellung phylogenetischer Korrelationen sind mittlerweile eine Vielzahl unterschiedlicher Computerprogramme entwickelt worden. Eine Alternative zu den aufgrund der hohen Anschaffungs- und Unterhaltungskosten nicht flächendeckend verbreiteten Softwarevarianten stellen die im Internet frei verfügbaren und in der Regel kostenlos angebotenen Lösungen dar. Um dem Anwender einen Überblick über praktikable Vorgehensweisen zu verschaffen, wurde die Datenbearbeitung exemplarisch mittels alternativer Anwendungen dargestellt.

Teilbereich 4: Etablierung eines universellen PCR-RFLP-Intimintypisierungsverfahrens

Die eigenen Untersuchungen und der sich fortlaufend erweiternde Kenntnisstand über neue Intiminvarianten bzw. -subtypen zeigten, dass ein leistungsfähiges Typisierungsverfahren eine Vielzahl unterschiedlicher *eae*-

Gensequenzen erfassen und gleichzeitig Potenzial zum einfachen Nachweis weiterer Variationen beinhalten sollte.

Basierend auf den von RAMACHANDRAN et al. (2003) bereitgestellten Primersystemen wurde durch restriktionsenzymatischen Verdau mittels der Restriktionsendonukleasen *Fok* I, *Rsa* I und *Hae* III ein universelles Typisierungsverfahren zur Erfassung von alpha (α), beta (β), gamma (γ), delta (δ), epsilon (ϵ), zeta (ζ), eta (η), theta (θ), iota (ι), kappa (κ), lambda (λ), my (μ), ny (ν), pi (π), rho (ρ) und sigma (σ) etabliert. Da für die xi (ξ)- und omikron (\omicron)-Ausprägungen zum Zeitpunkt der Arbeiten nur partielle Sequenzen vorlagen, konnten diese nicht in die Analyse zur Erstellung eines RFLP-Profiles einbezogen werden. Mit dem beschriebenen System wurden n=27 mit der oben beschriebenen Methode nicht typisierbare *E. coli*-Serovare untersucht.

Teilbereich 5: Darstellung der Zell-Zell-Interaktion

Um die unterschiedlichen Mechanismen der zellulären Interaktion von *E. coli* zu veranschaulichen, wurden Zelladhäsionsversuche mit dem EPEC-Referenzstamm E 2348/69 O127:H6 durchgeführt. Die unspezifische, Fimbrien-vermittelte Adhäsion konnte in Kombination mit humanen Mundschleimhautzellen demonstriert werden. Zur Darstellung der spezifischen, Intimin-bedingten lokalisierten Adhäsion wurde der „Fluorescent-Actin Staining“ (FAS)-Test unter Verwendung von HEp-2-Zellen eingesetzt.

Anmerkung: Letztgenannter Test konnte durch die freundliche Unterstützung Herrn Prof. Rolf Bauerfeinds in den Räumen und Einrichtungen des Institutes für Hygiene und Infektionskrankheiten der Tiere der Justus-Liebig-Universität Gießen durchgeführt werden. Herrn Prof. Rolf Bauerfeind, Herrn PD Dr. Christian Menge und Herrn Dr. Elmar Schröpfer sei an dieser Stelle für die praktische Unterstützung und Betreuung gedankt.

3.2 Material und Methoden

3.2.1 EINGESETZTE REFERENZSTÄMME

Die zur Etablierung der PCR-Systeme verwendeten Referenzstämme sind in der **Tabelle 3.2.1-1** unter Angabe der Bezugsquelle und der Genbankzugriffsnummer wiedergegeben. Dabei handelte es sich um in internationalen Arbeiten charakterisierte Isolate mit nahezu vollständig klassifizierten Bakteriengenomen und Virulenzprofilen.

Die Isolate O103 PMK3 und EHEC O26 8489/67 wurden freundlicherweise von Prof. Helge Karch, Institut für Hygiene der Westfälischen Wilhelms-Universität in Münster, zur Verfügung gestellt

In der **Tabelle 9.1.1-1** sind die zusätzlich zur Datenbankabfrage über NCBI®-BLAST für die Sequenzanalyse der verwendeten Primer eingesetzten Referenzstämme zusammengefasst.

Tabelle 3.2.1-1: Etablierung der PCR-Systeme nach ADU-BOBIE et al. (1998) und OSWALD et al. (2000) unter Berücksichtigung verwendeter Positiv- und Negativkontrollstämmen unter Angabe der Bezugsquelle und der Genbankzugriffsnummer

Intimintyp*	Eingesetzter Referenzstamm	Virulenzprofil**	Genbankzugriffs-Nr.***	Herkunft/Referenz
(alpha) α -Intimin	EPEC O127:H6 E 2348/69	<i>eae</i> , <i>vtx</i> -	M58154	BÜLTE, Institut für Tierärztliche Nahrungsmittelkunde, Gießen
(beta) β -Intimin	EHEC O26:H7 8489/67	<i>eae</i> , <i>vtx</i> -	AF200363 für REDC-1 (O15:H7)	KARCH, Institut für Hygiene, Münster
(gamma) γ -Intimin	EHEC O157:H7 EDL 933	<i>eae</i> , <i>vtx</i> 1, <i>vtx</i> 2, 2c, <i>hly</i>	Z11541.1	BÜLTE, Institut für Tierärztliche Nahrungsmittelkunde, Gießen
(delta) δ -Intimin	Nicht vorhanden	Keine Angaben	Y13112 für ICC 95/O86:H34	ADU-BOBIE et al. (1998a)
(epsilon) ϵ -Intimin	O103 PMK3	<i>eae</i> , <i>vtx</i> 1	AF 116899 für O103:H2 PMK5	KARCH, Institut für Hygiene, Münster
Negativkontrolle	<i>E. coli</i> ATCC 25922	<i>eae</i> -, <i>vtx</i> -		BÜLTE, Institut für Tierärztliche Nahrungsmittelkunde, Gießen

- * Für delta-Intimin stand kein Referenzstamm zur Verfügung
 ** Hausintern bestimmt, die Referenzen der PCR-Verfahren werden im Text erläutert
 *** Die Genbankzugriffs.-Nr. bezieht sich auf den eingesetzten Prüfstamm bzw. auf ein entsprechendes Serovar mit identischem Intimintyp

3.2.2 UNTERSUCHTE PRÜF-/WILDSTÄMME

In die Untersuchungen wurden insgesamt 317 Prüfstämme einbezogen. Darin waren insgesamt 117 (37,2%) Isolate aus humanen Stuhlproben, 114 (36,0%) Isolate aus Lebensmitteln, 64 (20,2%) und zehn (3,2%) aus Rinder- bzw. Schafkot, vier (1,3%) aus Schweinekot bzw. Sektionsmaterial der Tierart Schwein gewonnene Serovare und ein (0,3%) Isolat aus Pferdekot Isolat enthalten. Bei sechs (1,8%) der Stämme lagen keine Angaben zur Ausgangsmatrix vor. In der **Tabelle 3.2.2-1** wurde die Zusammensetzung der Prüfstämme in Abhängigkeit von der Ausgangsmatrix zugeordnet.

Die Serotypisierung war bereits zuvor durch das „Nationale veterinärmedizinische Referenzlaboratorium für *Escherichia coli*“ am Bundesinstitut für Risikobewertung (BfR), Standort Dessau erfolgt.

Tabelle 3.2.2-1: Zusammensetzung der untersuchten Prüfstämme (n=317 [%])

	Anzahl und Herkunft einbezogener Prüfstämme (n [%])						
	humane Stuhlproben	Lebensmittel	Rinderkot	Schafkot	Schweinekot/-sektionsmaterial	Pferdekot	Keine Angabe
Gesamtzahl 317	118 (37,2)	114 (36,0)	64 (20,2)	10 (3,2)	4 (1,3)	1 (0,3)	6 (1,8)

3.2.3 MATERIAL UND LABORGERÄTE

a) Allgemeine Laborausstattung

- Auflichtmikroskop (UV-Licht), Fa. ZEISS, Modell B552 mit Leica-Aufsatzkamera
- Auslaufmesspipetten nach DIN 12696, Genauigkeit Klasse B und AS
- Brutschrank (37°C, Fa. Memmert)
- Bunsenbrenner (Gasi, Typ 3.340 102, Fa. Schütt Labortechnik)
- Einwegpetrischalen gemäß DIN 58940
- Feinwaage (Sartorius Typ E 5500 S, Fa. Sartorius)
- Gefriertruhe (-25°C)
- Kulturröhrchen nach DIN 12395 mit Verschlusskappe aus Aluminium
- Kühlschrank (+2 bis +4°C)
- Kühlzentrifuge (Typ 5415C, Fa. Eppendorf)
- Magnetrührer „RCT-Basic“, Fa. KIKA Labordiagnostic
- Messzylinder aus Glas nach DIN 12680, Genauigkeit Klasse B
- Messkolben aus Glas nach DIN 12664, Genauigkeit Klasse A
- Ständer für Kulturröhrchen
- MICROBANK™ (Fa. Mast Diagnostica, Produkt-Nr. 291601)
- Platinösen
- Pipetten (Typ: Research®, Varipip®, Referenz der Fa. Eppendorf 100-1.000 µl, 50-200µl, 10-100µl, 0,5-10µl und 0,1-2,5 µl)
- Pipettierhilfen (Peleusball)
- Pipettenspitzen (Fa. Eppendorf, Volumina 10µl, 200µl und 1.000µl)
- Pipettenspitzen mit Filter (Fa. Nerbe Plus, Volumen 10µl)

- Reagenzglas-Schüttelgerät (IKA-VIBRO-FIX VF2, Fa. Ika Labortechnik)
- Schüttelwasserbad (Typ 1083, Fa. Gfl)
- Trocknungsschrank (40°C, Fa. Memmert)
- Vortexer/Minishaker (Typ: MS1, Ika-Werke)
- Zentrifuge (Tischzentrifuge, Typ 5403, Fa. Eppendorf)
- Zentrifugenröhrchen (50 ml) mit Schraubdeckel (Fa. Roth, Produkt-Nr. N465.1)

b) Nährmedien und Supplemente

- Brain Heart Infusion (BHI) broth, Hirn-Herz-Bouillon (Fa. Merck, Produkt-Nr. 1.10493)
- Plate-Count-(PC)-Agar, Keimzählagar (Fa. Merck, Produkt-Nr. 1.05463)
- Dulbecco´s modified Eagle Medium (Fa. GIBCO, Produkt-Nr. 31855-023)
- Hank´s Salt Solution 1 x (Fa. Biochem KG, Produkt-Nr. L2045)

c) Molekularbiologischer PCR-Bedarf

- DNeasy[®] Tissue-Kit (50 Reaktionen, Produkt-Nr. 69504, Fa. Qiagen)
- Einmalhandschuhe TouchNTuff-Nitril (Fa. Ansell, Produkt-Nr. 92-60)
- Einmal-Filterhalter, 0,2 µm hydrophob, steril, pyrogenfrei (Fa. Schleicher&Schüll, Produkt- Nr. FP030/8)
- Einmal-Injektionsspritzen (30 ml) nach DIN12009, BS 5081, ISO 7886 (Fa. Braun, Produkt-Nr. 304-M. E.)
- Kühlboxen (Iso-Therm[®]-System, Fa. Eppendorf)
- PCR-Reaktionsgefäßständer (Fa. Eppendorf)
- Reaktionsgefäßständer (Fa. Eppendorf)
- UV-Kabinett (825-UCV/22, Fa. Plas-Labs)

- Reagenzgefäße (Multi SafeSeal Tubes, 1,7 ml, Fa. Peqlab)
- Reagenzgefäße (Safe-Seal Reagiergef. 2 ml, Fa. Sarstedt)
- Reagenzgefäße für PCR (100µl, 200µl)
- Sicherheitswerkbank (LaminAir HB 2448, Fa. Heraeus Instruments)
- Netzteil MBT (Typ: 57042, Fa. MicroBiotecBrand)
- Netzteil Transphor/Electrophoresis DC Power Supply (Fa. HSI Hoefer, Produkt-Nr. 95-2021)
- Elektrophoresekammer (Typ Blue Manne 100 und 200, Fa. Blueline)
- Thermocycler (TouchDown™ Temperature Cycling System, Fa. Hybaid)
- Thermocycler (GeneAmp PCR-System 9600, Fa. Perkin Elmer)
- Thermocycler (Biometra Trio-Thermoblock mit Heizdeckeleinheit (Fa. Biometra, Produkt.-Nr. 050-000)
- Thermomixer (Typ 5436, Fa.Eppendorf)
- Fotoanlage (MP 4*, Fa Polaroid)
- Ultraviolet Transilluminator (Mighty Bright UVTM-25, Fa. Hoefer Scientific Instruments)
- Ultraviolet Transilluminator (UVC 500 Crosslinker , Fa. Hoefer Scientific Instruments)
- Temperierbares Wasserbad (Fa. GFL)
- Temperierbares Wasserbad (Fa. Köttermann)

3.2.4 CHEMIKALIEN UND REAGENZIEN

- Agarose Standard (Fa. Q Bio gene, Produkt-Nr. AGAH0500)
- AmpliTaq® DNA Polymerase mit GeneAmp® 10x Puffer (15 mM MgCl₂, Fa. Applied Biosystems, Produkt-Nr. N808-0152)
- AmpliTaq Gold® DNA Polymerase mit GeneAmp® 6x Puffer (15 mM MgCl₂, Fa. Applied Biosystems, Produkt-Nr. 4311814)
- Platinum® Taq Polymerase (Fa. Invitrogen, Produkt-Nr. 10966-026)
- Borsäure /Fa. Merck, Produkt-Nr. 1.00165)
- Bromphenolblau-Lösung (Produkt-Nr. 1.08122.0005, Fa. Merck)

- DNeasy[®] Tissue-Kit (50 Reaktionen, Produkt-Nr. 69504, Fa. Qiagen)
- dATP, Na-salt, PCR Grade (Fa. Merck, Produkt-Nr. 1 934 511)
- dCTP, Na-salt, PCR Grade (Fa. Merck, Produkt-Nr. 1 934 520)
- dGTP, Na-salt, PCR Grade (Fa. Merck, Produkt-Nr. 1 934 538)
- dTTP, Na-salt, PCR Grade (Fa. Merck, Produkt-Nr. 1 934 549)
- DNA Molecular Weight Marker XIII (50-750 Bp, Fa. Roche Diagnostics, Produkt-Nr.: 1 721 925)
- DNA Molecular Weight Marker XIV (100-1500 Bp, Fa. Roche Diagnostics, Produkt-Nr.: 1 721 933)
- Ethanol Absolut (Fa. Merck, Produkt-Nr. 1.00983)
- Ethidiumbromid-Lösung (1%ige wäßrige Lösung, Fa. Merck, Produkt-Nr. 1.11608.0030,)
- Expand High Fidelity PCR-System 100 Units (Fa. Roche Diagnostics, Produkt-Nr.: 1732641)
- Ficoll (Fa. Sigma, Produkt-Nr. F-2878)
- Fluorescein Isothiocyanat (Fa. Sigma, Produkt-Nr. F7250)
- Isopropanol (Fa. Merck, Produkt-Nr. 1.09634)
- MetaPhor[®] Agarose (Fa. BMA[®], Produkt-Nr. 50180)
- Restriktionsendonuklease *Taq* I (Produkt-Nr. 404128, Fa. Roche Diagnostics)
- Restriktionsendonuklease *Fok* I (Produkt-Nr. 1 004 816, Fa. Roche Diagnostics)
- Restriktionsendonuklease *Rsa* I (Produkt-Nr. 729 124, Fa. Roche Diagnostics)
- Restriktionsendonuklease *Hae* III (Produkt-Nr. 693 936, Fa. Roche Diagnostics)
- Salzsäure 25% reinst. (Fa Merck 312)
- Tris-hydroxymethylaminomethan (TRiS Ultra Qualität) (Fa. Roth, Produkt-Nr. 5429 3)
- Trinatriumcitrat-Dihydrat (Fa Merck 1 06448)
- TripleMaster PCR System (Fa. Eppendorf, Produkt-Nr. 0032 008.216)

3.2.5 HERSTELLUNG VON PUFFERN UND GEBRAUCHSLÖSUNGEN

3.2.5.1 Puffer und Lösungen für die Polymerasekettenreaktion

Herstellung des Nukletid (engl. „dNTP-Mix“ je 10 mM):

Jeweils:

10µl dATP, Na-salt, PCR Grade

10µl dCTP, Na-salt, PCR Grade

10µl dGTP, Na-salt, PCR Grade

10µl dTTP, Na-salt, PCR Grade

werden mit 60µl steril filtriertem Aqua bidest gemischt und autoklaviert.

3.2.5.2 Agarosen und Puffer für die konventionelle Gelelektrophorese

10x TBE (Tris-(hydroxymethyl)aminomethan-Borsäure, EDTA-Puffer (1 Liter):

60,60 g (0,5 M) Tris- (hydroxymethyl)aminomethan

30,90 g (0,5 M) Borsäure

2,92 g (10 mM) EDTA (Titriplex II) =

1x TBE (Tris-(hydroxymethyl)aminomethan-Borsäure, EDTA-Puffer (2 Liter):

200 ml 10x TBE- Puffer werden in einen 2 Liter- Glasrundkolben überführt und bis zur Marke mit steril filtriertem Aqua bidest aufgefüllt.

1,5%-Agarose-/3%-Metaphorgele:

1,35 g Agarose Standard (Fa. Q Bio gene, Produkt-Nr. AGAH0500) bzw. 2,7 g MetaPhor® Agarose (Fa. BMA®, Produkt-Nr. 50180) werden in 90 ml 1x TBE aufgenommen. Nach 30-minütigem Quellen erfolgt die Erhitzung in der Mikrowelle (Typ: Dimension 4, Fa. Panasonic) bis zum Ausbleiben von Schlieren oder Luftblasen. Anschließend wird unter Rühren (Magnetrührer „RCT-Basic“, Fa. KIKA Labordiagnostic) auf +45°C abgekühlt und der Ansatz in die abgedichtete Gelkammer gegossen. Nach etwa 30-minütigem Abbinden wird die Gelkammer in die Elektrophoresekammer mit dem Laufpuffer (1x TBE) verbracht. Nach

5minütigem Einwirken des Puffers können die Taschenschablonen gezogen werden.

Probenauftragspuffer/Bromphenolblau-Lösung

1. Herstellung der 15%igen Ficoll-Lösung

1,5 g	Ficoll
10 ml	Aqua bidest/sterilfiltriert

2. Herstellung der Stammlösung

10 ml	15%ige Ficoll-Lösung
0,25 ml	5%ige Bromphenolblaulösung

3. Herstellung der Gebrauchslösung für die Gelelektrophorese

2µl	Stammlösung
8µl	PCR-Amplifikat

Färbebad/Ethidiumbromid-Bad

50µl 1,0%ige Ethidiumbromidlösung
1,5 l Aqua bidest

3.2.6 REVITALISIERUNG UND ANZÜCHTUNG DER BAKTERIENKULTUREN

Die Stammhaltung der untersuchten *E. coli* erfolgte in NUNC®-Röhrchen in BHI-Glycerol bzw. mit dem MICROBANK™-System (Fa. Mast Diagnostica, Produkt-Nr. 291601).

Zur Revitalisierung wurde das Inokulum in 5 ml Hirn-Herz-Bouillon/Brain-Heart-Infusion (BHI)-Bouillon (Fa. Merck, Produkt-Nr. 1.10493) überimpft und anschließend bei 37°C für 24 h aerob inkubiert. Als Festmedium wurde Plate-Count (PC)-Agar (Fa. Merck, Produkt-Nr. 1.05463) mittels Dreiösenausstrich beimpft und anschließend bei 37°C für 24 h bebrütet. Die Herstellung der Nährmedien erfolgte nach Angaben des Herstellers.

3.3 POLYMERASEKETTENREAKTION (PCR)

3.3.1 EXTRAKTION GENOMISCHER DNA AUS BAKTERIENKULTUREN

Um ein zeit- und kostensparendes Prozedere zu gewährleisten, wurden drei Einzelkolonien vom PC-Agar abgenommen und in 300µl Aqua dest. eingerieben. Nach gründlichem Vortexen folgte das thermische Aufbrechen der Zellwand durch Kochen bei 100°C für 10 min im Wasserbad (Fa. Köttermann). Die Separation des Zelldetritus wurde durch einen abschließenden Zentrifugationsschritt in der Eppendorf-Tischzentrifuge bei 13.000 Umdrehungen pro Minute (engl. „rounds per minute [rpm]“) für 1 min gewährleistet. Der klare Überstand enthält die genomische DNA, die zur Minimierung nachteiliger Hydrolysereaktionen bei +4°C gelagert und maximal sieben Tage nach Gewinnung verwendet wurde.

3.3.2 EXTRAKTION GENOMISCHER DNA AUS FLÜSSIGMEDIEN

Um für die PCR-Optimierungsschritte, die fortlaufenden Typisierungsreaktionen und abschließenden Sequenzierungsarbeiten DNA hoher Qualität einsetzen zu können, wurde die DNA der Referenzstämme und der zur Sequenzierung verwendeten Isolate mittels DNeasy[®] Tissue-Kit (Produkt-Nr. 69504, Fa. Qiagen) gemäß dem Protokoll zur Aufbereitung Gram-negativer Bakterien isoliert. Dabei erlaubt die Bindung der Nukleinsäuren an die Silikaoberfläche der eingesetzten Spinsäulen eine zuverlässige Aufreinigung des Ausgangsmaterials. Dazu wurden Übernachtskulturen in BHI-Bouillon angelegt. Zur fortlaufenden Verwendung wurden 15µl-Aliquots bei -25°C tiefgefroren. Die Anzüchtung der Stämme erfolgte in BHI-Bouillon (Fa. Merck, Produkt-Nr. 1.10493) für 24 h bei 37°C. Je 1,5 ml des inkubierten Nährmediums wurde in sterile Reaktionsgefäße gegeben und für 10 min bei 7.500 rpm (5.000xg) zentrifugiert.

Das Sediment wurde zur Zellwandlyse und Proteinentfernung mit 180µl ATL-Puffer aus dem DNeasy[®] Tissue Kit (Produkt-Nr. 69504, Fa. Qiagen) unter Zusatz von 20µl Proteinase K aus dem Kitsystem versetzt und anschließend für 30 min bei 55°C im Thermomixer inkubiert. Zur Steigerung der bakteriellen Lyse wurde anschließend gut durchmischt, 200 ml AL-Puffer aus dem Kitsystem zugegeben und erneut für 10 min bei 70°C im Wasserbad (Fa. GFL) inkubiert.

Nach Zugabe von 200µl 96%-igem Ethanol und intensiver Durchmischung wurde der Ansatz durch einen „DNeasy“-Filtereinsatz mit Silicamembran mittels Zentrifugation für 1 min bei 6.000xg (8.000 rpm) in ein 2 ml-Reaktionsgefäß abfiltriert. Die an die Membran gebundene DNA wurde 2x mit je 500µl AW-Puffer gewaschen und dann nach dem ersten Waschdurchgang bei 6.000xg (8.000 rpm) für 1 min bzw. nach dem zweiten Waschdurchgang für 3 min zum Trocknen der Silicamembran zentrifugiert. Der erhaltene Überstand wurde verworfen und der „DNeasy“-Filtereinsatz mit Silicamembran in ein neues 2 ml-Reaktionsgefäß verbracht. Das Eluieren der an die Membran gebundenen DNA erfolgte mittels 30µl AE-Puffer. Dazu wurde der Elutionspuffer direkt auf die Säule gegeben und 5 min bei Raumtemperatur inkubiert. Die Separation des DNA-haltigen Eluates erfolgte erneut durch Zentrifugation bei 6.000xg (8.000 rpm). Zur Maximierung der DNA-Ausbeute wurde der Elutionsschritt mit 30µl AE-Puffer wiederholt.

3.3.3 EIGENSCHAFTEN DER EINGESETZTEN POLYMERASE-TYPEN

Für die Methoden nach ADU-BOBIE et al. (1998a) und OSWALD et al. (2000) wurden insgesamt fünf verschiedene rekombinante, thermostabile Polymerasen hinsichtlich ihrer Amplifikationseffizienz überprüft. Dabei wurden neben der subjektiv ermittelten Bandenintensität unter Berücksichtigung der Zyklenzahl auch das Auftreten und die Intensität von Fehlbanden berücksichtigt. Die abschließende Optimierung erfolgte über die bekannten Parameter Primer-Anlagerungstemperatur (engl. „Annealing“-temperatur“) und Magnesiumchloridkonzentration. Ziel der vergleichenden Analysen war die Bereitstellung einer

möglichst kostensparenden Zusammensetzung des Reaktionsansatzes (engl. „Mastermix“).

Dabei wurden konventionelle als auch sog. „hot start“-Formulierungen der aus *Thermus aquaticus* gewonnenen Taq-Polymerase eingesetzt. Insbesondere die von circa 2,2 bis 2,8 Kb reichenden Amplifikate der SK1- und LP-Primer-Kombinationen erforderten den Einsatz von Enzymkombinationen mit 3'-5'-Exonukleaseaktivität. Diese als Korrekturlesen (engl. „proofreading“) bezeichnete Eigenschaft wird im Expand High Fidelity PCR-System (Fa. Roche Diagnostics, Produkt-Nr. 1732641) durch Zugabe der aus *Thermococcus gorgonarius* isolierten rekombinanten Tgo-Polymerase erreicht. Das TripleMaster® PCR System (Fa. Eppendorf, Produkt-Nr. 0032 008.216) verfügt dabei über eine ähnliche Zusammensetzung. Aus patenschutzrechtlichen Gründen wird das die „Proofreading“-Komponente besteuernde Enzym vom Hersteller nicht deklariert. Es ist jedoch davon auszugehen, dass es sich um einen weiteren Vertreter der Typ-B-Familie der Archae-Gruppe handelt. Sehr häufig werden Modifikationen der aus *Pyrococcus furiosus* erhaltenen Pfu-Polymerase verwendet. Eine zusammenfassende Aufstellung der eingesetzten Polymeraseformulierungen, ihrer Eigenschaften und den laut den Herstellern vorgesehenen Anwendungsgebieten wird in **Tabelle 3.3.3-1** wiedergegeben.

Tabelle 3.3.3-1: Eigenschaften der verwendeten Polymerase-Typen

Bezeichnung	Enthaltene(s) Enzym(e)	hot start (+/-)	Isoliert aus folgendem Mikroorganismus	Anwendungsgebiet
AmpliTa ^q [®] DNA-Polymerase	<i>Taq</i> mit 5'-3'-Exonukleaseaktivität	-	<i>Thermus aquaticus</i>	Standard PCR ≤ 3 Kb Amplifikatgröße
AmpliTa ^q Gold [®] DNA-Polymerase	<i>Taq</i> mit 5'-3'-Exonukleaseaktivität	+	<i>Thermus aquaticus</i>	Standard PCR ≤ 3 Kb Amplifikatgröße
Platinum [®] <i>Taq</i> -Polymerase	<i>Taq</i> mit 5'-3'-Exonukleaseaktivität	+	<i>Thermus aquaticus</i>	Standard PCR ≤ 5 Kb Amplifikatgröße
Expand High Fidelity PCR System	<i>Taq</i> mit 5'-3'-Exonukleaseaktivität	-	<i>Thermus aquaticus</i>	Standard PCR ≤ 5 Kb Amplifikatgröße
	<i>Tgo</i> mit 5'-3'-Exonukleaseaktivität und Proofreading	-	<i>Thermococcus gorgonarius</i>	
TripleMaster [®] PCR System	<i>Taq</i> mit 5'-3'-Exonukleaseaktivität	-	<i>Thermus aquaticus</i>	Produkte mit hohem GC-Anteil, Sekundärstrukturen High Fidelity/Proof Reading PCR (≤ 5 Kb Amplifikatgröße) Long Range PCR (> 40 Kb)

3.3.4 DARSTELLUNG DER PCR-PRODUKTE MITTELS AGAROSE-GELELEKTROPHORESE

Die Herstellung der 1,5%-igen Agarose- bzw. 3%-igen Metaphorgele erfolgte wie unter **Kap. 3.2.5.1** beschrieben. Um eine gleich bleibende Qualität der Agarosegele zu gewährleisten, wurden Metaphorgele bis 48h nach der Herstellung verwendet. Die Lagerung erfolgte in der Gelelektrophoresekammer im Laufpuffer. Mit Agarose Standard produzierte Gele wurden bis zu sieben Tage nach der Herstellung eingesetzt. Für den restriktionsenzymatischen Verdau vorgesehene PCR-Amplifikate wurden nicht mit Bromphenolblau-Stammlösung versetzt. Hier wurden je Amplifikat ca. 2µl des Farbstoffs auf Parafilm aufgebracht und durch „Auf- und Abziehen“ in der Pipettenspitze mit 8µl des PCR-Produktes durchmischt.

3.3.5 ETHIDIUMBROMIDFÄRBUNG UND FOTODOKUMENTATION

Im Anschluß an die Gelelektrophorese erfolgte die Visualisierung der Amplifikate durch Färbung des Agarosegels für 20 min in einer wässrigen Ethidiumbromid-Lösung (5 µg/ml, Sigma) im Schüttelwasserbad. Überschüssiges Ethidiumbromid wurde durch 20minütiges Schwenken in Aqua dest. bei Raumtemperatur entfernt. Zur Auswertung wurde der UV-Transilluminator (Mighty Bright UVTM-25, Fa. Hoefer Scientific Instruments) eingesetzt. Zur Dokumentation auf Polaroidfilm mit der manuellen Fotoanlage (MP 4*, Fa. Polaroid) wurden Blendengrößen von „11“ bzw. „16“ vorgewählt.

3.3.6 PRIMER-SYSTEME

In den **Tabellen 3.3.7-1, 3.3.8-1 bis -3** sind die für die einzelnen Arbeitsschritte eingesetzten Primer-Systeme unter Berücksichtigung der resultierenden PCR-Produktgröße und der Angabe der zugrunde liegenden Publikation zusammengefasst. Die Primer-Synthese wurde durch die Fa. MWG-BIOTECH (Ebersberg, Deutschland) übernommen.

3.3.7 PRIMER-SYSTEM ZUM NACHWEIS DES EAE-GENS

KARCH et al. (1993) entwickelten das Primer-Paar „SK1“ und „SK2“ zur Amplifikation des konservierten *eae*-Genabschnittes enteropathogener und enterohämorrhagischer *E. coli*. Das PCR-System wurde in vorangegangenen Arbeiten sowohl institutsintern als auch international umfangreich validiert (**Kap. 5.5**), so dass auf eine weitere Austestung hinsichtlich falsch-positiver respektive -negativer Ergebnisse verzichtet werden konnte. Der Abgleich der Oligonukleotidsequenzen mit den Kenndaten der Genbank des National Center for Biotechnology Information (NCBI[®]), Bethesda, USA bestätigte die sichere Erfassung der in dem detektierten Genabschnitt hochkonservierten *eae*-Genträger. **Tabelle 3.3.7-1** beinhaltet die Primer-Positionierung, die -sequenzen und die PCR-Produktgröße (Bp).

Tabelle 3.3.7-1: Charakteristika der Primer zur Detektion des *eae*-Gens nach KARCH et al. (1993)

Primer-Bezeichnung	Primer-Positionierung*	Primer-Sequenzen	PCR-Produktgröße (Bp)
„SK1“ (Vorwärts-Primer)	26-46	5'-CCC GAA TTC GGC ACA AGC ATA AGC-3'	881
„SK2“ (Rückwärts-Primer)	879-903	5'-CCC GGA TCC GTC TCG CCA GTA TTC G-3'	

* Angabe bezogen auf den Referenzstamm EHEC O157:H7 EDL 933, Genbankzugriffs-Nr. Z11541

Die Zusammensetzung des Reaktionsansatzes ist in der folgenden Übersicht zusammengefasst.

Komponenten	Volumina je 25µl-Ansatz [µl]	Endkonzentration
Aqua bidest	17,1	
10 x Polymerase-Puffer (15 mM MgCl ₂)	2,5	1x
dNTP-Mix (je 10 mM)	0,5	0,2 mM
Primer „SK1“ (10 mM)	1,0	0,4 mM
Primer „SK2“ (10 mM)	1,0	0,4 mM
Polymerase Ampli Taq Gold® (5 U/µl) b	0,4	2 U
Proben-DNS	2,5	

Eine Konzentrationsbestimmung der Proben-DNS erfolgte nicht, da die DNS aus bakteriellen Reinkulturen gewonnen wurde. Bei Fehlen einer erwarteten PCR-Produktes wurde der PCR-Ansatz wiederholt.

Im Folgenden wird die Herstellung und Aliquotierung des PCR-REaktionsansatzes am Beispiel der PCR nach KARCH et al. (1993) beschrieben. Für die übrigen PCR-Systeme wurde unter Einbeziehung der in den Einzelübersichten der Mastermixkomponenten aufgeführten Reagenzien und Volumina und der Thermocycler-Profile **Tabelle 3.3.10-1** entsprechend verfahren. Der Ansatz des Mastermixes erfolgte im PCR-Kabinett. Zur Minimierung des Risikos einer DNA-Kontamination wurde UV-Licht verwendet. Die Einzelkomponenten wurden zum Auftauen und während der Arbeiten in Kühlboxen (Iso-Therm[®]-System, Fa. Eppendorf) bei +4°C gelagert bzw. im Falle der Polymerasen erst unmittelbar vor Gebrauch aus dem Gefrierschrank entnommen und in die Kühlboxen verbracht. Für den Reaktionsansatz wurden je PCR-Reaktion in 2 ml-Reagenzgefäße (Fa. Sarstedt) 17,1µl Aqua bidest vorgelegt. Anschließend wurden 2,5µl GeneAmp[®] Puffer mit 15 mM MgCl₂ (Fa. Applied Biosystems, Produkt-Nr. 4311814) hinzugefügt. Zuvor war der Puffer nach vollständigem Auftauen gemischt worden. Während der Lagerung und der Homogenisation in den Deckel gelangte Flüssigkeitsanteile wurden durch kurzes Anzentrifugieren in der Eppendorf Tischzentrifuge reponiert. Daran anschließend erfolgte eine erneute Homogenisation durch Auf- und Ab-Pipettieren in der Eppendorf-Pipette. Von dem wie unter „Herstellung von Puffern und Gebrauchslösungen“ angesetzten dNTP-Mix (je 10 mM) wurden 0,5µl je Reaktionsansatz hinzugefügt. Die vorgeschaltete Homogenisation erfolgte hierbei nur durch vorsichtiges Auf- und Ab-Pipettieren. Jeweils 1µl in 10 mM konzentrierter Gebrauchslösung vorliegenden Vorwärts-Primers „SK1“ und des Rückwärts-Primers „SK2“ wurden anschließend hinzugefügt.

Abschließend wurden 0,4µl AmpliTaq Gold[®] DNA Polymerase (Fa. Applied Biosystems, Produkt-Nr. 4311814) hinzugefügt. Der für ein Reaktionsvolumen von 25µl ausgelegte Mastermix wurde vor der Aliquotierung von 22,5µl je Reaktionsansatz in 200 bzw. 100µl fassende, auf Tiefkühlboxen gelagerten PCR-Reaktionsgefäße durch vorsichtiges „Auf- und Ab-Pipettieren“ homogenisiert.

Die von PC-Agar bzw. aus BHI-Bouillon gewonnene DNA wurde anteilig mit 2,5µl hinzugefügt. Nach Verschluss der PCR-Reaktionsgefäße wurden diese unter fortlaufender Kühlung in den Thermocycler TouchDown™ Temperature Cycling System (Fa. Hybaid) bzw. Biometra Trio-Thermoblock mit Heizdeckeinheit (Fa. Biometra) verbracht. Für jede PCR wurden Positiv- und Negativkontrollen (EHEC EDL 933 O157:H7 und *E. coli* apathogen AF25922) sowie Leerwertansätze des Mastermixes mitgeführt.

3.3.8 PRIMER-SYSTEME ZUR INTIMINTYPISIERUNG

ADU-BOBIE et al. (1998a) differenzierten vier verschiedene *eae*-Genvarianten, die sie als α -, β -, δ - und γ -Intiminvarianten bezeichneten. Vorausgegangen war eine immunologische Charakterisierung von α - und β -Intimin bei EPEC-Stämmen (ADU-BOBIE et al., 1998b). Die zur Bestätigung verwendeten Primer-Kombinationen sehen einen universellen Rückwärts-Primer „IntRu“ vor. Die Charakteristika, einschließlich des in dieser Arbeit entwickelten Primer-Paares zum Nachweis von ε -Intimin, sind nachfolgend zusammengefaßt (**Tabelle 3.3.8-1**).

Tabelle 3.3.8-1: Charakteristika der Primer zur Typisierung des *eae*-Gens nach ADU-BOBIE et al. (1998a) und der in dieser Arbeit entwickelten Oligonukleotide zur Erfassung von epsilon-Intimin

Primer-Bezeichnung*	Erfasster Intimintyp	Primer-Positionierung	Primer-Sequenzen	PCR-Produktgröße (Bp)
Int-A	α -Intimin	2820-2803	5'-CCT TAG GTA AGT TAA GT-3'	557
Int-B	β -Intimin	2097-2113	5'-TAA GGA TTT TGG GAC CC-3'	562
Int-Y	γ -Intimin	2098-2115	5'-ACA AAC TTT GGG ATG TTC-3'	546
Int-D	δ -Intimin	126-142	5'-TAC GGA TTT TGG GGC AT-3'	562
Int-Ru	Universalrückwärts-Primer	2643-2625	5'-TTT ATT TGC AGC CCC CCA T-3'	
epsilonF	ϵ -Intimin	2393-2413	5'-GGC-GCT-CTC-AAG-ATC-CAA-GTG-3'	340
epsilonR		2732-2712	5'-TGT-TTA-TCA-TCC-TCA-GTT-TGC-3'	

* „Int-Ru/Int-Y“: Angabe bezogen auf den Referenzstamm EHEC O157:H7 EDL 933, Genbankzugriffs-Nr. Z11541

„Int-A“: Bezogen auf Referenzstamm EPEC O127.6 E2348/69, Genbankzugriffs-Nr. M58154

„Int-B“: Bezogen auf Referenzstamm EPEC O15:H⁻ RDEC-1, Genbankzugriffs-Nr. U60002

„Int-D“: Bezogen auf Referenzstamm *E. coli* O86:H34, Genbankzugriffs-Nr. Y133112

„epsilonF/epsilonR“: Bezogen auf Referenzstamm EHEC O103:H2 PMK5, Genbankzugriffs-Nr. AF116899

Die Herstellung und Aliquotierung des nachfolgend gelisteten Reaktionsansatzes erfolgte wie in **Kap. 3.3-7** beschrieben. Als Thermocycler wurde das GeneAmp PCR-System 9600 (Fa. Perkin Elmer) eingesetzt.

Komponenten	Volumina je 50µl-Ansatz [µl]	Endkonzentration
Aqua bidest.	31,0	
10x Polymerase-Puffer (15 mM MgCl ₂)	5,0	1x (1,5 mM)
dNTP-Mix (je 10 mM)	1,0	0,2 mM
Primer „Int-Ru“ (10 mM)	5,0	1 mM
Primer „Int-A, Int-B, Int-Y, Int-D“ (10 mM)	5,0	1 mM
Polymerase Ampli [®] Taq (5 U/µl)	0,5	2,5 U
Proben-DNS	2,5	

OSWALD et al. (2000) konnten die Existenz einer fünften, der ϵ -Intiminvariante, zeigen. Sie verwendeten „SK1“ als universellen Vorwärts-Primer in Kombination mit den nachfolgend aufgeführten Oligonukleotiden (**Tabelle 3.3.8-2**).

Tabelle 3.3.8-2: Charakteristika der Primer zur Typisierung des *eae*-Gens nach OSWALD et al. (2000)

Primer-Bezeichnung	Erfasster Intimintyp	Primer-Positionierung*	Primer-Sequenzen	PCR-Produktgröße (Bp)
SK1**	Universalvorwärts-Primer	26-46	5'-CCC GAA TTC GGC ACA AGC ATA AGC-3'	
LP2***	α -Intimin	2803-2829	5'-CCC GAA TTC TTA TTT TAC ACA AGT GGC -3'	2807
LP3***	γ -Intimin	2788-2814	5'- CCC GAA TTC TTA TTC TAC ACA AAC CGC-3'	2792
LP4	β -Intimin	2283-2309	5'- CCC GTG ATA CCA GTA CCA ATT ACG GTC-3'	2287
LP5	ϵ -Intimin	2605-2630	5'- AGC TCA CTC GTA GAT GAC GGC AAG CG-3'	2608

- * „SK1/LP3“: Angabe bezogen auf den Referenzstamm EHEC O157:H7 EDL 933, Genbankzugriffs-Nr. Z11541
 „LP2“: Bezogen auf Referenzstamm EPEC O127.6 E2348/69, Genbankzugriffs-Nr. M58154
 „LP4“: Bezogen auf Referenzstamm EPEC O15:H⁻ RDEC-1, Genbankzugriffs-Nr. U60002
 „LP5“: Bezogen auf Referenzstamm EHEC O103:H2 PMK5, Genbankzugriffs-Nr. AF116899
- ** „SK1“: Erstmalig beschrieben von KARCH et al. (1988)
- *** LP2/LP3: Erstmalig beschrieben von SCHMIDT et al. (1993)

Da die Autoren selbst keine vollständigen Angaben zur Zusammensetzung des Mastermixes angaben, wurde auf die Angaben von GALLIEN et al. (2000) sowie die Herstellerangaben zum Einsatz der jeweiligen Polymerase zurückgegriffen. Die Zusammensetzung des Reaktionsansatzes ist nachfolgend dargestellt.

Komponenten	Volumina je 50µl- Ansatz [µl]	Endkonzentration
Aqua bidest.	25,79	
10x Polymerase-Puffer (ohne MgCl ₂)	5,0	1x
25 mM MgCl ₂	5,0	2,5 mM
dNTP-Mix (je 10 mM)	1,0	0,2 mM
Primer „SK1“ (10 mM)	5,0	1 mM
Primer „LP2, LP3, LP4, LP5“ (10 mM)	5,0	1 mM
Polymerase Expand High Fidelity (3,5U/µl)	0,71	2,5 U
Proben-DNS	2,5	

RAMACHANDRAN et. al (2003) beschrieben die Amplifikation der Intiminvarianten alpha, beta, gamma, delta, epsilon, zeta, eta, theta, iota, kappa, lambda, my, ny, xi, pi, rho und sigma mittels eines universellen Vorwärts-Primers „EaeVF“ und drei Rückwärts-Primern „EaeVR, EaeZetaVR“ und „EaelotaVR“. Die Charakteristika dieser Primer werden in **Tabelle 3.3.8-3** zusammengefasst.

Tabelle 3.3.8-3: Charakteristika der Primer zur Typisierung des *eae*-Gens nach RAMACHANDRAN et al. (2003)

Primer-Bezeichnung	Erfasster Intimintyp	Primer-Positionierung*	Primer-Sequenzen	PCR-Produktgröße (Bp)**
EaeVF	Universalvorwärts-Primer	1975-1992	5'- AGY ATT ACT GAG ATT AAG -3'	
EaeVR	Alle Typen	2806-2823	5'- AAA TTA TTY TAC ACA RAY -3'	834-876
EaeZetaVR		2803-2820	5'- AGT TTA TTT TAC GCA AGT -3'	
EaelotaVR		2800-2819	5'- TTA AAT TAT TTT ATG CAA AC -3'	

* „EaeVF/EaeVR“: Angabe bezogen auf den Referenzstamm EHEC O157:H7 EDL 933, Genbankzugriffs-Nr. Z11541

„EaeZetaVR“: Angabe bezogen auf den Referenzstamm P 256 O84:H, Genbankzugriffs-Nr. AY520905

„EaelotaVR“: Angabe bezogen auf den Referenzstamm P 256 O84:H, Genbankzugriffs-Nr. AY520905

** Abhängig vom amplifizierten Intimintyp

Der Einsatz der Oligonukleotide erfolgte nach dem hier dargestellten Schema:

Komponenten	Volumina je 50µl-Ansatz [µl]	Endkonzentration
Aqua bidest.	31,0	
10x Polymerase-Puffer (15 mM MgCl ₂)	5,0	1 x
dNTP-Mix (je 10 mM)	1,0	0,2 mM
Primer „EaeVR“ (10 mM)	5,0	1 mM
Primer „EaeVR, EaeZetaR, EaelotaR“ (10 mM)	5,0	1 mM
Polymerase Ampli [®] Taq (5 U/µl)	0,5	2,5 U
Proben-DNS	2,5	

3.3.9 PRIMER-SYSTEME ZUR SEQUENZIERUNG DES *EAE*-GENS

Für die schrittweise Amplifikation und Sequenzierung des *eae*-Gens der verotoxinbildenden Isolate P 146 O84:H⁻, isoliert aus Schafkot, und P 256 O156:H⁻, isoliert von der Oberfläche eines Schafschlachtettkörpers, wurden neben der Primer-Kombination „SK1/SK2“ eigene Oligonukleotide zur partiellen Amplifikation der benachbarten Genabschnitte im konservierten 5'- und im hypervariablen 3'-Ende entwickelt. Das verwendete Thermocyclerprofil basierte auf der Vorgabe von ABDULMAWJOOD et al. (2001) und gewährleistete die Amplifikation von bis zu 800 Bp großen PCR-Produkten. Die Mastermixzusammensetzung entspricht dem PCR-System nach ADU-BOBIE et al. (1998a).

3.3.10 THERMOCYCLER-PROFILE

Die Thermocycler-Profile werden für die vier zum Nachweis bzw. zur Intimingentypisierung verwendeten PCR-Systeme sowie die Primer-Systeme zur Sequenzierung unter Angabe der Primer-Kombinationen und der Referenz in **Tabelle 3.3.10-1** aufgeführt.

Tabelle 3.3.10-1: Thermocycler-Profile zur Intimintypisierung und Sequenzierung

Zielgen	Primer-Akronym	Thermocycler-Profil	Referenz
eae-Gen, konservierte Region	SK1/SK2	30 Zyklen a´ 95°C/42 s, 52°C/60 s und 72°C/40 s ^{1, 2}	KARCH et al. (1988),
alpha-Intimin	Int-Ru/Int-A	30 Zyklen a´ 94°C/45 s, 48°C/60 s und 72°C/50 s ^{2, 3}	ADU-BOBIE et al. (1998a), modifiziert
beta-Intimin	Int-Ru/Int-B	30 Zyklen a´ 94°C/45 s, 48°C/60 s und 72°C/50 s ^{2, 3}	
delta-Intimin	Int-Ru/Int-Y	30 Zyklen a´ 94°C/45 s, 48°C/60 s und 72°C/50 s ^{2, 3}	
gamma-Intimin	Int-Ru/Int-D	30 Zyklen a´ 94°C/45 s, 55°C/60 s und 72°C/50 s ^{2, 3}	
epsilon-Intimin	epsilonF/epsilonR	30 Zyklen a´ 94°C/45 s, 55°C/60 s und 72°C/50 s ^{2, 3}	Eigene Untersuchungen
alpha-Intimin	SK1/LP2	30 Zyklen a´ 94°C/60s, 55°C/60 s, 72°C/120 s ^{2, 3}	LP4, LP5: OSWALD et al. (2000) LP2, LP3: SCHMIDT et al. (1993) SK1:KARCH et al. (1988) jeweils modifiziert
beta-Intimin	SK1/LP4	30 Zyklen a´ 94°C/60s, 55°C/60 s, 72°C/120 s ^{2, 3}	
gamma-Intimin	SK1/LP3	30 Zyklen a´ 94°C/60s, 55°C/60 s, 72°C/120 s ^{2, 3}	
epsilon-Intimin	SK1/LP5	30 Zyklen a´ 94°C/60s, 55°C/60 s, 72°C/120 s ^{2, 3}	
zeta- Intimin	EaeVR/EaeZetaVR	35 Zyklen a´ 94°C/60s, 41°C/60 s, 72°C/60 s ^{2, 3}	RAMACHANDRAN et al. (2003), modifiziert
iota- Intimin	EaeVR/EaelotaVR		
übrige Intiminvarianten	EaeVR/EaeVF		
Schrittweise Amplifikation des eae-Gens	-	35 Zyklen a´ 93°C/30s, 52°C/45 s, 72°C/45 s ^{2, 4}	ABDULMAWJOOD et al. (2001)

¹ Initiale Denaturierung: 95°C/420 s

² Finale Extension: 72°C/300 s

³ Initiale Denaturierung: 94°C/300 s

⁴ Initiale Denaturierung: 93°C/180 s

3.4 Restriktionsfragmentlängenpolymorphismus (RFLP)

3.4.1 AUSWAHL UND EIGENSCHAFTEN DER RESTRIKTIONS- ENDONUKLEASEN

Die Auswahl geeigneter Restriktionsendonukleasen bzw. -enzyme erfolgte unter Verwendung der Software CloneManger™ (Fa. Scientific & Educational Software) bzw. MapDraw™ (Fa. DNASStar). Die für die Sequenzanalyse eingesetzten Referenzsequenzen sind unter Angabe der Genbank-Zugriffsnummer der **Tabelle 9.1.1-1** im Anhang zu entnehmen. **Tabelle 3.4.1-1** fasst die Nomenklatur, Schnittstellen und die Klassifikation der verwendeten Enzyme zusammen.

Für die Produktüberprüfung der durch eigene Primer um epsilon-Intimin erweiterten PCR nach ADU-BOBIE et al. (1998a), gewährleistete die Enzymkombination aus den beiden palindromisch schneidenden Enzymen *Taq* I (Produkt-Nr. 404128, Fa. Roche Diagnostics) und *Fok* I (Produkt-Nr. 1 004 816, Fa. Roche Diagnostics) eine sichere Differenzierung der überprüften fünf *eae*-Amplifikate (**Tabelle 3.4.1-2**).

Für eine exakte restriktionsenzymatische Analyse der mit den Universal-Primern nach RAMACHANDRAN et al. (2003) generierten PCR-Produkte aus den zu erwartenden 19 verschiedenen Intiminuntereinheiten erwies sich die genannte Kombination als nicht ausreichend. Daher wurden die palindromisch schneidenden Endonukleasen *Fok* I (Produkt-Nr. 1 004 816, Fa. Roche Diagnostics), *Rsa* I (Produkt-Nr. 729 124, Fa. Roche Diagnostics) und *Hae* III (Produkt-Nr. 693 936, Fa. Roche Diagnostics) nach EDV-gestützter Analyse ausgewählt (**Tabelle 3.4.1-3**).

Tabelle 3.4.1-1: Nomenklatur, Schnittstellen und Klassifikation der eingesetzten Restriktionsendonukleasen für die PCR-Produktüberprüfung der modifizierten PCR nach ADU-BOBIE et al. (1998a) und der PCR nach RAMACHANDRAN et al. (2003)

Bezeichnung der Restriktionsendonukleasen	Isoliert aus	Reihenfolge der Entdeckung im Organismus	Schnittstellen
<i>Taq</i> I	<i>Thermus aquaticus</i>	1	$\begin{array}{c} \downarrow \\ \text{TCGA}^* \\ \text{A}^*\text{GCT} \\ \uparrow \end{array}$
<i>Fok</i> I	<i>Flavobacterium okeanokoides</i>	1	$\begin{array}{c} \text{GGA}^*\text{TG}(\text{N})_9 \\ \text{CTA}^*\text{C}(\text{N})_{13} \end{array}$
<i>Rsa</i> I	<i>Rhodopseudomonas sphaeroides</i>	1	$\begin{array}{c} \text{GT} \downarrow \text{A}^*\text{C}^* \\ \text{C}^*\text{A}^*\text{TG} \\ \uparrow \end{array}$
<i>Hae</i> III	<i>Haemophilus aegypticus</i>	3	$\begin{array}{c} \text{GG} \downarrow \text{C}^*\text{C}^\circ \\ \text{C}^\circ\text{C}^*\text{ATG} \\ \uparrow \end{array}$

* / ° Erkennung der Schnittstelle kann durch Methyltransferaseaktivität gehemmt werden. Dabei können sowohl 5- bzw. 6-Methyladenin, als auch 4-Methylcytosin inhibieren.

Tabelle 3.4.1-2: Produktüberprüfung (n=5) der PCR nach ABU-BOBIE et al. (1998a): Eingesetzte Restriktionsenzyme, Anzahl der Schnittstellen und resultierende Fragmentgrößen

Enzym*	Anzahl der Schnittstellen	Erwartete Fragmentgrößen (Bp) bezogen auf die Intiminuntereinheiten				
		alpha	beta	gamma	delta	epsilon
<i>Taq</i> I	2, 0	489, 68	450, 112	319, 227	keine Schnittstelle	294, 46
<i>Fok</i> I	je 2	326, 231	501, 61	480, 66	404, 97, 61	254, 86

* *Taq* I: Isoliert aus *Thermus aquaticus*, erstes (I) bei diesem Mikroorganismus beschriebenes Restriktionsenzym

Fok I: Isoliert aus *Floavobacterium okeanokoites*, erstes (I) bei diesem Mikroorganismus beschriebenes Restriktionsenzym

Tabelle 3.4.1-3: Produktüberprüfung (n=19) der PCR nach RAMACHAN-DRAN et al. (2003): Generierte Amplifikatgrößen und erwartete Größe der Restriktionsfragmente in Abhängigkeit vom ausgewählten Restriktionsenzym

Fortl. Nr.	Intimin-untereinheit	Amplifikatgröße	Erwartete Größe der Restriktionsfragmente (Bp)		
			<i>Fok I</i>	<i>Rsa I</i>	<i>Hae III</i>
1	alpha 1	849	454, 395	725, 84, 40	-*
2	alpha 2	849	279, 200, 172, 111, 89	590, 135, 84, 40	-
3	beta 1	849	669, 180	528, 246, 75	435, 214, 200
4	beta 2	849	0	321/447/790	-
5	gamma 1	834	453, 234, 83, 64	402, 432	-
6	gamma 2	834	456, 396, 147, 35	405, 354, 78	660, 174
7	epsilon 1	846	730, 116	744, 102	-
8	epsilon 2	846	640, 210, 46	744, 132	744, 132
9	eta	876	673, 170	774, 102	-
10	iota 1	843	64	525, 318	362, 307, 174
11	kappa	849	408, 261, 116, 64	231, 201, 159, 120, 84, 54	-
12	lambda	846	779, 64	441, 318, 87	589, 257
13	my	843	763, 80	525, 318	614, 229
14	ny	876	422, 279, 111, 64	774, 102	-
15	pi	845	519, 172, 159	744, 101	-
16	rho	876	454, 422	753, 123	499, 376, 23
17	sigma	871	711, 83, 61, 16	0	592, 168, 111
18	theta	834	456, 196, 147, 35	405, 354, 78	660, 174
19	zeta	846	671, 121, 54	345, 279, 135, 87	-

* schneidet das Produkt nicht

Tabelle 3.4.1-4: Differenzierung der Intiminuntereinheiten (n=19) anhand der erwarteten Fragmentgrößen (+/- mit dem Restriktionsenzym differenzierbar bzw. nicht differenzierbar)

Fortlfd. Nr.	Intiminuntereinheit	Anhand der erwarteten Fragmentgrößen eindeutig differenzierbar (+) bzw. nicht differenzierbar (-)			Eindeutig bestimmt über (n) Restriktionsenzyme
		<i>Fok</i> I	<i>Rsa</i> I	<i>Hae</i> III	
1	alpha 1	+	+	-	2
2	alpha 2	+	+	-	2
3	beta 1	+	+	+	3
4	beta 2	-	+	-	1
5	gamma 1	+	+	-	2
6	gamma 2	+	+	+	3
7	epsilon 1	+	-	-	1
8	epsilon 2	+	-	+	2
9	eta	+	-	-	1
10	iota 1	-	+	+	2
11	kappa	+	+	-	2
12	lambda	-	+	+	2
13	my	+	+	+	3
14	ny	+	+	-	2
15	pi	+	-	-	1
16	rho	+	+	+	3
17	sigma	+	-	+	2
18	theta	+	+	+	3
19	zeta	+	+	-	2
	Summe eindeutig differenzierter Intiminuntereinheiten von Gesamt (n/n _{Gesamt})	16/19	14/19	9/19	

Der alleinige Einsatz des Restriktionsenzym *Fok I* gewährleistete die Unterscheidung von 14, der der Endonuklease *Rsa I* von 16 und der von *Hae III* von neun der insgesamt 19 in die Analysen einbezogenen Intiminvarianten und der jeweils enthaltenen Subtypen (**Tabelle 3.4.1-3** und **Tabelle 3.4.1-4**). Mit der genannten Enzymkombination wurden fünf (beta 1, gamma 2, my, rho, theta) der 19 darzustellenden Untereinheiten durch die Restriktionsfragmentlängenpolymorphismen von jeweils drei, zehn Varianten durch je zwei (alpha 1, alpha 1, beta 2, gamma 1, epsilon 2, iota 1, kappa, lambda, ny, sigma, zeta) und vier Varianten durch je ein Enzym (beta 2, epsilon 1, eta, pi) eindeutig bestimmt.

Dabei ist zu berücksichtigen, dass in dieser Auflistung die Varianten „xi“ und „omikron“ (BLANCO et al., 2003 [nur in der Sequenzdatenbank publiziert] und 2004) aufgrund zum Zeitpunkt der Arbeiten fehlender Sequenzinformationen im amplifizierten Genabschnitt unberücksichtigt blieben. Dies gilt ebenso für delta-Intimin (ADU-BOBIE et al.; 1998a), das als kappa-Variante anzusprechen ist und das als iota 2-Ausprägung einzustufende my-Intimin. Gamma 2- und theta-Intimin liefern identische Schnittmuster, da es sich um identische Typen handelt (**Kap. 4.6**).

3.4.2 REAKTIONSANSATZ UND GELELEKTROPHORETISCHE AUSWERTUNG

Für jede Restriktionsendonuklease wurden Reaktionsansätze mit n=25µl Endvolumen unter Zugabe von jeweils 10µl PCR-Amplifikat eingesetzt. Als Endkonzentration der Restriktionsendonukleasen wurden 5 U/25µl-Ansatz eingesetzt. Die Einzelkomponenten wurden nach sorgfältigem Mischen der jeweiligen Ausgangslösungen bei +4°C gelagerten 2ml-Eppendorf-Gefäßen pipettiert und durch vorsichtiges Auf- und Abziehen der Eppendorfpipette gemischt. Anschließend erfolgte die Aliquotierung in 2 ml-Eppendorf-Gefäße. Der Einsatz der Pufferformulierung und die Inkubation erfolgten nach Herstellervorgaben (**Tabelle 3.4.2-1**)

Für die gelelektrophoretische Auswertung sind für die Reihenuntersuchungen 1,5%-ige-Standard- und 3%-Metaphoragarosegele zur exemplarischen Fotodokumentation verwendet worden. Das 10 bzw. 15µl umfassende Auftragsvolumen beinhaltet jeweils 1/3-Volumenanteil Bromphenolblaugebrauchslösung. Die Herstellung der Agarosegele, die Visualisierung mit Ethidiumbromid und die abschließende Fotodokumentation wurden wie zuvor beschrieben durchgeführt (3.3.4, 3.3.5).

Tabelle 3.4.2-1: Reaktionsansatz für den restriktionsenzymatischen Verdau

Mastermixkomponenten	Volumen [µl]	Inkubationsbedingungen
PCR-Produkte (ca. 340, 560 bzw. 840 Bp)	10	
SuRE/Cut Puffer M bzw. L*	2,5 (FC 1x)	
<i>Taq</i> I**	0,5 (FC 5 U/25µl)***	3 h bei 65°C
<i>Fok</i> I**	0,5 (FC 5 U/25µl)	3 h bei 37°C
<i>Rsa</i> I**	0,5 (FC 5 U/25µl)	3 h bei 37°C
<i>Hae</i> III**	0,5 (FC 5 U/25µl)	3 h bei 37°C
Aqua dest.	11, 25	
Gesamtvolumen	25	

* SuRE/Cut Puffer L für *Rsa* I

** Ausgangskonzentration 10 U/µl

*** Endkonzentration/Final Concentration (FC)

3.5 Sequenzanalyse

Die Arbeiten gliederten sich in zwei Teilabschnitte. Im ersten Arbeitsabschnitt wurden die generierten Amplifikate des *eae*-Gens der Isolate P 146 und P 256 sequenziert, die erhaltenen Rohdaten aufbereitet und mittels Sequenzdatenbankanalyse mit den bislang beschriebenen Intiminvarianten in Beziehung gebracht.

Der zweite Arbeitsabschnitt beinhaltete die vergleichende Betrachtung der unterschiedlichen Intiminvarianten zur Bestimmung der phylogenetischen Relationen. Neben der Nukleotidabfolge innerhalb der berücksichtigten *eae*-Genvarianten wurden zusätzlich die resultierenden Aminosäuresequenzen ermittelt. Für die Erstellung der Stammbäume wurden zur Erhöhung der Aussagekraft unterschiedliche Distanz-basierte-Methoden und die „Maximum-Likelihood“-Berechnung verwendet. Zur Berücksichtigung der Einflussfaktoren Rekombination, Hybridisierung und Gentransfer wurden zusätzliche Algorithmen eingesetzt, die als Distanz-basierte-Methoden die Darstellung der phylogenetischen Analysen in Form von Netzwerken ermöglichten. Im Anschluss daran wurden die erhaltenen Bäume (engl. „trees“) evaluiert.

Anhand des Sequenzvergleichs der Nukleotidsequenzen wurden Untersuchungen zur Evolution der Intiminvarianten unter Einbeziehung der konservierten und hypervariablen Genabschnitte durchgeführt. Durch die Berücksichtigung der Aminosäuresequenzen konnten zusätzlich die funktionellen Intimin-Domänen einbezogen und Aussagen über den vorliegenden Selektionsmechanismus getroffen werden. Dazu wurden die synonymen, d. h. die Mutationen der Nukleotidsequenz, die nicht zu einer Veränderung der resultierenden Aminosäuresequenz führen und nonsynonymen, d. h. die Mutationen der Nukleotidsequenz, die zu einer Veränderung der resultierenden Aminosäuresequenz führen, einschließlich der Rekombinationen und polymorphen Sequenzabschnitte innerhalb der *eae*-Gene ermittelt. Unter Bezug auf die Gensequenzen und funktionellen Domänen erfolgte im Anschluss die grafische Darstellung.

Die notwendigen Parameter sind unter Einbeziehung unterschiedlicher, für die vorliegenden Fragestellungen optimierter Softwaresysteme ermittelt worden. Zielsetzung war, die Berechnungsmodelle in Form von über das Internet oder direkten Kontakt mit den Entwicklern öffentlich frei zugänglichen Programmen nutzen zu können, um eine validierte, kostengünstige und einfach zu handhabende Alternative zu kommerziellen Programmpaketen anbieten zu können.

3.5.1 VERWENDETE COMPUTERPROGRAMME

Die eingesetzten Computerprogramme und Datenbanken werden vorab in der **Tabelle 3.5.1-1** zusammengefasst. Neben den Anwendungen für die Sequenzanalyse wird auch die Software für die Erarbeitung des universellen PCR-Typisierungsschemas aufgeführt.

Tabelle 3.5.1-1: Computerprogramme und Datenbanken unter Berücksichtigung des Verwendungszweckes

Programm	Verwendete Optionen	Bezugsquelle/Internetadresse	Referenz
Anwendungen, die direkt über das Internet betrieben werden (Online-Anwendungen)			
<i>BANKIT</i>	Übermittlung von Sequenzdaten an die NCBI®-Gendatenbanken	http://www.ncbi.nlm.nih.gov/BankIt/	Sequenzverwaltung des National Center for Biotechnology Information (NCBI), Bethesda, USA
<i>BLAST</i>	Vergleichende Nukleotid- und Proteinsequenzdatenbankabfrage	www.ncbi.nlm.nih.gov/BLAST	ALTSCHUL et al. (1990): "Basic local alignment search tool." J. Mol. Biol. <u>215</u> , 403-410.
<i>GeneScan</i>	Identifizierung kompletter Genstrukturen in genomischer DNA	http://genes.mit.edu/GENSCAN:Html	BURGE und KARLIN (1997): Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. <u>268</u> , 78-94
<i>GENOMOME</i>	Darstellung des Genoms von <i>E. coli</i> EDL 933	www.ncbi.nlm.nih.gov	Datenbanken des National Center for Biotechnology Information (NCBI), Bethesda, USA
<i>LOCUS LINK/ENTREZ GENE</i>	Darstellung der Genlokalisierung im Genom, Genaufbau	www.ncbi.nlm.nih.gov	

Fortsetzung Tabelle 3.5.1-1: Computerprogramme und Datenbanken unter Berücksichtigung des Verwendungszweckes

Programm	Verwendete Optionen	Bezugsquelle/Internetadresse	Referenz
Anwendungen, die direkt über das Internet betrieben werden (Online-Applikationen)			
<i>NUCLEOTIDE</i>	Nukleotidsequenzdatenbankabfrage	http://www.ncbi.nlm.nih.gov	Datenbanken des <u>N</u> ational <u>C</u> enter for <u>B</u> iot <u>e</u> chnology <u>I</u> nformation (NCBI), Bethesda, USA
<i>STRUCTURE</i>	Proteinstrukturdatenbank	http://www.ncbi.nlm.nih.gov	
<i>TIGR</i>	<u>C</u> omprehensive <u>M</u> icrobial <u>R</u> esource (CMR): Bakterielle Genomdatenbank	http://www.tigr.org/CMR2	PETERSON et al. (2001): The Comprehensive Microbial Resource. <i>Nuc. Acids Res.</i> <u>29</u> , 123-125.
Anwendungen, die auf lokalen Computersystemen installiert werden („Stand alone“-Applikationen)			
<i>BioEdit V7</i>	Alignment, Bearbeitung von Sequenzrohdaten	http://www.mbio.ncsu.edu/BioEdit/bioedit.html	HALL (1999): BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. <i>Nucl. Acids. Symp. Ser.</i> <u>41</u> , 95-98.
<i>Chromas V 1.43</i>	Darstellung von Sequenzrohdaten im BIN-Format	http://www.basic.nwu.edu/biotools/Chromas.html	Betrieben durch „School of Biomolecular and Biomedical Science and Technology, Griffith University, Brisbane, Queensland, Australia“

Fortsetzung Tabelle 3.5.1-1: Computerprogramme und Datenbanken unter Berücksichtigung des Verwendungszweckes

Programm	Verwendete Optionen	Bezugsquelle/Internetadresse	Referenz
Anwendungen, die auf lokalen Computersystemen installiert werden („Stand alone“-Applikationen)			
Clone Manager™	Auswahl von Restriktionsendonukleasen	http://www.scienced.com/sescat.htm	Scientific & Educational Software, 600 Pinner Weald Way Ste 202, Cary, NC 27513 USA
<i>Cn3D 4.1</i>	Darstellung von 3D-Proteinstrukturen	http://www.ncbi.nlm.nih.gov/structure/CN3/cn3d.shtml	Produziert durch das National Center for Biotechnology Information (NCBI), Bethesda, USA
<i>DnaSP.4.00</i>	Darstellung, Berechnung von Nukleotidpolymorphismen und Sequenzvariationen	http://www.ub.es/dnasp	ROZAS. und ROZAS (1995): DnaSP, DNA sequence polymorphism: an interactive program for estimating Population Genetics parameters from DNA sequence data. <i>Compu. Applic. Biosci.</i> 11 , 621-625
EditSeq™	Sequenzbearbeitung und -translation	http://www.dnastar.com	Programmpaket „Lasergene, DNASTAR Inc. Wisconsin, USA“
<i>HAPLOT</i>	Vergleichende grafische Darstellung von Sequenzpolymorphismen	http://www.shiga-tox.net/cgi-bin/stec/haplot	Betrieben durch “Microbial Evolution Laboratory der Michigan State University“
<i>HON-NEW</i>	Bewertung nonsynonymer Mutationen anhand der Veränderungen der Aminosäureeigenschaften (Ladung, Größe, Hydrophobie)	Persönliche Zusendung	ZHANG, J. (2000): Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. <i>J. Mol. Evol.</i> 50 , 56-68

Fortsetzung Tabelle 3.5.1-1: Computerprogramme und Datenbanken unter Berücksichtigung des Verwendungszweckes

Programm	Verwendete Optionen	Bezugsquelle/Internetadresse	Referenz
Anwendungen, die auf lokalen Computersystemen installiert werden („Stand alone“-Applikationen)			
MapDraw™	Auswahl von Restriktionsendonukleasen	http://www.dnastar.com	Programmpaket „Lasergene, DNASTAR Inc. Wisconsin, USA“
MEGA 3	Alignment, Analyse von Sequenzpolymorphismen, Phylogenetische Stammbäume	http://www.mega-software.net/	KUMAR, S. et al. (2004a): MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. in Bioinformatics <u>5</u> , 15-163
MegAlign™	Alignment, Phylogenetische Analyse	http://www.dnastar.com	Programmpaket „Lasergene, DNASTAR Inc. Wisconsin, USA“
MS Oligo V. 4.0	Primer-Design	http://www.oligo.net/dnasis:Htm	Programmpaket „OLIGO, Primer Analysis Software, Molecular Biology Insights Inc., USA“
PAML 3.14	Maximum Likelihood Kalkulationen	http://www.abcus.gene.ucl.ac.uk/software/	YANG (1997): PAML: a program package for phylogenetic analysis by maximum likelihood. Comp.Appl. BioSciences <u>13</u> , 555-556
Primer Express™	Primer-Design	http://www.appliedbiosystems.com	Applied Biosystems, Darmstadt, Germany
PSFIND	Markierung von Polymorphismen, Konvertierung in HAPLOT-Formate	http://www.shiga-tox.net/cgibins/stec/psfind	Betrieben durch “Microbial Evolution Laboratory der Michigan State University“
RPD V. 2.0	Detektion, Analyse von Rekombination	http://darwin.uvigo.es/rdp/rdp.html	MARTIN und RYBICKI (2000): RPD: Detection of recombination amongst aligned sequences. Bioinformatics <u>16</u> , 562-563

Fortsetzung Tabelle 3.5.1-1: Computerprogramme und Datenbanken unter Berücksichtigung des Verwendungszweckes

Programm	Verwendete Optionen	Bezugsquelle/Internetadresse	Referenz
Anwendungen, die auf lokalen Computersystemen installiert werden („Stand alone“-Applikationen)			
SCR3	Bewertung nonsynonymer Mutationen anhand der Änderungen der Aminosäureeigenschaften (z. B. Ladung, Größe)	Persönliche Zusendung	HUGHES et al. (1990): Positive Darwinian Selection Promotes Charge Profile Diversity in the Antigen-binding Cleft of Class I Major-Histocompatibility-Complex Molecules.“ Mol. Biol. Evol. <u>7</u> , 5 15-524; HUGHES und FRIEDMAN (2000): Evolutionary Diversification of Protein-Coding Genes of Hantaviruses. Mol. Biol. Evol. <u>17</u> , 1558-1568
SeqMan™	Vergleich von Sequenzrohdaten	http://www.dnastar.com	Programmpaket „Lasergene, DNASTAR Inc. Wisconsin, USA“
<i>SplitsTree</i> 3.2 <i>SplitsTree</i> 4.0b	Grafische Darstellung phylogenetischer Beziehungen	http://www-ab.informatik.uni-tuebingen.de/software/splits/welcome_en.html	HUSON (1998): SplitsTree: Analyzing and visualising evolutionary data. Bioinf. <u>14</u> , 68-73; HUSON und BRYANT (2006): Application of Phylogenetic Networks in Evolutionary Studies. Mol. Biol. Evol. <u>23</u> , 254-267, Epub 2005

3.5.2 SEQUENZIERUNG

Die Sequenzierung der mittels Qiaquick[®] PCR-Purification-Kit (Fa. Qiagen) aufgereinigten Amplifikate wurde durch die Fa. Seqlab durchgeführt. Die mittels BioEdit[®] V. 7.0.3 (HALL, 1999) editierten Rohdaten in Form von Einzelstrangsequenzen wurden mittels SeqMan[™] V. 5.0 (Fa. DNASTar) verglichen. Die Online-Software BANKIT wurde zur Übermittlung der Sequenzen an die NCBI[®]-Nukleotidbank eingesetzt.

Zum Vergleich mit den weiteren Intiminvarianten erfolgte der Transfer an BLAST (ALTSCHUL et al. 1990) der NCBI- und die TIGR-Datenbank (PETERSON et al., 2001).

Zur Festlegung des offenen Leserasters (engl. „open reading frames“) wurden neben der BLAST-Datenbankabfrage EditSeq und zur Bestätigung GeneScan eingesetzt.

3.5.3 VERWENDETE REFERENZSEQUENZEN

Für die Darstellung der phylogenetischen Relationen wurden 51 *eae*-Gen-Referenzsequenzen aus der NCBI[®]-Nukleotidsequenzdatenbank eingesetzt. Darin enthalten waren neben der Sequenz des *eae*-Gens von *Citrobacter freundii* (Genbank-Zugriffs-Nr.: L11691.1) und *E. coli* O115 auch die im Rahmen dieser Arbeiten generierten Sequenzen der Isolate O156:H⁻ (Genbank-zugriffs-Nr.: AY520904) und O84:H⁻ (Genbank-zugriffs-Nr.: AY520905). Bei der Auswahl der Sequenzen wurden nach Möglichkeit ausschließlich vollständig sequenzierte Intiminvarianten verwendet. Soweit vorhanden, wurden neben den Angaben über das Habitat bezugnehmende Veröffentlichungen berücksichtigt. Eine genaue Übersicht über die eingesetzten Isolate ist der **Tabelle 9.1.1-1** zu entnehmen. Für die Berechnung der Parameter Sequenz-Übereinstimmungsmatrix, der synonymen und nonsynonymen Mutationen, Rekombination und Selektion wurden ausschließlich vollständig sequenzierte *eae*-Gene eingesetzt.

Im Falle der beiden in den eigenen Arbeiten sequenzierten Stämme konnten 95% der Nukleotidabfolge ermittelt werden. Die fehlenden Anteile lagen im hochkonservierten 5'-Bereich. Eine Verfälschung der Ergebnisse durch fehlende Sequenzinformationen war daher nicht zu erwarten. Soweit notwendig wurden die resultierenden Proteinsequenzen aus der NCBI®-Proteindatenbank übernommen.

3.5.4 SEQUENZVERGLEICH DER INTIMINVARIANTEN

Für den Sequenzvergleich (engl. „alignment“) der Intimingene wurde der Algorithmus „ClustalW“ (THOMPSON et al., 1994) der in den Programmpaketen MEGA3 (KUMAR et al. 2004a) und BioEdit V. 7 (HALL, 1999) enthalten ist, eingesetzt. Beide Programme verwenden unterschiedliche Eingabeformate. Daher wurden die Daten für MEGA3 im Seq- (engl. „sequence“)- und für BioEdit V. 7 im Fasta (engl. „FAST-All“)-Format editiert. Für die Kalkulation des mehr als zwei Sequenzen umfassenden (engl. „multiple alignment“) Vergleichs wurden die Parameter wie folgt gesetzt: Gap opening penalties (GOP) 15 und Gap extension penalty (GEP) 6,66. Die Einflußfaktoren „GOP“ und „GEP“ dienen als Berechnungsgrundlage für den Einbau neuer Lücken bzw. die Verlängerung von Lücken während des Vergleichs zweier Sequenzen. Hohe GOP-Werte repräsentieren hohe Kosten für das Einfügen von Lücken. Die Frequenz der Lücken wird reduziert. Entsprechend führen hohe GEP-Werte zur Bevorzugung kurzer Lücken. Als Matrize (engl. „DNA Weight Matrix“) wurde „ClustalW 1.6“ verwendet. Der Parameter „Transition Weight 0,5“ dient der Gewichtung von Transitionen und Transversionen. Die Sequenzvergleichsdaten wurden anschließend manuell überprüft und ggf. editiert.

3.5.5 ABLEITUNG MOLEKULARGENETISCHER PHYLOGENIE

Die graphische Visualisierung der phylogenetischen Relationen erfolgte zunächst anhand zweier Distanz-basierter-Algorithmen. Dabei handelte es sich um die modifizierte Methode nach NEI und GOJOBORI (NEI und GOJOBORI, 1986; ZHANG et al., 1998) und das Verfahren nach TAMURA und NEI (1993). Beide Berechnungsgrundlagen fußen auf der Neighbor Joining-Methodik (SAITOU und NEI, 1987). Als dritter Ansatz wurde die „Maximum Likelihood-Berechnung, implementiert in DNAML 3.14 als Bestandteil des Programmes PAML 3.14 (YANG, 1997a) verwendet.

Die Software MEGA3 ermöglicht die Anwendung beider Distanz-Modelle anhand des zuvor erstellten multiplen Alignments. Die Analyseparameter sind in den nachfolgenden Tabellen zusammengefasst.

Tabelle 3.5.5-1: Parameter zur Berechnung der p-Distanz zur Erstellung des phylogenetischen Baumes unter Verwendung der modifizierten Methodik nach NEI und GOJOBORI (1986, [p-Distanz])

Parameter	Wert/Belegung
Anzahl der Taxa	51
Datentyp	Kodierende Nukleotidsequenzen
Phylogenetische Evaluierung	Bootstrap (1000 Wiederholungen, seed 35562), Interior Branch Test
Einbezogene Sequenzpositionen: Lücken/fehlende Daten im Alignment	Vollständiger Ausschluss
Verhältnis der Transitionen/Transversionen (R)	1,8
Einbezogene Substitutionen	pN* (nur nonsynonyme)
Substitutionsmuster der Taxa (engl. „pattern among lineages“)	Homogen
Sequenzvariationen innerhalb der Alignmentpositionen (engl. „rates among sites“)	Uniform

* pN= Anzahl nonsynonymer Substitutionen (Nd)/Anzahl möglicher nonsynonymer Substitutionen (Nd)

Tabelle 3.5.5-2: Parameter zur Erstellung des phylogenetischen Baumes nach TAMURA und NEI (1993)

Parameter	Wert/Belegung
Anzahl der Taxa	51
Datentyp	Kodierende Nukleotidsequenzen
Phylogenetische Evaluierung	Bootstrap (1000 Wiederholungen, seed 64238), Interior Branch-Test
Einbezogene Sequenzpositionen: Lücken/fehlende Daten im Alignment	Vollständiger Ausschluss
Codon Positionen	1., 2., 3. und nicht kodierende Anteile
Einbezogene Substitutionen	d*: Transitionen und Transversionen
Substitutionsmuster der Taxa (engl. „pattern among lineages“)	Homogen
Sequenzvariationen innerhalb der Alignmentpositionen (engl. „rates among sites“)	Uniform

* d= Anzahl der Transitionen und Transversionen

Das Verhältnis Transitionen/Transversionen (R) wurde nach TAMURA und NEI (1993) ermittelt. Um eine Unterbewertung synonymer Mutationen zu vermeiden, wurden ausschließlich vierfach veränderte Positionen im Alignment (engl. „4-fold-degenerated-sites“) für die Berechnung verwendet. Diese Positionen beinhalten ausschließlich synonyme Nukleotidsubstitutionen.

Zur Darstellung der Sequenzunterschiede wurden die unkorrigierten p-Distanzen der nonsynonymen Substitutionen (modifizierte Methodik nach NEI und GOJOBORI“) bzw. der Parameter „d“ (Anzahl der Transitionen und Transversionen) nach TAMURA und NEI (1993) herangezogen.

Zur Erstellung einer Sequenzübereinstimmungsmatrix (engl. „sequence-identity-matrix“) wurde der mit der Software BioEdit V 7.0 erhaltene Vergleich der Referenzsequenzen eingesetzt. Die gewonnenen Daten werden im Anhang in der **Tabelle 9.1.1-2** wiedergegeben.

Die Darstellung phylogenetischer Netzwerke wird über die Splits-Zerlegung nach BANDELT und DRESS (1992) ebenso wie über das Neighbor-Net nach BRYANT und MOULTON (2004) ermöglicht. Dabei wurden neben dem gesamten *eae*-Gen der 25 Referenzsequenzen, der konservierte - (bis zur Position 2112) und der hypervariable Sequenzbereich (ab der Position 2113) in die Analysen einbezogen. Die Berechnungsgrundlagen sind Bestandteil der Software Splits Tree 3.2 (HUSON, 1997) bzw. SplitTree 4beta (HUSON und BRYANT, 2006). Das mit BioEdit V. 7 erstellte Alignment im FASTA-Format konnte mit der aktualisierten Version direkt eingelesen werden. Eine vorhergehende Umwandlung in den auch als Ausgabekodierung verwendeten „NEXUS“-Dateityp war damit nicht notwendig. Ebenso konnte die Neighbor-Net-Berechnung nach BRYANT und MOULTON (2004) eingesetzt werden. Da es sich um eine fortlaufend aktualisierte und in der Entwicklung befindliche beta-Version handelte, wurden die Ergebnisse mit der zum Zeitpunkt der Berechnungen aktuellen Version SplitsTree 3.2. überprüft. Um die Übersichtlichkeit der Abbildungen zu erhöhen, wurden die Genbankzugriffsnummern durch die Bezeichnungen der Intimintypen ersetzt.

Die statistische Evaluierung der erhaltenen Bäume erfolgte für die Distanzbasierten Verfahren, einschließlich der Split-Zerlegung nach BANDELT und DRESS (1992) sowie HUGHSON 1998 sowie der Neighbor-Net-Darstellung nach BRYANT und MOULTON (2004), anhand des Bootstrapping-Testes (EFRON, 1979; FELSENSTEIN 1985). Ergänzend wurde Darstellung der Baumdiagramme der als t-Test konzipierte „Interior Branch Test“ (NEI et al., 1985) eingesetzt. Ausgehend von den Astlängen und der dazugehörigen Standardabweichung (s) wurde anhand einer Bootstrap-Prozedur die Signifikanz berechnet. Werte größer oder gleich 95% wurden als signifikant positiv zu bewerten. Es wurden jeweils 1000 Wiederholungen zur Kalkulation der Pseudo-Bootstrap-Bäume eingesetzt.

Für die Überprüfung der „Maximum Likelihood-Kalkulation“ kam der χ^2 -Test zum Einsatz. Dabei wurden als signifikant positive Relationen Sequenzzuordnungen mit $p < 0,01$ herangezogen.

3.5.6 ABLEITUNG POPULATIONSGENETISCHER PARAMETER

Zur näheren Charakterisierung der genetischen Variationen wurden weitere populationsgenetische Parameter abgeleitet. Dazu gehörte zunächst eine Charakterisierung des als Berechnungsrundlage verwendeten Alignments mit $n=25$ Referenzsequenzen. Die Software DnaSP (DNA Sequence Polymorphism V. 4.00.5, ROZAS et al. 2004) verwendete hierzu das mit BioEdit V. 7 berechnete Alignment im FASTA-Format. Über eine interaktive Benutzeroberfläche wurden zunächst die kodierenden Sequenzbereiche für das gesamte Gen (Position 1 bis 2858), den konservierten 5'-Bereich (Position 1 bis 2112) und die hypervariable 3'-Region (Position 2113 bis 2858) festgelegt. Ferner wurden die resultierenden funktionellen Regionen unter Anwendung der Einteilung nach McGRAW et al. (1999) für die konservierte Zentraldomäne (CD) und die periplasmatische - (PP)-, transmembranöse - (TM)- und die vier extrazellulären Domänen D0 bis D3 nach LUO et al. (2000) berücksichtigt. Um aussagekräftige Ergebnisse für den gesamten Genbereich zu erzielen, sind ausschließlich vollständig sequenzierte Intimingene berücksichtigt worden. Zur weiteren Bearbeitung erfolgte eine Konvertierung in das Text (txt)-Format. Zur Charakterisierung des Datenpools wurden neben der Anzahl berücksichtigter Alignmentpositionen (abzgl. der Lücken), Anzahl variierender Positionen (S), Gesamtzahl der Mutationen (Eta) und der G+C-Gehalt der jeweiligen Sequenzabschnitte ermittelt. Eine detaillierte Aufschlüsselung der ermittelten Mutationen konnte durch die Hinzuziehung der „Singleton Polymorphic Sites (SPM), als der Anzahl der Positionen im Alignment mit mindestens 2, 3 oder 4 identischen Nukleotiden, die zu keiner Änderung der Aminosäurekodierung führen“ erreicht werden. Zusätzlich wurden die „Parsimony-Informative Sites (PIPM), als Anzahl der Positionen im Alignment mit mindestens 2, 3 oder 4

identischen Nukleotiden, die zur Änderung der Aminosäurencodierung führen, hinzugezogen. Zur Veranschaulichung wurden die Sequenzpolymorphismen graphisch unter Berücksichtigung der funktionellen Genregionen aufbereitet. Zur Aufbereitung des Sequenzvergleichs kamen die Applikationen PSFIND und HAPLOT zum Einsatz. PSFIND diente zur Konvertierung der Dateiformate. HAPLOT erzeugt eine „Postscript-Datei“, die anschließend in Textdokumente eingefügt werden kann.

Zur Ermittlung des Selektionsmechanismus dienen die Variablen p_S oder K_S als Quotient aus der Anzahl synonyme Substitutionen und den synonymen Alignmentposition und p_N oder K_A als Quotient der Anzahl nonsynonymer Substitutionen und den nonsynonymen Alignmentposition. Die Kalkulation der Parameter erfolgte vergleichend mit dem modifizierten Algorithmus nach NEI und GOJOBORI (1986) unter Berücksichtigung der Standardabweichung (s) über MEGA3. Beide Werte repräsentieren unkorrigierte p -Distanzen. Um Veränderungen über den gesamten Genbereich detailliert darstellen zu können, wurde neben der Berechnung für die gesamten Genabschnitte zusätzlich eine sog. „sliding window“-Analyse vorgenommen. Dazu wurde anhand eines Leseraster von 100 Basen eine abschnittsweise Analyse des Alignments eingesetzt. Lücken innerhalb des Sequenzvergleichs blieben unberücksichtigt. Exemplarisch erfolgte die graphische Darstellung der Parameter p_S (K_S), p_N (K_A) als Punkt-Diagramm für zeta (ζ)-Intimin mit Micrococcal™ Origin™ V. 5.0. Diese Vorgehensweise ist der statistischen Bewertung des adaptiven Selektionsmechanismus z. B. anhand eines Z-Testes für große Datenmengen überlegen, da die Analyse abschnittsweise erfolgen kann. Schwankungen des Verhältnisses von p_N/p_S können berücksichtigt werden. Die Kalkulation der Distanzen erfolgte über die modifizierte Methode nach NEI und GOJOBORI (NEI und GOJOBORI, 1986, ZHANG et al., 1998). Zur Berechnung der Standardabweichung (s) kam das Bootstrap-Verfahren mit 500 Wiederholungen zum Einsatz. Der für die festgelegten funktionellen Domänen ermittelte Selektionsmechanismus wurde mit dem Codon-basierten exakten Test auf Unabhängigkeit (FISHER, 1973) überprüft.

Bei der Manifestation von nonsynonymen Mutationen ist der mögliche Einfluss der Veränderungen der Aminosäuresequenz auf die Eigenschaften des Proteins von Bedeutung. Zu diesem Zweck erfolgte eine weitere Bewertung der nonsynonymen Positionen als p_{NR} und p_{NC} . Radikale nonsynonyme Mutationen (p_{NR}) beinhalten Aminosäuren, die zu einer Veränderung der Eigenschaften im Bezug auf die Ladung, die Größe und die hydrophoben Eigenschaften im Vergleich zur Ursprungssequenz führen. Hingegen bewirken konservierte nonsynonyme Mutationen (p_{NC}) die Synthese von Aminosäuren mit identischen Eigenschaften in Relation zur Ausgangssequenz. Zur Kalkulation der Kenngrößen standen die Anwendungen SCR3 (HUGHES et al., 1990, HUGHES und FRIEDMAN, 2000) und HON-NEW (ZHANG, 2000) zur Verfügung. Beide Programme wurden freundlicherweise durch die Prof. A. L. HUGHES (Department of Biology, 208 Mueller Laboratory, The Pennsylvania State University, University Park, Pennsylvania 16802) und J. ZHANG (Department of Ecology and Evolutionary Biology, University of Michigan, 3003 Natural Science Building, 830 North University Ave., Ann Arbor, MI 48109) als Windows NT-Versionen zur Verfügung gestellt. Die auf Basis der modifizierten Methode nach NEI und GOJOBORI (NEI und GOJOBORI, 1986; ZHANG et al., 1998) durchgeführten Kalkulationen berücksichtigen im Gegensatz zur Erstentwicklung durch HUGHES et al. (1990) die Möglichkeit transitionaler Abhängigkeiten durch das tatsächliche Verhältnis „R“ der Transitionen und Transversionen. Die Berechnung des Parameters „R“ erfolgte wie zuvor beschrieben (**Kap. 3.5.5**). Als Eingabeformat diente der Sequenzvergleich im „FASTA-“ (SCR3) und „SEQ-Format“ (HON-NEW). Microsoft® Excel 2002 diente zur Präsentation der Daten, die den Anhangstabellen (**Tabelle 9.1.2-1 bis 4**) beigefügt wurden.

Der Nachweis genetischer Rekombination erfolgte in zwei Schritten. Zunächst wurde die ζ -Intiminreferenzsequenz anhand des mit 25 Sequenzen erstellten Alignments mit verschiedenen Algorithmen als globale, d. h. alle Sequenzen einschließende Analyse, auf das Vorliegen rekombinanter Regionen oder sog. Bruchpunkte, die als möglicher Grenzpunkt eines rekombinanten Sequenzbereiches dienen, untersucht. Dabei handelte es sich im einzelnen um

den Maximum Chi-Quadrat Test (SMITH, 1992; POSADA und CRANDALL, 2001a), das Chimaera-Verfahren (POSADA und CRANDALL, 2001b), die SiScan (Sister Scanning)-Methode (GIBBS et al., 2000), das RPD (Recombination Detection Programm)-Verfahren (MARTIN und RYBICKI, 2000), den BOOTSCAN-Test (SALMINEN et al., 1995) und die GENECONV (Gene Conversion)-Methode (PADIDAM et al., 1999). Zielsetzung war die Festlegung rekombinanter Bereiche bei gleichzeitiger Ermittlung möglicher Vorläufersequenzen in der Entwicklung des ζ -Intimings. Dabei wurden nur Ergebnisse berücksichtigt, die mit mindestens drei Verfahren in einer wandernden Fensteranalyse von 20 Codons generiert wurden. Als Cut off-Wert diente $p > 0,05$. Die p-Wertkalkulation erfolgte anhand der Bonferroni-Korrektur (<http://darwin.uvigo.es/rdp/rdp.html>) unter der Annahme, dass 5% aller Wahrscheinlichkeiten zufällig ermittelt wurden.

In einem zweiten Schritt erfolgte eine sensitivere Untersuchung des ζ -Intimings anhand des Maximum Chi Quadrat-Testes (SMITH, 1992; POSADA und CRANDALL, 2001a) gegen die manuell als Referenzsequenzen ausgewählten α -, β -, γ - und δ -Intimingsene. ζ -Intimin wurde als potentielle Tochtersequenz gegen jeweils zwei der genannten möglichen Vorläufersequenzen in einer sog. Triplettanalyse mit der Signifikanz $p > 0,01$ untersucht. Zur Analyse wurde eine variable Fenstergröße gewählt, um eine Anpassung an die von der Berechnung ausgenommenen monomorphen Positionen zu ermöglichen. Als Ergebnis werden sog. Bruchpunkte „ k_{Max} “ ermittelt, die mögliche Anfangs- oder Endpunkte genetischer Rekombination aufzeigen. Der MaxChi-Test und die zuvor genannten Algorithmen sind Bestandteile der Software RPD (Recombination Detection Program) V. 2.0. Als Eingabeformat wurden mit BioEdit V. 7 erstellte Alignierungen der phylogenetischen Untergruppen im „FASTA-Format“ eingelesen.

Die ermittelten populationsgenetischen Parameter einschließlich des Verwendungszweckes und der jeweiligen Software wurden in der **Tabelle 3.5.6-1** zusammengestellt.

Tabelle 3.5.6-1: Abgeleitete populationsgenetische Parameter

Ermittelte Parameter		Eingesetzte(s) Computerpro- gramm(e)
Charakterisierung der Alignmentdaten		
Anzahl berücksichtigter Alignmentpositionen (abzgl. der Lücken)		DnaSP V 4.00.5
S	Anzahl variierender Positionen	
Eta	Gesamtzahl der Mutationen	
Guanin- und Cytosin (G+C)-Gehalt		
SPM (2)	Singleton Polymorphic Sites, Anzahl der Positionen im Alignment mit mindestens 2 identischen Nukleotiden, die zu <u>keiner</u> Änderung der Aminosäurekodierung führen	DnaSP V 4.00.5
PIPM (2)	Parsimony-Informative Sites, Anzahl der Positionen im Alignment mit mindestens 2 identischen Nukleotiden, die zur Änderung der Aminosäurekodierung führen	
SPM (3)	Siehe SPM (2)	
PIPM (3)	Siehe PIPM (2)	
SPM (4)	Siehe SPM (2)	
PIPM (4)	Siehe PIPM (2)	
Ermittlung des Selektionsmechanismus (I)		
p_s, K_s	Anzahl synonyme Substitutionen/synonyme Alignmentposition	MEGA3 unter Verwendung der modifizierte Methode nach NEI und GOJOBORI (NEI und GOJOBORI, 1986; ZHANG et al., 1998)
p_N, K_a	Anzahl nonsynonymer Substitutionen/nonsynonyme Alignmentposition	
$p_N=p_s$	Neutrale Selektion	abschnittsweise Analyse mit MEGA3
$p_N < p_s$	Negative Selektion	
$p_N > p_s$	Positive Selektion	

Fortsetzung Tabelle 3.5.6-1: Abgeleitete populationsgenetische Parameter

Ermittelte Parameter		Eingesetzte(s) Computerpro- gramm(e)
Ermittlung des Selektionsmechanismus (II)		
p_{NC}	Konservierte nonsynonyme Substitutionen	SCR3, HON-NEW
p_{NR}	Radikale nonsynonyme Substitutionen	
$p_{NC}=p_{NR}$	Neutrale Selektion	
$p_{NC}<p_{NR}$	Positive Selektion	
$p_{NC}>p_{NR}$	Negative Selektion	
Rekombination		
k_{max}	Lokalisation innerhalb eines Alignments mit der höchsten Wahrscheinlichkeit ($p>0,05$ bzw. $p>0,01$) der Rekombination	RPD V. 2.0

3.6 Zelladhäsion

Für die Darstellung der unspezifischen, fimbrienvermittelten und der spezifischen, intiminvermittelten Adhäsion des Referenzstammes EPEC E 2348/69 O127:H6 wurden Zellversuche mit humanen Mundschleimhautzellen und der Fluorescent-Actin-Staining-(FAS)-Test in Anlehnung an KNUTTON et al. (1987 und 1989) durchgeführt. Der FAS-Test wurde in den Räumen und unter Verwendung der Einrichtung des Institutes für Hygiene und Infektionskrankheiten der Tiere, JLU Gießen, durchgeführt.

Der Ablauf der Adhäsionsversuche mit Mundschleimhautzellen gliederte sich wie folgt:

Anzüchtung der Bakterien:

Übernachtskultur in Luria Bertani (LB)*- Bouillon

*Wahl des Mediums zur Stressexposition der Bakterien im Vergleich zur Brain-Heart-Infusion-Bouillon (BHI)-Bouillon

Überführung von 1 ml Bakteriensuspension in LB- Bouillon und Herstellung einer 1:40 Verdünnung, Schüttelinkubation bei 37°C für 2 h

Bakterienfärbung:

1 ml Suspension mit 100µl Fluoresceinisothicyanatlösung (FITC) (1:10 verdünnt) im Schüttelinkubator bei 37°C für 1h

Gewinnung der Mundschleimhautzellen mittels Spatel und Auswaschen in 10 ml Dulbecco´s modified Eagle Medium (Fa. GIBCO, Produkt-Nr. 31855-023)

Abzentrifugieren der Zellen bei 27°C und 1500 pm für 10 min

2 Waschschritte:

Dekantieren des Überstandes, Resuspension in 10 ml Dulbecco´s modified Eagle Medium, erneute Zentrifugation der Zellen

Abzentrifugation der gefärbten Bakterien bei 800 rpm für 3 min

Dekantieren des Überstandes und Resuspension in 1 ml physiologischer Kochsalzlösung

Zweimalige Wiederholung der letzten beiden Schritte

zuletzt: Resuspension des Bakterienpellets in 200µl physiologischer Kochsalzlösung

Eigentlicher Adhärenzversuch:

Zugabe von 200 ml Zell- zur Bakteriensuspension

Schüttelinkubation für 3 h bei 37°C

Abtrennung der ungebundenen Bakterien:

Zugabe des gesamten Ansatzvolumens auf 1 ml 1:50 verdünntes Percoll

Fluoreszenzmikroskopische Auswertung mittels Ölimmersion

Die Auswertung erfolgte durch Auszählung von mindestens 5 Zellen/Objektträger. Die Auswertung erfolgte zusätzlich nach 4, 5 und 6 h, ein Einsatz von 5, 6, bzw. 8 Stunden in LB-Bouillon inkubierten Bakterien wurde durchgeführt.

4 ERGEBNISSE

4.1 NACHWEIS UND TYPISIERUNG DES *EAE*-GENS

4.1.1 AMPLIFIKATION DES KONSERVIERTEN 5'-BEREICHS

Mit der Primer-Kombination „SK1“ und „SK2“ nach KARCH et al. (1993) konnten alle 251 *Escherichia coli*-Isolate als „*eae*-positiv“ klassifiziert werden. Damit bestätigten sich die im Rahmen der Charakterisierung der Isolate bei Aufnahme in die institutsinterne *Escherichia coli*-Stammsammlung ermittelten und hasuintern hinterlegten Ergebnisse. Mit den genannten Oligonukleotiden wurde ein 881 Bp umfassendes PCR-Produkt generiert. Die **Abbildung 4.1.1-1** verdeutlicht exemplarisch die Spezifität der PCR-Amplifikation. **Abbildung 4.1.1-2** zeigt die Selektivität des Primer-Paares im Bezug auf die bislang beschriebenen Sequenzen der Intiminvarianten alpha (α) bis sigma (σ) inklusive der in den eigenen Untersuchungen als zeta (ζ)- Intiminträger sequenzierten Isolate (mit ζ 1 HS und ζ 2 HS gekennzeichnet). Beide Oligonukleotide beinhalten am 5'-Ende Zielgen-unspezifische Anteile (in **Abbildung 4.1.1-2** mit unterbrochenen Linien hervorgehoben). So findet sich eine vollständige Übereinstimmung der gesamten Sequenz für „SK1“ nur bei zwei epsilon-Intiminträgern. Betrachtet man die spezifische Bindung an die Zielregion, unabhängig von den Intiminvarianten, sind bei „SK1“ nur die ersten 15 bzw. bei „SK2“ die ersten 16 Basen für diesen Vorgang verantwortlich.

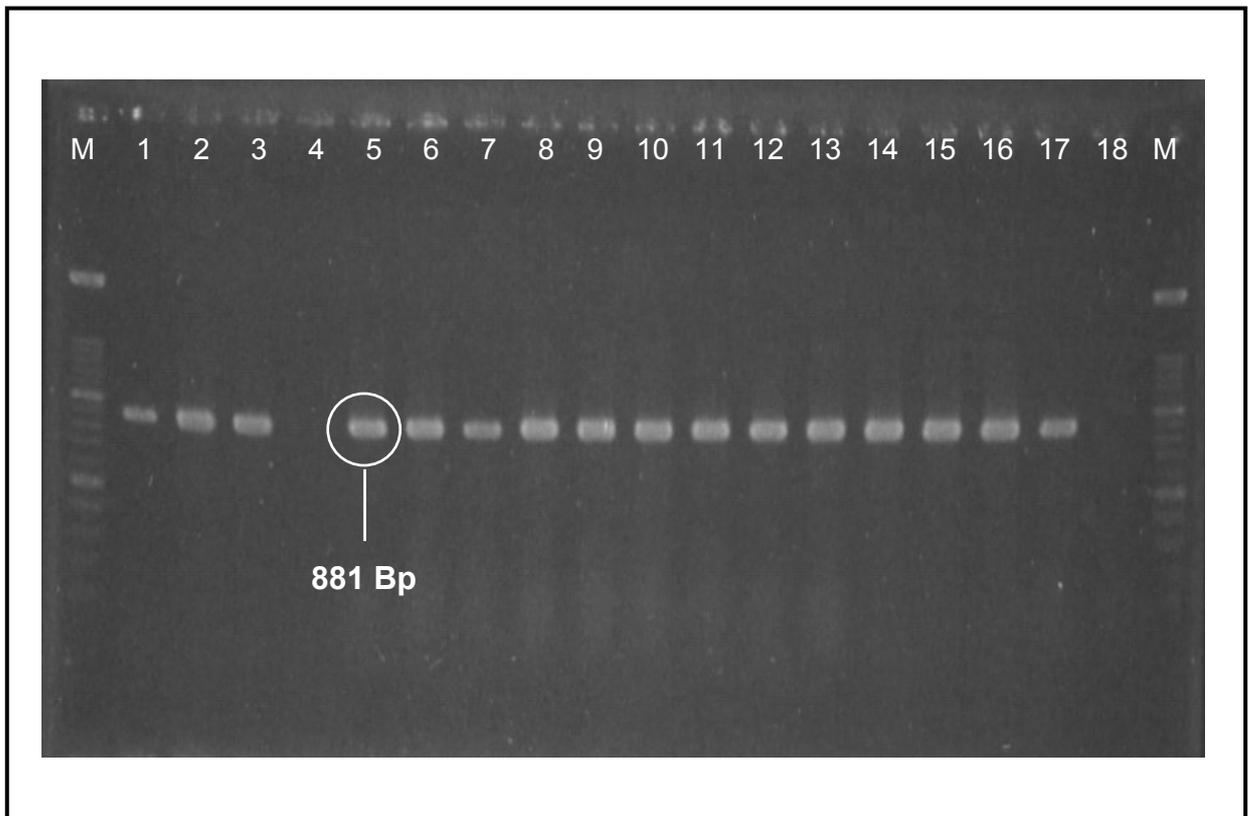


Abbildung 4.1.1-1: Beispielhafte Darstellung der Amplifikation der konservierten Region des *eae*-Gens mit den Oligonukleotid-Primern „SK1/SK2“ nach KARCH et al. (1993) und SCHMIDT et al. (1994). Spur 4 beinhaltet den negativen Kontrollstamm *E. coli* ATCC 25922 und Spur 18 den Leerwert. Als Längenstandard diente der Marker (M) XIV.

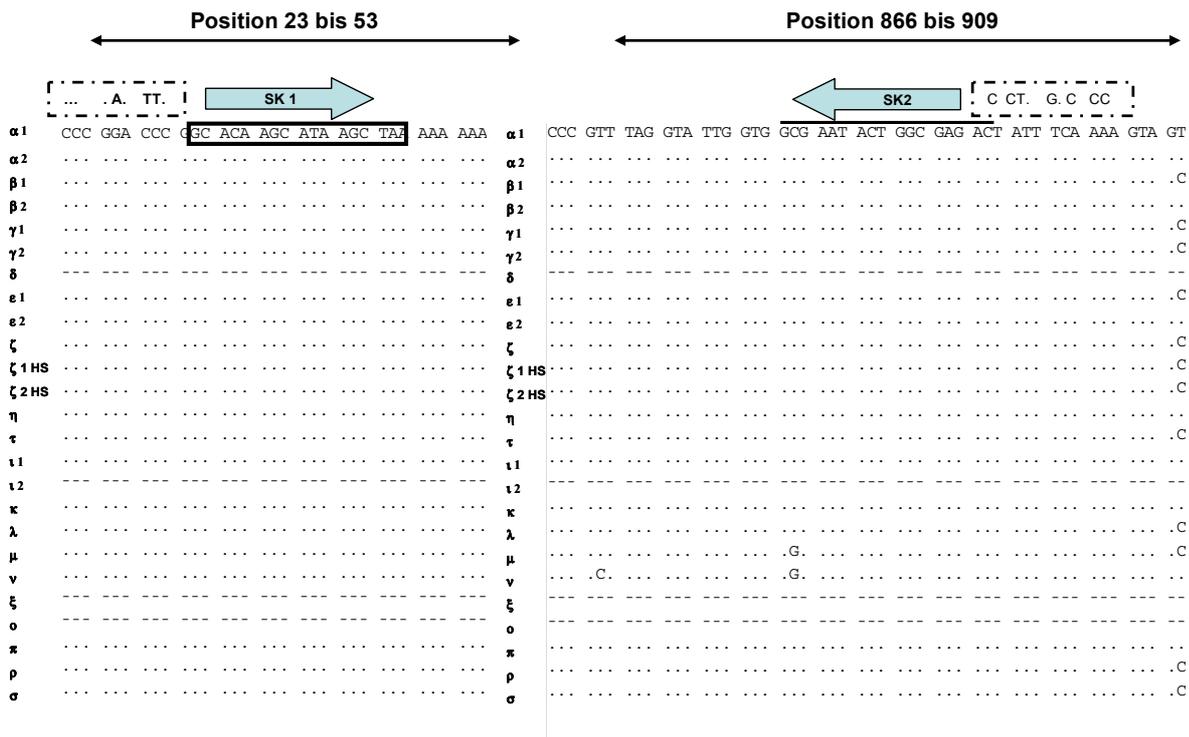


Abbildung 4.1.1-2: Lokalisation der Oligonukleotid-Primer „SK1/SK2“ nach KARCH et al. (1993) und SCHMIDT et al. (1994) innerhalb des Alignments der bislang beschriebenen Intiminvarianten alpha (α) bis sigma (σ). Die Ziffern hinter den Intimintypen (1 und 2) kennzeichnen die jeweilige Untereinheit. Punkte symbolisieren homologe Sequenzbereiche, Bindestriche repräsentieren Lücken im Alignment. In durchgezogenen Linien sind die sequenzspezifischen Anteile der Primer eingefasst.

4.1.2 SEQUENZANALYSE ZUR TYPISIERUNG ANHAND DES HYPERVARIABLEN 3'-BEREICHES

In den eigenen Untersuchungen wurde zunächst die Sequenzanalyse der unterschiedlichen Oligonukleotid-Primerpaare anhand der in **Tabelle 9.1.1-1** gelisteten Referenzstämme der oben genannten Intimingene vorgenommen. Anschließend erfolgte die Überprüfung der Spezifität im PCR-Einsatz. Nach Vergleich der dabei erzielten Ergebnisse wurde das um ein für epsilon (ϵ)-Intimin spezifisches Primerpaar erweiterte PCR-Typisierungsverfahren nach ADU-BOBIE et al. (1998a) für die weiteren Screening-Untersuchungen eingesetzt.

Eine Bewertung und Einstufung der verwendeten Nomenklatur wird unter „**Kap. 4.6**“ vorgenommen.

ADU-BOBIE et al. (1998a) verwendeten zum Nachweis von alpha (α)-, beta (β)-, gamma (γ)- und delta (δ)-Intimin den Universalrückwärts-Primer „Int-Ru“ in Kombination mit den jeweils spezifischen Vorwärts-Primern „Int-A, Int-B, Int-Y“ und „Int-D“. In **Abbildung 4.1.2-1** und **Abbildung 4.1.2-2** ist die Lokalisation der Primer im Alignment aller bekannten Intiminuntereinheiten dargestellt. Dabei sind oberhalb der eingefügten Trennlinie die zum Zeitpunkt der PCR-Systementwicklung vorliegenden Sequenzinformationen abgebildet. Alle Primer liegen in der ab Position 2112 (ADU-BOBIE et al. 1998a; ZHANG et al., 2002) beginnenden hypervariablen Region des *eae*-Gens: Der Rückwärts-Primer „Int-Ru“ vollständig, bei den Vorwärts-Primern nur das jeweilige 3'-Ende. Kreuzreaktionen waren in der mittels NCBI®-BLAST durchgeführten vergleichenden Datenbankabfrage nicht feststellbar. Der Vorwärts-Primer „Int-B“ ermöglichte auch die Detektion des nachfolgend durch OSWALD et al. (2000) beschriebenen β 2-Intiminsubtyps. Der Rückwärts-Primer „Int-RU“ zeigte bei dieser Untereinheit allerdings eine Fehlbasenpaarung (engl. „mismatch“) an Position 2 des 3'-Endes („A“ ersetzt durch „G“).

Um ϵ -Intiminträger, erstmalig bei O103-Serovaren identifiziert, typisieren zu können, wurde das spezifische Primer-Paar epsilonF/epsilonR auf Basis der in den **Abbildung 4.1.2-3** und **Abbildung 4.1.2-4** dargestellten Sequenzvergleiche abgeleitet. Die Alternative, einen spezifischen Vorwärts-Primer in Kombination mit „Int-Ru“ zu etablieren, wurde aufgrund des an erster Position des 3'-Endes bestehenden Mismatches („A“ anstatt „T“) nicht in Betracht gezogen.

Die Primer-Kombination erwies sich in der Sequenzdatenbankabfrage als ϵ -Intimin-spezifisch und ermöglichte auch die Amplifikation der von RAMACHANDRAN et al. (2003) bei O2:H19 (Genbank-Zugriffs-Nr AF530554) dargestellten ϵ_2 -Intiminuntereinheit. Kreuzreaktionen mit den beschriebenen neuen Intiminvarianten waren anhand des Sequenzvergleichs nicht zu erwarten und traten im Versuch nicht auf.

OSWALD et al. (2000) kombinierten aus vorangegangenen Arbeiten bekannte Oligonukleotidpaarungen mit neu entwickelten Primern. Als universeller Vorwärts-Primer diente der im konservierten *eae*-Genabschnitt lokalisierte und zum Nachweis des Vorhandenseins des Intimin-Gens verwendete „SK1“-Primer. Die von SCHMIDT et al. (1993) bereits in Kombination mit dem Primer „LP1“ verwendeten Oligonukleotide „LP2“ und „LP3“ (α - bzw. γ -Intimin) wurden um „LP4“ und „LP5“ (β - bzw. ϵ -Intimin) erweitert. Die resultierenden Amplifikate rangieren zwischen 2608 Bp für ϵ - und 2807 Bp bei α -Varianten. Mit „LP2“ und „LP3“ wurde der gesamte hypervariable Bereich einbezogen (**Abbildung 4.1.2-5**, **Abbildung 4.1.2-6**, **Abbildung 4.1.2-7**). In der **Abbildung 4.1.2-7** sind die außerhalb der codierenden Region des *eae*-Gens gelegenen Sequenzanteile am 5'-Ende von „LP2“ und „LP3“ durch gepunktete Umrandungen hervorgehoben. Der anschließend durchgeführte restriktionsenzymatische Verdau erlaubt die Differenzierung von β_1 - und β_2 -Subtypen. Im BLAST-Vergleich waren keine Kreuzreaktionen feststellbar.

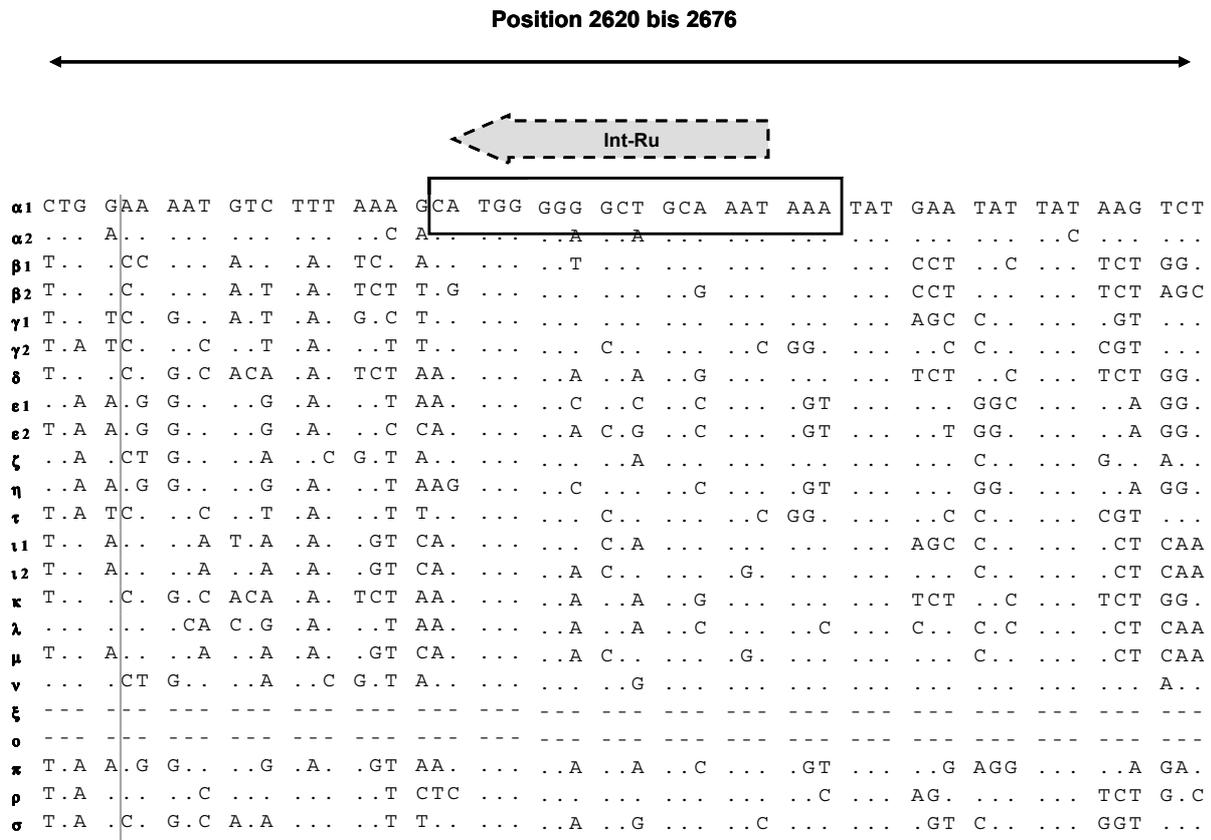


Abbildung 4.1.2-2: Lokalisation des Universalrückwärts-Primers „Int-Ru“ nach ADU-BOBIE et al. (1998a) innerhalb des Sequenzvergleichs der bislang beschriebenen Intiminvarianten alpha (α) bis sigma (σ). Zum Zeitpunkt der Veröffentlichung der Primer-Systeme waren nur die 4 bzw. 5 Intiminvarianten oberhalb des Trennstriches bekannt. Die Ziffern hinter den Intimintypen (1 und 2) kennzeichnen die jeweilige Untereinheit. Punkte symbolisieren homologe Sequenzbereiche, Bindestriche repräsentieren Lücken im Sequenzvergleich.

Position 2686 bis 2742

$\alpha 1$	ATA	ATT	TCA	TGG	GTA	CAA	CAA	ACA	GCT	CAA	GAT	GCG	AAG	AGT	GGT	GTT	GCA	AGT	ACA
$\alpha 2$GC	..T	A.A	GTG	...	AT.	..CA
$\beta 1$	T.G	.C.	G.T	...	A.T	A..	...	T.T	T..	TCT	..A	CA.	TCA	TCAA	T..	..C	...
$\beta 2$	T.G	.C.	A.T	...	A.T	A..	...	T.C	T..	TCT	..A	CTC	TCA	TCAA	T..
$\gamma 1$C.	G.T	...	A.T	A..	..G	...	T..	AGT	..G	CA.	CGT	TC.	..A	..A	T..	..C	..T
$\gamma 2$CA	G.T	...	A.T	AC.	...	T.T	.AA	GCT	...	AAA	.TA	TCA	..A	..A	T..	..C	..T
δ	T.G	.CA	G.T	...	A.T	AC.	...	T.C	T..	TCT	..G	TTA	CCA	TCGA	T..
$\epsilon 1$CA	G..	...	AC.	..GT	.AG	G.T	...	AAA	C.A	..AA	...	TGG	A.T
$\epsilon 2$CA	G..	...	AC.	..GA	G.T	...	ATA	CCT	..AA	...	TGG	A.T
ζ	..G	..A	..G	...	A..	A..A	..G	...	AT.	.G.	CAG	..A
ηCA	G..	...	ACTT	..G	GCT	...	AAA	C.A	TC.	...	TGG	A.T
ϑCA	G.T	...	A.T	AC.T	.AA	GCT	...	AAA	.TA	TCA	..A	..A	T..	..C	..T
$\iota 1$.AT	GC.	.GG	ATA	CAG	...	ACG	GAG	CAG	G.T	A.A	AAA	G.C	..C	.TC	.CA	A.G	.CA	TAT
$\iota 2$.AC	GC.	.GG	ATA	CAG	...	---	---	---	---	---	---	---	---	---	---	---	---	---
κ	T.G	.CA	G.T	...	A.T	AC.	...	T.C	T..	TCT	..G	TTA	CCA	TCGA	T..
λ	.C.	GCA	.GG	AC.	T.G	...	ACT	.GT	.A.	G.T	.TG	AA.	..A	G.G	.TC	ACC	AGT	.C.	TAC
μ	.AC	GC.	.GG	ATA	CAG	...	ACT	GAT	AAG	G.T	A.A	AGA	G.A	G.C	.TC	.CA	A.T	.CA	TAT
νA	..G	...	A..	A..C	.A.	G.TT.AG	..CC	...
ξ	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
\omicron	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
π	...	GCA	G..	...	AC.T	.AG	G..	...	AAA	..A	GCA	...	TGG	A.T
ρCA	G..	...	A.T	T.T	.G.	TC.	...	T.T	.GC	G..	..A	..G	A.T	.CA	...
σ	...	CCG	G.C	...	A.T	A..C	.AA	AGT	...	AAA	..T	TC.	..A	..A	T..	..CC	..T

Abbildung 4.1.2-4: Lokalisation des abgeleiteten Rückwärts-Primers „epsilonR“ nach SCHÖNENBRÜCHER et al. (2002) innerhalb des Sequenzvergleichs der bislang beschriebenen Intimvarianten alpha (α) bis sigma (σ). Zum Zeitpunkt der Veröffentlichung der Primer-Systeme waren nur die 4 bzw. 5 Intimvarianten oberhalb des Trennstriches bekannt. Die Ziffern hinter den Intimtypen (1 und 2) kennzeichnen die jeweilige Untereinheit. Punkte symbolisieren homologe Sequenzbereiche, Bindestriche repräsentieren Lücken im Sequenzvergleich.

Position 2571 bis 2624



```

α1 GAA TAC ATG TAA GAA TTT TGG AGG AAA GTT GCC GTC TTC TCA GAA TGA ACT GGA
α2 ... ..T ... G.. ... ..G A.. A.. C.. ... ..G. ... ..A.
β1 A.. C.A ... ..C AAA .A. C.C ... T.. AG. ACA G.. AA. AG. ACT .T. ..C
β2 A.. .A ... CC. ..C AAA ... ..C .GC T.. AG. ACA G.. AA. AG. AAT .T. ..C
γ1 C.T .TG --- --- --C AAA AAT TT- --- A.. A.. A.. CA. A.. ..C G.T .T. .TC
γ2 CTT .TG --- --- --T GGA AAT TT- --- A.. A.. A.. C.. A.. ..C G.T .T. ATC
δ A.. C.A G.. CCG A.C AA. .A. T.C ... T.. AG. AC. G.. AA. ..G .AT .T. ..C
ε1 A.C A.A ... .G. A.T GAA .A. T.. TCG T.. A.. ... A.. .AC A.G ... GT. AA.
ε2 .GC A.. ... ..G A.T GAA .AA T.C .CG C.. ... ..A.. .AC ..G ... G.. AA.
-----
ζ ..G .T. ... CC. ..G C.. GA. C.. G.G ... ..A.. ... ..A.. A.. ... A.C
η AGC A.. ... C.G C.C GAA .A. T.. TCG C.. A.. ... A.. .AC ..G ... G.. AA.
τ CTT .TG --- --- --T GGA AAT TT- --- A.. A.. A.. C.. A.. ..C G.T .T. ATC
ι1 TGG ..T ... .TC A.. .AA ..C T.C .T. T.. AT. TGT C.. AGT AG. CAG TT. .A.
ι2 T.G ..T ... .TT AGG AAA .A. T.C TC. T.. AT. TG. C.. AGT TG. CAC TT. .A.
κ A.. C.A G.. CCG A.C AA. .A. T.C ... T.. AG. AC. G.. AA. ..G .AT .T. ..C
λ TTC ..T T.. CGG A.. AA. .AA T.. CTC ... A.. A.. ... AAT TGC A.. G.. ...
μ T.G ..T ... .TT AGG AAA .A. T.C TC. T.. AT. TG. C.. AGT TG. CAC TT. .A.
ν ..G .T. ... CC. ..G C.. GA. ... ..G ... A.. A.. ... ..A.G ... ..C
ξ
o
π A.. A.. G.. ..G T.T .AA .A. T.. TCG T.. A.. ..T A.. .AC ... ..GT. AA.
ρ C.. GCA G.. CC. A.C AAA AAA T.. ..G AC. A.. A.. ... A.. .CG ... .T. A..
σ C.T .TG --- --- --T AAA AAT TC- --- T.. A.. A.. C.. A.. ... G.T .T. A.C

```

Abbildung 4.1.2-6: Lokalisation des Rückwärts-Primers „LP5“ zum Nachweis von ϵ -Intimin nach OSWALD et al. (2000) innerhalb des Sequenzvergleichs der bislang beschriebenen Intiminvarianten alpha (α) bis sigma (σ). Zum Zeitpunkt der Veröffentlichung der Primer-Systeme waren nur die 4 bzw. 5 Intiminvarianten oberhalb des Trennstriches bekannt. Die Ziffern hinter den Intimintypen (1 und 2) kennzeichnen die jeweilige Untereinheit. Punkte symbolisieren homologe Sequenzbereiche, Bindestriche repräsentieren Lücken im Sequenzvergleich.

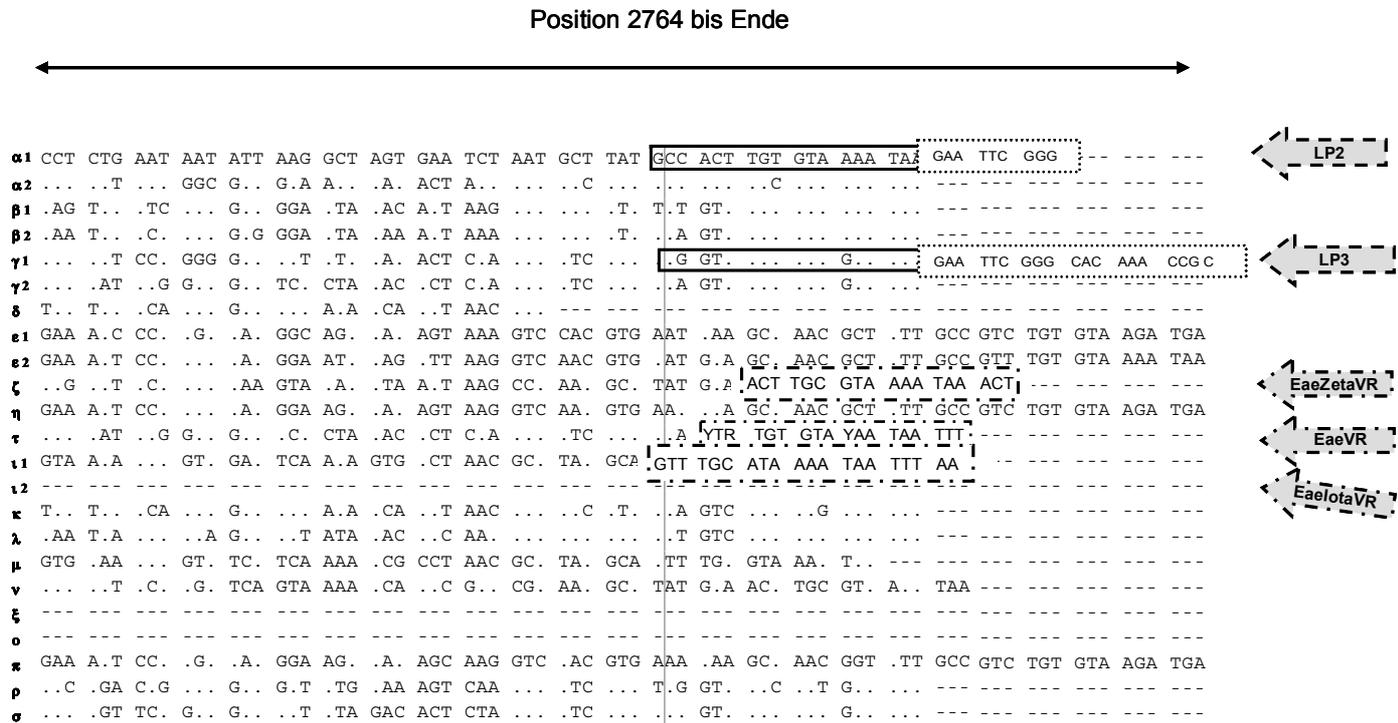


Abbildung 4.1.2-7: Lokalisation der Rückwärts-Primer „LP2“ und „LP3“ zum Nachweis von α - und γ -Intimin nach OSWALD et al. (2000) und der Primer „EaeVR2 (Universalrückwärts-Primer, alle Untereinheiten), „EaeZetaVR“ (Rückwärts-Primer für ζ -Intimin), „EaelotaVR“ (Rückwärts-Primer für ι -Intimin nach RAMACHANDRAN et al. [2003]) innerhalb des Sequenzvergleichs der bislang beschriebenen Intiminvarianten alpha (α) bis sigma (σ). Die Ziffern 1 und 2 hinter den Intimintypen kennzeichnen die jeweilige Untereinheit. Punkte symbolisieren homologe Sequenzbereiche, Bindestriche repräsentieren Lücken im Sequenzvergleich.

4.2 ETABLIERUNG UND OPTIMIERUNG DER PCR-TYPISIERUNGSVERFAHREN

4.2.1 EXEMPLARISCHE DARSTELLUNG DER PCR-ERGEBNISSE

Nachfolgend werden die mit den Primer-Kombinationen „Int-Ru/Int-A“ (alpha-Intimin), „Int-Ru/Int-B“ (beta-Intimin), „Int-Ru/Int-Y“ (gamma-Intimin) und „epsilonF/epsilonR“ (epsilon-Intimin) erzielten Ergebnisse unter Berücksichtigung der verwendeten Restriktionsendonukleasen dargestellt.

Alle Oligonukleotide generierten das spezifische Amplifikat im erwarteten Größenordnungsbereich (Bp). Die Restriktionsendonukleasen *Taq* I und *Fok* I gewährleisteten die Produktbestätigung anhand der zuvor berechneten Fragmentgrößen. Diese betragen für alpha-Intimin bei einer Amplifikatgröße von 557 Bp für *Taq* I 489 und 68 Bp sowie für *Fok* I 326 und 231 Bp. Analog dazu wurde das 562 Bp-große beta-Intiminprodukt durch *Taq* I in 450 und 112 Bp und durch *Fok* I in 501 und 61 Bp-große Fragmente unterteilt (**Abbildung 4.2.1-2**).

Die Generierung eines Amplifikates mit den an delta-Intimin angepassten Primern war bei keinem der Stämme möglich.

Für die 546 Bp des gamma-Intimins ergab sich ein Schnittmuster von 319 und 227 Bp für *Taq* I sowie von 480 und 66 Bp für *Fok* I (**Abbildung 4.2.1-3**). Das 66 Bp-große Fragment konnte auf dem dargestellten Agarosegel nicht visualisiert werden. Dies ist evtl. auf eine unzureichende Fragmentkonzentration zurückzuführen, die im Gel nicht dargestellt wurde.

Auch das 340 Bp große epsilon-Amplifikat zeigte das erwartete Schnittmuster durch *Fok* I mit 294 und 46 Bp sowie *Taq* I mit 254 und 86 Bp (**Abbildung 4.2.1-5** und **Abbildung 4.2.1-6**).

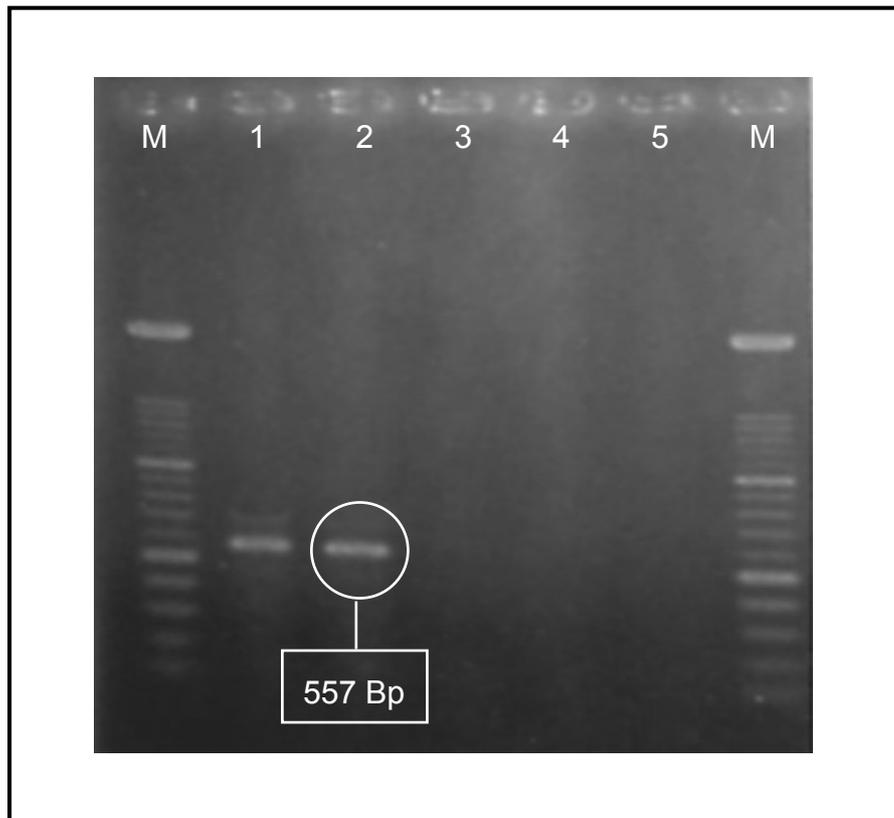


Abbildung 4.2.1-1: Exemplarische Darstellung der Amplifikation von α -Intimin mit der Primer-Kombination „Int-Ru/Int-A“ nach ADU-BOBIE et al. (1998a).

Spur 1 und 2: Positivkontrolle α -Intimin: Referenzstamm EPEC O127:H6 E 2348/69, Spur 3: EHEC O157:H7 EDL 933 (γ -Intimin) Spur 4: O26:H11 (β -Intimin), Spur 5: Leerwert. (M): Marker XIV

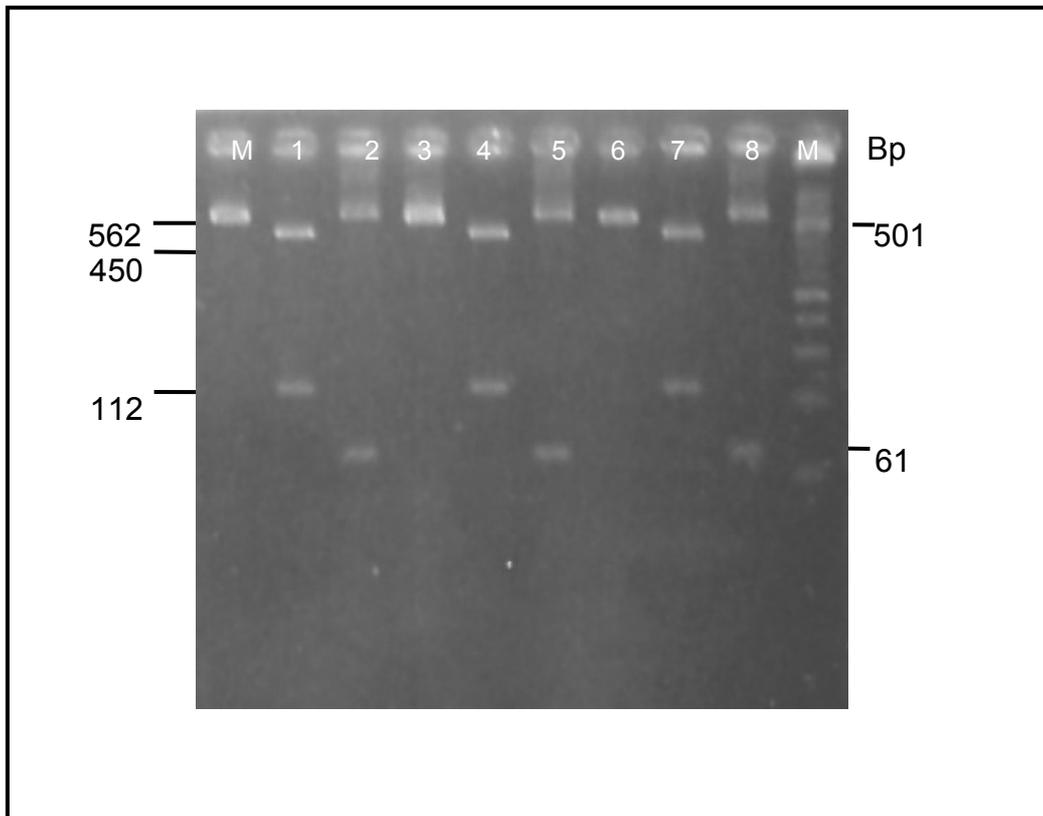


Abbildung 4.2.1-2: Exemplarische Darstellung der Amplifikation von β -Intimin mit der Primer-Kombination „Int-Ru/Int-B“ nach ADU-BOBIE et al. (1998a) und des restriktionsenzymatischen Verdaus mit *Taq* I und *Fok* I.

Spur 1, 4 und 7: β -Intimin ungeschnitten, Spur 2, 5 und 8: β -Intimin geschnitten mit *Taq* I, Spur 3, 6 und 8: β -Intimin geschnitten mit *Fok* I, (M): Marker XIII

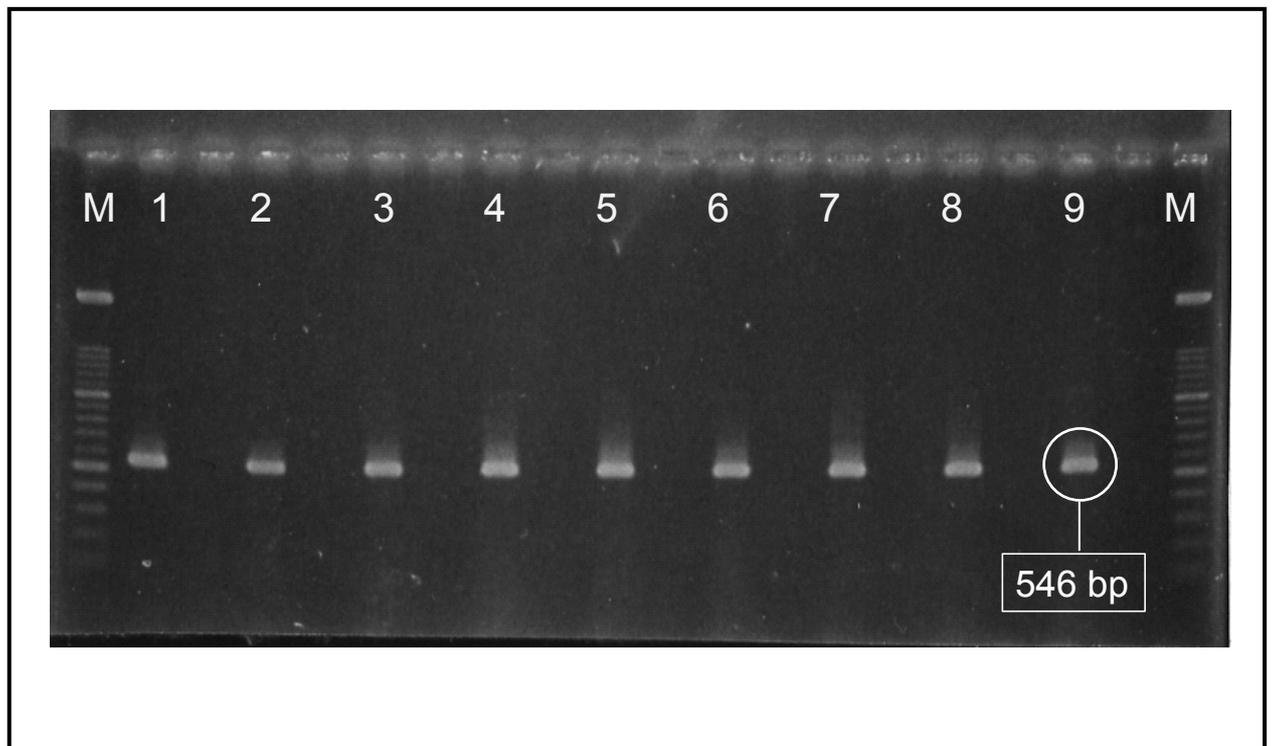


Abbildung 4.2.1-3: Exemplarische Darstellung der Amplifikation von γ -Intimin mit der Primer-Kombination „Int-Ru/Int-Y“ nach ADU-BOBIE et al. (1998a)

Spur 1 bis 9: γ -Intimin ungeschnitten, (M): Marker XIV

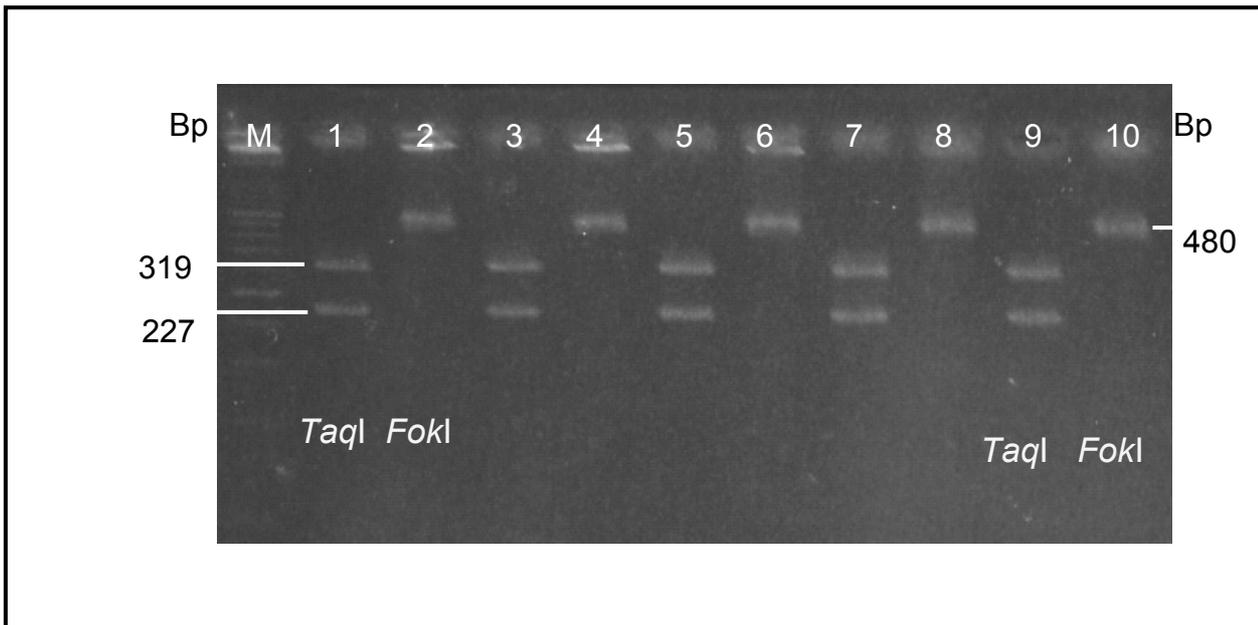


Abbildung 4.2.1-4: Restriktionsenzymatischer Verdau von γ -Intimin mit *Taq* I und *Fok* I.

Spur 1, 3, 5, 7 und 9: γ -Intimin geschnitten mit *Taq* I, Spur

2, 4, 6, 8 und 10: γ -Intimin geschnitten mit *Fok* I, (M):

Marker XIII

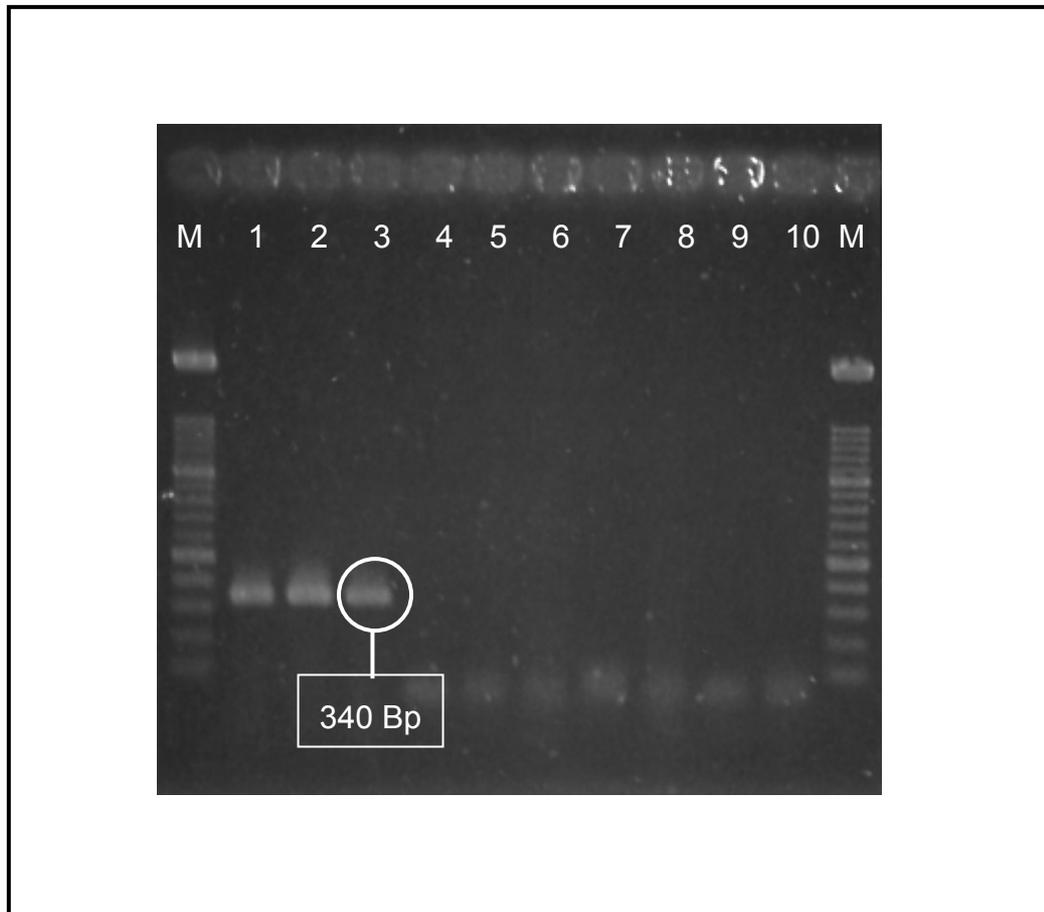


Abbildung 4.2.1-5: Exemplarische Darstellung der Amplifikation von ϵ -Intimin mit der Primer-Kombination „epsilonF/epsilonR“.

Spur 1: Referenzstamm PMK3 O103, Spur 2: Plath P 61,
Spur 3: 2704/94 O8:H⁻, Spur 5: EHEC O157:H7 EDL933,
Spur 6: EPEC O127:H6 E 2348/69, Spur 7: Petra 139
Osp:H⁻, Spur 8: C6 KJS O111:H, Spur 9: Sal 4/LXI/1
O65:H20, Spur 10: Leerwert Mastermix, (M): Marker XIV

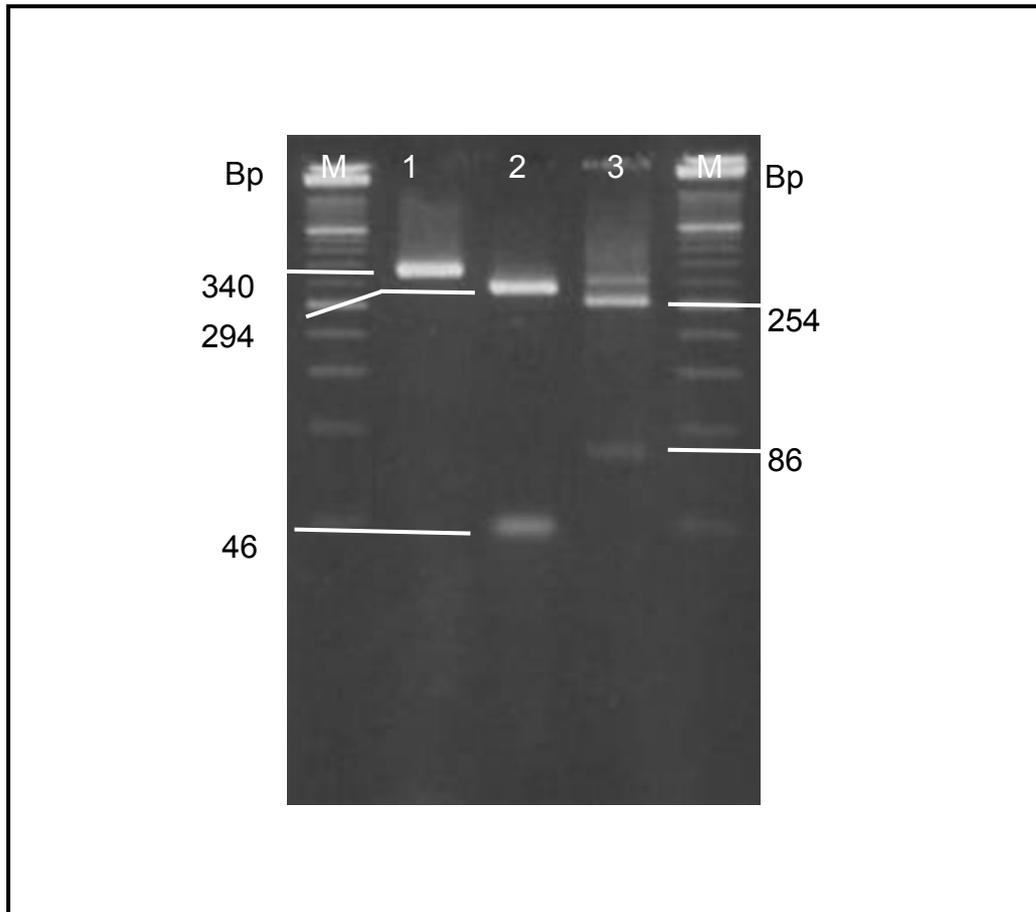


Abbildung 4.2.1-6: Restriktionsenzymatischer Verdau von ϵ -Intimin generiert mit der Primer-Kombination „epsilonF/epsilonR“.

Spur 1: Referenzstamm PMK3 O103, Spur 2: Geschnitten mit *Fok* I, Spur 3: Geschnitten mit *Taq* I, (M): Marker XIII

4.3 ETABLIERUNG EINER UNIVERSELLEN RFLP-TYPISIERUNGSMETHODE

Bezieht man alle Intiminsubtypen einschließlich der zunächst als eigenständig klassifizierten Untereinheiten, die sich mittlerweile als identisch erwiesen haben, mit ein, so waren zum Zeitpunkt der Arbeiten 19 verschiedene Intiminvarianten bekannt. Auf Basis der Oligonukleotid-Primer nach RAMACHANDRAN et al. (2003) wurde ein Restriktionsfragment-Längenpolymorphismus (RFLP) gestütztes Differenzierungsschema erarbeitet. Für den enzymatischen Verdau kamen abweichend von der Originalarbeit, die die Enzyme *Alu* I, *Rsa* I und *Cfo* I verwendete, die Restriktionsendonukleasen *Fok* I, *Rsa* I und *Hae* III zum Einsatz, da mit der erstgenannten Enzymkombination die Darstellung von 19 Varianten nicht möglich war.

4.3.1 ANALYSE DER VERWENDETEN OLIGONUKLEOTID-PRIMER-KOMBINATIONEN

Der Universalvorwärts-Primer „EaeVF“ wurde in die konservierte Region des *eae*-Gens gelegt. Aus dem Alignment-Vergleich ergaben sich am 3'-Ende Fehlbasenpaarungen für die Intiminuntereinheiten epsilon2 („G“ ersetzt durch „T“), iota2 bzw. my („G“ ersetzt durch „A“) und rho („AG“ ersetzt durch „CC“). Aufgrund der sehr stabilen Bindung kam der letztgenannten Fehlbindung die größte Bedeutung zu. Die Fehlbasenpaarungen „T“ bzw. „A“ konnten durch den stabilen „ATTAA“-Komplex aufgefangen werden. Der Primer wurde als degenerierter (engl. „degenerative“) Primer an Position 3, gemessen ab dem 5'-Ende, mit „Y“ versehen. An dieser Stelle wurden somit die Basen „C“ oder „T“ eingesetzt (**Abbildung 4.3.1-1**).

Die Lokalisation der Rückwärts-Primer „EaeVR, EaeZetaVR“ und „EaelotaVR“ ist in **Abbildung 4.1.2-7** dargestellt. „EaeVR“ beinhaltet als degeneriertes Oligonukleotid durch die Konfiguration „RAY“ (Leserichtung: 5´nach 3´) ein 3´-Ende bestehend aus „AAC, GAC, AAT“ bzw. „GAT“ und gewährleistete so die Amplifikation der Intiminvarianten alpha bis epsilon sowie kappa, lambda, rho und sigma durch vollständige Übereinstimmung am 3´-Ende.

Es wurde erkennbar, dass „EaeZetaVR“ spezifisch an die Sequenz für zeta-Intimin-Stämme adaptiert wurde. Das galt analog für „EaelotaVR“, der für die Amplifikation von iota-Intimin entwickelt wurde.

4.3.2 EXEMPLARISCHE DARSTELLUNG DER PCR-ERGEBNISSE

Die **Abbildung 4.3.2-1** und **Abbildung 4.3.2-2** veranschaulichen die mit den Oligonukleotid-Primern nach RAMACHANDRAN et al. (2003) generierten PCR-Produkte. In den **Abbildungen 4.3.2-3** bis **4.3.2-6** wurde das jeweilige restriktionsenzymatische Profil, ermittelt durch den Einsatz von *Fok I*, *Rsa I* und *Hae III*, dargestellt.

Für die exemplarisch analysierten Referenzstämme 8489/69 O26:H⁻, EDL 933 O157:H7 und PMK3 O103 ergab sich das erwartete Bild. Das 849 Bp-große beta-Intiminamplifikat ergab mit *Fok I* eine Fragmentierung in 669 und 180 Bp große Teilstücke, mit *Rsa I* traten 3 Banden von 528, 246 und 75 Bp Größe auf. *Hae III* schnitt in 435, 214 und 200 Bp-große Teilstücke (**Abbildung 4.3.2-3**).

Abbildung 4.3.2-3 Analog dazu wurde das 834 Bp-große gamma-Intiminprodukt durch *Fok I* in 453, 234, 83 und 64 Bp, mittels *Rsa I* in zwei 402 und 432 große Fragmente unterteilt, *Hae III* schnitt erwartungsgemäß nicht (**Abbildung 4.3.2-4**).

Für das 876 Bp enthaltene epsilon-Intimin ergab sich ein Schnittmuster von 730 und 116 Bp bei *Fok I* sowie mit 744 und 102 Bp durch *Rsa I*. *Hae III* schnitt diese Sequenz nicht. Damit war die Einstufung als epsilon1-Subtyp möglich (**Abbildung 4.3.2-5**).

Darüber hinausgehend dokumentiert die **Abbildung 4.3.2-6** die Differenzierung der Prüfstämme C6 KJS O111:H als theta-, Sal 4/LXI/1 O65:H20 als beta- und Petra 139 Osp:H⁻ als zeta-Intiminträger.

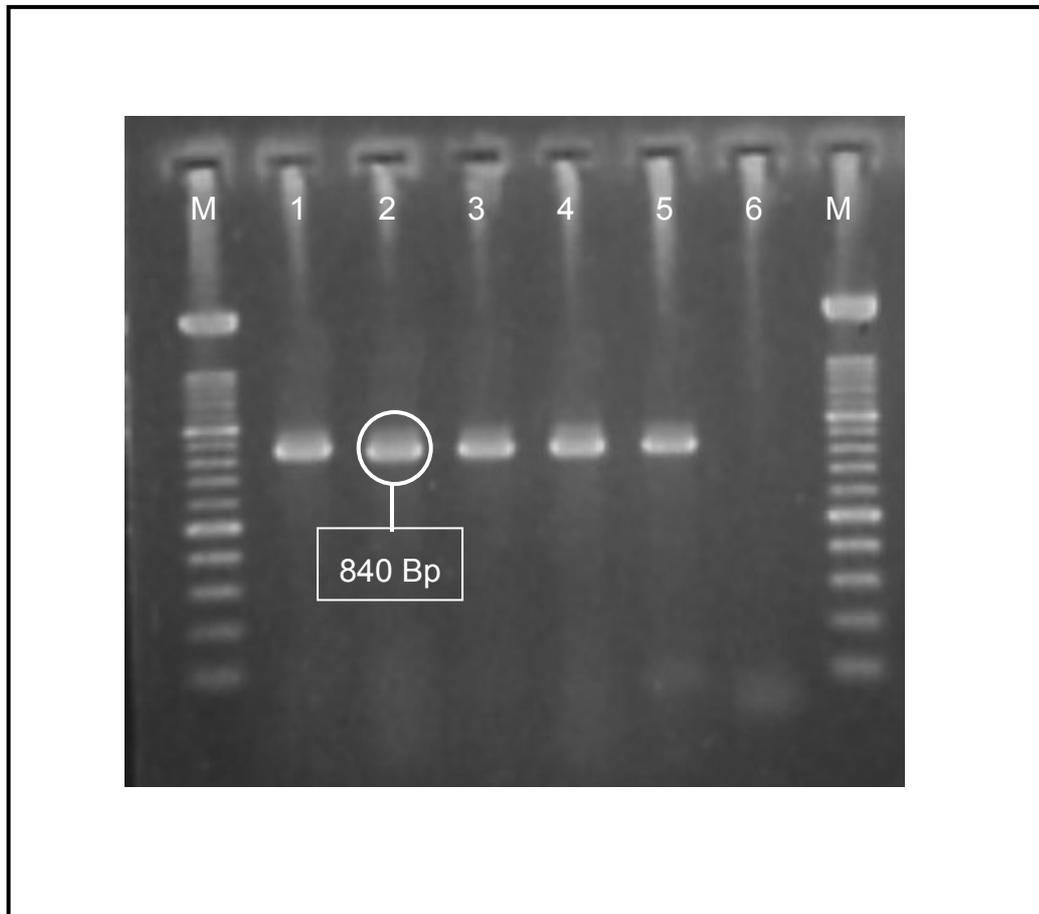


Abbildung 4.3.2-1: Exemplarische Darstellung der universellen Amplifikation verschiedener Intimingene mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003).

Spur 1, 2: Referenzstamm EDL 933 O157:H⁻, Spur 3: Referenzstamm PMK3 O103, Spur 4: 8489/67 O26:H⁻, Spur 5: C6KJS O111:H, Spur 6: LW, (M): Marker XIV



Abbildung 4.3.2-2: Exemplarische Darstellung der Amplifikation von ζ -Intimingenen mit der Primer-Kombination „EaeVF/EaeZetaR“ nach RAMACHANDRAN et al. (2003).

Spur 1, 2: Referenzstamm P 146 O84:H⁻, Spur 3: P 145 84:H31, Spur 4: P139 Osp:H⁻, Spur 5: LW, (M): Marker XIV

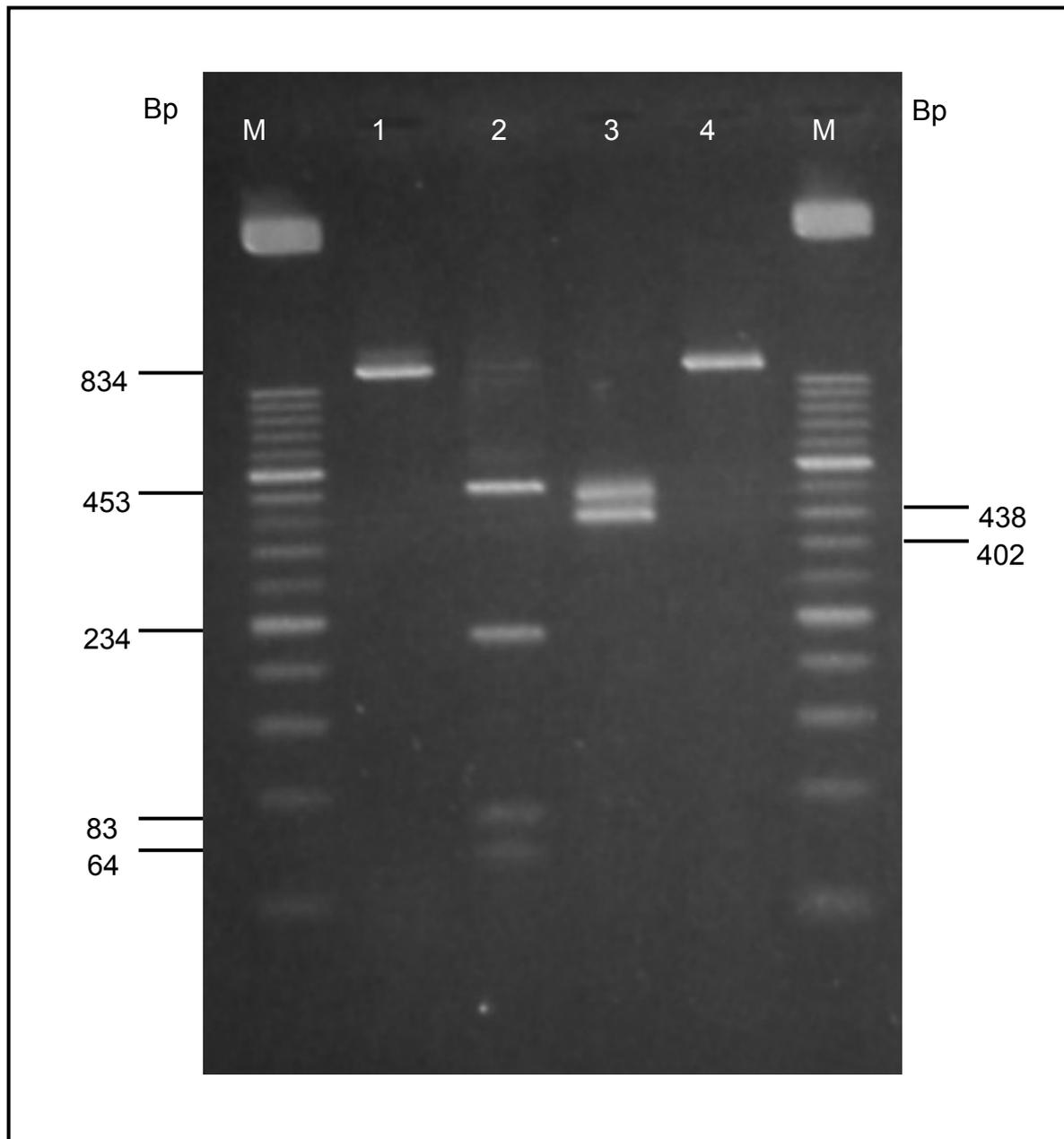


Abbildung 4.3.2-4: Exemplarische Darstellung des restriktionsenzymatischen Verdaus von γ -Intimin nach Amplifikation mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003).

Spur 1: Referenzstamm EDL 933 O157:H7 ungeschnitten,
Spur 2: Geschnitten mit Fok I, Spur 3: Geschnitten mit Rsa I,
Spur 4: Hae III (schneidet γ 1-Intimin nicht), (M): Marker XIII

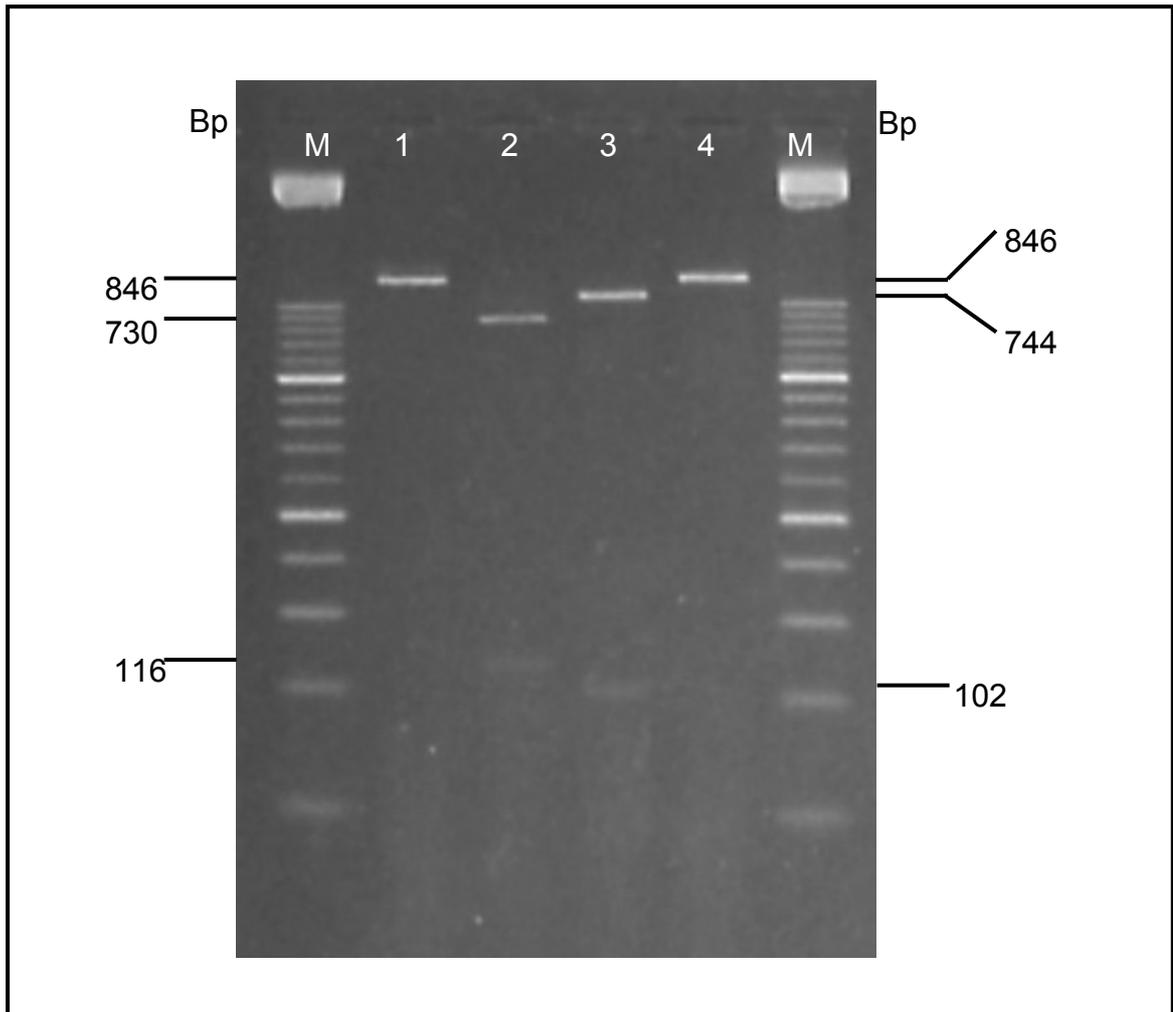


Abbildung 4.3.2-5: Exemplarische Darstellung des restriktionsenzymatischen Verdaus von ϵ -Intimin nach Amplifikation mit der Primer-Kombination „EaeVF/EaeVR“ nach RAMACHANDRAN et al. (2003).

Spur 1: Referenzstamm PMK3 O103 ungeschnitten, Spur 2: Geschnitten mit *Fok* I, Spur 3: Geschnitten mit *Rsa* I, Spur 4: *Hae* III (schneidet ϵ 1-Intimin nicht), (M): Marker XIII

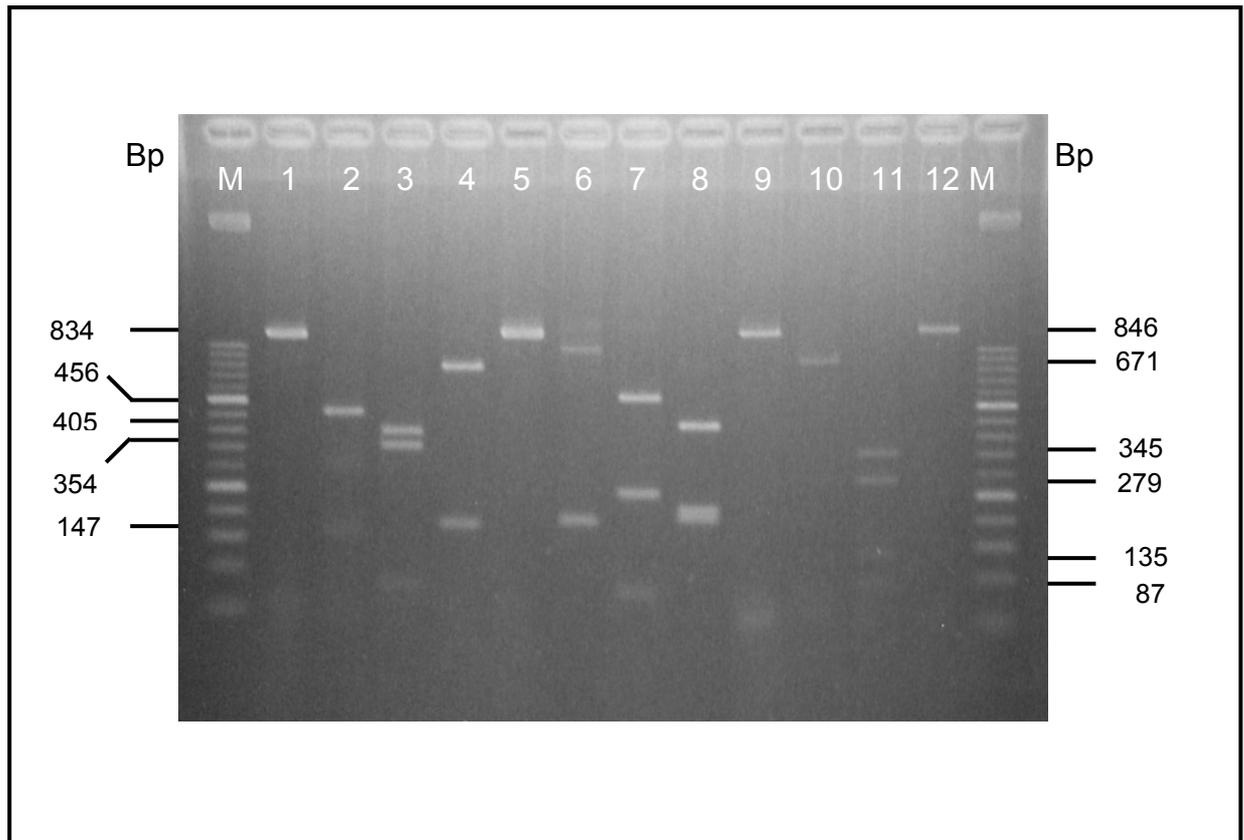


Abbildung 4.3.2-6: Exemplarische Darstellung des restriktionsenzymatischen Verdaus von θ -, β - und ζ -Intimin nach Amplifikation mit den Primer-Kombinationen „EaeVF/EaeVR“ und „EaeVF/EaeZetaR“ nach RAMACHANDRAN et al. (2003).

Spur 1: Prüfstamm C6 KJS O111:H ungeschnitten,

Spur 5: Prüfstamm Sal 4/LXI/1 O65:H20 ungeschnitten,

Spur 9: Prüfstamm Petra 139 Osp:H⁻,

Spur 2, 6, 10: Geschnitten mit *Fok* I,

Spur 3, 7, 11: Geschnitten mit *Rsa* I,

Spur 4, 8: Geschnitten mit *Hae* III ,

Spur 12: *Hae* III (schneidet bei ζ -Intimin nicht),

(M): Marker XIII

4.4 ERGEBNISSE DER INTIMINTYPISIERUNG EINBEZOGENER *E. COLI*-STÄMME

4.4.1 KLASSIFIZIERUNG *EAE*-POSITIVER *E. COLI*-STÄMME

In die Intimintypisierung wurden insgesamt 317 *eae*-positive Stämme einbezogen. Sie setzten sich aus 158 im Rahmen des BMG-Projektes (Aktenzeichen: 422-7030-56/48) gewonnenen und weiteren 159 Stämmen der institutseigenen Stammsammlung zusammen (**Tabelle 3.2.2-1**). Ihre Zuordnung zu den Habitaten „Stuhlproben von HUS/HC-Patienten, Lebensmitteln, Rinder-, Schaf- und Pferdekot sowie Schweinekot bzw. -sektionsmaterial“ wurde in **Tabelle 4.4.2-3** zusammengefasst. Bei 124 Stämmen wurde kein *stx*-Gen nachgewiesen. Sie wurden daher den EPEC zugewiesen. Die 193 mit wenigstens einem Verotoxin-Gen ausgestatteten Stämme wurden als EHEC klassifiziert.

4.4.2 SEROVAREN DER *EAE*-POSITIVEN *E. COLI*-STÄMME

Die in sieben Habitate gruppierten 317 *eae*-positiven Isolate waren 56 Serovaren zuzuordnen. Sie umfassten 31 O-Serogruppen, 17 H-Serovaren und 30 O:H-Serovaren. Insgesamt überwogen Stämme der vier Serogruppen O111 (6), O26 (11), O156 (37) und O157 (197), die damit 251 (79,2%) der untersuchten Isolate darstellten. Es überwogen vier H-Serovaren: H11 (3), H2 (7) sowie H⁻ (15). H7 wurde nur bei den 82 untersuchten O157-Stämmen typisiert.

Table 4.4.2-1: Zuordnung unterschiedlicher Serovare (n=56) zu den Intimintypen

Int- α^*	n	Int- β^*	n	Int- δ	n	Int- γ^{**}	n	Int- ϵ^*	n	Int- ζ	n	Int- θ	n	Int- κ	n	Nicht typisiert**	n
O127:H6	1	NT	1	Nicht nachgewiesen	1	O111:H-	1	Or:H-	1	Ont:H-	1	O1:H2	1	O107:H11	1	Ont:Hnt	1
		Ont:H11	1		1	O111:H2	1	O8:H-	1	Osp:H-	1	O76:H7Ont	1			OX177:H+	1
		O15:H17	1		1	O128:H2	1	O11:H2	1	O84:H-	2	O111:H	1			O2:H-	1
		O65:H20	1		1	O145:H-	1	O103:H3	1	O84:H31	2					O10:H4	1
		O111H2	1		1	O55:H-	2	O103:H2	14	O156:H21	1					O11:H2	1
		O15:Hnt	2		29	O157:H-	29			O156:H26	1					O15:H-	1
		O15:H2	2		82	O157:H7	82			O156:H-	7					O.22:H8	1
		O70:H-	2		86	O157	86			O156:H25	11					O48:H-	1
		O70:H11	2													O49:H41	1
		O111:H-	2													O56:H-	1
		O119:H-	2													O62:H32	1
		O26:H-	4													O111:H-	1
		O26:H11	6													O145:H-	1
		O118:H-	7													O145:H16	2
																O156:H26	1
																O145:H16	4
																O156:H47	11

* Soweit nicht anders angegeben, handelt es sich um den jeweiligen Subtyp 1

** Die Angabe „nicht typisiert“ bezieht sich darauf, dass die Isolate mit dem modifizierten PCR-System nach ADU-BOBIE et al. (1998a) keiner der 5 untersuchten Intimintypen zugeordnet werden konnten

grau unterlegt: Serovare, die über verschiedene Intimintypen verfügen

Tabelle 4.4.2-2: Anteil der Intiminvarianten (n[%]) an den 317 untersuchten Prüfstämmen

Intimin	n	[%]
α	1	0,32
β	34	11,04
δ	0	0
γ	203	64,03
ϵ	18	5,68
ζ	26	8,20
θ	3	0,95
κ	1	0,32
Nicht typisiert	31	9,46
Gesamtzahl untersuchter <i>E. coli</i> -Stämme	317	100,0

Tabelle 4.4.2-3: Verteilung (n) der Intiminuntereinheiten, gruppiert nach Habitaten

Habitat	Intimin									Summe
	Int- α	Int- β	Int- δ	Int- γ	Int- ϵ	Int- ζ	Int- θ	Int- κ	Nicht typisiert	
Stuhlproben von HUS/HC-Patienten	1	7	0	101	4	0	0	0	5	118
Lebensmittel	0	13	0	54	9	15	3	1	19	114
Rinderkot	0	11	0	41	3	4	0		5	64
Schafkot (n=10)	0	2	0	0	0	7	0	0	1	10
Schweinekot/-sektionsmaterial	0	1	0	0	2	0	0	0	1	4
Pferdekot	0	0	0	1	0	0	0	0	0	1
Keine Angabe	0	0	0	6	0	0	0	0	0	6
Summe	1	34	0	203	18	26	3	1	31	317

4.4.3 TYPISIERUNG DER INTIMINGENE

Mit den modifizierten PCR-Systemen nach ADU-BOBIE et al. (1998a) und dem bei 27 *Escherichia coli*-Stämmen eingesetzten Typisierungsverfahren auf der Basis von RAMACHANDRAN et al. (2003) konnten 145 (91,8%) der aus dem BMG-Projekt einbezogenen *E. coli*-Stämme einem Intimintyp zugeordnet werden.

Bei den übrigen 159 Isolaten gelang die Typisierung bei 142 Stämmen zu 89,3%. Insgesamt konnten mit beiden Verfahren 95,9% der *E. coli* einem Intimintyp zugewiesen werden.

Die insgesamt 31 nicht zuzuordnenden Isolate setzten sich aus 14 unterschiedlichen Serogruppen zusammen, dabei überwog das Serovar O156:H47 mit 11 Stämmen dieses Typus (**Tabelle 1.3.5-2**). Zuvor stellten sie sich mit der Methode nach ADU-BOBIE et al. (1998a) als alpha-, beta-, gamma-, delta und epsilon-Intimin negativ dar. Es ist davon auszugehen, dass bei einem erweiterten Screening mittels des vorgestellten RFLP-Profiles eine Zuordnung zu anderen Intiminvarianten möglich wäre.

4.5 SEQUENZIERUNG DES *EAE*-GENS

4.5.1 ANALYSE DER ROHDATEN

Mit den verwendeten Primer-Systemen konnten jeweils 2675 von 2817 Basenpaaren der *eae*-Gene der Serovare O156:H⁻ P 146 (Genbank-Zugriffs-Nr. AY520904) und O84:H⁻ P 256 (Genbank-Zugriffs-Nr. AY520905) generiert werden. Durch Einbeziehung benachbarter Gensequenzabschnitte wurde das hypervariable 3'-Ende des Intimingens (ab Nukleotidposition 2112) vollständig sequenziert und das Ende des offenen Leserasters (engl. „open reading frame [orf]“) anhand der Poly-A-Region und des Stoppsignales „*taa/uaa*“ festgelegt.

Die Sequenz des konservierten 5'-Endes wurde bis auf 142 Bp bestimmt. Insgesamt lagen anteilig 95% der *eae*-Gene zur vergleichenden Analyse vor.

Im BLAST-Sequenzdatenbankvergleich zeigte sich eine 100%-ige Übereinstimmung mit dem erstmalig von JORES et al., (2002) beschriebenen zeta (ζ)-Intimin. Die charakterisierten Serovare O84.NM (Genbankzugriffs-Nr. AJ298279) und O84:H? (Genbankzugriffs-Nr. AJ275089) und die durch ZHANG et al. (2002b) sowie TARR und WHITTAM (2002) als zeta (ζ)-Intiminträger definierten O84:H4 (Genbankzugriffs-Nr. AJ271407), O111:HNM (Genbankzugriffs-Nr. AF449416) und O111:H9 (Genbankzugriffs-Nr. AF449417) waren Bestandteil der dargestellten Analysen. Die Charakteristika der für die Sequenzanalyse eingesetzten sieben zeta (ζ)-Isolate gibt **Tabelle 4.5.1-1** wieder. Eine exemplarische Wiedergabe der schrittweisen Amplifikation der beiden Intimingene zeigt **Abbildung 4.5.1-1**.

Tabelle 4.5.1-1: Tabellarische Übersicht der Charakteristika der für die Sequenzanalyse verwendeten zeta (ζ)-Intimingene

Intimin-variante	Größe der vorliegenden Nukleotidsequenz (orf-Länge) [Bp]	G+C-Gehalt [%]	Genbankzugriffs-nr.	Serovar
zeta (ζ)	2817	42,39	AF449416	O111:HNM
zeta (ζ)	2817	42,39	AF449417	O111:H9
zeta (ζ)	2817	42,60	AJ271407	O84:H4
zeta (ζ)	859*	39,38	AJ275089	O84:H?
zeta (ζ)	2817	42,63	AJ298279	O84:NM
zeta (ζ)	2675**	43,18	AY520904	O156:H ⁻
zeta (ζ)	2675**	43,18	AY520905	O84:H ⁻

* Partielle Sequenzierung beinhaltet das 3'-Ende (ab Position 2112)

** Partielle Sequenzierung beinhaltet vollständiges 3'-Ende (ab Position 2112) und Anteile des 5'-Endes

Die orf-Größe des zeta-Intimingens betrug 2817 Bp. Der G+C-Gehalt lag bei den nahezu vollständig vorliegenden Referenzsequenzen zwischen 42,39 und 43,18% (**Tabelle 4.5.1-1**). Die prozentuale Übereinstimmung der genannten zeta-Nukleotidsequenzen (n=7) wurde in **Tabelle 4.5.1-2** zusammengefasst. Die Nukleotidsequenzen des O86:H⁻ - und des O156:H⁻ -Isolates sind zu 100% identisch und wiesen mit 99,9% bzw. 99,8% die größte Übereinstimmung zu den Serovaren O84:NM (Genbank-Zugriffs-Nr. AJ298279) und O84:H4 (Genbank-Zugriffs-Nr. AJ271407) auf. Die Serogruppe O111 differiert hingegen um 1,4%.

Tabelle 4.5.1-2: Tabellarische Übersicht der prozentualen Sequenzdistanzen (Übereinstimmung, Unterschiede) der zeta (ζ)-Intiminvarianten (n=7) erstellt mit MegAlign™ (Fa. DNastar)

Nr. ζ -Intimin-stamm	Übereinstimmung der Sequenzen in [%]								Genbank-zugriffs-nr.
	1	2	3	4	5	6	7		
1		99,4	98,6	98,6	98,6	98,5	98,5	1	AF449416
2	0,0		98,6	98,6	98,6	98,5	98,5	2	AF449417
3	1,4	1,4		99,9	100,0	99,9	99,9	3	AJ271407
4	1,4	1,4	0,1		100,00	99,8	99,8	4	AJ275089
5	1,4	1,4	0,0	0,0		99,9	99,9	5	AJ298279
6	1,4	1,4	0,1	0,1	0,0		100,00	6	AY520904
7	1,4	1,4	0,1	0,1	0,0	0,0		7	AY520905
	1	2	3	4	5	6	7		
	Differenz der Sequenzen in [%]								

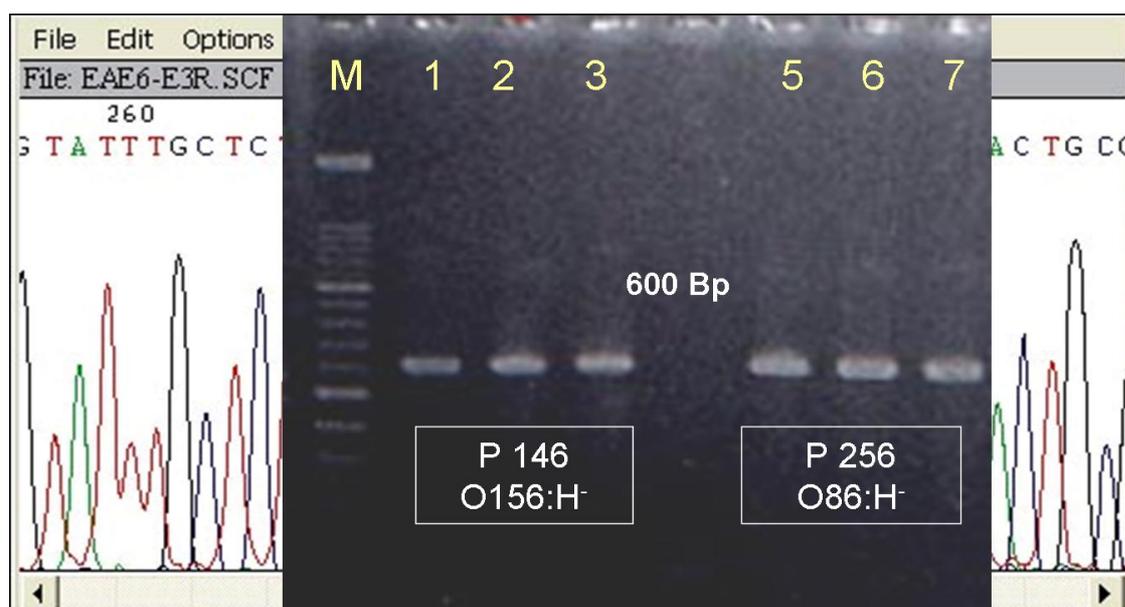


Abbildung 4.5.1-1: Exemplarische Darstellung der schrittweisen Amplifikation des *eae*-Gens der untersuchten Isolate P 146 und P 256

4.6 ABLEITUNG MOLEKULARGENETISCHER PHYLOGENIE

4.6.1 DENDROGRAMME

Die phylogenetischen Relationen wurden als linearisierte und ungewurzelte Bäume dargestellt. *Citrobacter freundii* (AB040740) und *Escherichia coli* O115 wurden als Außengruppe gewählt. Anhand des Modells nach NEI und GOJOBORI ließen sich unter Verwendung der nonsynonymen p-Distanzen folgende Intimingruppierungen unter Berücksichtigung einer 100%igen Bootstrap-Unterstützung bilden: Beta (β)₁ und Beta (β)₂ mit kappa (κ) und delta (δ), alpha (α)₁ mit omikron (\omicron) und zeta (ζ) einschließlich alpha (α)₂ und nu (ν), iota (ι)₁ und iota (ι)₂ mit my (μ), lambda (λ), epsilon (ϵ)₁ mit eta (η), xi (ξ) und epsilon (ϵ)₂ mit pi (π). Gamma (γ)₁ mit sigma (σ) und gamma (γ)₂ mit theta (θ) und rho (ρ) bildeten ebenfalls eigene Gruppierungen (**Abbildung 4.6.1-1**). Beta (β)₂-Intiminträger wiesen die engste Verwandtschaft zur Außengruppe auf. Die Sequenz-Übereinstimmungs-Matrix (**Tabelle 9.1.1-2**) gab Aufschluss über die korrekte Wiedergabe der im Sequenzvergleich ermittelten Relationen. Das untersuchte ζ -Intimin zeigte, wie im Dendrogramm dargestellt, die größte Übereinstimmung mit α , ν und \omicron auf. Mit Ausnahme der Intiminvarianten iota (ι)₁ und iota (ι)₂ mit my (μ), lambda (λ) und rho (ρ) entsprach die Darstellung der Matrize. Anhand des Bootstrap-Testes (48 und 36%) wurde ebenfalls deutlich, dass die Phylogenie nicht exakt wiedergegeben werden konnte. Anhand des „Interior Branch“- Testes (**Abbildung 4.6.1-2**) wurde die Wiedergabegenauigkeit der berechneten Astlängen und damit des Zeitpunktes der phylogenetischen Auseinanderentwicklung überprüft. Mit als signifikant einzustufenden Werten oberhalb von 95% wurde die oben vorgenommene Einteilung der Untergruppen bestätigt. Die Abgrenzung von λ - und ι -Intimin war nicht signifikant.

Sofern die Unterstützung für die weitere Verzweigung innerhalb der Subtypen unterhalb der 95%-Grenze lag, wurden Werte größer als 66% ermittelt. Da eine exakte Kalkulation nur für vollständig vorliegende Sequenzen vorgenommen werden kann, wurden partiell vorliegende Referenzsequenzen soweit möglich von der Berechnung ausgenommen.

Anmerkung: Für δ -Intimin lag nur ein partieller Sequenzeintrag vor, was bei der abschließenden Bewertung des Ergebnisses (36%) zu berücksichtigen war. Dies galt ebenso für die ζ -Intimingene (47%), da hier zur Kalkulation die in den eigenen Arbeiten mit 95%-Anteil vorliegenden Sequenzen einbezogen wurden.

Die **Abbildung 4.6.1-3** zeigt das Dendrogramm unter Verwendung des Modells nach TAMURA und NEI. Die Gruppierung der Intimingene war anhand der Bootstrap-Kalkulation als mit der zuvor getroffenen Einteilung übereinstimmend anzusehen. Die Gruppierung von ρ - (40%) und λ -Intimin (69%) erfuhr nur unzureichende Unterstützung. Der „Interior-Branch“-Test (**Abbildung 4.6.1-4**) bestätigte die signifikante Wiedergabe der Astlängen.

Die Maximum-Likelihood-Kalkulation (**Abbildung 4.6.1-5**) stimmte mit der durch die beiden beschriebenen Verfahren ermittelten Gruppierung überein.

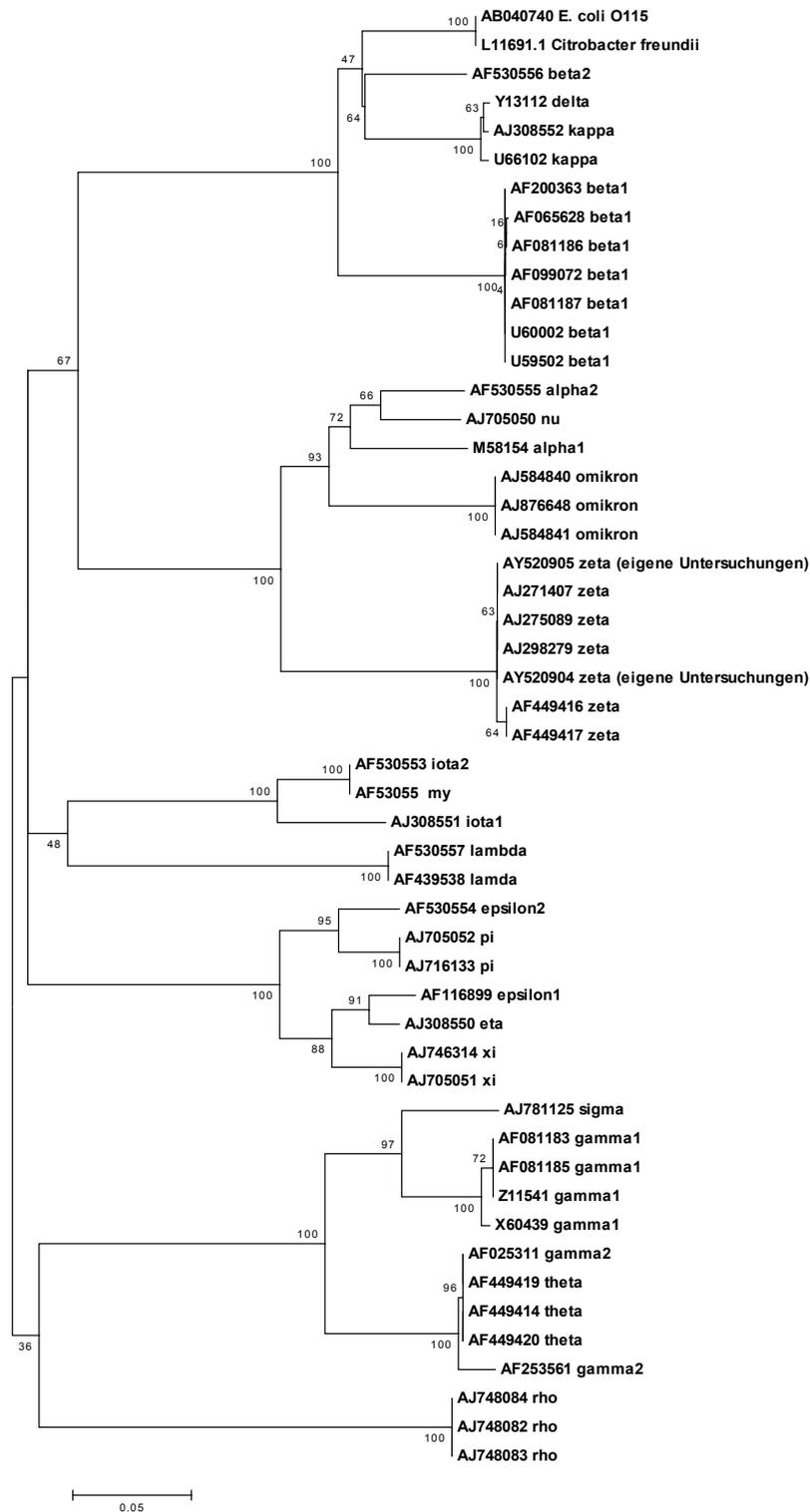


Abbildung 4.6.1-1: Linearisierter, ungewurzelter Baum der Intiminvarianten unter Verwendung des modifizierten Modells nach NEI und GOJOBORI (p-Distanz) in der Neighbor-Joining-Darstellung. Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent.

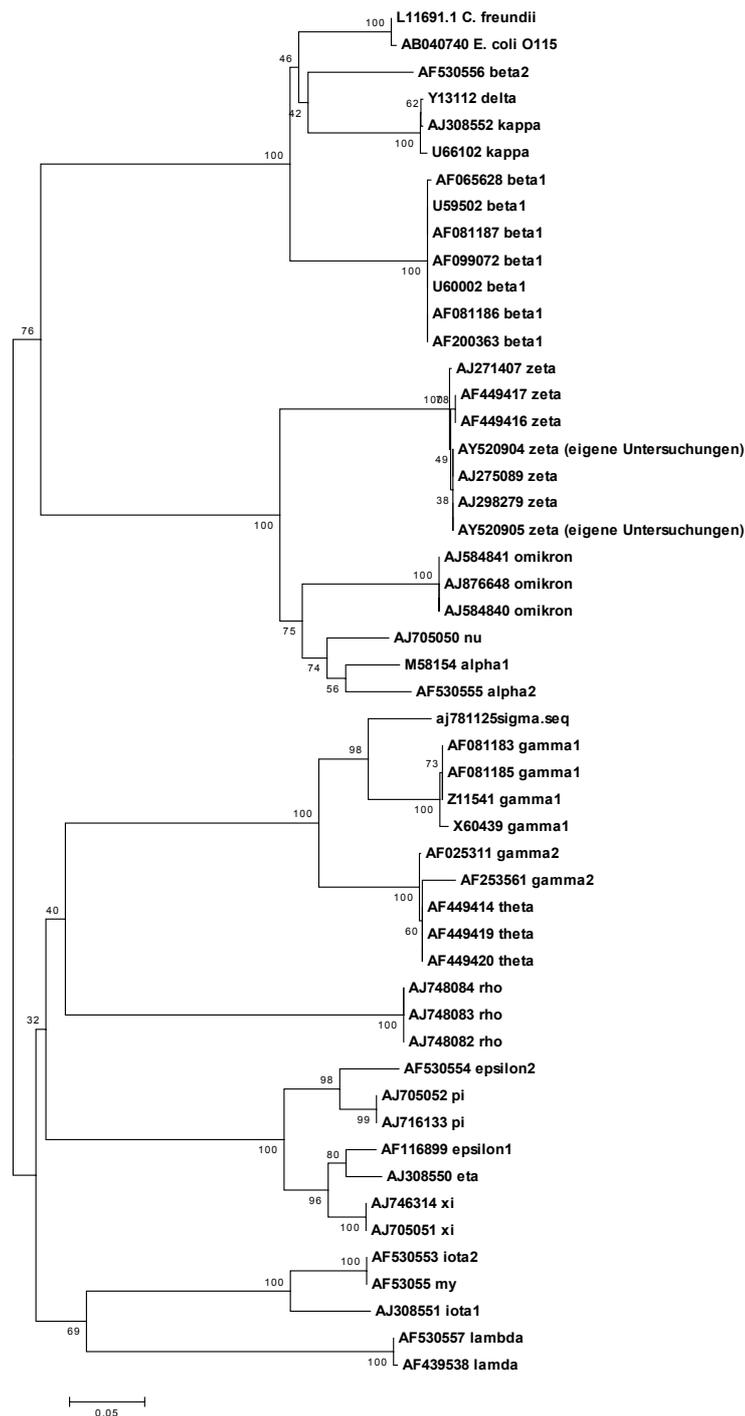


Abbildung 4.6.1-3: Linearisierter, ungewurzelter Baum der Intiminvarianten unter Verwendung des Modells nach TAMURA und NEI in der Neighbor-Joining- Darstellung. Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent.

4.6.2 NETZWERKDARSTELLUNG

Für die Darstellung des phylogenetischen Netzwerkes wurde die Eignung der Split-Zerlegung nach BANDELDT und DRESS (1992) und des Neighbor-Net-Verfahrens nach BRYANT und MOULTON (2004) für 25 Referenzsequenzen überprüft. Die Analysen berücksichtigten neben der gesamten Gensequenz, den konservierten 5'- und den hypervariablen 3'-Abschnitt. Anhand des Modells nach BRYANT und MOULTON (2004) konnten zwischen den abschnittsweise untersuchten Sequenzen einschließlich λ - und ρ -Intimin netzwerkartige Relationen nachgewiesen werden (**Abbildung 4.6.2-1**, **Abbildung 4.6.2-2**, **Abbildung 4.6.2-3**). Die phylogenetische Gruppierung differiert bei Berücksichtigung des Gesamtgens nicht (**Abbildung 4.6.2-1**), ρ -Intimin wird allerdings - im Unterschied zur Darstellung des 3'-Bereiches - mit μ - und σ -Intimin gruppiert. Neben der hohen Bootstrap-Unterstützung (dargestellte Werte: $\geq 90\%$) sind die prozentualen Fit-Werte zu berücksichtigen. Sie sind als Maß für die Wiedergabegenauigkeit, d. h. die in der Darstellung erfassbaren Splits und damit die Eignung der Methode zur korrekten Analyse der Datenmatrix im Hinblick auf nichtlineare Evolutionsereignisse, wie z. B. Rekombination, anzusehen und betragen für den gesamten Intimingenbereich 95,72%, den konservierten 5'-Abschnitt 96,79% und die hypervariable 3'-Region 98,51%. Die detaillierte Auflistung der berechneten „Splits“ ist der **Tabelle 9.1.2-1** zu entnehmen. Die Split-Zerlegung nach BANDELDT und DRESS (1992) und HUGHSON (1998) generiert für den 5'-Intimingenbereich (**Abbildung 4.6.2-4**) eine buschartige Darstellung, der Fit-Wert betrug bei hohen Bootstrap-Werten nur 49,87%. Die Analyse der hypervariablen Region (**Abbildung 4.6.2-5**) ergab eine erkennbar geringere Anzahl dargestellter „splits“ bei einem Fit-Wert von 63,91%.

(Anmerkung: Die Bewertung der Daten aus der Split-Zerlegung erfolgte in Absprache mit Herrn Prof. D. Huson, Lehrstuhl für Algorithmen in der Bioinformatik, Universität-Tübingen)

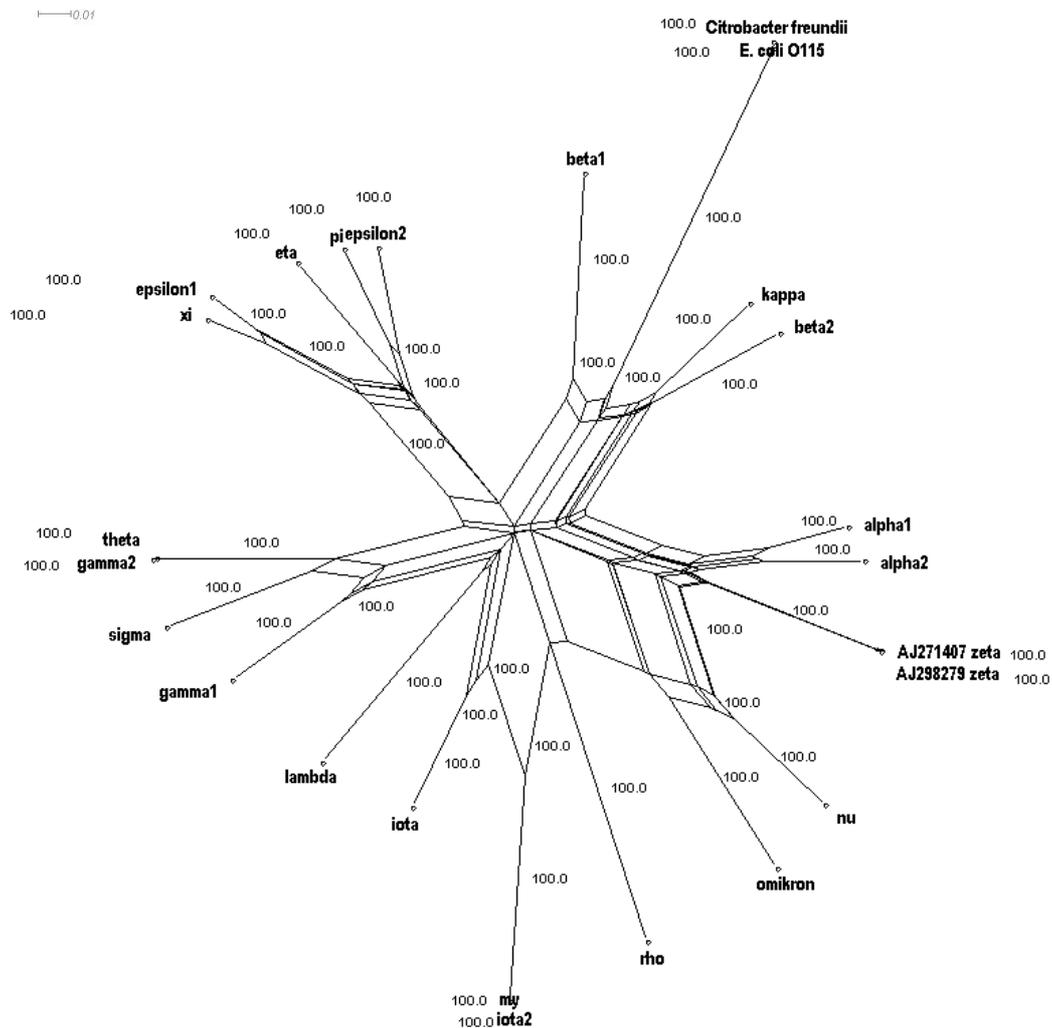


Abbildung 4.6.2-1: Neighbor-Net-Darstellung des phylogenetischen Netzwerkes unter Berücksichtigung der gesamten kodierenden Intimingensequenz anhand des Modells nach BRYANT und MOULTON (2004). Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent. Der Fit-Wert der Darstellung beträgt 95,72.

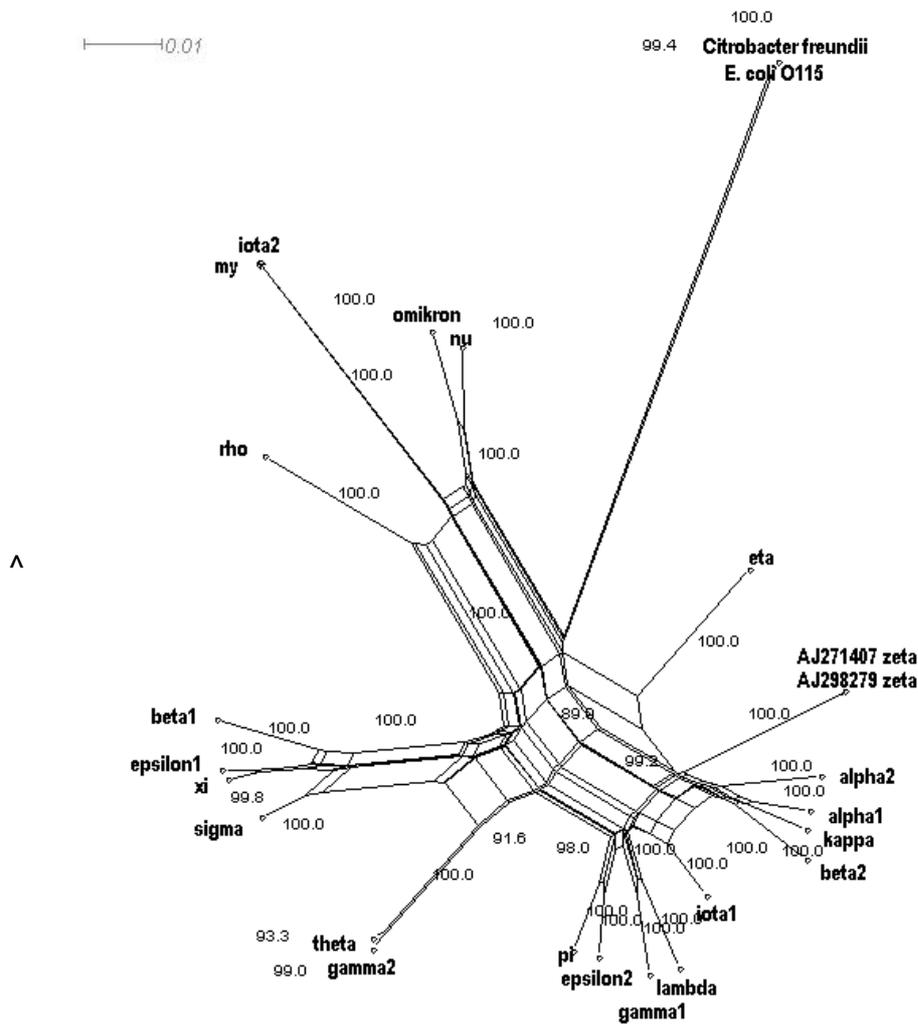


Abbildung 4.6.2-2: Neighbor-Net-Darstellung des phylogenetischen Netzwerkes der konservierten 5'-Intimingenabschnitte anhand des Modells nach BRYANT und MOULTON (2004). Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent. Der Fit – Wert der Darstellung beträgt 96,79.

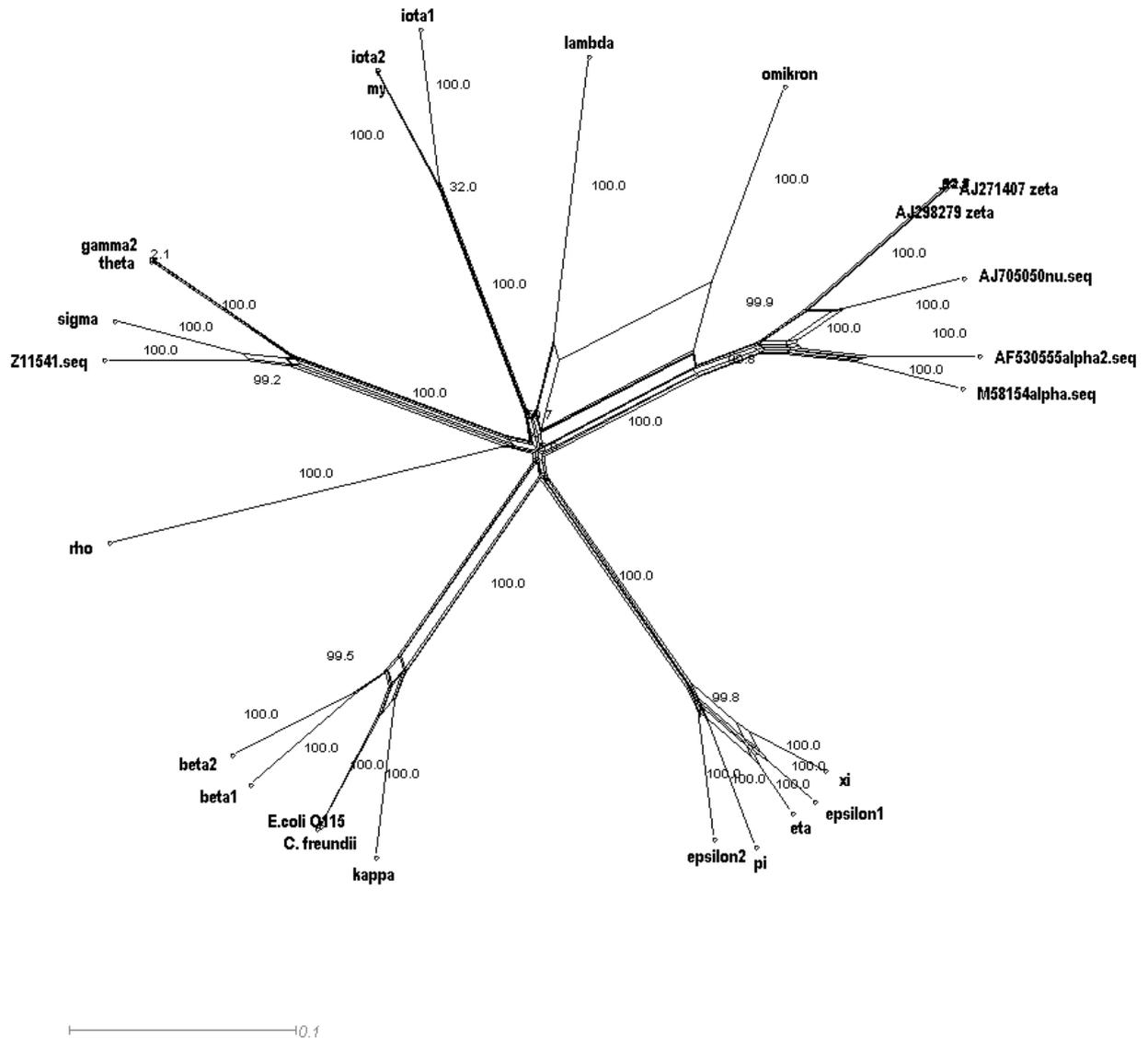


Abbildung 4.6.2-3: Neighbor-Net-Darstellung des phylogenetischen Netzwerkes der hypervariablen 3'-Intimingenregionen anhand des Modells nach BRYANT und MOULTON (2004). Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent. Der Fit-Wert der Darstellung beträgt 98,51.

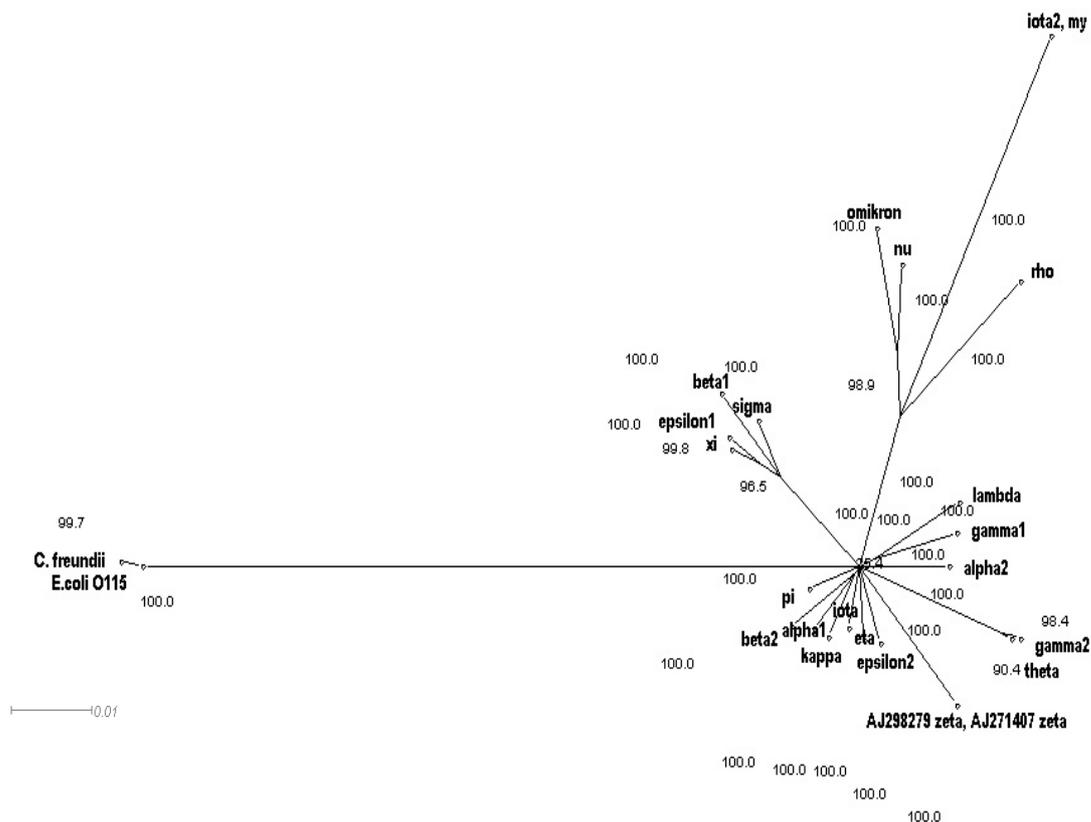


Abbildung 4.6.2-4: Darstellung des phylogenetischen Netzwerkes der konservierten 5'-Intimingenenabschnitte anhand der Split-Zerlegung nach BANDEL and DRESS (1992) und HUSON (1998). Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent. Der Fit –Wert der Darstellung beträgt 49,87.

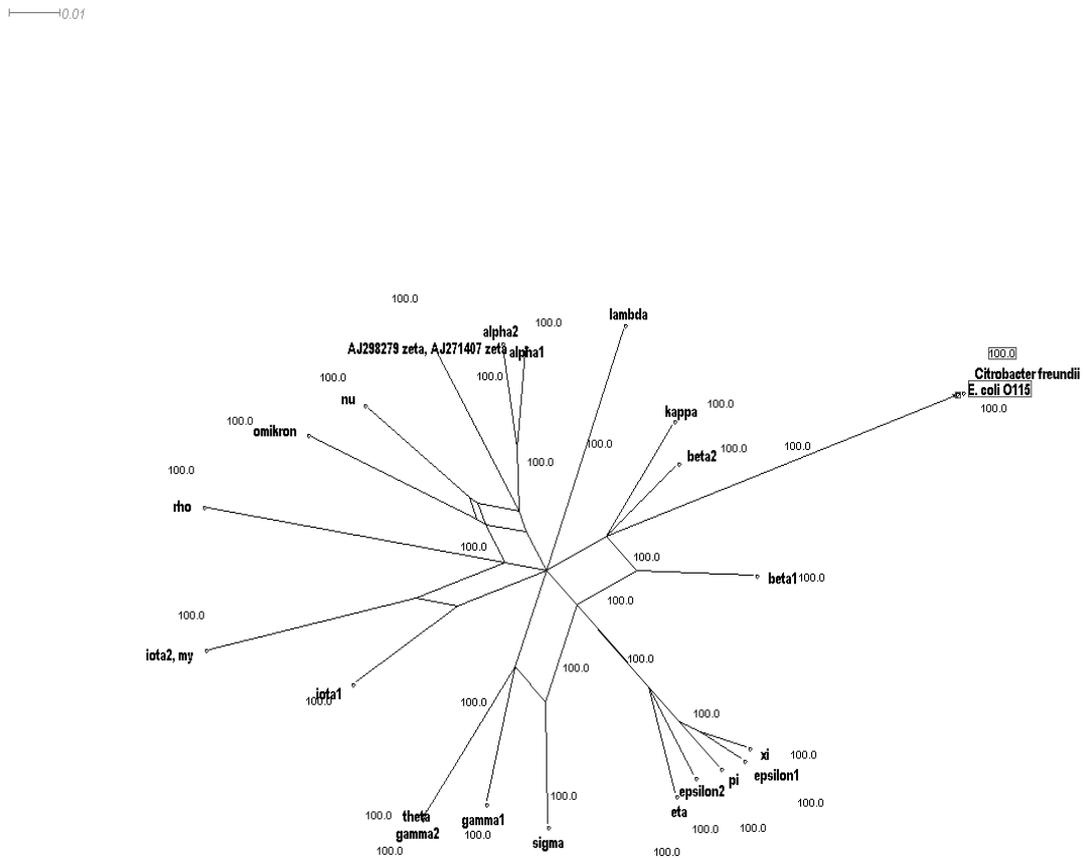


Abbildung 4.6.2-5: Darstellung des phylogenetischen Netzwerkes unter Berücksichtigung der gesamten Intimingensequenz anhand der Split-Zerlegung nach BANDELT und DRESS (1992) sowie HUSON (1998). Die Zahlen an den Kanten sind Bootstrap-Werte in Prozent. Der Fit –Wert der Darstellung beträgt 63,91.

4.7 ABLEITUNG POPULATIONSGENETISCHER PARAMETER

4.7.1 SEQUENZVERGLEICH DER INTIMINUNTEREINHEITEN

Zur Charakterisierung des analysierten Sequenzdatenpools wurden unter Verwendung verschiedener Computerprogramme und -algorithmen Analysen zum Nachweis des Selektionsmechanismus und der genetischen Rekombination vorgenommen. Dabei berücksichtigten die Untersuchungen die Einteilung in konservierte und hypervariable Genabschnitte und die funktionellen periplasmatischen (PP), transmembranösen (TM) und extrazellulären Domänen (D0 bis D3). Die ermittelte Zonierung ist in der **Abbildung 4.7.1-1** unter Verwendung des Proteinsequenzvergleichs wiedergegeben. ζ -Intimin diente als Bezugssequenz. Hierzu identische Aminosäuren wurden durch Punkte repräsentiert, heterologe in Form des Standardcodes wiedergegeben. Die **Abbildung 4.7.1-3** veranschaulicht polymorphe Sequenzabschnitte anhand schwarzer Zonierungen. Zur Erstellung der Genkartierung diente das Alignment der als Berechnungsgrundlage eingesetzten 25 Nukleotidsequenzen. Nachfolgend sind die anhand der Vorgaben von McGRAW et al. (1999) und LUO et al. (2000) ermittelten und in den beiden genannten Abbildungen veranschaulichten Einteilungen als Übersicht zusammengefasst.

Domänen	Position innerhalb des Alignments*	
	Nukleotid-	Proteinsequenz
Konservierter 5'-Bereich	1-2112	1-704
Hypervariabler 3'-Bereich	2113-2856	705-952
PP	1-560	1-186
TM	561-1650	187-550
D0	1677-1970	559-656
D1	1971-2251	657-756
D2	2252-2516	757-846
D3	2517-2856	847-952

* Bezugssequenz: ζ -Intimin

Auf Basis des über den ClustalW 1.6 berechneten Sequenzvergleichs konnte eine Sequenz-Übereinstimmungsmatrix erstellt werden (**Tabelle 9.1.1-2**), die folgende Aussagen über die Ähnlichkeit der 25 berücksichtigten Intiminuntereinheiten und -subtypen ermöglicht: Die Außengruppen *E. coli* O115 und *C. freundii* zeigen eine Übereinstimmung von 0,99 (99%). Mit 0,87 (87%) besteht die größte Ähnlichkeit mit β 2- und κ -Intimin. Die Untergliederung von α -, β -, γ - und ε -Intimin in jeweils zwei Subtypen wurde durch eine hohe Übereinstimmung der betrachteten Sequenzen widergespiegelt, die für alle vier Varianten oberhalb von 0,91 (91%), bei α -Intimin sogar bei 0,96 (96%) lag. Die Relation von α -// ζ -Intimin lag bei >91% für die vollständig sequenzierten ζ -Referenzen und bei 87% für die in den eigenen Arbeiten zu 95% generierten Serovare O84:H⁻ (AY52094) und O156:H⁻ (AY520905). Auf die Relation der ζ -Intimingene wurde im Abschnitt „**Sequenzierung des eae-Gens**“ eingegangen. Innerhalb des Datenpools wurde eine Übereinstimmung >90% ebenfalls zwischen β 2-// κ -Intimin (94%), γ 1- (91%) bzw. γ 2 (99,6%)// τ -Intimin, γ 1- (92%) bzw. γ 2- und τ -(92%)// σ -, Intimin erreicht. ε 1- (92%) bzw. ε 2 (93%)// η -Intimin, ν -// \omicron - (92%) und ξ -// π - (92%) lagen ebenfalls oberhalb des genannten Schwellenwertes. δ -Intimin besaß als parziell vorliegende Sequenz mit 24 und 23% die größte Übereinstimmung mit den κ -, β 2 und β 1-Varianten. Ein ausschließlicher Vergleich der übereinstimmenden Genbereiche ergab Identitätsraten von 94 und 93%. ρ - und λ -Intimin wiesen überwiegend einheitliche Relationen von etwa 83% - mit den Maximalwerten bei γ - bzw. bei λ -Intimin - mit dem Maximalwert von 88% bei ι 1 sowie mit 86 bis 87% bei α -, β 2 und γ auf.

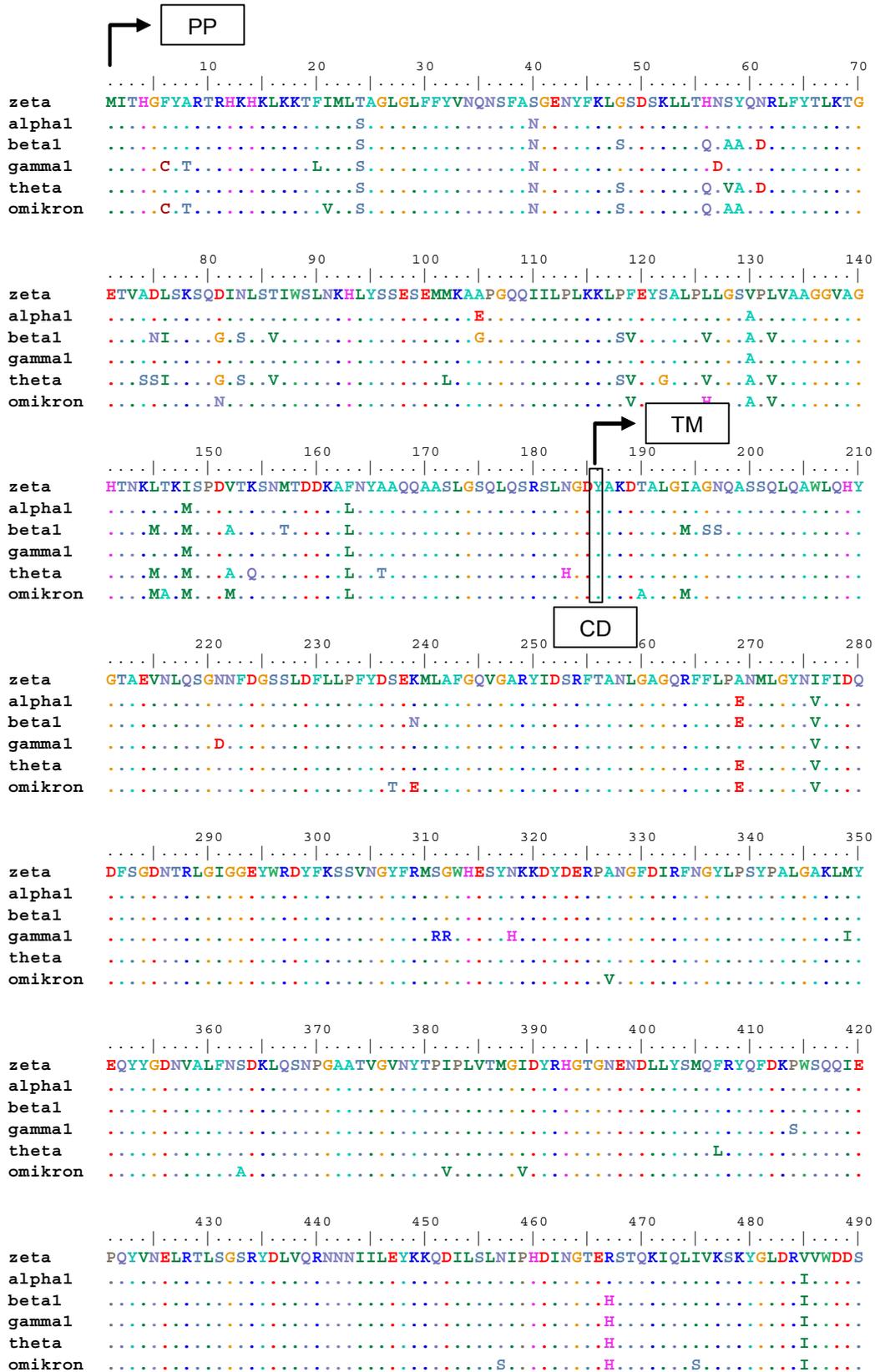


Abbildung 4.7.1-1:

Sequenzvergleich der Proteinsequenzen. Hypervariable, konservierte und funktionelle Regionen werden unter Verwendung des zeta (ζ)-Intimins als Referenz dargestellt. (Seite 1 von 2)

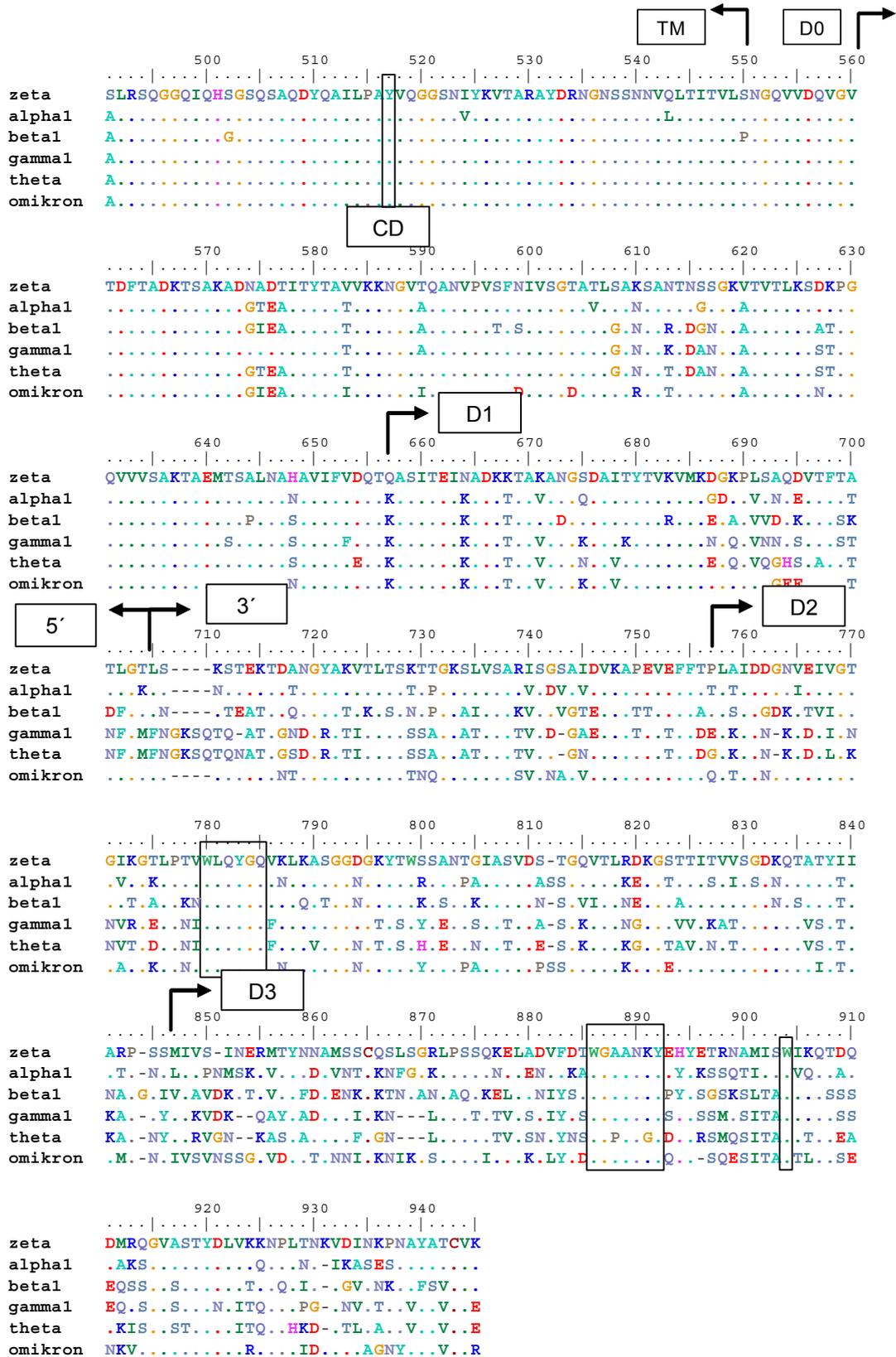


Abbildung 4.7.1-1: Seite 2 von 2

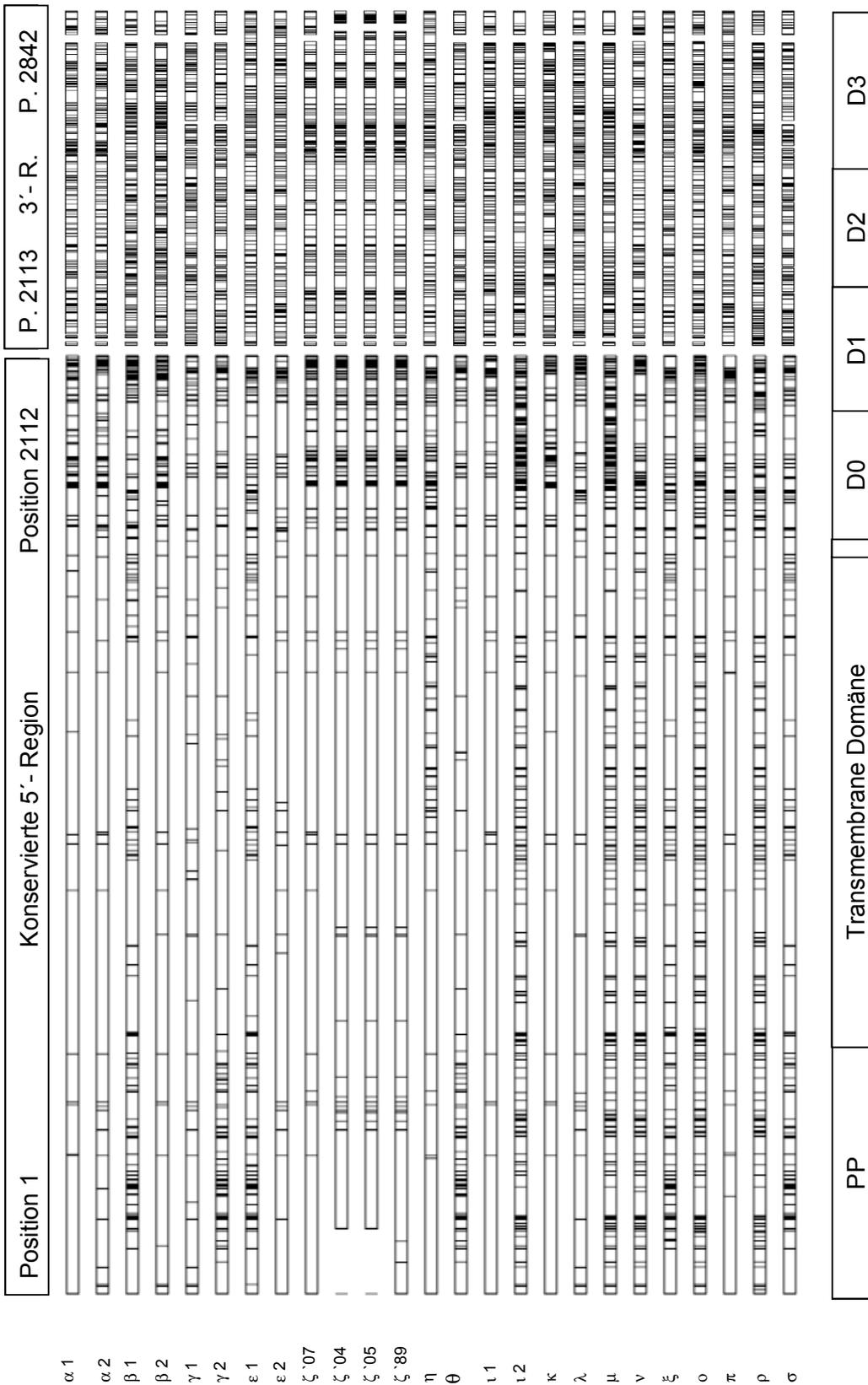


Abbildung 4.7.1-2: Visualisierung polymorpher Nukleotidpositionen innerhalb des Sequenzvergleichs. Hypervariable, konservierte und funktionelle Regionen dienen der Orientierung.

4.7.2 CHARAKTERISIERUNG DES SEQUENZVERGLEICHS DER REFERENZSEQUENZEN

Der anhand von 25 Referenzsequenzen, die mit *E. coli* O115 und *Citrobacter freundii* zwei Vorläufer in der Intiminentwicklung beinhalteten, erstellte Sequenzvergleich wurde manuell überprüft und anhand des G+C-Gehaltes, der Anzahl unterschiedlicher Mutationen und der Alignmentpositionen ohne Lücken charakterisiert. Letztgenannter Parameter gibt Auskunft über die Qualität des Alignments, da ein möglichst hoher Sequenzanteil der einzelnen Intiminvarianten in die Berechnungen einfließen muss. Insgesamt umfasste der Sequenzvergleich 2858 kodierende Positionen mit 2771 (96,96%) Positionen nach Abzug der 87 (3,04%) Lücken, d. h. Alignmentlokalisationen, die vollständig oder teilweise keine Nukleotide beinhalteten). Der G+C-Gehalt betrug im Gesamtdurchschnitt 42,4%. Die Anzahl variabler sog. polymorpher Positionen wurde mit $S=1050$, die Gesamtzahl der Mutationen mit $\text{Eta}=1607$ ermittelt. Die polymorphen Positionen (S) überwiegen mit 559 im konservierten 5'-Bereich im Gegensatz zu 491 Lokalisationen im hypervariablen 3'-Abschnitt (**Tabelle 4.7.2-1**). Während im konservierten Abschnitt die Anzahl der Positionen mit mindestens zwei identischen Nukleotiden (SPM (2) und PIPM (2)) überwiegt, ist die Anzahl der Positionen mit 3 und 4 identischen Nukleotiden mit Ausnahme der SPM (4) im hypervariablen Genbereich größer.

Über die Nukleotidfrequenzen gibt das Diagramm in **Abbildung 4.7.2-1** Auskunft. Hier wurde die Verteilung der Nukleotidbasen Adenin, Cytosin, Guanin und Thymin in Verbindung mit der Desoxyribose und dem dazugehörigen Phosphatrest im gesamten Sequenzbereich und unter Berücksichtigung der Positionierung an erster, zweiter oder dritter Stelle innerhalb des Codons dargestellt.

Tabelle 4.7.2-1: Verteilung der polymorphen Alignmentpositionen unter Berücksichtigung des Genabschnittes

Alignment positionen	Anzahl der Alignmentpositionen pro Genabschnitt	
	Konservierter 5'-Bereich	Hypervariabler 3'-Bereich
SPM (2)	63	31
PIPM (2)	397	153
SPM (3)	0	4
PIPM (3)	81	170
SPM (4)	0	0
PIPM (4)	18	133
Gesamtzahl	559	491

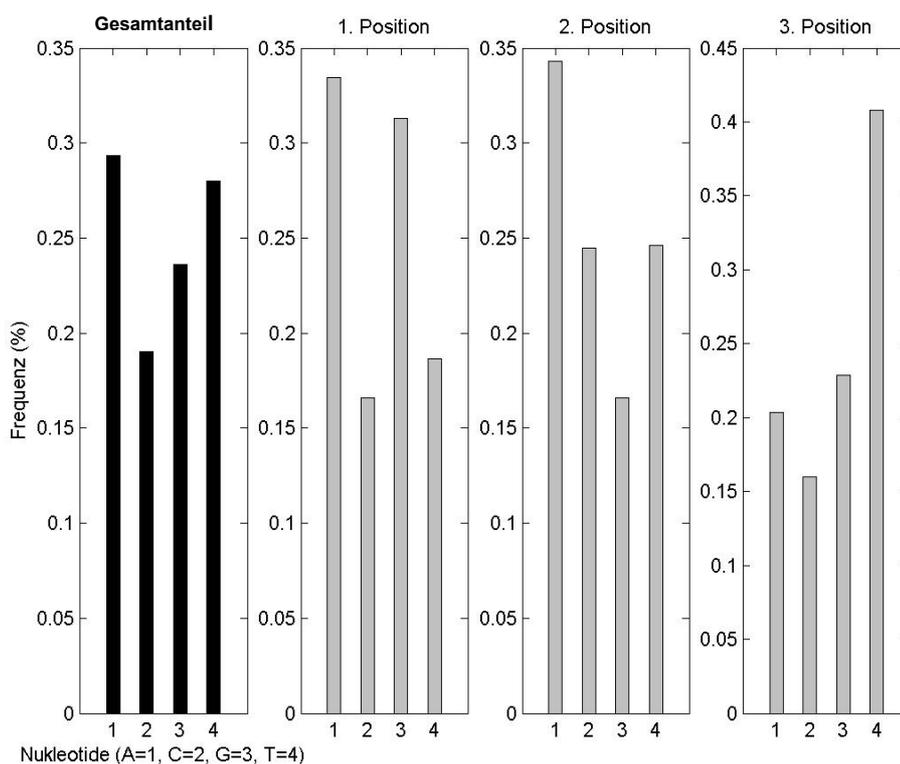


Abbildung 4.7.2-1: Nukleotidfrequenzen [%] innerhalb der kodierenden Sequenzregionen

4.7.3 DARSTELLUNG DES SELEKTIONSMECHANISMUS

Für die Darstellung des Selektionsmechanismus wurden das Verhältnis der Transitionen und Transversionen im 5'- und 3'-Sequenzbereich ermittelt. Die sich anschließende Kalkulation der Parameter p_s (K_s)= Anzahl der synonymen Substitutionen/Anzahl synonyme Alignmentpositionen und p_N (K_a)= Anzahl der nonsynonymen Substitutionen/Anzahl nonsynonymer Alignmentpositionen wurde zur Abklärung des Vorliegens einer „neutralen Selektion ($p_N=p_s$)“, der „negativen Selektion ($p_N<p_s$)“ und der „positiven Selektion ($p_N>p_s$)“ verwendet. Die Analyse erfolgte dabei für ζ -Intimin als abschnittsweise (engl. „sliding window“)-Kalkulation durch Betrachtung von jeweils 30 Basen und in Form des Gesamtdurchschnitts bezogen auf die funktionellen Domänen. In einem zweiten Schritt wurden zusätzlich nonsynonyme Mutationen im Hinblick auf die Veränderungen der Eigenschaften „elektrische Ladung“ und „Polarität“ des Proteins überprüft.

Die Betrachtung der Anzahl der Transitionen und der Transversionen gegen die Sequenzdistanzen ergab folgendes Bild: Im 5'-Bereich war bereits ab einer Distanz von 0,03 ein geringgradiges, letztlich bis zu einer Differenz von 0,025 ansteigendes Überwiegen der Transitionen erkennbar (**Abbildung 4.7.3-1**). Im 3'-Sequenzabschnitt lag die Zahl der Transversionen bis zu einer Distanz von ca. 0,4 oberhalb der Transitionen. Die Differenz lag zwischen 0,02 und 0,1. Der in der **Abbildung 4.7.3-2** dargestellte Sachverhalt berücksichtigte durch die gewählte Skalierung die größere Distanz und die höhere Anzahl der Transitionen und Transversionen im hypervariablen Bereich.

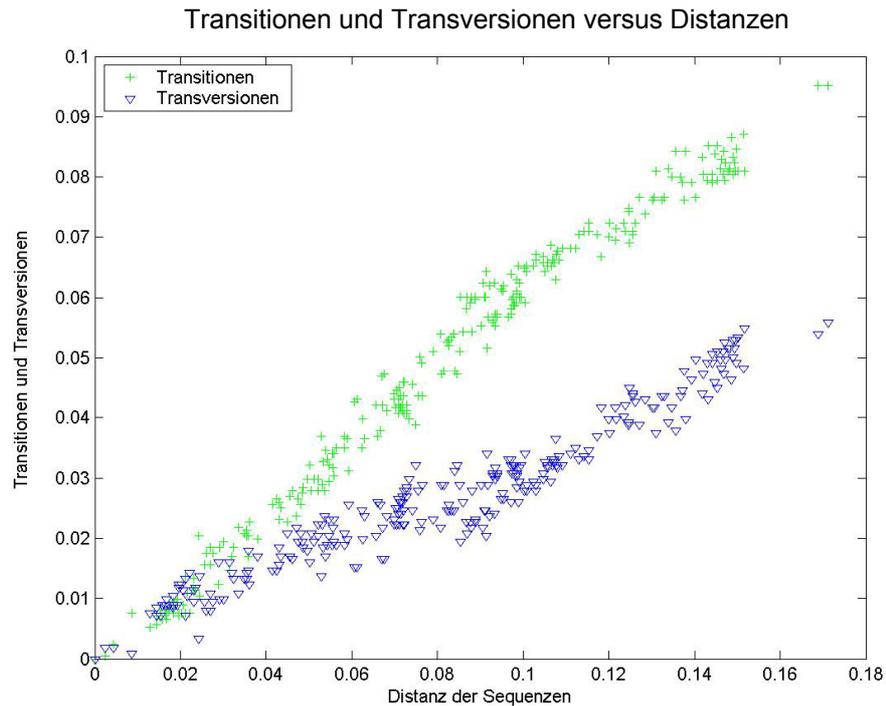


Abbildung 4.7.3-1: Darstellung der Anzahl der Transitionen und Transversionen in Abhängigkeit von der Distanz im 5'-Sequenzbereich

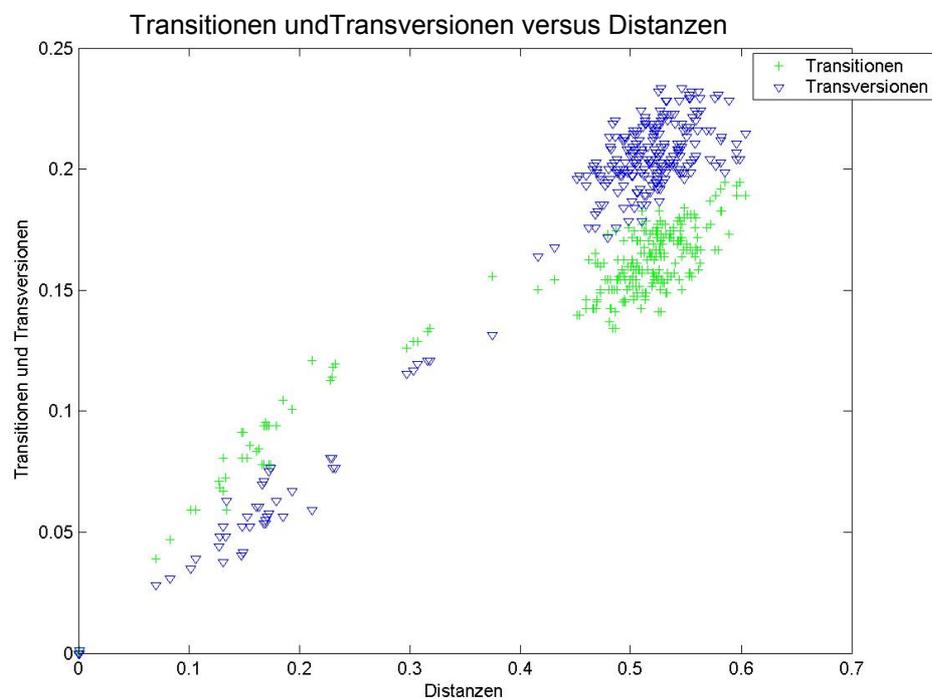


Abbildung 4.7.3-2: Darstellung der Anzahl der Transitionen und Transversionen in Abhängigkeit von der Distanz im 3'-Sequenzbereich

Die abschnittsweise Ermittlung der Parameter p_S (K_s) und p_N (K_a) mit der modifizierten Methode nach NEI und GOJOBORI (NEI und GOJOBORI, 1986; ZHANG et al., 1998) wurde grafisch für den Vergleich des ζ -Intimings mit den phylogenetisch eng verwandten α -, ν - und o -Intiminvarianten durchgeführt (**Abbildung 4.7.3-3**, **Abbildung 4.7.3-4**). Im Vergleich mit α -Intimin überwogen mit Ausnahme des Sequenzbereiches zwischen Codon 1 bis 80 und 460 bis 560 die synonymen Proportionen (p_S). Ein deutlicher Anstieg der p_S -Werte wurde erst ab Position 570 und über den gesamten hypervariablen Bereich deutlich (ab Position 704). Gegenüber ν - und o -Intimin wurde über den gesamten Sequenzbereich eine insgesamt höhere p_S -Frequenz erkennbar. p_N überwiegt in der Gegenüberstellung ν - und o -Intimin ebenfalls im Positionsbereich 60-80, lokalisiert in der periplasmatischen Domäne (PP). Ab Position 930 überwog p_S gegenüber p_N . Im Vergleich mit allen drei Referenzsequenzen war ein p_N -Anstieg im Codonbereich zwischen 850 bzw. 880 (α -Intimin) im Grenzbereich der extrazellulären Domänen D2 und D3 erkennbar. In allen betrachteten Funktionsabschnitten gilt mit Ausnahme der genannten Regionen $p_S > p_N$ ein überwiegend negativer, bereinigender Selektionsdruck.

Tabelle 4.7.3-1: Variation der Verhältnisse von p_S und p_N [s] in den Genabschnitten und funktionellen Domänen (Einteilung nach LUO et al., 2000)

Genabschnitt, Domäne	$p_S \times 100$ [s]	$p_N \times 100$ [s]	Verhältnis
5'-Abschnitt	15,0 [0,8]	4,0 [0,4]	3,8
3'-Abschnitt	52,9 [1,0]	28,5 [1,6]	1,9
Gesamtes Gen	25,0 [1,0]	10 [1,0]	2,5
PP	14,2 [1,6]	3,8 [0,2]	3,7
TM	10,0 [0,9]	1,1 [0,2]	9,1
D0	27,1 [2,4]	8,6 [1,3]	3,2
D1	49,0 [2,4]	21,3 [2,3]	2,3
D2	51,3 [2,4]	23,5 [2,4]	2,2
D3	53,7 [2,6]	35,4 [2,7]	1,5

Bezogen auf die funktionellen Domänen (**Tabelle 4.7.3-1**) unterschieden sich die betrachteten Abschnitte in den absoluten Werten für „ p_S und p_N “ deutlich. Dabei überwog p_S . Das Verhältnis p_S/p_N sank im hypervariablen 3'-Bereich im Vergleich zur konservierten 5'-Region um 50% (von 3,8 auf 1,9) bei einem gleichzeitigen 3,5fachen p_S - und 7,1fachen p_N -Anstieg. Die transmembranöse Domäne (TM) verfügte mit einem Verhältnis von 9,1 über die höchste Gewichtung der synonymen p-Distanz. Absolut betrachtet lagen die Werte für p_S und p_N unterhalb denen der anderen Domänen. Innerhalb der extrazellulären Domänen „D0 bis D3“ nahm die Zahl der Substitutionen ab „D1“ (Faktoren 1,5 für p_S und 2,4 für p_N) zu. Das Verhältnis p_S/p_N sank von 3,2 (D0) auf 1,5 (D3), was durch die insgesamt 4,1fache Zunahme des p_N -Wertes hervorgerufen wurde. Während die Domänen „D1“ und „D2“ beide Parameter nur etwa um das 1,05- bis 1,1fache anstiegen, betrug der Faktor für die Zunahme von p_N von D auf D3 das 1,5fache (Faktor p_S 1,04). Das Vorliegen der negativen Selektion mit $p_S > p_N$ konnte anhand des exakten Testes auf Unabhängigkeit (FISHER, 1973) mit $p > 0,05$ bestätigt werden.

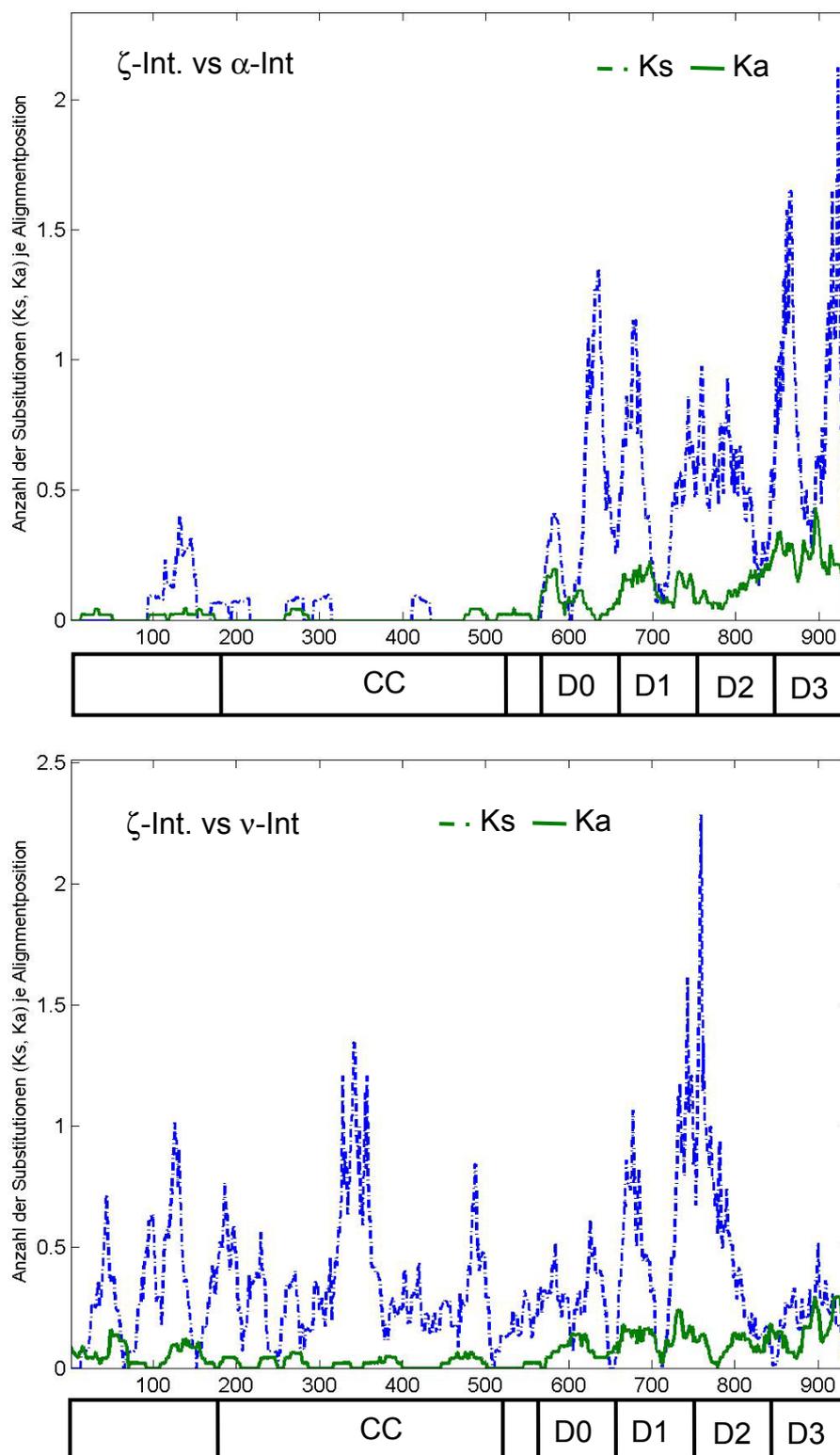


Abbildung 4.7.3-3: Proportion der synonymen (Ks)- und nonsynonymen (Ka)-Substitutionen des zeta (ζ)-Intimingens im Vergleich mit den alpha (α)- und nu (ν)-Intimingenen

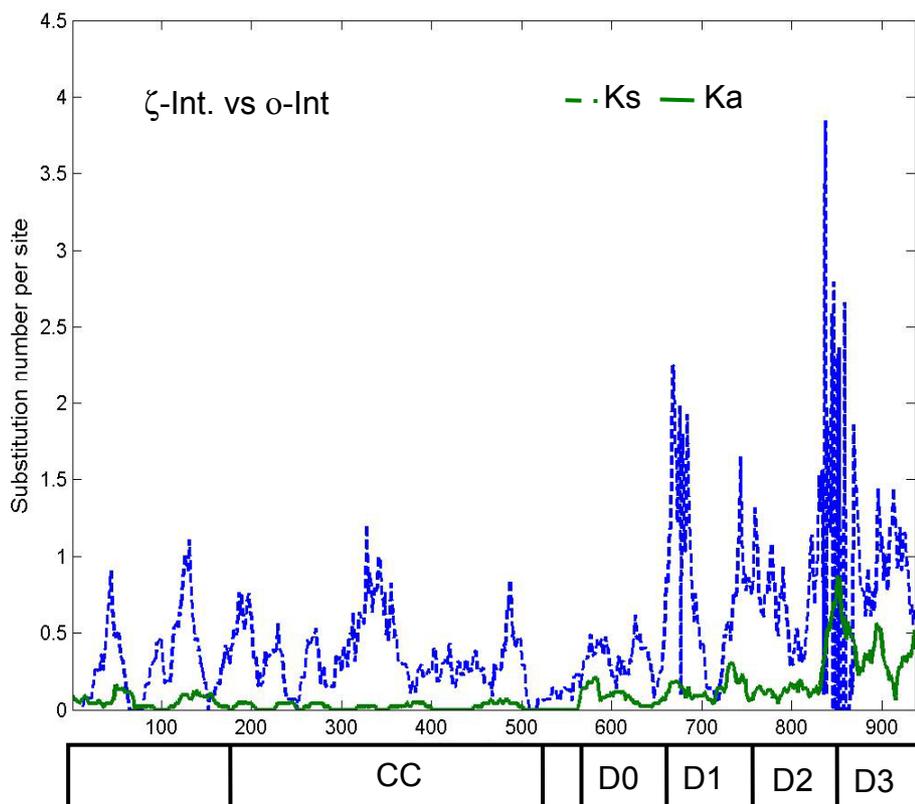


Abbildung 4.7.3-4: Proportion der synonymen (Ks) und nonsynonymen (Ka)-Substitutionen des zeta (ζ)-Intimings im Vergleich zum omikron (o)-Intimingen

Ein weiterer Faktor für die Art des Selektionsdruckes stellt die mögliche Veränderung der Proteineigenschaften Ladung und Polarität und damit die Funktionalität des Anheftungsfaktors dar. Die anhand des Alignments für den konservierten und den hypervariablen Sequenzbereich durchgeführten Berechnungen berücksichtigten für SCR3 (HUGHES et al., 1990; HUGHES und FRIEDMANN, 2000) und HON-NEW (ZHANG, 2000) das zuvor ermittelte Verhältnis der Transitionen/Transversionen R von 1.2. Beide Algorithmen ergaben übereinstimmende Resultate, daher wurden für die graphische Auswertung nur die mit HON-NEW errechneten Werte p_{NR} gegen p_{NC} für die Parameter Ladung und Polarität wiedergegeben.

Um die Übersichtlichkeit zu erhöhen, wurden nur die in den **Tabellen 9.1.3-1 bis 4** gelisteten Daten berücksichtigt. Unterschiede zu den anderen Sequenzen bestanden nicht. In der hypervariablen 3'-Region überwogen Substitutionen, die zu einer Ladungsveränderung führten. Im Unterschied dazu wurden unter dem Gesichtspunkt Polarität konservative nonsynonyme Mutationen (p_{NC}) bevorzugt, die nicht zu einer Änderung des Status quo führen. Im konservierten 5'-Abschnitt überwogen für beide Parameter konservative Veränderungen. Die in der Grafik dargestellte Winkelhalbierende repräsentiert gleichgewichtete Mutationsparameter (**Abbildung 4.7.3-5**).

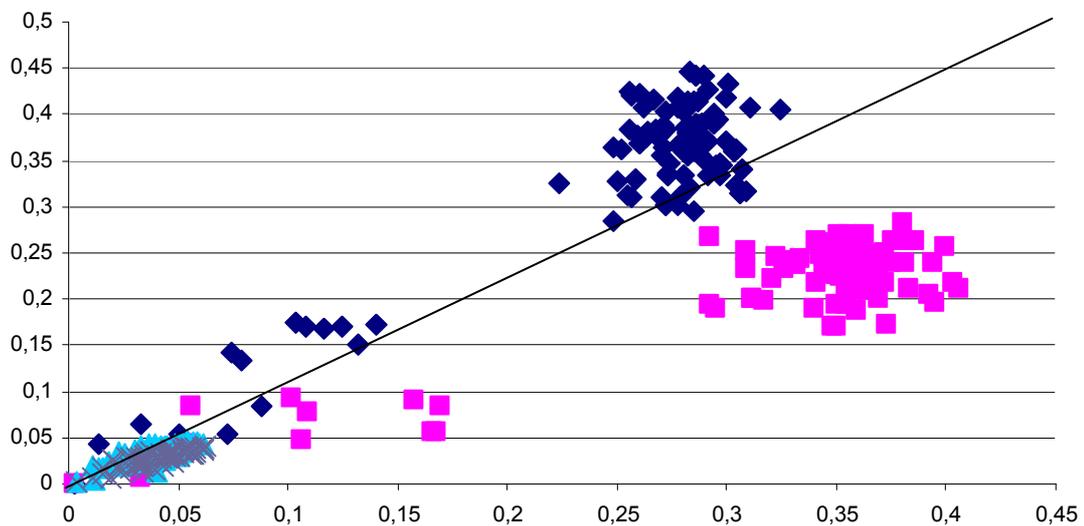


Abbildung 4.7.3-5: Relation der radikalen nonsynonymen Substitutionen (p_{NR}) gegen die konservativen nonsynonymen Substitutionen (p_{NC})

- 3'-Ende: ◆ Ladung
 ■ Polarität
- 5'-Ende: ▲ Ladung
 X Polarität

4.7.4 GENETISCHE REKOMBINATION

Die globale Überprüfung des Sequenzvergleichs ergab sechs als „rekombinant“ im Sinne der Voraussetzung „ $p > 0,05$ “ eingestufte Regionen innerhalb des ζ -Intimingens (**Tabelle 4.7.4-1**). Die Regionen zwischen den Nukleotidpositionen „1819 bis 2292“ und „1737 bis 2292“ konnten jeweils durch fünf Verfahren (Bootscan, Chimaera, MaxChi, RPD und SiScan), der Abschnitt „1014 bis 1590“ durch vier Modelle (Bootscan, Chimaera, MaxChi und RPD) identifiziert werden. Mit MaxChi und SiScan wurde auch der Bereich „2208 bis 2805“ als „wahrscheinlich rekombinant“ eingestuft. Mittels GeneConv wurden die Bereiche von „0-1034“ und „52 bis 1719“ als rekombinant eingestuft. Eine Bestätigung durch andere Algorithmen konnte in der manuellen Überprüfung nicht gezeigt werden. Die den jeweiligen Regionen zugeordneten Vorläufersequenzen präsentierten sich uneinheitlich. Folgende Zuordnungen wurden für die vier mehrfach identifizierten Abschnitte getroffen: Region „1014 bis 1590“: beta1 (2x), gamma 1 (n=1), iota1 (n=2), pi (n=1) und *E. coli* O115 (n=2). Für den Abschnitt 1737-2292: beta1 (n=2), epsilon1 (n=1), iota1 (n=2), omikron (n=3), rho (n=2) und *E. coli* O115 (n=1). Für die Partition „1819-2307“ gamma1 (n=1), eta (n=1) lambda (n=1), pi (n=1), xi (n=3) und *C. freundii* (n=3). Als wahrscheinliche evolutive Vorläufer für den Sequenzbereich „2208 bis 2805“ wurden beta1 (n=1), epsilon1 (n=2) und omikron (n=1) angegeben.

Die k_{\max} -Lokalisation unter Angabe der paarweisen Sequenzdifferenz im Bezug zu ζ -Intimin sind der Grafik in **Abbildung 4.7.4-1** zu entnehmen. Im paarweisen Sequenzvergleich konnten für α 1-Intimin sechs, für β 1- und γ 1-Intimin je acht und o-Intimin fünf „wahrscheinliche Bruchpunkte (K_{\max})“ mit $p > 0,01$ identifiziert werden. Dabei wurde ein mosaikartiger Aufbau der Intimingene erkennbar. Neben den Positionsangaben der Bruchpunkte wurden die prozentualen Sequenzunterschiede der jeweiligen Referenz im Vergleich mit ζ -Intimin angegeben. Bei allen Sequenzen konnten wahrscheinliche rekombinante

Regionen im Alignment an den Positionen 1691 ($\alpha 1$), 1715 ($\beta 1$), 1736 ($\gamma 1$) und 1744 (\circ) ermittelt werden. Für die Varianten „ $\alpha 1$ “ und „ $\beta 1$ “ wurden wahrscheinliche Rekombinationsereignisse auch im Positionsbereich 2241 bzw. 2242 berechnet. Rekombinante Ereignisse wurden in Übereinstimmung mit der globalen Analyse unter Berücksichtigung des gesamten Alignments auch für die Positionen „979, 990 und 1066“ (Intiminvarianten $\alpha 1$, $\beta 1$ und $\gamma 1$) ermittelt. Die Sequenzdistanzen stiegen dabei entsprechend der k_{Max} -Lokalisation im 3'-Bereich an. Hervorzuheben sind allerdings die Positionsabschnitte „128 bis 360 bis 460“ bei $\beta 1$ -Intimin und „1-178“ bei 0-Intimin. Hier ergaben sich Distanzen von 12,5%, 14,9% und 13,9%, eingebettet in sich anschließende Regionen mit deutlich geringeren Sequenzunterschieden (2,4%, 4,8% und 7,6%).

Tabelle 4.7.4-1: Mögliche rekombinante Regionen innerhalb des ζ -Intimings mit $p > 0,05$

Algorithmus	Position der rekombinanten Region		Mögliche Vorläufersequenz
	Anfang	Ende	
Bootscan	1014	1590	gamma1, pi
	1737	2292	beta1, iota1?
	1819	2307	<i>C. freundii</i> , xi
Chimaera	1014	1590	<i>E. coli</i> O115, beta1
	1737	2292	rho, omikron?
	1819	2307	<i>C. freundii</i> , pi?
GeneConv	0	1034	eta, ?
	520	1719	theta, ?
MaxChi	1014	1540	<i>E. coli</i> O115, iota1?
	1737	2292	rho, omikron?
	1819	2307	xi, eta?
	2208	2266	epsilon1, beta1
RDP	1014	1590	beta1, kappa
	1737	2292	beta1, <i>E. coli</i> O115
	1819	2324	lambda, <i>C. freundii</i>
	2208	2805	kappa, omikron
SiScan	1737	2292	omikron, epsilon1
	1819	2307	gamma1, xi

?: Identifizierung der wahrscheinlichen Vorläufersequenz nicht eindeutig

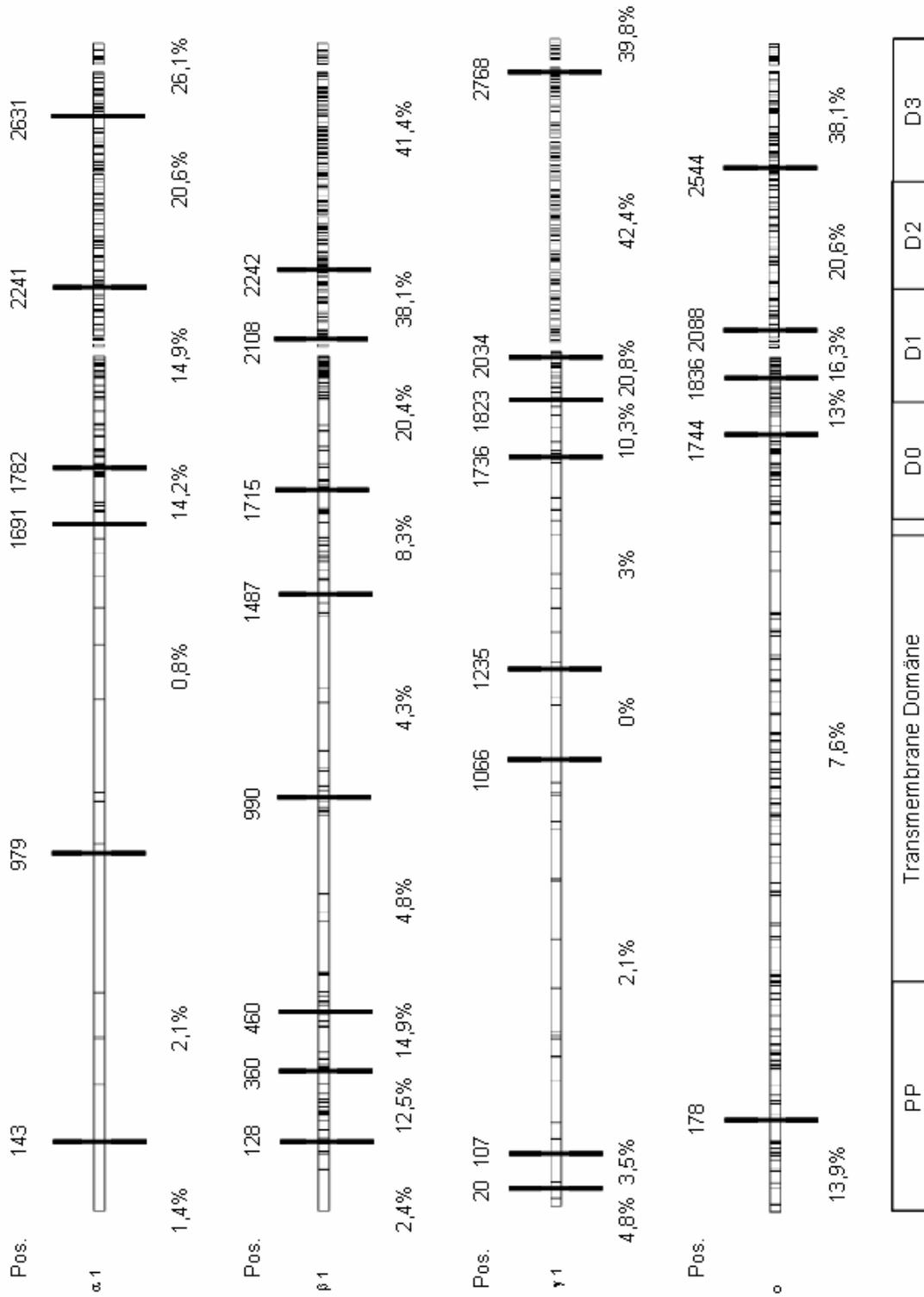


Abbildung 4.7.4-1: Manuelle Überprüfung wahrscheinlicher rekombinanter Regionen ($p > 0,01$) im paarweisen Sequenzvergleich mit ζ-Intimin

4.8 DARSTELLUNG DER ZELLULÄREN INTERAKTION

4.8.1 VISUALISIERUNG DER UNSPEZIFISCHEN, FIMBRIEN-VERMITTELTEN ADHÄSION

Mit der beschriebenen Versuchsanordnung konnte die unspezifische und über Fimbrien vermittelte Adhäsion des Referenzkeimes EPEC E 2348/69 O127:H6 an humane Mundschleimhautzellen gezeigt werden. Dargestellt ist jeweils eine Mundschleimhautzelle mit den angelagerten *E. coli*-Zellen (**Abbildung 4.8.1-1**). Im linken Ausschnitt sind anhand der Zellränder zwei überlagerte Mundschleimhautzellen erkennbar. Der Zellkern der aufliegenden Zelle wurde deutlich dargestellt. Im rechten Bild ist eine separate Mundschleimhautzelle erkennbar.

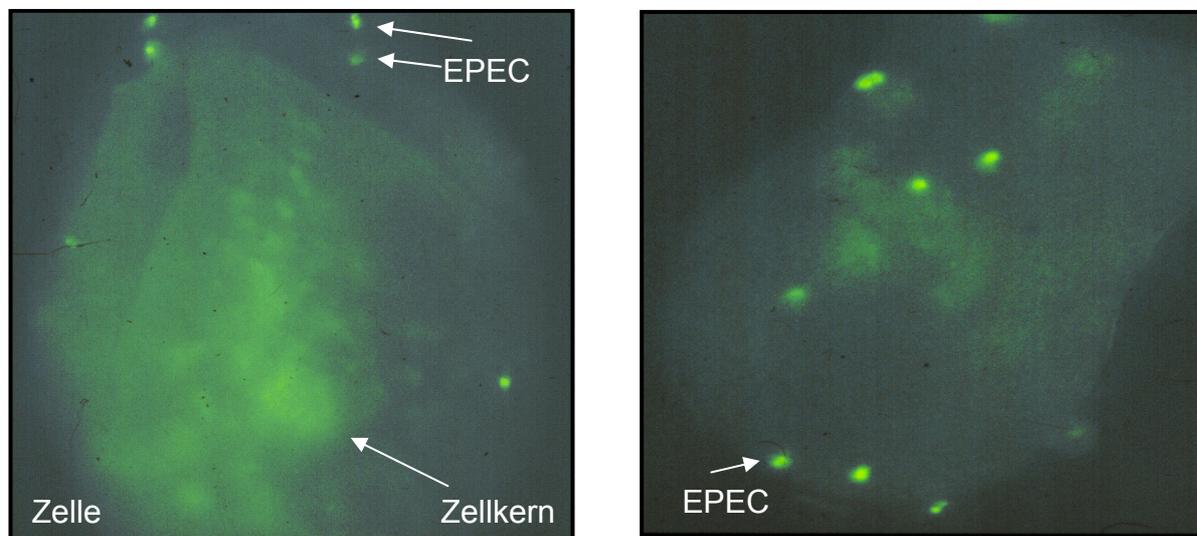


Abbildung 4.8.1-1: Unspezifische, fimbrienvermittelte Adhäsion von EPEC O127:H6 an humane Mundschleimhautzellen (Fluoreszenzmikroskop, Vergrößerung: 400fach)

4.8.2 VISUALISIERUNG DER SPEZIFISCHEN, INTIMIN-VERMITTELTEN ADHÄSION

Die **Abbildung 4.8.1-1 bis 3** dokumentieren die spezifische, intiminvermittelte Anlagerung der *E. coli*-Zellen nach Blockade der unspezifischen Adhäsionsmechanismen. Die Aktinpolymerisation unterhalb der Zelle und die parallele Ausrichtung der Aktinfasern als sog. Stressfasern sind deutlich erkennbar.

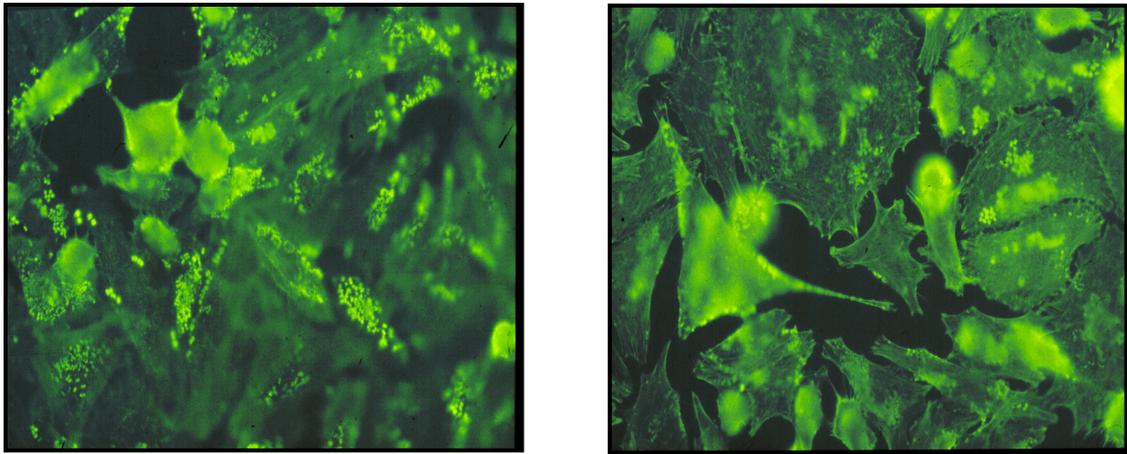
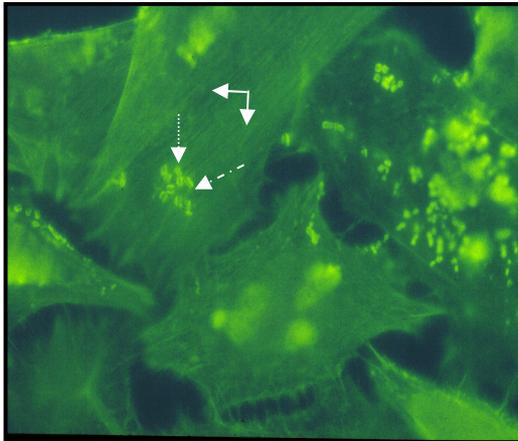


Abbildung 4.8.2-1: Lokalisierte Adhäsion des EPEC E2348/69 O127:H6 an HEp2-Zellen. (links: Vergrößerung: 400fach; rechts: Teilausschnitt, Vergrößerung: 600fach)



- ↖ Stressfasern
- ⋮ Aktinpolymerisation
- ▲ Sitz des EPEC E 2348/69
O127:H6

Abbildung 4.8.2-2: Darstellung der Stressfasern und der Aktinakkumulation.
Zentrale Lücken innerhalb der Aktinhügel sind der Sitz des
EPEC E 2348/69 O127:H6

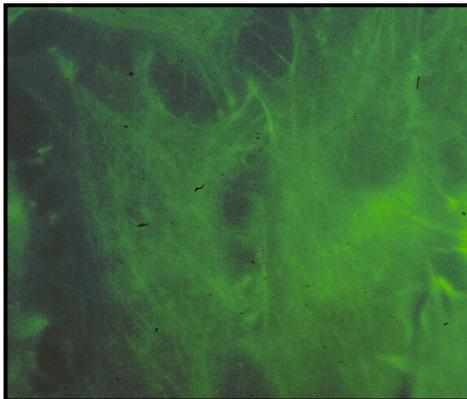


Abbildung 4.8.2-3: Kontrolle: Unbeimpfte HEp2-Zellen

5 DISKUSSION

In den eigenen Untersuchungen sollten die *eae*-Gene von *E. coli* aus unterschiedlichen Habitaten typisiert werden. Neben der Etablierung eines einfachen, routinetauglichen PCR-Verfahrens stand die Darstellung der phylogenetischen Relation unter Einbeziehung der genetischen Rekombination und des auf die *eae*-Varianten einwirkenden Selektionsmechanismus im Vordergrund. Hierfür wurden vergleichend verschiedene Evolutionsmodelle berücksichtigt. Zusätzlich wurden zwei *E. coli*-Isolate mit bislang nicht typisiertem Anheftungsfaktor sequenziert. Modellhaft wurde die Adhäsion des EPEC-Referenzstammes E 2348/69 an Zellkultursystemen dargestellt.

5.1 TYPISIERUNG DER *EAE*-GENE

Für den allgemeinen Nachweis des *eae*-Gens wurde das Oligonukleotid-Primerpaar „SK1/SK2“ (KARCH et al., 1993) in der konventionellen PCR eingesetzt. In der vergleichenden Sequenzanalyse konnte zuvor die sichere Amplifikation aller zum Zeitpunkt der Untersuchungen über das Internet verfügbaren *eae*-Genvarianten durch 100%ige Übereinstimmung der jeweils auf die Zielsequenz abgestimmten ersten 17 Basen der Primer-Sequenzen (421-2) gezeigt werden. Lediglich für die Intiminuntereinheiten delta (δ), iota (η) 2, xi (ξ) und omikron (\omicron) konnten keine exakten Aussagen getroffen werden, da für diese Varianten in den Sequenzdatenbanken keine vollständig sequenzierten Gene zugänglich waren. Entsprechende Daten standen erst zum Zeitpunkt der phylogentischen Auswertung zur Verfügung. Dabei zeigten sich keine Abweichungen. Für den Einsatz der Primer-Kombination ist die Verwendung sog. „Hot start“-Polymerasen wie der AmpliTaq Gold[®] DNA Polymerase mit GeneAmp[®] 6xPuffer (15 mM MgCl₂, Fa. Applied Biosystems, Produkt-Nr. 4311814) vorzuziehen, da in parallel durchgeführten Versuchen mit konventionellen Taq-Polymerasen (hier: AmpiTaq, Produkt-Nr. N808-0160, Fa. Applied Biosystems) uneinheitliche Fehlbanden auftraten.

Falsch negative Reaktionen waren nicht zu erwarten, da alle bekannten Sequenzen im dargestellten Bereich hochkonserviert und somit identisch sind. In der Praxis bestätigte sich die hohe Zuverlässigkeit des Primer-Paares, da alle als *eae*-positiv ausgewählten Isolate des Datensatzes reproduzierbare Ergebnisse lieferten. Zusätzlich zu den eigenen Ergebnissen ist die Routinetauglichkeit mehrfach belegt worden. So untersuchten SCHMIDT et al. (1994) und BÜLTE (2001) 122 bzw. 150 *E. coli*-Stämme ohne Berücksichtigung der Intiminvarianten. OSWALD et al. (2000), GALLIEN et al. (2000) und ZHANG et al. (2002b) belegten die Detektion einschließlich der Varianten eta, iota, kappa und theta. BEUTIN et al. (2004 und 2005) verwendeten zusätzlich eine DNA-Hybridisierungssonde, die den Sequenzbereich zwischen den beiden Oligonukleotiden enthält und erzielten übereinstimmende Ergebnisse. Dabei berücksichtigten BEUTIN et al. (2004) 677 Humanisolate. KRAUSE et al. (2005) untersuchten mehr als 2000 Isolate von Haussäugetieren.

Zum Zeitpunkt der Aufnahme der Arbeiten waren durch die Arbeiten von ADU-BOBIE et al. (1998a) und OSWALD et al. (2000) vier bzw. fünf Intiminvarianten beschrieben und molekulargenetisch gemäß der griechischen Nomenklatur charakterisiert worden. Dabei gelang der Arbeitsgruppe um OSWALD neben der Amplifikation von alpha (α)-, beta (β)- und gamma (γ) die Charakterisierung der fünften, der epsilon (ϵ)-Intiminvariante. Für delta (δ)-Intimin wurde keine Primer-Kombination entwickelt, sie wurde als β 2-Intimin eingeordnet. Die von beiden Autorengruppen eingesetzten PCR-Verfahren unterschieden sich in den Primer-Sequenzen und den resultierenden PCR-Produktgrößen (**Tabelle 3.3.8-1**, **Tabelle 3.3.8-2**). Beide Arbeitsgruppen verwiesen auf die wahrscheinliche Existenz weiterer Intiminvarianten. Der eigene Sequenzvergleich belegte die Möglichkeit zur spezifischen Detektion durch entsprechend geeignete Oligonukleotide. Der Vorwärts-Primer „Int-B“ ermöglichte auch die Detektion des nachfolgend durch OSWALD et al. (2000) beschriebenen β 2-Intiminsubtyps. Zusätzlich konnten die untersuchten Stämme auf das Vorhandensein von Intimin- δ überprüft werden. Der Rückwärts-Primer „Int-RU“ zeigte bei dieser Untereinheit eine Fehlbasenpaarung an Position 2 des 3'-Endes („A“ ersetzt durch „G“).

Im Versuch erwies sich dieser Austausch nicht als nachteilig, da die als β -Träger beschriebene Stämme (**Tabelle 9.1.1-1**) erfasst wurden. In den Vorversuchen wurde das Typisierungsverfahren nach ADU-BOBIE et al. (1998a) als robustes Prozedere etabliert. Kreuzreaktionen traten mit den in die Etablierung einbezogenen Referenzstämmen (**Tabelle 3.2.1-1**) und 100 Prüfstämmen nicht auf. Zusätzlich gewährleisteten die ausgewählten Restriktionsendonukleasen eine zuverlässige Differenzierung der PCR-Produkte. Alle Reaktionen konnten mit der im Vergleich kostengünstigeren AmpliTaq[®] DNA Polymerase Fa. Applied Biosystems, Produkt-Nr. N808-0152) durchgeführt werden. Die zusätzlich abgeleiteten Oligonukleotide „epsilonF/epsilonR“ amplifizierten, auch unter Berücksichtigung der später beschriebenen Intimine ζ - bis σ , selektiv ϵ -Intimin. Das 340 Bp-große Amplifikat beinhaltet zur Differenzierung von ϵ 1 und ϵ 2 relevante Anteile. BLANCO et al. (2005) stellten in der Genbank die Sequenzen von ϵ 3- und ϵ 4-Trägern bereit (Genbankzugriffsnummern: AJ876649 bis AJ876651, nur Sequenzeinträge publiziert), die ebenfalls amplifiziert werden können.

Die von OSWALD et al. (2000) eingesetzten Primer-Kombinationen erzielten insbesondere bei β - und ϵ -Intimin Fehlbanden im Größenordnungsbereich zwischen 800 und 1200 Bp. Eine Optimierung durch Variation der PCR-Konditionen gelang nicht. Für die Generierung der bis zu 2870 Bp-großen Amplifikate war ferner entgegen der Originalliteratur die Verwendung eines Polymerasemixes mit Proofreading-Eigenschaften erforderlich, da die Effizienz der eingesetzten *Taq*-Polymerasen nicht ausreichte (Daten hier nicht gezeigt). Das mit Proofreading-Aktivität versehene Expand High Fidelity PCR-System (Fa. Roche-Diagnostics, Produkt-Nr. 1732641) war für den geplanten Probenumfang zu kostenintensiv. GALLIEN et al. (2000) benötigten ebenfalls eine hochwertigere Polymerase.

Zusammenfassend lag der Vorteil des Verfahrens nach ADU-BOBIE et al. (1998a) in der Verwendung als einfache und kostengünstige Screeningmethode. Das diagnostische Potential wurde durch eigene Oligonukleotide um ϵ -Intimin und durch eine zusätzliche enzymatische Produktdifferenzierung zur Erfassung weiterer

Untereinheiten erweitert. Die Methodik nach OSWALD et al. (2000) offerierte aufgrund der den hypervariablen Bereich nahezu vollständig beinhaltenden Amplifikate, die zuverlässigere Ermittlung neuer Untereinheiten, schied aber aufgrund der mangelnden Robustheit und der hohen Kosten aus.

RAMACHANDRAN et al. (2003) offerierten mit der Etablierung eines RFLP-PCR-Verfahrens einen einfachen Ansatz zur Berücksichtigung der bis dahin bekannten Intimine α_1 , α_2 , β , γ , κ , ε , ν , ι , λ , τ und ζ . Zusätzlich klassifizierten sie μ -, ν - und ξ - Intimin. Die Arbeitsgruppe lieferte damit eine solide technische Grundlage. GEUE et al. (2004) belegten zudem die Routinetauglichkeit des Verfahrens, insbesondere auch zur Darstellung des ζ -Intimins bei der Typisierung von bovinen EHEC. Die auf Basis der Arbeit von OSWALD et al. (2000) veröffentlichte Weiterentwicklung durch ZHANG et al. (2002b) schied aufgrund der im oberen Abschnitt dargestellten Nachteile aus. Zielsetzung der eigenen Arbeiten war die Optimierung des RFLP-Verfahrens im Hinblick auf die notwendige Unterscheidung der Intimine β_1 und β_2 , γ_1 und γ_2 , ε_1 und ε_2 sowie die Einbeziehung von π -, ρ - und σ -Intimin. Bei der Überprüfung der durch RAMACHANDRAN et al. (2003) veröffentlichten Gendatenbankeinträge fiel auf, dass das als ξ -Variante beschriebene Rinderkotisolat Ont:Hnt (Genbankzugriffsnummer: AF530556) dort als β_2 -Subtyp eingetragen war. Analog fanden sich die Isolate Or:H⁻ (Ziegenkot, AF530553, publiziert als μ -Intimin) als ι_2 und O2 related H.19 (Ziegenkot, AF530554, publiziert als ν -Intimin) als ε_2 -Typ. Daher wurde für die Bestimmung des RFLP auf die vollständigen Sequenzen der Isolate O10:H⁻ (Sequenzeintrag: BLANCO et al., 2004, **Tabelle 9.1.1-1**) zurückgegriffen. Die vollständigen Sequenzen für μ -, ξ - und σ -Intimin standen zu diesem Zeitpunkt noch nicht zur Verfügung (vgl. Nomenklatur).

In der Überprüfung der Oligonukleotidsequenzen konnte „EaeVR“ als degeneriertes Oligonukleotid mit der Konfiguration „RAY“ (Leserichtung: 5'→3') des 3'-Endes bestehend aus „AAC, GAC, AAT“ bzw. „GAT“ die Amplifikation der Intiminvarianten alpha bis epsilon sowie kappa, lambda, rho und sigma durch vollständige Übereinstimmung am 3'-Ende gewährleisten.

„EaeZetaVR“ wurde spezifisch an die Sequenz für zeta-Intimin-Stämme adaptiert. Das gilt analog für „EaelotaVR“, der für die Amplifikation von iota-Intimin entwickelt wurde. Für die erweiterte Produktüberprüfung war die Neuentwicklung des Typisierungsschemas notwendig. Die Restriktionsenzyme *Fok* I, *Rsa* I und *Hae* III erbrachten dabei die eindeutige Bestimmung von 15 der 19 Untereinheiten (79%) durch mindestens zwei Polymorphismen. *Fok* I und *Rsa* I lieferten mit 16 sowie 14 eindeutig unterschiedenen Varianten das höchste Differenzierungspotenzial. Das bereits durch GANNON et al. (1993) verwendete *Hae* III sicherte die zuverlässige Unterscheidung durch sich stark unterscheidende Einzelprofile bei $\varepsilon 2$ -, $\iota 1$ und σ -Intimin als zweites Schlüsselenzym bzw. bei $\beta 1$ -, $\gamma 2$, μ , ρ und τ -Intimin tragenden Stämmen als drittes Enzym ab. Damit wurde die Notwendigkeit der dritten Restriktionsendonuklease zur Erweiterung des diagnostischen Aussagekraft unterstrichen.

Zusammenfassend konnte mit dem entwickelten RFLP-Verfahren die zusätzliche Differenzierung von sieben Intiminvarianten respektive -untereinheiten erreicht werden. Damit war ein Abgleich der Originalliteratur mit der aktualisierten Nomenklatur möglich. Im Vergleich zu JENKINS et al. (2003), die eine sog. „Long-Range“-RFLP-PCR-Verfahren unter Einsatz der Primer „Inf/IntR“ nach ADU-BOBIE et al. (1998a) mit vier Enzymen zur Erfassung von lediglich acht Varianten (bis ζ -Intimin) beschrieben, konnten mit nur drei Endonklesen im eigenen Verfahren 11 zusätzliche Varianten unterschieden werden.

5.2 SEQUENZIERUNG DES *EAE*-GENS

In den eigenen Untersuchungen konnten die *stx*-negativen EPEC O84:H⁻ (AY520905) und O126:H (AY520904) als ζ -Intiminträger identifiziert werden. Es konnten jeweils 2675 Bp (95%) des offenen Leserasters gewonnen werden. Die sequenzielle Übereinstimmung mit dem erstbeschreibenden STEC/VTEC O84.NM (entspricht O84:H⁻, JORES et al., 2003) beträgt über 99% (**Tabelle 4.5.1-2**). Das Serovar O156 (MANSFIELD et al., 2001, AF301015) war bislang lediglich als ε -Intimin tragender Stamm bei Rhesusaffen bzw. α -Intiminvariante aus Rinderkot isoliert worden (BLANCO et al., 2005). JORES et al. (2003) typisierten ζ -Intimin zusätzlich bei den STEC-Serovaren O84:H₂, O92:NM, O119:H₂₅ und O10:NM. ZHANG et al. (2002b) wiesen ζ -Intimin bei O118:H₅ nach. Die Übereinstimmung zum ζ -Intimingen des EPEC O111:H₉ (AF449417) und O111:H_{NM} (AF449416) sequenziert durch TARR und WHITTAM (2002) lag bei 98,6%. ζ -Intimin wies über die gesamte Sequenz betrachtet die größte Übereinstimmung zu α -Intimin auf.

5.3 PHYLOGENETISCHE ANALYSEN

Ausgangspunkt der phylogenetischen Analysen, die Erfassung möglicher Selektionsmechanismen und der genetischen Rekombination war die Generierung eines aussagekräftigen, qualitativ hochwertigen Sequenzvergleichs aller bislang beschriebenen Intiminvarianten. Die durch unzureichende Alignierung verursachte Variation der phylogenetischen Bäume kann dabei die Schwankungsbreite, die durch verschiedene Baum-darstellende (engl. „tree-building“)-Methoden entsteht, deutlich übersteigen (MORRISON und ELLIS, 1997).

5.3.1 CHARAKTERISIERUNG DES ALIGNMENTS

Für die eigenen Untersuchungen wurde der Algorithmus ClustalW (THOMSON et al., 1994) verwendet, der in viele kommerzielle sowie in die frei erhältlichen Programme MEGA3 (KUMAR et al., 2004a) und BioEdit V7 (HALL, 1999) übernommen wurde. Damit war die einfache Handhabung durch untereinander kompatible Dateiformate, sowohl während der Bearbeitung des Alignments als auch bei der Übernahme in die weiteren Programme, sichergestellt. McCLURE et al. (1994) und THOMPSON et al. (1999) evaluierten unterschiedliche Alignment-Algorithmen und konnten dabei die Überlegenheit globaler Verfahren belegen. Die letztgenannte Untersuchung erfolgte anhand des objektiven Leistungstest (engl. „Benchmarking“)-Systems BALiBASE und klassifizierte CLUSTALW zusammen mit PRRP (GOTOH, 1996) und SAGA (NOTREDAME und HIGGINS, 1996) als sehr gut geeignete Verfahren. PRRP und SAGA beinhalten bereits die „iterative Überprüfung“ der Ergebnisse des Sequenzvergleichs. Kurz zusammengefasst, handelt es sich dabei um eine schrittweise Überprüfung der innerhalb der Alignmentprozedur wegen ihrer größten Übereinstimmung zuerst gruppierten Sequenzen. Unterbleibt diese Korrektur, kann es insbesondere beim Vergleich stark übereinstimmender Sequenzen, wie den betrachteten *eae*-Genen, zu Berechnungsfehlern kommen (WALLACE et al., 2005). BRIFFEUIL et al. (1998) ermittelten ClustalW anhand von analysierten Proteinsequenzen in Gegenüberstellung mit sechs weiteren, über das Internet erhältlichen Methoden, als Methode mit der höchsten Selektivität und Sensitivität. RAGAVA et al. (2003) konnten zusätzlich die Eignung des ClustalW für die Analytik von Sequenzen mit hohem Übereinstimmungsgrad belegen. Selbst im Vergleich mit den modernen iterativen Verfahren, wie etwa ProbCons (DO et al., 2005), Muscle (EDGAR, 2004) oder T-Coffee (NOTREDAME et al., 2000), zählte CLUSTALW zu den vier leistungsstärksten Methoden (WALLACE et al., 2005). Die drei iterativen Alignmentverfahren erschienen zusätzlich weniger anwenderfreundlich, da die benötigten Dateiformate zunächst erstellt werden mussten und keine einfachen Benutzeroberflächen existierten.

Qualitätsmerkmal eines optimalen Sequenzvergleichs ist die Anzahl der Alignmentpositionen ohne Lücken (ROZAS et al., 2003). Für die 25 einbezogenen Referenzsequenzen mit offenen Leserahmen von 2805 (z. B. α -Intimin) bis 2847 (z. B. ϵ -Intimin) konnte ein Informationsgehalt bezogen auf die auswertbaren Sequenzvergleichsanteile von 97% realisiert werden (**Kap. 4.7.2**). Lediglich 3%, das entspricht 87 von 2858 Alignmentpositionen beinhalteten Lücken und wurden von der weiteren Analyse ausgeschlossen. Die Zuverlässigkeit des ClustalW wurde dadurch unterstrichen, dass ZHANG et al. (2002) einen geringfügig höheren Informationsgehalt von 97,4% erhielten. Dies war auf die geringere Anzahl von nur zehn, allerdings gleichzeitig stärker differierenden Sequenzen zurückzuführen. Als weitere Kenngröße des Informationsgehaltes sind die Anzahl polymorpher Alignmentpositionen „S“ mit 1050 zu werten, denen insgesamt 1607 Mutationen zugeordnet werden konnten. Als Maß für die größere Heterogenität ist der ca. siebenfach höhere Anteil variabler, informativer Positionen, die zu einer Veränderung der Aminosäuresequenz führen, zu werten. Hingegen überwiegen Positionen mit nur zwei Veränderungen im 5'-Bereich (**Tabelle 4.7.2-1**). Eine ausschließlich auf den hypervariablen 3'-Genbereich gestützte Phylogenetik kann zu ungenauen Ergebnissen führen, da der als hochkonserviert eingestufte 5'-Genbereich mit 496 PIPM im Vergleich zu 456 PIPM am 3'-Ende sogar über geringfügig mehr diskriminierende Information verfügte. Der zuvor ermittelte durchschnittliche Gesamt-G+C-Gehalt von 42,4% ließ sich durch die mit MEGA3 berechneten Nukleotidfrequenzen für Guanin mit 23% und Cytosin mit 19% innerhalb der kodierenden Sequenzregionen bestätigen (**Abbildung 4.7.2-1**). Veränderungen an der dritten Codonposition traten häufiger auf, da sie allgemein seltener zu einer Veränderung der Aminosäurenkodierung führen und daher weniger stark sanktioniert wurden (BARTON, 2005).

Eine Übersichtsdarstellung der polymorphen Alignmentpositionen beinhaltet **Abbildung 4.7.1-2**. Heterologe Abschnitte wurden durch schwarze Linien hervorgehoben. Für ζ -Intimin spiegelten sich die Unterschiede zur den Referenzsequenzen wieder.

Die von ZHANG et al. (2002b) beschriebenen Cluster polymorpher Positionen innerhalb der PP- und TM- sowie der D0-Domäne sind erkennbar.

5.3.2 STRUKTURELLE ASPEKTE

Die von LUO et al. (2000) ermittelte Einteilung in funktionelle Domänen konnte auch für ζ -Intimin auf Basis der Proteinsequenzen dargestellt werden. Interessanterweise verfügten alle untersuchten Sequenzen über die innerhalb der *eae*-Gene hochkonservierten Motive „WLQYYGQ“ und „WGAANKY“ (ADU-BOBIE et al., 1998a; BATCHELOR et al., 2000). Ihre Funktion ist nicht geklärt (ADU-BOBIE et al.; 1998a). LUO et al (2000) beschrieben sie als immunologisch unbedeutend. Ebenfalls darstellbar war die Aminosäure „W (Tryptophan)“ an Position 150 (**Abbildung 4.7.1-1**). DONNENBERG und KAPER (1991) selektierten diese Aminosäure für ihre Mutationsversuche. Ein Austausch gegen Alanin ermöglichte den Mutanten nach wie vor die ienge Zell-Zell-Verbindung (engl. „intimate adhearance“), allerdings konnten keine Restrukturierungen des Zytoskelettes eingeleitet werden. BATCHELOR et al. (2000) vermuteten, unterstützt durch die Strukturanalytik, keine direkte Beteiligung an der Bindungsregion, sondern einen stabilisierenden Effekt. ADU-BOBIE et al (1998a) beschrieben für EPEC mit den Aminosäuren W117, W136, W222 und W240 weitere konservierte Tryptophan-Positionen. Als essenziell für das Anheftungsvermögen des Intimins wurde die Disulfidbrücke zwischen Cystein 937 und Cystein 860 eingestuft (FRANKEL et al., 1994). LUO et al (2000) leiteten aus der Struktur keine direkte Beteiligung an der Anheftung ab, da die Disulfidbrücke zu weit von der Bindungsregion entfernt sei und nicht über den nach WEIS et al. (1992) für die Zuckerbindung essentiellen Calcium-Anteil verfüge. Die durch HARTLAND et al. (1999) angenommene gleiche Funktionsweise wie bei Invasin wird abgelehnt.

5.3.3 DARSTELLUNG DER BAUMDIAGRAMME

Für die Darstellung der phylogenetischen Relationen wurden bislang überwiegend Distanz-basierte Methoden eingesetzt. So verwendeten ZHANG et al. (2002b) die UPGMA (engl. „unweighted pair group method with arithmetic mean)-Methode, die dem Prinzip der konstanten Evolutionsraten folgend als unpräzise eingestuft wurde (**Kap. 2.7**). RAMACHANDRAN et al. (2003) verwendeten mit BIONJ (GASCUEL, 1997) eine Weiterentwicklung des Neighbor-Joining-Verfahrens, dem eine verbesserte Effizienz bei der Differenzierung stark ähnelnder Sequenzen zugesprochen wurde. TARR und WHITTAM setzten auf die modifizierte Darstellung nach NEI und GOJOBORI (NEI und GOJOBORI, 1986 und ZHANG et al., 1998). Unterschiede bestanden auch in der Verwendung von Nukleotid- oder den resultierenden Aminosäuresequenzen. Letztere wurden durch TARR und WHITTAM (2002) und RAMACHANDRAN et al. (2003) eingesetzt. Für die eigenen Arbeiten wurden 25 Nukleotidsequenzen verwendet, die - soweit möglich - vollständig sequenziert vorliegen mussten. Die hohe Zahl von Vergleichsequenzen minimierte mögliche Fehler in der Berechnung (BARTON, 2005). Auf die alleinige Verwendung des hypervariablen Abschnittes (JORES et al., 2003) wurde verzichtet, da die Untersuchung der informativen Alignmentpositionen auch dem 5'-Bereich einen hohen Informationsgehalt bescheinigte. In den eigenen Untersuchungen konnte gezeigt werden, dass sowohl die eingesetzten Distanz-basierten Verfahren (modifiziertes Modell nach NEI und GOJOBORI, Modell nach TAMURA und NEI) als auch die Maximum-Likelihood-Darstellung eine gute Wiedergabe der in der über das Alignment erstellten Distanzmatrix gewährleisteten (**Tabelle 9.1.1-2**). In Erweiterung der bislang existierenden Einteilung (ZHANG et al., 2002b) konnten die folgenden sechs Untergruppen zugeordnet werden: $\alpha/\nu/\omicron/\zeta$, $\beta/\delta/\kappa$ (mit der engsten Relation zu den Außengruppen *E. coli* O115 und dem *eae*-Genhomolog von *Citrobacter freundii*). $\sigma/\gamma/\tau$, ι/λ und $\epsilon/\pi/\eta/\xi$ bildeten die weiteren Untergruppen. ρ -Intimin

wurde in der Mehrzahl der Fälle γ_2 bzw. θ -Intimin zugeordnet. Allerdings ergab die verwendete Interior-Branch-Bewertung mitunter eine Eingruppierung zwischen α - und ι -Intimin. Die Neighbor-Net-Darstellung lieferte die Erklärung, da dort ebenfalls für das gesamte Gen sowie den konservierten Genabschnitt eine Zuordnung zu α - und ι -Intimin vorgenommen wurde. Für den hypervariablen Bereich ergab sich ebenfalls die Zuordnung zu γ_2/θ -Intimin. Die Netzwerkdarstellung zeigte den Einfluss möglicher rekombinanter Regionen. Zusätzlich konnte die problematische Zuordnung über die Distanzmatrix erklärt werden. ρ -Intimin wies zu allen Sequenzen mit etwa 82 bis 84% die gleiche p-Distanz auf, was eine eindeutige Zuordnung erschwerte. Methodisch wurde die Distanz-basierte Darstellung der Topologie als Bootstrap-consensus-Bäume nach FELSENSTEIN (1985) und EFRON et al. (1996) eingesetzt, da sie inkonstante Wiedergaben der phylogenetischen Relationen in Form von zwei oder mehr Baumdiagrammen vermeidet, die durch sich stark ähnelnde Sequenzen (vgl. ρ - oder λ -Intimin) bedingt sein können (engl. „tie trees“, BACKELJAU et al., 1996; TAKEZAKI, 1998). Zusätzlich wurde der Interior-Branch-Test angewendet. SITNIKOWA et al. (1995) bescheinigten diesem Test Vorzüge gegenüber dem Bootstrapping. Während RUSSO et al. (1996) keinen Unterschied im Leistungsvermögen zwischen dem Minimum-Evolution-Algorithmus ermittelten, konnten TAKAHASI und NEI (2000), GASCUEL et al. (2001) und WILLSON (2005) die Überlegenheit des Neighbor-Joining-Verfahrens gegenüber anderen Distanz-basierten Methoden, insbesondere dem Minimum-Evolution-Algorithmus, darstellen.

Die eigenen Untersuchungen zeigten deutlich, dass alle eingesetzten Methoden zu einer zuverlässigen Bewertung führten. Die seitens RAMACHANDRAN et al. (2003) zwar lediglich allgemein propagierte, bessere Auflösung durch das BIONJ-Verfahren ist somit nicht gegeben. Die hier verwendeten ergänzenden Darstellungsformen des Neighbor-Net-Verfahrens gewährleisteten eine deutlich zuverlässigere Bewertung der Relationen.

5.3.4 BEWERTUNG DER NOMENKLATUR

Seit der Erstbeschreibung des *eae*-Gens durch JERSE et al. (1990) wurden mehr als 100 *eae*-Gensequenzen verschiedener Serovare in den Gendatenbanken veröffentlicht. Dabei wurde nur ein kleinerer Teil in entsprechenden Journalpublikationen erwähnt. OSWALD et al. (2000) schlugen vor, neue Intiminvarianten erst ab einer Sequenzdistanz von $\geq 5\%$ zu benennen. Bei einer Übereinstimmung $>95\%$ sollte es sich um Intiminuntereinheiten handeln. In praxi bedeutet dies, dass zu einer endgültigen Zuordnung mehr als 90% der *eae*-Gensequenz vorliegen müssen, um exakte Aussagen treffen zu können. Daher wurde δ -Intimin zunächst als $\beta 2$ -Untereinheit eingeordnet (OSWALD et al., 2000) und später dem κ -Intimin zugewiesen (ZHANG et al., 2002b). BLANCO et al. (2005) lieferten die Komplettssequenz des O86:H34 (ADU-BOBIE et al., 1998a, ICC025) unter Beibehaltung der δ -Nomenklatur (AJ805027). Die Ähnlichkeit zu den bislang bekannten κ -Intimintypen lag bei über 99%, so dass die genannte Zuordnung gewährleistet blieb. JENKINS et al. (2003) verwiesen neben dem Stamm EPEC 4221 O49:HND (U66102, Sequenzeintrag: AN et al., 1997, vgl. **Tabelle 9.1.1-1**) zusätzlich auf EPEC 088:H5 als δ -Intiminvariante. Hierfür finden sich allerdings keine Sequenzeinträge, lediglich die Erläuterung bei PRESTON (2001). Die nachfolgende Übersicht fasst die existierenden Intimintypen und ihre Nomenklatur zusammen.

Bezeichnung	Erstbeschreibung*	Eigenständigkeit bestätigt
α (α_1, α_2)	ADU-BOBIE et al.;	+
β (β_1, β_2)	1998a; McGRAW et al., 1999	+
γ (γ_1, γ_2)		+
δ		+
ε ($\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$)	OSWALD et al., 2000	+
ζ	JORES et al., 2003	+
η (η_1, η_2)	ZHANG et al., 2002b	+
θ	TARR und WHITTAM, 2002	+
ι (ι_1, ι_2)		+
κ	ZHANG et al., 2002b	+
λ	AF439538**	+
μ	BLANCO et al., 2004	+
ν		+
ξ		+
π	AJ705052**	+
ρ	AJ748082**	+
\omicron	AJ876648 **	
σ	J781125**	+

* bezogen auf die Intiminvariante

** Nur Sequenzbereitstellung in der Gendatenbank

Zusammenfassend konnte für die Intiminvarianten α bis σ die seitens OSWALD et al. (2000) geforderte Distanzgrenze dargestellt werden. RAMACHANDRAN et al. (2003) veröffentlichten in ihrer Journalpublikation die Beschreibung der Intimine μ , ν und ξ . BLANCO et al. (2005) stellten nahezu zeitgleich eigene Sequenzen in der Gendatenbank bereit. Im direkten Vergleich entsprachen dabei μ/ι_2 , ν/ϵ_2 , ξ/β_2 . RAMACHANDRAN et al. (2003) verwendeten allerdings ebenfalls die Nomenklatur nach BLANCO et al. (2005) in der Sequenzdatenbank. Daher wurden diese Untereinheiten nicht separat aufgeführt. OSWALD et al. (2000) lieferten die Unterteilung von β - und γ -Intimin in zwei Untereinheiten. BLANCO et al. (2004) fanden Subtypen bei ν -Intimin und im Jahr 2005 ϵ_3 - und ϵ_4 -Intimin. Für die Angabe der orf-Größen und der Serovaren siehe **Tabelle 9.1.1-1**.

5.3.5 NETZWERKDARSTELLUNG

Die Netzwerkdarstellung phylogenetischer Relationen stellt eine sinnvolle Erweiterung im Vergleich zu den oben beschriebenen klassischen Baumdiagrammen dar. Eine vergleichende Bewertung der existierenden Verfahren existiert bislang nicht. POSADA und CRANDALL (2001b) hoben in ihrer vergleichenden Zusammenstellung neben der hohen Berechnungsgeschwindigkeit, die Möglichkeit zur statistischen Bewertung der Split-Zerlegung (BANDELT und DRESS, 1992) hervor. BRYANT und MOULTON (2004) sowie HUSON und BRYANT (2006) lieferten zusätzlich detaillierte Beschreibungen der in der neuesten Version Splits Tree 4.0 enthaltenen Berechnungsroutinen. Das eigene Vorgehen wurde zusätzlich unter alleiniger Verwendung der von ZHANG et al. (2002b) eingesetzten Sequenzen überprüft. Dabei konnten in Übereinstimmung mit den Autoren identische Ergebnisse erzielt werden (Daten hier nicht gezeigt). Allerdings fehlte in der Journalpublikation die Angabe der Fit-Werte, die eine Bewertung erleichtern hätte. Sie konnten mit Werten zwischen 60 und 80% ermittelt werden (Daten hier nicht gezeigt).

REID et al. (2000) bestätigten - anhand des Programms „SplitsTree 2.0“ - die bestehenden klonalen Gruppen der EPEC und EHEC (WHITTAM et al., 1993; WHITTAM und McGRAW, 1996; WHITTAM, 1998) anhand ihrer über MLTS gewonnenen Daten. BEUTIN et al. (2005) charakterisierten die MLTS-Profile von O103-Pathotypen. Die Split-Zerlegung ist daher als eine für die Darstellung evolutiver Verhältnisse bei *E. coli* etablierte Methodik einzustufen.

Sowohl für den als hoch konserviert einzustufenden 5'-Bereich der *eae*-Gene als auch für den hypervariablen 3'-Abschnitt wurden netzwerk- oder buschartige Relationen ermittelt. Die Betrachtung der Gesamtsequenzen lieferte ein identisches Bild. Alle Sequenzen verfügten damit über rekombinante Bereiche, die eine korrekte phylogenetische Aufschlüsselung in Form eines einfachen, exakt verzweigenden Baumes erschwerten. Zusätzlich bestätigt die Neighbor-Net-Darstellung nach BRYANT und MOULTON (2004) die bereits zuvor über die klassischen Distanz- bzw. Maximum Likelihood-Darstellungen erhaltenen Relationen (**Abbildung 4.6.1-1** bis **Abbildung 4.6.1-5**). Gleichzeitig lieferte sie sowohl in der abschnittsweisen als auch in der Betrachtung der Gesamtsequenzen Hinweise für die Schwierigkeit der Eingruppierung des λ -Intimins. Das in den eigenen Untersuchungen bestimmte ζ -Intimin wird korrekt dem α -Intimin als der ähnlichsten Vergleichssequenz zugeordnet. Die Zuverlässigkeit der Kalkulation anhand einer nahezu vollständigen Einbeziehung aller erhaltenen Werte und damit die qualitativ hohe Aussagekraft der Ergebnisse belegen die mit >95% (im Falle der Darstellung des 3'-Endes von >98%) guten Fit-Werte. Sie sind als Maß für die Wiedergabegenauigkeit, d. h. die in der Darstellung erfassbaren „splits“ und damit die Eignung der Methode zur korrekten Analyse der Datenmatrix im Hinblick auf nichtlineare Evolutionsereignisse, wie z. B. Rekombination, anzusehen.

ZHANG et al. (2002b) erhielten in ihren Untersuchungen für die Intiminvarianten α bis κ ebenfalls netzwerkartige Verzweigungen für alle drei Genabschnitte. Mit insgesamt 25 einbezogenen Referenzsequenzen berücksichtigen die eigenen Arbeiten auch die evolutiven Vorläufersequenzen bzw. mit *Citrobacter freundii* auch ein *eae*-Genhomolog und die zusätzlich bekannten Intiminvarianten μ bis

σ . Weiterhin werden die relevanten Untereinheiten einbezogen. Es konnte gezeigt werden, dass die einfache Split-Zerlegung nach BANDELT und DRESS (1992) sowie HUSON (1998) für die korrekte Wiedergabe rekombinanter Einflüsse bei mehr als zehn Referenzsequenzen nicht geeignet ist. Dies wird an den niedrigen Fit-Werten von <50% (**Abbildung 4.6.2-4**) bzw. <64% (**Abbildung 4.6.2-5**) deutlich. Im Fall der **Abbildung 4.6.2-4** ist sogar keine klare Darstellung der Relationen mehr möglich. Die seitens BRYANT und MOULTON (2004) erarbeitete höhere Auflösung des Neighbor-Net-Verfahrens war der einfachen Split Zerlegung bei den sich insgesamt sehr stark ähnelnden Intimingensequenzen überlegen (Prof. D. HUSON, Lehrstuhl für Algorithmen in der Bioinformatik, Universität Tübingen, persönliche Mitteilung). Die Bootstrap-Evaluierung (FELSENSTEIN, 1985) unter Angabe der Werte von 100% zeigt, dass mögliche Fehler innerhalb des Datensets, bedingt durch eine zu geringe Sequenzanzahl bzw. stark variierende Sequenzlängen, vermieden wurden (HUSON und BRYANT, 2006). Für die abschließende Bewertung der Darstellung war die alleinige Verwendung der Bootstrap-Werte allerdings nicht ausreichend, da sie lediglich ein Maß für die Wiederholpräzision der Analyse darstellten. Die Darstellungen nach BANDELT und DRESS (1992) wurden ebenfalls durch hohe Bootstrap-Werte unterstützt, ihre Fit-Werte lieferten allerdings nur eine mangelhafte Unterstützung des Modells und bewerteten es als für diese Fragestellung „systematisch falsch“ (HUSON und BRYANT, 2006).

5.3.6 SELEKTION UND GENETISCHE REKOMBINATION

Bakterielle Oberflächenproteine zeigen einen hohen Grad an Sequenzpolymorphismen (LI et al., 1995; WHITTAM, 1995). Als mögliche Ursache wird das Vorliegen der positiven Selektion ($p_N - p_S > 0$) angesehen. Der hier bevorzugte Selektionsmechanismus wurde anhand der Parameter p_S und p_N dargestellt. Die Bezeichnungen K_s bzw. K_a werden synonym in Maximum-Likelihood-Verfahren verwendet (**Kap. 2.4.3; Kap. 4.7.3**; YANG et al., 1997a; 1997b). Unter Berücksichtigung der durch LUO et al. (2000) ermittelten

funktionellen Einteilung konnte anhand der Einbeziehung aller 25 Referenzsequenzen belegt werden, dass sowohl für die Gesamtsequenz als auch die einzelnen Regionen der Intimogene $p_S > p_N$ signifikant überwog ($p > 0,05$). Der Quotient p_S/p_N sank dabei von neun in die Zellmembran durchspannenden Bereich, auf 1,5 in der extrazellulären Domäne D3 (483-1). Eine mögliche Fehlberechnung des Verhältnisses p_S/p_N wurde durch die Verwendung der modifizierten Berechnung nach NEI und GOJOBORI (1986) und ZHANG et al. (1998) durch die Angabe des Parameters R verhindert (**Kap. 2.4.2**). Wird das Verhältnis nicht berücksichtigt, kann es zu einer Unterbewertung der synonymen Positionen und damit einer fehlerhaften Selektionsbewertung kommen. HUGHES und FRIEDMAN (2000) konnten den Einfluss des Parameters R bei der Analyse von Hantavirussequenzen und im Jahr 2005 durch Vergleich von mehr als 4000 proteinkodierenden Genen von *Saccharomyces cerevisiae* und *Saccharomyces paradoxus* darstellen. Gemäß YANG und NIELSEN (2000) wurde zusätzlich das Verhältnis nur über die vierfach veränderten informativen Positionen PIPM (4) kalkuliert, da diese nur Transitionen enthalten. Zweifach veränderte informative Substitutionen PIPM (2), beinhalten Transitionen und Transversionen (LI, 1993) und waren daher für die Berechnung ungeeignet. Die abschnittsweise Betrachtung des Selektionsmechanismus bei ζ -Intimin im Vergleich mit den stark ähnlichen α -, σ - und ν -Intimingenen ergab ebenfalls eine Dominanz der negativen, bereinigenden Selektion. Die zuvor als Gesamtdurchschnitt bewerteten Ergebnisse ließen sich auch im Detail auch auf das einzelne Gen bei übereinstimmenden p_S/p_N -Relationen übertragen. So fanden sich in der periplasmatischen (PP)-, der transmembranösen (TM)- und der extrazellulären Domäne „D3“ positiv selektierte Bereiche. TARR und WHITTAM (2002) erhielten bei ihrer Gegenüberstellung von θ -, γ -, α - und β -Intimin ebenfalls diese Ergebnisse. Obwohl in Übereinstimmung mit den genannten Autoren und ZHANG et al (2002b) Hinweise für eine positive, breit gefächerte (engl. „diversifying“) Selektion darstellbar waren, war der Einfluss einer Absättigung der Sequenz durch vielfache Substitutionen, die weitere Veränderungen maskieren können, zu berücksichtigen (HUGHES et al., 1990). Im paarweisen

Vergleich der *eae*-Gene wurden sehr stark übereinstimmende bzw. stark divergierende Sequenzen gegenübergestellt. Erstgenanntes führte zu einer Erschwerung der statistischen Absicherung. Um die negative Selektion weiter zu charakterisieren, wurde die Auswirkung von p_N -Mutationen auf die Aminosäuresequenzen betrachtet. HUGHES et al. (1990) entwickelten ein Verfahren zur Klassifizierung der p_N als sog. „konservative p_{NC} “, also die Ladung bzw. die Polarität nicht verändernde Aberrationen oder „radikale p_{NR} “ Veränderungen. ZHANG (2000) evaluierte an 47 Zellkern-Gensequenzen das Programm HON-NEW, das die zuvor dargestellten Anforderungen an R berücksichtigte. HUGHES und FRIEDMAN (2000) optimierten ihr bereits durch TARR und WHITTAM (2002) verwendetes Programm SCR3. Es konnte gezeigt werden, dass im hypervariablen Bereich eindeutig radikale Mutationen, die zu einer Veränderung der Ladung anstatt der Polarität führten, bevorzugt wurden (**Abbildung 4.7.3-5**). Identische Beobachtungen konnten für die Histokompatibilitätsgene *HLA-A* und *HLA-B* (HUGHES, 1995), Defensine (HUGHES und YEAGER, 1997) und fünf Oberflächenproteine von *Plasmodium falciparum* (HUGHES und VERRA, 1998) dargestellt werden. Der Parameter „Ladung“ erwies sich in allen Untersuchungen als Hauptfaktor für die erfolgreiche Adhäsion des Proteins an Zielmoleküle. Für die *eae*-Gene und die korrespondierenden Tir (CLARKE et al., 2003) wurde unterschiedliches Bindungsverhalten beschrieben (ADU-BOBIE et al., 1998a). TARR und WHITTAM konnten mit der ursprünglichen SCR3-Version nach HUGHES et al. (1990) identische Relationen darstellen. Im konservierten Abschnitt zeigte sich ebenfalls in Übereinstimmung mit den genannten Autoren keine Bevorzugung von Ladungs- oder Polaritätsveränderungen. Derartige Varianzen würden zu einer Beeinflussung der Membranpassage führen.

Die Weiterentwicklungen in HON-NEW und SCR3 wurden als methodisch gut geeignet für die Untersuchung stark ähnlicher Gene eingestuft (SMITH, 2003). Der Autor bewertete die Arbeiten von ZHANG et al. (2002b) zur Darstellung der positiven Selektion bei pankreatischen Ribonukleasegenen als Beleg für die Notwendigkeit der weiterentwickelten Verfahren. Gleichzeitig wird kritisch auf

die Anfälligkeit der Methoden für Variationen der statistischen Parameter hingewiesen.

Als dritter Ansatz spiegeln die **Abbildungen 4.7.3-1** und **4.7.3-2** das Verhältnis der Transitionen zu den Transversionen wider. Transitionen treten in den meisten Genen häufiger auf als Transversionen (FITCH, 1967, GOJOBORI et al., 1982, KOCHER und WILSON, 1991). Für die 5'-Enden der *eae*-Gene bestätigte sich diese Einschätzung. Die Hypervariabilität der 3'-Enden wurde im Gegensatz dazu überwiegend durch Transversionen hervorgerufen, die ab einer Distanz von 0,4 etwa um den Faktor zwei überwogen. Hieraus ließ sich neben der höheren Variabilität auch eine gerichtete Selektion ableiten, die im konservierten Abschnitt die Aufrechterhaltung des Status quo zeigte und im hypervariablen Bereich Veränderungen der bestehenden Konformation erlaubte.

Zielsetzung der Untersuchungen zur genetischen Rekombination war die Bewertung des ζ -Intimings in Relation zu den bislang bekannten Intimvarianten und -subtypen bei gleichzeitiger Berücksichtigung möglicher Vorläufersequenzen. Ferner sollte die mögliche Mosaikstruktur des ζ -Intimins im direkten Vergleich mit den paarweise als Vorläufersequenzen definierten α -, β -, γ - und σ -Intiminen ermittelt werden (**Tabelle 4.7.4-1**). Für die Detektion der genetischen Rekombination innerhalb des *eae*-Gens bei *E. coli* wurde bislang immer nur ein Verfahren eingesetzt. So verwendeten McGRRAW et al. (1999) sowie TARR und WHITTAM (2002) mit dem Maximum Chi-Quadrat-Verfahren (SMITH 1992, 1999), entgegen den Empfehlungen zur Analytik genetischer Rekombination (MARTIN und RYBICKI, 2000) nur eine Methodik. Für die eigenen Untersuchungen wurde daher mit der Verwendung von Bootscan, Chimarea, Genconv, Maxchi, RDP und Siscan auf insgesamt sechs Modelle zurückgegriffen (**Kap. 2.6.3**). Neben den Evaluierungsstudien von DROUIN et al. (1999), WIUF et al. (2001) und BROWN et al. (2001) wurde dabei insbesondere auf die Arbeiten von POSADA und CRANDALL (2001a) gesetzt. Letztere bewerteten vergleichend 14 unterschiedliche Berechnungsmodelle anhand vorgegebener Computersimulationen oder Alignments bekannter

Bezugsgene mit unterschiedlichen rekombinanten Anteilen (POSADA 2002). Basierend auf den Parametern statistische Aussagekraft und der Anzahl falsch-positiver Reaktionen wurden die Algorithmen Chimaera (POSADA und CRANDALL 2001b), MaxChi (SMITH, 1992) und Geneconv (PADIDAM et al., 1999) als sensitivste und aussagekräftigste Verfahren ermittelt. Eine Sequenzdivergenz von 5 bis 20% wurde für aussagekräftige Ergebnisse als notwendig angesehen. Das kostenlose Programm RDP (Recombination Detection Program) Version 2.0 (MARTIN und RYBICKI, 2000 und MARTIN et al., 2005) ermöglichte die vergleichende Anwendung der drei zuvor als besonders geeignet evaluierten Systemen. Das Leistungsspektrum der eingesetzten Algorithmen unterscheidet sich nach der Möglichkeit, ausschließlich mögliche rekombinante Genabschnitte zu identifizieren bzw. die Genauigkeit der Analyse durch die Festlegung von Anfangs- und Endpunkten (engl. „breakpoints“) und gesamten Bereichen bzw. mögliche evolutive Vorläufersequenzen (engl. „parent-daughter-identification“) zu erhöhen.

Die globale Überprüfung des Alignments ergab sechs als rekombinant im Sinne der Voraussetzung $p > 0,05$ eingestufte Regionen innerhalb des ζ -Intimings (**Tabelle 4.7.4-1**). Die Regionen zwischen den Nukleotidpositionen „1819 bis 2292“ und „1737 bis 2292“ konnten jeweils durch fünf Verfahren, der Abschnitt „1014 bis 1590“ durch vier Modelle identifiziert werden. Die k_{\max} -Lokalisation unter Angabe der paarweisen Sequenzdifferenz im Bezug zu ζ -Intimin (**Abbildung 4.7.4-1**) ergaben im paarweisen Sequenzvergleich für $\alpha 1$ - Intimin $n=6$, für $\beta 1$ - und $\gamma 1$ -Intimin $n=8$ und σ -Intimin $n=5$ wahrscheinliche Bruchpunkte (k_{\max}) mit $p > 0,01$. Die Sequenzdistanzen stiegen dabei entsprechend der k_{\max} -Lokalisation im 3'-Bereich an. Die Positionsabschnitte „128 bis 360 bis 460“ bei $\beta 1$ -Intimin und „1-178“ bei σ -Intimin bedurften besonderer Beachtung. Hier ergaben sich Distanzen von 12,5%, 14,9% und 13,9%, eingebettet in sich anschließende Regionen mit deutlich geringeren Sequenzunterschieden (2,4%, 4,8% und 7,6%).

Zusammenfassend konnten die in der globalen Analyse für die in **Tabelle 4.7.4-1** dargestellten Bruchpunkte bestätigt werden. Zusätzliche rekombinante Regionen wurden im konservierten 5'-Abschnitt aufgezeigt. Die berechneten Vorläufersequenzen sind zu vernachlässigen, da keine einheitliche und - bis auf wenige Ausnahmen - nur wenige Mehrfachidentifizierungen erreicht wurden. McGRAW et al. (1999) lieferten erstmalig die Darstellung der α -, β - und γ -Intimogene als Mosaikprofile. TARR und WHITTAM (2002) fanden ebenfalls rekombinante Regionen im 5'- und 3'-Abschnitt beim Vergleich von θ - mit α -, γ - und β -Intimin. Für eine abschließende Bewertung siehe **Kap. 5.5**.

5.4 BEWERTUNG DER VERWENDETEN COMPUTERPROGRAMME

BioEdit V7 (HALL, 1999) erwies sich als ein gut zusammengestelltes Programmpaket. Der Im- und Export der benötigten Dateiformate funktionierte ebenso reibungslos, wie die Bedienung der einfach gehaltenen Benutzeroberfläche unter Windows XP. Der wesentliche Vorteil der Programmversion Mega3 lag in der einfachen Handhabung und der kurzen Berechnungszeiten. Eine zusätzliche Erweiterung um Maximum-Likelihood-Methoden wäre von Vorteil. Beide Programmpakete erwiesen sich als sinnvolle Alternativen zu den kommerziell verfügbaren Anwendungen (z. B. DNASTar, Fa. Lasergene). Letztere weisen bislang eine schlechte Kompatibilität zu anderen Programmen auf, was die Überprüfung der Ergebnisse schwierig und sehr zeitaufwendig macht. DnaSP verfolgte mit der Analytik der Alignments einen sinnvollen Ansatz. Allerdings erwies sich das Programm als störungsanfällig und war dadurch insbesondere bei den enthaltenen grafischen Darstellungsmöglichkeiten unbrauchbar. Die Netzwerkdarstellung über Splits Tree V.4.0 wurde hinsichtlich der möglichen Algorithmen (z. B. Neighbor-Net) umfangreich erweitert. Zusätzlich wurde die Ergebnisausgabe und damit die Einarbeitung der Daten in wissenschaftliche Beiträge verbessert. Die in den eigenen Arbeiten eingesetzte Java-Plattform erwies sich als wesentlich robuster im Vergleich zur

Vorversion. Insbesondere das Einlesen von Dateiformaten wurde verbessert. Im Gegensatz dazu handelt es sich bei PAML V.3.14, HON-NEW und SCR3 sowie Happlot um einfache, auf die mittlerweile überholte MS-DOS-Ebene zugreifende Programme. Sie erfordern alle eine intensive Einarbeitung in die verwendeten Dateiformate. Ein direkter Einsatz des eigenen Datensets ohne die Etablierung der mitgelieferten Testprogramme erhöht den Fehlereintrag durch falsche Datenkonvertierungen. Das Programmpaket RDP V.2.0 lieferte eine einfache Zusammenstellung verschiedener Algorithmen und enthielt insbesondere eine anwenderfreundliche Version des häufig verwendeten MaxChi. Alle überprüften Verfahren gehen deutlich über das Potenzial einer kommerziellen Anwendung hinaus. Der direkte Kontakt zum Autor (SCR3, HON-NEW, Splits Tree) bzw. der Zugriff auf umfangreiche Bedienungsanleitungen (RDP, Mega, BioEdit, PAML) erleichterten die Einarbeitungsphase. Aus eigener Erfahrung ist auf die benötigte Computerleistung hinzuweisen, die je nach Art der verwendeten Programme stark variierte und zu unterschiedlichen Ergebnissen führte.

5.5 BEDEUTUNG DER VARIATIONEN INNERHALB DER INTIMINGENE

Das *eae*-Gen und das resultierende Protein Intimin gehören zu den bestbeschriebenen Virulenzfaktoren der EPEC und VTEC. Die Literaturabfrage der Datenbank „Pubmed“ ergab mehr als 400 Treffer unter dem Suchbegriff „Intimin“. Die Suchkriterien „intimin and typing“ sowie „intimin and characterization“ wurde 31 bzw. 38 mal bedient. Für eine Bewertung der Verteilung der Intiminvarianten innerhalb der Serovare und -typen wurde insbesondere auf Arbeiten zurückgegriffen, die neben der Intimintypisierung weitere Virulenzfaktoren ermittelten, wie etwa „Saa“ oder die Verankerung des LEE.

Das *eae*-Gen erlangte besondere Bedeutung, weil es signifikant mit komplizierten und schwerwiegenden Erkrankungsverläufen der HC und des HUS korreliert (BEUTIN et al., 1998, BOERLIN et al., 1999). BEUTIN et al. (2004) wiesen das Virulenzgen bei 95,2% der HUS-Isolate, aber nur bei 66,7% der von Patienten mit unblutiger Diarrhoe gewonnenen Isolate nach. Zuvor hatten bereits AGIN und WOLF im Jahr 1997 auf das mögliche Auftreten einheitlicher Intiminvarianten beim Menschen sowie bei Hasen und Schweinen hingewiesen.

PHILLIPS und FRANKEL (2000), PHILLIPS et al. (2000) sowie FITZHENRY et al. (2002a) belegten für das γ -Intimin bei *E. coli* O157 eine ausschließliche Bindung an das follikelassoziierte Epithel (FAE) der Peyerschen Platten. Hingegen verursachte das α -Intimin bei EPEC O127:H6 die Besiedelung des FAE und des Dünndarms. Der Effekt war innerhalb einzelner Mutanten übertragbar. So konnte durch Translokation des γ -Intimins in EPEC, das VTEC-Verteilungsmuster hervorgerufen werden und umgekehrt (FITZHENRY et al., 2002b). Die Autoren konnten anhand von *E. coli* O55 ebenfalls im Jahr 2002 zeigen, dass bereits eine kleine Mutation der stark differierenden α - und γ -Intiminvarianten zur Veränderung des Gewebetropismus führt. Damit ist der Einfluss weiterer Faktoren auf die Anheftungslokalisation zu vermuten. Für die ϵ -Intimin tragenden Pathovaren O103:H⁻ und O103.H2 konnten sie in Übereinstimmung mit den γ -Intimin tragenden-Stämmen ebenfalls eine überwiegend follikelassoziierte Adhäsion nachweisen (FITZHENRY et al., 2003). MARCHES et al. (2000) hatten bereits zuvor bei Versuchen mit O103-Pathotypen die Notwendigkeit des Intimin-Tir-Komplexes für die Auslösung der Erkrankung bei Kaninchen untersucht. HARTLAND et al. (2000) gehen gleichzeitig von einer möglichen Wirtsspezifität, verursacht durch den Intimintyp, aus. Verschiedene Autoren verweisen auf eine stammbezogene Zuordnung der Intiminvariante (ADU-BOBIE et al., 1998a; OSWALD et al. 2000; REID et al., 2000; TARR und WHITTAM, 2002). Dies konnte zusätzlich in den eigenen Arbeiten für ausschließlich als β -Intimin-Varianten ermittelte O26-Stämme bestätigt werden. Für den VTEC Prototypen O157 war diese

Zuordnung nicht möglich. Hier schien γ -Intimin mit dem H-Antigen (H7, H-) korreliert zu sein. OSWALD et al. (2000) zeigten allerdings bereits das Vorkommen von α -Intimin bei O157:H⁻ und O157:H45 Humanisolaten. RAMACHANDRAN et al. (2003) ermittelten α -Intimin überwiegend bei H6- und β -Intimin vor allem bei H2-, H11- und H-Antigenen. In den eigenen Untersuchungen gehörten die β -Intimin tragenden Isolate mehrheitlich diesen H-Antigen Gruppen an. Für H- bzw. H2 wurden allerdings ebenfalls ζ - und γ -Intimin nachgewiesen.

Hinsichtlich der Nachweisraten konnte β -Intimin, wie auch aus den eigenen Ergebnissen erkennbar, am häufigsten isoliert werden. Es erschien nicht spezifisch für EPEC oder EHEC und wurde bei Isolaten von Rindern, Schweinen, Menschen und Hasen detektiert. α -Intimin wurde zunächst ausschließlich bei humanen EPEC (O127:H6, O142:H6, O55:H6) isoliert (OSWALD et al., 2000). Die Bedeutung des Rindes als Reservoir humanpathogener HC und HUS Stämme ist gut belegt (PRADEL et al., 2000; WELLS et al., 1991; ZSCHÖCK et al., 2000; BEUTIN et al., 1993 und MONTENEGRO et al., 1990). Die Prävalenz des *eae*-Gens betrug nach NAYLOR et al (2005) 16,5% bei Rindern und 44% bei Schafen. Dabei handelte es sich überwiegend um β - und γ -Isolate. Insgesamt wurde die Prävalenz der ζ -Intimin tragenden O-Antigentypen O84, O92 und O150 als gering eingestuft (JORES et al., 2003). JENKINS et al. (2003) konnten bei einer 34 Ausbruchsstämme aus dem Vereinigten Königreich umfassenden Typisierung keine O26- bzw. O111-Pathotypen ermitteln. Beide wurden häufig bei gesunden Tieren nachgewiesen (WIELER et al., 1996; KOBAYASHI et al., 2001; BLANCO et al., 2004a, 2004b). Allerdings wurde der später durch TARR und WHITTAM charakterisierte Ausbruchsstamm (Einwohner der engl. Stadt Sheffield erkrankten bei einem Auslandsaufenthalt in Frankreich) EPEC O111:H9 als ζ -Intimin-Isolat bestimmt. KRAUSE et al. (2005) ermittelten anhand von 2165 *E. coli*-Stämmen *eae*-Gen-Prävalenzen von 19,2% bei Schafen, 17,6% bei Schweinen und 10,4% bei Rindern sowie 7,2% bei Hunden, 6,5% bei Katzen und nur 2,3% bei Puten. Sie wiesen mit 92% überwiegend atypische EPEC

nach, die sie überwiegend als β 1- und θ -Intiminvarianten klassifizierten. β - und γ -Intimin gehören zu den Hauptvertretern der humanpathogenen EHEC. Unter den nicht-O157-Serovaren dominieren ε - β - und γ -Intimin (KARCH et al., 2005). BEUTIN et al. (2004) ermittelten bei der VTEC-Charakterisierung humanpathogener Isolate ein signifikant höheres Vorkommen *eae*-positiver Isolate bei Kindern unter sechs Jahren. Auch hier dominierten *E. coli* vom β 1 bis γ 1 und ε -Intimintyp. Die ermittelten Serovaren (O26, O145, O157, O118 und O129) wiesen eine enge Verbindung zum Vektor Rind auf. Gleichzeitig konnten sie eine deutliche Assoziation der hochvirulenten Stämme bei der Kombination aus *vt2*- und *eae*-Gen aufzeigen. Die eigenen Ergebnisse zeigen ebenfalls hohe Nachweisraten von β - und γ -Intimin bei Stuhlproben von HUS/HC-Patienten, Lebensmitteln und Rinderkot. Der hohe Anteil an γ -Intimin war allerdings auch auf den hohen Anteil von O157:H7- und O157:H⁻-Stämmen zurückzuführen. Das ζ -Intimigen wurde innerhalb der Lebensmittel und aus Rinderkot häufig bei O156- und O84-Serovaren (**Tabelle 4.4.2-1**, **Tabelle 4.4.2-3**) nachgewiesen, dabei handelt es sich allerdings ausnahmslos um *vtx*-negative Stämme. Allerdings konnte das ε -Intimin mehrfach bei O103-Pathotypen detektiert werden. Hinsichtlich des allgemeinen Vorkommens des *eae*-Gens ermittelte BÜLTE (2001) Genfrequenzen von 92% bei HUS/HC-Isolaten, 81,3% bei symptomlosen Ausscheidern und lediglich 7% bei Stämmen bovinen Ursprungs bzw. 10,9% lebensmittelassoziierter Isolate. Zusammengefasst belegten die eigenen Intimintypverteilungen das breite Vorkommen von β - und γ - Intimin, wenngleich ζ - und auch in einem Fall κ -Intimin nachgewiesen wurde. Dabei handelte es sich mit O107:H11 um Stämme ohne Verotoxinbildungsvermögen. Als θ -Intiminvariante wurde unter anderem der Ausbruchstamm C6 KJS O111:H bestätigt. Eine gute Übereinstimmung ergab sich für ε -Intimin, dass wie erwartet bei O103-, O8- und O11- und O-Serovaren nachgewiesen wurde.

Phylogenetisches Potential wird dem Intimingen als Marker für den LEE zugewiesen (NATARO und KAPER, 1998; TARR und WHITTAM, 2002). EHEC und EPEC lassen sich in klonalen Gruppen zusammenfassen (WHITTAM et al., 1993, DONNENBERG und WHITTAM, 2001). Die klonalen Gruppen EPEC 1 (Verankerung des LEE in *se/C*) und EHEC 1 (Verankerung des LEE in *se/C*) unterscheiden sich stark, EPEC 2 und EHEC 2 (beide Verankerung in *pheU*) sind dagegen enger verwandt. Innerhalb der EHEC 1 gilt O55:H7 als anerkannter Vorläufer des O157:H7. Beide verfügen somit über γ -Intimin. JORES et al. (2001) und RUMER et al. (2003) fanden bei eng verwandten Stämmen einheitliche LEE-Verankerungen und stark ähnliche Intimintypen. RUMER et al. (2003) wiesen nach, dass β -Intimin gruppenübergreifend verankert in *pheU* und *pheV* auftritt. Hieraus folgerten JORES et al. (2004), dass der LEE in der Vergangenheit erworben wurde. Ab hier sind mehrere Vorgehensweisen denkbar: Zum einen die evolutive Weiterentwicklung hin zu den heute ermittelten Intimingenen und zum anderen können rekombinante Einflüsse zur Veränderungen der Intimingene bei identischen Isolaten führen. Letztlich konnte nicht genau geklärt werden, ob ein Stamm unabhängig von den übrigen Serovaren seiner Gruppe eine LEE-Ausprägung erworben hat, was aufgrund der Bewertung der Evolution unter dem sog. Parsimonie-Kriterium (geringster Aufwand) das Vorkommen von ζ -Intimin bei O111:H9 und der *se/C*-Verankerung des LEE erklären könnte (TARR und WHITTAM, 2002). Gleichzeitig sind sowohl der LEE als auch die *eae*-Gene Mosaikbausteine (SANDNER et al., 2001 und MCGRAW et al., 1999). Vor diesem Hintergrund bewerteten REID et al. (2000) die Veränderungen des LEE und des Intimins als *in situ* mutations- und rekombinationsbedingt.

Sowohl in der Veterinär- als auch der Humanmedizin werden die variablen Intimingene als gute Immunogene angesehen. Antikörper lassen sich bereits frühzeitig nachweisen (FRANKEL et al., 1995; ADU-BOBIE et al., 1998a). LI et al. (2000) wiesen Antikörper gegen Intimin bei HUS-Patienten acht Tage post infectionem nach. GANSHERHOFF et al. (1999) begannen mit der Gewinnung von Antiseren gegen O157- und O26-Serovaren. DEAN-NYSTROM et al. (2002) erzielten Erfolge bei der Immunisierung von Muttersauen mit O157-

Antiseren. JUDGE et al. (2004) konnten über die Vakkzinierung die Ausscheidungsrate von *E. coli* O157 bei Mäusen reduzieren. Aufgrund der nahezu ubiquitären Verbreitung und der Vielfalt der Serovaren, insbesondere in Rinderbeständen, erscheint ein derartiger Ansatz allerdings wenig Erfolg versprechend und sehr kostenintensiv.

6 SCHLUSSFOLGERUNGEN

Das vorgestellte RFLP-Verfahren ermöglicht die einfache und schnelle Typisierung von mindestens 19 Intiminvarianten, einschließlich ihrer beschriebenen Untereinheiten. Es kann somit auch als akkreditierte Methodik für die Routinediagnostik und epidemiologische Untersuchungen Verwendung finden.

Zur Abgrenzung neuer Intiminvarianten unter Beibehaltung der griechischen Nomenklatur, erscheint der Cut off-Wert von 95% nach OSWALD et al. (2000) sinnvoll. Die Ergebnisse der Intimintypisierung untermauern die hohen Nachweisraten für β - und γ -Intimin. Die sequenzierten ζ -*eae*-Gene sind apathogenen Stämmen zuzuordnen.

Für die Darstellung der phylogenetischen Relationen sollten nach Möglichkeit vollständig sequenzierte Gene eingesetzt werden, da ansonsten wertvolle Informationen zur diskriminierenden Bewertung verloren gehen können. Methodisch konnten keine Unterschiede in der Effizienz der Distanz-basierten und der Maximum-Likelihood-Darstellung ermittelt werden. Als Grundvoraussetzung für die korrekte Wiedergabe der Phylogenie ist das qualitativ hochwertige Alignment anzusehen. Die hochauflösende Netzwerkdarstellung nach BRYANT und MOULTON (2004) stellt eine sinnvolle Ergänzung zur Klassifizierung problematischer Sequenzabschnitte dar.

Die bislang untersuchten 19 Intiminvarianten unterliegen alle einem negativen, d. h. bereinigenden Selektionsdruck. In der hypervariablen Bindungsregion werden überwiegend Mutationen bevorzugt, die zu einer Ladungsveränderung anstatt einer Polaritätsbeeinflussung führen. Im konservierten Abschnitt werden weder Ladung noch Polarität begünstigt. Auch für ζ -Intimin konnte das Vorliegen genetischer Rekombination mit sechs unterschiedlichen Verfahren gezeigt werden. Für die abschließende Bewertung der Evolution des *eae*-Gens respektive des LEE sind allerdings weitere Untersuchungen zu den Komponenten des LEE notwendig.

7 ZUSAMMENFASSUNG

Attaching und Effacing Läsionen sind das pathognomische Kennzeichen (engl. „hallmark“) der EPEC und VTEC/STEC. Der Intimintypisierung auf Basis des hypervariablen 3'-Genabschnitts mittels PCR kommt dabei eine wesentliche diagnostische und epidemiologische Bedeutung zu. So konnten bislang einschließlich der Untereinheiten mehr als 20 verschiedene Intimintypen differenziert werden (α bis σ). Ziel der vorliegenden Arbeit war es, anhand der Typisierung des *eae*-Gens von 317 *E. coli*-Stämmen unterschiedlicher Herkunft einen Beitrag zur - möglicherweise habitatassoziierten - Verbreitung der Intiminvarianten zu liefern. Die Typisierung von α -, β -, γ -, δ - und ε -Intimin erfolgte anhand des PCR-Verfahrens nach ADU-BOBIE et al. (1998), das hierfür um ein Primer-Paar zur Erfassung von ε -Intimin erweitert wurde. Die Produktüberprüfung wurde mittels restriktionsenzymatischem Verdau vorgenommen. Anschließend wurde, basierend auf den Oligonukleotiden von RAMACHANDRAN et al. (2003) ein PCR-RFLP-Verfahren zur Erfassung der übrigen Intiminvarianten entwickelt. Damit konnte im Vergleich zur Originalliteratur die zusätzliche Differenzierung von sieben Intiminvarianten respektive -untereinheiten erreicht werden. Mit 203 (64,03%) der Stämme entfiel der Hauptanteil auf γ -Intimin. 34 (11,04%) der Stämme wurden als β -Intiminvarianten identifiziert. Die Stämme wurden aus Stuhlproben von HUS/HC-Patienten, Lebensmitteln und Rinderkot gewonnen. Die eigenen Untersuchungen belegten das breite Vorkommen von β - und γ -Intimin, wengleich ζ - und auch in einem Fall κ -Intimin nachgewiesen wurde. Dabei handelte es sich mit O107:H11 um Stämme ohne Verotoxinbildungsvermögen. Als θ -Intiminvariante wurde unter anderem der Ausbruchstamm C6 KJS O111:H bestätigt. ε -Intimin wurde wie erwartet bei O103-, O8- und O11- und Or-Serovaren nachgewiesen. Der hohe Anteil an γ -Intimin ist allerdings auch auf den hohen Anteil von O157:H7 und O157:H zurückzuführen.

Die phylogenetische Analyse berücksichtigte neben Distanz-basierten- auch ein Maximum Likelihood-Verfahren und zeigte die Eignung aller drei Verfahren für die differenzierte Darstellung. In Erweiterung der bislang existierenden Einteilung (ZHANG et al., 2002b) konnten die folgenden sechs Untergruppen zugeordnet werden: $\alpha/\nu/\omicron/\zeta$, $\beta/\delta/\kappa$ (mit der engsten Relation zu den Außengruppen *E. coli* O115 und dem *eae*-Genhomolog von *Citrobacter freundii*). $\sigma/\gamma/\theta$, ι/λ und $\epsilon/\pi/\eta/\xi$ bilden die weiteren Untergruppen. ρ -Intimin wurde in der Mehrzahl der Fälle γ - bzw. θ -Intimin zugewiesen. Zusätzlich wurden die *eae*-Gene zweier aus Schafkot isolierter *E. coli*-Stämme sequenziert und als ζ -Intimin identifiziert. Die Netzwerkdarstellung anhand eines hochauflösenden Algorithmus ergab zusätzliche Hinweise für das Vorliegen genetischer Rekombinationen innerhalb der *eae*-Gene. Die Rekombinationsanalyse, vorgenommen mit sechs verschiedenen Verfahren, konnte mögliche rekombinante Bereiche aufzeigen. Ferner wurde belegt, dass alle *eae*-Gene unter einem negativen, bereinigenden Selektionsdruck stehen. Auf Proteinebene wurden dabei im hypervariablen Bereich Ladungsveränderungen gegenüber Variationen der Polarität bevorzugt. Die phylogenetischen Analysen stehen in guter Übereinstimmung mit der aktuellen wissenschaftlichen Literatur und erweitern den Kenntnisstand hinsichtlich der zum Zeitpunkt der Untersuchungen bekannten Intiminvarianten sowie der Anwendbarkeit hochauflösender Analyseverfahren. Somit wurde eine Bewertung der bestehenden Nomenklatur ermöglicht, die die Beibehaltung des Cut off-Wertes nach OSWALD et al. (2000) zu Grunde legt.

8 SUMMARY

Attaching and Effacing (AE)-lesions are the histopathological hallmark of enteropathogenic *Escherichia coli* (EPEC) and verocytotoxin-producing *Escherichia coli*/shigatoxin-producing *Escherichia coli* (VTEC/STEC). The *E. coli* attaching and effacing (*eae*)-gene encodes the protein intimin. Intimin mediates the intimate adherence to the enterocyte host cell. The subtyping of the attaching and effacing (*eae*)-gene has been proven to be of high diagnostic, epidemiologic and phenotypic value. More than 19 different *eae*-gene subtypes have been identified so far. The nomenclature is based on the Greek alphabet.

The aim of the presented thesis was the assessment of the distribution of *eae*-gene subtypes in *E. coli* strains with a special focus on a possible source association. For this purpose 317 *E. coli* strains from five different sources have been analyzed using a restriction fragment length polymorphism polymerase chain reaction (RFLP-PCR). The differentiation of the five intimin variants (α , β , γ , δ and ϵ) described at the first, was based on the method described by ADU-BOBIE et al. (1998a) and an own oligonucleotide primer pair for ϵ -intimin. The procedure described by RAMACHANDRAN et al. (2003) was augmented with a new RFLP typing scheme to address the remaining 14 gene variants. This means an improvement of the existing methodologies by the addition of seven new subtypes. 203 (64.03%) *E. coli* strains were classified as γ -intimin. 34 (11.04%) of the strains belonged to β -Intimin. The results confirm the large incidence of γ - and β -variants in strains from HUS/HC-patients, human food and bovine faeces. However, the rate of γ -intimin is also a consequence of the large amount of O157 strains used in this study. The RFLP analysis showed no sequence variations in the O157 PCR products. In addition, the two strains O86:H⁻ and O156:H⁻, both isolated from ovine faeces, have been identified as ζ -intimin variants by sequencing. More than 90% of the *eae* genes including the

complete 3'-region were determined and submitted to NCBI gene bank (AY 520904, AY520905). Additional H-antigen types were classified as ζ -intimin strains by the RFLP typing. ε -Intimin was detected in O103-, O8-, O11- and Or-serovars. The outbreak strain C6 KJS O111:H was identified as θ -intimin variant. κ -Intimin has been associated with a non verotoxin producing O107:H11 strain.

The phylogenetic analysis included two distance based- and one maximum-likelihood procedure. All three algorithms gave correct and reproducible results and allowed the correct estimation of the phylogenetic relationships. Six different subgroups were identified. The first group containing α / ν / o - and the sequenced ζ -intimin, the second one with β -/ δ -/ κ -intimin showed the closest relationship to the out groups of *E. coli* O115 and the *eae*-gene homologue of *Citrobacter freundii*. σ -/ γ -/ θ -, ι -/ λ - and the ε -/ π -/ η -/ ξ -intimin variants represented the groups three, four and five. ρ -intimin was classified as most closely related to γ 2-/ θ -intimin. Phylogenetic network analysis using the high resolution neighbor-net-algorithm revealed a network like structure. This showed the possible influence of recombination. Recombination events could subsequently be confirmed by six different mathematical models using ζ -intimin as a template. Furthermore the studies identified a negative/purifying selection for the complete coding sequences by using two different procedures. Also a more detailed estimation was done for the resulting protein sequences. The hypervariable 3'-region showed a preference of mutations, which resulted in a variation of the charge. Using the polarity as a criterion, mutations without a change of the status quo occurred more often. In contrast to that, the conserved 5'-region showed a clear overhang of conserved sequence variations for both parameters.

The phylogenetic analysis represented a good conformance with the literature. They enlarge the existing knowledge by considering all known 19 intimin subtypes being identified at the time of the studies and by using new, adopted methods. In addition an assessment of the scientific nomenclature for several

gene bank entries using the cut off value described by OSWALD et al. (2000) was enabled.

9 ANHANG

9.1 PHYLOGENETISCHE ANALYSE

Nachfolgend sind die für die phylogenetischen Analysen verwendeten Daten und Ergebnisse aufgeführt.

9.1.1 VERWENDETE REFERENZSTÄMME UND -SEQUENZEN

Tabelle 9.1.1-1: Für die phylogenetische Analyse und die Berechnung populationsgenetischer Parameter eingesetzte Referenzstämme

Genbank-zugriffs-codes	Stammbezeichnung/Serovar	Intimin-variante	Größe (Bp)	isoliert aus	Fundstelle
L11691.1	4280/ <i>Citrobacter freundii</i>	-	2811	Keine Angaben	SCHAUER und FALKOW, 1993
AB040740	MPEC/O115a,c:K(B)	alpha	2811	Keine Angaben	OKUTANI et al., 2000
AF022236	E2348/69/O127:H7	alpha	2820	Ausbruch, Kindergarten United Kingdom	ELLIOT et al., 1998
AF530555	k. A./O125:H6	alpha 2	2820	Gesundes Baby	RAMACHANDRAN et al., 2003
AF065628	1390/k.A.	beta	2820	Keine Angaben	AN et al., 1997
AF081186	DEC 11a, CDC 2254-5/O128:H2	beta	2820	Keine Angaben	MCGRAW et al. , 1999
AF081187	DEC 12a, Orskov F1-50/O111:H2	beta	2820	Keine Angaben	MCGRAW et al., 1999
AF099072	REPEC B10/k. A.	beta	971	Kaninchenkot	KREJANY, 2001
AF200363	RDEC-1/O15:H-	beta	2820	Kaninchenkot	ZHU et al., 2001
U60002	RDEC-1/O15:H-	beta	2820	Kaninchenkot	AGIN et al., 1996
AJ275091	IHIT0084/k. A.	beta	862	Kälberkot	ZEHMKE, 2001
U59502	REPEC 84/110/1/k. A.	beta	2820	Keine Angaben	KREJANY, 2001
AF530556	k. A./Ont:Hnt	beta 2	2820	Rinderkot	RAMACHANDRAN et al., 2003
AF530554	k. A./O2 related:H19	epsilon 2	2847	Ziegenkot	RAMACHANDRAN et al., 2003
AF116899	PMK5/O103:H2	epsilon1	2847	Stuhl, Human	OSWALD et al., 2000
AF081185	ECOR 37/O:H-	gamma	2805	Krallenaffe	MCGRAW,E.A. et al., 1999

Fortsetzung Tabelle 9.1.1-1: Für die phylogenetische Analyse und die Berechnung populationsgenetischer Parameter eingesetzte Referenzstämme

Genbank-zugriffs-codes	Stammbezeichnung/Serovar	Intimin-variante	Größe (Bp)	isoliert aus	Fundstelle
AF081184	DEC 5d; Karch 493/89/ O55:H7	gamma	2805	Mensch, Sri Lanka, 1965	McGRAW,E.A. et al., 1999
AF081183	DEC 3f/O157:H ⁻	gamma	2805	Keine Angaben	McGRAW,E.A. et al., 1999
Z11541	EDL 933/O157:H7	gamma	2805	Keine Angaben	YU und KAPER, 1992
AF025311.	O111.H ⁻	gamma 2	2808	HUS-Patient	VOSS et al., 1998
Y13112	k. A./O86:H34	delta	825	Keine Angaben	ADU-BOBIE et al., 1998a
AJ308551	7476/96/O145:H4	jota	2814	Stuhl, Human	ZHANG et al., 2002b
AF530553	k. A./OR:H ⁻	jota2	2814	Ziegenkot	RAMACHANDRAN et al., 2003
U66102	4221/k. A.	kappa	2823	Hundekot	AN et al., 1997
AF439538	97-3/k. A.	lamda	831	Keine Angaben	PAN, 2003
AF530557	k. A./O2/74:H ⁻	lamda	2817	Rinderkot	RAMACHANDRAN et al., 2003
AJ705049	EPEC-373/ O55:H51	my	2808	Stuhl, Human	BLANCO et al., 2004
AJ705050	EAEC-IH1229a/ O10:H ⁻	ny	2823	Stuhl, Human	BLANCO et al., 2004
AJ584841	AEEC-IH1301a/ O84:H ⁻	omikron	449	Stuhl, Human	BLANCO et al., 2004
AJ584840	AEEC-IH2997f/ O129:H ⁻	omikron	449	Stuhl, Human	BLANCO et al., 2004
AJ876648	AEEC-IH2997f/ O129:H ⁻	omikron	2820	Stuhl, Human	BLANCO et al., 2004
AJ716133	EPEC-113.4/ONT	pi	571	Stuhl, Human	BLANCO et al., 2004
AJ705052	AEEC-191.2/k. A.	pi	2847	Stuhl, Human	BLANCO et al., 2004
AJ748084	AEEC-M170a/ O180:H ⁻	rho	2829	Schafkot	BLANCO et al., 2004
AJ748082	AEEC-K226a/ O149:H-	rho	2829	Ziegenkot	BLANCO et al., 2004

Fortsetzung Tabelle 9.1.1-1: Für die phylogenetische Analyse und die Berechnung populationsgenetischer Parameter eingesetzte Referenzstämme

Genbank-zugriffs-codes	Stammbezeichnung/Serovar	Intimin-variante	Größe (Bp)	isoliert aus	Fundstelle
AJ748083	AEEC-K227a/O149:H-	rho	2829	Ziegenkot	BLANCO et al., 2004
AJ781125	EPEC-EC74699/O86:K61:H-	sigma	2808	Gewebeproben toter Vögel	BLANCO et al., 2004
AF449414	412/55/O111:H?	theta	2808	Keine Angaben	TARR und WHITTAM, 2002
AF253561	CL559/O91	gamma 2, theta	2808	Ziegenkot	CID et al., 2001
AF449420	DEC8b; 3030A-86/O111:H8	theta	2808	Keine Angaben	TARR und WHITTAM, 2002
AF449419	DEC8a; 2198-77/O111:NM	theta	2808	Mensch, USA, 1977	TARR und WHITTAM, 2002
AF449415	3215-99;F6627/O111:H8	theta	2808	Keine Angaben	TARR und WHITTAM, 2002
AJ746314	k. A./O71.H-	xi	575	Schafkot	BLANCO et al., 2004
AJ705051	STEC-B49/O80:H-	xi	2847	Rinderkot	BLANCO et al., 2004
AJ298279	537/89/O84:NM	zeta	2817	Keine Angaben	JORES et al., 2003
AJ275089	k. A./O84:K?	zeta	859	Keine Angaben	ZEHMKE, 2001
AJ271407	4795/97/O84:H4	zeta	2817	Keine Angaben	ZHANG et al., 2002
AF449417	921-B4/O111:H9	zeta	2817	Keine Angaben	TARR und WHITTAM, 2002
AF449416	9084-83/k. A.	zeta	2817	Keine Angaben	TARR und WHITTAM, 2002
AY520905	k. A./O84.H-	zeta	2675	Schafkot	SCHOENENBRUECHER et al., 2003
AY520904	k. A./O156.H-	zeta	2675	Schafkot	SCHOENENBRUECHER et al., 2003
AJ308550	11201/O125:H-	eta	2847	Keine Angaben	ZHANG et al., 2002b
X60439	CL-8/O157:H7		2805	Keine Angaben	BEEBAKHEE et al., 1999
AF081182	DEC 3a; CDC 3299-85/O157:H7		2805	Mensch, USA, 1985	MCGRAW et al., 1999

Tabelle 9.1.1-2: Sequenz-Übereinstimmungs-Matrix, erstellt anhand des CLUSTALW 1.6-Sequenzvergleichs

Sequenzen	alpha1	alpha2	beta1	beta2	gamma1	gamma2	delta	epsilon1	epsilon2	*04zeta	*05zeta	*07zeta
alpha1		0,955	0,833	0,879	0,856	0,843	0,179	0,827	0,855	0,865	0,865	0,914
alpha2	0,955		0,829	0,869	0,861	0,84	0,178	0,821	0,854	0,865	0,865	0,914
beta1	0,833	0,829		0,908	0,829	0,838	0,232	0,864	0,829	0,777	0,777	0,827
beta2	0,879	0,869	0,908		0,852	0,837	0,24	0,82	0,855	0,81	0,81	0,858
gamma1	0,856	0,861	0,829	0,852		0,913	0,18	0,829	0,861	0,796	0,796	0,844
gamma2	0,843	0,84	0,838	0,837	0,913		0,18	0,84	0,842	0,778	0,778	0,826
delta	0,179	0,178	0,232	0,24	0,18	0,18		0,178	0,179	0,18	0,18	0,172
epsilon1	0,827	0,821	0,864	0,82	0,829	0,84	0,178		0,916	0,767	0,767	0,815
epsilon2	0,855	0,854	0,829	0,855	0,861	0,842	0,179	0,916	0,916	0,803	0,803	0,852
AY520904 zeta	0,865	0,865	0,777	0,81	0,796	0,778	0,18	0,767	0,803			1
AY520905 zeta	0,865	0,865	0,777	0,81	0,796	0,778	0,18	0,767	0,803	1		0,948
AJ271407zeta	0,914	0,914	0,827	0,858	0,844	0,826	0,172	0,815	0,852	0,948	0,948	
AF449416zeta	0,919	0,917	0,83	0,865	0,844	0,83	0,172	0,816	0,853	0,936	0,936	0,986
AF449417zeta	0,919	0,917	0,83	0,865	0,844	0,83	0,172	0,816	0,853	0,936	0,936	0,986
AJ275089zeta	0,24	0,241	0,19	0,188	0,18	0,18	0,56	0,183	0,186	0,32	0,32	0,304
eta	0,851	0,843	0,819	0,849	0,839	0,828	0,182	0,921	0,931	0,787	0,787	0,836
theta	0,844	0,84	0,84	0,838	0,913	0,996	0,18	0,841	0,844	0,779	0,779	0,827
iota1	0,875	0,866	0,835	0,869	0,884	0,856	0,182	0,835	0,87	0,802	0,802	0,852
iota2	0,815	0,813	0,815	0,814	0,809	0,813	0,177	0,808	0,814	0,759	0,759	0,806
kappa	0,88	0,87	0,886	0,94	0,856	0,842	0,29	0,821	0,855	0,807	0,807	0,857
lambda	0,864	0,867	0,839	0,862	0,87	0,844	0,184	0,84	0,866	0,81	0,81	0,859
my	0,815	0,813	0,815	0,814	0,809	0,813	0,177	0,808	0,814	0,759	0,759	0,806
nu	0,889	0,897	0,821	0,826	0,82	0,815	0,184	0,814	0,815	0,846	0,846	0,893
xi	0,825	0,82	0,867	0,824	0,831	0,842	0,183	0,973	0,916	0,769	0,769	0,816
omikron	0,869	0,872	0,818	0,824	0,811	0,815	0,178	0,809	0,808	0,815	0,815	0,861
pi	0,858	0,852	0,832	0,853	0,851	0,839	0,18	0,922	0,952	0,797	0,797	0,847
rho	0,826	0,821	0,823	0,819	0,837	0,829	0,178	0,831	0,824	0,769	0,769	0,814
sigma	0,829	0,829	0,866	0,825	0,917	0,92	0,179	0,867	0,832	0,772	0,772	0,821
<i>E. coli</i> O115	0,803	0,798	0,86	0,869	0,79	0,79	0,246	0,79	0,793	0,752	0,752	0,796
<i>C. freundii</i>	0,801	0,796	0,858	0,867	0,787	0,788	0,246	0,788	0,791	0,75	0,75	0,794

Fortsetzung Tabelle 9.1.1-2: Sequenz-Übereinstimmungs-Matrix, erstellt anhand des CLUSTALW 1.6-Sequenzvergleichs

Sequenzen	*89zeta	eta	theta	iota1	iota2	kappa	lambda	my	nu	xi	omikron	pi
alpha1	0,24	0,851	0,844	0,875	0,815	0,88	0,864	0,815	0,889	0,825	0,869	0,858
alpha2	0,241	0,843	0,84	0,866	0,813	0,87	0,867	0,813	0,897	0,82	0,872	0,852
beta1	0,19	0,819	0,84	0,835	0,815	0,886	0,839	0,815	0,821	0,867	0,818	0,832
beta2	0,188	0,849	0,838	0,869	0,814	0,94	0,862	0,814	0,826	0,824	0,824	0,853
gamma1	0,18	0,839	0,913	0,864	0,809	0,856	0,87	0,809	0,82	0,831	0,811	0,851
gamma2	0,18	0,828	0,996	0,856	0,813	0,842	0,844	0,813	0,815	0,842	0,815	0,839
delta	0,56	0,182	0,18	0,182	0,177	0,29	0,184	0,177	0,184	0,183	0,178	0,18
epsilon1	0,183	0,921	0,841	0,835	0,808	0,821	0,84	0,808	0,814	0,973	0,809	0,922
epsilon2	0,186	0,931	0,844	0,87	0,814	0,855	0,866	0,814	0,815	0,916	0,808	0,952
AY520904 zeta	0,32	0,787	0,779	0,802	0,759	0,807	0,81	0,759	0,846	0,769	0,815	0,797
AY520905 zeta	0,32	0,787	0,779	0,802	0,759	0,807	0,81	0,759	0,846	0,769	0,815	0,797
AJ271407zeta	0,304	0,836	0,827	0,852	0,806	0,857	0,859	0,806	0,893	0,816	0,861	0,847
AF449416zeta	0,3	0,843	0,83	0,859	0,806	0,864	0,856	0,806	0,893	0,818	0,863	0,851
AF449417zeta	0,3	0,843	0,83	0,859	0,806	0,864	0,856	0,806	0,893	0,818	0,863	0,851
AJ275089zeta	0,182	0,182	0,179	0,18	0,182	0,183	0,19	0,182	0,256	0,184	0,224	0,184
eta	0,182	0,829	0,829	0,849	0,825	0,855	0,846	0,825	0,835	0,915	0,833	0,933
theta	0,179	0,829	0,829	0,857	0,814	0,843	0,845	0,814	0,816	0,843	0,816	0,84
iota1	0,18	0,849	0,857	0,891	0,891	0,873	0,884	0,891	0,827	0,837	0,827	0,869
iota2	0,182	0,825	0,814	0,891	0,816	0,816	0,826	1	0,851	0,811	0,852	0,811
kappa	0,183	0,855	0,843	0,873	0,816	0,865	0,865	0,816	0,832	0,827	0,829	0,859
lambda	0,19	0,846	0,845	0,884	0,826	0,865	0,865	0,826	0,834	0,841	0,845	0,869
my	0,182	0,825	0,814	0,891	0,826	0,816	0,826	0,826	0,851	0,811	0,852	0,811
nu	0,256	0,835	0,816	0,827	0,851	0,832	0,834	0,851	0,814	0,814	0,915	0,813
xi	0,184	0,915	0,843	0,837	0,811	0,827	0,841	0,811	0,814	0,81	0,81	0,918
omikron	0,224	0,833	0,816	0,827	0,852	0,829	0,845	0,852	0,915	0,81	0,807	0,807
pi	0,184	0,933	0,84	0,869	0,811	0,859	0,869	0,811	0,813	0,918	0,807	0,828
rho	0,185	0,829	0,83	0,83	0,846	0,819	0,833	0,846	0,855	0,836	0,849	0,828
sigma	0,183	0,819	0,922	0,844	0,814	0,83	0,84	0,814	0,822	0,871	0,818	0,834
<i>E. coli</i> O115	0,183	0,794	0,791	0,795	0,783	0,873	0,799	0,783	0,805	0,792	0,795	0,789
<i>C. freundii</i>	0,183	0,792	0,789	0,793	0,782	0,871	0,797	0,782	0,803	0,79	0,793	0,787

Fortsetzung Tabelle 9.1.1.1-2: Sequenz-Übereinstimmungs-Matrix, erstellt anhand des CLUSTALW 1.6-Sequenzvergleichs

Sequenzen	<i>E. coli</i> O115	<i>C. freundii</i>
alpha1	0,803	0,801
alpha2	0,798	0,796
beta1	0,86	0,858
beta2	0,869	0,867
gamma1	0,79	0,787
gamma2	0,79	0,788
delta	0,246	0,246
epsilon1	0,79	0,788
epsilon2	0,793	0,791
AY520904 zeta	0,752	0,75
AY520905 zeta	0,752	0,75
AJ271407zeta	0,796	0,794
AF449416zeta	0,796	0,794
AF449417zeta	0,796	0,794
AJ275089zeta	0,183	0,183
eta	0,794	0,792
theta	0,791	0,789
iota1	0,795	0,793
iota2	0,783	0,782
kappa	0,873	0,871
lambda	0,799	0,797
my	0,783	0,782
nu	0,805	0,803
xi	0,792	0,79
omikron	0,795	0,793
pi	0,789	0,787
rho	0,793	0,791
sigma	0,793	0,791
<i>E. coli</i> O115		0,997
<i>C. freundii</i>	0,997	

Auflistung der Distanzmatrix zur Erstellung des Maximum-Likelihood-Baumes im „Phylip“-Format

((((AF530554ep:0.02518,(AJ716133pi:0.00007,AJ705052pi:0.00007):0.02409):0.00773,(AJ308550et:0.03295,((AJ746314xi:0.00007,AJ705051xi:0.00007):0.01638,AF116899ep:0.01093):0.04226):0.00882):0.06010,(((AJ748083rh:0.00007,AJ748082rh:0.00007):0.00007,AJ748084rh:0.00007):0.11911,((Z11541gamm:0.00178,(X60439gamm:0.00107,AF081183ga:0.00007):0.00007):0.00007,AF081185.g:0.00071):0.02606,(aj781125si:0.03985,(AF025311O1:0.00252,(AF253561O9:0.00675,((AF449414th:0.00007,AF449419th:0.00007):0.00007,AF449420th:0.00007):0.00007):0.00101):0.04136):0.02276):0.03785):0.02400,((AF439538.l:0.01149,AF530557la:0.00041):0.06357,((AF530553my:0.00007,AF530553io:0.00007):0.09611,AJ308551io:0.02167):0.04237):0.02004):0.01230):0.01505,(((U66102kapp:0.00908,Y13112part:0.00119):0.00007,AJ308552ka:0.00007):0.02842,(((L11691.1Ci:0.00178,AB040740Ec:0.00038):0.09558,((U60002beta:0.00007,AF200363be:0.00007):0.00032,((AF065628be:0.00564,AF099072pa:0.00007):0.00007,(AF081187be:0.00007,AF081186be:0.00007):0.00007):0.00007):0.05671):0.02607,AF530556be:0.02179):0.01560):0.04740,(((AF449416ze:0.00007,AF449417ze:0.00007):0.00484,(AJ275089ze:0.00007,((AY520904HS:0.00007,AY520905HS:0.00007):0.00031,(AJ271407ze:0.00007,AJ298279ze:0.00007):0.00007):0.00007):0.00866):0.03678,(((AJ584841om:0.00007,AJ584840pa:0.00007):0.00007,AJ876648om:0.00007):0.06470,AJ705050nu:0.02373):0.05331):0.01260,(M58154alph:0.02340,AF530555al:0.02317):0.01357):0.04862):0.01295);

Tabelle 9.1.1-3: Berechnung des wahrscheinlichsten Baumes mit DNA-ML V.3.6a3 (enthalten in PAML)

Bezugspunkte Gensequenzen und Verzweigungen		Distanzlängen	Approximierte Konfidenzlimits		Signifi- kanz
2	AF530555al	0.02317	(0.01702,	0.02923)	**
2	1	0.01357	(0.00833,	0.01879)	**
1	6	0.04862	(0.03914,	0.05813)	**
6	4	0.02800	(0.02011,	0.03586)	**
4	18	0.06010	(0.04966,	0.07055)	**
18	7	0.00773	(0.00376,	0.01167)	**
7	AF530554ep	0.02518	(0.01898,	0.03138)	**
7	39	0.02409	(0.01800,	0.03017)	**
39	AJ716133pi	0.00007	(zero,	infinity)	
39	AJ705052pi	0.00007	(zero,	0.00188)	
18	10	0.00882	(0.00435,	0.01330)	**
10	AJ308550et	0.03295	(0.02587,	0.04003)	**
10	25	0.04226	(0.03429,	0.05022)	**
25	40	0.01638	(0.01128,	0.02140)	**
40	AJ746314xi	0.00007	(zero,	infinity)	
40	AJ705051xi	0.00007	(zero,	0.00176)	
25	AF116899ep	0.01093	(0.00672,	0.01520)	**
4	13	0.01230	(0.00598,	0.01872)	**
13	19	0.02400	(0.01604,	0.03202)	**
19	42	0.11911	(0.10477,	0.13345)	**
42	41	0.00007	(zero,	infinity)	
41	AJ748083rh	0.00007	(zero,	infinity)	
41	AJ748082rh	0.00007	(zero,	infinity)	
42	AJ748084rh	0.00007	(zero,	infinity)	
19	5	0.03785	(0.02897,	0.04667)	**
5	45	0.02606	(0.01905,	0.03307)	**
45	47	0.00007	(zero,	0.00028)	
47	Z11541gamm	0.00178	(0.00022,	0.00334)	**
47	46	0.00007	(zero,	infinity)	
46	X60439gamm	0.00107	(zero,	0.00228)	**
46	AF081183ga	0.00007	(zero,	infinity)	
45	AF081185.g	0.00071	(zero,	0.00173)	*
5	21	0.02276	(0.01633,	0.02921)	**
21	aj781125si	0.03985	(0.03183,	0.04782)	**
21	11	0.04136	(0.03322,	0.04950)	**
11	AF025311O1	0.00252	(0.00065,	0.00442)	**
11	48	0.00101	(zero,	0.00225)	**
48	AF253561O9	0.00675	(0.00375,	0.00983)	**
48	43	0.00007	(zero,	infinity)	
43	44	0.00007	(zero,	infinity)	

Fortsetzung Tabelle 9.1.1-3: Berechnung des wahrscheinlichsten Baumes mit DNA-ML V.3.6a3 (enthalten in PAML)

Bezugspunkte (Gensequenzen und Verzweigungen)		Distanzlängen	Approximierte Konfidenzlimits		Signifi- kanz
44	AF449414th	0.00007	(zero,	infinity)	
44	AF449419th	0.00007	(zero,	infinity)	
43	AF449420th	0.00007	(zero,	infinity)	
13	15	0.02004	(0.01300,	0.02703)	**
15	38	0.06357	(0.05313,	0.07401)	**
38	AF439538.l	0.01149	(0.00418,	0.01883)	**
38	AF530557la	0.00041	(zero,	0.00164)	
15	17	0.04237	(0.03337,	0.05128)	**
17	16	0.09611	(0.08368,	0.10854)	**
16	AF530553my	0.00007	(zero,	infinity)	
16	AF530553io	0.00007	(zero,	infinity)	
17	AJ308551io	0.02167	(0.01528,	0.02809)	**
6	14	0.04740	(0.03781,	0.05699)	**
14	36	0.02842	(0.02129,	0.03550)	**
36	37	0.00007	(zero,	0.00189)	
37	U66102kapp	0.00908	(0.00548,	0.01266)	**
37	Y13112part	0.00119	(zero,	0.00356)	*
36	AJ308552ka	0.00007	(zero,	0.00097)	
14	3	0.01560	(0.00986,	0.02142)	**
3	22	0.02607	(0.01906,	0.03309)	**
22	23	0.09558	(0.08302,	0.10813)	**
23	L11691.1Ci	0.00178	(0.00019,	0.00338)	**
23	AB040740Ec	0.00038	(zero,	0.00119)	
22	20	0.05671	(0.04697,	0.06647)	**
20	32	0.00032	(zero,	0.00100)	
32	U60002beta	0.00007	(zero,	infinity)	
32	AF200363be	0.00007	(zero,	infinity)	
20	33	0.00007	(zero,	0.00089)	
33	35	0.00007	(zero,	0.00171)	
35	AF065628be	0.00564	(0.00286,	0.00844)	**
35	AF099072pa	0.00007	(zero,	infinity)	
33	34	0.00007	(zero,	infinity)	
34	AF081187be	0.00007	(zero,	infinity)	
34	AF081186be	0.00007	(zero,	infinity)	
3	AF530556be	0.02179	(0.01548,	0.02804)	**
1	8	0.01260	(0.00722,	0.01804)	**
8	26	0.03678	(0.02894,	0.04472)	**

Fortsetzung Tabelle 9.1.1-3: Berechnung des wahrscheinlichsten Baumes mit DNA-ML V.3.6a3 (enthalten in PAML)

Bezugspunkte (Gensequenzen und Verzweigungen)		Distanzlängen	Approximierte Konfidenzlimits		Signifi- kanz
26	27	0.00484	(0.00212,	0.00754)	**
27	AF449416ze	0.00007	(zero,	infinity)	
27	AF449417ze	0.00007	(zero,	infinity)	
26	24	0.00866	(0.00514,	0.01225)	**
24	AJ275089ze	0.00007	(zero,	infinity)	
24	28	0.00007	(zero,	0.00251)	
28	29	0.00031	(zero,	0.00097)	
29	AY520904HS	0.00007	(zero,	infinity)	
29	AY520905HS	0.00007	(zero,	infinity)	
28	9	0.00007	(zero,	0.00080)	
9	AJ271407ze	0.00007	(zero,	0.00026)	
9	AJ298279ze	0.00007	(zero,	infinity)	
8	12	0.05331	(0.04405,	0.06263)	**
12	30	0.06470	(0.05456,	0.07496)	**
30	31	0.00007	(zero,	infinity)	
31	AJ584841om	0.00007	(zero,	infinity)	
31	AJ584840pa	0.00007	(zero,	infinity)	
30	AJ876648om	0.00007	(zero,	0.00265)	
12	AJ705050nu	0.02373	(0.01705,	0.03039)	**
2	M58154alph	0.02340	(0.01735,	0.02958)	**

* signifikant positiv P < 0.05

** signifikant positiv P < 0.01

Empirisch ermittelte Basenfrequenzen:

A 0.29784

C 0.19017

G 0.23323

T(U) 0.27876

Verhältnis der Transitionen und Transversionen: 1.2

9.1.2 DATEN DER NETZWERKDARSTELLUNG

Tabelle 9.1.2-1: Daten der Neighbor-Net-Darstellung nach BRYANT und MOULTON (2004) für den gesamten Intimingebereich

Dimensionen: Anzahl der Taxa=25 Anzahl der Splits=66
 Eigenschaften: Fit=95.72 zyklisch
 Zyklen 1 10 24 3 18 20 13 17 14 16 7 21 6 12 25 8 11 19 9 5 23 22 15 4 2;

MATRIX		
Anzahl und Größe der Splits	Distanzen	Einbezogene Referenzsequenzen
[1, size=1]	1,88E+03	1 2 3 4 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[2, size=2]	3,94E-02	1 2 3 4 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21 22 23 25,
[3, size=3]	4,10E-03	1 2 4 5 6 7 8 9 11 12 13 14 15 16 17 18 19 20 21 22 23 25,
[4, size=4]	2,11E-03	1 2 4 15,
[5, size=2]	1,36E-02	1 2,
[6, size=1]	2,13E-02	1
[7, size=2]	4,06E+03	1 2 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 25,
[8, size=5]	1,22E+03	1 2 4 10 15,
[9, size=1]	2,63E-02	1 2 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[10, size=2]	6,43E-03	1 2 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[11, size=5]	2,47E-02	1 2 4 5 6 7 8 9 10 11 12 14 15 16 19 21 22 23 24 25,
[12, size=6]	5,25E-03	1 2 4 10 15 24,
[13, size=5]	8,19E+03	1 2 4 10 24,
[14, size=4]	9,16E+00	1 2 10 24,
[15, size=1]	4,31E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 19 20 21 22 23 24 25,
[16, size=12]	1,37E+03	1 2 3 4 5 9 10 11 15 19 22 23 24,
[17, size=5]	9,05E-03	1 2 3 10 24,
[18, size=4]	1,40E-03	1 3 10 24,
[19, size=1]	6,68E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 21 22 23 24 25,
[20, size=10]	3,57E-03	1 2 3 4 10 15 18 22 23 24,
[21, size=6]	1,74E-02	1 2 3 10 18 24,
[22, size=5]	1,22E-03	1 3 10 18 24,
[23, size=1]	9,78E+02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 18 19 20 21 22 23 24 25,
[24, size=2]	4,72E-02	1 2 3 4 5 6 7 8 9 10 11 12 14 15 16 18 19 20 21 22 23 24 25,
[25, size=3]	2,85E-02	1 2 3 4 5 6 7 8 9 10 11 12 15 16 18 19 20 21 22 23 24 25,
[26, size=1]	9,98E+01	1 2 3 4 5 6 7 8 9 10 11 12 14 15 16 17 18 19 20 21 22 23 24 25,
[27, size=1]	2,64E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 15 16 17 18 19 20 21 22 23 24 25,
[28, size=2]	3,79E-03	1 2 3 4 5 6 7 8 9 10 11 12 13 15 17 18 19 20 21 22 23 24 25,
[29, size=3]	4,39E-03	1 2 3 4 5 6 8 9 10 11 12 13 15 17 18 19 20 21 22 23 24 25,
[30, size=1]	5,24E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 17 18 19 20 21 22 23 24 25,
[31, size=2]	2,80E-03	1 2 3 4 5 6 8 9 10 11 12 13 14 15 17 18 19 20 21 22 23 24 25,
[32, size=1]	2,82E-02	1 2 3 4 5 6 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[33, size=2]	5,29E-03	1 2 3 4 5 6 8 9 10 11 12 13 14 15 16 17 18 19 20 22 23 24 25,
[34, size=4]	2,80E-02	1 2 3 4 5 8 9 10 11 13 14 15 16 17 18 19 20 22 23 24 25,
[35, size=1]	3,28E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 22 23 24 25,
[36, size=5]	1,08E-02	1 2 3 4 5 7 9 10 11 13 14 15 16 17 18 19 20 22 23 24,
[37, size=1]	1,83E-03	1 2 3 4 5 6 7 8 9 10 11 13 14 15 16 17 18 19 20 21 22 23 24 25,
[38, size=2]	3,55E-02	1 2 3 4 5 7 8 9 10 11 13 14 15 16 17 18 19 20 21 22 23 24 25,
[39, size=1]	2,67E-03	1 2 3 4 5 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[40, size=1]	1,33E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24,
[41, size=2]	2,23E-02	1 2 3 4 5 6 7 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24,
[42, size=3]	1,80E+03	1 2 3 4 5 6 7 9 10 12 13 14 15 16 17 18 19 20 21 22 23 24,
[43, size=4]	2,89E-03	1 2 3 4 5 6 7 9 10 12 13 14 15 16 17 18 20 21 22 23 24,
[44, size=5]	2,59E-02	1 2 3 4 5 6 7 10 12 13 14 15 16 17 18 20 21 22 23 24,
[45, size=6]	5,82E-03	1 2 3 4 6 7 10 12 13 14 15 16 17 18 20 21 22 23 24,
[46, size=10]	1,33E-03	1 2 3 6 7 10 12 13 14 16 17 18 20 21 24,
[47, size=1]	1,13E-02	1 2 3 4 5 6 7 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[48, size=2]	1,43E-03	1 2 3 4 5 6 7 9 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[49, size=4]	2,43E-03	1 2 3 4 5 6 7 10 12 13 14 15 16 17 18 20 21 22 23 24 25,
[50, size=1]	3,29E-02	1 2 3 4 5 6 7 8 9 10 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[51, size=3]	7,17E+02	1 2 3 4 5 6 7 8 10 12 13 14 15 16 17 18 20 21 22 23 24 25,
[52, size=1]	2,23E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20 21 22 23 24 25,
[53, size=2]	9,50E-03	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 20 21 22 23 24 25,
[54, size=1]	2,27E-02	1 2 3 4 5 6 7 8 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[55, size=1]	4,36E-02	1 2 3 4 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[56, size=3]	4,66E-03	1 2 3 4 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 24 25,
[57, size=5]	2,61E-03	1 2 3 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21 24 25,
[58, size=1]	2,46E-03	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 24 25,
[59, size=2]	7,83E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 24 25,
[60, size=3]	2,63E-03	1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21 24 25,
[61, size=4]	5,32E+03	1 2 3 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21 24 25,
[62, size=1]	1,12E-03	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 23 24 25,
[63, size=1]	2,74E-02	1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21 22 23 24 25,
[64, size=3]	3,73E-03	1 3 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21 22 23 24 25,
[65, size=1]	3,06E-02	1 2 3 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,
[66, size=1]	1,75E-02	1 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25,

9.1.3 BERECHNUNG DES SELEKTIONSMECHANISMUS

Exemplarisch werden die mit HON-NEW (ZHANG, 2000) generierten Daten radikaler (p_{NR}) und konservativer (p_{NC}) Mutationen angegeben.

Tabelle 9.1.3-1: 3'-Bereich: Berechnung von „ p_{NR} “ für den Parameter „Polarität“

	alpha 1	alpha 2	beta 1	beta 2	gamma 1	gamma 2	epsilon 1	epsilon 2
alpha 1		0.084	0.230	0.198	0.207	0.234	0.241	0.241
alpha 2	0.022		0.221	0.190	0.208	0.248	0.246	0.227
beta 1	0.036	0.036		0.044	0.194	0.204	0.247	0.225
beta 2	0.034	0.033	0.016		0.201	0.217	0.241	0.219
gamma 1	0.034	0.034	0.035	0.035		0.055	0.211	0.218
gamma 2	0.037	0.038	0.036	0.037	0.019		0.228	0.226
epsilon 1	0.037	0.038	0.039	0.039	0.037	0.038		0.092
epsilon 2	0.037	0.036	0.036	0.037	0.038	0.038	0.024	
zeta HS	0.023	0.023	0.039	0.037	0.035	0.038	0.037	0.037
zetaW	0.023	0.023	0.039	0.037	0.036	0.039	0.037	0.037
eta	0.037	0.037	0.037	0.037	0.037	0.036	0.016	0.022
theta	0.037	0.037	0.034	0.035	0.019	0.000	0.037	0.038
iota 1	0.035	0.037	0.038	0.042	0.039	0.040	0.038	0.039
iota 2	0.036	0.038	0.038	0.041	0.038	0.038	0.037	0.037
lambda	0.035	0.037	0.035	0.036	0.034	0.034	0.037	0.035

	zeta HS	eta	theta	iota 1	iota 2	lambda
alpha 1	0.090	0.237	0.234	0.217	0.231	0.204
alpha 2	0.084	0.242	0.242	0.245	0.255	0.233
beta 1	0.253	0.227	0.197	0.239	0.239	0.193
beta 2	0.218	0.228	0.211	0.282	0.269	0.187
gamma 1	0.228	0.218	0.055	0.236	0.223	0.172
gamma 2	0.250	0.209	0.000	0.249	0.224	0.170
epsilon 1	0.251	0.039	0.225	0.226	0.237	0.200
epsilon 2	0.239	0.078	0.226	0.241	0.241	0.198
zeta HS		0.264	0.247	0.257	0.264	0.228
zetaW	0.006	0.270	0.260	0.263	0.270	0.245
eta	0.038		0.209	0.229	0.244	0.201
theta	0.038	0.036		0.246	0.224	0.171
iota 1	0.038	0.038	0.039		0.047	0.193
iota 2	0.037	0.038	0.038	0.018		0.189
lambda	0.035	0.036	0.034	0.036	0.036	

Tabelle 9.1.3-2: 3'-Bereich: Berechnung von „ p_{NC} “ für den Parameter „Polarität“

	alpha 1	alpha 2	beta 1	beta 2	gamma 1	gamma 2	epsilon 1	epsilon 2
alpha 1		0.056	0.330	0.353	0.364	0.351	0.349	0.360
alpha 2	0.013		0.321	0.340	0.365	0.343	0.345	0.350
beta 1	0.033	0.032		0.055	0.355	0.392	0.366	0.369
beta 2	0.035	0.033	0.013		0.369	0.403	0.368	0.360
gamma 1	0.033	0.033	0.033	0.033		0.168	0.383	0.372
gamma 2	0.032	0.032	0.033	0.033	0.024		0.365	0.365
epsilon 1	0.031	0.032	0.035	0.034	0.033	0.032		0.102
epsilon 2	0.031	0.032	0.034	0.034	0.033	0.032	0.017	
zeta HS	0.024	0.022	0.030	0.032	0.032	0.031	0.031	0.032
zetaW	0.022	0.023	0.029	0.031	0.032	0.031	0.032	0.032
eta	0.031	0.032	0.035	0.035	0.033	0.032	0.014	0.020
theta	0.032	0.032	0.033	0.033	0.023	0.003	0.032	0.032
iota 1	0.035	0.033	0.035	0.033	0.032	0.032	0.034	0.033
iota 2	0.034	0.032	0.033	0.032	0.033	0.033	0.031	0.031
lambda	0.031	0.030	0.035	0.035	0.034	0.033	0.031	0.030

	zeta HS	zetaW	eta	theta	iota 1	iota 2	lambda
alpha 1	0.184	0.161	0.359	0.347	0.369	0.370	0.321
alpha 2	0.170	0.158	0.350	0.346	0.369	0.349	0.309
beta 1	0.309	0.292	0.363	0.395	0.394	0.375	0.350
beta 2	0.341	0.326	0.367	0.406	0.380	0.356	0.359
gamma 1	0.350	0.348	0.359	0.166	0.367	0.367	0.373
gamma 2	0.344	0.341	0.357	0.003	0.359	0.352	0.350
epsilon 1	0.372	0.351	0.047	0.367	0.350	0.332	0.312
epsilon 2	0.381	0.359	0.109	0.362	0.353	0.331	0.317
zeta HS		0.033	0.383	0.346	0.400	0.376	0.345
zetaW	0.010		0.363	0.344	0.386	0.351	0.323
eta	0.033	0.033		0.355	0.354	0.334	0.312
theta	0.031	0.032	0.032		0.361	0.355	0.348
iota 1	0.034	0.034	0.034	0.032		0.106	0.292
iota 2	0.032	0.032	0.031	0.033	0.020		0.295
lambda	0.031	0.031	0.031	0.032	0.031	0.029	

Tabelle 9.1.3-3: 3'-Bereich: Berechnung von „ p_{NR} “ für den Parameter „Ladung“

	alpha 1	alpha 2	beta 1	beta 2	gamma 1	gamma 2	epsilon 1	epsilon 2
alpha 1	0.000	0.053	0.336	0.347	0.334	0.361	0.361	0.375
alpha 2	0.018	0.000	0.301	0.295	0.316	0.343	0.355	0.356
beta 1	0.043	0.042	0.000	0.054	0.384	0.413	0.371	0.388
beta 2	0.043	0.041	0.018	0.000	0.408	0.441	0.362	0.355
gamma 1	0.042	0.041	0.044	0.043	0.000	0.171	0.391	0.416
gamma 2	0.043	0.043	0.044	0.045	0.030	0.000	0.384	0.422
epsilon 1	0.041	0.043	0.043	0.044	0.040	0.041	0.000	0.142
epsilon 2	0.042	0.042	0.043	0.042	0.040	0.041	0.028	0.000
zeta HS	0.031	0.032	0.040	0.041	0.041	0.042	0.041	0.043
zetaW	0.029	0.032	0.041	0.042	0.041	0.043	0.041	0.043
eta	0.042	0.043	0.043	0.045	0.041	0.041	0.020	0.027
theta	0.043	0.043	0.044	0.045	0.031	0.000	0.041	0.042
iota 1	0.044	0.045	0.046	0.046	0.041	0.043	0.043	0.042
iota 2	0.045	0.046	0.045	0.045	0.042	0.045	0.040	0.040
lambda	0.039	0.040	0.043	0.043	0.042	0.042	0.039	0.038

	zeta HS	zetaW	eta	theta	iota 1	iota 2	lambda
alpha 1	0.173	0.150	0.381	0.360	0.402	0.404	0.302
alpha 2	0.170	0.169	0.362	0.346	0.392	0.386	0.329
beta 1	0.310	0.310	0.380	0.419	0.441	0.413	0.368
beta 2	0.334	0.334	0.374	0.447	0.433	0.390	0.374
gamma 1	0.314	0.324	0.384	0.174	0.360	0.341	0.378
gamma 2	0.343	0.345	0.378	0.000	0.388	0.366	0.362
epsilon 1	0.395	0.370	0.065	0.388	0.383	0.355	0.312
epsilon 2	0.414	0.387	0.134	0.424	0.421	0.377	0.327
zeta HS	0.000	0.043	0.419	0.345	0.405	0.427	0.335
zetaW	0.016	0.000	0.400	0.348	0.408	0.408	0.322
eta	0.044	0.044	0.000	0.381	0.387	0.364	0.310
theta	0.043	0.044	0.041	0.000	0.391	0.366	0.364
iota 1	0.045	0.045	0.043	0.043	0.000	0.084	0.325
iota 2	0.045	0.044	0.042	0.045	0.024	0.000	0.285
lambda	0.040	0.041	0.039	0.042	0.039	0.037	0.000

Tabelle 9.1.3-4: 3'-Bereich: Berechnung von „ p_{NC} “ für den Parameter „Ladung“

	alpha 1	alpha 2	beta 1	beta 2	gamma 1	gamma 2	epsilon 1	epsilon 2
alpha 1	0.000	0.072	0.273	0.274	0.297	0.283	0.285	0.289
alpha 2	0.016	0.000	0.278	0.285	0.309	0.293	0.288	0.282
beta 1	0.026	0.028	0.000	0.050	0.256	0.282	0.300	0.283
beta 2	0.027	0.028	0.012	0.000	0.262	0.286	0.304	0.288
gamma 1	0.027	0.029	0.026	0.026	0.000	0.108	0.289	0.267
gamma 2	0.027	0.028	0.027	0.027	0.020	0.000	0.283	0.260
epsilon 1	0.030	0.030	0.031	0.030	0.029	0.028	0.000	0.074
epsilon 2	0.030	0.029	0.029	0.029	0.027	0.027	0.016	0.000
zeta HS	0.020	0.018	0.027	0.027	0.029	0.028	0.029	0.028
zetaW	0.020	0.018	0.027	0.027	0.029	0.028	0.029	0.028
eta	0.030	0.030	0.029	0.029	0.028	0.026	0.011	0.017
theta	0.027	0.028	0.027	0.027	0.019	0.003	0.028	0.027
iota 1	0.028	0.027	0.026	0.028	0.028	0.027	0.029	0.027
iota 2	0.028	0.027	0.027	0.027	0.029	0.028	0.028	0.027
lambda	0.029	0.027	0.027	0.028	0.027	0.027	0.028	0.027

	zeta HS	zetaW	eta	theta	iota 1	iota 2	lambda
alpha 1	0.140	0.132	0.282	0.280	0.272	0.280	0.272
alpha 2	0.125	0.116	0.287	0.291	0.293	0.281	0.258
beta 1	0.279	0.270	0.282	0.278	0.290	0.285	0.260
beta 2	0.280	0.273	0.290	0.283	0.301	0.294	0.263
gamma 1	0.306	0.304	0.272	0.104	0.303	0.307	0.269
gamma 2	0.295	0.298	0.268	0.003	0.287	0.278	0.252
epsilon 1	0.296	0.291	0.033	0.282	0.268	0.270	0.255
epsilon 2	0.287	0.282	0.079	0.256	0.257	0.259	0.250
zeta HS	0.000	0.014	0.300	0.295	0.324	0.291	0.291
zetaW	0.007	0.000	0.294	0.297	0.311	0.278	0.283
eta	0.029	0.029	0.000	0.264	0.271	0.271	0.257
theta	0.028	0.028	0.026	0.000	0.286	0.282	0.248
iota 1	0.028	0.028	0.029	0.027	0.000	0.088	0.224
iota 2	0.027	0.027	0.028	0.028	0.017	0.000	0.248
lambda	0.028	0.028	0.028	0.026	0.026	0.028	0.000

Tabelle 9.1.3-5: Serovarenverteilung unter Berücksichtigung der nachgewiesenen Intiminvarianten

tRNA	Intimintype	O-Antigen	H-Antigen	Habitat
<i>selC</i>	α	OX3, O51, O55, O63, O85, O103, O111, O125, O127, O132, O142, O156, O157, ONT	2, 4, 6, 16, 29, 34, 45, 49, NM, NT	Hund, Mensch
<i>selC</i> <i>pheU, pheV</i>	β, δ	O2, O5, 07, O11, O14, O15, O16, O17, O18, O20, O26, O33, O39, O45, O49, O65, O68, O70, O75, O81, O86, O88, O90, O91, O 103, O104, O108, O109, O110, O111, O114, O118, O119, O121, O123, O126, O127, O128, O128ab, O132, O145, O153, O157, O167, ONT, OR, OX3	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 16, 17, 18, 20, 21, 25, 26, 30, 32, 34, 45, 54, NM, NT	Rind, Huhn, Hund, Mensch, Schwein, Ziege, Meersschwein, Hase, Reh, Schaf
<i>pheU, pheV</i>	ϵ	O4, O8, O11, O15, O45, O80, O96, O102, O103, O121, O157, O165, O166, O172, ONT, OR	2, 3, 18, 19, 21, 25, 49, NM	Mensch, Rind, Schaf

Fortsetzung Tabelle 9.1.3-5: Serovarenverteilung unter Berücksichtigung der nachgewiesenen Intiminvarianten

tRNA	Intimintype	O-Antigen	H-Antigen	Habitat
	η	O2 , O12, O125	19, 49, NM	Mensch
<i>selC, pheV</i>	γ, θ	O1 , O2 , O3, O5 , O8 , O20 , O25, O28 , O33 , O35, O51 , O55 , O70, O76, O76.H7Ont , O84 , O86 , O98 , O101, O103 , O105 , O108 , O111 , O116, O118 , O119 , O125 , O127 , O128 , O145 , O153 , O153/21, O156 , O156 related, O157 , O172, ONT, OR	1 , 2 , 6, 7 , 8, 11, 11/21, 19, 21, 25, 28, 31, 33, 34, 35, 37, 40, 45, 51, NM , NT	Büffel, Mensch, Rind, Schaf
ND	ι	O98, o145, ONT, OR	2 , 4 , 8 , NT	Mensch, Schaf
ND	κ	O37, O49 , O86 , O107 , O118 , O157 , ONT, OR	5 , 11 , 42, NM	Mensch, Schaf
ND	ξ	ONT	NT	Rind
ND	λ	O145, O2/74	NM	Mensch, Rind
ND	μ	OR	NM	Schaf
ND	ν	O2-related	19	Schaf
<i>selC, pheV</i>	ζ	O5 , O8 , O28 , O49 , O84 , O85 , O91 , O92, O98, O104 , O105 , O108 , O111 .	1 , 2 , 4 , 9 , 11 , 21 , 25 , 26 , 31 , 34 , NM	Mensch, Rind, Schaf

rot unterlegt: Erstzuordnung durch die eigenen Arbeiten

rote Schrift: Erstzuordnungen durch andere Autoren, bestätigt in den eigenen Arbeiten

hellblaue Schrift: LEE verankert in *selC*

fett: O- bzw. H-Antigen mehrfach vorkommend

10 LITERATURVERZEICHNIS

- ABBOTT, S. L., J. O'CONNOR, T. ROBIN, B. L. ZIMMER und J. M. JANDA (2003)
Biochemical properties of a newly described *Escherichia* species, *Escherichia albertii*. J. Clin. Microbiol. 41, 4852-4854
- ABDUCH FABREGA, V. L., A. J. PIANTINO FERREIRA, F. REIS DA SILVA PATRICIO, C. BRINKLEY und I. C. SCALETSKY (2002)
Cell-detaching *Escherichia coli* (CDEC) strains from children with diarrhea: identification of a protein with toxigenic activity. FEMS Microbiol. Lett. 217, 191-197
- ABDULMAWJOOD, A. und M. BÜLTE (2001)
Snail Species Identification by RFLP-PCR and Designing of Species-Specific Oligonucleotide Primers. J. Food Sci. 66, 1287-1293
- ABE, A., M. DE GRADO, R. A. PFUETZNER, C. SANCHEZ-SANMARTIN, R. DEVINNEY, J. L. PUENTE, N. C. STRYNADKA und B. B. FINLAY (1999)
Enteropathogenic *Escherichia coli* translocated intimin receptor, Tir, requires a specific chaperone for stable secretion. Mol. Microbiol. 33, 1162-1175
- ABE, K., S. YAMAMOTO und K. SHINAGAWA (2002)
Economic impact of an *Escherichia coli* O157:H7 outbreak in Japan. J. Food Prot. 65, 66-72
- ADU-BOBIE, J., G. FRANKEL, C. BAIN, A. G. GONCALVES, L. R. TRABULSI, G. DOUCE, S. KNOTTON und G. DOUGAN (1998a)
Detection of intimins α , β , γ and δ , four intimin derivatives expressed by attaching and effacing microbial pathogens. J. Clin. Microbiol. 36, 662-668
- ADU-BOBIE, J., L. R. TRABULSI, M. M. CARNEIRO-SAMPAIO, G. DOUGAN und G. FRANKEL (1998b)
Identification of Immunodominant Regions within the C-Terminal Cell Binding Domain of Intimin α and Intimin β from Enteropathogenic *Escherichia coli*. Infect. Immun. 66, 5643-5649
- AGIN, T. S. und M. K. WOLF (1997)
Identification of a family of intimins common to *Escherichia coli* causing attaching-effacing lesions in rabbits, humans, and swine. Infect. Immun. 65, 320-326

- AGIN, T. S., J. R. CANTEY, E. C. BOEDECKER und M. K. WOLF (1996)
Characterization of the *eaeA* gene from rabbit enteropathogenic *Escherichia coli* strain RDEC-1 and comparison to other *eaeA* genes from bacteria that cause attaching–effacing lesions. *FEMS Microbiol. Lett.* 144, 249-258
- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS und D. J. LIPMAN (1990)
Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410
- BACKELJAU, T., L. DEBRUYN, H. DEWOLF, K. JORDAENS, S. VAN DONGEN und B. WINNEPENNINCKX (1996)
Multiple UPGMA and neighbor-joining trees and the performance of some computer packages. *Mol. Biol. Evol.* 13, 309-313
- BANDELT, H. J. und A. W. M. DRESS (1992)
Split decomposition: a new and useful approach to phylogenetic analysis of distance data. *Mol. Phylogenet. Evol.* 1, 242-252
- BARTON, G. J. (2005)
Creation and Analysis of Protein Multiple Sequence Alignments, In: BAXEVANIS, A. D. UND B. F. F. OUELLETTE (Hrsg.) *Bioinformatics: A practical Guide to the Analysis of Genes and Proteins*, 3. Auflage, John Wiley & Sons, New Jersey, USA, pp. 325-340
- BATCHELOR, M., S. PRASANAN, S. DANIELL, S. REECE, I. CONNERTON, G. BLOOMBERG, G. DOUGAN, G. FRANKEL und S. MATTHEWS (2000)
Structural basis for recognition of the translocated intimin receptor (Tir) by intimin from enteropathogenic *Escherichia coli*. *EMBO J.* 19, 2452-2464
- BEEBAKHEE, G., M. LOUIE, J. DE AZAVEDO und J. BRUNTON (1992)
Cloning and nucleotide sequence of the *eae* gene homologue from enterohemorrhagic *Escherichia coli* serotype O157:H7. *FEMS Microbiol. Lett.* 70, 63-68
- BERTIN, Y., K. BOUKHORS, N. PRADEL, V. LIVRELLI und C. MARTIN (2001)
Stx2 subtyping of Shiga toxin-producing *Escherichia coli* isolated from cattle in France: Detection of a new Stx2 subtype and correlation with additional virulence factors. *J. Clin. Microbiol.* 39, 3060-3065

BETTELHEIM, K. A. (2005)

Serotypes of vero toxin-producing *Escherichia coli* reported in the literature, apart from those belonging to serogroup O157

<http://www.microbionet.com.au/frames/feature/vtec/brief01.html>, Stand: 03.12.2005

BEUTIN, L., D. GEIER, H. STEINBRÜCK, S. ZIMMERMANN und F. SCHEUTZ (1993)

Prevalence and some properties of verotoxin (Shiga-like toxin)-producing *Escherichia coli* in seven different species of healthy domestic animals. J. Clin. Microbiol. 31, 2483–2488

BEUTIN, L., S. ZIMMERMANN und K. GEIER (1998)

Human Infections with Shiga Toxin-Producing *Escherichia coli* Other Than Serogroup O157 in Germany. Emerg. Infect. Dis. 4, 635-639

BEUTIN, L., O. MARCHES, K. A. BETTELHEIM, K. GLEIER, S. ZIMMERMANN, H. SCHMIDT und E. OSWALD (2003)

HEp-2 cell adherence, actin aggregation, and intimin types of attaching and effacing *Escherichia coli* strains isolated from healthy infants in Germany and Australia. Infect. Immun. 71, 3995-4002

BEUTIN, L., G. KRAUSE, S. ZIMMERMANN, S. KAULFUSS und K. GLEIER (2004)

Characterization of Shiga toxin-producing *Escherichia coli* strains isolated from human patients in Germany over a 3-year period. J. Clin. Microbiol. 42, 1099-1108

BEUTIN, L., S. KAULFUSS, S. HEROLD, E. OSWALD und H. SCHMIDT (2005)

Genetic analysis of enteropathogenic and enterohemorrhagic *Escherichia coli* serogroup O103 strains by molecular typing of virulence and housekeeping genes and pulsed-field gel electrophoresis. J. Clin. Microbiol. 43, 1552-1563

BIELASZEWSKA, M., M. FELL, L. GREUNE, R. PRAGER, A. FRUTH, H. TSCHAPE, M. A. SCHMIDT und H. KARCH (2004)

Characterization of cytolethal distending toxin genes and expression in shiga toxin-producing *Escherichia coli* strains of non-O157 serogroups. Infect. Immun. 72, 1812-1816.

BIELASZEWSKA, M., B. SINHA, T. KUCZIUS und H. KARCH (2005)

Cytolethal Distending Toxin from Shiga Toxin-Producing *Escherichia coli* O157 Causes Irreversible G₂/M Arrest, Inhibition of Proliferation, and Death of Human Endothelial Cells. Infect. Immun. 73, 552-562

BILGE, S. S., C. R. CLAUSEN, W. LAU und S. L. MOSELEY (1989)

Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. *J. Bacteriol.* 171, 4281-4289

BLANCO, J., E. A. GONZALES, P. ESPINOSA, M. BLANCO, J. I. GARABAL und M. P. ALONSO (1992)

Enterotoxigenic and necrotizing *Escherichia coli* in human diarrhoea in Spain. *Eur. J. Epidemiol.* 8, 548-552

BLANCO, M., J. E. BLANCO, A. MORA, G. DAHBI, M. P. ALONSO, E. A. GONZALEZ, M. I. BERNARDEZ und J. BLANCO (2004a)

Serotypes, virulence genes, and intimin types of Shiga toxin (verotoxin)-producing *Escherichia coli* isolates from cattle in Spain and identification of a new intimin variant gene (*eae-xi*). *J. Clin. Microbiol.* 42, 645-651

BLANCO, M., N. L. PADOLA, A. KRUGER, M. E. SANZ, J. E. BLANCO, E. A. GONZALEZ, G. DAHBI, A. MORA, M. I. BERNARDEZ, A. L. ETCHEVERRIA, G. H. ARROYO, P. M. LUCCHESI, A. E. PARMA und J. BLANCO (2004b)

Virulence genes and intimin types of Shiga-toxin-producing *Escherichia coli* isolated from cattle and beef products in Argentina. *Int. Microbiol.* 7, 269-76

BLANCO, M., S. SCHUMACHER, T. TASARA, C. ZWEIFEL, J. E. BLANCO, G. DAHBI, J. BLANCO und R. STEPHAN (2005)

Serotypes, intimin variants and other virulence factors of *eae* positive *Escherichia coli* strains isolated from healthy cattle in Switzerland. Identification of a new intimin variant gene (*eae-eta2*). *BMC Microbiol.* 5, 23-33

BOERLIN, P., S. A. MCEWEN, F. BOERLIN-PETZOLD, J. B. WILSON, R. P. JOHNSON und C. L. GYLES (1999)

Associations between Virulence factors of Shiga Toxin-Producing *Escherichia coli* and Disease in Humans. *J. Clin. Microbiol.* 37, 497-503

BRIFFEUIL, P., G. BAUDOIX, C. LAMBERT, X. DE BOLLE, C. VINALS, E. FEYTMANS und E. DEPIEREUX (1998)

Comparative analysis of seven multiple protein sequence alignment servers: clues to enhance reliability of predictions. *Bioinf.* 14, 357-366

BRINKMANN, F. S. L. (2005)

Phylogentic Analysis, In: BAXEVANIS, A. D. UND B. F. F. OUELLETTE (Hrsg.) *Bioinformatics: A practical Guide to the Analysis of Genes and Proteins.*, 3. Auflage, John Wiley & Sons, New Jersey, USA, pp. 365-392

BRYANT, D. und V. MOULTON (2004)

Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol. Biol. Evol.* 21, 255-265

- BROWN, C. J., E. C. GARNER, A. K. DUNKER und P. JOYCE (2001)
The Power to Detect Recombination Using the Coalescent. *Mol. Biol. Evol.* 18, 1421–1424
- BURGE, C. und S. KARLIN (1997)
Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268, 78-94
- BÜLTE, M. (2000)
Versuch einer lebensmittelrechtlichen Bewertung von Verotoxin-bildenden *E. coli* (VTEC) und enterohämorrhagischen *E. coli* (EHEC). 53. Arbeitstagung des Arbeitskreises Lebensmittelhygienischer Tierärztlicher Sachverständiger (ALTS), Berlin, 14.-15.06.2000
- BÜLTE, M. (2001)
Nachweis und Charakterisierung von Verotoxin-bildenden *Escherichia coli*-Stämmen (VTEC) aus unterschiedlichen Habitaten. *Berl. Münch. Tierärztl. Wschr.* 114, 473-477
- BÜLTE, M. und S. HECKÖTTER (1997)
Vorkommen und Bedeutung von O157 und anderen verotoxinbildenden *E. coli* bei Tieren und in Lebensmitteln. *Mitt. Gebiete Lebensm. Hyg.* 88, 665-680
- BÜRK, C., R. DIETRICH, G. ACAR, M. MORAVEK, M. BÜLTE und E. MÄRTLBAUR (2003)
Identification and characterization of a new variant of Shiga toxin 1 in *Escherichia coli* ONT:H19 of bovine origin. *J. Clin. Microbiol.* 41, 2106-2112
- CALDERWOOD, S. B., D. W. K. ACHESON, G. T. KEUSCH, T. J. BARRETT, P. M. GRIFFIN, N. A. STROCKBINE, B. SWAMINATHAN, J. B. KAPER, M. M. LEVINE, B. S. KAPLAN, H. KARCH, A. D. O'BRIEN, T. G. OBRIG, Y. TAKEDA, P. I. TARR und I. K. WACHSMUTH (1996)
Proposed new nomenclature for SLT(VT) family. *ASM News* 62, 118–119
- CAMPELLONE, K. G. und J. M. LEONG (2003)
Tails of two Tirs: actin pedestal formation by enteropathogenic *E. coli* and enterohemorrhagic *E. coli* O157:H7. *Curr. Opin. Microbiol.* 6, 82-90
- CANTARELLI, V. V. , A. TAKAHASHI, Y. AKEDA, K. NAGAYAMA und T. HONDA (2000a)
Interaction of enteropathogenic or enterohemorrhagic *Escherichia coli* with HeLa cells results in translocation of cortactin to the bacterial adherence site. *Infect. Immun.* 68, 382-386

CANTARELLI, W., A. TAKAHASHI, I. YANAGIHARA, Y. AKEDA, K. IMURA, T. KODAMA, G. KONO, Y. SATO und T. HONDA (2000B)

Talin, a host cell protein, interacts directly with the translocated intimin receptor, Tir, of enteropathogenic *Escherichia coli*, and is essential for pedestal formation. *Cell Microbiol.* 3, 745-751

CANTEY, J. R. und D. S. HOSTERMAN (1979)

Characterization of colonization of the rabbit gastrointestinal tract by *Escherichia coli* RDEC-1. *Infect. Immun.* 26, 1099-1103

CAPRIOLI, A., V. FALBO, L. G. RODA, F. M. RUGGERI und C. ZONA (1983)

Partial purification and characterization of an *Escherichia coli* toxic factor that induces morphological cell alterations. *Infect. Immun.* 39, 1300-1306

CID, D., J. A. RUIZ-SANTA-QUITERIA, I. MARIN, R. SANZ, J. A. ORDEN, R. AMILS und R. DE LA FUENTE (2001)

Association between intimin (*eae*) and EspB gene subtypes in attaching and effacing *Escherichia coli* strains isolated from diarrhoeagenic lambs and goat kids. *Microbiology* 147, 2341-2353

CLARK, C. G., S. T. JOHNSON, R. H. EASY, J. L. CAMPBELL und J. G. RODGERS (2002)

PCR for detection of *cdt*-III and the relative frequencies of cytolethal distending toxin variant-producing *Escherichia coli* isolates from humans and cattle. *J. Clin. Microbiol.* 40, 2671-2674

CLARKE, S. C., R. D. HAIGH, P. P. E. FREESTONE und P. H. WILLIAMS (2004)

Virulence of Enteropathogenic *Escherichia coli*, a Global Pathogen. *Clin. Microbiol. Rev.* 16, 365-378

COPE, L. D., S. LUMBLEY, J. L. LATIMER, J. KLESNEY-TAIT, M. K. STEVENS, L. S. JOHNSON, M. PURVEN, R. S. MUNSON, JR., T. LAGERGARD, J. D. RADOLF und E. J. HANSEN (1997)

A diffusible cytotoxin of *Haemophilus ducreyi*. *Proc. Natl. Acad. Sci.* 94, 4056-4061

CORTES-BRATTI, X., T. FRISAN und M. THELESTAM (2001)

The cytolethal distending toxins induce DNA damage and cell cycle arrest. *Toxicon* 39, 1729-1736

DAYHOFF, M. O., R. M. SCHWARTZ und B. C. ORCUTT (1978)

A model of evolutionary change in proteins. In: DAYHOFF, M. O (Hrsg.) *Atlas of Protein Sequence and Structure*, 1. Aufl., National Biomedical Research Foundation, Washington, USA, pp. 345-362

- DE RYCKE, J. und E. OSWALD (2001)
Cytolethal distending toxin (CDT): a bacterial weapon to control host cell proliferation? FEMS Microbiol. Lett. 203, 141-148
- DE RYCKE, J., A. MILON und E. OSWALD (1999)
Necrotoxic *Escherichia coli* (ntec): two emerging categories of human and animal pathogens. Vet. Res. 30, 221-233
- DE VINNEY, R., J. L. PUENTE, A. GAUTHIER, D. GOOSNEY und B. B. FINLAY (2001)
Enterohaemorrhagic and enteropathogenic *Escherichia coli* use a different Tir-based mechanism for pedestral formation. Mol. Microbiol. 41, 1445-1458
- DEAN-NYSTROM, E. A., L. J. GANSHEROFF, M. MILLS, H. W. MOON und A. D. O'BRIEN (2002)
Vaccination of pregnant dams with intimin (_{O157}) protects suckling piglets from *Escherichia coli* O157:H7 infection. Infect. Immun. 70, 2414-2418
- DEIBEL, C., S. KRAMER, T. CHAKRABORTY und F. EBEL (1998)
EspE, a novel secreted protein of attaching and effacing bacteria, is directly translocated into infected host cells, where it appears as a tyrosine-phosphorylated 90 kDa protein. Mol. Microbiol. 28, 463-474
- DO, C. B., M. S. MAHABHASHYAM, M. BRUDNO und S. BATZOGLOU (2005)
ProbCons: Probabilistic consistency-based multiple sequence alignment. Genome Res. 15, 330-340
- DONNENBERG, M. S. und J. B. KAPER (1991)
Construction of an *eae* deletion mutant of enteropathogenic *Escherichia coli* by using a positive-selection suicide vector. Infect Immun. 59, 4310-4317
- DONNENBERG, M. S. und J. B. KAPER (1992)
Enteropathogenic *Escherichia coli*. Infect Immun. 60, 3953-3961
- DONNENBERG, M. S. und T. S. WHITTAM (2001)
Pathogenesis and evolution of virulence in enteropathogenic and enterohemorrhagic *Escherichia coli*. J. Clin. Invest. 107, 539-548
- DONNENBERG, M. S., J. U. YU und J. B. KAPER (1993a)
A second chromosomal gene necessary for intimate attachment of enteropathogenic *Escherichia coli* to epithelial cells. J. Bacteriol. 175, 4670-4680
- DONNENBERG, M. S., C. O. TACKET, S. P. JAMES, G. LOSONSKY, J. P. NATARO, S. S. WASSERMAN, J. B. KAPER und M. M. LEVINE (1993b)
Role of the *eaeA* gene in experimental enteropathogenic *Escherichia coli* infection. J. Clin. Invest. 92, 1412-1417

DONNENBERG, M. S., S. TZIPORI, M. L. MCKEE, A. D. O'BRIEN, J. ALROY und J. B. KAPER (1993c)

The role of the *eae* gene of enterohemorrhagic *Escherichia coli* in intimate attachment in vitro and in a porcine model. *J. Clin. Invest.* 92, 1418-1424

DROUIN, G., F. PRAT, M. ELL und G. D. P. CLARK (1999)

Detecting and characterizing gene conversions between multigene family members. *Mol. Biol. Evol.* 16, 1369-1390

DUBREUIL, J. D. (1997)

Escherichia coli STb enterotoxin. *Microbiology* 143, 1783-1795

DYTOC, M., B. GOLD, M. LOUIE, M. HUESCA, L. FEDORKO, S. CROWE, C. LINGWOOD, J. BRUNTON und P. SHERMAN (1993)

Comparison of *Helicobacter pylori* and attaching-effacing *Escherichia coli* adhesion to eukaryotic cells. *Infect. Immun.* 61, 448-456.

EDGAR, R. C. (2004)

MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792-1797

EDWARDS, A. W. F. und L. L. CAVALLI-SFORZA (1963)

The reconstruction of evolution. *Heredity* 18, 553

EFRON, B. (1979)

Bootstrap methods: another look at the jackknife. *The Annals of Statistics* 7, 1-26

EFRON, B., E. HALLORAN und S. HOLMES (1996)

Bootstrap confidence levels for phylogenetic trees. *Proc. Natl. Acad. Sci.* 93, 13429-13434

ELLIOTT, S. J., L. A. WAINWRIGHT, T. K. MCDANIEL, K. G. JARVIS, Y. DENG, L.-C. LAI, B. P. MCNAMARA, M. S. DONNENBERG und J. B. KAPER (1998)

The complete sequence of the locus of enterocyte effacement (LEE) from enteropathogenic *E. coli* E2348/69. *Mol. Microbiol.* 28, 1-4

ELLIOTT, E. J., R. M. ROBINS-BROWNE, E. V. O'LOUGHLIN, V. BENNETT-WOOD, J. BOURKE, P. HENNING, G. G. HOGG, J. KNIGHT, H. POWELL, D. REDMOND und CONTRIBUTORS TO THE AUSTRALIAN PAEDIATRIC SURVEILLANCE UNIT (2001)

Nationwide study of haemolytic uraemic syndrome: clinical, microbiological, and epidemiological features. *Arch. Dis. Child.* 85, 125-131

ENDO, T., K. IKEO und T. GOJOBORI (1996)

Large-scale search for genes on which positive selection may operate. *Mol. Biol. Evol.* 13, 685-690

- EWERS, C, T. JANSSEN und L. H. WIELER (2003)
Avian pathogenic *Escherichia coli* (APEC). Berl. Münch. Tierärztl. Wschr. 116, 381-95
- FALBO, V., M. FAMIGLIETTI und A. CAPRIOLI (1992)
Gene block encoding production of cytotoxic necrotizing factor 1 and hemolysin in *Escherichia coli* isolates from extraintestinal infections. Infect. Immun. 60, 2182-2187
- FANG, G. D., A. A. M. LIMA, C. V. MARTINS, J. P. NATARO und R. L. GUERRANT (1995)
Etiology and epidemiology of persistent diarrhea in northeastern Brazil: a hospitalbased, prospective, case-control study. J. Pediatr. Gastroenterol. Nutr. 21, 137-144
- FELSENSTEIN, J. (1978)
Cases in which parsimony or compatibility methods will be positively misleading. Syst. Zool. 27, 401-410
- FELSENSTEIN, J. (1981)
Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17, 368-376
- FELSENSTEIN, J. (1985)
Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783-791
- FELSENSTEIN, J. (1997)
An alternative least squares approach to inferring phylogenies from pairwise distances. Syst. Biol. 46, S101-S111
- FENG, D. F. UND R. F. DOOLITTLE (1987)
Progressive sequence alignment as a prerequisite to correct phylogentic trees. J. Mol. Evol. 25, 351-360
- FISHER, R. A. (1973)
Statistical Methods and Scientific Inference , 3. Auflage. Oxford University Press, Oxford, England
- FITCH, W. M. (1967)
Evidence suggesting a non-random character to nucleotide replacements in naturally occurring mutations. J. Mol. Biol. 26, 499-507
- FITCH, W. M. (1997)
Networks and viral evolution. J. Mol. Evol. 44, S65–S75

FITZHENRY, R. J., D. J. PICKARD, E. L. HARTLAND, S. REECE, G. DOUGAN, A. D. PHILLIPS und G. FRANKEL (2002a)

Intimin type influences the site of human intestinal mucosal colonisation by enterohaemorrhagic *Escherichia coli* O157:H7. *Gut* 50, 180-5

FITZHENRY, R. J., S. REECE, L. R. TRABULSI, R. HEUSCHKEL, S. MURCH, M. THOMSON, G. FRANKEL und A. D. PHILLIPS (2002b)

Tissue tropism of enteropathogenic *Escherichia coli* strains belonging to the O55 serogroup. *Infect. Immun.* 70, 4362-4368

FITZHENRY, R. J., M. P. STEVENS, C. JENKINS, T. S. WALLIS, R. HEUSCHKEL, S. MURCH, M. THOMSON, G. FRANKEL und A. D. PHILLIPS (2003)

Human intestinal tissue tropism of intimin epsilon O103 *Escherichia coli*. *FEMS Microbiol. Lett.* 218, 311-316

FIVAZ, M. und F. G. VAN DER GOOT (1999)

The tip of a molecular syringe. *Trends Microbiol.* 7, 341-343

FRANKEL, G., D. C. A. CANDY, P. EVEREST und G. DOUGAN (1994)

Characterization of the C-terminal domains of intimin like proteins of enteropathogenic and enterohemorrhagic *Escherichia coli*, *Citrobacter freundii*, and *Hafnia alvei*. *Infect Immun.* 62, 1835-1842

FRANKEL, G., D. C. CANDY, E. FABIANI, J. ADU-BOBIE, S. GIL, M. NOVAKOVA, A. D. PHILLIPS und G. DOUGAN (1995)

Molecular characterization of a carboxy-terminal eukaryotic-cell-binding domain of intimin from enteropathogenic *Escherichia coli*. *Infect. Immun.* 63, 4323-4328

FRANKEL, G., A. D. PHILIPS, M. NOVAKOVA, M. BATCHELOR, S. HICKS und G. DOUGAN (1998)

Generation of *Escherichia coli* intimin-derivatives with differing biological activities using site-directed mutagenesis of the intimin C-terminus domain. *Mol. Microbiol.* 29, 559-570

FREEMAN, N. L., D. V. ZURAWSKI, P. CHOWRASHI, J. C. AYOUB, L. HUANG, B. MITTAL, J. M. SANGER und J. W. SANGER (2000)

Interaction of the enteropathogenic *Escherichia coli* protein, translocated intimin receptor (Tir), with focal adhesion proteins. *Cell. Motil. Cytoskeleton.* 47, 307-318

FRIEDRICH, A. W., M. BIELASZEWSKA, W. L. ZHANG, M. PULZ, T. KUCZIUS, A. AMMON und H. KARCH (2002)

Escherichia coli harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J. Infect. Dis.* 185, 74-84

FRIEDRICH, A. W., J. BORELL, M. BIELASZEWSKA, A. FRUTH, H. TSCHAPE und H. KARCH (2003)

Shiga toxin 1c-producing *Escherichia coli* strains: phenotypic and genetic characterization and association with human disease. *J. Clin. Microbiol.* 41, 2448-2453

GALLIEN, P., H. KARCH, C. MUCH, H. STEINRUCK, S. LEHMANN, M. TIMM, H. RICHTER, K. W. PERLBERG und D. PROTZ (2000)

Subtyped *eae*-genes in Shigatoxin-producing *Escherichia coli* (STEC) - Occurrence in raw or undercooked food samples and comparison of isolates from faecal samples and stool samples. *Fleischwirtschaft* 80, 84-89

GANNON, V. P. J., C. TEERLING, S. A. MASRI und C. L. GYLES (1990)

Molecular cloning and nucleotide sequence of another variant of the *Escherichia coli* Shigalike toxin II family. *J. Gen. Microbiol.* 136, 1125-1135

GANNON, V. P. J., M. RASHED, R. K. KING und E. J. G. THOMAS (1993)

Detection and characterization of the *eae* gene of Shiga-like toxin-producing *Escherichia coli* using polymerase chain reaction. *J. Clin. Microbiol.* 31, 1268-1274

GANSHERHOFF, L. J., M. R. WACHTEL und A. D. O'BRIEN (1999)

Decreased adherence of enterohemorrhagic *Escherichia coli* to HEp-2 cells in the presence of antibodies that recognize the C-terminal region of intimin. *Infect. Immun.* 67, 6409-6417

GAREIS, M., R. PICHNER, N. BREY und H. STEINRÜCK (2000)

Nachweis Verotoxin-bildender *E. coli* (VTEC) bei gesunden Mitarbeitern eines Fleisch verarbeitenden Betriebes. *Bundesgesundhbl.* 43, 781-787

GASCUEL, O. (1997)

BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol Biol Evol.* 14, 685-695

GASCUEL, O., BRYANT, D. und F. DENIS (2001)

Strengths and limitations of the minimum evolution principle. *Syst. Biol.* 50, 621-627

GASSER, C., E. GAUTHIER, A. STECK, R. E. SIEBENMANN und R. OECHSLIN (1955)

Hämolytisch-urämisches Syndrom: bilaterale Nierenrindennekrosen bei akuten erworbenen hämolytischen Anämien. *Schweiz. Med. Wochenschr.* 85, 905-909

GERBER, A., H. KARCH, F. ALLERBERGER, H. M. VERWEYEN und L. B. ZIMMERHACKL (2002)

Clinical course and the role of Shiga toxin-producing *Escherichia coli* infection on the hemolytic-uremic syndrome in pediatric patients, 1997-2000, in Germany and Austria: a prospective study. *J. Infect. Dis.* 186, 493-500

- GEUE, L., T. SELHORST, C. SCHNICK, B. MINTEL und F. J. CONRATHS (2004)
Untersuchungen zum humanen Gefährdungspotenzial von potenziellen enterohämorrhagischen *E. coli* (EHEC) isoliert in deutschen Rinderbeständen. EHEC-Workshop, 22.-24. Juli 2004, Wildbad Kreuth
- GIBBS, M. J., J. S. ARMSTRONG und A. J. GIBBS (2000)
Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinf.* 16, 573–582
- GOJOBORI, T., W. H. LI und D. GRAUR (1982)
Patterns of nucleotide substitutions in pseudogenes and functional genes. *J. Mol. Evol.* 18, 360-369
- GOLL, M. (2005)
Nachweis und DNA-Fingerprinting von *Escherichia coli* O157-Stämmen bei Pferden. Vet. Med. Diss., Justus-Liebig-Universität Gießen
- GOOSNEY, D. L., R. DEVINNEY, R. A. PFUETZNER, E. A. FREY, N. C. STRYNADKA und B. B. FINLAY (2000)
Enteropathogenic *E. coli* translocated intimin receptor, Tir, interacts directly with alpha-actinin. *Curr. Biol.* 10, 735-738
- GOTOH, O. (1996)
Significant improvement in accuracy of multiple protein sequence alignments by iterative refinement as assessed by reference to structural alignments. *J. Mol. Biol.* 264, 823-838-592
- GREEN, B. A., R. J. NEILL, W. T. RUYECHAN und R. K. HOLMES (1983)
Evidence that a new enterotoxin of *Escherichia coli* which activates adenylate cyclase in eucaryotic target cells is not plasmid mediated. *Infect. Immun.* 41, 383-390
- GUNZBURG, S. T., B. J. CHANG, S. J. ELLIOTT, V. BURKE und M. GRACEY (1993)
Diffuse and enteroaggregative patterns of adherence of enteric *Escherichia coli* isolated from aboriginal children from the Kimberley region of Western Australia. *J. Infect. Dis.* 167, 755-758
- HALL, T. A. (1999)
BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41, 95-98.
- HAMBURGER, Z. A., M. S. BROWN, R. R. ISBERG und P. J. BJORKMAN (1999)
Crystal structure of invasins: a bacterial integrin-binding protein. *Science* 286, 291-295
- HAHN, H., D. FALKE, S. H. E. KAUFMANN und U. ULLMANN (1999)
Medizinische Mikrobiologie und Infektiologie. Springer-Verlag, Berlin-Heidelberg-New York, 3. Auflage, p. 251

- HARTLAND, E. L., M. BATCHELOR, R. M. DELAHAY, C. HALE, S. MATTHEWS, G. DOUGAN, S. KNUSTON, I. CONNERTON und G. FRANKEL (1999)
Binding of intimin from enteropathogenic *Escherichia coli* to Tir and to host cells. *Mol. Microbiol.* 32, 151-158
- HARTLAND, E. L., V. HUTER, L. M. HIGGINS, N. S. GONCALVES, G. DOUGAN, A. D. PHILLIPS, T. T. MACDONALD und G. FRANKEL (2000)
Expression of intimin gamma from enterohemorrhagic *Escherichia coli* in *Citrobacter rodentium*. *Infect. Immun.* 68, 4637-4646
- HASEGAWA, M., H. KISHINO und T. YANO (1985)
Dating of the human ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160-174
- HOLT, J. G., N. R. KRIEG, P. H. A. SNEATH, J. T. STALEY und S. T. WILLIAMS (1994)
Bergey's Manual of Determinative Bacteriology. Williams & Wilkins Verlag, 9. Auflage, pp.179-180
- HUGHES, A. L (1995)
Origin and evolution of HLA class I pseudogenes. *Mol. Biol. Evol.* 12, 247-258
- HUGHES, A. L und R. FRIEDMAN (2000)
Evolutionary Diversification of Protein-Coding Genes of Hantaviruses. *Mol. Biol. Evol.* 17, 1558-1568
- HUGHES, A. L. und R. FRIEDMAN (2005)
Variation in the pattern of synonymous and nonsynonymous difference between two fungal genomes. *Mol. Biol. Evol.* 22, 1320-1324
- HUGHES, A. L. und M. YEAGER (1997)
Molecular evolution of the vertebrate immune system. *Bioessays* 19, 777-786
- HUGHES, A. L. und F. VERRA (1998)
Ancient polymorphism and the hypothesis of a recent bottleneck in the malaria parasite *Plasmodium falciparum*. *Genetics* 150, 511-513
- HUGHES, A. L., T. OTA und M. NEI (1990)
Positive Darwinian Selection Promotes Charge Profile Diversity in the Antigen-binding Cleft of Class I Major-Histocompatibility-Complex Molecules. *Mol. Biol. Evol.* 7, 515-524
- HUSON, D. H. (1998)
SplitsTree: Analyzing and visualising evolutionary data. *Bioinf.* 14, 68-73

HUSON, D. H. und D. BRYANT (2006)

Application of Phylogenetic Networks in Evolutionary Studies. *Mol. Biol. Evol.* 23, 254-267, Epub 2005

HUYS, G., M. CNOCKAERT, J. M. JANDA und J. SWINGS (2003)

Escherichia albertii sp. nov., a diarrhoeogenic species isolated from stool specimens of Bangladeshi children. *Int. J. Syst. Evol. Microbiol.* 53, 807-810

HYMA, K. E., D. W. LACHER, A. M. NELSON, A. C. BUMBAUGH, J. M. JANDA, N. A. STROCKBINE V. B. YOUNG und T. S. WHITTAM (2005)

Evolutionary genetics of a new pathogenic *Escherichia* species: *Escherichia albertii* and related *Shigella boydii* strains. *J. Bacteriol.* 187, 619-628

ISBERG, R. R., D. L. VOORHIS und S. FALKOW (1987)

Identification of invasins: a protein that allows enteric bacteria to penetrate cultured mammalian cells. *Cell* 28, 769-778.

ITO, H., A. TERAJ, H. KURAZONO, Y. TAKEDA und M. NISHIBUCHI (1990)

Cloning and nucleotide sequencing of Verotoxin 2 variant genes from *Escherichia coli* O91:H21 isolated from a patient with the hemolytic uremic syndrome. *Microb. Pathog.* 8, 47-60

ITOH Y, I. NAGANO, M. KUNISHIMA und T. EZAKI (1997)

Laboratory investigation of enteroaggregative *Escherichia coli* O untypeable: H10 associated with a massive outbreak of gastrointestinal illness. *J. Clin. Microbiol.* 35, 2546-2550.

JANKA, A., M. BIELASZEWSKA, U. DOBRINDT, L. GREUNE, M. A. SCHMIDT und H. KARCH (2003)

Cytolethal distending toxin gene cluster in enterohemorrhagic *Escherichia coli* O157:H⁻ and O157:H7: Characterization and evolutionary considerations. *Infect. Immun.* 71, 3634-3638

JELACIC, J. K., T. DAMROW, G. S. CHEN, S. JELACIC, M. BIELASZEWSKA, M. CIOL, H. M. CARVALHO, A. R. MELTON-CELSA, A. D. O'BRIEN und P. I. TARR (2003)

Shiga toxin-producing *Escherichia coli* in Montana: bacterial genotypes and clinical profiles. *J. Infect. Dis.* 188, 719-729

JENKINS, C., A. J. LAWSON, T. CHEASTY, G. A. WILLSHAW, P. WRIGHT, G. DOUGAN, G. FRANKEL und H. R. SMITH (2003)

Subtyping intimin genes from enteropathogenic *Escherichia coli* associated with outbreaks and sporadic cases in the United Kingdom and Eire. *Mol. Cell. Probes* 17, 149-156

- JERSE, A. E., J. YU, B. D. TALL und J. B. KAPER (1990)
A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. Proc. Natl. Acad. Sci. 87, 7839-7843
- JERSE, A. E. und J. P. KAPER (1991)
The *eae* gene of enteropathogenic *Escherichia coli* encodes a 94-kilodalton membrane protein the expression of which is influenced by the EAF plasmid. Infect. Immun. 59, 4302-4309
- JERSE, A. E., K. G. GICQUELAIS, und J. B. KAPER (1991)
Plasmid and chromosomal elements involved in the pathogenesis of attaching and effacing *Escherichia coli*. Infect. Immun. 59, 3869-3875
- JORES J, L. RUMER , S. KIESSLING, J. B. KAPER und L. H. WIELER (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, 75-79
- JORES, J., K. ZEHRKE, J. EICHBERG, L. RUMER und L. H. WIELER (2003)
Description of a novel intimin variant (type zeta) in the bovine O84:NM verotoxin-producing *Escherichia coli* strain 537/89 and the diagnostic value of intimin typing. Exp. Biol. Med. 228, 370-376
- JORES, J., L. RUMER und L. H. WIELER (2004)
Impact of the locus of enterocyte effacement pathogenicity island on the evolution of pathogenic *Escherichia coli*. Int. J. Med. Microbiol. 294, 103-113
- JUDGE, N. A., H. S. MASON und A. D. O'BRIEN (2004)
Plant cell-based intimin vaccine given orally to mice primed with intimin reduces time of *Escherichia coli* O157:H7 shedding in feces. Infect. Immun. 72, 168-175
- JUKES, T. H. und C. R. CANTOR (1969)
Evolution of protein molecules. In: H. N. MUNRO (Hrsg.) *Mammalian protein metabolism* Academic Press, New York, USA, , pp. 21-132
- KALMAN, D., O. D. WEINER, D. L. GOOSNEY, J. W. SEDAT, B. B. FINLAY, A. ABO und J. M. BISHOP (1999)
Enteropathogenic *E. coli* acts through WASP and Arp2/3 complex to form actin pedestals. Nat. Cell Biol. 1, 389-391
- KAPER, J. B., J. P. NATARO und H. L. MOBLEY (2004)
Pathogenic *Escherichia coli*. Nat. Rev. Microbiol. 2, 123-140

KARCH, H., H. BOHM, H. SCHMIDT, F. GUNZER, S. ALEKSIC und J. HEESEMANN (1993)

Clonal structure and pathogenicity of Shiga-like toxin-producing, sorbitol-fermenting *Escherichia coli* O157:H⁻. J. Clin. Microbiol. 31, 1200-1205

KARCH, H., P. I. TARR und M. BIELASZEWSKA (2005)

Enterohaemorrhagic *Escherichia coli* in human medicine. Int. J. Med. Microbiol. 295, 405-418

KARMALI, M. A., B. T. STEELE, M. PETRIC und C. LIM (1983)

Sporadic cases of haemolytic-uraemic syndrome associated with faecal cytotoxin and cytotoxin-producing *Escherichia coli* in stools. Lancet i, 619-620

KARMALI, M. A., M. PETRIC, C. LIM, P. C. FLEMING, G. S. ARBUS und H. LIOR (1985)

The association between idiopathic hemolytic uremic syndrome and infection by verotoxin-producing *Escherichia coli*. J. Infect. Dis. 151, 775-782

KAUFFMANN, F. (1966)

The bacteriology of Enterobacteriaceae. Munksgaard Verlag, Kopenhagen, p. 397

KELLY, G., S. PRASANAN, S. DANIELL, K. FLEMING, G. FRANKEL, G. DOUGAN, I. CONNERTON und S. MATTHEWS (1999)

Structure of the cell-adhesion fragment of intimin from enteropathogenic *Escherichia coli*. Nat. Struct. Biol. 6, 313-318.

KENNY, B., R. DE VINNEY, M. STEIN, D. J. REINSCHIED, E. A. FREY und B. B. FINLAY (1997)

Enteropathogenic *E. coli* (EPEC) transfers its receptor for intimate adherence into mammalian cells. Cell 91, 511-520

KIMURA, M. (1980)

A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16, 11-120

KNUDSEN, B., M. M. MIYAMOTO, P. J. LAIPIS und D. N. SILVERMAN (2003)

Using evolutionary rates to investigate protein functional divergence and conservation. A case study of the carbonic anhydrases. Genetics 164, 1261-1269

KNUTTON, S. (2003)

Microscopic methods to study STEC. Analysis of the attaching and effacing process. Methods Mol. Med. 73, 137-149

- KNUTTON, S., D. R. LLOYD und A. S. MCNEISH (1987)
Adhesion of enteropathogenic *Escherichia coli* to human intestinal enterocytes and cultured human intestinal mucosa. *Infect. Immun.* 55, 69-77
- KNUTTON S., T. BALDWIN, P. H. WILLIAMS und A. S. MCNEISH (1989)
Actin accumulation at sites of bacterial adhesion to tissue culture cells: basis of a new diagnostic test for enteropathogenic and enterohemorrhagic *Escherichia coli*. *Infect. Immun.* 57, 1290-1298
- KNUTTON, S., I. ROSENSHINE, M. J. PALLEN, I. NISAN, B. C. NEVES, C. BAIN, C. WOLFF, G. DOUGAN und G. FRANKEL (1998)
A novel *EspA*-associated surface organelle of enteropathogenic *Escherichia coli* involved in protein translocation into epithelial cells. *EMBO J.* 17, 2166-2176
- KOBAYASHI, H., J. SHIMADA, M. NAKAZAWA, T. MOROZUMI, T. POHJANVIRTA, S. PELKONEN und K. YAMAMOTO (2001)
Prevalence and characteristics of Shiga toxin-producing *Escherichia coli* from healthy cattle in Japan. *Appl. Environ. Microbiol.* 67, 484-489
- KOCHER, T. D. und A. C. WILSON (1991)
Sequence evolution of mitochondrial DNA in humans and chimpanzees: Control region and a protein-coding region. In: Osawa, S. und T. Honjo (Hrsg.) *Evolution of life* 1. Auflage, Springer-Verlag, NY, USA, pp. 391-413
- KONOWALCHUK, J., J. I. SPEIRS und S. STAVRIC (1977)
Vero Response to a Cytotoxin of *Escherichia coli*. *Infect. Immun.* 18, 775-779
- KOSAKOWSKY-POND, S. L., S. D. FROST und S. V. MUSE (2005)
HyPhy: hypothesis testing using phylogenies. *Bioinformatics.* 21, 676-679
- KOTETISHVILI, M., O. C. STINE, A. KREGER, J. G. MORRIS JR und A. SULAKVELIDZE (2002)
Multilocus sequence typing for characterization of clinical and environmental *Salmonella* strains. *J. Clin. Microbiol.* 40, 1626-1635
- KRAUSE, G., S. ZIMMERMANN und L. BEUTIN (2005)
Investigation of domestic animals and pets as a reservoir for intimin- (*eae*) gene positive *Escherichia coli* types. *Vet. Microbiol.* 106, 87-95
- KUMAR, H. S., I. KARUNASAGAR, I. KARUNASAGAR, T. TEIZOU, K. SHIMA und S. YAMASAKI (2004a)
Characterization of Shiga toxin-producing *Escherichia coli* (STEC) isolated from seafood and beef. *FEMS Microbiol. Lett.* 233, 173-178

KUMAR, S. K. TAMURA und M. NEI (2004b)

MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* 5, 150-163

LARA-TEJERO, M. und J. E. GALAN (2000)

A bacterial toxin that controls cell cycle progression as a deoxyribonuclease I-like protein. *Science* 290, 354-357

LEUNG, P. H. M., J. S. M. PEIRIS, W. W. S. NG, R. M. ROBINS-BROWNE, K., A. BETTELHEIM und W. C. YAM (2003)

A newly discovered verotoxin variant, VT2g, produced by bovine verocytotoxigenic *Escherichia coli*. *Appl. Environ. Microbiol.* 69, 7549-7553

LEVINE, M. M. (1987)

Escherichia coli that cause diarrhea: enterotoxinogenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. *J. Infect. Dis.* 155, 377-389

LEVINE, M. M., J. G. XU, J. B. KAPER, H. LIOR, V. PRADO, B. TALL, J. NATARO, H. KARCH und K. WACHSMUTH (1987)

A DNA probe to identify enterohemorrhagic *Escherichia coli* of O157:H7 and other serotypes that cause hemorrhagic colitis and hemolytic uremic syndrome. *J. Infect. Dis.* 156, 175-182

LI, W. H. (1993)

Unbiased estimation of the rates of synonymous and nonsynonymous substitutions. *J. Mol. Evol.* 36, 96-99

LI, W. H. (1997)

Molecular Evolution. In: LI, W. H., (Hrsg.), 1. Auflage), Sinauer associates, Sunderland, USA, pp. 478

LI, Y., E. FREY, A. M. MACKENZIE und B. B. FINLAY (2000)

Human response to *Escherichia coli* O157:H7 infection: antibodies to secreted virulence factors. *Infect. Immun.* 68, 5090-5095

LOCKING, M., L. ALLISSON, L. RAE und M. HANSON (2003)

VTEC in Scotland 2002: enhanced surveillance and reference laboratory data. *SCIEH Weekly Report* 37, 304-307

LOUIE, M., J. C. S. DE AVEZADO, M. Y. C. HANDELSMAN, C. G. CLARK, B. ALLY, M. DYTOC, P. SHERMAN und J. BRUNTON (1993)

Expression and Characterization of the *eaeA* Gene Product of *Escherichia coli* Serotype O157:H7. *Infect Immun.* 61, 4085-4092

LOUIE, M., J. C. S. DE AZAVEDO, R. CLARKE, A. BORCZYK, H. LIOR, M. RICHTER und J. BRUNTON (1994)

Sequence heterogeneity of the *eae* gene and detection of verotoxin-producing *Escherichia coli* using serotype-specific primers. *Epidemiol. Infect.* **112**, 449-461

LUDWIG, K., V. SARKIM, M. BITZAN, M. A. KARMALI, C. BOBROWSKI, H. RUDER, R. LAUFS, I. SOBOTTKA, M. PETRIC, H. KARCH und D. E. MÜLLER-WIEFEL (2002)

Shiga toxin-producing *Escherichia coli* infection and antibodies against Stx2 and Stx1 in household contacts of children with enteropathic hemolytic-uremic syndrome. *J. Clin. Microbiol.* **40**, 1773–1782

LUO, Y., E. A. FREY, R. A. PFUETZNER, A. L. CREAGH, D. G. KNOEHEL, C. A. HAYNES, B. B. FINLAY und N. C. STRYNADKA (2000)

Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex. *Nature* **405**, 1073-1077

MANSFIELD, K. G., K. C. LIN, J. NEWMAN, D. SCHAUER, J. MACKKEY, A. A. LACKNER und A. CARVILLE (2001)

Identification of enteropathogenic *Escherichia coli* in simian immunodeficiency virus-infected infant and adult rhesus macaques. *J. Clin. Microbiol.* **39**, 971-976

MARQUES, L. R. M., C. M. ABE, P. M. GRIFFIN und T. A. T. GOMES (1995)

Association between alpha-hemolysin production and HeLa cell-detaching activity in fecal isolates of *Escherichia coli*. *J. Clin. Microbiol.* **33**, 2707-2709

MARCHES, O., J. P. NOUGAYREDE, S. BOULLIER, J. MAINIL, G. CHARLIER, I. RAYMOND, P. POHL, M. BOURY, J. DE RYCKE, A. MILON und E. OSWALD (2000)

Role of tir and intimin in the virulence of rabbit enteropathogenic *Escherichia coli* serotype O103:H2. *Infect. Immun.* **68**, 2171-2182

MARTIN, D. P. und E. RYBICKI (2000)

RDP: detection of recombination amongst aligned sequences. *Bioinf.* **16**, 562-563

MARTIN, D. P., C. WILLIAMSON und D. POSADA (2005)

RDP2: recombination detection and analysis from sequence alignments. *Bioinf.* **21**, 260–262

MCCLURE, M. A., T. K. VASI und W. M. FITCH (1994)

Comparative analysis of multiple protein-sequence alignment methods. *Mol. Biol. Evol.* **11**, 571-592 Erratum in: *Mol. Biol. Evol.* **11**, 811

MCDANIEL, T. K. und J. B. KAPER (1997)

A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Mol. Microbiol.* 23, 399-407

MCDANIEL, T. K., K. G. JARVIS, M. S. DONNENBERG und J. B. KAPER (1995)

A genetic locus of enterocyte attachment and effacement conserved among diverse enterobacterial pathogens. *Proc. Natl. Acad. Sci.* 92, 1664-1668

MCGRAW, E. A., J. LI, R. K. SELANDER und T. S. WHITTAM (1999)

Molecular evolution and mosaic structure of alpha, beta, and gamma intimins of pathogenic *Escherichia coli*. *Mol. Biol. Evol.* 16, 12-22

MELTON-CELSA, A. R., S. C. DARNELL und A. D. O'BRIEN (1996)

Activation of Shiga-like toxins by mouse and human intestinal mucus correlates with virulence of enterohemorrhagic *Escherichia coli* O91:H21 isolates on orally infected, streptomycin-treated mice. *Infect. Immun.* 64, 1569-1576

MELTON-CELSA, A. R., J. E. ROGERS, C. K. SCHMITT, S. C. DARNELL und A. D. O'BRIEN (1998)

Virulence of Shiga toxin-producing *Escherichia coli* (STEC) in orally-infected mice correlates with the toxin produced by the infecting strain. *Jpn. J. Med. Sci. Biol.* 51 S, S108-S114

MEYER, T., H. KARCH, J. HACKER, H. BOCKLAGE und J. HEESEMANN (1992)

Cloning and sequencing of an Shiga-like toxin II-related gene from *Escherichia coli* O157:H7 strain 7279. *Zbl. Bakt.* 276, 176-188

MONTENEGRO, M. A., M. BÜLTE, T. TRUMPF, S. ALEKSIC, G. REUTER und E. BULLING (1990)

Detection and characterization of fecal verotoxin-producing *Escherichia coli* from healthy cattle. *J. Clin. Microbiol.* 28, 1417-1421

MOON, H. W., S. C. WHIPP, R. A. ARGENZIO, M. M. LEVINE und R. A. GIANELLA (1983)

Attaching and effacing activities of rabbit and human enteropathogenic *Escherichia coli* in pig and rabbit intestines. *Infect. Immun.* 41, 1340-1351

MORRISON, D. A. und J. T. ELLIS (1997)

Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of apicomplexa. *Mol. Biol. Evol.* 14, 428-441

- MULLIS, K. B. und F. A. FALOONA (1987)
Specific synthesis of DNA *in vitro* via polymerase-catalyzed chain reaction. Meth. Enzymol. 155, 335-350
- MUZA-MOONS, M. M., A. KOUTSOURIS UND G. HECHT (2003)
Disruption of cell polarity by enteropathogenic *Escherichia coli* enables basolateral membrane proteins to migrate apically and to potentiate physiological consequences. Infect. Immun. 71, 7069-7078
- NAGY B. und P. Z. FEKETE (1999)
Enterotoxigenic *Escherichia coli* (ETEC) in farm animals. Vet. Res. 30, 259-284
- NAGY B. und P. Z. FEKETE (2005)
Enterotoxigenic *Escherichia coli* in veterinary medicine. Int. J. Med. Microbiol. 295, 443-454
- NATARO, J. P., I. C. A. SCALETSKY, J. B. KAPER, M. M. LEVINE und L. R. TRABULSI (1985)
Plasmid-mediated factors conferring diffuse and localized adherence of enteropathogenic *Escherichia coli*. Infect. Immun. 48, 378-383
- NATARO, J. P., J. B. KAPER, R. ROBINS-BROWNE, V. PRADO, P. VIAL und M. M. LEVINE (1987)
Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. Pediatr. Infect. Dis. J. 6, 829-831
- NATARO, J. P. und J. B. KAPER (1998)
Diarrheagenic *Escherichia coli*. Clin. Microbiol. Rev. 11, 142-201
- NATARO, J. P., T. STEINER und G. GUERRANT (1998)
Enterotoxigenic *Escherichia coli*. Emerg. Infect. Dis. 4, 251-261
- NAYLOR, S. W., D. L. GALLY und J. C. LOW (2005)
Enterohaemorrhagic *E. coli* in veterinary medicine. Int. J. Med. Microbiol. 295, 419-441
- NEI, M., J. C. STEPHENS und N. SAITOU (1985)
Methods for computing the standard errors of branching points in an evolutionary tree and their application to molecular data from humans and apes. Mol. Biol. Evol. 2, 66-85
- NEI, M. und T. GOJOBORI (1986)
Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418-426

- NOTREDAME, C. und D. G. HIGGINS (1996)
SAGA: sequence alignment by genetic algorithm. *Nucleic Acids Res.* 24, 1515-1524
- NOTREDAME, C., HIGGINS, D. G. und J. HERINGA (2000)
T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302, 205-217
- O'BRIEN, A. D. und G. D. LA VECK (1983)
Purification and characterization of a *Shigella dysenteriae* 1-like toxin produced by *Escherichia coli*. *Infect. Immun.* 40, 675-683
- O'BRIEN, A. D. und R. K. HOLMES (1987)
Shiga and Shiga-like Toxins. *Microbiol. Rev.* 51, 206-220
- O'BRIEN, A. D., M. R. THOMPSON, J. R. CANTEY und S. B. FORMAL (1977)
Production of a *Shigella dysenteriae*-like toxin by pathogenic *Escherichia coli*. *Abstr. Annu. Meet. Am. Soc. Microbiol.* B-103, p. 32
- OHMURA-HOSHINO, M., S.-T. HO, H. KURAZONO, K. IGARASHI, S. YAMASAKI und Y. TAKEDA (2003)
Genetic and immunological analysis of a novel variant of Shiga toxin 1 from bovine *Escherichia coli* strains and development of bead-ELISA to detect the variant toxin. *Microbiol. Immunol.* 47, 717-725
- ORDEN, J. A., D. CID, J. A. RUIZ-SANTA-QUITERIA, S. GARCIA, S. MARTINEZ und R. DE LA FUENTE (2002)
Verotoxin-producing *Escherichia coli* (VTEC), enteropathogenic *E. coli* (EPEC) and necrotoxicogenic *E. coli* (NTEC) isolated from healthy cattle in Spain. *J. Appl. Microbiol.* 93, 29-35
- ØRSKOV, F. und I. ØRSKOV (1984)
Serotyping of *Escherichia coli*. In: BERGAN, T. (Hrsg.) *Methods in Microbiol.* 15, Academic Press, London, pp. 43-112
- OSWALD, E., H. SCHMIDT, S. MORABITO, H. KARCH, O. MARCHES und A. CAPRIOLI (2000)
Typing of intimin genes in human and animal enterohemorrhagic and enteropathogenic *Escherichia coli*: characterization of a new intimin variant. *Infect. Immun.* 68, 64-71

PABST, W. L., M. ALTWEGG, C. KIND, S. MIRLANIC, D. HARDEGGER und D. NADAL (2003)

Prevalence of enteroaggregative *Escherichia coli* among children with and without diarrhea in Switzerland. *J. Clin. Microbiol.* 41, 2289-2293

PADIDAM, M., S. SAWYER und C. M. FAUQUET (1999)

Possible emergence of new geminiviruses by frequent recombination. *Virology* 265, 218-225

PAIVA DE SOUSA, C. und J. D. DUBREUIL (2001)

Distribution and expression of the *astA* gene (EAST1 toxin) in *Escherichia coli* and *Salmonella*. *Int. J. Med. Microbiol.* 291, 15-20

PATON, A. W., J. C. PATON, M. W. HEUZENROEDER, P. N. GOLDWATER und P. A. MANNING (1992)

Cloning and nucleotide sequence of a variant Shiga-like toxin II gene from *Escherichia coli* OX3:H21 isolated from a case of sudden infant death syndrome. *Microb. Pathog.* 13, 225-236

PATON, A. W., J. C. PATON und P. A. MANNING (1993)

Polymerase chain reaction amplification, cloning and sequencing of variant *Escherichia coli* Shiga-like toxin type II operons. *Microb. Pathog.* 15, 77-82

PATON, A. W., P. SRIMANOTE, M. C. WOODROW und J. C. PATON (2001)

Characterization of *Saa*, a novel autoagglutinating adhesin produced by locus of enterocyte effacement-negative Shiga-toxigenic *Escherichia coli* strains that are virulent for humans. *Infect. Immun.* 69, 6999-7009

PERES, S. Y., O. MARCHES, F. DAIGLE, J.-P. NOUGAYREDE, F. HERAULT, C. TASCA, J. DE RYCKE und E. OSWALD (1997)

A new cytolethal distending toxin (CDT) from *Escherichia coli* producing CNF2 blocks HeLa cell division in G2/M phase. *Mol. Microbiol.* 24, 1095-1107

PERNA, N. T., G. F. MAYHEW, G. POSFA, S. ELLIOTT, M. S. DONNENBERG, J. B. KAPER und F. R. BLATTNER (1998)

Molecular evolution of a pathogenicity island from enterohemorrhagic *Escherichia coli* O157:H7. *Infect. Immun.* 66, 3810-3817

PETERSON, J. D., L. A. UMayAM, T. M. DICKINSON, E. K. HICKEY und O. WHITE (2001)

The Comprehensive Microbial Resource. *Nucleic Acids Research* 29, 123-125

PHILLIPS, A. D. und G. FRANKEL (2000)

Intimin-mediated tissue specificity in enteropathogenic *Escherichia coli* interaction with human intestinal organ cultures. *J. Infect. Dis.* 181, 1496-1500

PHILLIPS A. D., S. NAVABPOUR, S. HICKS, G. DOUGAN, T. WALLIS und G. FRANKEL (2000)

Enterohaemorrhagic *Escherichia coli* O157:H7 target Peyer's patches in humans and cause attaching/effacing lesions in both human and bovine intestine. *Gut*

PICKETT, C. L., D. L. COTTLE, E. C. PESCI und G. BIKAH (1994)

Cloning, sequencing, and expression of the *Escherichia coli* cytolethal distending toxin genes. *Infect. Immun.* 62, 1046-1051

PIERARD, D., G. MUYLDERMANS, L. MORIAU, D. STEVENS und S. LAUWERS (1998)

Identification of new verocytotoxin type 2 variant B-subunit genes in human and animal *Escherichia coli* isolates. *J. Clin. Microbiol.* 36, 3317–3322

PIVA, I. C., A. L. PEREIRA, L. R. FERRAZ, R. S. N. SILVA, A. C. VIEIRA, J. E. BLANCO, M. BLANCO, J. BLANCO und L. G. GIUGLIANO (2003)

Virulence markers of enteroaggregative *Escherichia coli* isolated from children and adults with diarrhea in Brasília, Brazil. *J. Clin. Microbiol.* 41, 1827-1832

POSADA, D. (2002)

Evaluation of methods for detecting recombination from DNA sequences: empirical data. *Mol. Biol. Evol.* 19, 708-717

POSADA, D. und K. A. CRANDALL (2001a)

Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol. Evol.* 16, 37-45

POSADA, D. und K. A. CRANDALL (2001b)

Evaluation of methods for detecting recombination from DNA sequences: Computer simulations. *Proc. Natl. Acad. Sci.* 98, 13757-13762

PRADEL, N., V. LIVRELLI, C. DE CHAMPS, J. B. PALCOUX, A. REYNAUD, F. SCHEUTZ, J. SIROT, B. JOLY und C. FORESTIER (2000)

Prevalence and characterization of Shiga toxin-producing *Escherichia coli* isolated from cattle, food, and children during a one-year prospective study in France. *J. Clin. Microbiol.* 38, 1023–1031

QUINTERO BOTERO, C. L. (2003)

Phäno- und Genotypisierung von *Escherichia coli* O157-Stämmen aus unterschiedlichen Habitaten. *Vet. Med. Diss.*, Justus-Liebig-Universität Gießen

QUINTO, E. J. und A. CEPEDA (1996)

Presence of CNF-producing *Escherichia coli* strains in soft cheese. *Arch. Lebensmittelhyg.* 47, 81-104

RAGHAVA, G. P. S., S. M. SEARLE, P. C. AUDLEY, J. D. BARBER und G. J. BARTON (2003)

OXBench: a benchmark for evaluation of protein multiple sequence alignment accuracy. *BMC. Bioinformatics* 4, 47-70

RAMACHANDRAN, V., K. BRETT, M. A. HORNITZKY, M. DOWTON, K. A. BETTELHEIM, M. J. WALKER und S. P. DJORDJEVIC (2003)

Distribution of intimin subtypes among *Escherichia coli* isolates from ruminant and human sources. *J. Clin. Microbiol.* 41, 5022-5032

REID, S. D., C. J. HERBELIN, A. C. BUMBAUGH, R. K. SELANDER und T. S. WHITTAM (2000)

Parallel evolution of virulence in pathogenic *Escherichia coli*. *Nature* 406, 64-67

REUTER, G. (1996)

Mikrobiologie des Fleisches. In: WEBER, H. (Hrsg.), *Mikrobiologie der Lebensmittel, Fleisch und Fleischerzeugnisse*, 1. Aufl., Behr, Hamburg, pp. 1-108

REVAZISHVILI, T., M. KOTETISHVILI, O. C. STINE, A. S. KREGER, J. G. MORRIS JR und SULAKVELIDZE (2004)

Comparative analysis of multilocus sequence typing and pulsed-field gel electrophoresis for characterizing *Listeria monocytogenes* strains isolated from environmental and clinical sources. *J. Clin. Microbiol.* 42, 276-285

RILEY, L. W., R. S. REMIS, S. D. HELGERSON, H. B. MCGHEE, J. G. WELLS, B. R. DAVIS, R. J. HERBERT, E.S. OLCOTT, L. M. JOHNSON, N. T. HAGRETT, P. A. BLAKE und M. L. COHEN (1983)

Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. *N. Engl. J. Med.* 308, 681-685

RILEY, D. G., J. T. GRAY, G. H. LONERAGAN, K. S. BARLING und C. C. CHASE, JR (2003)

Escherichia coli O157:H7 prevalence in fecal samples of cattle from a southeastern beef cow-calf herd. *J. Food Prot.* 66, 1778-1782

RKI (ROBERT KOCH-INSTITUT, 1996)

Übersicht: Die Pathovaren von *Escherichia coli* beim Menschen. *Epid. Bull.* 30, 205-206

RKI (ROBERT KOCH-INSTITUT, 2004)

Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2003. http://www.rki.de/cln_006/nn_226928/DE/Content/Infekt/Jahrbuch/Jahrbuch__2003,templateId=raw,property=publicationFile.pdf/Jahrbuch_2003

RKI (ROBERT KOCH-INSTITUT, 2005)

Aktuelle Statistik meldepflichtiger Infektionskrankheiten. *Epid. Bull.* 3, 18-20

RKI (ROBERT KOCH-INSTITUT) und BGVV (BUNDESINSTITUT FÜR GESUNDHEITLICHEN VERBRAUCHERSCHUTZ UND VETERINÄRMEDIZIN, 2001)

EHEC-Infektionen–Erkennung, Verhütung und Bekämpfung. *Bundesgesundhbl.* 44, 1146–1148

ROBERTS, J. A., P. A. UPTON und G. AZENE (2000)

Escherichia coli O157:H7; an economic assessment of an outbreak. *J. Public Health Med.* 22, 99-107

ROLLE, M. und A. MAYR (1993)

Medizinische Mikrobiologie, Infektions- und Seuchenlehre für Tierärzte, Biologen, Agrarwissenschaftler und Interessierte aus benachbarten Fachgebieten. Enke-Verlag, Stuttgart, 6. Auflage, pp. 579-593

ROSENSHINE, I., M. S. DONNENBERG, J. B. KAPER und B. B. FINLAY (1992)

Signal transduction between enteropathogenic *Escherichia coli* (EPEC) and epithelial cells: EPEC induces tyrosine phosphorylation of host cell proteins to initiate cytoskeletal rearrangement and bacterial uptake. *EMBO J.* 11, 3551-3560

ROZAS, J. und R. ROZAS (1995)

DnaSP, DNA sequence polymorphism: an interactive program for estimating Population Genetics parameters from DNA sequence data. *Comput. Applic. Biosci.* 11, 621-625

ROZAS, J., J. C. SANCHEZ-DELBARRIO, X. MESSEGUER und R. ROZAS (2003)

DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19, 2496-2497

RUMER, L., J. JORES, P. KIRSCH, Y. CAVIGNAC, K. ZEHMKE und L. H. WIELER (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, 463-475

RUSSO, C. A. M., N. TAKEZAKI und M. NEI (1996)

Efficiencies of Different Genes and Different Tree-building Methods in Recovering a Known Vertebrate Phylogeny. *Mol. Biol. Evol.* 13, 525-536

RUSSO, T. A. und J. R. JOHNSON (2000)

Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *J. Infect. Dis.* 181, 1753-1754, Epub 2000

RZHETSKY, A. und M. NEI (1992a)

Statistical properties of the ordinary least-squares, generalized least-squares, and minimum-evolution methods of phylogenetic inference. *J. Mol. Evol.* 35, 367-375

RZHETSKY, A. und M. NEI (1992b)

A simple method for estimating and testing minimum- evolution trees. *Mol. Biol. Evol.* 9, 945-967

SAITOU, N. und M. NEI (1987)

The Neighbor-joining Method: A new Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.* 4, 406-425

SALEMI, M., T. DE OLIVEIRA, V. COURGNAUD, V. MOULTON, B. HOLLAND, S. CASSOL, W. M. SWITZER und A. M. VANDAMME (2003)

Mosaic genomes of the six major primate lentivirus lineages revealed by phylogenetic analyses. *J. Virol.* 77, 7202-7213

SALMINEN, M. O., J. K. CARR, D. S. BURKE und F. E. MCCUTCHAN (1995)

Identification of breakpoints in intergenotypic recombinants of HIV type 1 by Bootscanning. *AIDS Res. Hum. Retroviruses* 11, 1423-1425

SANDNER, L., L. E. EGUIARTE, A. NAVARRO, A. CRAVIOTO und V. SOUZA (2001)

The elements of the locus of enterocyte effacement in human and wild mammal isolates of *Escherichia coli*: evolution by assemblage or disruption? *Microbiology* 147, 3149-3158

SAVARINO, S. J., A. MCVEIGH, J. WATSON, A. CRAVIOTO, J. MOLINA, P. ECHERVERRIA, M. K. BHAN, M. M. LEVINE und A. FASANO (1996)

Enteroaggregative *Escherichia coli* heat-stable enterotoxin is not restricted to enteroaggregative *E. coli*. *J. Infect. Dis.* 173, 1019-1022

SCALETSKY, I. C. A., S. H. FABBRICOTTI, R. L. B. CARVALHO, C. R. NUNES, H. S. MARANHÃO, M. B. MORAIS und U. FAGUNDES-NETO (2002)

Diffusely adherent *Escherichia coli* as a cause of acute diarrhea in young children in northeast Brazil: a case-control study. *J. Clin. Microbiol.* 40, 645-648

SCHAUER, D. B. und S. FALKOW (1993a)

Attaching and effacing locus of a *Citrobacter freundii* biotype that causes transmissible murine colonic hyperplasia. *Infect Immun.* 61, 2486-2492

SCHAUER, D. B. UND S. FALKOW (1993b)

The *eae* gene of *Citrobacter freundii* biotype 4280 is necessary for colonization in transmissible murine colonic hyperplasia. *Infect. Immun.* 61, 4654-4661

SCHMIDT, H., B. PLASCHKE, S. FRANKE, H. RUESSMANN, A. SCHWARZKOPF, J. HEESEMANN und H. KARCH (1994)

Differentiation in virulence patterns of *Escherichia coli* possessing *eae* genes. *Med. Microbiol. Immunol.* 183, 23-31

SCHMIDT, H., J. SCHEEF, S. MORABITO, A. CAPRIOLOI, L. H. WIELER und H. KARCH (2000)

A new Shiga toxin 2 variant (Stx2f) from *Escherichia coli* isolated from pigeons. *Appl. Environ. Microbiol.* 66, 1205–1208

SCHMITT, C. K., M. L. MCKEE und A. D. O'BRIEN (1991)

Two copies of Shiga-like toxin II-related genes common in enterohemorrhagic *Escherichia coli* strains are responsible for the antigenic heterogeneity of the O157:H⁻ strain E32511. *Infect. Immun.* 59, 1065-1073

SCHOOLNIK, G. K. (1993)

Intimin and the intimate attachment of bacteria to human cells. *J. Clin. Invest.* 92, 1117-1118

SCOTLAND, S. M., H. R. SMITH und B. ROWE (1985)

Two Distinct Toxins Active on Vero Cells from *Escherichia coli* O157
Lancet i, 885-886

SCOTT, D. A. und J. B. KAPER (1994)

Cloning and sequencing of the genes encoding *Escherichia coli* cytolethal distending toxin. *Infect. Immun.* 62, 244-251

SCVPH (SCIENTIFIC COMMITTEE ON VETERINARY MEASURES RELATING TO PUBLIC HEALTH, 2003)

Opinion of the Scientific committee on veterinary measures relating to public health on verotoxigenic *E. coli* (VTEC) in foodstuffs, 21-22.01.2003,
http://europa.eu.int/comm/food/fs/sc/scv/outcome_en.html

SEARS C.L. und J. B. KAPER (1996)

Enteric bacterial toxins: mechanisms of action and linkage to intestinal secretion. *Microbiol. Rev.* 60, 167-215

SHARIFF, M., M. K. BHAN, S. KNUTTON, B. K. DAS, S. SAINI und R. KUMAR (1993)

Evaluation of the fluorescence actin staining test for detection of enteropathogenic *Escherichia coli*. *J. Clin. Microbiol.* 31, 386-389

SHAW, R. K., S. DANIELL, F. EBEL, G. FRANKEL und S. KNUTTON (2001)

EspA filament-mediated protein translocation into red blood cells. *Cell Microbiol.* 3, 213-222

SIEGLER, R. L. (1995)

The hemolytic uremic syndrome. *Pediatr. Clin. North. Am.* 42, 1505-1529

- SILVESTRO, L. M. CAPUTO, S. BLANCATO, L. DECASTELLI, A. FIORAVANTI, R. TOZOLLI, S. MORABITO und A. CAPRIOLI (2004)
Asymptomatic carriage of verocytotoxin-producing *Escherichia coli* O157 in farm workers in Northern Italy. *Epid. Infect.* 132, 915-919
- SINCLAIR, J. F. und A. D. O'BRIEN (2002)
Cell surface-localized nucleolin is a eukaryotic receptor for the adhesin intimin-gamma of enterohemorrhagic *Escherichia coli* O157:H7. *J. Biol. Chem.* 277, 2876-2885
- SITNIKOVA T, A. RZHETSKY und M. NEI (1995)
Interior-branch and bootstrap tests of phylogenetic trees. *Mol. Biol. Evol.* 12, 319-333
- SMITH, J. M. (1992)
Analyzing the mosaic structure of genes. *J. Mol. Evol.* 34, 126-129
- SMITH, J. M. (1999)
The detection and measurement of recombination from sequence data. *Genetics* 153, 1021-1027
- SMITH, N.G. (2003)
Are radical and conservative substitution rates useful statistics in molecular evolution? *J. Mol. Evol.* 57, 467-478
- SNEATH, P. H. A. und R. R. SOKAL (1973)
Numerical taxonomy. The Principles and Practice of Numerical Classification. p.573, Freeman, 1. Auflage, San Francisco, CA , USA
- SOKAL, R. R. und C. D. MICHENER (1958)
A statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* 28, 1409-1438
- STALEY, T., E. W. JONES und L. D. CORLEY (1969)
Attachment and penetration of *Escherichia coli* into intestinal epithelium of the ileum in newborn pigs. *Am. J. Pathol.* 56, 371-392
- STEPHAN, R., S. RAGETTI und F. UNTERMANN (2000)
Prevalence and characteristics of verotoxin-producing *Escherichia coli* (VTEC) in stool samples from asymptomatic human carriers working in the meat processing industry in Switzerland. *J. Appl. Microbiol.* 88, 335-341
- STROCKBINE, N. A., L. R. M. MARQUES, J. W. NEWLAND, H. WILLIAMS SMITH, R. K. HOLMES und A. D. O'BRIEN (1986)
Two toxin-converting phages from *Escherichia coli* O157:H7 strain 933 encode antigenically distinct toxins with similar biologic activities. *Infect. Immun.* 53, 135-140

SUERBAUM, S., J. M. SMITH, K. BAPUMIA, G. MORELLI, N. H. SMITH, E. KUNSTMANN, I. DYREK und M. ACHTMANN (1998)

Free recombination within *Helicobacter pylori*. Proc. Natl. Acad. Sci. 95, 12619-12624

SUZUKI, M., F. KONDO, Y. ITO, M. MATSUMOTO, M. HATA, H. OKA, M. TAKAHASHI und K. SAKAE (2004)

Identification of a Shiga-toxin type I variant containing an IS1203-like element, from Shigatoxin producing *Escherichia coli* O157:H7. FEMS Microbiol. Lett. 234, 63-67

SWEENEY, N. J., P. KLEMM, B. A. MCCORMICK, E. MOLLER-NIELSEN, M. UTLEY, M. A. SCHEMBRI, D. C. LAUX und P. S. COHEN (1996)

The *Escherichia coli* K-12 *gntP* gene allows *E. coli* F-18 to occupy a distinct nutritional niche in the streptomycin-treated mouse large intestine. Infect. Immun. 64, 3497-3503.

SWOFFORD, D. L., G. J. OLSEN, P. J. WADDEL und D. M. HILLIS (1996)

Phylogenetic interference. In: Molecular Systematics. HILLIS, D. M., C. MORITZ und B. K. MABLE (Hrsg.), 2. Auflage, Sinauer Associates, Sunderland MA, USA, pp. 407-514

TAKAHASHI, K. und M. NEI (2000)

Efficiencies of fast algorithms of phylogenetic inference under the criteria of maximum parsimony, minimum evolution and maximum likelihood when a large number of sequences are used. Mol. Biol. Evol. 17, 1251-1258

TAKEZAKI, N. (1998)

Tie trees generated by distance methods of phylogenetic reconstruction. Mol. Biol. Evol. 15, 727-737

TAKEZAKI, N. und M. NEI (1996)

Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. Genetics 144, 389-399

TAMURA, K. und M. NEI (1993)

Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10, 512-526

TARR, C. L. und T. S. WHITTAM (2002)

Molecular evolution of the intimin gene in O111 clones of pathogenic *Escherichia coli*. J. Bacteriol. 184, 479-487

TAVECHIO, A. T., L. R. MARQUES, C. M. ABE und T. A. GOMES (2004)

Detection of cytotoxic necrotizing factor types 1 and 2 among fecal *Escherichia coli* isolates from Brazilian children with and without diarrhea. Mem. Inst. Oswaldo Cruz. 99, 81-83

TAYLOR, W. R. (1987)

Multiple sequence alignment by a pairwise algorithm. *Comput. Appl. Biosci.* 3, 81-87

THOMPSON, J. D., D. G. HIGGINS und T. J. GIBSON (1994)

Clustalw: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* 22, 4673-4680

THOMPSON, J. D., F. PLEWNIAK und O. POCH (1999)

A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Res.* 27, 2682-2690

TOTH, I., E. OSWALD, J. G. MAINIL, M. AWAD-MASALMEH und B. NAGY (2000)

Characterization of intestinal *cnf1+* *Escherichia coli* from weaned pigs. *Int. J. Med. Microbiol.* 290, 539-542

TRAN VAN NHIEU G., R. BOURDET-SICARD, G. DUMENIL, A. BLOCKER und P. J. SANSONETTI (2000)

Bacterial signals and cell responses during *Shigella* entry into epithelial cells. *Cell Microbiol.* 2, 187-193

TZIPORI, S., F. GUNZER, M. S. DONNENBERG, L. DE MONTIGNY, J. B. KAPER und A. DONOHUE-ROLFE (1995)

The role of the *eaeA* gene in diarrhea and neurological complications in a gnotobiotic piglet model of enterohemorrhagic *Escherichia coli* infection. *Infect. Immun.* 63, 3621-3627

WAINWRIGHT, L. A. und J. B. KAPER (1998)

EspB and EspD require a specific chaperone for proper secretion from enteropathogenic *Escherichia coli*. *Mol. Microbiol.* 27, 1247-1260

WALLACE, I. M., O. O'SULLIVANN und D. G. HIGGINS (2005)

Evaluation of iterative alignment algorithms for multiple alignment. *Bioinf.* 21, 1408-1414

WEINSTEIN, D. L., M. P. JACKSON, J. E. SAMUEL, R. K. HOLMES und A. D. O'BRIEN (1988)

Cloning and sequencing of a Shiga-like toxin type II variant from *Escherichia coli* strain responsible for edema disease of swine. *J. Bacteriol.* 170, 4223-4230

WEIS, W. I., K. DRICKAMER und W. A. HENDRICKSON (1992)

Structure of a C-type mannose-binding protein complexed with an oligosaccharide. *Nature* 360, 127-134

WELLS, J. G., L. D. SHIPMAN, K. D. GREENE, E. G. SOWERS, J. H. GREEN, D. N. CAMERON, F. P. DOWNES, M. L. MARTIN, P. M. GRIFFIN, S. M. OSTROFF, *ET AL*

Isolation of *Escherichia coli* serotype O157:H7 and other Shiga-like-toxin-producing *E. coli* from dairy cattle. *J. Clin. Microbiol.* 29, 985–989

WHITTAM, T. S. (1995)

Genetic population structure and pathogenicity in enteric bacteria, In: BAUMBERG, S., J. P. W. YOUNG, E. M. H. WELLINGTON und J. R. SAUNDERS (Hrsg.), *Population genetics of bacteria*. 1. Auflage Cambridge University Press, Cambridge, England, p. 217-245

WHITTAM, T. S. (1998) Kapitel 20

In: *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* strains (KAPER, J. B. und A. D. O'BRIEN), ASM, Washington DC, USA, p. 195-209

WHITTAM, T. S. und E. A. MCGRAW (1996)

Clonal analysis of EPEC serogroups. *Rev. Microbiol.* 27 (Suppl. 1), 7–16

WHITTAM, T. S., M. L. WOLFE, I. K. WACHSMUTH, F. ORSKOV, I. ORSKOV und R. A. WILSON (1993)

Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea. *Infect. Immun.* 61, 1619-1629

WHO (WORLD HEALTH ORGANIZATION, 1997)

Prevention and control of enterohaemorrhagic *Escherichia coli* (EHEC) infections. Report of a WHO consultation, Genf, Schweiz, 28. April-1. Mai 1997

WHO/FSF/FOS/97.6

WHO (WORLD HEALTH ORGANIZATION, 1998)

Zoonotic non-O157 Shiga toxin-producing *Escherichia coli* (STEC). Report of a WHO scientific working group meeting. Berlin, Deutschland, 23.-26. Juni 1998 WHO/CSR/APH/98.8

WIELER, L. H., B. BUSSE, H. STEINRUCK, L. BEUTIN, A. WEBER, H. KARCH und G. BALJER (2000)

Enterohemorrhagic *Escherichia coli* (EHEC) strains of serogroup O118 display three distinctive clonal groups of EHEC pathogens. *J. Clin. Microbiol.* 38, 2162-2169

WIELER, L. H., E. VIELER, C. ERPENSTEIN, T. SCHLAPP, H. STEINRUCK, R. BAUERFEIND, A. BYOMI und G. BALJER (1996)

Shiga toxin-producing *Escherichia coli* strains from bovines: association of adhesion with carriage of *eae* and other genes. *J. Clin. Microbiol.* 34, 2980-2984

- WILLSON, S. J. (2005)
Minimum evolution using ordinary least-squares is less robust than neighbor-joining. *Bull. Math. Biol.* 67, 261-279
- WIUF, C., CHRISTENSEN, T. und J. HEIN (2001)
A simulation study of the reliability of recombination detection methods. *Mol. Biol. Evol.* 18, 1929-1939
- YANG, Z. (1994)
Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* 39, 306-314
- YANG, Z. (1997a)
PAML: a program for package for phylogenetic analysis by maximum likelihood. *CABIOS* 13, 555-556
- YANG, Z. (1997b)
How often do wrong models produce better phylogenies? *Mol. Biol. Evol.* 14, 105-108
- YANG, Z. und A. YODER (1999)
Estimation of the transition/transversion rate bias and species sampling. *J. Mol. Evol.* 48, 274-283
- YANG, Z. und R. NIELSEN (2000)
Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* 17, 32-43
- YU, J. und J. B. KAPER (1992)
Cloning and characterization of the *eae* gene of enterohaemorrhagic *Escherichia coli* O157:H7. *Mol. Microbiol.* 6, 411-417
- ZHANG J. (2000)
Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J. Mol. Evol.* 50, 56-68
- ZHANG, J. und X. GU (1998)
Correlation between the substitution rate and rate variation among sites in protein evolution. *Genetics* 149, 1615-1625
- ZHANG, J., H. F. ROSENBERG und M. NEI (1998)
Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc. Natl. Acad. Sci.* 95, 3708-3713
- ZHANG, W., M. BIELASZEWSKA, T. KUCZIUS und H. KARCH (2002a)
Identification, characterization, and distribution of a Shiga toxin 1 gene variant (*stx1c*) in *Escherichia coli* strains isolated from humans *J. Clin. Microbiol.* 40, 1441-1446

ZHANG, W. L., B. KÖHLER, E. OSWALD, L. BEUTIN, H. KARCH, S. MORABITO, A. CAPRIOLI, S. SUERBAUM und H. SCHMIDT (2002a)

Genetic diversity of intimin genes of attaching and effacing *Escherichia coli* strains. *J. Clin. Microbiol.* 40, 4486-4492

ZSCHÖCK, M., H. P. HAMANN, B. KLOPPERT und W. WOLTER (2000)

Shiga-toxin-producing *Escherichia coli* in faeces of healthy dairy cows, sheep and goats: prevalence and virulence properties. *Lett. Appl. Microbiol.* 31, 203–208

ZWEIFEL, C., J. E. BLANCO, M. BLANCO, J. BLANCO und R. STEPHAN (2004)
Serotypes and virulence genes of ovine non-O157 Shiga toxin-producing *Escherichia coli* in Switzerland. *Int. J. Food Microbiol.* 95, 19-27

Gesetze und Verordnungen

ANONYMOUS (2000)

Gesetz zur Verhütung und Bekämpfung von Infektionskrankheiten beim Menschen (Infektionsschutzgesetz-IfSG), zuletzt geändert am 06.08.2002
Bundesgesetzbl. Teil I, 1045

ANONYMOUS (2005)

Verordnung (EG) Nr. 2073/2005 der Kommission vom 15. November 2005 über mikrobiologische Kriterien für Lebensmittel Amtsblatt der europäischen Union vom 22.12.2005 L338/1 bis 26

Internetadressen

<http://ww.abacus.gene.ucl.ac.uk/software/paml.html>

<http://www.genes.mit.edu/GENSCAN.html>

<http://www.basic.nwu.edu/biotools/Chromas.html>

<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>

<http://www.ncbi.nlm.nih.gov>

<http://www.ncbi.nlm.nih.gov/BankIt/>

<http://www.ncbi.nlm.nih.gov/BLAST>

<http://www.tigr.org/CMR2>

<http://www.darwin.uvigo.es/rdp/rdp.html>

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben.

Herrn Univ.-Prof. Dr. M. Bülte danke ich für die Überlassung des Themas, die gewährte Unterstützung bei den Arbeiten und die Möglichkeit, mich umfassend weiterbilden zu können.

Besonders möchte ich mich bei Herrn Dr. Amir Abdulmawjood bedanken. Seine fortwährende Unterstützung über das rein Fachliche hinaus und die partnerschaftliche Zusammenarbeit eröffneten mir viele Perspektiven.

Bei Frau Karin Simon und Frau Claudia Walter möchte ich mich für die gute Zusammenarbeit und ihre technische Unterstützung über die vergangenen Jahre hinweg bedanken.

Herrn Prof. Dr. D. Huson, Herrn Prof. W. Hughes und Herrn Prof. J. Zhang danke ich für die hilfreichen Erläuterungen zum Gebrauch der Programme Splits-Tree, SCR3 und HON-NEW.

Bei meiner Familie und meinen Freunden möchte ich mich herzlich für ihr Verständnis gegenüber dem wiederholt fehlenden Tierarzt bedanken.

Herzlich möchte ich mich bei Miriam Schönenbrücher und Raphael Gasper bedanken, die trotz eigenem Stress und knapper Zeit immer noch Zeit für das Korrekturlesen gefunden haben und besonders in der Schlussphase in vielen Momenten für mich eingesprungen sind.

Der größte Dank gilt meinen Eltern Erich und Hannelore Schönenbrücher und meiner Frau Vanessa. Liebe Eltern, danke für eure beständige Unterstützung auf allen Ebenen.

Vanessa, ohne deine Hilfe und dein Verständnis bei der Bewältigung unseres „gemeinsamen Projektes“ wäre die Umsetzung nicht möglich gewesen. Zeit für neue Ziele!

Liebe Alrun, schön, dass du da bist.

