

Aus dem Max von Pettenkofer-Institut für Hygiene und Medizinische Mikrobiologie
der Ludwig-Maximilians-Universität München

Vorstand: Prof. Dr. Dr. J. Heesemann

**Charakterisierung einer horizontal transferierten Region
im *Escherichia coli* Stamm ECOR31**

Dissertation

zum Erwerb des Doktorgrades der Medizin

an der Medizinischen Fakultät der

Ludwig-Maximilians-Universität zu München

vorgelegt von

Bettina Preisach

aus Arad

2008

Mit Genehmigung der Medizinischen Fakultät
der Universität München

Berichterstatter: PD Dr. med. Sören Schubert

Mitberichterstatter: Prof. Dr. Hans-Gustav Klobeck
PD Dr. Johannes Liese

Dekan: Prof. Dr. med. Dr. h.c. M. Reiser, FACR, FRCR

Tag der mündlichen Prüfung: 11.12.2008

Inhaltsverzeichnis

Inhaltsverzeichnis	2
Abkürzungsverzeichnis	5
1. EINLEITUNG	6
1.1. Plastizität des bakteriellen Genoms	6
1.2. Horizontaler Gentransfer	9
1.2.1. Gentransfer durch Transformation	10
1.2.2. Gentransfer durch Transduktion	10
1.2.3. Gentransfer durch Konjugation	11
1.3. Genomische Inseln (GI) / Pathogenitätsinseln (PAI)	13
1.4. <i>Yersinia</i>-HPI, Bedeutung und Struktur	15
1.5. Die Verbreitung der HPI und Struktur bei <i>E. coli</i>	17
1.6. Zielsetzung der Arbeit	19
2. MATERIAL UND METHODEN	21
2.1. Material	21
2.1.1. Geräte	21
2.1.2. Chemikalien und Kleinmaterialien	22
2.1.3. Nährmedien	23
2.1.4. Stämme	23
2.1.5. Vektoren	25
2.1.6. Plasmide und Cosmide	25
2.1.7. Synthetische Oligonukleotide	25
2.2. Methoden	27
2.2.1. Isolierung chromosomaler DNA	27
2.2.2. Isolierung chromosomaler DNA mit Phenol	27
2.2.3. Schnellisolierung von Plasmid-DNA	28
2.2.4. Isolierung von Plasmid und Cosmid DNA	29
2.2.5. Schneiden von DNA mit Restriktionsendonukleasen	30
2.2.6. Dephosphorylierung mit alkalischer Phosphatase (CIAP Behandlung)	31

2.2.7. Agarose Gel-Elektrophorese von DNA	31
2.2.8. Elution von DNA-Fragmenten aus einem Agarosegel	32
2.2.9. Polymerase Kettenreaktion (PCR).....	32
2.2.10. LA-Polymerase-Kettenreaktion	33
2.2.11. Klonierung von DNA-Fragmenten	36
2.2.12. Klonierung von PCR-Fragmenten	37
2.2.13. Transformation elektrokompetenter <i>E. coli</i>	37
2.2.14. In vitro Transposon-Mutagenese	38
2.2.15. Herstellung einer Cosmid-Bank	40
2.2.16. Screening der Cosmid Bank, Herstellung der Colony-Blot-Filter	42
2.2.17. Markierung von Gensonden.....	42
2.2.18. Colony Blot-Hybridisierung	43
2.2.19. RNA-Isolation	44
2.2.20. Reverse Transkriptase Polymerasekettenreaktion (RT-PCR).....	45
2.2.21. Sequenzierung	46
2.3. Software	48
3. ERGEBNISSE	49
3.1. Untersuchung des Insertionsortes der <i>Yersinia</i>-HPI im <i>E. coli</i> Stamm ECOR31	49
3.2. Untersuchung der unbekanntenen Region X mit einer Cosmid-Bank von ECOR31	53
3.2.1. DNA-Bereich A.....	56
3.2.1.1. Sequenzanalyse DNA-Bereich A.....	56
3.2.1.2. Homologienvergleiche DNA-Bereich A auf DNA- und Proteinebene	57
3.2.1.3. Homologienvergleiche DNA-Abschnitt A mit pLVPK von <i>Klebsiella pneumoniae</i> (A Y378100) und mit <i>Mesorhizobium loti</i> MAFF303099 (AP003013)	62
3.2.1.4. Transkriptionsuntersuchungen ABC-Transporter (RT-PCR).....	64
3.2.2. DNA-Bereich B.....	66
3.2.2.1. Sequenzanalyse DNA-Bereich B.....	66
3.2.2.2. Homologienvergleiche DNA-Bereich B auf DNA- und Proteinebene	67
3.2.2.3. Homologienvergleiche DNA-Abschnitt B mit dem Plasmid pLVPK von <i>Klebsiella pneumoniae</i>	72
3.2.3. DNA-Bereich C.....	74
3.2.3.1. Sequenzanalyse DNA-Bereich C	74
3.2.3.2. Homologienvergleiche DNA-Bereich C auf DNA- und Proteinebene	74
3.2.4. DNA-Bereich D.....	77
3.2.4.1. Sequenzanalyse DNA-Bereich D	77
3.2.4.2. Homologienvergleiche DNA-Bereich D auf DNA- und Proteinebene	77

3.2.4.3. Homologienvergleich DNA-Abschnitt D mit dem Plasmid pLVPK von <i>Klebsiella pneumoniae</i>	81
3.2.4.4. Transkriptionsuntersuchungen ABC-Transporter (RT-PCR).....	82
3.2.5. DNA-Bereich E.....	84
3.2.5.1. Sequenzanalyse DNA-Bereich E.....	84
3.2.5.2. Homologienvergleiche DNA-Bereich E auf DNA- und Proteinebene	85
3.2.5.3. Homologienvergleiche DNA-Abschnitt E mit dem Plasmid pLVPK von <i>Klebsiella pneumoniae</i>	88
3.3. Untersuchung verschiedener Stämme auf das Vorhandensein von RegX.....	89
3.3.1. Anzahl und Art der untersuchten Bakterien-Stämme	89
3.3.2. RegX-positiver Stamm: <i>Klebsiella pneumoniae</i> Isolat 2495.....	90
4. DISKUSSION	93
Bakterielle Vielfalt.....	93
Bedeutung des horizontalen Gentransfers	94
4.1. Die Region X (RegX), eine Clusterregion für den Metallionenstoffwechsel.....	97
4.1.1. ABC-Transporter zur Aufnahme von Metallionen	98
4.1.2. Aufnahmesystem für Fe(II) und Transkriptionsregulatoren der Fur-Familie.....	100
4.2. Strukturelle Besonderheiten der Region X und deren Verbreitung	104
4.2.1. Ameliorationsmechanismen bakterieller Genome	104
4.2.2. Ko-Lokalisation der <i>Yersinia</i> -HPI und der Region X.....	106
4.2.3. Übertragung der DNA-Region RegX mittels horizontalem Gentransfer	107
4.2.4. Verbreitung der DNA-Region RegX bei der Familie der <i>Enterobacteriaceae</i>	110
Zusammenfassung der Arbeit.....	111
Literaturverzeichnis	112
Anhang	126
Lebenslauf.....	136

Abkürzungsverzeichnis

A. dest.	destilliertes Wasser (lat.: ' <u>a</u> qua <u>d</u> estillata')
aa	Aminosäuren (' <u>a</u> mino <u>a</u> cids')
Abb.	<u>A</u> bbildung
ADP	<u>A</u> denosin <u>d</u> iphosphat
ATP	<u>A</u> denosin <u>t</u> riphosphat
bp	Basenpaare (' <u>b</u> ase <u>p</u> airs')
Da / kDa	<u>D</u> alton / <u>K</u> ilo- <u>D</u> alton
ddNTP	<u>D</u> idesoxyribon <u>u</u> kleosid <u>t</u> riphosphat
dNTP	<u>D</u> esoxyribon <u>u</u> kleosid <u>t</u> riphosphat
DR	Wiederholungssequenz ' <u>d</u> irect <u>r</u> epeat'
GI	Genomische Insel (' <u>g</u> enomic <u>i</u> sland')
HPI	' <u>h</u> igh- <u>p</u> athogenicity <u>i</u> sland'
IS	Insertionselement (' <u>i</u> nsertion <u>e</u> lement')
kb	<u>K</u> ilob <u>a</u> senpaare
KBE / CFU	<u>k</u> olonie <u>b</u> ildende <u>E</u> inheiten (' <u>c</u> olony <u>f</u> orming <u>u</u> nits')
LB-Agar / -Medium	' <u>L</u> uria <u>B</u> ertani'- Agar / - Medium
OD	optische <u>D</u> ichte
ORF	offenes Leseraster (' <u>o</u> pen <u>r</u> eading <u>f</u> rame')
Orf	Translationsprodukt eines offenen Leserasters (Polypeptid)
PAI	Pathogenitätsinsel (' <u>p</u> athogenicity <u>i</u> sland')
rpm	Umdrehungen pro Minute (' <u>r</u> ounds <u>p</u> er <u>m</u> inute')
RT-PCR	<u>R</u> everse <u>T</u> ranskriptase-Polymerase-Kettenreaktion
SDS	' <u>s</u> odium <u>d</u> odecyl <u>s</u> ulfate'
SSC	' <u>s</u> tandard <u>s</u> aline <u>c</u> itrate'
tRNA	<u>T</u> ransfer- <u>R</u> ibon <u>u</u> kleinsäure
U	Einheiten (' <u>u</u> nits')
Ybt	Yersiniabactin-Siderophor

1. Einleitung

1.1. Plastizität des bakteriellen Genoms

Seit mehreren Millionen Jahren gibt es Bakterien auf der Erde. Sie haben sich ubiquitär verbreitet, und selbst lebenswidrige Nischen kolonisiert, wie beispielsweise heiße Quellen, hochkonzentrierte Salzlaken, schwefelsaure Schlammlöcher oder das arktische Eis.

Auch der menschliche Körper wird von mehreren Billionen Mikroorganismen kolonisiert. Sie besiedeln die Haut und Schleimhäute, wo sie als Symbiont oder Kommensale dem Körper nützlich sind^{56, 165}. Gleichsam sind Bakterien als Infektionserreger bekannt. Sie können mittels verschiedener Virulenzfaktoren in den Körper eindringen, den humoralen und zellulären Abwehrmechanismen entgehen, sich vermehren und den Körper durch eine Infektion schädigen^{31, 49, 52, 165}. Diese bakterielle Anpassungsfähigkeit mit Adaptation an neue Lebensräume spiegelt ein sehr wandlungsfähiges bakterielles Genom wieder.

Durch verbesserte und vereinfachte Sequenzierungsmethoden konnten seit 1995 Genome von mehr als 330 verschiedenen Organismen sequenziert werden, darunter hauptsächlich bakterielle Genome. Hieraus ergab sich eine ganz neue Einsicht in die genetischen Strukturen apathogener und pathogener Mikroorganismen. Die erste Erkenntnis war, dass das bakterielle Erbgut eine mosaikartige Struktur aufweist¹⁷³. Betrachtet man verschiedene Bakterien, so lässt sich zudem beobachten, dass die Größe der Genome sehr stark variiert. Der Größenumfang reicht von einem der kleinsten Genomen, von *Mycoplasma genitalium* mit 580 Kilobasenpaaren (kb), bis zu dem 10-fach größeren Genom von *Pseudomonas aeruginosa*¹⁶¹. Die Genome können jedoch hinsichtlich ihrer Größe nicht nur zwischen den Spezies, sondern auch innerhalb einer Spezies variieren¹¹⁵. Pathogene *Escherichia coli*-Stämme haben ein größeres Genom als kommensale *E. coli*, wobei der Größenunterschied bis 1000 kb betragen kann^{11, 12}. Was also macht den Unterschied aus?

Alle Bakterien besitzen eine Erbgut-Mindestausstattung, den so genannten 'Kern-Genpool' ('*core genepool*'). Dieser umfasst einen relativ umschriebenen und konstanten Anteil an Genen, deren Produkte in lebenswichtige Stoffwechselprozesse des Bakteriums involviert sind, wie DNA-, Glukose- und Protein-Stoffwechsel^{112, 114}. Der so genannte flexible Genpool ('*flexible genepool*') umfasst Gene, die nicht unmittelbar zum Überleben nötig sind, beispielsweise Antibiotikaresistenzen oder Virulenzfaktoren, die das Überleben der Bakterien im Wirt ermöglichen^{59, 61}. Der flexible Genpool entsteht durch die Modifikation des bestehenden Genoms und die Aufnahme fremder DNA^{63, 90, 102}. Im Gegensatz zum kommensalen, apathogenen *E. coli* verursachen beispielsweise *Escherichia*-Stämme vom Serotyp O157:H7 Ausbrüche von hämorrhagischer Kolitis und hämolytisch-urämischem Syndrom⁹⁸. *E. coli* O157:H7 erhält durch die Infektion mit einem Bakteriophagen Gene zur Produktion von Shiga-Toxin, einem Toxin von Shigellen^{127, 131}. Das Shiga-Toxin stoppt nach rezeptorvermittelter Aufnahme in menschliche Enterozyten die Proteinbiosynthese und führt zu blutigen Durchfällen mit hoher Mortalität.

Zu den Mechanismen, die das bakterielle Genom verändern und formen, gehören Punktmutationen, Umschichtungen ('*rearrangements*') über homologe Rekombination und *site*-spezifische Rekombinationen^{19, 43}.

Punktmutationen beschreiben den Austausch oder die Deletion einzelner Basen innerhalb einer Gensequenz. Darüber können Gene ganz ausgeschaltet werden oder bei Punktmutationen im Promotorbereich eine Änderung der Genexpression erzielt werden³⁶. Besonders bei der Entwicklung von Antibiotikaresistenzen spielen Punktmutationen eine wichtige Rolle. Mutationen im Gyrase- und Topoisomerase-Gen von *E. coli* führen beispielsweise zur Resistenz gegenüber Fluorchinolon-Antibiotika⁷⁰.

Homologe Rekombinationen können im Bereich von Regionen mit hoher Homologie der Nukleotidsequenz stattfinden. Vereinfacht dargestellt kommt es bei dem Prozess, an dem verschiedene Proteine beteiligt sind, nach einer *Crossover*-Überlagerung gleichartiger DNA-Sequenzen zu einer Umordnung im Genom ('*rearrangement*')¹⁵⁵. Es können Inversionen, Duplikationen und Deletionen resultieren. Die homologe Rekombination ist bei der Antigenvariation essentiell, kann aber auch ganz neue Phänotypen erzeugen¹⁰². *Neisseria gonorrhoeae* kann die Primärstruktur des Pilins

der Haftpili mit großer Häufigkeit über homologe Rekombination modifizieren⁹⁹. Interessanterweise führt nicht immer nur ein Zugewinn von DNA zu vermehrter Virulenz. Bei der Spezies *Shigella*, die eng verwandt ist mit *E. coli*, bewirkte die Deletion eines Genom-Abschnittes mit Verlust der Lysin-Dekarboxylase-(LDC)-Aktivität eine deutliche Steigerung der Virulenz¹⁰³.

Vergleichsuntersuchungen zwischen *E. coli* K12 und hochpathogenen *E. coli*-Stämmen zeigen in den pathogenen Stämmen eine höhere Rate an Rekombinations-Ereignissen, was vermutlich an Mutationen in den Mismatch-Reparatur-Genen liegt¹⁰⁹.

Während Punktmutationen lediglich eine Modifikation der vorhandenen DNA vornehmen, können die homologe Rekombination und vor allem die *site*-spezifische Rekombination die Aufnahme fremder DNA in das Genom vermitteln. Bei der *site*-spezifischen Rekombination fungieren bestimmte DNA-Sequenzen als Erkennungssequenz ('*attachment site*') für die Integration von fremder DNA, wie beispielsweise des Bakteriophagen Lambda⁶².

Bakterien profitieren vom Zugewinn fremder DNA und diese macht einen wesentlichen Teil des flexiblen Genpools aus¹¹⁵. DNA-Transfer ist auch unter weit entfernt verwandten Stämmen möglich und bietet die Gelegenheit, in einem Schritt komplexe neue Stoffwechsel-Eigenschaften zu inkorporieren, die zur Entstehung bakterieller Vielfalt und neuer Arten beitragen¹⁸¹.

Nach erfolgreichem DNA-Transfer kann der Rezipient von der fremden oder artverwandten DNA nur profitieren, wenn es zur Integration der DNA in das Genom des Rezipienten kommt oder die DNA in einem Plasmid dauerhaft beherbergt wird. Zudem müssen die erhaltenen Gene adäquat exprimiert werden können¹¹⁵. Stellt die erworbene DNA einen Selektionsvorteil für den Rezipienten dar, wird die Region einem Adaptationsprozess unterzogen, dem Wirtsgenom in der Aminosäurekodierung angeglichen und über Ausschaltung oder Mutationen der Mobilisierungsgene stabilisiert^{60, 90, 115}. Einem nicht unerheblichen Anteil der erworbenen Gene wird sich das Bakterium gleichsam aber wieder entledigen, denn jedes Genom unterliegt einem Selektionsdruck, um eine gewisse Größe nicht zu überschreiten^{2, 3}.

1.2. Horizontaler Gentransfer

Eine Übertragung von Genen außerhalb des vertikalen Transfers über Zellteilung wird als horizontaler oder lateraler Gentransfer bezeichnet. Dieser ist auch über Artgrenzen hinweg möglich und erfolgt üblicherweise über Transformation, Transduktion und Konjugation, siehe Unterpunkt 1.2.1. bis 1.2.3. Bekannte Vehikel horizontalen Gentransfers sind Phagen, Plasmide und andere translozierbare Elemente, wie Integrons, Transposons und Insertions-Elemente (IS)^{6, 91, 182}. Über die Ergebnisse der Genomsequenzierungen und biomathematischen Methoden kann diese akquirierte DNA aufgespürt werden⁹. Trotz des universellen DNA-Codes unterscheiden sich die verschiedenen Bakterien-Gattungen u.a. durch den Anteil der Basen Guanin und Cytosin (G+C-Gehalt) auf Grund der Präferenz unterschiedlicher Triplets zur Darstellung der Aminosäuren^{81, 113}. Weitere Indizien für neu erworbene DNA sind in das Genom integrierte rudimentäre Vehikel des horizontalen Gentransfers, wie Transferorigins fremder Plasmide, die Erkennungssequenzen ('*attachment sites*') für Bakteriophagen oder translozierbare Elemente, wie Transposons und Insertionselemente⁴³. Man schätzt, dass zwischen 8 bis 18% des Genoms von *Escherichia coli* über horizontalen Gentransfer erworben wurde^{90, 115}.

Die Bedeutung, die horizontalem Gentransfer zukommt, wächst mit der Möglichkeit des Rezipienten die neuen Gene zu nutzen, um beispielsweise eine neue Stoffwechseleigenschaft zu erwerben und somit eine (neue) ökologische Nische zu erschließen¹²⁰. Bakterien können bereits 'etablierte' Eigenschaften von anderen Stämmen entlehnen und damit das eigene Genom für spezielle Anforderungen präparieren.

Die klinische Bedeutung des horizontalen Gentransfers steigt besonders mit dem Erwerb von Funktionen, die mit einer erhöhten Pathogenität einhergehen und die es Stämmen ermöglichen im menschlichen Körper zu wachsen, ihn zu schädigen und Abwehrmechanismen des Körpers zu entgehen^{8, 57}. Der Erreger der Diphtherie, *Corynebacterium diphtheriae*, ist beispielsweise erst nach Infektion mit einem Bakteriophagen in der Lage das schädigende Diphtherie-Toxin zu bilden. Dieser Phage überträgt die Gene, die zur Synthese des Translations-hemmenden Toxins nötig sind¹³⁵ (siehe Unterpunkt 1.2.2.). So führt die Diphtherie zu einer Lokalinfektion

an den Tonsillen und zu einer systemischen Infektion, mit Parenchymdegeneration vor allem im Herzmuskel, in der Leber, in den Nieren und den Nebennieren ⁸².

1.2.1. Gentransfer durch Transformation

Unter Transformation wird die Aufnahme freier DNA durch Bakterien verstanden.

Die Transformation war der erste Gen-Transfer-Prozess, der für Bakterien beschrieben wurde. Bereits 1928 fand Frederick Griffith heraus, dass die Fähigkeit der Kapselbildung zwischen verschiedenen Pneumokokken-Stämmen übertragen werden kann. Dass dabei die DNA transferiert wird konnte später von O. Avery beschrieben werden ^{6, 174}. Neben Streptokokken zählen *Neisseria gonorrhoeae*, *Bacillus subtilis*, *Haemophilus influenzae* und *Helicobacter pylori* zu den Bakterien, die ständig 'kompetent' sind fremde DNA aufzunehmen ⁴⁷. Durch spezielle Erkennungssequenzen kann diese über Rekombination in das Genom aufgenommen werden ⁴⁷. Der Gentransfer durch Transformation bleibt jedoch meist auf kleine DNA-Fragmente begrenzt und eignet sich damit beispielsweise zur Antigenvariation, aber weniger zum Zugewinn ganzer Gencluster.

1.2.2. Gentransfer durch Transduktion

Die Schlüsselvehikel der Transduktion sind Bakteriophagen, die über unspezifische und spezifische Transduktion eine Übertragung von Genen zwischen Bakterien ermöglichen, ohne dass diese dazu in direkten Kontakt zueinander treten. Die Transduktion durch Bakteriophagen ist ein wesentlicher Faktor beim Gentransfer, zumal Bakteriophagen viele Jahre auch in zellfreiem Milieu überleben können und die DNA im Phagenkopf gut geschützt ist vor Nukleasen und organischen Lösungsmittel ¹⁹. Die Infektion mit einem Bakteriophagen kann für das Bakterium den Zugewinn wichtiger Virulenzdeterminanten bedeuten. So sind unter anderem folgende Toxine auf Bakteriophagen kodiert: Diphtherie-Toxin (*C. diphtheriae*), das Exotoxin A (*Streptococcus pyogenes*), Cholera-Toxin (*Vibrio cholerae*) und Botulinum-Toxin (*Clostridium botulinum*).

Die unspezifische Transduktion wird durch virulente Bakteriophagen vermittelt. Bei der Phagenvermehrung in der Bakterienzelle (lytischer Zyklus) kann versehentlich bakterielle DNA des Wirtes statt Bakteriophagen-DNA in den Phagenkopf verpackt werden. Bei der Infektion eines neuen Bakteriums durch diese Phagen kann die DNA in das neue Bakteriengenom über Rekombination eingebaut werden. Sofern es sich um ein Transposon handelt, kann sich dieses selbst integrieren¹⁵⁶. Beispiele sind die *E. coli*-Phagen P1, T1 und T4, sowie der Phage P22 von *Salmonella*^{19, 182}.

Temperente Phagen wie der Bakteriophage λ , der typischerweise *E. coli* infiziert, ermöglichen die spezifische Transduktion. Diese Phagen durchlaufen einen lysogenen Zyklus. Die Phagen-DNA wird an bestimmten Anhaftungs-Stellen ('*attachment-sites*') in das Chromosom des Wirtes integriert und als Prophage über die Wirtszelle vermehrt¹⁴¹. Die Wirtszelle ist dadurch meist gegen eine weitere Superinfektion durch einen Bakteriophagen der gleichen Art geschützt. Vor dem Absterben der Wirtszelle wird die Replikation des Phagen induziert. Erfolgt das Ausschneiden der Phagen-DNA aus dem bakteriellen Genom unpräzise, kann ein Teil der Wirts-DNA mit ausgeschnitten werden. Diese wird bei einer erneuten Infektion auf einen anderen Bakterien-Stamm übertragen¹⁵⁶.

1.2.3. Gentransfer durch Konjugation

Bei der Konjugation ist im Gegensatz zur Transformation und zur Transduktion ein Zellkontakt zwischen Donor und Rezipient notwendig. Typische Vehikel sind selbst-übertragbare (konjugative) Plasmide oder mobilisierbare Plasmide und konjugative Transposons¹²⁶. Das benötigte Paarungssystem, wie z.B. der F-Faktor für die Ausbildung von Sexpili von *E. coli*, kann sowohl im Genom, als auch auf Plasmiden kodiert sein.

Die Konjugation kommt insbesondere bei gramnegativen Stäbchen (*Enterobacteriaceae*), bei Staphylokokken und Enterokokken vor¹⁹. Beispiele für gut charakterisierte Virulenzplasmide, die in der Medizin eine große Bedeutung haben, sind das EAF-Plasmid beim enteropathogenen *E. coli* (EPEC)¹⁶⁶, das Plasmid pO157 des enterohämorrhagischen *E. coli* O157:H7²⁴ und die Virulenzplasmide von *Yersinia pestis* (pPCP1, pCD1, pMT1)^{67, 132}. Die erwähnten Plasmide sind nicht-

konjugativ und werden mit Hilfe konjugativer Plasmide mobilisiert. Konjugativer Gentransfer wurde vor allem auch bei Plasmiden untersucht, die Antibiotikaresistenzen kodieren³⁴.

Generell ist der Austausch von Plasmiden eine wichtige Quelle für Genom-Variationen.

1.3. Genomische Inseln (GI) / Pathogenitätsinseln (PAI)

Im Genom vieler Bakterien wurden große horizontal transferierte Regionen identifiziert, die charakteristische strukturelle Gemeinsamkeiten aufweisen. Diese Genomischen Inseln (*genomic island* = GI) werden je nach Ausstattung, Vorkommen und Nutzen für den Rezipienten als Fitness-Inseln, Symbiose-Inseln, Metabolische Inseln und Pathogenitätsinseln (*pathogenicity island* = PAI) beschrieben und sind durch die folgenden Charakteristika gekennzeichnet^{27, 44, 58}:

1. Die durchschnittliche Größe der GIs liegt zwischen 10 und 200 kb.
2. GIs weisen den G+C-Gehalt und die Aminosäurecodon-Bevorzugung des Ursprungstammes auf und lassen sich dadurch identifizieren.
3. GIs werden oft von '*direct repeats*' (DR) begrenzt. Dies sind identische DNA-Sequenzen, deren Größe zwischen 9 und 135 bp liegt.
4. GIs sind häufig neben tRNA-Genen lokalisiert. Die tRNA Gene derselben Aminosäure sind bei den unterschiedlichen Bakterien-Spezies hochkonserviert und eignen sich als Integrationsstellen von Fremd-DNA. Eine Reihe von Phagen, wie P4 von *E. coli*, P22 von *Salmonella enterica* und HP1 von *Haemophilus influenzae* inserieren in oder neben tRNA-Genen.
5. GIs enthalten oft funktionelle Gene, die beispielsweise für Mobilitätsfaktoren kodieren, wie Insertions-Elemente (IS), Integrasen und Transposasen. So sind viele Inseln nicht nur transferier- und mobilisierbar, sondern unterliegen auch einer Instabilität. Über homologe Rekombination kann es zu dem Verlust der gesamten Insel oder von Teilbereichen derselben kommen.
6. Es wird vermutet, dass es sich um zu unterschiedlichen Zeitpunkten, von verschiedenen Stämmen erworbene DNA handelt, die sich an bestimmten Regionen im Chromosom sammelt ('*hotspots*') und die weiteren Umstrukturierungen über IS-Elemente unterliegt. IS-Elemente fungieren als neuer Integrationspunkt für DNA und als Werkzeuge der Rekombination mit anderen IS-Elementen im Genom. Daraus entsteht die mosaikartige Struktur der GIs.

Von besonderem Interesse für die medizinische Mikrobiologie sind die Pathogenitätsinseln ('*pathogenicity Island*', PAI), die neben den typischen Charakteristika einer Genomischen Insel, auch virulenzassoziierte Determinanten beinhalten, darunter Toxine, Adhäsine, Sekretionssysteme und Eisenaufnahmesysteme^{57, 105}. PAIs kommen im Genom pathogener Stämme vor und nicht, oder nur selten, bei apathogenen Stämmen derselben oder nahe verwandter Spezies⁶⁰.

Trotzdem sie weit verbreitet sind, weiß man bislang wenig über die genauen Mechanismen des Transfers Genomischer Inseln.

Pathogenitätsinseln wurden sowohl bei grampositiven als auch bei gramnegativen Stämmen beschrieben und mittlerweile bei nahezu allen humanpathogenen Bakterien nachgewiesen^{58, 105}. PAIs verleihen pathogenen Bakterien die Fähigkeit, die angeborenen Abwehrmechanismen des Körpers zu umgehen¹⁸¹. Eisenaufnahmesysteme unterschiedlicher Art sind unter den Virulenzfaktoren, die auf Pathogenitätsinseln gefunden wurden, von großer Bedeutung. So besitzt *Shigella flexneri* das Eisen (III)-Dicitrat – Aufnahmesystem (Fec), das auf der ~66 kb großen SRL PAI liegt⁹⁷. Der Pit-Eisenaufnahmetransporter bei Pneumokokken ist auf der Pathogenitätsinsel PPI-1 kodiert und ein Beispiel bei grampositiven Erregern¹⁸.

Bei verschiedenen klinischen Infektionen spielen Eisen-Ionen als Kofaktoren metabolischer Schlüsselenzyme eine wichtige und kritische Rolle^{23, 107}. Bei einigen Erregern konnte gezeigt werden, dass deren Pathogenität direkt mit der Eisenaufnahme korreliert^{133, 167}. Die chemische Löslichkeit von Eisen ist bei neutralem pH-Wert und unter aeroben Bedingungen sehr gering⁸⁹. Durch körpereigene Chelatoren wird sie in den extrazellulären Flüssigkeiten des menschlichen Körpers weiter vermindert, so dass dort bakterielles Leben nahezu unmöglich wird¹⁰⁷. Insbesondere pathogene Bakterien haben daher verschiedene Eisenaufnahme-Transportsysteme entwickelt, um freies, aber auch an Wirtproteine wie Ferritin, Transferrin, Lactoferrin, Hämoglobin und Myoglobin gebundenes Eisen aufzunehmen¹⁷². Unter anderem wird Eisen über niedermolekulare Eisenchelatoren, so genannte Siderophore, komplexiert und anschließend in die Bakterienzelle aufgenommen^{37, 107}.

1.4. *Yersinia*-HPI, Bedeutung und Struktur

Ein typisches Beispiel für ein siderophorvermitteltes Eisenaufnahmesystem ist das Yersiniabactin-System, das durch die erstmals bei humanpathogenen *Yersinien* entdeckte Pathogenitätsinsel *Yersinia* High-Pathogenicity Island kodiert wird (*Yersinia*-HPI)⁶⁶. Unter eisendepletierten Bedingungen führt das Vorhandensein der *Yersinia*-HPI zur Aufnahme von Eisenionen über das synthetisierte Siderophor Yersiniabactin²⁶. Die Inaktivierung der HPI-Gene führt im Mausversuch zu einer verminderten Virulenz und bestätigt ein Einfluss der HPI mit dem Siderophor-System für die Expression eines hochpathogenen Phänotyps^{35, 137}. Gleichsam konnte durch Übertragung der *Yersinia*-HPI-Gene auf einen gering-pathogenen Stamm eine erhöhte Virulenz beim Rezipienten-Stamm erzielt werden¹²⁵.

Die *Yersinia*-HPI ist ein mobiles genetisches Element. Das 30-45 kb große Segment enthält eine Kernregion ('core'), die 11 Gene zur Synthese (*irp1-5*, *irp9*), für den Transport (*irp6*, *irp7*) und die Aufnahme (*fyuA*) des Yersiniabactin-Siderophors (Ybt) umfasst, sowie ein Regulator-Gen (*ybtA*)^{17, 22, 27, 128, 129}.

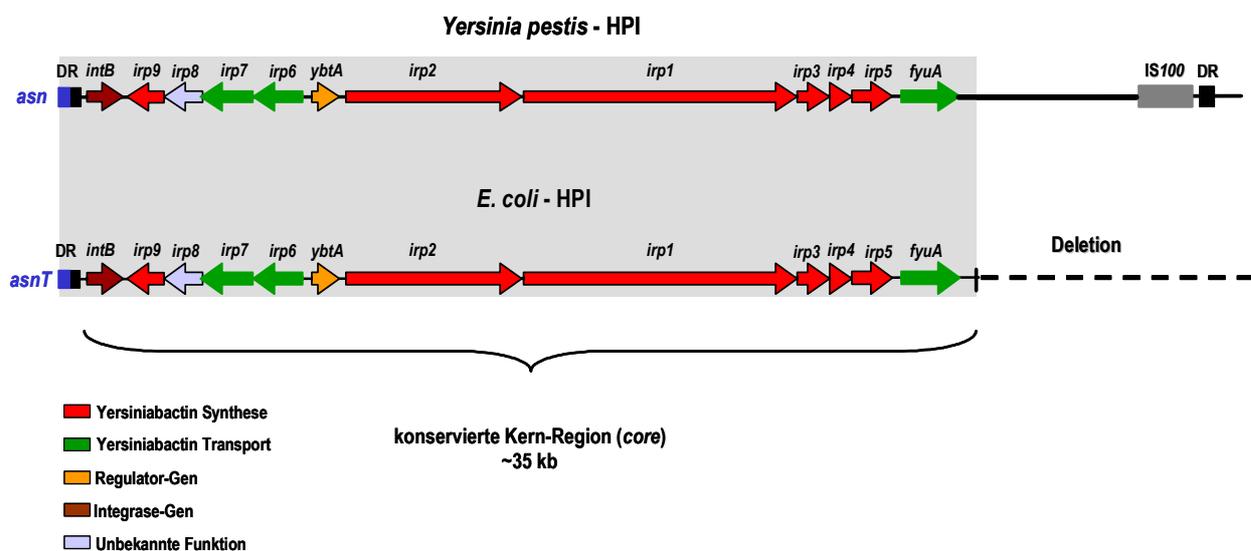


Abbildung 1: Struktur der HPI in *Yersinia pestis* und *E. coli* mit Kern-Region ('core') und variablem, wirtsspezifischem Bereich. Die Funktion der farblich dargestellten Gene ist beschrieben. Mit DR ('direct repeat') sind die Wiederholungssequenzen beschrieben, die typischerweise eine Genomische Insel umranden. Bei *E. coli* fehlt der Abschnitt stromabwärts von *fyuA*, inklusive dem IS-Element und dem zweiten DR.

Neben diesem etwa 30 kb großen konservierten und funktionellen Anteil der HPI gibt es noch einen wirtsspezifischen Bereich, der variieren kann ¹⁴⁹ (Abbildung 1).

Die *Yersinia*-HPI erfüllt die Kriterien einer Genomischen Insel, mit unterschiedlichem G+C-Gehalt und der Insertion in einem tRNA-Gen (*asn*-tRNA = Asparaginsäure-tRNA). Sie trägt ein Integrase-Gen, ähnlich dem Bakteriophagen P4 und wird zum umgebenden Chromosom durch kleine Wiederholungssequenzen (DR = 'direct repeats'), die eine Duplikation des 3'-Endes des *asn*-tRNA Gens darstellen, abgegrenzt ⁶⁰.

Die mobile Form der HPI ist ein zirkuläres DNA-Molekül, eine nicht-replikative Funktionseinheit, die über einen sequenzspezifischen Integrations-Mechanismus zur Interaktion mit konservierten Zielsequenzen in der Lage ist ¹³⁸ (Abbildung 2). Es konnte an verschiedenen Beispielen gezeigt werden, dass Genomische Inseln aus dem Chromosom desintegrieren und wieder reintegrieren, sowie auf andere Stämme übertragen werden können ^{45, 134}. Vermittelt wird die Rekombination durch die HPI-Integrase, eine 'site'-spezifische Rekombinase, aus der Gruppe der Thyrosin-Kinasen. Die HPI integriert ins Wirtsgenom, indem die Integrase mit zwei DNA-Sequenzpaaren interagiert, *attB* und *attP*, in Analogie zu den gut untersuchten 'attachment sites' bei der Rekombination von lambdoiden Phagen ^{62, 69}. Erstere (*attB*) ist die chromosomale Erkennungssequenz (BOB'), im Fall der HPI aus dem Bereich von *asn*-tRNA-Genen. Die zweite Sequenz (*attP*) ist auf der Pathogenitätsinsel selbst lokalisiert (POP') (Abbildung 2).

Neben einer chromosomalen Insertion ist prinzipiell auch das Einfügen in ein Plasmid möglich, da tRNA-Gene bei *Salmonella* beispielsweise auch auf Plasmiden vorkommen ¹²¹.

Der horizontale Gentransfer der HPI gelang der Arbeitsgruppe um Lesic und Carniel bei verschiedenen *Y. pestis*- und *Y. pseudotuberculosis*-Serotypen ⁹⁴.

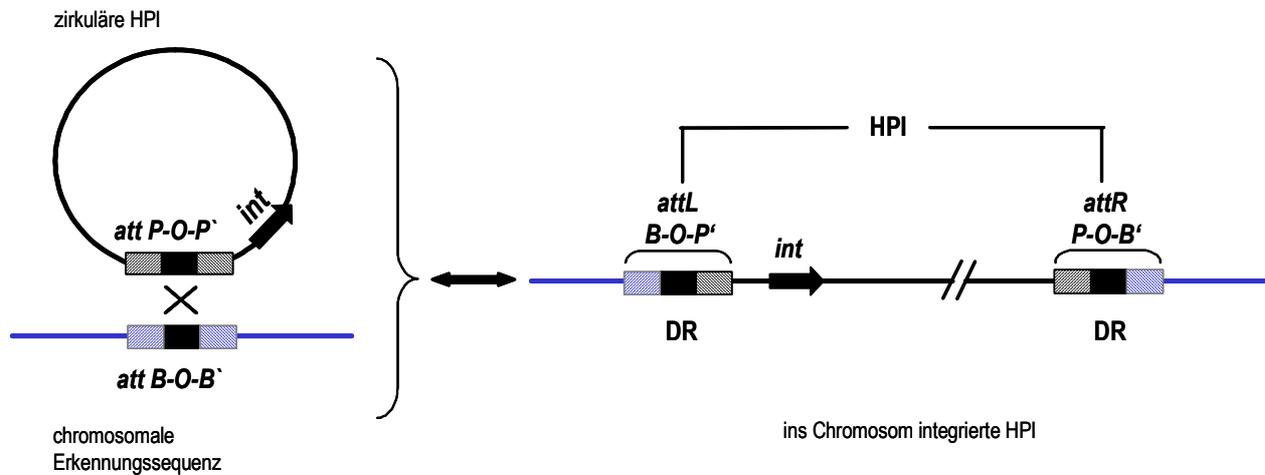


Abbildung 2: Vereinfachtes Modell der *site*-spezifischen Rekombination am Beispiel der HPI. B-O-B' repräsentiert die chromosomale Erkennungssequenz für die zirkuläre Form der HPI. Dabei stellt die Region O das 3'-Ende der tRNA und einen Teil des 'direct repeat' dar. Die entsprechende 'attachment site' *att* auf der HPI ist P-O-P'. Nach der Rekombination ist die HPI zwischen zwei neu rekombinierten attachments *attL* und *attR* integriert.

1.5. Die Verbreitung der HPI und Struktur bei *E. coli*

Die HPI konnte bei pathogenen, vereinzelt aber auch bei nicht-pathogenen Vertretern der Familie der *Enterobacteriaceae* nachgewiesen werden. Sie ist insbesondere bei pathogenen Isolaten von *Citrobacter* sp., *Klebsiella* sp., *Enterobacter* sp., *Salmonella enterica* non-I Serotyp, *Serratia liquefaciens* und bei *E. coli* verbreitet^{7, 88, 146}.

Interessanterweise wird die HPI unter den Pathovaren von *E. coli* vor allem bei enteroaggregativen (EAEC) und extraintestinal-pathogenen *E. coli* (ExPEC) angetroffen^{77, 151, 152}. Die HPI ist in *E. coli* hochkonserviert, mit einer Sequenzidentität zwischen 98% und 100% zur Kernregion der HPI von *Y. pestis*¹³⁹. Bei einem Großteil der Isolate wurde zudem Ybt nachgewiesen, was für die erhaltene Funktionalität der Insel spricht¹⁴⁶. Im Gegensatz zu der HPI der Yersinien fehlt in *E. coli* typischerweise das 3'-Ende der Pathogenitätsinsel, jenseits des Genabschnitts *fyuA*¹⁴⁹ (Abbildung 1).

Die *Yersinia*-HPI spielt eine große Rolle bei der Ausbreitung extraintestinaler Infektionen. Ihr Vorhandensein erwies sich als höchster Prädiktor für extraintestinale Virulenz im Vergleich zu sonstigen Virulenzfaktoren wie Fimbrien, Invasine und andere siderophorvermittelte Eisenaufnahmesysteme¹⁴⁸.

ExPEC-Stämme sind in etwa 90% der Fälle die Erreger bei urogenitalen Infektionen der Frau, der weltweit häufigsten bakteriellen Entzündung⁵⁴. Aber auch die Cholezystitis, Appendizitis, Peritonitis, postoperative Wundinfekte und Sepsis zählen zu *E. coli*-bedingten Infektionen¹⁴³, die meist über die normale Darmflora durch Kontamination, Schmierinfektionen, oder Durchwanderungsinfektion bei Schrankenstörungen entstehen. ExPEC Stämme stellen etwa 20% der fäkalen *E. coli* Isolaten dar. Sie besiedeln den Darm ohne durch gastrointestinale Symptome aufzufallen⁷⁴ und besitzen ein weites Spektrum an Virulenzfaktoren, wie Adhäsine, Toxine und Eisenaufnahmesysteme, die zu extraintestinaler Pathogenität führen^{14, 75}.

Um die Verbreitung und die Struktur der HPI analysieren zu können, arbeitet unsere Arbeitsgruppe mit einer phylogenetischen Stamm-Sammlung von *E. coli*, die ECOR-Sammlung (*E. coli* Reference Collection). Howard Ochmann und K. Selander erstellten diese 1983 aus unterschiedlichsten Isolaten, nach Auswahl mittels '*multilocus enzyme electrophoresis*', um eine möglichst breite Genomvariation einer Spezies darstellen zu können¹¹⁷. Die 72 Stämme aus einer Auswahl von 2600 *E. coli*, von Menschen und Tieren aus unterschiedlichen Kontinenten, weisen untereinander die größtmögliche Diversität an 11 Enzym-loci auf und unterscheiden sich ebenso hinsichtlich ihrer Pathogenität.

Die *Yersinia*-HPI lässt sich vorwiegend bei den Stämmen nachweisen, die neben einer Akkumulation weiterer Virulenzfaktoren und Pathogenitätsinseln, sowohl die *E. coli*-Stämme mit den größten Genomen als auch der höchsten Pathogenität sind und häufig extraintestinale Infektionen verursachen¹⁴⁸.

1.6. Zielsetzung der Arbeit

Der horizontale Gentransfer ist für den größten Anteil am flexiblen Genpool verantwortlich. Die übertragenen Gene können neue Stoffwechsel-Eigenschaften oder Virulenzfaktoren vermitteln und die Lebensweise und die Pathogenität eines Stammes verändern.

Mit strukturellen und biomathematischen Untersuchungen kann horizontal transferierte DNA aufgespürt werden. Bei der Integration fremder DNA werden bestimmte Lokalisationen im Genom besonders bevorzugt und daher als Insertions-*'hotspots'* bezeichnet. Hierzu gehören beispielsweise tRNA-Gene, die als Integrationsort für viele bekannte Genomische Inseln dienen¹⁴¹.

Die Annahme liegt nahe, dass in der Umgebung von tRNA-Genen horizontal transferierte DNA gefunden werden kann. Um diese Hypothese zu überprüfen, wurde in Vorarbeiten bei einer großen Auswahl verschiedener *E. coli*-Stämme der Insertionsort der horizontal übertragenen *Yersinia*-Pathogenitätsinsel (HPI) untersucht. Die HPI, ein Prototyp Genomischer Inseln, ist im Normalfall in ein tRNA-Gen inseriert. Das auf der Insel kodierte Siderophor-Eisenaufnahmesystem ist ein wichtiger Virulenzfaktor^{125, 137}.

Bei allen untersuchten Stämmen, bis auf eine Ausnahme, ist die HPI in das Genom integriert, ohne Nachweis horizontal transferierter DNA in der Umgebung des tRNA-Gens. Lediglich beim *E. coli*-Stamm ECOR31 fehlt das tRNA-Gen als typischer Insertionsort. Es stellt sich die Frage, inwieweit die HPI unabhängig von tRNA-Genen auch in andere Strukturen inserieren kann?

Die Region, die der *Yersinia*-HPI in *E. coli* ECOR31 angrenzt, wird RegX genannt und soll strukturell, inhaltlich (Annotationen) sowie bezüglich ihrer Verbreitung unter gramnegativen Bakterien untersucht werden.

Das Ziel der vorliegenden Arbeit ist

1. die Untersuchung des atypischen Insertionsortes der *Yersinia*-HPI im *E. coli*-Stamm ECOR31.
2. die strukturelle Charakterisierung der angrenzenden, möglicherweise horizontal transferierten DNA-Region (RegX) in ECOR31 und die Beschreibung der aktuellen Annotationen in den Datenbanken zu RegX.
3. die Untersuchung der Prävalenz der Region, sowie der vorliegenden atypischen Insertion der HPI, bei gramnegativen Bakterien.

2. Material und Methoden

2.1. Material

2.1.1. Geräte

Brutschrank	WTC Binder
Elektrophoresezubehör	Horizontal-Plexiglaskammern: aus eigener Werkstatt des Max von Pettenkofer-Instituts
	Spannungsquelle: Bio-Rad [®] , Power Pac 200/300
Hybridisierungsöfen	Stratagene Personal Hyb
Kühlzentrifuge	Sorvall: Rotor SL-50
PCR-Thermocycler	Eppendorf: Mastercycler Gradient [®]
	Applied Biosystems: GeneAmp [®] PCR System 9700
Photometer	Pharmacia: Ultraspec [®] Plus
Röntgenfilm-Exponierkassetten	CawoX-omatic, Kodak
Röntgenfilm-Entwicklungsgerät	Fujifilm FPM 100A
Schüttelinkubator groß	Certomat [®] H, B.Braun Biotech International
	New Brunswick Scientific Rolodrum
Schüttelinkubator klein	Eppendorf Thermomixer Komfort
Tischinkubator	Unitec HB-130
Tischzentrifuge	Eppendorf centrifuge 5417 C
Tischzentrifuge, gekühlt	Eppendorf centrifuge 5417 R
Vakuumblot-Gerät	Pharmacia LKB 2016 VacuGene
Vakuum-Zentrifuge	Speedvac DNA 110/Savant
Elektroporationsgerät	Gene Pulser [®] Transfection Apparatus BIO RAD
UV-Crosslinker	Bachhofer
pH-Meter	Mettler, Toledo 320 pH-Meter
Sterilwerkbank (Flow)	Hereaus, Herasafe HS12
Transiluminator	Herolab, UVT-20M/W
Videophotoanlage	Herolab, E.A.S.Y. (Enhanced Analysis System)
Sequenzierer	ABI Prism; Perkin Elmer, Model 373A und 377

2.1.2. Chemikalien und Kleinmaterialien

Firma	Material
Amersham	Hybond N+ positively charged Nylon Membrane
Schleicher & Schuell	PROTRAN BA 85 Cellulosenitrat-Transfer-Membranen, Porengröße 0.45 μm , 132 mm Whatman Paper
Invitrogen life technologies	<i>Bam</i> HI, <i>Eco</i> RI, , <i>Xba</i> I (10 U/ μl), <i>Not</i> I (15 U/ μl)
GIBCO BRL	<i>Cl</i> aI, <i>Bg</i> II, <i>Sma</i> I, <i>Hind</i> III, <i>Pst</i> I (10 U/ μl)
MBI Fermentas	<i>Acc</i> 65I (10 U/ μl), <i>Psu</i> I = <i>Xho</i> II (5 U/ μl) GeneRuler™ 100 bp DNA Ladder Plus, 0,5 mg DNA/ml GeneRuler™ 1 kb DNA Ladder, 0,5 mg DNA/ml
Boehringer Mannheim GmbH	<i>Sau</i> 3A I (4 U/ μl)
TAKARA SHUZO CO., LTD	LA Taq (5 U/ μl)
PAN Biotech GmbH	Panscript Polymerase, dNTP Solutions
Carl Roth GmbH + Co	Agarose NEEO
INC Biomedicals, Inc.	Glycerol
ABI Applied Biosystems	Big Dye-Sequenzierreagenz
Promega GmbH	RNA-free DNase (1 U/ μl)

Antibiotika

Antibiotikum	Endkonz. [$\mu\text{g/ml}$]	Firma
Kanamycin	50	Serva Electrophoresis GmbH
Ampicillin	100	SIGMA CHEMICAL CO.
Trimethoprim	75	SIGMA CHEMICAL CO.

2.1.3. Nährmedien

LB (Luria Bertani)-Flüssig-Medium	10 g/l Bacto Trypton 5 g/l Bacto Hefeextrakt 10 g/l NaCl → 1 l Aqua demin., pH 7,5
LB-Agar	LB-Medium 15 g/l Bacto Agar, pH 7.5
NB-Medium	8 g/l Nutrient Broth 5 g/l NaCl
NBD-Medium (Dipyridyl-Medium)	NB Medium + 200 µM α-α'-Dipyridyl
SOC-Medium	0,5 g NaCl 5 g Bacto Hefeextrakt 20 g Bacto Trypton 2,5 mM KCl 20 mM Glukose → 1 l Aqua demin., pH 7,0

Columbia Blutagar und Müller Hinton-Agar wurden als Fertignährböden von der Firma Becton Dickinson bezogen.

2.1.4. Stämme

Stamm (<i>E. coli</i>)	Eigenschaft	Referenz
ECOR31	HPI positiv, Antibiotikaresistenzen: str ^r , spc ^r , sul ^r , ap ^r , kn ^r , tet ^r	104, 117
DH 5α	endA1 hsdR17 (rK-mK+) supE44 THI-1 recA1 gyrA (NalR) relA1 Δ(lacZYA- argF)U169 (φ80lacZdM15)	GIBCO BRL®
DH10B	F- mcrA Δ(mrr-hsdRMS-mcrBC) φ80ΔlacZΔM15 ΔlacX74 deoR recA1 endA1 araΔ139 D(ara, leu)7697 galU galK1- rpsL nupG	GIBCO BRL®
XL1-Blue MR	Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac	Stratagene®

Stämme für das Screening nach der DNA-Region 'RegX'

Stamm	Gesamt-Anzahl	Blut-Kultur	Urin-Kultur	Sonstige
<i>Enterobacteriaceae</i>				
<i>Escherichia coli</i>	111	50	60	1
<i>Escherichia coli</i> (ECOR) ¹⁾	72			
<i>Klebsiella pneumoniae</i>	59	24	13	22
<i>Enterobacter cloacae</i>	38	6	9	23
<i>Enterobacter species</i>	5			5
<i>Klebsiella oxytoca</i>	23	14	4	5
<i>Proteus mirabilis</i>	17	6	9	2
<i>Yersinia enterocolitica</i> ²⁾	15			
<i>Yersinia pseudotuberculosis</i> ²⁾	10			
<i>Yersinia pestis</i> ²⁾	10			
<i>Citrobacter koseri</i>	14	4	2	8
<i>Citrobacter freundii</i>	20	16	2	2
<i>Serratia marcescens</i>	24	8	3	13
<i>Serratia species</i>	11			5 + 6 ³⁾
<i>Morganella morganii</i>	8	6		2
<i>Pseudomonaceae</i>				
<i>Pseudomonas aeruginosa</i>	77	1	9	67
<i>Pseudomonas species</i>	16	4		12 ³⁾

alle nicht gekennzeichneten Stämme wurden aus der Diagnostik des Max von Pettenkofer-Institutes bezogen

¹⁾ ECOR Stammsammlung, von Ochman, H. und Selander, R.K.¹¹⁷

²⁾ chromosomale DNA dieser Stämme wurde freundlicherweise von Dr. A. Rakin, Max von Pettenkofer-Institut München, bereitgestellt

³⁾ die Stämme wurden freundlicherweise von Dr. G. Berg, Institut für molekulare Physiologie und Biotechnologie, Abteilung Mikrobiologie, der Universität Rostock, bereitgestellt

2.1.5. Vektoren

Vektor	Replikon	Marker	Referenz
pCR [®] -XL-TOPO [®]	ColE1	Zeo ^r , Kn ^r , <i>lacZα</i> , <i>ccdB</i> , <i>Plac</i>	Invitrogen [™] life technologies
SuperCos1	ColE1, SV40	Neo ^r , Ap ^r , <i>cos</i>	Stratagene [®]
pBluescript II KS+	ColE1, f1+ ori	Ap ^r , <i>lacZ</i>	Stratagene [®]

2.1.6. Plasmide und Cosmide

Klon	Vektor	Insertion	Referenz
pBP-LA Klone aus der Cosmidbank	TOPO	4,1 kb PCR-Fragment, TA-cloning	diese Arbeit
p17-A7	SuperCos1	ECOR31-DNA-Fragment, <i>Bam</i> HI	diese Arbeit
p17-D6	SuperCos1	ECOR31-DNA-Fragment, <i>Bam</i> HI	diese Arbeit
p14-F1	SuperCos1	ECOR31-DNA-Fragment, <i>Bam</i> HI	diese Arbeit

2.1.7. Synthetische Oligonukleotide

Name	Verwendung	Primersequenz	Schmelz-Temp. (°C)
UP HPi comp 1	= S2; LA-PCR	5'ATCAGACAGGGAGATGGCAGGATAGGC-3'	79°C
UP HPi comp 2	= S1; LA-PCR	5'TACGGATACCTTCACGCTGCTGTCGCGC-3'	77°C
LA-PCR, C1	LA-PCR	5'GTACATATTGTCGTTAGAACGCGTAATACGACTCA-3'	93°C
LA-PCR, C2	LA-PCR	5'CGTTAGAACGCGTAATACGACTCACTATAGGGAGA-3'	97°C
RegX-521.for	Screening	5'-TTGGGCTGACTGGTATTAGTGGC-3'	62,4°C
RegX-358.for	Screening	5'-CGAGTTTTTCCACCGCATCAACG-3'	62,4°C
primer island-PI(+).for	Sequenzierung	5'-CAGGACATTGGATGCTGAGAATTCG-3'	63°C
primer island-PI(-).rev	Sequenzierung	5'-CAGGAGCCGTCTATCCTGCTTGC-3'	66°C
T3	Seq SuperCos	5'-ATTAACCCTCACTAAAGGGA-3'	45,5°C
T7	Seq SuperCos	5'-GTAATACGACTCACTATAGGGC-3'	45,5°C

Name	Verwendung	Primersequenz	Schmelz-Temp. (°C)
M13-20	Sequenzierung	5'-GTAAAACGACGGCCAGT-3'	47°C
M13-rev	Sequenzierung	5'-GGAAACAGCTATGACCATG-3'	54,5°C
RegX.5781.for	Screening <i>K. pneum.</i> RT-PCR	5'-TCGCAAGGAAGCATGAAAACAGA-3'	58,9°C
E31-RegX-5.for	Screening <i>K. pneum.</i> RT-PCR	5'-GACGCCCGACCTGGTATTGCTTGATGAACC CAC-3'	73°C
17-A7.13949.for	Screening <i>K. pneum.</i>	5'-GCGCTGACGGAAGTTATAGAT-3'	57,9°C
17-A7.14884.rev	Screening <i>K. pneum.</i>	5'-TGGCGGCGCTTCTTCGTCGTA-3'	63,7°C
17-A7.25773.for	Screening <i>K. pneum.</i>	5'-TCACTGAGCTATGTTGGGATA-3'	55,9°C
17A7Cont23.128. rev	Screening <i>K. pneum.</i>	5'-GCCATAAATCCGCTGTAGCTG-3'	59,8°C
17-A7.19742.rev	Screening <i>K. pneum.</i>	5'-CACTGGGGTATGACCTGATGG-3'	61,8°C
17-A7.Con11.for	Screening <i>K. pneum.</i>	5'-ATTGCGCCAGTGTACTCAC-3'	57,9°C
17-A7.3094.for	Screening <i>K. pneum.</i>	5'-CCGAAATGGTGGTGAGAAAG-3'	59,8°C
17-A7.3500.rev	Screening <i>K. pneum.</i>	5'-ACGGCGACTGAAAGAAT-3'	50,4°C
17-A7.3852.for	Screening <i>K. pneum.</i>	5'-GGACGGCATCACTGACGAA-3'	58,8°C
17-A7.4416.rev	Screening <i>K. pneum.</i>	5'-AACTGGGCTACTGTATTGGCA-3'	57,9°C
RegX-6089.rev	RT-PCR	5'-CGCAACGCCTTTCCTGTATGG-3'	61,8°C
RegX-ORF1-RT.for	RT-PCR	5'-GAAGTGACGCCGGATACGCTAA-3'	62,1°C
RegX-ORF1-RT.rev	RT-PCR	5'-CCTCAATTATGGCCGTAGGCATC-3'	62,4°C
RegX-ORF4-RT.for	RT-PCR	5'-CATATTCTGAACAACCGGGAAG-3'	58,4°C
RegX-ORF4-RT.rev	RT-PCR	5'-GCATCCACCACAGTGACCAT-3'	59,4°C
Membr.prot.rev	RT-PCR	5'-CGCCCAGTCCGTGAAA-3'	54,3°C
sitA.for	RT-PCR	5'-GAACCGCTGCGTAAACAAAT-3'	55,3°C
sitB.for	RT-PCR	5'-GCGTAGAGCGGGTAAAGAAGA-3'	59,8°C
sitB.rev	RT-PCR	5'-AGGCATCCCGAGAAGTGAAAT-3'	57,9°C
sitC.for	RT-PCR	5'-TCTGTATTGCGGTGACCATAG-3'	57,9°C
sitC.rev	RT-PCR	5'-CCTCCGACCATTGCCGAAATC-3'	61,8°C
sitD.for	RT-PCR	5'-CTTACTGTTGTCGCCACGCTA-3'	59,8°C
sitD.rev	RT-PCR	5'-GCCCAACCTTTCAGCACGA-3'	58,8°C
RegX.ORF2- 8117.for	RT-PCR	5'-GAAATACGGCCTGTAAAGCAC-3'	57,9°C
RegX.ORF3- 9413.rev	RT-PCR	5'-TTTGCACTGACTTCGGACA-3'	54,5°C

2.2. Methoden

2.2.1. Isolierung chromosomaler DNA

Protokoll: *PureGene DNA Purification Kit*[®] der Firma Genra Systems, Minneapolis, USA

Durchführung: 1,5 ml einer Übernachtskultur wurden abzentrifugiert und das Pellet in 600 µl *Cell Lysis Solution* aufgenommen und 5 min bei 80°C inkubiert. Das Lysat wurde dann mit 3 µl RNase (Konz: 1-10 mg/ml) versetzt und nach vorsichtigem Mischen für 45 min bei 37°C gelagert.

Die noch enthaltenen Proteine wurden durch Zugabe von 200 µl *Protein Precipitation Solution* und starkem Schütteln gefällt und durch Zentrifugieren am Gefäßboden gesammelt. Der die DNA enthaltende Überstand konnte mit 100%-igem Isopropanol gefällt und mit 70%-igem Ethanol gewaschen werden. In einem letzten Schritt wurde die DNA in 100 µl *DNA Hydratation Solution* resuspendiert.

2.2.2. Isolierung chromosomaler DNA mit Phenol

Für die Erstellung der Genbank wurde die chromosomale DNA mit einer Methode extrahiert, die eine größere Ausbeute gewährleistet. Die Bakterienzellen werden mit Lysozym aufgelöst und zum Abbau der Zellproteine mit Proteinase K versetzt. Durch Phenolextraktion werden die Zellwand- und Zellmembranbestandteile entfernt und die DNA kann anschließend durch einfache Ethanol-fällung aus dem Lysat extrahiert und in Aqua dest. resuspendiert werden. Zur Entfernung der RNA-Reste muss eine Behandlung mit RNase durchgeführt werden.

Durchführung: Das Bakteriensediment einer 40 ml Bakterienkultur (OD_{600} 0,8) wurde in 5 ml TE-Puffer aufgenommen, mit 50 µl RNase Stammlösung (10 mg/ml) versetzt und eine Stunde bei -20°C gelagert. Zu der gefrorenen Bakteriensuspension wurde eine Spatelspitze Lysozym gegeben und 1h bei 37°C inkubiert. Anschließend wurden

1 ml STEP-Lösung und Proteinase K in einer Endkonzentration von 1 mg/ml hinzugefügt und die Lösung wurde im Wasserbad bei 50°C ca. 12 h inkubiert. Die Aufreinigung erfolgte nach Zugabe von 6 ml TRIS-gepufferten Phenols und 20 ml TE-Puffer. Nach vorsichtigem Schütteln und Abzentrifugieren (20 min bei 10000 rpm) wurde die DNA aus der wässrigen Phase durch Ethanol (2 Volumen absolutes Ethanol und 0,1 Volumen 3 M NaAcetat pH 5,0) gefällt und ließ sich auf Glasstäbe wickeln. Nach der weiteren Aufreinigung, in Chloroform-Isoamylalkohol (24:1) und 70% Ethanol, wurde die DNA nach Lufttrocknung in sterilem H₂O aufgenommen.

Reagenzien:

1x TE-Puffer, pH 8,0:	50 mM TRIS/HCl 50 mM EDTA
STEP-Lösung, pH 8,0:	500 µl 10% SDS 500 µl 1 M TRIS/HCl 8 ml 0,5 M EDTA 1 ml steriles H ₂ O

2.2.3. Schnellisolierung von Plasmid-DNA

Durch ein alkalisches Milieu werden bakterielle Zellwandproteine und die DNA denaturiert. Nur der geschlossenen, zirkulären Plasmid-DNA gelingt eine Renaturierung nach Neutralisation der Lösung. Die chromosomale DNA bleibt unlöslich und kann abzentrifugiert werden. Die im Überstand verbliebene Plasmid-DNA wird an eine Glasmilch-Matrix gebunden und mit Ethanol gewaschen. Nach Lufttrocknung wird die DNA in H₂O aus der Matrix gelöst.

Protokoll: *Miniprep Express*TM Protokoll der Firma Qbiogene, Inc, Carlsbad, USA

Durchführung: 1 ml einer Übernachtskultur der Bakterien wurde abzentrifugiert und das Sediment in 100 µl *Lösung I* resuspendiert, mit 200 µl *Lösung II* versetzt, und vorsichtig vermischt. Nach Zugabe von 150 µl *Lösung III* und vorsichtigem Schütteln wurde die Suspension 2 min zentrifugiert und der Überstand mit 400 µl *Miniprep*

*Express Matrix*TM vermischt. Die Glasmilch mit der daran gebundenen Plasmid-DNA wurde abzentrifugiert, mit 500 µl 80 % Ethanol gewaschen und nach erneutem Zentrifugieren und Trocknen schließlich in sterilem H₂O aufgenommen.

Reagenzien:

Lösung I:	50 mM Glukose 10 mM EDTA 25 mM Tris, pH 8,0
Lösung II:	0,2 N NaOH 1% SDS
Lösung III:	3 M Natriumacetat 3 M Essigsäure

2.2.4. Isolierung von Plasmid und Cosmid DNA

Die Isolierung großer Mengen reiner Plasmid- oder Cosmid-DNA erfolgte nach der Anleitung des *Nucleobond*[®] *AX-Kits* über Austauscher-Säulen.

Protokoll: *Nucleobond*[®] *AX-Kit* der Firma Macherey & Nagel, Düren, Deutschland

Durchführung: 100 ml einer über Nacht angewachsenen Bakterienkultur wurden zentrifugiert und in 4 ml *Puffer S1* aufgenommen. Nach Zugabe des gleichen Volumens von *Puffer S2* wurden die Lösungen vorsichtig vermischt und 5 min bei Raumtemperatur inkubiert. 4 ml *Puffer S3* wurden zugegeben und die Lösungen wurden vorsichtig durch Umdrehen des Gefäßes gemischt, bis sich eine homogene Suspension gebildet hatte. Nach 5-minütiger Lagerung auf Eis wurde die Suspension für 25 min bei 4°C und einer Geschwindigkeit von 10000 x g zentrifugiert. Die zwischenzeitlich mit 2 ml *N2 Puffer* equilibrierte NUCLEOBOND[®] AX 100 Säule wurde mit dem Überstand beladen. Nach zweimaligem Durchlauf, zur Erhöhung der Ausbeute, wurde die Säule mit 2 x 4 ml *N3 Puffer* gewaschen und schließlich die

Plasmide durch den *Puffer N5* eluiert. Die Plasmid-DNA wurde mit 70 % Isopropanol gefällt, mit 100%-igem Ethanol gewaschen und in Aqua dest. aufgenommen.

Die Isolierung der Cosmid-DNA erfolgte entsprechend dem obigen Protokoll, allerdings unter Einsatz des doppelten Volumens der *Puffer S1, S2, S3* und 3 x 4 ml des Waschpuffers *N3*.

2.2.5. Schneiden von DNA mit Restriktionsendonukleasen

Restriktionsendonukleasen spalten doppelsträngige DNA an einer für das jeweilige Restriktionsenzym spezifischen Basensequenz. Auf diese Weise lässt sich Proben-DNA fragmentieren und nach Schneiden größerer Mengen gleichartiger DNA können die Fragmente durch Gel-Elektrophorese dargestellt werden. Ebenfalls können mit dieser Methode Plasmide und Cosmide linearisiert werden. Dieser Vorgang ist Voraussetzung für eine spätere Ligation.

Durchführung: 1 µg DNA wurde mit 2-5 Units (U) der im Materialenteil aufgeführten Restriktionsenzymen und den dazugehörigen Puffern verdaut. Die geschnittene DNA wurde nach gelelektrophoretischer Auftragung mit Hilfe des *GeneClean II Kits* aus dem Gel aufgereinigt.

Partialverdau chromosomaler DNA

Zur Konstruktion einer Cosmid-Bank wurde chromosomale DNA des *E. coli* Stammes ECOR31 einem Partialverdau mit *Sau3A I* unterzogen. Dieses Enzym spaltet die Sequenz GATC und produziert einzelsträngige Überhänge von vier Basen. *Sau3A I* schneidet bei einem G/C-Gehalt von 50% statistisch alle 256 Basen. Durch Variation von Zeit und Menge des eingesetzten Enzyms, erhält man DNA-Fragmente in einem bestimmten Größenintervall. Zur Erstellung der Cosmid-Bank mit dem *SuperCos1 Cosmid Vector Kit* von Stratagene® sollten die DNA-Fragmente zwischen 35 kb und 42 kb groß sein.

2.2.6. Dephosphorylierung mit alkalischer Phosphatase (CIAP Behandlung)

Um nach einem Restriktionsverdau eine Selbstligation eines Vektors oder DNA-Fragmenten zu verhindern, wurden die Phosphatreste am 5'-Ende der Fragmente mit alkalischer Phosphatase aus Kälberdarm (*calf intestinal alkaline phosphatase* = CIAP) entfernt.

Durchführung: In einem Ansatz entsprechend einem Restriktionsverdau wurde die zu behandelnde DNA mit 0,01 U alkalischer Phosphatase pro pmol Enden und dem zugehörigen 1x Puffer eine Stunde bei 37°C inkubiert und anschließend mit 2 µl 0,5 M EDTA 20 min bei 65°C inaktiviert. Die pmol-Angabe wurde nach Anleitung der Firma Promega auf folgende Weise berechnet:

$$(\mu\text{g DNA} / \text{kb-Größe der DNA}) \times 3,04 = \text{pmol Enden}$$

Die dephosphorylierte DNA wurde vor der Weiterverwendung, z.B. für eine Ligation, durch Phenolextraktion oder Gelextraktion aufgereinigt.

2.2.7. Agarose Gel-Elektrophorese von DNA

Durchführung: Die Agarose wurde in dem für die Elektrophorese als Laufpuffer benutzten 1xTAE-Puffer aufgekocht und nach Abkühlung auf 55°C mit 0,5 µg/ml Ethidiumbromidlösung versetzt. Anschließend wurde die Agarose in horizontale Gelelektrophoresekammern gegossen. Ein, in diesem Zustand in die noch flüssige Agarose eingesetzter Gelkamm, bildet nach Erstarren der zwischen 0,4% und 2%-igen Agarose die Taschen für die DNA-Proben. Die mit 1/10 Volumen 10x Probenpuffer versetzten DNA-Lösungen wurden bei einer Feldstärke zwischen 0,6 – 3 V/cm aufgetrennt und anschließend unter UV-Licht fotografiert. Durch paralleles Auftragen eines DNA-Markers konnte die Größe der Fragmente ermittelt werden.

Reagenzien:

1x TAE-Puffer:	12,1 g/l Tris, 30 ml 96% Essigsäure/l, 18,6 g/l Na ₂ EDTA
Probenpuffer:	20% Ficoll 400 0,25% Bromphenol-Blau 1% SDS 0,1 M Na ₂ EDTA, pH 8

2.2.8. Elution von DNA-Fragmenten aus einem Agarosegel

Die Elution einzelner solitären Banden nach Auftrennung durch Gel-Elektrophorese erfolgte entsprechend dem Protokoll des *GENECLEAN II Kit* der Firma Qbiogene, Inc., Carlsbad, USA. Die hierdurch gereinigte DNA konnte für Ligationsexperimente eingesetzt werden.

2.2.9. Polymerase Kettenreaktion (PCR)

Die PCR dient der Amplifikation eines ausgewählten DNA-Abschnittes und kann sowohl von einem Gesamtzelllysat von Bakterien als auch von isolierter DNA als Ausgangsmaterial erfolgen.

Durchführung:Herstellen des bakteriellen Zelllysats ('gekochte Zellen'):

Eine Kolonie wurde in 400 µl Aqua dest. eingerieben und bei 100°C für 10 min erhitzt.

Isolierte DNA wurde für die PCR im Falle von chromosomaler DNA 1:100 und im Falle von Plasmid- bzw. Cosmid-DNA 1:50 verdünnt.

50 µl eines PCR-Ansatzes enthalten:

5 µl bakterielles Zelllysat bzw. verdünnte DNA
5 µl enzyspezifischer 10x Puffer
je 0,2 mM dATP, dCTP, dGTP, dTTP
1 µl forward Primer (Konz. 20 pmol/µl)
1 µl reverse Primer (Konz. 20 pmol/µl)
2-2,5 U Taq-Polymerase (Pan-Taq, LA-Taq)
Aqua bidest. ad 50 µl

Nach einer einmaligen Denaturierung der DNA bei 94° für 4 min erfolgt die Amplifikation des zwischen den beiden Primern liegenden DNA-Fragments in 35 Zyklen, beginnend mit einem ersten Denaturierungsschritt: 94°C für 45 sec. Es folgt für 45 sec. die Primerbindung ('*annealing*'), abhängig von ihrer Schmelztemperatur. Die DNA-Polymerisation findet für 0,5 – 2 min bei 72° statt. Die PCR endet nach 35 Zyklen und einem zusätzlichen Polymerisierungsschritt von 10 min. Die Ergebniskontrolle erfolgte durch Auftragen von ca. 1/10 des Ansatzes auf Agarose-Gele.

2.2.10. LA-Polymerase-Kettenreaktion

Mit Hilfe des *LA-PCR in vitro Cloning Kits* von Takara Shuzo Co., Ltd. lassen sich unbekannte DNA-Bereiche amplifizieren und klonieren, die einer bekannten Region benachbart sind. Das Prinzip beruht auf der Anwendung von Nukleotidkassetten und einer speziellen Polymerase (LA Taq), die in der Lage ist, selbst relativ lange Fragmente zu amplifizieren. Die so genannten Nukleotidkassetten sind doppelsträngige synthetische Oligonukleotide, die an einem Ende eine Restriktionsschnittstelle aufweisen und zwei Primerbindungsstellen für die Primer C1 und C2 beinhalten, letzterer etwas weiter stromabwärts gelegen als C1 (Abbildung 3). Die isolierte chromosomale DNA des Bakteriums wird mit dem entsprechenden Restriktionsenzym verdaut und die passende Nukleotidkassette wird an die einzelnen DNA-Fragmente des bakteriellen Genoms ligiert. Aus dem bekannten DNA-Bereich werden die Primer S1 und S2 für die PCRs ausgewählt. Nachdem das 5'-Ende der Nukleotidkassette nicht phosphoryliert ist (roter Kreis in Abbildung 3)

kommt es dort zu einer Unterbrechung des Einzelstranges bei der PCR. Erst nach der erfolgreichen Synthese des anderen DNA-Stranges, durch Primer aus der bekannten Region, kommt es zur Amplifikation des gewünschten Fragments.

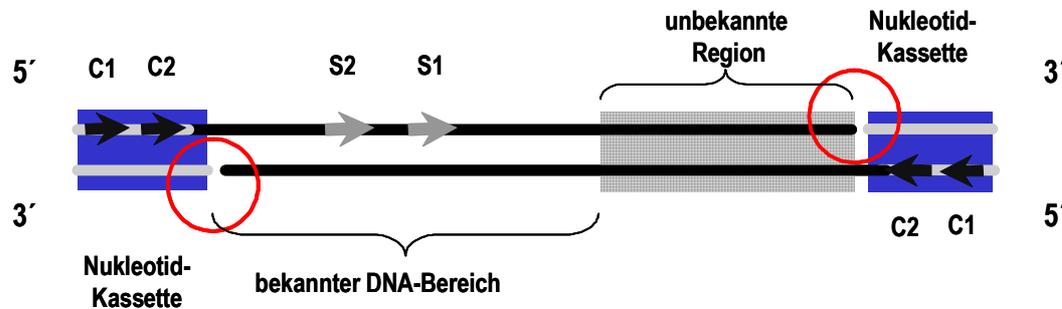


Abbildung 3 Darstellung des Prinzips des LA-PCR-Cloning

Die Spezifität und Ausbeute dieser Amplifikationen wird mit einer zweiten PCR (*nested PCR* = verschachtelte PCR) verbessert. Hierfür werden die Primer C2 und S2 verwendet. Die Größe des erhaltenen Fragments ist abhängig von der Entfernung der Erkennungs-DNA-Sequenz für das entsprechende Restriktionsenzym.

Nukleotidkassetten werden für folgende Restriktionsenzyme mitgeliefert: *Sau3AI*, *EcoRI*, *HindIII*, *PstI*, *Sall*, und *XbaI*.

Protokoll: *LA-PCR in vitro Cloning Kit Manual*, Takara Shuzo Co., Ltd., Kyoto, Japan

Durchführung: Von je 5 μg Gesamt-DNA aus einer Phenolextraktion des *E. coli* Stammes ECOR31 wurde ein 50 μl Restriktionsverdau mit ca. 50 U folgender Enzyme angesetzt: *Sau3A I*, *EcoRI*, *HindIII*, *PstI*, *Sall* und *XbaI*. Die geschnittene DNA wurde ethanolgefällt und in 10 μl sterilem H_2O aufgenommen. Zur Ligation an die entsprechenden Kassetten wurde 5 μl der vorbereiteten DNA mit 2,5 μl der jeweiligen, im Kit mitgelieferten Kassette, sowie 15 μl der *Ligationslösung I* und 7,5 μl der *Ligationslösung II* 30 min bei 16°C inkubiert. Nach erneuter Ethanolgefällung wird die DNA in 5 μl sterilem H_2O aufgenommen.

1 μl dieser DNA wurde mit 33,5 μl sterilem H_2O 10 min bei 94°C erhitzt und diente als Ausgangsmaterial für die 1. PCR mit den Primern S1 und C1:

- + 5 μl Mg^{2+} -versetzter 10x LA Puffer
- + 8 μl dNTP-Mix
- + 1 μl Primer C1 (10 pmol)
- + 1 μl Primer S1 (10 pmol)
- + 0,5 μl TaKaRa LA Taq (entspricht 2,5 U)

Programm:

30 Zyklen : $\left\| \begin{array}{l} 94^{\circ}\text{C}: 30 \text{ sec} \\ 55^{\circ}\text{C}: 2 \text{ min} \\ 72^{\circ}\text{C}: 8 \text{ min} + 5 \text{ sec pro Zyklus} \end{array} \right.$

72°C: 20 min

4°C: ∞

Mit 1 μl des 1:10 verdünnten Produkts dieser 1. PCR erfolgte, mit den Primern C2 und S2, die 2. PCR:

- + 5 μl Mg^{2+} -haltiger 10x LA Puffer
- + 8 μl dNTP-Mix
- + 1 μl Primer C2 (10 pmol)
- + 1 μl Primer S2 (10 pmol)
- + 33.5 μl steriles H_2O
- + 0,5 μl TaKaRa LA Taq (entspricht 2,5 U)

Programm:

94°C: 4 min

30 Zyklen : $\left\| \begin{array}{l} 94^{\circ}\text{C}: 30 \text{ sec} \\ 58^{\circ}\text{C}: 2 \text{ min} \\ 72^{\circ}\text{C}: 8 \text{ min} + 5 \text{ sec pro Zyklus} \end{array} \right.$

72°C: 20 min

4°C: ∞

Das Produkt der 2. PCR wurde auf ein Agarosegel aufgetragen. Bei Vorhandensein des gewünschten DNA-Fragments wurde die DNA-Bande ausgeschnitten, die DNA aufgereinigt und das Produkt in den Vektor pCR[®]-XL-TOPO[®] kloniert.

2.2.11. Klonierung von DNA-Fragmenten

Durch Restriktionsverdau entstandene DNA-Fragmente lassen sich in spezielle Plasmide, so genannte Vektoren, einbauen. Wird dieses zirkuläre DNA-Konstrukt in ein geeignetes Bakterium transferiert, kann es dort beliebig vermehrt werden.

Der Vektor wird mit dem Restriktionsenzym linearisiert, durch welches das einzufügende Fragment geschaffen wurde und mit alkalischer Phosphatase behandelt. Dies führt zur Abspaltung überstehender Phosphatresten an den freien Enden des Vektors und verhindert so eine Selbstligation. Der nächste Schritt ist die Ligation, bei der die beiden DNA-Fragmente, Vektor und Probe, miteinander verklebt werden. Rekombinante Plasmide können z.B. mittels Elektroporation in geeignete Bakterienstämme transformiert werden. Die Überprüfung einer Klonierung erfolgt mittels Antibiotika-Selektion und ggf. Sequenzierung der inserierten DNA-Region.

Protokoll: *DNA Ligation Kit* von Takara Shuzo Co., Ltd, Kyoto, Japan

Durchführung: Die Vektor-DNA wurde mit dem entsprechenden Enzym geschnitten und mit alkalischer Phosphatase behandelt. Die einzusetzende Menge in µg wurde nach Anleitung der Firma Promega nach folgender Formel berechnet, wobei das Verhältnis zwischen Vektor-DNA und Insert-DNA immer 3:1 betrug:

$$\text{Quotient} \frac{\text{Vektor}}{\text{Insert}} = \frac{[(\text{ng Vektor-DNA} \times \text{kb Größe des Inserts}) / \text{kb Größe des Vektors}] \times \text{molarer Quotient}}{\text{Insert}}$$

Das zu klonierende DNA Fragment wurde in 5-10 µl H₂O oder TE-Puffer aufgenommen. Nach Zugabe der gleichen Menge *Solution I* (enthält die DNA-Ligase) und vorsichtigem Mischen, wurden Vektor und potentiell Insert für 30 min bei 16°C

inkubiert. Die Ligations-Mischung kann danach sofort zur Transformation benutzt werden.

2.2.12. Klonierung von PCR-Fragmenten

Der TOPO-Vektor bietet als Besonderheit am 3'-Ende einen Thymidinüberhang und eine kovalent gebundene Topoisomerase, weshalb keine Enzymzugabe zur Ligation nötig ist. Nachdem die mit Taq-Polymerase amplifizierten PCR Produkte probenunabhängig am 3'-Ende einen Adenosinrest haben, können PCR-Produkte ohne weiteren Aufwand in den TOPO-Vektor ligiert werden.

Protokoll: TOPO[®] XL PCR Cloning Kit von Invitrogen[™], Carlsbad, USA

Durchführung: Das PCR-Produkt wurde auf einem 0,8%-igem Agarosegel elektrophoretisch aufgetrennt und anschließend nach Anleitung des *GeneClean II Kits* aufgereinigt. 4 µl des gereinigten PCR Produkts wurden mit 1 µl pCR[®]-XL-TOPO Vektor vorsichtig vermischt und 5 min bei Raumtemperatur inkubiert. Die Reaktion wurde mit 1 µl 6x TOPO[®] Cloning Stop Solution beendet und jeweils 2 µl der Suspension wurden pro Transformation in elektrokompente Zellen verwendet.

2.2.13. Transformation elektrokompenter *E. coli*

Zirkuläre Plasmide wurden mittels Elektroporation zur Vermehrung in *E. coli*-Stämme gebracht. Das Prinzip dieser Methode beruht darauf, dass die DNA in einer Bakteriensuspension hoher Zelldichte kurzzeitig einem starken elektrischen Feld ausgesetzt wird. Die Zellwand wird durchlässig und die Plasmid-DNA kann ins Zellinnere diffundieren. Der Empfängerstamm muss für diese Prozedur vorbereitet werden.

Herstellung der kompetenten Zellen: 1 l LB-Medium wurde mit dem entsprechenden Antibiotikum versetzt und im Volumenverhältnis 1:50 mit einer Bakterien-Übernacht-

Kultur beimpft und bei 37°C bis zu einer OD von 0.7 – 0.9 inkubiert. Nach 15 Minuten Kühlung auf Eis wurde die Kultur 30 min bei 4300 rpm in der vorgekühlten Zentrifuge bei 4°C abzentrifugiert. Der Überstand wurde verworfen und die Zellpellets in 300 ml 4° C-kaltem, sterilem Aqua dest. resuspendiert und erneut 15 min auf Eis gekühlt. Nach einer 20-minütigen Zentrifugation unter den oben beschriebenen Bedingungen wurden die Sedimente in 150 ml 4° C-kaltem, sterilem Aqua dest. aufgenommen, erneut 15 min auf Eis gekühlt und abzentrifugiert. Die Zellsedimente wurden anschließend in 50 ml gekühltem 10%-igem Glycerol resuspendiert. Wie oben beschrieben wurden die Zellen auf Eis gekühlt und anschließend bei 4°C abzentrifugiert. Zur Lagerung bei – 80°C wurden die Pellets in 1.5-3 ml kaltem 10%-igem Glycerol resuspendiert und in Eppendorfcups aliquotiert.

Durchführung der Elektroporation: 2-3 µl des Ligationsansatzes wurden mit 50 µl der elektrokompetenten Bakterien gemischt und in einer auf Eis gekühlten 1 mm-Küvette entsprechend der Anleitung der Firma BIO RAD® elektroporiert (U=1,8 kV, C=25 µF, R=200 Ω). Die Zellen wurden anschließend in 1 ml, auf 37°C vorgewärmtes, SOC Medium aufgenommen und nach Inkubation von einer Stunde bei 37°C auf antibiotikahaltigen LB-Agar-Platten ausgestrichen.

2.2.14. In vitro Transposon-Mutagenese

Mit dem *Primer Island Transposition Kit* lassen sich in vitro Transposon-Insertionen durchführen. Die Methode kann verwendet werden um Gene auszuschalten, oder auch um größere Inserts in Plasmiden zu sequenzieren.

Das hier verwendete Transposon AT-2 enthält ein Gen, das für die Dihydrofolat-Reduktase kodiert und somit eine Resistenz gegen das Antibiotikum Trimethoprim vermittelt. Die in vitro Insertion in das Plasmid erfolgt nach dem Zufallsprinzip durch eine Exzisions-Ligations-Reaktion.

Die Plasmide wurden durch Elektroporation in einen *E. coli* Stamm (DH10B) transformiert und die Insertion des Transposons durch Antibiotikaselektion (Tp^r) überprüft. Nach Isolierung der Plasmid-DNA kann eine Sequenzierung bidirektional erfolgen. Für die Sequenzierung verwendet man Primer, die beiden Enden des

Transposons entstammen und Bestandteile des Kits sind (primer island-PI(+).for, primer island-PI(-).rev).

Protokoll: *Primer Island Transposition Kit*, PE Applied Biosystems, Darmstadt, Deutschland

Durchführung: Die Reagenzien des Kits wurden langsam aufgetaut und es wurde ein 20 µl Ansatz vorbereitet aus:

- 12 µl deionisiertes H₂O
- 4 µl 5x Transposase Puffer
- 1 µl Plasmid DNA (1 µg)
- 1 µl AT-2 Transposon (0,2 µg)
- 2 µl Transposase

Der Ansatz wurde nach vorsichtigem Mischen für eine Stunde bei 30°C inkubiert und anschließend mit 5 µl *Stop-Puffer* 30 min bei 65°C inaktiviert. Die DNA wurde im Anschluss daran durch eine Ethanol-fällung gereinigt und in 20 µl sterilem, deionisiertem H₂O aufgenommen. 1 µl der DNA wurde in 50 µl des elektrokompetenten *E. coli* Stammes DH10B transformiert.

Die DH10B Zellen wurden in verschiedenen Verdünnungsstufen (10^{-1} , 10^{-2} , 10^{-3}) auf vorgewärmten Trimethoprim- und Kanamycin-haltigen Müller Hinton Agarplatten ausgestrichen und über Nacht bei 37°C bebrütet.

Reagenzien:

- | | |
|-----------------------|--|
| Stop Puffer (10 µl): | 2,75 µl deionisiertes H ₂ O |
| | 5,0 µl 0,5 M EDTA |
| | 1,0 µl 10% SDS |
| | 1,25 µl Proteinase K (40 mg/ml) |

2.2.15. Herstellung einer Cosmid-Bank

Mit einer Cosmid Bank gelingt es das gesamte Genom eines Bakterienstammes in Fragmente überschaubarer Größe repräsentativ zu klonieren. Hierzu wird die chromosomale DNA des zu untersuchenden Stammes mit dem Enzym *Sau3A I* partial verdaut und in DNA-Stücke zwischen 30 und 42 kb fragmentiert. Diese werden über eine *Bam*HI-Schnittstelle in den Cosmid-Vektor SuperCos1 ligiert. Der Vektor kann auf Grund seiner speziellen *cos*-sites in einen Phagen verpackt werden. Ein *E. coli*-Stamm, z.B. XL1-Blue MR, wird von den cosmidtragenden Phagenpartikeln transfiziert. In der Bakterienzelle wird das Cosmid als rezirkuläres DNA-Molekül extrachromosomal repliziert und in dieser Plasmidform auf die Tochterzellen weitergegeben.

Die rekombinanten Cosmid-tragenden *E. coli* konnten auf Antibiotika-haltigen Agarplatten selektioniert werden. Die so erhaltene Cosmid-Bibliothek kann durch Colony-Hybridisierung mit spezifischen Sonden durchsucht werden, um große DNA-Regionen zu isolieren, die den betreffenden Bereich (Sonde) umschließen.

Protokoll: *SuperCos1 Cosmid Vector Kit* von Stratagene[®], La Jolla, USA

Durchführung:

a) Vorbereitung der Vektor DNA: 25 µg SuperCos1 DNA wurde nach Protokoll mit *Xba*I geschnitten und somit linearisiert, dephosphoryliert und mit *Bam*HI in 2 Fragmente von 1,1 kb und 6,5 kb Größe gespalten. Nach jedem Verdau, sowie nach der Dephosphorylierung mit alkalischer Phosphatase, wurde eine Phenol-Aufreinigung und Ethanolfällung durchgeführt und die DNA in sterilem H₂O aufgenommen.

b) Vorbereitung der chromosomalen DNA: Die mittels Phenol-Extraktion isolierte chromosomale DNA des *E. coli* Stammes ECOR31 wurde mit *Sau3A I* partial verdaut. Um Fragmente von 30-42 kb zu erhalten wurde die genaue Restriktions-Zeit und Enzym-Menge in einer Testreihe ermittelt, (siehe Unterpunkt 2.2.5.).

100 µg DNA wurden in einem Endvolumen von 1 ml mit 1 U *Sau3A I* 12 min bei 37°C angedaut und anschließend mit 15 µl 0,5 M EDTA 10 min bei 70°C inaktiviert. Vor

und nach Dephosphorylierung mit 45 U alkalischer Phosphatase wurde die DNA Phenol-gereinigt und Ethanol-gefällt.

Ligation: 2,5 µg der partial verdauten und dephosphorylierten chromosomalen DNA wurden in 1 µg vorbereiteten Vektor ligiert, unter Zugabe von 1-2 Weiss U der T4 DNA-Ligase und dem entsprechenden Puffer.

c) Verpackung in den Phagen

Protokoll: *Gigapack[®] III Gold Packaging Extract* von Stratagene[®]

Durchführung: 4 µl der Ligation wurden mit dem *Packaging Extract* aus dem *Packaging Kit* vorsichtig vermischt und 2 h bei Raumtemperatur gelagert. Anschließend wurden 500 µl *SM-Puffer* und 20 µl Chloroform zugegeben. Der Überstand mit dem Phagen wurde dann titriert.

d) Infektion mit dem Phagen: Der *E. coli* Stamm XL1-Blue MR wurde gemäß dem Protokoll für die Infektion mit dem Phagen vorbereitet. Angezchtet in 50 ml LB-Medium mit 10 mM MgSO₄ und 0,2% Maltose (OD₆₀₀ < 1,0), wurde das Sediment in 10 mM MgSO₄ auf eine OD₆₀₀ von 0,5 verdünnt und mit Verdünnungsreihen von 1:10 und 1:50 mit der Mischung aus dem verpackten Phagen (*packaging reaction*) eine Stunde bei 37°C inkubiert. Die danach nochmals in frischem LB-Medium aufgenommenen Bakterien wurden auf Ampicillin-LB-Agarplatten ausplattiert und über Nacht bei 37°C inkubiert.

e) Amplifizierung der Cosmid-Bank: Die verpackte Cosmid-DNA wurde in einer Konzentration von ca. 50 000 KBE/Röhrchen mit der gleichen Menge des *E. coli* Stammes XL1-Blue MR (wie bei der Titration auf eine OD₆₀₀ von 0,5 vorbereitet) mit 4 Volumen LB-Medium eine Stunde bei 37°C inkubiert. Die Sedimente wurden in 500 µl LB-Medium aufgenommen und auf 150-ml Ampicillin- und Kanamycin-haltigen LB-Agarplatten bei 37°C über Nacht inkubiert. Auf die Platten wurden nacheinander 2 x 3 ml LB-Medium pipettiert und die Kolonien mit einem sterilen Drigalski-Spatel abgelöst. Die Abschwemmung aller Platten wurde in einem 50-ml Falcontube gesammelt und mit sterilem Glycerol auf eine Endkonzentration von 20% gebracht, und nachfolgend bei -80°C aufbewahrt.

2.2.16. Screening der Cosmid Bank, Herstellung der Colony-Blot-Filter

Es empfiehlt sich ca. 20 Platten mit insgesamt 50 000 KBE zu screenen.

150-ml-LB-Agar-Platten mit Ampicillin und Kanamycin wurden mit Nitrocellulose-Transfer-Membranen ausgelegt und 500 μ l der $1:10^{-5}$ verdünnten Cosmid-Bank wurden auf den Filtern ausplattiert und bei 37°C über Nacht angezchtet. Auf die mit Kolonien bewachsenen Filtern wurde durch Auflage neuer Filter ein spiegelverkehrter Abdruck gemacht. Beide Filter wurden auf LB-Agarplatten 6 h bei 37°C inkubiert und der erste Filter wurde als Originalplatte (*Masterplate*) aufbewahrt. Der 2. Filter, *Replica-Filter*, wurde zunächst auf 3 MM Whatman Papier 30 sec mit 0,5 M NaOH denaturiert, anschließend für weitere 30 sec auf in 1 M Tris-HCl-Lösung (pH 7,6) getränktes Whatman Papier gelegt. Nach weiteren 30 sec mit einer Mischung aus 1 M Tris-HCl-Lösung (pH 7,6) und 1,5 M NaCl wurden die Kolonien von dem Filter entfernt und die Membran mit der gleichen Lösung noch mal gereinigt. Der Blot wurde getrocknet und zur Fixierung der DNA mit UV-Licht einer Energie von 120 mJ bestrahlt.

Reagenzien:

SM-Puffer (1 l):

- 5,8 g NaCl
- 2,0 g $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$
- 50 ml 1 M Tris-HCl (pH 7,5)
- 5,0 ml 2% Gelatine
- deionisiertes H_2O ad 1 l, autoklaviert

2.2.17. Markierung von Gensonden

Einzelsträngige DNA-Fragmente werden mit einem Nonanukleotidgemisch einer vielfältigen Anzahl an Basensequenzkombinationen (*random primer*) versetzt. Diese Primer binden an die DNA-Einzelstränge und wirken als Startmoleküle für die Neusynthese des Stranges durch das Klenow-Enzym. Dieses Enzym baut neben den hinzugegebenen vier unmarkierten Nukleotiden dATP, dCTP, dGTP, dTTP auch das Fluorescein-markierte dUTP an Stelle von dTTP ein. Die Fluorescein-markierte

DNA kann durch ein Antikörperkonjugat und nachfolgender Chemilumineszenz-Reaktion nachgewiesen werden.

Protokoll: *Gene Images Random Prime Labelling Module-KitTM*, Amersham, Piscataway, USA

Durchführung: 1 µg des aufgereinigten PCR-Produkts wurde in 34 µl H₂O aufgenommen und durch Erhitzen im Wasserbad denaturiert. Nach Abkühlen auf Eis wurden 10 µl Nukleotid-Mix, 5 µl Primer und 1 µl Klenow-Enzym (5 U/µl) dazu pipettiert und eine Stunde bei 37°C inkubiert. Die Sonde wurde entweder gleich verwendet oder bei –20°C gelagert.

2.2.18. Colony Blot-Hybridisierung

Der Nachweis bestimmter DNA-Sequenzen beruht auf der Anlagerung einzelsträngiger, Fluoreszein-markierter DNA (Gensonde) an die komplementäre Sequenz der zu untersuchenden, an eine Trägermembran fixierten einzelsträngigen DNA. Unspezifische Bindungen werden durch eine Blockierlösung minimiert. Ein Fluoreszein-Antikörper (Peroxidase-Konjugat) bindet an die markierte DNA und durch Zusatz von Detektionslösung wird die enzymatische Reduktion von Peroxid mit der Oxidation von Luminol gekoppelt. Das Luminol gelangt vorübergehend in einen angeregten Zwischenzustand, wobei es im Übergang zum Grundzustand Licht emittiert. Die Emission kann autoradiographisch durch Schwärzung eines Röntgenfilms festgehalten werden.

Durchführung: Die Membran wurde nach Fixieren der DNA durch UV-Belichtung, zunächst mit 0,125 ml/cm² vorgewärmter *Hybridisierungslösung* bei 60°C unter konstantem Drehen prähybridisiert. Die markierte Gensonde wurde inzwischen durch 5-minütiges Erhitzen in einem Wasserbad denaturiert, auf Eis gestellt und nach 30 bis 60 min Prähybridisierung in eine Glasröhre pipettiert. Im Hybridisierungssofen erfolgte so bei ständigem Drehen und 60°C die Hybridisierung über Nacht. Nach zweimaligem Waschen mit 2-5 ml/cm² der *Waschlösungen 1* und *2* bei 60°C,

erfolgten alle weiteren Schritte bei Raumtemperatur. Die Membran wurde kurz mit *Puffer A* gespült und danach, für mindestens 30 min, mit einer 20-fachen Verdünnung von *Blockierungsreagenz* (Kit-Bestandteil) in *Puffer A*, inkubiert (0,5 ml/cm² *PufferA*). Nach der Blockierung der Membran erfolgte die zweistündige Inkubation mit dem *Antikörper-Konjugat* (*Anti-Fluorescein-HRP-Konjugat* 1000-fach verdünnt in *Puffer A* (0,25 ml/cm²) mit 0,5% BSA). Das überschüssige Konjugat wurde in 3 Waschschritten à 10 min mit einer 0,3%-igen *Tween-20-Puffer A-Lösung* entfernt. Die Membran wurde daraufhin mit einem Gemisch, aus den im Kit enthaltenen *detection solution 1* und *2*, bedeckt (0,125 ml/cm²) und von beiden Seiten in eine Folie geschlossen. Die Lichtemission der in der Röntgenkassette befestigten Membran wird auf dem aufgelegten Röntgenfilm dokumentiert.

Reagenzien:

Hybridisierungslösung:	5x SSC (' <u>s</u> tandard <u>s</u> aline <u>c</u> itrate')
	0,1% SDS (' <u>s</u> odium <u>d</u> odecyl <u>s</u> ulfate')
	5% Dextransulfat
	1/20 Vol. Blockierungsreagenz (mitgeliefert)
Waschlösung 1:	1x SSC
	0,1% SDS
Waschlösung 2:	0,5x SSC
	0,1% SDS
Puffer A, pH 7,5:	100 mM Tris-HCl
	600 mM NaCl

2.2.19. RNA-Isolation

Um die Gentranskription zu analysieren benötigt man als Ausgangsmaterial RNA. Bakterienzellen des *E. coli* Stammes ECOR31 wurden zu diesem Zwecke mit Lysozym und einem mitgelieferten *Lysepuffer* behandelt. Die freiwerdenden Nukleinsäuren werden an die Silica-Matrix der Filter gebunden, woraufhin die DNA

durch Inkubation mit einem *DNase-Puffer-Gemisch* eliminiert wird. Die RNA wird in mehreren Waschschrinen gesäubert und kann schließlich durch Zugabe von *Elutionspuffer* aus dem Filter gelöst werden.

Protokoll: *High Pure RNA Isolation Kit* der Firma Boehringer Mannheim, Deutschland

Durchführung: 1 ml Bakterienkultur mit einer $OD_{600} = 0,8 - 0,9$ wird 5 min bei 5000 rpm abzentrifugiert und das Sediment in 200 μ l 10 mM Tris pH 8,0 resuspendiert. Nach Zugabe von 4 μ l Lysozym (Stammlösung 50 mg/ml) wurde die Lösung 10 min bei 37°C inkubiert und danach mit 400 μ l *Lysepuffer* versetzt. Die Nukleinsäuren binden an die Silica-Matrix der Filter (15 sec bei 10000 rpm zentrifugieren) und wurden mit einem Gemisch aus 10 μ l DNase (1 U/ μ l) und 90 μ l des dazugehörigen Puffers eine Stunde bei Raumtemperatur inkubiert. Danach wurde nacheinander eine Waschung mit je 500 μ l der beiden *Waschpuffer I* und *II* aus dem Kit durchgeführt. Nach nochmaliger Säuberung mit 200 μ l *Waschpuffer II* (10 min abzentrifugieren bei 14000 rpm) erfolgte die Elution in 75 μ l *Elutionspuffer* (zentrifugieren: 1 min bei 10000 rpm).

2.2.20. Reverse Transkriptase Polymerasekettenreaktion (RT-PCR)

Die RT-PCR ist eine sensitive Methode, um die Transkription eines Gens (messenger RNA) zu untersuchen. Die zuvor isolierte, DNA-freie RNA wird mit einer reversen Transkriptase in cDNA umgeschrieben und das Produkt mit Hilfe einer Taq-Polymerase in einer ‚normalen‘ PCR amplifiziert.

Protokoll: *Access RT-PCR Introductory System*, Promega, Heidelberg, Deutschland

Durchführung: Vor der PCR wird eine zusätzliche DNase Behandlung durchgeführt, um eventuell noch vorhandene DNA abzubauen. Die RNA, zwischen 1 pg und 1 μ g, wurde dazu in einem Volumen von 40 μ l mit 10 μ l *AMV/Tfl 5x-Puffer* aus dem Kit, 2 μ l DNase (1 U/ μ l) und sterilem Aqua dest. eine Stunde bei 37°C inkubiert und danach 7 min bei 75°C inaktiviert. Anschließend wurde der RT-PCR Ansatz wie folgt pipettiert:

1 μ l dNTPs (je 10 mM dATP, dCTP, dGTP, dTTP), Endkonz. 0,2 mM
 1 μ l forward Primer (20 pmol)
 1 μ l reverse Primer (20 pmol)
 2 μ l 25 mM MgSO₄
 1 μ l AMV Reverse Transkriptase (5 U/ μ l)
 1 μ l Panscript Taq-Polymerase (5 U/ μ l)
 Aqua bidest. ad 50 μ l

Um zu kontrollieren, dass die RNA-Proben DNA-frei sind und positive PCR-Ergebnisse ausschließlich auf spezifische RNA zurückzuführen sind, wurden analoge RT-PCR-Reaktionen ohne initialen reversen Transkriptase-Zusatz durchgeführt (Negativkontrolle).

Die Proben wurden zunächst 45 min bei 55°C inkubiert und anschließend wurde folgende PCR durchgeführt:

94°C	2 min		
94°C	30 sec		30 Zyklen
57°C bzw. 56°C	1 min		
68°C	2 min		
68°C	7 min		
4°C	∞		

2.2.21. Sequenzierung

Direkte Sequenzierungen wurden mit der Methode nach Sanger ¹⁴⁵ durchgeführt. Dabei wird in einer Sequenzier-PCR mittels eines Primers aus dem bekannten DNA-Bereich einer der beiden komplementären DNA-Stränge verlängert. Zusätzlich zu den gewöhnlichen Desoxyribonukleosidtriphosphaten (dNTPs) werden in einem bestimmten Verhältnis auch Didesoxyribonukleosidtriphosphate (ddNTP) beigemischt (ddATP, ddCTP, ddGTP, ddTTP). Diese sind mit unterschiedlichen

Farbstoffen fluoreszinmarkiert und besitzen keine 3'-Hydroxygruppe. Daher kommt es nach deren Einbau in den neu synthetisierten DNA-Strang zu einem Abbruch, da eine Verknüpfung mit der Phosphatgruppe des nächsten Nukleotids nicht mehr möglich ist. Es entstehen damit eine Reihe unterschiedlich langer DNA-Fragmente. In einem automatischen Sequenziergerät (ABI Prism; Model 373A und 377 von Perkin Elmer) werden die erhaltenen DNA-Fragmente mittels Kapillarelektrophorese aufgetrennt und die ddNTPs am Ende jedes DNA-Fragmentes werden mit Hilfe eines Lasers zur Fluoreszenz angeregt. Das Fluoreszenzlicht wird von einer Kamera registriert und über eine spezielle Software in vierfarbige Chromatogramme umgewandelt. Jedem Didesoxynukleotid ist hierbei eine andere Farbe zugeordnet (A = grün, T = rot, C = blau, G = schwarz).

Protokoll für die Sequenzierungs-PCR:

4 µl BD-Mix
 + 1 µl Primer (5 pmol/µl)
 + DNA (1,5 µg Plasmid DNA oder 0,3 µg aufgereinigtes PCR-Produkt)
 ad 20 µl Aqua dest.

PCR-Programm:

96°C	10 sec		25 Zyklen
x°C	5 sec		
60°C	4 min		
4°C	∞		

wobei x = Annealing-Temperatur, Berechnung: $x = (G + C) \times 4 + (A + T) \times 2 - 2$

Anschließend wird das PCR-Produkt mittels Ethanol-Fällung aufgereinigt, getrocknet und nach Anleitung des Sequenziergerätes für die Detektion aufbereitet.

2.3. Software

Die Analyse der Nukleotidsequenzen wurde mit dem Programmpaket DNASTAR (Firma Lasergene, Version 5.02) mit den Programmen SeqMan, GeneQuest, MapDraw, MegAlign und EditSeq durchgeführt.

Der G+C-Gehalt der Sequenzen wurde mit dem Programm EditSeq bestimmt. Zur Ermittlung offener Leseraster (ORF = Open-Reading-Frame) wurde ein Start- und Stopp-Codon, sowie eine Mindestgröße von 50 Aminosäuren (150 Basenpaare) vorausgesetzt.

Homologienvergleiche und die Bestimmung konservierter Motive erfolgten über die veröffentlichten Sequenzen der Sequenzdatenbank EMBL/GenBank mit der BLAST-Software (<http://www.ncbi.nlm.nih.gov/BLAST>). Eine ausführliche Tabelle, inklusive der Beschreibung der Zugangsdaten (Accession-Nr), befindet sich im Anhang (Tabelle 1).

3. Ergebnisse

3.1. Untersuchung des Insertionsortes der *Yersinia*-HPI im *E. coli* Stamm ECOR31

Die Pathogenitätsinsel *Yersinia*-HPI ist bei *Escherichia coli* wie bei Yersinien in der Genregion der Asparaginsäure-tRNA (*asn*-tRNA) integriert^{21, 149}. *E. coli* beherbergt in seinem Genom vier *asn*-tRNA-Gen-Kopien, *asnV*, *asnU*, *asnW* und *asnT* (Abbildung 4). In Vorarbeiten wurde gezeigt, dass die HPI ausschließlich in die *asnT*-tRNA-Kopie integriert ist¹⁴⁹. Die 17 Basenpaare (bp) am 3'-Ende des *asn*-tRNA-Gen sind die Erkennungssequenz einer *site*-spezifischen Integration der HPI in das bakterielle Genom (in Abbildung 4 schwarz dargestellt). Diese Sequenz wird bei der Eingliederung der Insel dupliziert und begrenzt als 'direct repeat' (DR) ein Ende der HPI, siehe auch Unterpunkt 1.4. Das *asn*-tRNA-Gen wird bei dem Prozess nicht zerstört und flankiert die HPI als intaktes Gen am Integrase-Ende (*int*) der HPI.

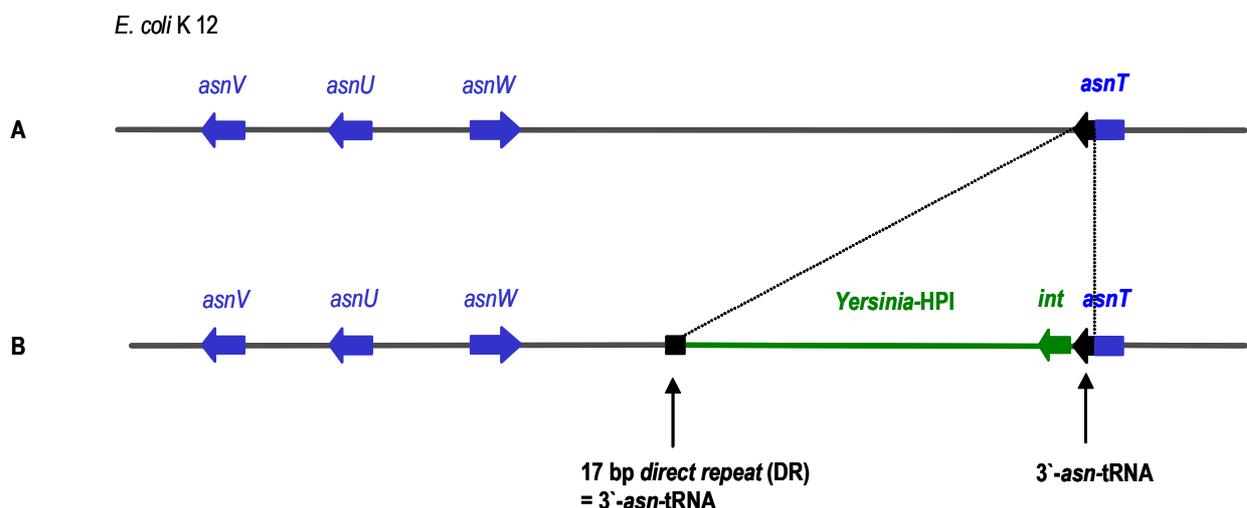


Abbildung 4: *E. coli* K12 mit 4 Kopien des *asn*-tRNA-Gens (A). Die *Yersinia*-HPI ist üblicherweise in die *asnT*-tRNA-Kopie integriert (B). Das intakte *asnT*-tRNA-Gen flankiert die HPI am Integrase-Ende (*int*).

Bei Untersuchungen des Integrationsortes der *Yersinia*-HPI bei *E. coli* wurde nur ein einziger Stamm isoliert, bei dem die HPI nicht in das *asnT*-tRNA-Gen integriert ist: ECOR31 aus der *E. coli* Referenzsammlung¹¹⁷. Im Genom von ECOR31 finden sich zwei HPI-Kopien, wobei eine in das *asnV*-tRNA-Gen inseriert hat¹⁴⁷. Eine zweite Kopie konnte in einem unbekanntem DNA-Bereich lokalisiert werden.

Zur DNA-Charakterisierung des unbekanntem Insertionsortes der *Yersinia*-HPI in ECOR31 wurde der *LA PCR in vitro cloning Kit* benutzt. Mit dieser Technik gelangen, ausgehend von einer bekannten Region, PCR-Amplifikationsprodukte die in einen komplett unbekanntem DNA-Bereich hineinreichen. Nach Aufreinigung des PCR-Produktes kann dieses in einen Vektor kloniert werden und steht für Sequenzierungen und für weitere Versuche zu Verfügung.

Wie im Methoden-Abschnitt beschrieben, wurde dabei die jeweilige enzyspezifische Nukleotidkassette an die mit dem Restriktionsenzym *EcoRI* geschnittene chromosomale DNA des *E. coli* Stammes ECOR31 ligiert. Nach der Ligation folgte eine verschachtelte PCR ('*nested PCR*') mit jeweils zwei Primern von der Nukleotidkassette (C1 und C2) und dem *intB*-Bereich der *Yersinia*-HPI (Primer: UP HPIcomp1 und UP HPIcomp2) (Abbildung 3 und Abbildung 5).

Das Ergebnis ist ein ca. 4.1 kb langes PCR Produkt (Abbildung 5), das nach Aufreinigung aus dem Agarosegel in den *TOPO[®] XL PCR*-Vektor gebracht wurde. Das so entstandene Plasmid pBP-LA wurde zur Kontrolle zunächst vom flankierenden Vektor aus ansequenziert (Primer: M13 und T7) und anschließend nach Transposon-Mutagenese mit dem *Primer Island Transposition Kit* komplett sequenziert.

E. coli ECOR 31

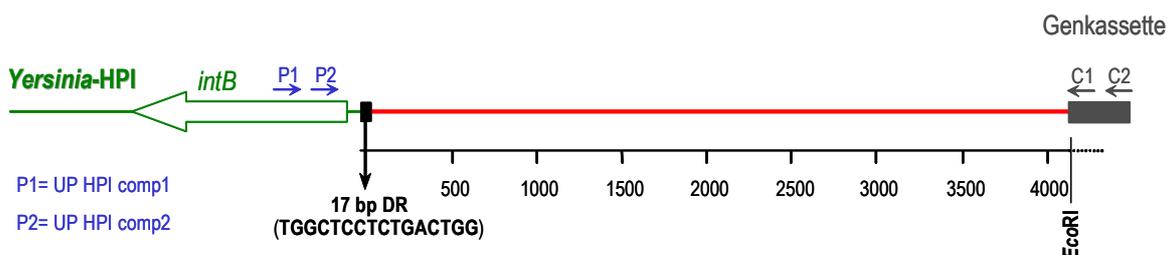


Abbildung 5: Integrationsort der HPI in *E. coli* ECOR31: 4,1 kb LA-PCR-Produkt (rot dargestellt) mit *EcoRI* Schnittstelle, Genkassette, PCR-Primern (P1, P2, C1, C2) und angrenzender HPI.

Die so erhaltene 4162 bp lange DNA-Sequenz wurde mit Nukleotid-Datenbankeinträgen verglichen. Die höchsten Homologien ergaben sich zur DNA von *Mesorhizobium loti* (55% identische Nukleotide) (Stand 2002). Der G+C-Gehalt der Region weicht mit 47,07% von dem G+C-Gehalt von *E. coli* mit 50,8% ab⁴⁶. Die in diesem Bereich enthaltenen offenen Leseraster wiesen konservierte Motive von ABC-Transportern für divalente Metallionen auf. Die genauere Beschreibung der einzelnen ORFs folgt im Abschnitt 3.2.1. Auf Grund der fehlenden Homologien zu *E. coli*-DNA wird der gesamte unbekannte DNA-Bereich ab dem 17 bp *direct repeat* im Folgenden **Region X (RegX)** genannt.

Der Insertionsort der HPI im *E. coli* Stamm ECOR31 ist ungewöhnlich, da es sich nicht um ein tRNA-Gen handelt. Um der Frage nachzugehen, ob die *Yersinia*-HPI in eine Sequenz mit hoher Homologie zu einem tRNA-Gen 'fehlinsertiert' ist, wurde die angrenzende DNA auf hohe Homologien untersucht. Bis auf die 17 Basenpaare '*direct repeat*' ergaben sich keine höheren Homologien zum *asn*-tRNA-Gen, dem gewöhnlichen Insertionsbereich der *Yersinia*-HPI (Abbildung 6).

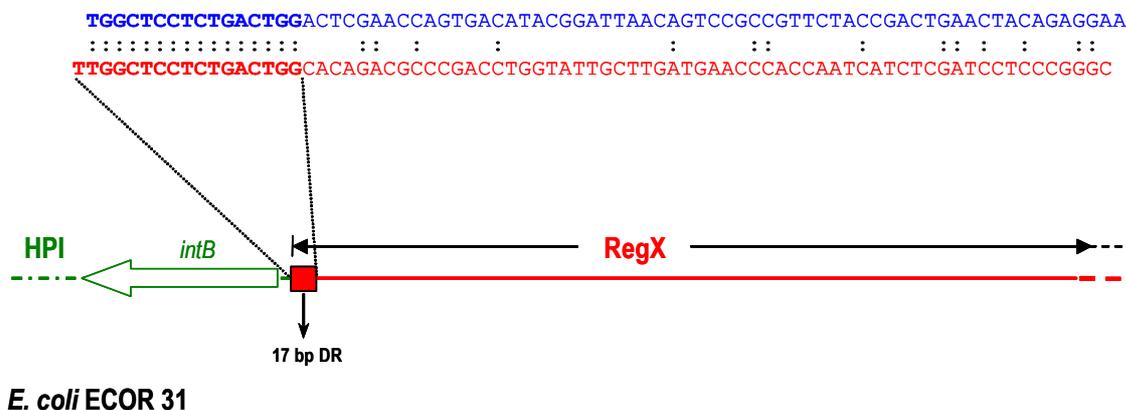


Abbildung 6: DNA-Sequenz-Vergleich der Insertionsstelle der HPI in *E. coli* ECOR31 = RegX, inklusive 17 bp DR (rot) mit dem sonst üblichen Insertionsort der HPI, dem Gen für *asn*-tRNA aus *E. coli* (blau).

Zusammenfassend betrachtet weist der *E. coli*-Stamm ECOR31 einen bislang unbekanntem, atypischen Insertionsort der *Yersinia*-HPI auf. Die Sequenz der Region im Insertionsbereich (RegX) hat keine Homologien zur bislang veröffentlichten *E. coli*-DNA. Vielmehr weist der G+C-Gehalt darauf hin, dass die Region von einem anderen Bakterium stammt und mittels horizontalen Gentransfers übertragen wurde. Im Rahmen der vorliegenden Dissertationsarbeit soll daher der gesamte Insertionsbereich der HPI in *E. coli* ECOR31 sequenziert und genauer charakterisiert werden.

3.2. Untersuchung der unbekanntenen Region X mit einer Cosmid-Bank von ECOR31

Der ungewöhnliche Insertionsort RegX der *Yersinia*-HPI in ECOR31 ist eine bislang unbekanntene DNA-Region. Die ersten Sequenzierungen ergaben Hinweise auf ABC-Transporter-Proteine von nicht verwandten Mikroorganismen. Dies deutet darauf hin, dass die Region horizontal transferiert ist und Gen-Cluster kodiert, die den Stamm möglicherweise mit Virulenz- oder Stoffwechsel-Eigenschaften ausstatten.

Um einen größeren Bereich der Region charakterisieren zu können und um mit Original-DNA statt PCR-generierter DNA zu arbeiten, wurde eine Genom- bzw. Cosmid-Bank des *E. coli* Stammes ECOR31 mit Hilfe des *SuperCos1 Cosmid Vector Kit* und des *Gigapack[®] III Gold Packaging Extract* angelegt (siehe Methoden, Abschnitt 2.2.15 f). Hierzu wurde die mit dem Restriktionsenzym *Sau3A I* auf DNA-Fragmente um 40 kb partial verdaute ECOR31-DNA in das linearisierte und mit *BamHI* geschnittene SuperCos1-Cosmid ligiert. Das Ligationsgemisch wurde dann mit Phagenpartikeln aus dem *Gigapack[®] III Gold Packaging Extract* inkubiert und in Phagenköpfe verpackt. So konnte anschließend der *E. coli*-Stamm XL1-Blue MR durch die Cosmid-tragenden Phagen transfiziert werden (Abbildung 7).

Es wurde eine Cosmid-Bibliothek von insgesamt 50 000 Cosmid-tragenden *E. coli*-Klonen über Colony-Blot-Filter mit RegX-spezifischen Gensonden gescreent. Dazu wurde die Cosmid-Bank auf Nitrocellulose Filtern ausgestrichen, die auf Ampicillin-Kanamycin-LB-Agar-Platten aufgebracht waren. Von dem Original-Filter mit den angewachsenen Kolonien konnte ein Replica-Filter angefertigt werden, der nach der Denaturierung der DNA mittels Hybridisierung, von einer markierten Gensonde ausgewertet wurde. Diese Gensonde wurde aus dem bisher bekannten RegX-Abschnitt hergestellt. Hierzu wurde das 4.1 kb große RegX-Fragment aus dem Plasmid pBP-LA mit *EcoRI* herausgeschnitten und die Gensonde nach Aufreinigung markiert und fertig gestellt (*Gene Images Random Prime Labeling Module-KitTM*). Die den schwarzen Signalen auf dem Röntgenfilm entsprechenden Kolonien konnten von der Originalplatte entnommen und angezüchtet werden. Nach Isolierung der Cosmid-DNA wurden die positiven Klone nochmals durch eine PCR bestätigt.

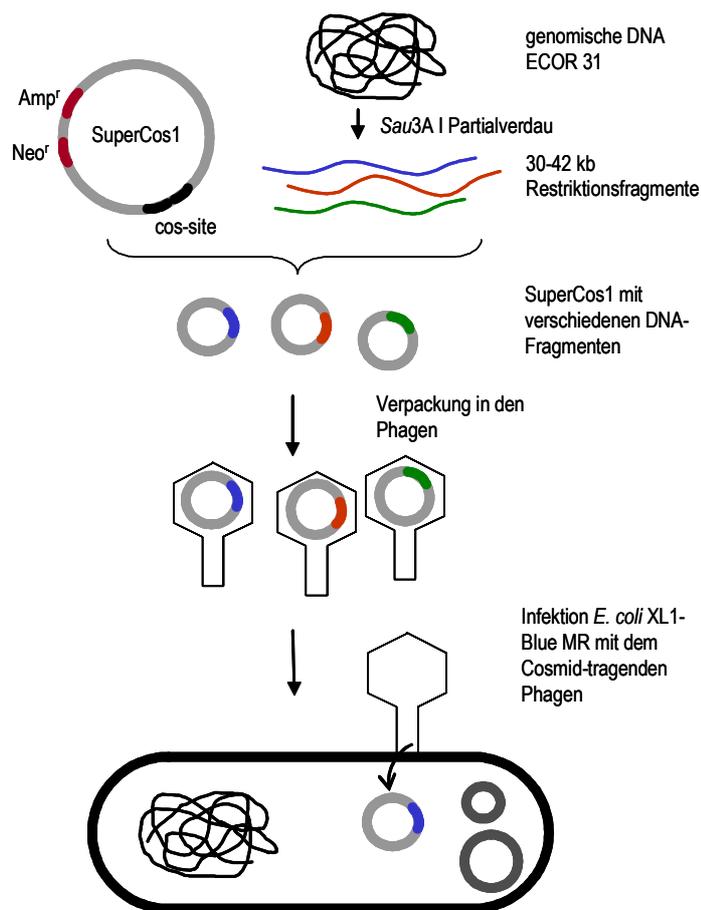


Abbildung 7: Herstellung einer Cosmid-Bank, schematische Darstellung.

Das Screening brachte insgesamt drei verschiedene RegX-positive Klone (p17-A7, p17-D6, p14-F1) hervor, die von den Originalplatten isoliert werden konnten. Mit den Primern T3 und T7 wurde der Randbereich der in die Cosmide eingeschlossenen DNA von ECOR31 bestimmt. Alle Cosmide schließen neben einem unbekanntem Bereich auch einen Teil der HPI mit ein.

Das HPI-Ende des Cosmids p17-A7 liegt im Bereich des *irp2*-Genes (Abbildung 8). Die beiden anderen Klone enthalten auch Teile des *irp1*-Genes. Erste Untersuchungen ließen vermuten, dass p17-A7 den größten Anteil unbekannter DNA trägt. Die Cosmid-DNA des Klons p17-A7 wurde daher in großer Menge isoliert, aufgereinigt und für die Primer-Walking-Sequenzierung in Zusammenarbeit mit der Firma Agowa GmbH, Berlin, vorbereitet. Nach Abschluss externer Sequenzierarbeiten wurden die Lücken in der DNA-Sequenz über PCR-Amplifikationen und eigene Sequenzierungen geschlossen.

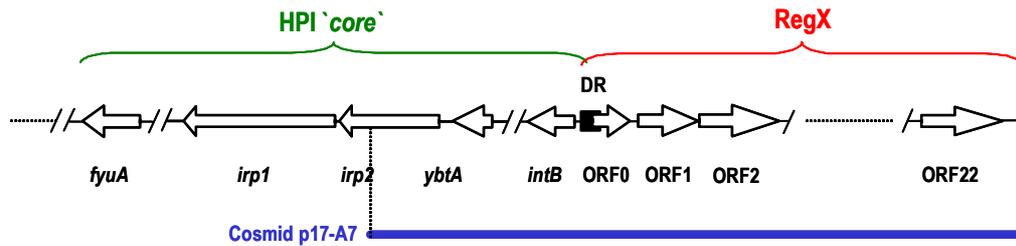


Abbildung 8: Übersicht DNA-Inhalt des RegX-positiven Cosmids p17-A7.

Die Sequenzdaten ergaben eine DNA-Region von 24196 Basenpaaren vom Beginn des 'direct repeat' (DR) bis zum Ende des ECOR31-DNA-Fragments im Cosmid p17-A7. Das erhaltene Fragment, RegX, umfasst 22 offene Leseraster (ORF) (Abbildung 9) und wird bei der genaueren Charakterisierung aus Gründen der Übersichtlichkeit in mehrere DNA-Abschnitte unterteilt: DNA-Bereich A bis E. Die Festlegung der einzelnen Bereiche erfolgte in Orientierung an potentiellen Gen-Clustern (Operons), die sich durch Homologienvergleiche mit den Datenbanken ergaben.

Die Auswertung der DNA-Sequenz erfolgte mit dem Programm-Paket Lasergene (DNASTAR) und mittels Homologienvergleichen zu den bisher veröffentlichten Sequenzen mit dem online verfügbaren NCBI Blast search Algorithmus (<http://www.ncbi.nlm.nih.gov/BLAST>), siehe auch Material und Methoden.

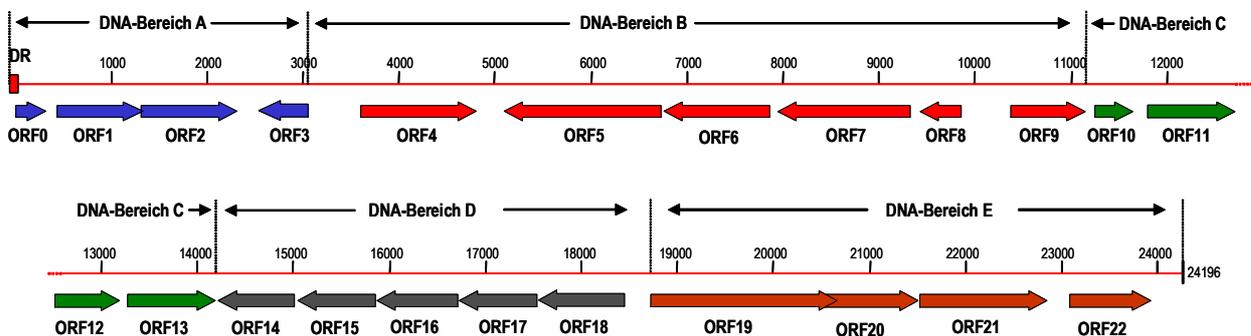


Abbildung 9: Übersicht der gesamten Region X ab dem 'direct repeat' DR, mit den ORFs 0 bis 22.

3.2.1. DNA-Bereich A

3.2.1.1. Sequenzanalyse DNA-Bereich A

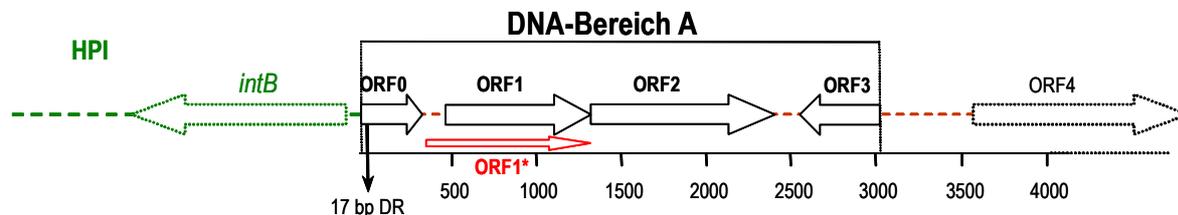


Abbildung 10: DNA-Bereich A mit den ORFs 0-3; Position bp 1 - bp 3042.

Im ersten Abschnitt, dem DNA-Bereich A, wird der an die HPI angrenzende Teil von RegX beschrieben. Die Einteilung erfolgte auf Grund eines potentiellen Gen-Clusters. Der Abschnitt umfasst den 17 bp großen '*direct repeat*' (DR) als Teil der Insertion der HPI. Anschließend lassen sich 4 offene Leseraster ORF0-3 (ORF = *open reading frame*) charakterisieren.

ORF0 (318 bp) beginnt überlappend zu den letzten 4 bp des DR und hat weder ein Startcodon, noch einen Promotorbereich oder eine Shine-Dalgarno-Sequenz. Genauso wie ORF1 und 2 ist es stromabwärts der HPI gerichtet. Das offene Leseraster 1 folgt nach 164 Nukleotiden und hat eine Größe von 813 bp. Da auch hier die Shine-Dalgarno-Sequenz fehlt, wurde die Sequenz genauer analysiert. Insbesondere durch Homologienvergleiche (siehe Abschnitt 3.2.1.2.) fiel dabei auf, dass es im Bereich einer poly-Adenosin-Wiederholungssequenz (*poly adenosine stretch*) zwischen Position bp 356 und bp 360 offenbar zu einer Deletion gekommen ist. Durch das Fehlen einer Adenosin-Base ist der vermutlich ursprüngliche ORF1* (Position bp 353 – bp 1309), der nur 18 bp nach ORF0 folgt, verkürzt (Abbildung 10). ORF1* würde sowohl ein Startcodon (ATG) als auch eine Shine-Dalgarno-Sequenz besitzen.

Unmittelbar an ORF1 schließt ORF2 (1020 bp) an. ORF3 dagegen ist den vorangegangenen offenen Leserastern entgegengerichtet und folgt im Abstand von 205 bp mit einer Größe von 504 bp. ORF3 hat das Startcodon 'TTG'. Eine Shine-Dalgarno-Sequenz ist im Promotorbereich von ORF2 und 3 nachweisbar.

3.2.1.2. Homologienvergleiche DNA-Bereich A auf DNA- und Proteinebene

Wie bereits im Abschnitt 3.1. angedeutet ergab ein initialer Vergleich der ersten 4000 bp der Region X mit veröffentlichten Sequenzen der NCBI-Daten-Bank (BLAST Algorithmus) keine Homologien zu bekannter *E. coli*-DNA (Stand 2002). Die höchsten Homologien (~ 55% Identität) konnten der DNA von *Mesorhizobium loti* zugeordnet werden.

Im Rahmen der schriftlichen Abfassung der Arbeit wurde die Homologiesuche aktualisiert (Stand Juli 2007).

Auf der Basis dieser neuen Homologienvergleiche bestehen auf DNA-Ebene für die ORFs 0-2 zwischen 94% und 96% Identität zu entsprechenden ORFs des Plasmids pLVPK von *Klebsiella pneumoniae* CG43³⁰ (Accessionnr AY378100.1, Tabelle 1 im Anhang), sowie zu ORFs des Plasmids pK2044 von *Klebsiella pneumoniae* NTUH-K2044¹⁶⁸. Auf Grund der Sequenzgleichheit der beiden Plasmide im RegX-orthologen Bereich wird im Folgenden nur auf das Plasmid pLVPK Bezug genommen.

ORF3 hat ebenfalls kurzstreckige Homologien zu *K. pneumoniae*. Für die 205 bp große DNA-Region zwischen ORF2 und 3 gibt es keine Homologien.

Die potentiellen Translationsprodukte der einzelnen ORFs werden im Folgenden **Orf** in Fettdruck und mit der entsprechenden Nummerierung bezeichnet.

Orf0 ist ein Polypeptid mit 105 größtenteils hydrophoben Aminosäuren und einem Molekulargewicht von 11,6 kDa. **Orf0** weist 96% Identität und 98% Ähnlichkeit der Aminosäuren zu dem ATP-bindenden Protein eines ABC-Eisentransporters von *Klebsiella pneumoniae* auf. Deutlich geringere Homologien, zwischen 60-70% Ähnlichkeit, zeigen sich z.B. zu der ATP-bindenden Komponente eines Eisen(III)-ABC-Transporters in *Erwinia carotovora subsp. atroseptica* SCRI1043 und *Vibrio splendidus* 12B01.

ABC-Transporter sind ubiquitär verbreitet und bei Prokaryonten wie Eukaryonten an diversen physiologischen Prozessen beteiligt. Sie sorgen vorwiegend für den Transport verschiedener Substanzen entlang biologischer Membranen und können dabei als Exporter, Importer, Rezeptoren oder Kanäle fungieren⁶⁸.

Sie sind aus mindestens vier Domänen aufgebaut, die als separate Polypeptide, aber auch in Form einer einzigen Polypeptidkette vorkommen können. Zwei Domänen durchspannen als alpha-Helix mehrfach die innere Membran bei gramnegativen Erregern und ebnet so den Weg für das zu transportierende Substrat (Permease, Transmembranproteine). Die zwei nukleotidbindenden Domänen sind über ATP-Hydrolyse für die Energiebereitstellung des transmembranösen Transportes verantwortlich. Zusätzlich zu diesen vier Hauptdomänen gibt es bei Prokaryonten oft ein periplasmatisches Bindungsprotein, welches eine hohe Affinität zum Substrat hat und mit den membrandurchspannenden Proteinen interagiert⁸⁶ (Abbildung 11).

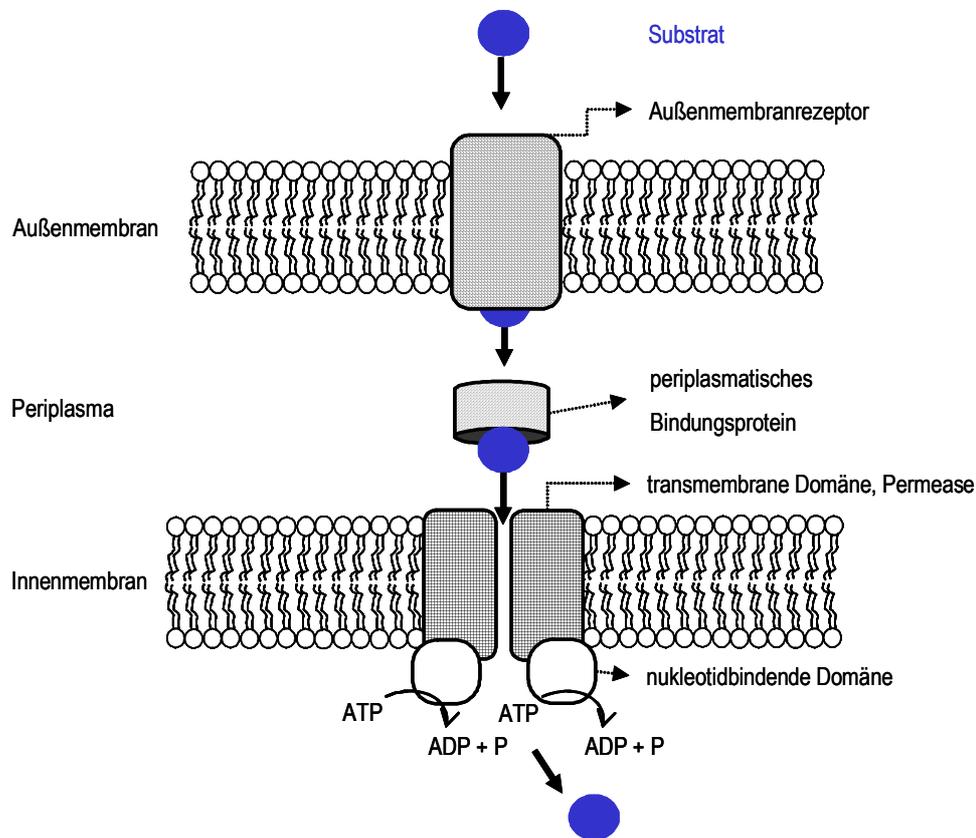


Abbildung 11: Modell eines ABC-Transporters in gramnegativen Bakterien

Das ATP-bindende Protein weist insgesamt fünf verschiedene stark konservierte Regionen auf, die auch in Transportern unterschiedlicher Klassen konserviert sind. Darunter das *Walker A*- und *Walker B*-Motiv, die bei dem größten Teil ATP-bindender Proteine vorkommen und direkt in die Bindung und die Hydrolyse von ATP involviert sind. Daneben existieren die spezifischeren Motive, das *Signature*-Motiv,

Histidin- und *Glutamin*-Loop, die die nukleotidbindende Domäne eines ABC-Transporters definieren⁹⁶.

Da ORF0 unvollständig ist, weist das potentielle Protein **Orf0** nur die C-terminal gelegenen *Walker B*- und *Histidin*-Motive auf (Abbildung 12). Das *Walker B* Motiv ist gekennzeichnet durch die, nach vier hydrophoben Aminosäuren (bei **Orf0**: Leucin, Valin, Leucin, Leucin) folgenden, Aminosäuren Asparaginsäure (D), Glutaminsäure (E), Prolin (P) und Tryptophan (T), in genannter Reihenfolge. Etwa 30 Aminosäuren C-terminal dieser Asparaginsäure erkennt man die zweite konservierte Struktur bei **Orf0**, den *Histidin-Loop*. Dieser ist gekennzeichnet durch 4 hydrophobe Aminosäuren (bei **Orf0**: Valin (V), Alanin (A), Valin (V), Leucin (L)) gefolgt von der Aminosäure Histidin und einer geladenen Aminosäure (bei **Orf0**: Asparaginsäure D). **Orf0** stellt den C-terminalen Anteil eines ATP-bindenden Proteins dar, ist aber ohne den N-terminalen Anteil nicht vollständig. Die anderen, oben genannten Motive *Walker A* (G X X G X G K S T), das *Signature* Motiv (L S G G) und der *Glutamin-Loop*, sind daher nicht nachweisbar.

		Signature		Walker B		His-Loop	
Orf0	1	-----	LAQTF	DLVLLDE	PTNHL..KGITV	VAVLH	D-LTLTEsFADRVLILLSQG 59
MalK	82	1LNRRKPRE	LSGGQR	ORVALGRAIVRKE	QVFLMDE	PLSNL..lgVTT	IYVTH dqvEAMTm-gDRIAVMNRG 216
PotG	119	fAKRKPHQ	LSGGQR	ORVALARSLAKRE	KLILLDE	PMGAL..vgVTC	VMVTH dqeEAMTm-aGRIAIMNRG 253
OpuBA	77	yVDRYPHE	LSGGQR	ORIGVLRALAAEP	PLILLMDE	PFGAL..lhKTT	VFVTH dmdEAIK1-aDRIVILKAG 212
CysA	77	lADRYPAQ	LSGGQR	ORVALARALAVEP	QILLLDE	PFGAL..lkFTS	VFVTH dqeEATEv-aDRVVMSQG 213
FecE	79	lAVRRLTE	LSGGQR	ORAFLLAMVLAQNT	PVLLLDE	PTTYL..qgKTV	VAVLH dlnQASRY-cDQLVVMANG 214
FepC	84	lADQSVDT	LSGGQR	ORAWIAMVLAQET	AIMLLDE	PTTWL..kgYTI	AAVLH dlnQACRY-aSHLIALREG 220
Rv2564	86	rMKHRPGD	MSGGQR	ORVAVARAIALDE	QLLLADE	PTAHL..gdRVV	VVATH d-sRMLP1-aDRVLELMPA 218
AbcX	74	lLGRIYSS	LSQGEK	QLLLIARSLMEDP	EIIILLDE	ATSGL..haPTI	LYVTH haeEITDk-mSHILLRRG 219

MalK (*Thermococcus litoralis*); PotG (*E. coli*), OpuBA (*Bacillus subtilis*), CysA (*E. coli*), FecE (*E. coli*), FepC (*E. coli*), Rv2564 (*Mycobacterium tuberculosis*), AbcX (*Streptococcus mutans*).

Abbildung 12: Orf0 mit den konservierten Motiven eines ATP-bindenden Proteins, *Walker B*-Motiv und *Histidin-Loop*.

Entsprechend zeigen die Polypeptide der ORFs 1 und 2 auf Proteinebene Homologien zu analogen Komponenten eines Eisentransporters, ebenfalls mit sehr hohen Homologien zu dem Plasmid pLVPK von *Klebsiella pneumoniae*³⁰.

94% der Aminosäuren von **Orf1** sind identisch zu einem Polypeptid (LV126), das als mögliches periplasmatisches Protein eines Fe³⁺-ABC Transporters beschrieben ist (Abbildung 11).

Hier ergeben sich bei der Suche nach konservierten Domänen Ähnlichkeiten zu den periplasmatischen Komponenten BtuF, FepB, FhuD, FecB entsprechender Eisentransport-Systeme und dem Rezeptor TroA.

Metall-bindende Proteine der Superfamilie TroA beispielsweise fungieren als initiale Rezeptoren in ABC-Transportern von Metallionen in Eubakterien. Ein typisches TroA Protein setzt sich aus zwei, über eine Einzelhelix verbundene, globuläre Unterbereiche zusammen und bindet die Metallionen zwischen diesen Domänen. FepB gehört als periplasmatische Komponente zu dem Fe³⁺-Hydroxamat-Transport-System. BtuF ist ein Cobalamin-bindendes Protein und gehört zur oben erwähnten TroA Superfamilie. Zusätzlich zu den bereits erwähnten Strukturmerkmalen besitzen diese Proteine eine repetitive HDH-Sequenz, d.h. ein histidinreiches Motiv.

Orf2 ist ein hydrophobes Polypeptid und stellt mit 329 von 333 identischen Aminosäuren zum Polypeptid LV125 von *Klebsiella pneumoniae* die Permease eines potentiellen Transporters, ähnlich FecCD, FepD oder BtuC (Abbildung 11). Die Permease ist das Transmembranprotein, die das Substrat entlang der Membran transportiert.

Vor der Veröffentlichung der Genom-Sequenz von pLVPK von *K. pneumoniae* lagen die höchsten Homologien auch von **Orf0**, **1** und **2** mit identischen Aminosäuren zwischen 44-63%, zu *Mesorhizobium loti* und *Pseudomonas aeruginosa* PA01.

Die Homologien zu Komponenten eines ABC-Transporters und die Ausrichtung der ORFs 0, 1 und 2 deuten auf ein Gencluster hin. ORF3 ist diesem Cluster entgegengerichtet. Das Protein **Orf3** mit 167 Aminosäuren zeigt deutlich niedrigere Homologien zu *Klebsiella pneumoniae*. Auf Proteinebene sind 90 von 165 Aminosäuren identisch zu dem putativen Protein LV197 auf pLVPK von *Klebsiella pneumoniae* mit einer Länge von 174 aa. Darüber hinaus ergeben sich nur geringe Homologien mit ca. 53% Ähnlichkeit und 32% Identität der Aminosäuren zu Regulatorproteinen in *Sodalis glossinidius*.

Interessanterweise finden sich in **Orf3** konservierte Strukturen aus der FUR Familie (= *ferric uptake regulator*). Dies sind Metallionen-abhängige Regulatorproteine (Repressor) deren Produkte bei der Aufnahme von Metallionen beteiligt sind, wie etwa Fur oder Zur (*zinc uptake regulator*). Eine typische Bindungsstelle für den Fe²⁺-Fur-Komplex, die so genannte Fur-Box mit der Sequenz

'GATAATGAT(AT)ATCATTATC' ist im Promotorbereich von ORF3 nicht nachweisbar.

Zusammengefasst enthält der vorliegende DNA-Abschnitt A ein potentielles Gen-Cluster mit hohen Homologien zum Plasmid pLVPK von *Klebsiella pneumoniae* und mit konservierten Domänen eines putativen ABC-Eisentransporters. Anteile des ATP-bindenden Proteins finden sich im unvollständigen ORF0. ORF1 kodiert offenbar für das periplasmatische Bindungsprotein/ den Rezeptor und ORF2 für die Permease. ORF3 hingegen gehört nicht zu diesem Cluster, kodiert jedoch vermutlich für ein putatives Regulatorprotein mit Ähnlichkeiten zu den Regulatoren Fur, Zur und Mur.

3.2.1.3. Homologienvergleiche DNA-Abschnitt A mit pLVPK von *Klebsiella pneumoniae* (AY378100) und mit *Mesorhizobium loti* MAFF303099 (AP003013)

Auf dem Plasmid pLVPK von *Klebsiella pneumoniae* CG43 gibt es ein den ORFs 0-2 entsprechendes Gen-Cluster LV127-LV125 (Homologien zwischen 94%-96%), dessen Translationsprodukte als Proteine eines möglichen ABC-Transporters beschrieben sind. Ein genauer Vergleich der beiden Regionen ist in der Abbildung 13 dargestellt.

Bemerkenswerterweise ist ORF0 (318 bp) weniger als halb so groß, wie das entsprechende offene Leseraster LV127 von *Klebsiella pneumoniae* (828 bp). Die hohen Übereinstimmungen auf Proteinebene beginnen erst ab Aminosäure-Position 171. Zusammen mit der Größe des offenen Leserasters legt dies die Vermutung nahe, dass ORF0 nicht vollständig ist.

Die Homologien zwischen LV126 und ORF1 beginnen unmittelbar nach ORF0, jedoch ist ORF1 auf Grund einer *frameshift*-Mutation im Bereich eines 6-Adenosin-Polymers kürzer als LV126 (schwarz gestrichelte Linie in Abbildung 13).

Nach ORF2, mit ähnlicher Größe wie LV125 (1035 bp), kommt es zu einem Abbruch der Homologie zu *Klebsiella pneumoniae*. Erst im Bereich von ORF3 bestehen Homologien über kurze DNA-Sequenzen von 39 bp und 97 bp, wie in Abbildung 13 dargestellt. Die Homologien betreffen das Gen LV197, das dem ABC-Transporter-Operon auf dem Plasmid pLVPK nicht benachbart ist.

Auch im Bereich des ABC-Transporters von *Mesorhizobium loti* (*mlr8227-mlr8229*), zu dem RegX vor Veröffentlichung der Plasmidsequenz von pLVPK die höchsten Homologien hatte, gibt es kein orthologes Gen für ORF3 (Abbildung 13). *mlr8230* kodiert zwar für einen putativen Transkriptionsregulator aus der Familie der Mar-Proteine (*multiple antibiotic resistance*), weist jedoch eine entgegengesetzte Leserichtung zu ORF3 auf und zeigt auch sonst keine Homologien zu ORF3.

Interessanterweise grenzt weder bei *K. pneumoniae* noch bei *M. loti* das ABC-Transporter-Operon an die *Yersinia*-HPI, wie es bei ECOR31 der Fall ist.

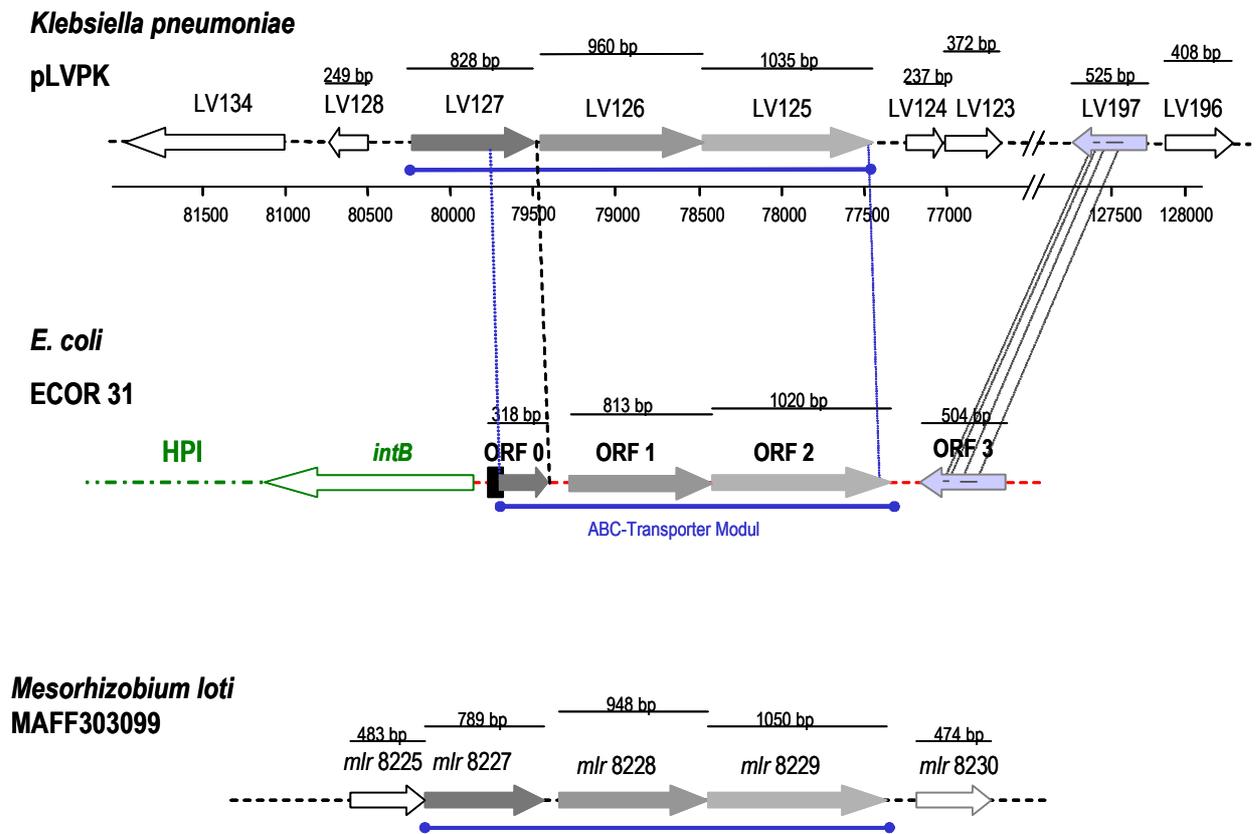


Abbildung 13: Homologievergleiche des DNA-Bereiches A mit dem Plasmid pLVPK von *Klebsiella pneumoniae* CG43 (AY378100) und *Mesorhizobium loti* MAFF303099 (AP003013). Das Cluster mit den ORFs 0-2 entspricht der ABC-Transporter-Region von pLVPK (LV127-LV125) und der von *M. loti* (*mlr8227-mlr8229*). Für ORF3 finden sich nur kleinere Homologiesequenzen aus dem Bereich LV197.

3.2.1.4. Transkriptionsuntersuchungen ABC-Transporter (RT-PCR)

Die strukturellen Untersuchungen des DNA-Bereichs A zeigen ein Cluster ORF0, 1 und 2 mit großer Ähnlichkeit zu einem ABC Transporter. Da ORF0 aber unvollständig ist und keinen Promotor hat, ist eine Expression fraglich.

Um die Transkription der ersten drei offenen Leseraster zu überprüfen, wurde eine Reverse-Transkriptase Polymerasekettenreaktion (RT-PCR) mit Primern aus diesem Bereich durchgeführt. Sofern die Gene exprimiert werden, ist der erste Schritt die Umschreibung der DNA in RNA (messenger RNA). Über Transkriptionsuntersuchungen mittels RT-PCR lassen sich daher erste Hinweise auf eine mögliche Genexpression erhalten.

Das Ausgangsmaterial der Untersuchung ist die RNA von ECOR31, die mit einer reversen Transkriptase in cDNA umgeschrieben wird. In einer PCR wird anschließend die cDNA mit Hilfe einer Taq-Polymerase und speziell ausgewählten Primern amplifiziert. Um DNA-Kontaminationen auszuschließen und sicherzustellen, dass positive PCR-Ergebnisse auf spezifische RNA zurückzuführen sind, wird zum einen vor der PCR eine zusätzliche DNase Behandlung durchgeführt, die eventuell noch vorhandene DNA abbaut. Zum anderen gibt es von jeder Probe eine Negativkontrolle ohne Reverse-Transkriptase-Zusatz.

Interessanterweise werden alle Gene des potentiellen ABC-Transporters transkribiert (Abbildung 14). Die Gel-Elektrophorese der Primerprodukte 2 und 3 zeigt die zu erwartenden Banden von 282 bp bzw. 311 bp in den Bereichen von ORF0 und ORF1 (Primerpaar 2: E31-RegX-5.for und RegX-6089.rev; Primerpaar 3: RegX-ORF1-RT.for/rev). Das bedeutet, dass auch die beiden unvollständigen Gene exprimiert werden. Da vor dem unvollständigen ORF0 kein eigener Promotor existiert, muss es im Bereich der angrenzenden HPI einen bidirektionalen Promotor geben. Mit dem Primerpaar 1 (RegX.5781.for und RegX-6089.rev), das den Übergang HPI - ORF0 amplifiziert, konnte diese Vermutung durch ein positives PCR-Ergebnis bestätigt werden (Abbildung 14).

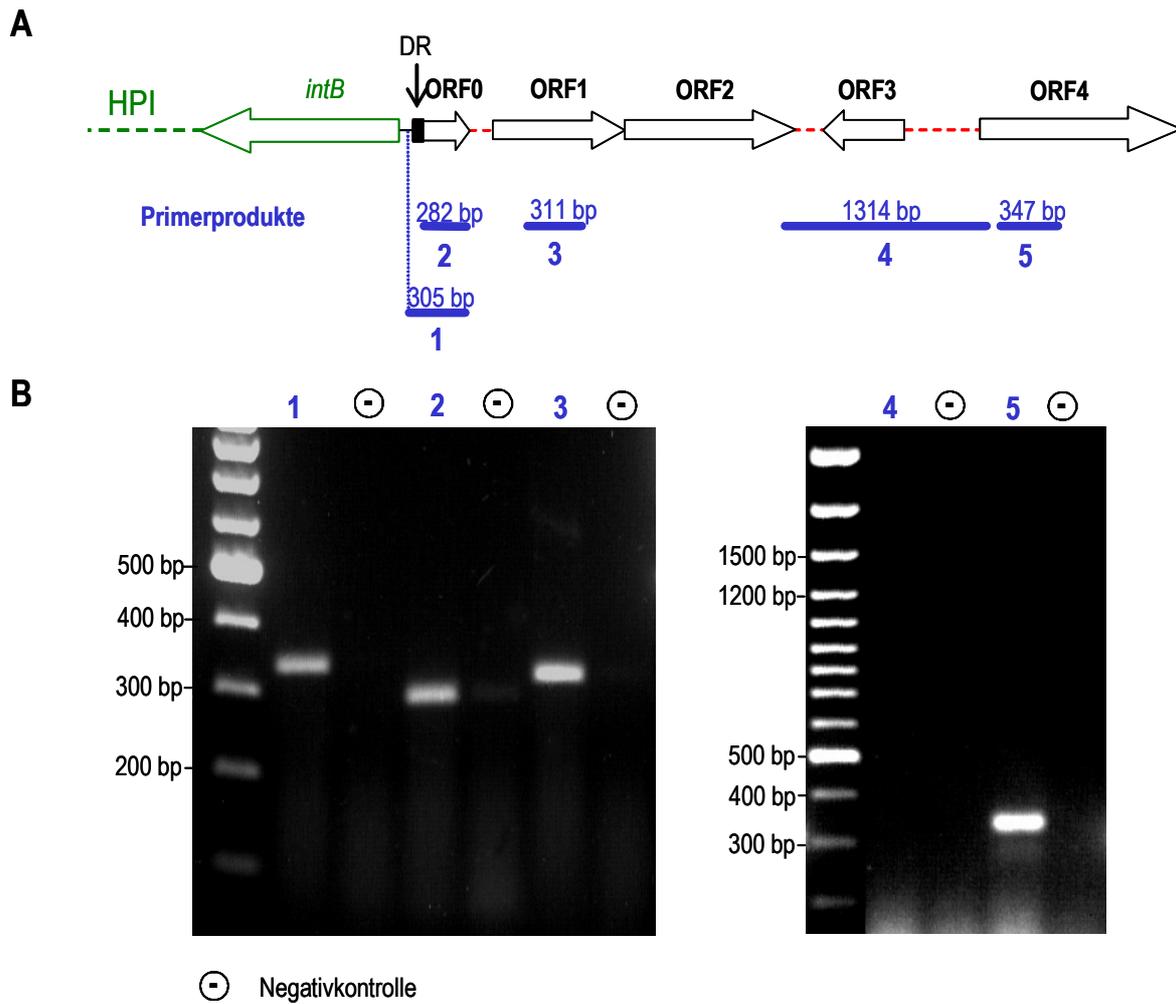


Abbildung 14: RT-PCR des ABC-Transporter-Clusters ORF0-2 (Gelelektrophorese). A: Lokalisation der Primerpaare 1 – 5 auf RegX. B: positives RT-PCR-Ergebnis für die Primerpaare 1, 2, 3 und 5.

Mit dem Primerpaar 5 (RegX-ORF4-RT.for/rev) wurde eine Transkription von ORF4 mit Hilfe der RT-PCR überprüft und durch eine Bande von 347 bp bestätigt. ORF4 wird im Abschnitt DNA-Bereich B besprochen (siehe Unterpunkt 3.2.2.).

Es handelt sich um zwei verschiedene Transkriptionseinheiten (mRNAs), wie durch das negative Ergebnis des Primerpaares 4 (RegX.ORF2-8117.for, RegX.ORF4-9413.rev) gezeigt wird.

3.2.2. DNA-Bereich B

3.2.2.1. Sequenzanalyse DNA-Bereich B

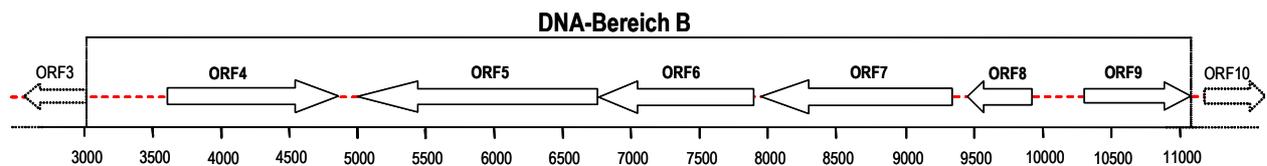


Abbildung 15: Übersicht DNA-Bereich B mit den ORFs 4 - 9; Position bp 3043 – bp 11169.

Der DNA-Bereich B ist zwischen der Position bp 3043 und bp 11169 lokalisiert und umfasst die ORFs 4-9. Alle offenen Leseraster besitzen ein Startcodon ('ATG'). Eine Shine-Dalgarno Sequenz fehlt lediglich bei ORF9. 562 bp nach ORF3 beginnt der ORF4 mit einer Größe von 1224 bp. Nach 266 bp folgen in entgegengesetzter Orientierung die ORFs 5, 6, 7 und 8. ORF5 ist mit 1650 bp das größte offene Leseraster. Sowohl zwischen ORF6 (1131 bp) und ORF7 (1383 bp), als auch zwischen ORF7 und ORF8 (453 bp) liegt ein nicht-kodierender Bereich von 70 respektive 81 bp. Das offene Leseraster 9 (801 bp) folgt erst nach 495 bp und ist ausgerichtet wie ORF4.

3.2.2.2. Homologievergleiche DNA-Bereich B auf DNA- und Proteinebene

Auf DNA-Ebene ergeben sich für den gesamten DNA-Bereich B hohe Homologien zum Plasmid pLVPK von *Klebsiella pneumoniae*.

94% der Aminosäuren des putativen Polypeptides **Orf4** sind identisch zu einem als CobW bezeichneten Protein von *K. pneumoniae*. Dieses wird von dem Gen LV195 auf dem Plasmid pLVPK kodiert. Geringere Homologien finden sich ebenfalls zu einem Cobalamin-Synthese-Protein, von *Burkholderia sp.* (73% Identität, Genregion *cobW-C*).

Beide besitzen, wie auch **Orf4**, als konservierte Struktur eine nukleotidbindende Domäne, den Glycin-reichen Loop (*P-Loop*, phosphatbindender-Loop), oder auch *Walker A*-Motiv genannt (Abbildung 16).

P-Loop, Walker A

		GXXXXGK	
Orf4	15	PVTVLS	GFLGAGK
YjiA	5	AVTLLT	GFLGAGK
SAV0450	5	PVTVLS	GYLGS GK
YciC	5	PVTVLS	GYLGAGK
BH1790	7	PVTVLS	GYLGAGK
P47K	11	PVTVLS	GFLGAGK
PA5535	6	PVTVLS	GFLGAGK
Orf1188	7	PVTVLS	GFLGAGK
XF_1830	13	PVTVFS	GFLGAGK
COG0523	6	PVTVLS	GFLGAGK
HypB/ <i>E. coli</i>	1	[18].LS	GFLGAGK
MK1546	14	[23].VL	GSIGSGK
HypB/ <i>Sin</i>	1	[11].IG	GPVGS GK

YjiA Protein (*E. coli*); SAV0450 (*Staphylococcus aureus* subsp. *aureus* Mu50); YciC Protein (*Bacillus subtilis*); BH1790 Protein (*Bacillus halodurans*); P47K Protein (*Pseudomonas chlororaphis*); PA5535 Protein (*Pseudomonas aeruginosa*); Orf1188 (*Rhodococcus sp.*), XF_1830 (*Xylella fastidiosa*), COG0523 (*Burkholderia pseudomallei*)

HypB (*Sinorhizobium meliloti* 1021), MK1546 (*Methanopyrus kandleri* AV19)

Abbildung 16: N-terminale Region von **Orf4** mit dem konservierten *P-Loop*- respektive *Walker A*-Motiv (Consensus 'GXXXXGK').

Diese konservierten Domänen werden auch bei anderen Proteinen gefunden, wie z.B. HypB und UreG. Bei diesen Proteinen handelt es sich um Nickel-bindende GTPasen, die an der Regulation und Reifung von Urease und Hydrogenase beteiligt sind¹⁰⁶. Diese Enzyme sind maßgeblich für die Virulenz von *Helicobacter pylori* verantwortlich⁴⁸, dem Verursacher von Gastritis und Ulcera des Magens und Duodenums.

Das CobW Protein, zu dem **Orf4** die größte Ähnlichkeit besitzt, ist in die Cobalaminbiosynthese involviert. Die Synthese von Cobalamin, oder Vitamin B12, ist auf Grund der Größe und Komplexität sehr aufwendig und beteiligt etwa 25 Enzyme¹⁴². Am besten untersucht ist die Synthese des Corrinoid-Ringes in *Pseudomonas denitrificans*. In dieser Spezies, in dem die Vitamin B12 Synthese nur in Anwesenheit von Sauerstoff stattfindet, wurde auch CobW beschrieben. Es ist ein Bestandteil des *cob*-Operons und ist in die Reduktion von Kobalt involviert³². Neben dem aeroben Stoffwechselweg können Bakterien auch unter anaeroben Bedingungen Vitamin B12 herstellen. Man geht davon aus, dass die Synthese von Vitamin B12 entstanden ist, um die Fermentation kleinerer Moleküle in Abwesenheit eines externen Elektronenakzeptors zu ermöglichen¹⁴².

Bei **Orf5** handelt es sich um ein mögliches Fe²⁺-Transporter-Protein. Diese Einschätzung beruht auf vorhandenen Ähnlichkeiten zu Eisen-Transporter-Proteinen, z.B. bei *Klebsiella pneumoniae* (488 von 549 Aminosäuren identisch) und bei Cyanobakterien. Zudem findet man auch charakteristische konservierte Bereiche mit Homologien zu **FeoB**, dem Fe²⁺-Eisentransporter Protein von *Escherichia coli*. Das 700-800 Aminosäuren lange integrale Membranprotein FeoB ist zwar größer als **Orf5**, das nur 549 Aminosäuren umfasst. Ähnliche Proteine nach BLAST weisen teilweise sogar noch geringere Größen auf (z.B. Protein SA2369 von *Staphylococcus aureus*, NP_375695). FeoB besteht aus zwei Haupt-Regionen, einer hydrophilen N-terminalen Region mit GTPase-Aktivität und einer hoch-konservierten Guanin Nukleotid-bindenden Sequenz, die für die Eisenaufnahme benötigt wird. Der zweite Bereich umfasst eine hydrophobe C-terminale Region, mit 7-12 membrandurchspannenden α -helikalen Proteindomänen, von denen sieben konserviert sind¹⁰¹. Der hochkonservierte Guanin Nukleotid-bindende Bereich befindet sich in den ersten 160 aa und wird durch fünf Loop-Regionen, die GTPase-Motive enthalten, charakterisiert¹⁰¹. Bei **Orf5** sind davon das G1-Motiv

(GXXXXGK(S/T)), das G3-Motiv (DXXG) und teilweise das G4-Motiv (NXXD) vorhanden (Abbildung 17). Vom G5-Motiv ((T/G)(C/S)A) ist nur die Aminosäure Alanin konserviert. Im mittleren Teil des Proteins findet sich eine so genannte *gate*-Struktur, die eine Membran-Pore darstellt. Der C-terminale, hydrophobe Teil ist insgesamt weniger gut konserviert, dennoch gibt es Homologien zu den FeoB-Vergleichsproteinen.

		G1 (GXXXXGKs/t)	G3 (DXXG)	G4 (NXXD)	G5 (T/G) (C/S) A
Orf5	6	LM GLESAGKS TLFNILT..ICLV DTPG I..PQAKHVVILL TFAD KIRKGLLEV..PVLA VNA RE.			
Thcr	16	LI GSPNSGKT TLFNRLT..YHIT DLPG V..HMGLPVVVVL NQMD KLENHHLEL..PVIPI ISA YY.			
Strmut	6	LI GNPNSGKT SLFNLLT..ISVC DLPG I..ETGIPVTIAL NMSD VLKNQGKSI..PVVA TSA LK.			
Geobkau	6	LF GNPNTGKT SLFNCLT..VPVV DLPG V..EFGRPVLVIAL NMVD VAEKRGVKV..PVIPI VIA RT.			
Archfulg	1	MV GNPNVGKT SLLNALT..VTFV DLPG I..EFGIPMIVVL NMID EAEKEGIRI..PVPVK TSA VR.			
Baccer	8	LA GNPNTGKS TLFNLT..YTLI DLPG I..EMTSNVVICI NLID EAEKKGIVI..PVPVK ISA RN.			
Porging	39	LA GNPNTGKS TVFNALT..YKIV DLPG I..EITDRAVLCA NLMD EAEERNGVEI..PVPVG ASA RS.			
Solibus	28	LV GPPNSGKT TLFNRLT..VTIV DLPG I..ALGLPTLIIL NMAD DLRNRGGKL..PVALI ISA RA.			
Methcaps	6	LV GMPNTGKS TFFNRLT..AQVV DLPG I..RLGLPTVLLL NMAD EARHLGISI..PVIL LSA KY.			

Thiomicrospira crunogena XCL-2, *Streptococcus mutans* UA159, *Geobacillus kaustophilus* HTA426, *Archaeoglobus fulgidus* DSM 4304, *Bacillus cereus* G9241, *Porphyromonas gingivalis* W83, *Solibacter usitatus* Ellin6076, *Methylococcus capsulatus* str. Bath

Abbildung 17: Konservierte Motive des FeoB orthologen Polypeptides **Orf5**. Dargestellt sind FeoB-Proteine unterschiedlicher Spezies mit den Motiven G1 (GXXXXGKs/t), G3 (DXXG), G4 (NXXD) und G5 [(T/G)(C/S)A].

Orf6 ist ein 41,6 kDa großes Polypeptid und zu 98% identisch mit einem putativen Protein, das vom Gen LV193 auf dem Plasmid pLVPK von *Klebsiella pneumoniae* kodiert wird. Wesentlich geringere Homologien (44% identische Aminosäuren) ergeben sich zu möglichen Monooxygenasen von Cyanobakterien (*Crocospaera watsonii*) und NAD-bindenden DNA-Abschnitten von *Bacillus licheniformis* (40% Identität).

Kleinere konservierte Bereiche finden sich zu verschiedenen Reduktasen wie der Pyridin-Nukleotid-Disulfit Oxidoreduktase (Pyr_redox_2) von *Bacillus cereus* und *Bacillus thuringiensis*, der Thioredoxin-Reduktase (TrxB) von *Bacillus halodurans* und *B. clausii* sowie der Alkylhydroperoxid-Reduktase (AhpF) von *Bacillus subtilis* und *Salmonella* Typhimurium. Homologien gibt es ebenso zu TrkA, einem Flavoprotein das bei verschiedenen Bakterien, unter anderem bei *E. coli*, am Kalium-Transport beteiligt ist.

ORF7 kodiert ein Polypeptid (**Orf7**) mit Homologien ausschließlich zu Cysteinyl-tRNA Synthetasen verschiedener Spezies. Die höchsten Übereinstimmungen von 60% Identität und 74% Ähnlichkeit ergaben sich zur Cys-tRNA-Synthetase von *Erwinia carotovora* und *Shewanella sp.*

Aminoacyl-tRNA-Synthetasen sind für die sorgfältige Übersetzung der genetischen Information in Proteine essentiell. Sie sorgen für die Verknüpfung der t-RNA mit der dafür vorgesehenen Aminosäure. Letztere wird am Ribosom mit dem entstehenden Polypeptid verknüpft (Translation). Aminoacyl-tRNA Synthetasen müssen daher eine extrem hohe Substrat-Spezifität gewährleisten.

Die monomere Cysteinyl-tRNA-Synthetase gehört zur Familie der tRNA-Synthetasen der Klasse I⁵⁰.

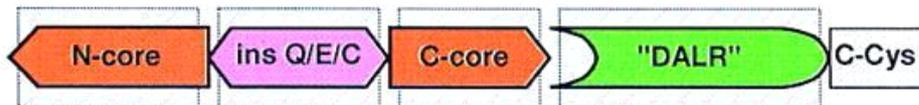


Abbildung 18: Konservierte Motive der Cysteinyl-tRNA-Synthetasen am Beispiel der Cys-tRNA-Synthetase von *E. coli*¹⁷⁸. Rot: katalytische 'Core'-Domäne mit den Motiven 'HIGH' und 'KMSK'. Grün: 'DALR'-Motiv.

Cysteinyl-tRNA-Synthetasen haben mehrere konservierte Motive gemeinsam (Abbildung 18), die sich auch bei Orf7 nachweisen lassen (Abbildung 19). Darunter ist das katalytische Zentrum, auch *Core*-Domäne genannt, das für die ATP-abhängige Bildung des enzymgebundenen Aminoacyladenylates verantwortlich ist. Es weist die charakteristischen Motive der Klasse I Synthetasen auf, mit dem so genannten *HIGH*- und dem *KMSK*-Motiv⁵⁰. Im C-terminalen Bereich findet man die DALR-Domäne, eine Region mit α -Helix-Struktur, die den Anticodon-Bereich darstellt^{50, 178}.

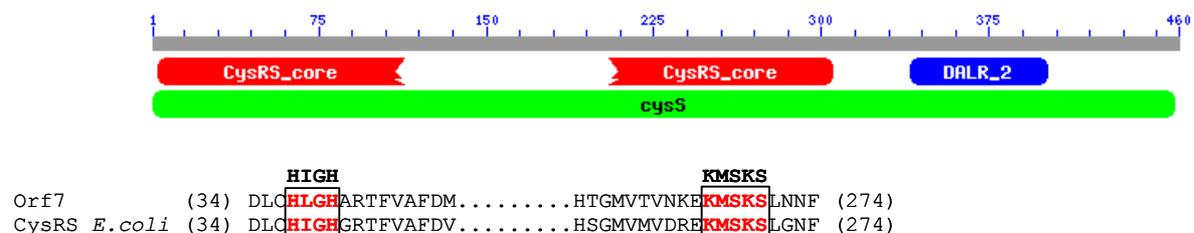


Abbildung 19: Konservierte Motive der putativen Cysteinyl-tRNA-Synthetase **Orf7**. Hervorgehoben sind die Motive der katalytischen *Core*-Domäne *HIGH* und *KMSK*.

Wie ORF3 kodiert auch ORF8 für ein Protein (**Orf8**) aus der Familie der Eisen- und Zink-Aufnahmeregulatoren (siehe Unterpunkt 3.2.1.2., DNA-Bereich A). Dabei handelt es sich um Transkriptionsregulatoren solcher Gene, die für Proteine kodieren, die an der Aufnahme von divalenten Metallionen beteiligt sind. Die im Vorfeld erwähnte Fur-Box-Sequenz (5'-GATAATGATAATCATTATC-3') dient als Zielsequenz für die Bindung des Fur-Proteins. Im Promotorbereich von ORF8 ist keine Fur-Box nachweisbar.

Orf8 weist eine Länge von 150 Aminosäuren auf. Interessanterweise bestehen Homologien zu zwei putativen Proteinen, die an zwei entfernten Stellen auf dem Plasmid pLVPK von *Klebsiella pneumoniae* kodiert werden. 98% der Aminosäuren von **Orf8** sind identisch zu dem putativen Protein des Genes LV192 und 65% der Aminosäuren sind identisch zu dem von LV119 kodierten Polypeptid.

Orf9 hat eine Länge von nur 266 Aminosäuren und zeigt wie das deutlich größere **Orf4** (407 aa) ebenfalls Homologien zum CobW-Protein (C-terminaler Bereich).

Wie bereits bei **Orf8** gibt es auch bei **Orf9** hohe Homologien zu zwei verschiedenen putativen Proteinen von *Klebsiella pneumoniae*, die an voneinander entfernten Stellen im Plasmid pLVPK kodiert sind. Homologien von 98%, respektive 96%-Aminosäure-Identität bestehen zu den Polypeptiden der Gene LV120 und LV191. Geringere Ähnlichkeiten (173 von 265 identischen Aminosäuren) gibt es zu einem putativen Protein aus *E. coli* Nissle, einem in der Medizin als Probiotikum eingesetzten *E. coli*-Stamm.

Zusammenfassend beinhaltet der DNA Bereich B zwei Gene, die in die Biosynthese von Cobalamin involviert sein können (ORF4 und ORF9), sowie ein mögliches Fe²⁺-Transporter-Protein ähnlich FeoB (**Orf5**). Die putative Reduktase **Orf6** könnte bei der Reduktion von Fe³⁺ zu Fe²⁺ beteiligt sein. Ähnlich wie **Orf3** zeigt auch **Orf8** Homologien zu eisenabhängigen Transkriptionsregulatoren wie Fur. **Orf7** hingegen hat ausschließlich Homologien zu Cysteinyl-tRNA-Synthetasen.

3.2.2.3. Homologienvergleiche DNA-Abschnitt B mit dem Plasmid pLVPK von *Klebsiella pneumoniae*

Die höchsten Homologien des DNA-Bereiches B wurden im Abschnitt 3.2.2.2. beschrieben. Bei genauerem Vergleich mit dem Plasmid pLVPK von *Klebsiella pneumoniae* werden mehrere strukturelle Auffälligkeiten deutlich (Abbildung 20).

Der DNA-Abschnitt B ist ab RegX-Position bp 3111 bis bp 11162 zu 93% identisch zu dem Bereich bp 127818 – bp 136364 von *Klebsiella pneumoniae* mit den Genen LV196, CobW, LV194, LV248, LV193, LV245, LV192 und LV191. Interessanterweise ergeben sich auch hohe Homologien zu einem anderen Abschnitt des Plasmids pLVPK. Wie in der Abbildung 20 dargestellt sind einzelne Sequenzabschnitte aus dem Bereich ORF3 und 4, sowie ORF8 und 9 zu 95%-identisch. Im Bereich ORF8 und 9 handelt es sich um eine Inversion der Gene. Die Sequenzen von LV191 und LV120, von *Klebsiella pneumoniae* weisen trotz der Lokalisation an verschiedenen Stellen im Plasmid eine Sequenzidentität von 97% auf. Kleinere, mosaikartige Homologiesequenzen finden sich im Bereich zwischen ORF3 und 4.

Zudem sind mit LV196 und LV194 auf pLVPK zwei Gene beschrieben, für die kein orthologes offenes Leseraster auf RegX ermittelt wurde. Bei fast identischer DNA-Sequenz wurde die RegX-Sequenz genau betrachtet. Für beide Gene gibt es auch auf RegX ein mögliches orthologes Gen. Ein zusätzliches Nukleotid (Cytosin) an Position bp 3328 führt zu einer Rasterverschiebung (*frameshift*) und verhindert die korrekte Translation des LV196-orthologen Genes. Im Bereich des LV194-orthologen Genes bestehen im 3'- und 5'-Bereich hohe Sequenzidentität. Diese Homologie fehlt im zentralen Bereich des Genes und gibt so Hinweise auf ein stattgefundenes Rekombinationsereignis. Eine Shine-Dalgarno-Sequenz, als Zeichen einer Ribosomenbindungsstelle transkribierter mRNA, lässt sich vor dem LV194-orthologen Gen nachweisen.

Das *cobW*-orthologe Gen ORF9 (801 bp) weist eine deutlich geringere Größe auf als die Gene mit der höchsten Homologie auf pLVPK, LV191 und LV120 (1239 bp). ORF9 unterscheidet sich auch von ORF4 (1224 bp), welches ebenfalls Ähnlichkeiten zum *cobW*-Gen zeigt. Bei genauer Betrachtung sind jedoch DNA-Homologien

stromaufwärts von ORF9 nachweisbar. Eine Punktmutation nach Position bp 10167 führte offenbar zu einem *frameshift*, der einen verkürzten ORF9 entstehen ließ.

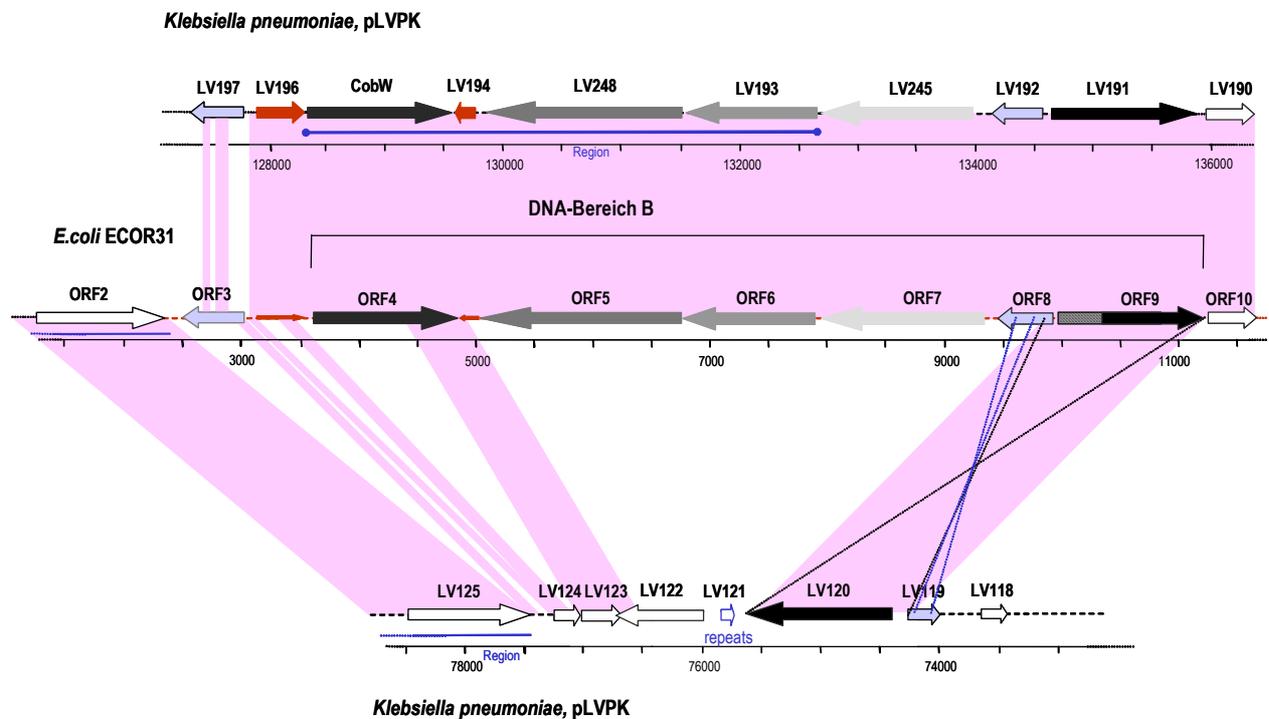


Abbildung 20: Homologien-Vergleiche DNA-Bereich B mit pLVPK von *Klebsiella pneumoniae*. Die rosa unterlegten Bereiche stellen die hohen Homologien von RegX zum Plasmid pLVPK von *K. pneumoniae* dar. Rot dargestellt sind die Gene LV196 und LV194 und die möglichen orthologen Gene auf RegX. ORF9 ist auf Grund eines *frameshifts* verkürzt. Der schwarz-schraffierte Teil stellt die Homologie zu den orthologen Genen dar.

Zusammenfassend ist festzustellen, dass trotz sehr hoher Sequenzhomologie von RegX zu dem Plasmid pLVPK von *Klebsiella pneumoniae*, die beiden Sequenzen nicht identisch sind. Beim Vergleich von RegX und pLVPK sind Punktmutationen nachzuweisen. Sie sorgen neben Inversionen und partiellen Duplikationen für die Plastizität dieser speziellen Genom-Regionen.

3.2.3. DNA-Bereich C

3.2.3.1. Sequenzanalyse DNA-Bereich C

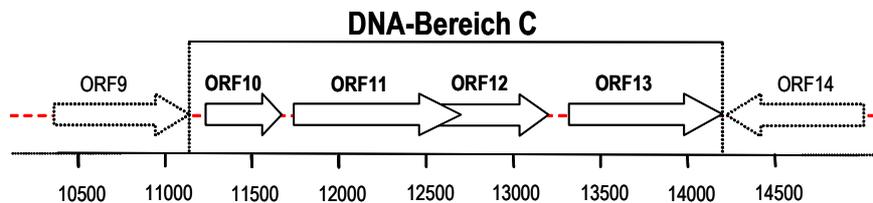


Abbildung 21: Übersicht DNA-Bereich C mit den ORFs 10-13; Position bp 11170 – bp 14207.

Der DNA-Bereich C besteht aus den vier ORFs 10, 11, 12 und 13 (Abbildung 21). Der Abschnitt beginnt unmittelbar nach dem Stoppcodon von ORF9, an Position bp 11170 und erstreckt sich bis Position bp 14207, nach dem Stoppcodon von ORF13. ORF10 folgt dem in DNA-Abschnitt B beschriebenen ORF9 nach 70 Nukleotiden und weist eine Größe von nur 423 bp auf. Nach weiteren 63 bp folgt ORF11 mit 936 bp und überlappend ORF12 an Position bp 12519 mit einer Größe von 693 bp. Nach 66 bp beginnt ORF13 mit 930 bp. Bis auf ORF13 weisen alle ORFs eine typische Shine-Dalgarno-Sequenz auf.

3.2.3.2. Homologienvergleiche DNA-Bereich C auf DNA- und Proteinebene

Auf DNA-Ebene ergeben sich für den gesamten DNA-Abschnitt C, von Position bp 11170 bis Position bp 14207 (3038 bp), hohe Homologien von bis zu 96% Identität zu dem Plasmid pLVPK von *Klebsiella pneumoniae*.

93% der Aminosäuren von **Orf10** sind identisch mit dem Polypeptid kodiert von LV190 von *Klebsiella pneumoniae*. Weitere Homologien bestehen zu Transkriptionsregulatoren verschiedener Spezies, z.B. von *Serratia proteamaculans*. Die größte Ähnlichkeit besteht zu DksA, einem DnaK Transkriptionsregulator (DnaK

Suppressor Protein). DnaK ist das am besten charakterisierte HSP-70 (*heat shock protein*, 70 kDa) Hitzeschockprotein und ein negativer Modulator der Hitzeschockantwort. Hitzeschockproteine dienen z.B. der Stabilisierung zellulärer Proteine unter Extrembedingungen, wie Hitze, Nahrungsmangel, UV-Bestrahlung etc., indem sie andere Proteine u.a. bei der Faltung oder dem Erhalt der Sekundärstruktur unterstützen. Sie sind ubiquitär und über Artgrenzen hinweg konserviert. DnaK ist das einzige HSP-70-ähnliche Hitzeschockprotein in *E. coli* und wurde entdeckt, da es für das Phagenwachstum essentiell ist. Es bildet mit DnaJ ein Operon⁷⁹.

Orf11 kodiert ein Polypeptid mit einer Länge von 311 Aminosäuren. Dieses ist bis auf 3 Aminosäuren praktisch identisch mit LV189 von *Klebsiella pneumoniae*. Beide gehören zu einer größeren Gruppe strukturgleicher Proteine, deren Funktion unbekannt ist (DUF198 und COG1469). Proteine aus dieser Gruppe wurden v.a. bei Pseudomonaden beschrieben.

Für ORF12 gibt es auf Proteinebene (**Orf12**) 196 von 198 Aminosäuren Übereinstimmung mit dem als potentielle Carboanhydratase beschriebenen Translationsprodukt von LV188 von *Klebsiella pneumoniae*. Hohe Ähnlichkeiten finden sich auch zu der Carboanhydratase von *Serratia proteamaculans* (82% Id.) und der verschiedener Pseudomonaden (~66% Id.).

Die Carboanhydratase gehört zu der Familie der Metalloproteine und war eines der ersten Enzyme, für das Metalle als Kofaktor beschrieben wurden. Die Enzymklasse ist ubiquitär in allen bekannten Organismen und besitzt eine große Strukturvielfalt. Die Carboanhydratase enthält als Kofaktor ein Zink-Ion (Zn^{2+}), das im aktiven Zentrum an drei Imidazolreste gebunden ist, die je von einer im Protein enthaltenen Aminosäure Histidin stammen. Die Carboanhydratase katalysiert die Reaktion: $CO_2 + H_2O \leftrightarrow HCO^{3-} + H^+$, eine der schnellsten enzymatischen Reaktionen überhaupt. Diese spielt im menschlichen Körper, bei Tieren, Pflanzen und Mikroorganismen eine große Rolle. *Helicobacter pylori* beispielsweise benutzt die Pufferfunktion bei der Besiedelung des Magens, angesichts des dort vorherrschenden niedrigen pH-Wertes¹⁰⁰.

Für **Orf13** finden sich Homologien zu dem Polypeptid LV187 von *Klebsiella pneumoniae* (97% Identität, 98% Ähnlichkeit). Eine genaue Funktion ist weder für das Polypeptid von *Klebsiella* noch für andere Proteine mit Ähnlichkeiten zu **Orf13** bekannt, wie z.B. *Ping_2266* von *Psychromonas ingrahamii* (siehe Tabelle 1 im Anhang). **Orf13** besteht vorwiegend aus hydrophoben Aminosäuren (118 aa hydrophob, 76 polar), was auf eine transmembrane Struktur deutet.

Zusammenfassend lässt sich zeigen, dass die RegX-Sequenz im Bereich des DNA-Abschnitts C nahezu identisch mit der Position 135872 - 138909 des Plasmids pLVPK von *Klebsiella pneumoniae* ist. Auch die Translationsprodukte der Gene LV187, LV188, LV189 und LV190 sind bis auf eine potentielle Carboanhydratase (LV188) nicht genauer beschrieben. **Orf10** enthält konservierte Domänen von Transkriptionsregulatoren der Hitzeschockproteine.

3.2.4. DNA-Bereich D

3.2.4.1. Sequenzanalyse DNA-Bereich D

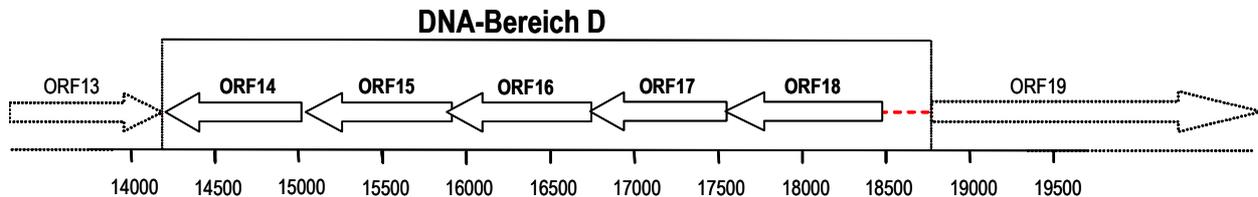


Abbildung 22: Übersicht DNA-Bereich D mit den ORFs 14-18; Position bp 14208 – bp18718.

Auf Grund der ORF-Ausrichtung, sowie der potentiellen Anordnung in einem Gencluster wurde die Region zwischen Position bp 14208 - bp 18718 als DNA-Bereich D definiert. Der Bereich besteht aus den fünf gleichgerichteten offenen Leserastern ORF14 - ORF18. Die ORFs sind in Richtung HPI angeordnet und überlappen einander größtenteils. Lediglich ORF14 und ORF15 sind 15 Nukleotide voneinander entfernt. Die Größe der einzelnen ORFs ist sehr ähnlich und liegt zwischen 810 - 915 bp (ORF14: 810 bp; ORF15: 828 bp; ORF16: 861 bp; ORF17: 822 bp; ORF18: 915 bp). Alle ORFs besitzen ein Startcodon ('ATG') und jeweils upstream davon eine Shine-Dalgarno Sequenz.

3.2.4.2. Homologienvergleiche DNA-Bereich D auf DNA- und Proteinebene

Die hohen Homologien aus den Abschnitten des DNA-Bereichs A bis C zwischen RegX und pLVPK reißen im DNA-Bereich D ab. Lediglich die ersten 738 bp vom DNA-Abschnitt D bis zur Position bp 14945, sind auf DNA-Ebene zu 93% mit dem Plasmid pLVPK von *Klebsiella pneumoniae* identisch. Die Homologie zu pLVPK besteht erst wieder am Ende des Bereichs D. Hier zeigen die letzten 78 bp 92% Identität. Für die 3695 Basenpaare (Position bp 14946 – bp 18640) in der Mitte des

Abschnittes gibt es auf DNA-Ebene Homologien von 70% Sequenzidentität zu *Shigella sonnei* und *Shigella flexneri*, sowie zu verschiedenen Serovaren von *Salmonella enterica*, ebenso zu *Enterobacter sp.* und *E. coli*.

Das von ORF14 kodierte Polypeptid **Orf14** hat 87% Aminosäuren-Identität mit einem Protein von *Klebsiella pneumoniae* (Genlocus LV186). Deutlich niedrigere Homologien (knapp 50% identische Aminosäuren) finden sich zu Salmonellen (*Salmonella enterica subsp. enterica* Serovar Paratyphi A und Typhimurium LT2). Diese beiden orthologen Proteine zu **Orf14** werden als 'putatives Protein der inneren Membran' angegeben. Konservierte Domänen finden sich analog der so genannten Protein-Familie DUF1460, einer Protein-Gruppe mit weiter Verbreitung unter verschiedenen Mikroorganismen, jedoch ohne bekannte Funktion.

Das Polypeptid **Orf15** hat eine Masse von 29,9 kDa und weist Homologien zu einer Gruppe von Transporterproteinen für divalente Metallionen auf, darunter Zink-, Eisen- und Mangan-Ionen. Die größten strukturellen Ähnlichkeiten (konservierte Bereiche) ergeben sich zu ZnuB, der transmembranen Domäne oder so genannten Permease eines ABC-Transporters für Mangan- und Zink-Ionen. Zweiwertige Metallionen sind bekannte und wichtige Kofaktoren in zentralen zellulären Stoffwechselprozessen und werden über strukturähnliche Transportsysteme befördert.

Die höchsten Homologien zu **Orf15** mit 192 von 267 identischen Aminosäuren ergeben sich zu einem SitD-orthologen Protein von *Enterobacter sp.* 638 kodiert von dem Gen Ent638_1114. SitD ist eines der beiden Membranproteine des SitABCD-Transporters für Mangan- und Eisen-Ionen (Mn^{2+} / Fe^{2+}), der bei unterschiedlichen Spezies verbreitet ist und strukturelle Gemeinsamkeiten mit ZnuB, der Permease des ZnuABC-Transporter-Systems besitzt^{33, 83, 180}. Die in ABC-Transportern meist doppelt vorkommenden Permeasen bestehen jeweils aus α -Helices, die die Membran mehrfach durchqueren und so sechs bis elf Membran-durchspannende Bereiche bilden^{68, 96}. Diese bewahren die Substrat-Spezifität, sind aber dennoch weniger gut konserviert als die Nukleotid-bindende Domäne der ABC-Transporter.

Der Anteil von über 50% hydrophoben Aminosäuren (149 aa) bei **Orf15** spricht für die Funktion als Membranprotein.

Orf16 ist ebenfalls ein hydrophobes Polypeptid. 229 der 280 Aminosäuren sind identisch mit einem transmembranen ABC-Transporter-Protein, SitC, in *Enterobacter sp.* und *Shigella flexneri 2a*. Ähnlich hohe Homologien bestehen zu Anteilen eines ABC-Transporters für Mangan- und Zink-Ionen bei Shigellen, zum SitC Protein aus *Escherichia coli*, *Klebsiella pneumoniae*, *Agrobacterium tumefaciens*, *Sinorhizobium meliloti* und anderen, sowie zu nicht weiter klassifizierten Eisentransportproteinen und ABC-Transporterproteinen bei verschiedenen Mikroorganismen.

Wie bereits bei **Orf15** handelt es sich auch bei **Orf16** um ein Protein der inneren Membran. Die konservierten Bereiche von **Orf16** stimmen wie bei **Orf15** am meisten mit ZnuB überein, der Permease des ABC-Transporters, welcher vorwiegend Mangan- und Zink-Ionen befördert.

Während **Orf15** und **Orf16** hohe Homologien zu den Permeasen eines ABC-Transporters haben, entspricht **Orf17** der ATP-bindenden Komponente, dem Energielieferanten des Transporters. Das Protein besteht aus gut konservierten Bereichen, mit den typischen Motiven einer ATP-bindende Domäne von ABC-Transportern^{86, 96} (Abbildung 23). Im Gegensatz zu dem N-terminal nicht vollständigen Polypeptid **Orf0** ist hier das Walker A-Motiv erkennbar (siehe auch Abschnitt 3.2.1.2.).

Die höchsten Homologien von **Orf17** auf Proteinebene bestehen zu SitB, der ATP-bindenden Komponente des SitABCD-Transporters bei Shigellen und bei *Escherichia coli* (jeweils 73% Identität).

Der letzte ORF im DNA-Bereich D kodiert das Protein **Orf18**, das mit knapp über 300 Aminosäuren etwas größer ist als die anderen Polypeptide. Es bestehen deutliche Ähnlichkeiten zu dem periplasmatischen Protein von ABC-Transportern. Die meiste Übereinstimmung, 235 identische Aminosäuren (79%), ergibt sich zum SitA-orthologen Protein von *Enterobacter sp.*, Genlocus Ent638_1111. Der konservierte Anteil von **Orf18** ist ebenfalls in dem Metall-bindenden Protein PsaA nachweisbar. Dieses Protein fungiert als initialer Rezeptor beim Transport von Mangan-Ionen via ABC-Transporter, aber auch als Oberflächen-Adhäsion in einigen Eubakterien. PsaA gehört zu der TroA-Superfamilie. Diese bestehen typischerweise aus zwei kugelförmigen Subdomänen, die über eine Einzel-Helix verbunden sind und dazwischen das Metall-Ion binden können. Die Proteine dieser Familie zeigen

überdies als Gemeinsamkeit eine 'low complexity region' mit einem metall-bindenden Histidin-reichen Motiv, die so genannte repetitive *HDH-Sequenz*³⁸. Auch ZnuA, der Aufnahmerezeptor für Zink-Ionen, gehört zu dieser Protein-Familie und zeigt daher ebenso Ähnlichkeiten zu **Orf18**.

			Walker A		Gin-Loop	
Orf17	8	[1].VAGLTVV.[7].LQNASFSVPKG.[1].NAALI	GVNGSGKS	TLFRAMMGF.[5].GTISL.[13].IAY	VPO	85
HI0361	9	[1].VNDVTVR.[7].IHNMTFSLNSG.[1].ICALV	GVNGSGKS	TLFKSIMGL.[5].GEIKL.[13].VAY	VPO	86
SitB	8	[1].VNDVTVT.[7].LRDASFTVPGG.[1].IAALV	GVNGSGKS	TLFKAIMGF.[5].GKISV.[13].VAY	VPO	85
MntA	11	[1].VDGVSVT.[7].LYNATCTVEPG.[1].ITALV	GPNGSGKS	TLFKSIMGF.[5].GRVRI.[13].MAY	VPO	88
TW686	8	[1].VSDLVSV.[6].MRDATFEVAG.[1].VCGIT	GMNGSGKS	SLFNSTVGN.[5].GSIST.[13].VAY	VPO	84
Chut02000344	11	[1].IHDLTVS.[6].LWNIDLTIKPG.[1].LAGIT	GPNGAGKS	TLIKTAMGL.[5].GYVKL.[11].VSY	VPO	85
Rrub02000534	16	[1].VTGASVT.[7].LRDASFEIPTG.[1].VTGLV	GINGSGKS	TLFKAIMGF.[5].GSIAI.[13].VAY	VPO	93
MntA/ZnuC	32	[1].VDDLTVV.[6].IWDVELEIPAG.[1].LVGIV	GPNGAGKS	TLKKAAMD.[5].GRVEV.[11].VGY	VPO	106
TroB	12	[1].VDDLTLA.[6].LWDVDVRIPEG.[1].IEAIT	GPNGAGKS	TLKKAIMGF.[5].GEVRV.[11].VAY	VPO	86

			Signature		Walker B	
Orf17	86	[7].FPVLVEDVVMGR.[14].RRIVNDALERVDM	LGGGQR	KRVFLARSIAQRG	KVILLDE	PFTGVDV 175
HI0361	87	[7].FPVSVYDVMGR.[14].KQKVQEQAMQ	LGGGQR	KRVFLARALAAQS	PIILLDE	PFTGVDV 176
SitB	86	[7].FPVLVEDVVMGR.[14].RQIVTDALERV	LGGGQR	KRVFLARALAAQS	OVILLDE	PFTGVDV 175
MntA	89	[7].FPVSVFDVVMGR.[14].RRLVMSLE	LGGGQR	KRAFLARALAAQS	KVILLDE	PFTGVDV 178
TW686	85	[7].FPISVREVVMGR.[14].RDKVTDAL	LGGGQR	KRAFLARALAAQS	KLLLLDE	PFAGVDK 174
Chut02000344	86	[7].FPASALDVVMGR.[14].KEVATECLR	LGGGQR	KRVFLARALAAQS	DLYFMDE	PFAGVDA 175
Rrub02000534	94	[7].FPVLVEDVVMGR.[14].KSAVDQAL	LGGGQR	KRVFLARALAAQS	RVILLDE	PFTGVDV 183
MntA/ZnuC	107	[7].FPVDALDVVTMGL.[14].RDLAREAL	LGGGQR	KRVFLARALVQNA	DLYLMDE	PLAAVDA 196
TroB	87	[7].FPPTVFDVVMGR.[14].KARAREAI	LGGGQR	KRVFLARALVQNA	DLYFMDE	PFQGVDA 176

			His-Loop	
Orf17	176	NTEKKIITLLGEL.[3].GSTI	VVSTH	NLTSVSEY.[1].DYTVLVKGTVLASG 223
HI0361	177	KTENAIVDLLQQL.[3].GHLI	LVSTH	NLGSVPDF.[1].DQVVMINRTVIAAG 224
SitB	176	KTEAKIISLLREL.[3].GKTM	LVSTH	NLGSVTF.[1].DYTMVKGTVLASG 223
MntA	179	KTEKGMIDLLMEL.[3].GHTI	LISTH	DLASISTF.[1].DHTILLNRTILAQG 226
TW686	175	KSEGMIVQLIRQL.[3].GCCV	LVSTH	DLNLTLHGL.[1].NEAILVKNRILMHN 222
Chut02000344	176	STEKTIIALLKEM.[3].GKTI	IVVHH	DLNSADEY.[1].DWILLNTRLIASG 223
Rrub02000534	184	TTESEIITLLRAL.[3].GRVM	LVSTH	DLGSVPEF.[1].DRAVLLNRTVLASG 231
MntA/ZnuC	197	ATEAAIIDLLRQM.[3].GKTA	VVIHH	DLQSVDPY.[1].DYVLLNMRVVASG 244
TroB	177	ATEQAIVTLLKTL.[3].GKTI	LVVHH	DLQTVAEY.[1].DRVLLNVRVIAEG 224

HI0361 (*Haemophilus influenzae*), SitB (*E. coli*), MntA (*Synechocystis sp. PCC 6803*), TW686 (*Tropheryma whipplei* TW08/27), Chut02000344=ZnuC (*Cytophaga hutchinsonii*), Rrub02000534=ZnuC (*Rhodospirillum rubrum*), MntA/ZnuC (*Rhodopirellula baltica* SH 1), TroB (*Treponema pallidum*)

Abbildung 23: Konservierte Motive von **Orf17**, der ATP-bindenden Komponente eines ABC-Transporters. *Walker A-* und *Walker B-* Motiv, *Histidin-Loop*, *Gin-Loop* und *Signature-Motiv* sind eingezeichnet.

Zusammenfassend kodiert der DNA-Bereich D einen weiteren ABC-Transporter für Metallionen. Er besteht aus den putativen Proteinen **Orf15**, **Orf16**, **Orf17** und **Orf18** und hat die höchsten Homologien zu ABC-Transportern für Zink, Mangan und Eisen unterschiedlicher Mikroorganismen (Homologien zwischen 73% bis 81% identische Aminosäuren). Lediglich **Orf14** hat hohe Homologien zu einem als Membran-Protein klassifizierten Protein von *Klebsiella pneumoniae*.

3.2.4.3. Homologienvergleich DNA-Abschnitt D mit dem Plasmid pLVPK von *Klebsiella pneumoniae*

Im Gegensatz zu den DNA-Abschnitten A, B und C, die hohe Homologien zu pLVPK von *Klebsiella pneumoniae* auf DNA- und Proteinebene zeigen, ist zwischen der Position bp 14946 und bp 18640 ein kompletter Abbruch der Sequenzähnlichkeiten nachweisbar. Dies betrifft das Gen-Cluster mit hohen Homologien zu ABC-Transportern, z.B. zum Znu-ABC- und zum SitABCD-Transporter.

Genauere Sequenzvergleiche mit dem Plasmid pLVPK zeigen die Homologien bis kurz vor den Beginn des ORFs14, respektive des ORFs LV186 auf dem Plasmid pLVPK (Abbildung 24). Im Bereich der ORFs 15, 16, 17 und 18 befindet sich auf dem Plasmid von *K. pneumoniae* nur der kleine ORF LV185. Auch die Sequenzvergleiche um ORF19, der dem DNA-Bereich E zugeordnet wird, zeigen einen Abschnitt auf, der bei *Klebsiella* fehlt.

E. coli ECOR 31

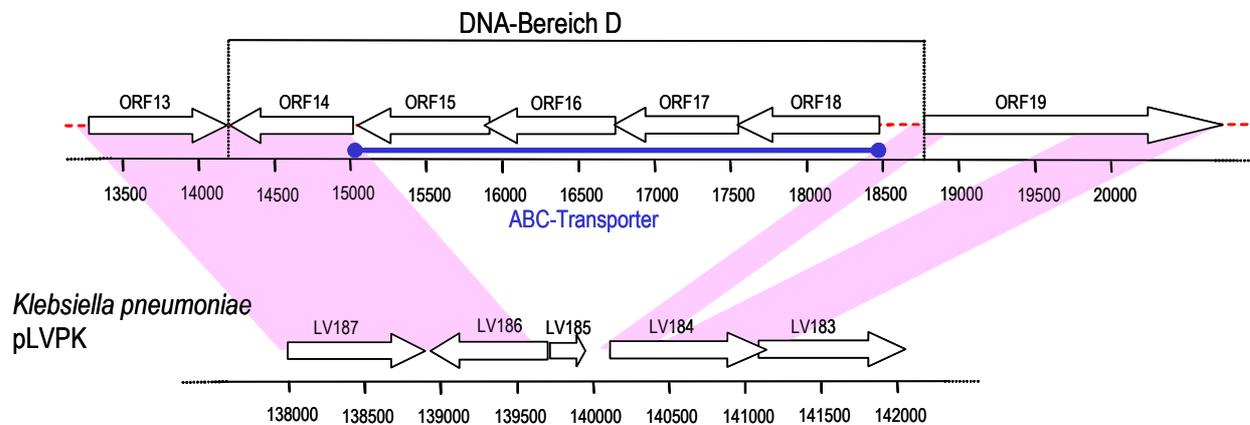


Abbildung 24: Sequenzvergleiche auf DNA-Ebene zwischen dem DNA-Bereich D von RegX und der Position bp 138000 - bp 142000 auf dem Plasmid pLVPK von *Klebsiella pneumoniae*. Rosa markiert sind DNA-Abschnitte hoher Homologien (>86% Sequenzidentität). Der Genbereich der für den *sitABCD*-ähnlichen ABC-Transporter kodiert, fehlt bei *Klebsiella pneumoniae*. Ebenso ergeben sich Homologie-Lücken im Bereich ORF19.

3.2.4.4. Transkriptionsuntersuchungen ABC-Transporter (RT-PCR)

Die Ergebnisse aus dem DNA-Bereich D lassen vermuten, dass das Gencluster ORF15-18 für einen ABC-Transporter kodiert. Um die Expression zu überprüfen, wurde analog zum Gencluster ORF0-2 eine RT-PCR durchgeführt (siehe auch Unterpunkt 3.2.1.4.).

Als Ausgangsmaterial der Untersuchung diene erneut die RNA von ECOR31, die mit einer reversen Transkriptase in cDNA umgeschrieben wird. Über ausgewählte Primer aus dem Bereich des Genclusters kann das Vorhandensein von RNA als erster Schritt für die Expression überprüft werden. Wie im Abschnitt 3.2.1.4., sowie im Material- und Methoden-Abschnitt beschrieben, wurden auch hier Negativkontrollen ohne Reverse-Transkriptase-Zusatz durchgeführt.

Bei der RT-PCR aus dem Bereich der ORFs 14 bis 18 wurden die Primerpaare so ausgewählt, dass das Produkt zwei ORFs überlappt (Abbildung 25). Auf diese Weise lässt sich eine polycistronische mRNA darstellen und eine Operonstruktur des Genclusters nachweisen.

Die RT-PCR mit den Primerpaaren 1 bis 4 (Abbildung 25) zeigt eine Transkription des gesamten ABC-Transporter-Clusters mit den einzelnen ORFs 18, 17, 16 und 15. Verwendet wurden die Primer sitA.for, sitB.for, sitB.rev, sitC.for, sitC.rev, sitD.for, sitD.rev, Membr.Prot.rev. Allerdings findet sich kein positives Signal mit einem Primerpaar aus dem Bereich ORF14/ORF15. Dies spricht für einen Abbruch bei der mRNA-Synthese nach ORF15. ORF14 zählt nicht zu dem gleichen mRNA-Transkript wie das restliche Cluster.

Wie bereits die Vergleiche mit der Genbank ergeben haben, ist ORF14 damit wahrscheinlich kein Teil des putativen ABC-Transporter-Operons, das aus den ORFs 18, 17, 16 und 15 besteht.

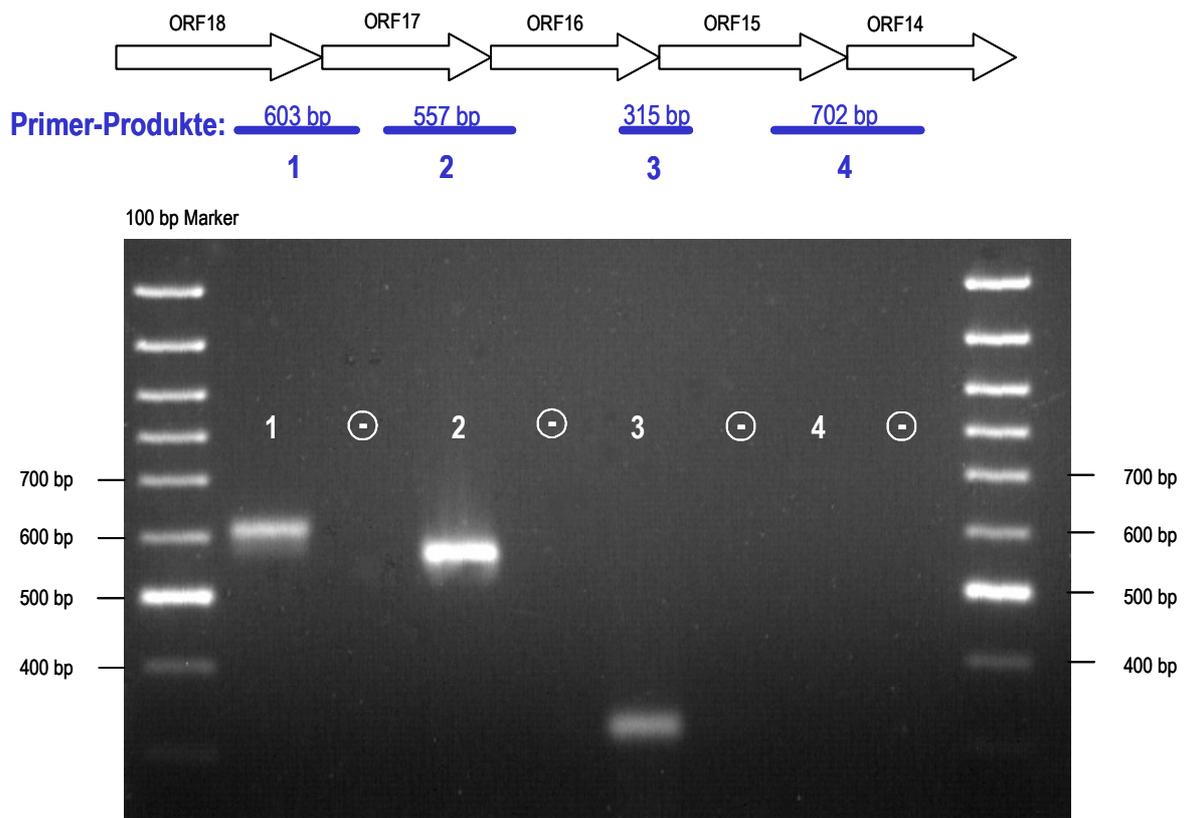


Abbildung 25: Reverse-Transkriptase-PCR (Gelelektrophorese). Die positiven PCR-Ergebnisse mit den Primerpaaren aus den Bereichen ORF18/17 (1), ORF17/16 (2) und ORF16/15 (3) bestätigen eine Operonstruktur des putativen ABC-Transporters ORF15-18.

3.2.5. DNA-Bereich E

3.2.5.1. Sequenzanalyse DNA-Bereich E

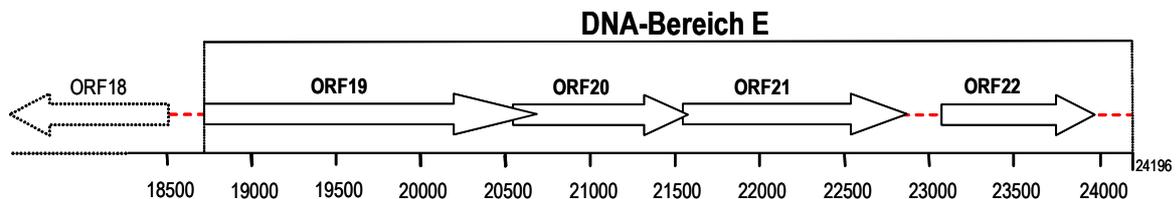


Abbildung 26: Übersicht DNA-Bereich E mit den ORFs 19 – 22; Position bp 18719 – bp 24196.

Der DNA-Bereich beschreibt die restlichen 5478 Basenpaare bis zum Ende der von uns definierten Region X (Position bp 18719 bis bp 24196). Alle ORFs sind dem vorherigen ABC-Transporter-Cluster (ORF15-18) entgegengerichtet und besitzen ein Startcodon ('ATG'). Eine Shine-Dalgarno-Sequenz lässt sich vor allen ORFs, außer vor ORF20 nachweisen.

252 bp vor ORF18 beginnt das offene Leseraster 19 mit einer Länge von 1959 bp. Es folgen ORF20 (975 bp), ORF21 (1347 bp) und nach 225 bp nicht-kodierender Region beginnt ORF22 an Position bp 23090 (858 bp). Die Position 24196 definiert das Ende von RegX im Cosmid p17-A7, die folgenden Sequenzen entsprechen dem Klonierungs-Vektor SuperCos1 (Stratagene).

3.2.5.2. Homologienvergleiche DNA-Bereich E auf DNA- und Proteinebene

Die Homologien auf DNA-Ebene für den DNA-Bereich E sind nicht durchgängig zu *Klebsiella pneumoniae* Plasmid pLVPK. Von Position bp 18719 bis bp 23964 sind die beiden Sequenzen zu 85% identisch. Eine Homologie-Lücke von etwa 1000 Basenpaaren besteht jedoch im Bereich von ORF19 und am Ende von RegX (Abbildung 29). In diesem Bereich gibt es lediglich kleinere homologe Abschnitte zu *E. coli*-DNA. Der direkte Übergang in den Klonierungs-Vektor definiert das Ende des Cosmids p17-A7 nach 24196 Basenpaaren.

Das Protein **Orf19** hat ausschließlich Ähnlichkeiten zu Threonyl-tRNA-Synthetasen anderer Mikroorganismen, vorwiegend von Proteobakterien, wobei sich mit 61% Aminosäuren-Identität die höchsten Homologien zu einem *Burkholderia pseudomallei* Isolat ergeben. Interessanterweise kodiert bereits ORF7 eine tRNA-Synthetase.

Die Threonyl-tRNA-Synthetase ist ein Homodimer und verantwortlich für die Anheftung der Aminosäure Threonin an die 3'-OH-Gruppe der Ribose der entsprechenden tRNA (Aminoacyl-tRNA-Synthetase der Klasse II). Die in der Mitte des Proteins liegende, konservierte katalytische Domäne (*Core*), sorgt für die ATP-abhängige Bildung des enzymgebundenen Aminoacyladenylats. An weiteren konservierten Motiven finden sich das N-terminale *TGS-Motiv*, das außer bei der Thr-tRNA-Synthetase auch bei GTPasen und beim Ribosom-assoziierten Stress-Antwort-Protein SpoT vorkommt¹⁷⁸. Die genaue Funktion der ca. 50 Aminosäuren großen Domäne ist nicht bekannt, man geht jedoch von einer regulatorischen Funktion bei der Bindung des Liganden aus¹⁷⁸. Am C-terminalen Ende von **Orf19** befindet sich die Domäne des Anticodon-Bindungsbereichs. Dieser macht die Spezifität aus und sorgt für den Transfer des aktivierten Aminoacyl-Restes an die spezifische tRNA. Erstaunlicherweise fehlt das N-terminale *HxxxH-Motiv* bei **Orf19** (Abbildung 27), das vermutlich hydrolytische Aktivität vermittelt.

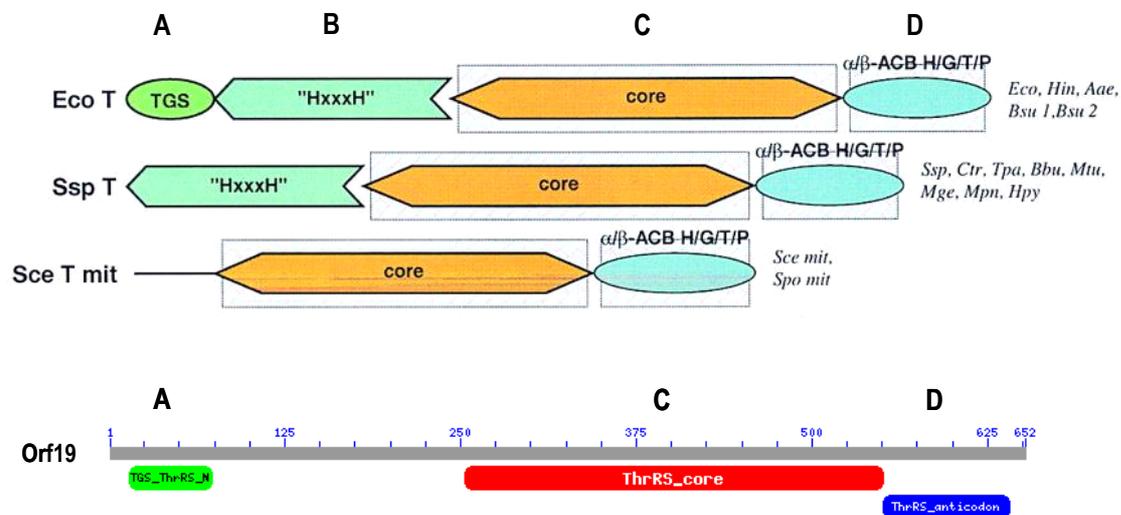


Abbildung 27: Domänen der Threonyl-tRNA-Synthetase verschiedener Spezies, nach Wolf Y.I. ¹⁷⁸, mit TGS- (A), HxxxH- (B), Core- (C) und Anticodon Motiv (D). Im unteren Bereich des Bildes sind die konservierten Bereiche von **Orf19** dargestellt. Hier fehlt das HxxxH-Motiv.

(Eco: *E. coli*, Ssp: *Synechocystis sp.*, Sce: *S. cerevisiae*)

Die höchsten Übereinstimmungen für **Orf20** ergeben sich zum hypothetischen Protein LV183 von *Klebsiella pneumoniae*, wobei 245 von 291 Aminosäuren identisch sind. Die Homologien zu verschiedenen putativen Proteinen unterschiedlicher Spezies der, marine Bakterien umfassenden, Gattung *Shewanella* sind nur halb so hoch. Für die entsprechenden orthologen Proteine sind keine Funktionen beschrieben. Über einen kurzen Peptid-Abschnitt finden sich 60-70 konservierte Aminosäuren mit Ähnlichkeiten zu Zink-Carboxypeptidasen. Die Homologien, respektive der konservierte Bereich, erstrecken sich jedoch nicht über das gesamte Protein.

Für **Orf21** ergeben sich die höchsten Homologien ebenfalls zu *Klebsiella pneumoniae*. 94% der 448 Aminosäuren sind identisch zu einem als PyrC beschriebenen Protein. PyrC ist eine Dihydroorotase, die in den Nukleotid-Transport und -Metabolismus involviert ist. Sie gehört, wie die Hydantoinase, Dihydropyrimidinase und die Allantoinase, mit denen **Orf21** konservierte Domänen gemeinsam hat, zu den zyklischen Amidohydrolasen ⁸⁷. Diese sind am Metabolismus

von Pyrimidinen und Purinen beteiligt. Die Dihydroorotase beispielsweise katalysiert die reversible Umwandlung von Carbamoylaspartat zu Dihydroorotat, eine Schlüsselreaktion in der Pyrimidin-Biosynthese. Hingegen ist die Allantoinase beim Abbau der Purine beteiligt und die Dihydropyrimidinase und Hydantoinase spielen eine wichtige Rolle beim Abbau der Pyrimidine.

Orf22 ist weitgehend identisch mit dem Protein auf dem Plasmid pLVPK von *Klebsiella pneumoniae*, das als mögliche Porphobilinogen-Synthase gilt (266 von 285 Aminosäuren identisch). **Orf22** gehört zu den konservierten Proteinen für die man 62-70% Aminosäurenidentität zu Porphobilinogen-Synthasen oder auch Delta-Aminolävulinsäure-Dehydratase verschiedener Bakterien erhält.

Die Porphobilinogen-Synthase (PBGS) oder auch Delta-Aminolävulinsäure-Dehydratase (ALAD) katalysiert die erste gemeinsame Reaktion in der Biosynthese der Tetrapyrrol-Pigmente wie Häm, Chlorophyll und Cobalamin¹³. Der gemeinsame Vorläufer Porphobilinogen wird dabei mit Hilfe der ALAD aus zwei Molekülen Aminolävulinsäure synthetisiert. Die Porphobilinogen-Synthase oder ALAD ist hoch konserviert, zeigt jedoch eine phylogenetische Variation bei der Metall-Auswahl des katalytischen Zentrums. Etwa die Hälfte der bisher sequenzierten PBGS-Gene besitzt Bindungsstellen für Zink¹³.

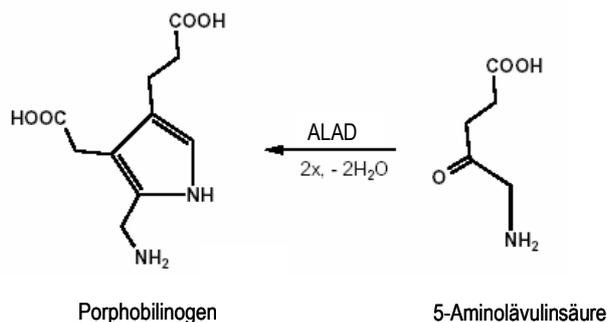


Abbildung 28: Synthese von Porphobilinogen als Vorläufer der Tetrapyrrole durch das Enzym Delta-Aminolävulinsäure-Dehydratase (ALAD).

Zusammenfassend kann festgestellt werden, dass die offenen Leseraster des DNA-Bereichs E für eine putative Threonyl-tRNA-Synthetase, Dihydroorotase und Delta-Aminolävulinsäure-Dehydrogenase oder Porphobilinogen-Synthase kodieren. Für die orthologen Proteine von **Orf20** sind keine Funktionen beschrieben.

3.2.5.3. Homologievergleiche DNA-Abschnitt E mit dem Plasmid pLVPK von *Klebsiella pneumoniae*

Zu dem Plasmid pLVPK von *Klebsiella pneumoniae* bricht die Homologie direkt nach ORF22 ab. Die letzten Basenpaare bis zum Ende von RegX im Cosmid p17-A7 haben keine Homologien zum Plasmid pLVPK von *Klebsiella pneumoniae* (Abbildung 29).

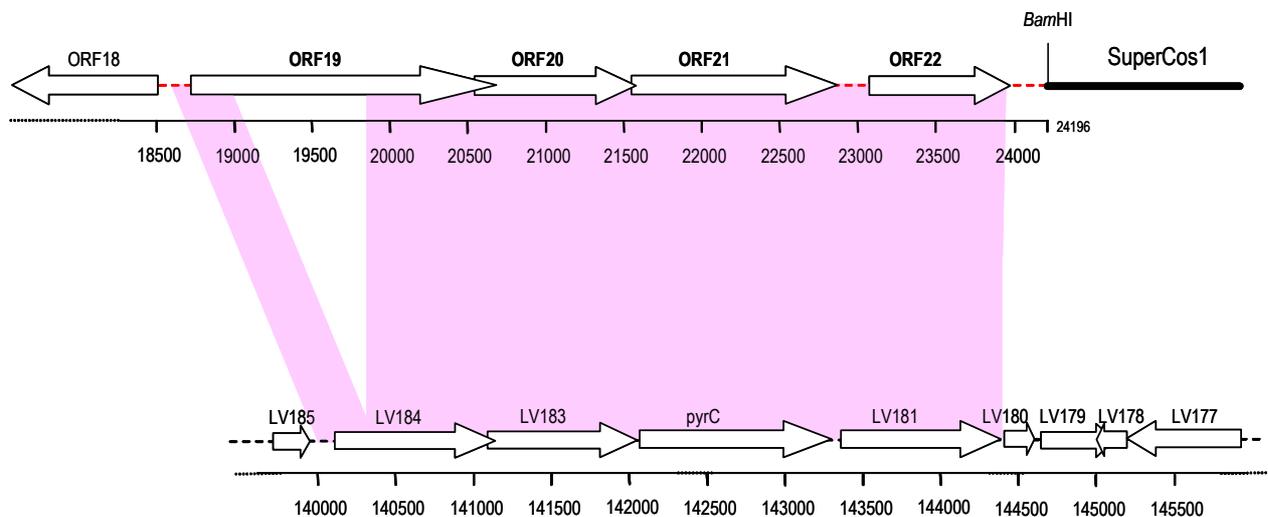


Abbildung 29: Vergleich DNA-Abschnitt E mit dem Plasmid pLVPK von *Klebsiella pneumoniae*.

Ab der Position bp 24196 ist die Sequenz 100% identisch zu dem Klonierungs-Vektor SuperCos1. Die letzten Nukleotide von RegX, respektive von ORF24 sind die Schnittstelle des Restriktionsenzym *Bam*HI ('GGATC').

3.3. Untersuchung verschiedener Stämme auf das Vorhandensein von RegX

Um der Frage nach der Herkunft und der Verbreitung dieser unbekannt Region nach zu gehen, wurden verschiedene Stämme mittels PCR nach dem Vorkommen von RegX untersucht. Die verwendeten Bakterienstämme wurden zum überwiegenden Teil aus der Routine-Diagnostik des Max von Pettenkofer-Instituts bezogen. Aus einer Einzelkolonie wurden entsprechend der Anleitung im Methoden-Teil Zellysate (gekochte Zellen) angefertigt. Mit Hilfe einer Polymerasekettenreaktion (PCR) mit den Primern RegX-521.for und RegX-358.rev aus dem DNA-Bereich A wurden die Bakterien gescreent.

Zur Untersuchung der Spezies *Yersinia* wurde chromosomale DNA verwendet.

3.3.1. Anzahl und Art der untersuchten Bakterien-Stämme

Insgesamt wurden 530 Bakterien-Stämme auf das Vorhandensein der Region X untersucht, davon 437 Mitglieder der Familie *Enterobacteriaceae* (Abbildung 30) und 93 Mitglieder der Familie *Pseudomonaceae* (siehe Abschnitt 2.1.4.).

Auf Grund der erstmaligen Charakterisierung von RegX in einem *E. coli* Stamm (ECOR31) wurde das Vorkommen dieser Region zunächst innerhalb dieser Spezies untersucht. Unter 111 *E. coli*-Isolaten, vorwiegend aus Blut- und Urin-Kulturen, sowie der ECOR-Stammsammlung mit 72 Stämmen, fanden sich bis auf ECOR31 keine RegX-positiven Stämme.

Die Suche wurde auf weitere Mitglieder der Familie *Enterobacteriaceae* (254 Stämme) ausgeweitet und nach früheren Homologien-Ergebnissen der ORFs 0 - 4 auch auf die Familie der *Pseudomonaceae*, von denen zusammen 93 Stämme untersucht wurden. Die DNA-Proben der Spezies *Yersinia* wurden freundlicherweise von Dr. A. Rakin (Max von Pettenkofer-Institut) in Form chromosomaler DNA zur Verfügung gestellt. Sechs pflanzenpathogene *Serratia*- und 12 *Pseudomonas*-Stämme wurden freundlicherweise von Dr. G. Berg (Institut für molekulare

Physiologie und Biotechnologie, Abteilung Mikrobiologie, der Universität Rostock), für die Untersuchung bereitgestellt.

Mit Hilfe der Screening-PCR konnte nur bei einem Stamm die Region X nachgewiesen werden. Dabei handelt es sich um die Spezies *Klebsiella pneumoniae*, dem Isolat 2495.

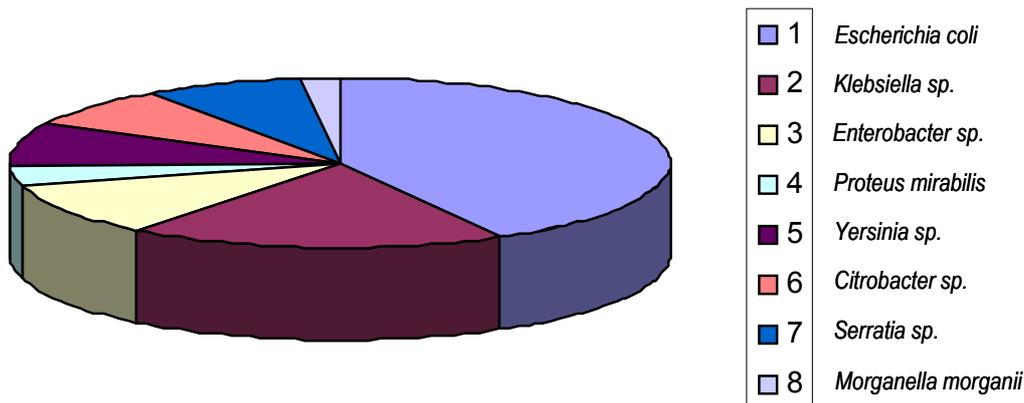


Abbildung 30: Untersuchte *Enterobacteriaceae* (gesamt 437).

3.3.2. RegX-positiver Stamm: *Klebsiella pneumoniae* Isolat 2495

Der RegX-positive Stamm *Klebsiella pneumoniae*, Isolat 2495, aus der Routine-Diagnostik des Max von Pettenkofer-Instituts für Mikrobiologie und Hygiene in München, ist ein Blut-Kultur-Isolat eines Sepsis-Patienten. Die Zuordnung zur Spezies *Klebsiella pneumoniae* wurde durch einen biochemischen Differenzierungstest (API, Firma bioMérieux) bestätigt.

Mittels PCR-Ergebnissen konnte der gesamte RegX-Bereich auf *K. pneumoniae* lokalisiert werden. Verwendet wurden Primerpaare aus den DNA-Bereichen A, B, C, D und E, sowie aus dem Bereich des 17 bp *direct repeat* (DR) und der Pathogenitätsinsel *Yersinia*-HPI (Abbildung 31). Sowohl die *Yersinia*-HPI, als auch die atypische Insertionsstelle, siehe Unterpunkt 3.1., sind bei *Klebsiella pneumoniae* 2495 nachweisbar. Interessanterweise ist der RegX-Bereich von *Klebsiella*

pneumoniae 2495 dem Abschnitt auf ECOR31 ähnlicher als dem Bereich auf dem Plasmid pLVPK von *Klebsiella pneumoniae* CG43³⁰.

Im Gegensatz zur der Region auf dem *Klebsiella*-Plasmid pLVPK wurden bei *Klebsiella pneumoniae* 2495 mit den PCR-Untersuchungen sowohl die *Yersinia*-HPI, als auch die atypische Insertionsstelle der HPI mit dem *direct repeat* (DR) und beispielsweise der ABC-Transporter im Bereich DNA-Bereich D bestätigt. Diese fehlen auf dem Plasmid pLVPK.

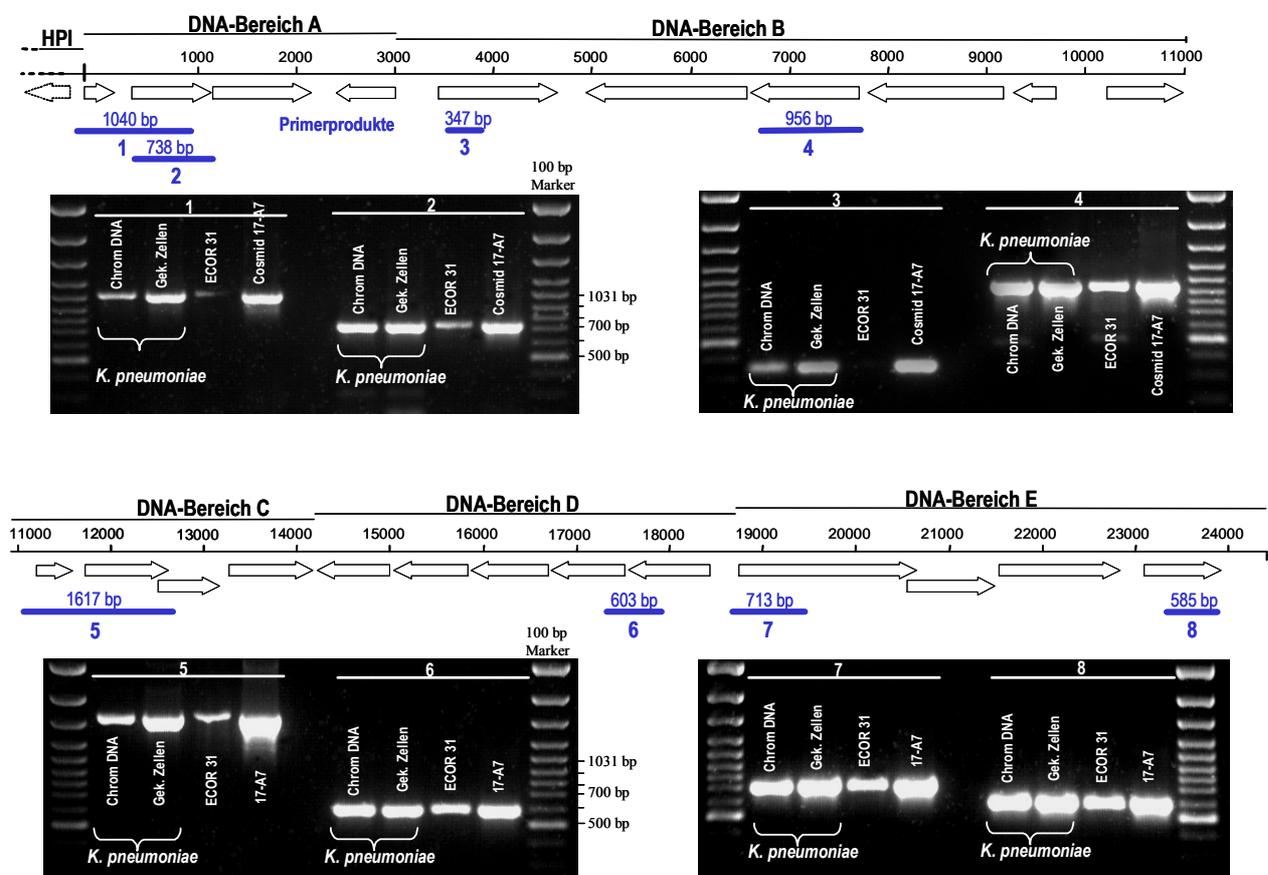


Abbildung 31: Ergebnisse der Gel-Elektrophorese aus den PCRs des isolierten RegX-positiven Stammes *Klebsiella pneumoniae* 2495. Der Stamm enthält die gesamte Region X, sowie die *Yersinia*-HPI mit der atypischen Insertionsstelle, wie in ECOR31.

Die Region RegX des *E. coli* Stammes ECOR31 ist auch bei anderen Spezies verbreitet. In der vorliegenden Arbeit gelang der Nachweis dieser Region bei einem von 59 untersuchten *Klebsiella pneumoniae* Stämmen. Das Vorliegen der gesamten

Region wurde mit PCR-Untersuchungen verschiedener Abschnitte innerhalb der Region X gezeigt. Hierbei gelang auch der Nachweis von Bereichen, die auf dem RegX-ähnlichen Abschnitt des vollständig sequenzierten Plasmids pLVPK von *Klebsiella pneumoniae* CG43 nicht vorhanden sind, wie z.B. die *Yersinia*-HPI und der Sit- oder Znu-ähnliche ABC-Transporter (ORF15-18). Die Verbreitung von RegX unter verschiedenen Spezies wie *Escherichia coli* und *Klebsiella pneumoniae*, die mosaikartige Struktur und die Variabilität bei verschiedenen Isolaten sind deutliche Indizien für einen horizontalen Gentransfer. Die Sequenzen der einzelnen PCR-Produkte weisen dabei auf eine hohe Sequenzhomologie zwischen *Klebsiella pneumoniae* 2495 und ECOR31 und sprechen somit für ein zeitlich kurz zurückliegendes Gentransfer-Ereignis.

4. Diskussion

Bakterielle Vielfalt

Der menschliche Körper wird von etwa einer Billiarde (10^{15}) Mikroorganismen kolonisiert, mehr als die Zahl menschlicher Zellen eines Körpers ($10^{13} - 10^{14}$)¹⁶⁵.

Ein Großteil dieser Mikroorganismen sind Bakterien, die als Teil der Haut- und Schleimhaut-Flora für den Körper nützlich sind⁵⁶. Die Darmflora beispielsweise ist unter anderem an den metabolischen Prozessen bei der Nährstoffverwertung beteiligt, schützt den Wirt vor dem Ausbreiten anderer Mikroorganismen und spielt eine Rolle bei der Entwicklung des Immun-Systems^{10, 85, 175, 177}.

Ist ein Bakterium mit Eigenschaften ausgestattet, die ihm die Fähigkeit verleihen eine Krankheit hervorzurufen, spricht man in der medizinischen Mikrobiologie und Infektiologie von pathogenen Bakterien. Diese stellen jedoch nur einen kleinen Teil der gesamten Mikroorganismen dar. Pathogenität kann aber auch als Selektionsvorteil angesehen werden¹⁴⁴, denn pathogene Keime unterscheiden sich von den nicht-pathogenen Vertretern der gleichen Spezies durch Mechanismen den menschlichen Körper zu kolonisieren und sich dort zu vermehren, sowie den Abwehrmechanismen des Immunsystems zu entkommen^{53, 116}. So verfügen beispielsweise Pathovaren von *Escherichia coli* wie ExPEC (extraintestinale pathogene E. coli) oder EHEC (enterohämorrhagische E. coli) im Gegensatz zum kommensalen *E. coli* über Virulenzfaktoren wie Entero- und Zytotoxine, Adhäsions- und Invasionsfaktoren, Siderophore und Polysaccharid-Kapseln^{73, 105, 119}. Diese Faktoren sind nicht für das unmittelbare Überleben der Bakterien nötig, doch Stämme mit diesen zusätzlichen Genen sind in der Lage eine neue Nische zu erobern, wie beispielsweise bei ExPEC den menschlichen Urogenitaltrakt^{41, 73}. Damit hat sich *E. coli*, zusätzlich zu seiner Bedeutung als Indikatorkeim für fäkale Verunreinigung von Wasser, in der Medizin einen Namen gemacht: *E. coli* ist einer der häufigsten Erreger bakterieller Infektionen des Menschen^{72, 76, 176}.

Bakterielle Vielfalt entsteht durch das sehr wandlungsfähige bakterielle Genom, das einem steten Umwandlungsprozess unterzogen ist ⁴³. Dabei erhält der einzelne Stamm die Möglichkeit, zusammen mit dem Genotypen in der Regel auch den Phänotypen zu ändern. Beispiele für diese Mechanismen sind die Variation der Oberflächenantigene durch Punktmutationen, wie bei den Typ 1-Fimbrien von *E. coli* ¹⁵⁷ oder den Pili von Neisserien ^{1, 99}. Durch Punktmutationen oder Deletionen entstehen feine Veränderungen bereits bestehender Strukturen, jedoch kommt es dadurch nur selten zum Erwerb ganz neuer Eigenschaften. Eine größere evolutive Bedeutung kommt dem horizontalen Gentransfer zu. Hier werden große DNA-Bereiche mit mehreren Genen oder Gengruppen (Cluster) gleichzeitig übertragen und können zu einer kompletten Änderung des Phänotyps führen.

Bedeutung des horizontalen Gentransfers

Horizontaler Gentransfer durch Transformation, sowie durch Bakteriophagen und Plasmide ist bereits seit längerem bekannt ^{6, 91, 110, 163, 182}. Virulenzplasmide und Plasmide mit Genen für Antibiotikaresistenzen spielen eine große Rolle bei Infektionserkrankungen und deren Therapie. Die Sequenzierung mehrerer bakterieller Genome führte zur Entdeckung großer chromosomaler DNA-Abschnitte, die auf Grund bestimmter Charakteristika als Genomische Inseln bezeichnet werden und zwischen Bakterien horizontal transferiert werden ⁶⁰ (siehe Abschnitt 1.3.).

Horizontaler Gentransfer Genomischer Inseln ermöglicht Bakterien eine schnelle Anpassung an wechselnde Umweltbedingungen ⁴⁴. Durch Übertragung größerer DNA-Pakete mit ganzen Gen-Clustern gibt er dem einzelnen Bakterien-Stamm die Gelegenheit eine 'Evolution in Quantensprüngen' zu vollziehen ⁵⁵. Entsprechend der Bedeutung, die die Genomischen Inseln für das Bakterium darstellen, unterscheidet man Fitness-Inseln, Symbiose-Inseln, Metabolische Inseln und Pathogenitäts-Inseln. Untersuchungen verschiedener Genomischer Inseln haben bereits einige Aufschlüsse über ihren Aufbau und die Verbreitung, sowie den dadurch entstandenen Phänotypen gegeben.

Die genauen Transfermechanismen bleiben weitgehend ungeklärt, ein wiederkehrendes Motiv ist jedoch die Insertion in bestimmte DNA-Sequenzen, vorwiegend Gene für Transfer-Ribonukleinsäure (tRNA). tRNA-Gene werden durch

ihre konservierte Struktur und das ubiquitäre Vorkommen als Insertions-*hotspots* gesehen¹⁴¹. Die Symbiose-Insel von *Mesorhizobium loti* beispielsweise inseriert in das Gen für Phenylalanin-tRNA (*phe*-tRNA)¹⁶² und die SHI-2 Insel von *Shigella flexneri*, die das Siderophor-Eisenaufnahmesystem Aerobactin kodiert, in die Selenocystein-tRNA (*selC*-tRNA)^{111, 171}. Im uropathogenen *E. coli* Stamm 536 (UPEC) sind alle vier Pathogenitätsinseln in tRNA-Gene integriert: PAI I₅₃₆ in das *selC*-tRNA Gen (Selenocystein-tRNA), PAI II₅₃₆ in *leuX* (Leucin-tRNA), PAI III₅₃₆ in *thrW* (Threonin-tRNA) und schließlich PAI IV₅₃₆ in das *asnT*-tRNA-Gen (Asparaginsäure-tRNA)⁴².

Die in verschiedenen Spezies strukturell und funktionell gut untersuchte Pathogenitätsinsel *Yersinia*-HPI ist in Gene für Asparaginsäure tRNA (*asn*-tRNA) inseriert^{21, 65, 149}. Sie kodiert die Biosynthese und den Transport des Siderophors Yersiniabactin und ist ein typisches Beispiel für eine Genomische Insel^{25, 28, 57, 150}. Die Insertion erfolgt über *site*-spezifische Rekombination. Im Genom einer Spezies existieren meist mehrere *asn*-tRNA-Gene. In *E. coli* beispielsweise gibt es vier genetisch nahezu identische Asparaginsäure-tRNA-Kopien, *asnT*, *asnW*, *asnU* und *asnV*. Während die HPI bei *Yersinia pseudotuberculosis* in allen *asn*-tRNA-Kopien angetroffen wird, ist bei *E. coli* nur eine von vier Asparaginsäure-tRNA-Genkopien bevorzugt, die *asnT*-tRNA^{21, 149}.

Von besonderer Bedeutung für das Verständnis des horizontalen Gentransfers und der Evolution ist daher die Untersuchung der Insertionsorte der HPI im Genom anderer Bakterien und vor allem die Bestimmung alternativer Insertionsorte der HPI. Im Vorfeld dieser Arbeit konnte ein *E. coli*-Stamm mit zwei atypischen Insertionsstellen der HPI isoliert werden. Der HPI-positive Stamm ECOR31 aus der *E. coli*-Referenzsammlung ECOR zeigte zusätzlich zu multiplen Antibiotikaresistenzen und dem Nachweis diverser Virulenzfaktoren^{15, 130} auch hinsichtlich der HPI Besonderheiten auf. Neben einer ungewöhnlichen Insertion in die *ansV*-tRNA-Kopie einer selbst-mobilisierbaren HPI-Kopie¹⁴⁷ existiert eine HPI-Kopie, deren Insertion kein tRNA-Gen ist. Der alternative Insertionsort ist vielmehr eine bislang unbekannte DNA-Region (RegX) (Abbildung 32).

In dieser Arbeit wurde der atypische Insertionsort der HPI in ECOR31 beschrieben, der erstmals keine *asn*-tRNA darstellt. Die angrenzende, bisher völlig unbekannte Region (RegX) wurde bezüglich ihrer Struktur und der Homologien auf DNA- und Protein-Ebene (Annotationen) charakterisiert. Über ein Screening vorwiegend anderer *Enterobacteriaceae* auf das Vorhandensein der neuen Region erhielten wir Aufschluss über das Vorkommen von RegX und der atypischen HPI-Insertion.

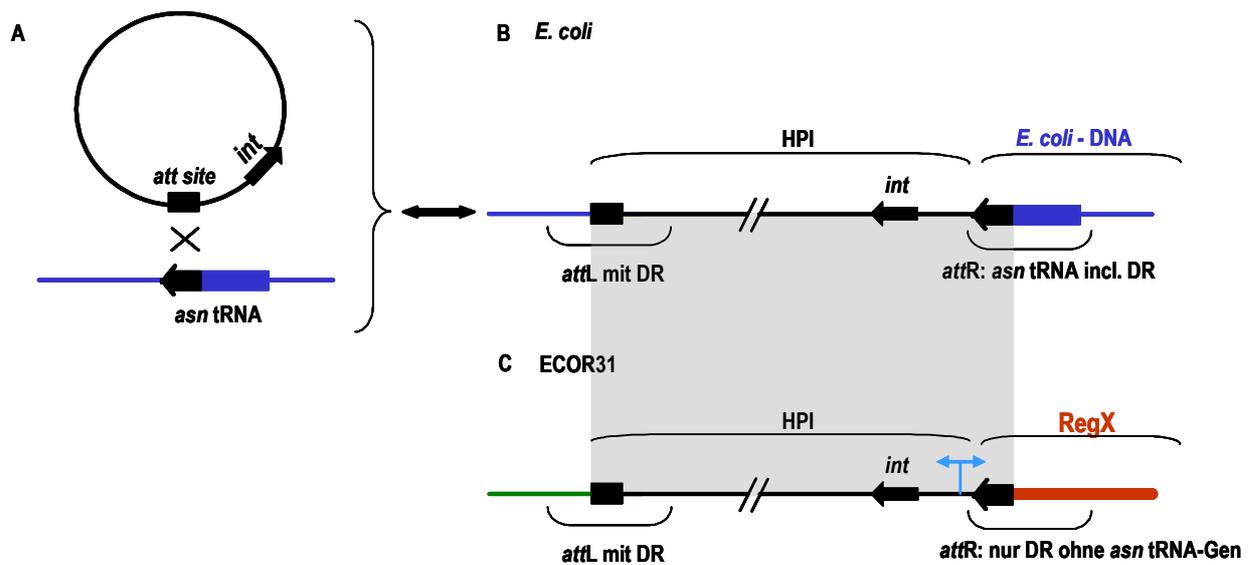


Abbildung 32: Homologe Rekombination bei der Integration der HPI über Erkennungssequenzen im Bereich des *asn*-tRNA-Gens (A). Bei der Integration in das Genom wird die HPI am 3'-Ende von einer partiellen *asn*-tRNA-Duplikation (DR = *direct repeat*) und am 5'-Ende vom gesamten *asn*-tRNA-Gen flankiert (B). Im *E. coli*-Stamm ECOR31 fehlt das *asn*-tRNA-Gen als typischer Insertionsort der HPI. Am 5'-Ende ist lediglich der typische DR nachweisbar (C). *att* = Erkennungs-Sequenz ('*attachment site*'); hellblau: bidirektionaler Promotor.

4.1. Die Region X (RegX), eine Clusterregion für den Metallionenstoffwechsel

Zur näheren Untersuchung des Insertionsortes der HPI im *E. coli* Stamm ECOR31 wurde zunächst mittels LA-PCR ein Fragment aus dem unbekanntem Bereich amplifiziert und mittels Transposon-Mutagenese sequenziert (siehe Unterpunkt 3.1.). Nachdem deutlich wurde, dass das gesamte LA-PCR-Fragment unbekannte DNA darstellt, wurde zur Sequenzierung eines größeren Bereiches eine Cosmid-Bank angefertigt. Neben der Isolation großer DNA-Fragmente besteht ein weiterer Vorteil der Cosmid-Bank darin, dass mit Original-DNA anstatt mit PCR-generierter DNA gearbeitet werden kann. Fehlerquellen wie Punktmutationen auf Grund verkehrten Nukleotid-Einbaus durch die Polymerase entfallen so. Zur Herstellung der Cosmid-Bank wurde die partial verdaute chromosomale DNA des Stammes ECOR31 in Cosmide kloniert. Übertragen in Phagenköpfen konnten die Cosmide in einen Labor-*E. coli*-Stamm eingebracht werden.

Nach dem Screening von insgesamt 20 Platten mit ca. 50 000 Klonen Cosmid-tragender *E. coli*, konnte das Cosmid p17A7 isoliert werden, das neben Teilen der *Yersinia*-HPI die DNA-Sequenz der angrenzenden Region X enthält. Das isolierte RegX Fragment besteht aus 24196 Basenpaaren, die für 22 offene Leseraster kodieren (ORF0-22). Zur besseren Beschreibung wurde die Region an Hand der Gen-Anordnung in 5 Abschnitte eingeteilt, die DNA-Bereiche A - E.

Durch strukturelle Analysen und DNA-Vergleiche konnten auf RegX mehrere Gene und Gen-Cluster identifiziert werden. Hierbei handelt es sich vorwiegend um Gene, die im Metabolismus eines Stammes von Bedeutung sind, aber auch Cluster, die als Virulenzfaktoren eine Rolle spielen können.

4.1.1. ABC-Transporter zur Aufnahme von Metallionen

Im DNA-Abschnitt A und D von RegX sind zwei ABC-Transporter kodiert, die offenbar in den Transport von Metallionen involviert sind.

Neben der bereits erwähnten wichtigen Funktion von Eisen sind auch andere divalente Metallionen, wie Zink oder Mangan wichtig für den bakteriellen Metabolismus^{38, 64, 84}. Zink ist für Bakterien ein essentielles Spurenelement und wird im aktiven Zentrum oder an einer Schlüsselposition von Enzymen benötigt. Sowohl im Intermediärstoffwechsel, wie beispielsweise bei der RNA-Polymerase¹⁵⁸, als auch bei Proteasen und Toxinen wird Zink im katalytischen Zentrum gefunden. Das Enterotoxin von *Bacteroides fragilis* und das *Clostridium botulinum* – Neurotoxin, das stärkste bekannte Toxin, sind Zink-Proteasen^{108, 153}.

Der durch die ORFs 0-2 kodierte Transporter weist konservierte Motive zu anderen ABC-Transportern auf. Innerhalb der einzelnen Proteine, ATP-bindendes Protein, periplasmatisches Bindungsprotein und Permease gibt es Homologien zu dem Cobalamin-Aufnahme-Transporter Btu, dem unspezifischen Metallionen-Aufnahmetransporter Tro (für Eisen, Zink und Mangan) oder den Hämin-Aufnahmetransporter Hut und Chu. Die höchsten Homologien ergeben sich zu Fe³⁺-Eisenaufnahme-ABC-Transportern und anderen Eisen-Aufnahme-Systemen wie der Fe³⁺-Siderophor-Komplex-Transporter Fhu.

Im *E. coli* Stamm ECOR31 grenzt die HPI unmittelbar an das Gen das für das ATP-bindende Protein des Transporters kodiert (ORF0). Obgleich ORF0 somit offenbar deletiert ist und keinen funktionellen Promotor besitzt, kann eine Transkription des Clusters ORF0-3 nachgewiesen werden. Dies ist am ehesten durch die Präsenz eines bidirektionalen Promotors vor dem *intB*-Gen der HPI erklärbar (Abbildung 32). Dennoch ist der ABC-Transporter auf Grund der Deletion vermutlich nicht funktionell.

Neben dem rudimentären ABC-Transporter (Orf0-2) findet man auf RegX aber auch einen ABC-Transporter mit vollständiger Gen-Zusammensetzung, ORF15-18. Dieser hat die größte strukturelle Ähnlichkeit zu zwei Transportsystemen für divalente Metallionen, dem Znu- und dem SitABC-Transporter.

Das Znu-Transport System wurde erstmals in *E. coli* beschrieben und ist in den Transport von Zink-Ionen (Zn²⁺) involviert (znu: zinc uptake)¹²⁴. ZnuA ist das

periplasmatische Protein, das die Zink-Ionen nach Passage der äußeren Membran und Eintritt in das Periplasma bindet. Der weitere Transport verläuft durch die innere cytoplasmatische Membran über das Integralprotein ZnuB, während die Energie durch Hydrolyse von ATP am cytoplasmatischen ATP-bindenden Protein ZnuC bereitgestellt wird (Abbildung 11). Die Gene werden von Zur reguliert, einem Protein ähnlich Fur bei Eisenaufnahmesystemen. Eine Inaktivierung von *znuA* und orthologen Genen führte zu einer Minderung der Wachstumsrate und Virulenz verschiedener pathogener Bakterien^{29, 95}.

Der *sit*-Locus besteht wie ORF15-18 aus vier Genen. Er kodiert einen ABC-Transporter für Eisen und Mangan^{83, 180}. Janakiraman und Mitarbeiter haben gezeigt, dass bei *Salmonella* Typhimurium die Transkription des *sit*-Operons unter Eisenmangelbedingungen und bei Invasion des intestinalen Epithels induziert wird⁷¹. Zudem konnte im Mausversuch demonstriert werden, dass die Virulenz von *sit*-Mutanten deutlich abgeschwächt ist^{16, 71}.

Das in dieser Arbeit untersuchte Gencluster ORF15-18 weist nicht nur hohe Homologien zu den erwähnten ABC-Transportern Sit und Znu auf. Mit einer RT-PCR konnte zudem eine Operon-Anordnung der Gene ORF15-18 und eine Transkription bestätigt werden, was auch auf eine intakte Funktion hindeutet.

Trotz der deutlichen Homologien zu dem Sit- und Znu-Transporter kann über die genaue Funktion des RegX-kodierten ABC-Transporters Orf15-18 zunächst nur spekuliert werden, da keine Funktionsuntersuchungen durchgeführt wurden. Es ist bekannt, dass viele divalente ABC-Transporter Spezifität für verschiedene Metallionen besitzen^{16, 38, 83}. Allerdings ist auffällig, dass in der Umgebung des ABC-Transporters auf RegX mehrere Zink-abhängige Enzyme kodiert sind, was auf eine Funktion als Zink-Transporter hindeutet.

Die Abbildung 33 gibt eine Übersicht von RegX, mit den potentiellen zinkabhängigen Enzymen, die auf dieser Region kodiert sind. **Orf12**, mit Homologien zur Carboanhydratase und **Orf21** mit Ähnlichkeiten zu PyrC, der Dihydroorotase, die in die De-novo-Synthese von Pyrimidinen involviert ist, benötigen beide Zink als Kofaktor¹⁶⁴. Zink ist auch für die enzymatische Aktivität der Porphobilinogen-Synthase (PBGS) oder auch Delta-Aminolävulinsäure-Dehydratase (ALAD), **Orf22**, essentiell. Diese spielt in der Biosynthese der Tetrapyrrol-Pigmente eine

entscheidende Rolle. Die Funktion von **Orf20**-ähnlichen Proteinen bleibt unklar, es ergeben sich jedoch konservierte Motive von Zink-Carboxypeptidasen.

Ob sich durch die Zink-abhängigen Proteine und Zink-Aufnahme-Systeme für den Stamm ECOR31 ein Selektions-Vorteil ergibt, z.B. während der Kolonisierung des Wirtes, müssen zukünftige Untersuchungen belegen.

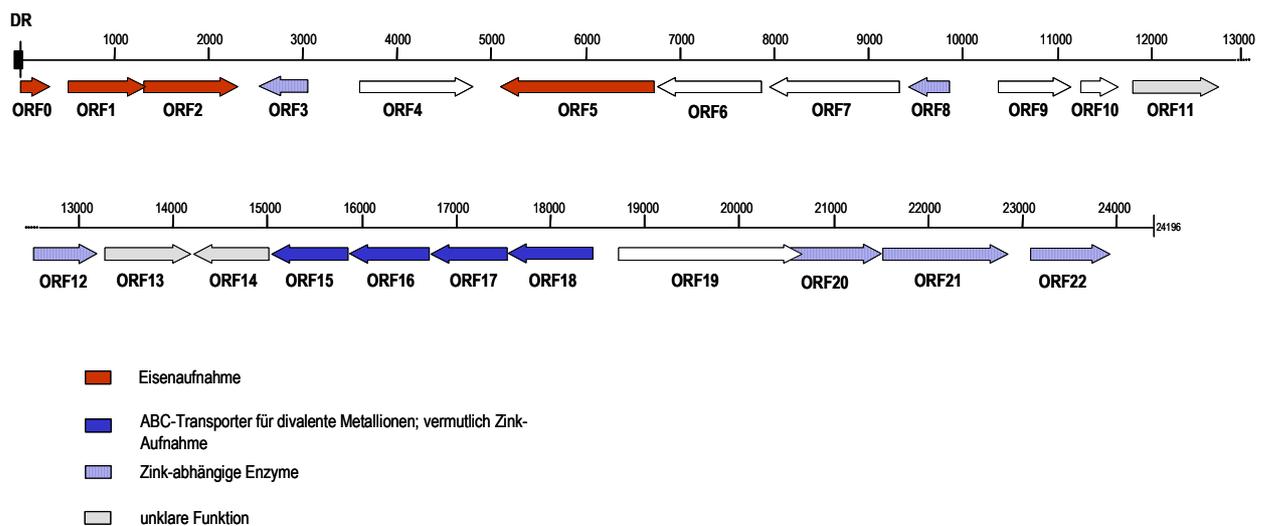


Abbildung 33: Übersicht der Region X mit den Metallionen-Aufnahme-Systemen und die dazu in Beziehung stehenden Gene.

4.1.2. Aufnahmesystem für Fe(II) und Transkriptionsregulatoren der Fur-Familie

Neben dem Fe³⁺-ABC-Eisentransporter wird auf RegX auch ein Protein mit großer Ähnlichkeit zum FeoB kodiert (**Orf5**). In Abwesenheit von Sauerstoff liegen Eisenionen in der gut wasserlöslichen reduzierten Form vor (Fe²⁺). Diese können beispielsweise über den Feo-Eisentransporter hocheffizient aufgenommen werden^{4, 170}. Die Feo-Eisentransporter stellen für viele Bakterien unter anaeroben Bedingungen eine wichtige Eisenaufnahmequelle der Zellen dar⁷⁸. Analoge Gene zu den zuerst bei *E. coli*⁷⁸ beschriebenen, anaerob induzierten Genen *feoA* und *feoB*, wurden bei verschiedenen Spezies nachgewiesen. In *E. coli* und *Salmonella* Typhimurium sind *feoB*-Mutanten in ihrer Fähigkeit den anaeroben Mäusedarm zu

kolonisieren eingeschränkt^{160, 169}. Die Virulenz von *Salmonella* zeigt sich jedoch nicht beeinflusst, was darauf hinweist, dass es dort kein essentielles Transportsystem während der Infektion ist. Aber auch in weiter entfernt verwandten Spezies zu *E. coli*, wie etwa bei *Helicobacter pylori*, findet man den Feo-Eisentransporter¹⁷⁰. Im Fall der Region X von ECOR31 liegt, wie bei *Helicobacter pylori*, das *feoB*-Gen ohne das kleinere, *feoA*-Gen vor. Das entsprechende FeoB-Protein hat in *Helicobacter pylori* eine entscheidende Rolle bei der Kolonisation und Infektion der gastralen Schleimhaut der Mäuse, sowie für das Wachstum und die Eisen-Aufnahme unter Eisenmangel-Zuständen¹⁷⁰. Dies ist ein weiteres Beispiel dafür, wie wichtig Eisen und damit Eisenaufnahmesysteme für die Pathogenität von Bakterien sind.

Neben Genen, die bei der Eisenaufnahme eine wichtige Rolle spielen, sind auf RegX auch Gene kodiert, die in die Eisenregulation eingreifen. ORF3 und ORF8 sind Gene mit Ähnlichkeiten zu Transkriptionsregulatoren aus der Fur-Familie. Fur ist ein aus 17-kDa Untereinheiten bestehendes Homodimer, das als Repressor in der Anwesenheit seines Co-Repressors Fe^{2+} die Transkription bestimmter Gene, z.B. Eisen-Aufnahme Systeme, Proteasen, Fumarase C, SodA, verschiedene Oxidoreduktasen und Dehydrogenasen unterdrückt^{4, 118} und die Expression Eisenhaltiger Proteine induziert⁴.

Bakterien regulieren ihren Eisen-Haushalt anhand der Verfügbarkeit von Eisen. In *E. coli* und in vielen anderen Bakterien geschieht dies durch das Fur-Protein (= *ferric-uptake regulator*), das die Eisen-abhängige Expression von mehr als 90 Genen in *E. coli* kontrolliert²⁰. Fur spielt zunehmend eine Rolle als globaler Regulator, nachdem zusätzlich zum Eisenhaushalt von Fur auch Gene kontrolliert werden, deren Funktion nicht mit dem Eisenhaushalt verknüpft sind, z.B. Atmung, Flagella Chemotaxis, Glykolyse, Phagen-DNA-Verpackung und DNA-Synthese^{159, 179}. Fur hat möglicherweise ebenso eine Fe^{2+} -Puffer-Funktion, da es in der Zelle in großer Zahl vorkommt, z.B. 5000 Kopien in der exponentiellen Phase bei *E. coli* und bis 10000 Kopien in der stationären Phase¹⁷⁹.

Die Eisenverfügbarkeit hat großen Einfluss auf die Gen-Expression von Bakterien. Im Genom von *Pseudomonas aeruginosa* gibt es 205 Eisen-regulierte Gene, wovon 118 durch Eisen-Mangel induziert und 87 reprimiert werden¹¹⁸. Viele von ihnen, die immerhin ein Anteil von 4% aller *Pseudomonas*-Gene ausmachen, sind Fur-

abhängig. Der Fe²⁺-Fur-Komplex bindet normalerweise zwischen der -35 und -10 *site* der DNA in der Promotorregion von Fur-reprimierten Genen, bekannt als die Eisen-Box oder Fur-Box, mit der Sequenz: GATAATGAT(A/T)ATCATTATC⁵¹. Auf RegX sind vor den ORFs 3 und 8 keine klassischen Fur-Bindungsstellen nachweisbar. Die beiden potentiellen *fur*-Gene sind jedoch in der Umgebung von Fe³⁺- und Fe²⁺-Aufnahmetransportern (**Orf0-2**, **Orf5**), sowie einer möglichen Oxidoreduktase (**Orf6**) lokalisiert.

Für Fur gibt es homologe Proteine in vielen grampositiven und gramnegativen Bakterien^{39, 64}, von denen manche, wie z.B. Zur und PerR, nicht als Eisen-abhängiger Repressor fungieren. Das Zur-Protein von *E. coli* ist ein Zink-abhängiger Repressor, PerR von *B. subtilis* hingegen reguliert die Genexpression in Abhängigkeit von Peroxid- und Mangan-Ionen^{20, 123}. Zur (zinc-uptake regulator) ist ein Repressor mit ca. 30% Ähnlichkeiten zu Fur¹²³. Wie auch Fur ist Zur weitverbreitet und manche als Fur annotierten Proteine sind möglicherweise Transkriptionsregulatoren für Zink⁶⁴. Im Promotorbereich des ABC-Transporters ORF15-18 findet sich allerdings keine Zur-Bindungssequenz, wie von Shinh beschrieben (TGaaAatgatTttCA)¹⁵⁴. Diese Zur-Box hat keine Ähnlichkeiten zur Fur-Box.

Die auf RegX kodierten Proteine **Orf4** und **Orf9** haben konservierte Motive mit dem CobW Protein gemeinsam, das bei der Cobalaminbiosynthese bei *Pseudomonas denitrificans* beteiligt ist. Nachdem sonst mehrere Gene zur Synthese von Cobalamin nötig sind und diese typischerweise in einem Operon geclustert sind³², ist hier eher davon auszugehen, dass die erwähnte Ähnlichkeit nicht mit einer intakten Funktion der Proteine einhergeht.

Die Datenbank-Vergleiche haben für die weiteren offenen Leseraster auf RegX zwei Gene ermittelt, die für zwei putative tRNA-Synthetasen (**Orf19**: Threonyl-tRNA-Synthetase, **Orf7**: Cysteinyl-tRNA-Synthetase) kodieren. Mehrere ORFs kodieren für Proteine mit teils konservierten Motiven aber unklarer Funktion, wie ORF11, 13 und 14.

Auffällig ist, dass auf RegX mehrere Gene gehäuft sind, die mit dem Transport von Metallionen und dessen Regulation verknüpft sind, oder Metallionen-abhängige Proteine kodieren.

Eine Übertragung der gesamten Region könnte dem Rezipienten somit eine metabolische Schlüsselstellung verleihen. Weitere Untersuchungen müssen zeigen, ob beispielsweise die auffällige Konzentration an Eisenaufnahme-Systemen mit dem ABC-Transporter **Orf0-2**, Feo-Transporter und der benachbarten HPI mit dem Siderophor-Eisenaufnahme-System lediglich den natürlichen Lebensraum des Stammes unter eisendepletierten Bedingungen widerspiegeln, oder sogar aktive Virulenzfaktoren darstellen.

4.2. Strukturelle Besonderheiten der Region X und deren Verbreitung

4.2.1. Ameliorationsmechanismen bakterieller Genome

Bei Vergleichen der aktuellen Datenbankeinträge (Stand 08/2007) besitzt die DNA-Region 'RegX' die höchsten Homologien zu Nukleotidsequenzen des Plasmids pLVPK des *Klebsiella pneumoniae*-Stammes CG43³⁰. Es handelt sich um ein Virulenzplasmid eines bakteriämischen Isolats von *Klebsiella pneumoniae*. Die Homologien von RegX zu pLVPK sind sowohl auf DNA- als auch auf Proteinebene sehr hoch (DNA: max. 97% Identität, Protein: 84% - 89% Identität), allerdings erstrecken sich diese Homologien nur über 77% der Gesamtsequenz. Die Sequenzen sind demnach nicht identisch.

Beim Vergleich der beiden DNA-Regionen fällt auf, welchen Wandlungen bakterielle Genome unterliegen. Die unterschiedlichen Mechanismen, die bei strukturellen Vergleichen der beiden DNA-Sequenzen deutlich werden, betreffen einerseits feine Veränderungen im bestehenden Genom, wie Punktmutationen und Basendeletionen und andererseits rekombinative Ereignisse. Ob eine Mutation das Genprodukt komplett unterbricht oder nur partiell beeinträchtigt (*leaky mutation*) hängt davon ab, an welcher Position des Aminosäurecodons die Mutation stattfindet, ob es eine *missense* Mutation ist oder eine *nonsense* Mutation und ob sich die Mutation im kodierenden Bereich, Promotor-Bereich oder im nicht-kodierenden Abschnitt befindet. Die meisten Basenpaarveränderungen führen dazu, dass in einem Protein eine Aminosäure durch eine andere ausgetauscht wird ('*missense*' Mutation oder nicht-synonyme Mutation). Die meisten Aminosäuren können jedoch durch verschiedene Codons verschlüsselt werden, wobei sich die synonymen Triplets meist nur an der 3. Stelle unterscheiden. So entstehen stille oder synonyme Mutationen (*silent mutation*) durch Punktmutationen an Position 3 des genetischen Codes, ohne dass sich die in das Polypeptid eingebaute Aminosäure ändert. Nonsense Mutationen erzeugen Stoppcodons (UAA, UAG, UGA).

Die Verteilung der Punktmutationen eines Gens zeigt dabei den Grad des Selektionsdrucks der auf das Genprodukt, respektive den damit verbundenen Phänotyp, einwirkt. Beispielsweise wurden im Gen ORF2, das für die Permease

eines ABC-Transporters kodiert, im Vergleich zu dem orthologen Gen LV125 auf pLVPK von *Klebsiella pneumoniae* insgesamt 39 Punktmutationen festgestellt. Davon führen 82% zu keiner Änderung der Aminosäure (synonyme Punktmutationen). Im Bereich ORF21 (orthologes Gen *pyrC* auf pLVPK) gibt es insgesamt 70 Punktmutationen, wovon 62% synonym sind. Für ORF5, dessen mögliches Translationsprodukt Ähnlichkeit zu FeoB besitzt, sind weniger als die Hälfte der Punktmutationen im Vergleich zu LV248 synonym.

Deletionen einzelner Basenpaare, wie beispielsweise im Bereich der poly-Adenosin-Wiederholungssequenz (*poly adenosine stretch*) zwischen Position bp 356 und bp 360 von ORF1 oder zwischen Position bp 10167 und bp 10168 vor ORF 9 führte zu einer Verkürzung beider Gene und damit zur Ausbildung von Pseudogenen. Darunter versteht man funktionslose Duplikate intakter Gene, die durch Mutationen entstanden sind. Pseudogene sind in den Annotationen oft unterrepräsentiert und deren mögliches Translationsprodukt wird häufig oft als '*hypothetical protein*' annotiert ⁹².

Rekombinative Ereignisse können zu größeren Umorganisationen im bakteriellen Genom führen: etwa zu Deletionen, Inversionen und Insertionen von DNA. Sie spielen eine herausragende Rolle beim horizontalen Gentransfer. Die Arbeitsgruppe um Dobrindt konnte eine hohe Genom-Heterogenität und genetische Diversität an verschiedenen *E. coli* Isolaten aufzeigen ⁴⁰.

Auch in der Region X ergeben sich Hinweise für Rekombinationsereignisse, wie im Bereich des orthologen Genes zu LV194, siehe DNA-Abschnitt B. Am auffälligsten jedoch ist der mosaikartige Aufbau mit zusätzlichen Genclustern, wie der Sit und Znu-ähnliche ABC-Transporter **Orf15-18** auf RegX, mit ca. 4 kb neuer DNA, oder die angrenzende *Yersinia*-HPI. Diese Veränderungen machen deutlich, wie sich die Sequenzen mit gleicher Abstammung in den beiden Stämmen weiterentwickelt haben.

4.2.2. Ko-Lokalisation der *Yersinia*-HPI und der Region X

Die Lokalisation der in dieser Arbeit beschriebenen Region X (RegX) neben der *Yersinia*-HPI ist bislang noch nicht beschrieben. Überhaupt wurde erstmals ein Bakterien-Stamm isoliert und charakterisiert, bei dem eine atypische Insertion der HPI vorliegt: die Insertion außerhalb eines tRNA-Gens. Legen wir die eingangs beschriebenen evolutiven Mechanismen bakterieller Genome zu Grunde, so ergeben sich mehrere potentielle Möglichkeiten, die eine derartige Konstellation erklären können.

Im Rahmen eines umfangreichen '*rearrangement*', also einer Umstrukturierung des bestehenden Genoms, kann es durch homologe Rekombination zu der Deletion eines größeren DNA-Abschnittes gekommen sein, der zu der Ko-Lokalisation der beiden Regionen geführt hat. Durch diese Deletion lässt sich der unmittelbare Übergang der *Yersinia*-HPI in ein Gen-Cluster eines ABC-Transporters auf RegX erklären, wobei damit sowohl der ABC-Transporter unvollständig ist, als auch die typische Insertionsstelle der HPI fehlt. Das nicht-vorhandene *asn*-tRNA-Gen zwischen HPI und RegX könnte eine Trennung der beiden Regionen oder eine Rezirkulation der HPI verhindern. Eine Deletion in der Erkennungssequenz für die Rekombination (*attachment site*) im Bereich der HPI kann zu einer Stabilisierung der Insel geführt haben. Ähnliche Stabilisierungsmechanismen sind auch durch Deletionen im Bereich des Integrase-Genes oder durch den Ersatz des *Int*-Promotors bekannt^{80, 136}. Bei *E. coli* ist die HPI auch über eine Deletion stromabwärts von *fyuA* stabilisiert¹⁵⁰ (Abbildung 1). Diese Option erscheint allerdings unwahrscheinlich, da die Deletion die komplette Pathogenitätsinsel, inklusive dem hoch-konservierten Randbereich *direct repeat* (DR), intakt lässt.

Das lässt eine zweite Hypothese plausibler erscheinen: eine Fehlinsertion der *Yersinia*-HPI im Bereich der Region X (ORF0), statt in ein *asn*-tRNA-Gen, wie sonst bei der *site*-spezifischen Rekombination der HPI üblich¹³⁸. Im Bereich ORF0 gibt es jedoch keine stärkeren Homologien zu tRNA-Genen. Eine tRNA-Gen ähnliche *attO-site* fehlt auch in der orthologen Region auf dem Plasmid pLVPK im Vergleichsstamm *Klebsiella pneumoniae* CG43, zu der sich hohe Homologien der restlichen Sequenz ergeben haben. Doch auch die RegX-analoge Region dieses Stammes ist vermutlich nicht die Ursprungs-Sequenz und könnte in diesem Bereich

verändert worden sein. Ob es sich bei RegX um eine eigenständige Genomische Insel handelt ist nicht abschließend zu klären, wobei typische Gene für Genomische Inseln, wie Integrase- und tRNA-Gene, IS-Elemente und Wiederholungssequenzen wie DR, fehlen.

4.2.3. Übertragung der DNA-Region RegX mittels horizontalem Gentransfer

Der stärkste Hinweis, dass RegX mittels horizontalem Gentransfer in den *E. coli* Stamm ECOR31 gelangt ist, ist die fehlende Übereinstimmung mit anderer *E. coli*-DNA, wobei mittlerweile mehrere *E. coli*-Genome vollständig sequenziert sind und zahlreiche DNA-Fragmente anderer *E. coli*-Stämme in den Datenbanken annotiert sind. Auch der G+C-Gehalt der Region von 47,9% entspricht nicht dem des Kern-Genoms von *E. coli* ⁴⁶ und spricht dafür, dass RegX von einem anderen Bakterium stammt. Die höchsten Homologien ergeben sich zu Vertretern einer anderen Spezies, zu *Klebsiella pneumoniae*.

Die Lokalisation einer RegX-homologen Region auf einem Plasmid bei *Klebsiella pneumoniae* ^{30, 168} unterstreicht die Möglichkeit einer Mobilisierung von Gattung zu Gattung, z.B. mit konjugativem Gentransfer. RegX ist in ECOR31, sowie in *Klebsiella pneumoniae* 2495, zudem vergesellschaftet mit der *Yersinia*-HPI, die horizontal übertragen wird ⁶⁰.

Die Frage, wie der horizontale Gentransfer jeweiliger Genomischer Inseln im Einzelfall von statten geht, bleibt bisher meist ungeklärt. Neben der 105 kb *clc* Insel von *Pseudomonas aeruginosa* ¹⁴⁰ und Symbiosis-Insel von *Mesorhizobium loti* ¹⁶² konnte experimentell auch der horizontale Gentransfer der *Yersinia*-HPI gezeigt werden. Lesic *et al.* demonstrierten die Notwendigkeit des Integrase Genes, sowie eines HPI-assoziierten Faktors *Hef* (HPI excision factor) für die Rekombination bei der Exzision der HPI von *Y. pseudotuberculosis* IP32637 ⁹³ und konnten unter bestimmten Bedingungen den horizontalen HPI-Transfer mit einer Frequenz von 10^{-8} von *Y. pseudotuberculosis* in Stämme von *Y. pestis* und *Y. pseudotuberculosis* Serotyp I demonstrieren ⁹⁴. Ein Transfer zu *Y. pseudotuberculosis* Serotyp II und IV, sowie zu *Y. enterocolitica* gelang nicht.

Antonienka und Rakin gelang der horizontale Gen-Transfer über ein konjugatives Shuttle-Plasmid ⁵. Die HPI von *Yersinia pseudotuberculosis* konnte aus dem Chromosom mobilisiert werden und in den Shuttle-Vektor, dem konjugativen Plasmid RP4, über *site*-spezifische Rekombination (*attO-site*) integrieren (Abbildung 2). Die so fixierte HPI konnte in einem zweiten Schritt konjugativ auf Rezipientenstämme übertragen werden. Im Rezipientenstamm konnte die HPI aus dem Shuttle-Vektor desintegrieren und in die *attachment sites* des Chromosoms (*asn*-tRNA-Genlocus) inserieren.

Von unserer Arbeitsgruppe wurde im *E. coli* ECOR31 eine weitere, besonders interessante Kopie der *Yersinia*-HPI entdeckt ¹⁴⁷. Im Gegensatz zu den bisher isolierten HPIs bei *E. coli* und bei *Yersinia sp.* enthält diese Kopie ein zusätzliches Fragment (Größe 35 kb). Diese Region kodiert für ein konjugatives Pilus- (*mating pair system*) und DNA-Prozessierungssystem ähnlich dem des Plasmids CloDF13 von *Enterobacter cloacae*. Die HPI von ECOR31 ist damit in der Lage, sich selbständig aus dem Chromosom herauszuschneiden (phagenartige Mobilisierung) und dann als zirkuläres, nicht replikatives Element auf konjugativem Weg in einen Rezipientenstamm zu übertragen. Daher ist die HPI von ECOR31 ein Vertreter der Familie der ICE Elemente (ICE = *integrative and conjugative element*). Offenbar stellt diese HPI-Form eine Art Urinsel oder Vorläufer-Inselle sämtlicher *E. coli* HPIs dar.

Die Gemeinsamkeit all dieser Ergebnisse ist eine Mobilisierung der Insel aus dem Genom mit Integrase und Excisionase und anschließendem Transfer in ein anderes Genom. Im Falle der in dieser Arbeit diskutierten *Yersinia*-HPI-Kopie in *E. coli* ECOR31 gehen wir davon aus, dass es sich um eine Insel handelt, die in ein natürliches Klebsiellen-Plasmid ähnlich pLVPK inseriert ist und fortan zusammen mit der nun benachbarten Region X übertragen werden kann. Dafür sprechen die hohen Homologien von RegX zu einer Region auf dem *Klebsiella*-Plasmid pLVPK, ohne benachbarte *Yersinia*-HPI und der in dieser Arbeit isolierte *Klebsiella pneumoniae* Stamm 2495, der strukturell identisch zu der HPI-RegX-Verschmelzungs-Region von ECOR31 ist (siehe Unterpunkt 4.2.4.).

In neueren, bisher unveröffentlichten Daten konnte unsere Arbeitsgruppe zeigen (Schubert, Bénédict, persönliche Mitteilung), dass die Region *Yersinia*-HPI-RegX auch bei weiteren *Klebsiella*-Isolaten vorkommt und plasmidvermittelt, also durch konjugativen Gentransfer, übertragen wird.

Die Bedeutung des Gentransfers über Plasmide, über Art- respektive Speziesgrenzen hinweg, hängt von den kodierten Genen ab. Nur drei Jahre nach dem Einführen der Cephalosporine der dritten Generation wurden bereits die ersten Resistenzen beschrieben. Diese so genannte ESBL-Gene (*extended spectrum β -lactamase*) werden über Plasmide verbreitet und sind oft zusammen mit anderen Resistenzgenen gruppiert. In den USA sprechen etwa 20% der Infektionen mit *Klebsiella pneumoniae* und ca. 30% der Infektionen mit *Enterobacter spp.* auf Intensivstationen nicht mehr auf Drittgenerations-Cephalosporine an¹²².

Denkt man sich den Transfer derartiger Regionen mit multipler Antibiotikaresistenz auf ein relevantes Virulenzplasmid, kann diese Kombination aggressive und zugleich schwer therapierbare Infektionserkrankungen hervorbringen. Welcher Selektionsvorteil sich jedoch aus dem Transfer der HPI – RegX-Verschmelzungsregion für *Enterobacteriaceae* ergibt, bleibt zu untersuchen.

4.2.4. Verbreitung der DNA-Region RegX bei der Familie der *Enterobacteriaceae*

Nach genauer Charakterisierung der Region X wurde auf Grund der Hinweise auf horizontalen Gentransfer, die Verbreitung auch bei anderen Bakterien untersucht. Da die Region aus einem *E. coli*-Stamm kommt, wurden Bakterien aus der Familie der *Enterobacteriaceae* und *Pseudomonaceae* ausgewählt. Dabei wurde in klinisch relevanten Isolaten aus Blut-, Urin- und Stuhlkulturen exemplarisch der Nachweis von RegX-Genen mittels PCR durchgeführt. Ziel des Screenings von insgesamt 437 *Enterobacteriaceae* (Abbildung 30) und 93 Pseudomonaden war es zu ermitteln, ob das Vorkommen von RegX ein häufiges Ereignis ist, oder ob diese Region nur sporadisch gefunden wird.

Die PCR-Ergebnisse waren bei 530 der untersuchten Stämme negativ und nur bei einem einzigen Stamm konnten Gene von RegX nachgewiesen werden. Besonders interessant dabei ist, dass dieser RegX-positive Stamm ein *Klebsiella pneumoniae*-Isolat aus der Blutkultur eines Patienten mit Sepsis ist, ähnlich also, wie der Stamm zu dem RegX die höchsten Homologien hat, *Klebsiella pneumoniae* CG43. Mit PCRs aus mehreren Abschnitten der Region und der HPI konnte gezeigt werden, dass die Region X und die orthologe Region auf *Klebsiella pneumoniae* 2495 strukturell gleich sind. Sämtliche Gene von RegX, d.h. auch Bereiche die bei der Region auf dem Plasmid pLVPK von *Klebsiella pneumoniae* CG43 fehlen, inklusive der *Yersinia*-HPI, sind in dem Isolat aus dem PCR-Screening nachweisbar.

In Zusammenfassung dieser Untersuchung und der Ergebnisse der Datenbanken kann man abschließend feststellen, dass die in dieser Arbeit neu beschriebene Region 'RegX' offenbar bei Klebsiellen beheimatet ist und über konjugativen Gentransfer verbreitet wird. Einschränkend muss angemerkt werden, dass es sich bei den untersuchten Isolaten um eine Sammlung aus einem Institut handelt, die sich damit auf einen bestimmten Einzugskreis belaufen. Das Ergebnis lässt sich daher nicht zwangsläufig auf andere Regionen übertragen. Ein fehlender Nachweis bei den beschriebenen Spezies spricht lediglich gegen ein sehr häufiges Vorkommen.

Zusammenfassung der Arbeit

Im *Escherichia coli* Stamm ECOR31 konnte erstmals eine atypische Insertionsstelle der weit verbreiteten Pathogenitätsinsel *Yersinia*-HPI beschrieben werden. Statt dem *asnT*-tRNA-Gen findet sich am Integrase-Ende der HPI eine weitere, offenbar horizontal transferierte Region, die RegX genannt wurde. Diese besitzt eine Größe von 24196 Basenpaaren und unterscheidet sich mit einem G+C-Gehalt von 47,9% vom *E. coli*-Kerngenom. Weder auf DNA noch auf Protein-Ebene existieren höhere Homologien zu *E. coli*-DNA.

Nach Anfertigung einer Cosmid-Bank wurde die gesamte Region sequenziert und annotiert. RegX weist eine mosaikartige Struktur auf, mit Punktmutationen, Deletionen und Insertionen. Sie besteht aus 22 offenen Leserastern und alle potentiellen Gene und deren Translationsprodukte wurden durch Vergleiche mit der NCBI-Datenbank charakterisiert. Die Transkription verschiedener Gen-Cluster und deren Operonstruktur wurde mittels RT-PCR nachgewiesen. Die DNA-Region enthält eine interessante Anhäufung von putativen Aufnahme-Systemen für divalente Metallionen, wie Eisen, Zink, Mangan. Zudem beheimatet die Region X Regulatorgene ähnlich zu Fur und Zur und Zink-abhängige Enzyme.

Im Gesamtvergleich der Region X ergeben sich die höchsten Homologien zu Teilen des Plasmids pLVPK von *Klebsiella pneumoniae* CG43³⁰. Anhand der strukturellen Unterschiede der Sequenzen mit Punktmutationen und Rekombinations-Ereignissen kann der anhaltende Wandel bakterieller Genome nachvollzogen werden.

Des Weiteren wurde die Verbreitung dieser Region unter klinisch relevanten *Enterobacteriaceae* und *Pseudomonaceae* untersucht. Von 530 gescreenten Bakterien konnte ein *Klebsiella pneumoniae* Stamm isoliert werden, der in der groben Struktur identisch zu der untersuchten Region von ECOR31 ist. In diesem Isolat, das aus der Blutkultur eines Patienten mit Sepsis stammt, konnte sowohl die gesamte Region X, als auch die benachbarte *Yersinia*-HPI und die atypische Insertionsstelle der HPI nachgewiesen werden. Es ist davon auszugehen, dass die untersuchte Region auf einem konjugativen *Klebsiella*-Plasmid lokalisiert ist und so zusammen mit der HPI horizontal zwischen verschiedenen Spezies übertragen werden kann.

Literaturverzeichnis

1. Achtman, M. Epidemic spread and antigenic variability of *Neisseria meningitidis*. *Trends Microbiol.* **3**, 186-192 (1995).
2. Andersson, J.O. & Andersson, S.G. Insights into the evolutionary process of genome degradation. *Curr. Opin. Genet. Dev.* **9**, 664-671 (1999).
3. Andersson, S.G. & Kurland, C.G. Reductive evolution of resident genomes. *Trends Microbiol.* **6**, 263-268 (1998).
4. Andrews, S.C., Robinson, A.K., & Rodriguez-Quinones, F. Bacterial iron homeostasis. *FEMS Microbiol. Rev.* **27**, 215-237 (2003).
5. Antonenka, U., Nolting, C., Heesemann, J., & Rakin, A. Horizontal transfer of *Yersinia* high-pathogenicity island by the conjugative RP4 *attB* target-presenting shuttle plasmid. *Mol. Microbiol.* **57**, 727-734 (2005).
6. Avery, O.T., Macleod, C.M., & McCarty, M. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* **149**, 297-326 (1979).
7. Bach, S., de Almeida, A., & Carniel, E. The *Yersinia* high-pathogenicity island is present in different members of the family *Enterobacteriaceae*. *FEMS Microbiol. Lett.* **183**, 289-294 (2000).
8. Barinaga, M. A shared strategy for virulence. *Science* **272**, 1261-1263 (1996).
9. Beiko, R.G. & Hamilton, N. Phylogenetic identification of lateral genetic transfer events. *Bmc Evolutionary Biology* **6**, (2006).
10. Bentley, R. & Meganathan, R. Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiol. Rev.* **46**, 241-280 (1982).
11. Bergthorsson, U. & Ochman, H. Heterogeneity of genome sizes among natural isolates of *Escherichia coli*. *J. Bacteriol.* **177**, 5784-5789 (1995).
12. Bergthorsson, U. & Ochman, H. Distribution of chromosome length variation in natural isolates of *Escherichia coli*. *Mol. Biol. Evol.* **15**, 6-16 (1998).
13. Bollivar, D.W. *et al.* Rhodobacter capsulatus porphobilinogen synthase, a high activity metal ion independent hexamer. *BMC. Biochem.* **5**, 17 (2004).

14. Bower, J.M., Eto, D.S., & Mulvey, M.A. Covert operations of uropathogenic *Escherichia coli* within the urinary tract. *Traffic*. **6**, 18-31 (2005).
15. Boyd, E.F. & Hartl, D.L. Nonrandom location of IS1 elements in the genomes of natural isolates of *Escherichia coli*. *Mol. Biol. Evol.* **14**, 725-732 (1997).
16. Boyer, E., Bergevin, I., Malo, D., Gros, P., & Cellier, M.F. Acquisition of Mn(II) in addition to Fe(II) is required for full virulence of *Salmonella enterica* serovar Typhimurium. *Infect. Immun.* **70**, 6032-6042 (2002).
17. Brem, D., Pelludat, C., Rakin, A., Jacobi, C.A., & Heesemann, J. Functional analysis of yersiniabactin transport genes of *Yersinia enterocolitica*. *Microbiology* **147**, 1115-1127 (2001).
18. Brown, J.S., Gilliland, S.M., & Holden, D.W. A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. *Mol. Microbiol.* **40**, 572-585 (2001).
19. Brunder, W. & Karch, H. Genome plasticity in Enterobacteriaceae. *Int. J. Med. Microbiol.* **290**, 153-165 (2000).
20. Bsat, N., Herbig, A., Casillas-Martinez, L., Setlow, P., & Helmann, J.D. *Bacillus subtilis* contains multiple Fur homologues: identification of the iron uptake (Fur) and peroxide regulon (PerR) repressors. *Mol. Microbiol.* **29**, 189-198 (1998).
21. Buchrieser, C., Brosch, R., Bach, S., Guiyoule, A., & Carniel, E. The high-pathogenicity island of *Yersinia pseudotuberculosis* can be inserted into any of the three chromosomal *asn tRNA* genes. *Mol. Microbiol.* **30**, 965-978 (1998).
22. Buchrieser, C. *et al.* The 102-kilobase *pgm* locus of *Yersinia pestis*: sequence analysis and comparison of selected regions among different *Yersinia pestis* and *Yersinia pseudotuberculosis* strains. *Infect. Immun.* **67**, 4851-4861 (1999).
23. Bullen, J.J., Rogers, H.J., Spalding, P.B., & Ward, C.G. Natural resistance, iron and infection: a challenge for clinical medicine. *J. Med. Microbiol.* **55**, 251-258 (2006).
24. Burland, V. *et al.* The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7. *Nucleic Acids Res.* **26**, 4196-4204 (1998).
25. Carniel, E. The *Yersinia* high-pathogenicity island. *Int. Microbiol.* **2**, 161-167 (1999).
26. Carniel, E. The *Yersinia* high-pathogenicity island: an iron-uptake island. *Microbes. Infect.* **3**, 561-569 (2001).

27. Carniel,E., Guilvout,I., & Prentice,M. Characterization of a large chromosomal "high-pathogenicity island" in biotype 1B *Yersinia enterocolitica*. *J. Bacteriol.* **178**, 6743-6751 (1996).
28. Carniel,E., Mercereau Puijalon,O., & Bonnefoy,S. The gene coding for the 190,000-dalton iron-regulated protein of *Yersinia* species is present only in the highly pathogenic strains. *Infect. Immun.* **57**, 1211-1217 (1989).
29. Chen,C.Y. & Morse,S.A. Identification and characterization of a high-affinity zinc uptake system in *Neisseria gonorrhoeae*. *FEMS Microbiol. Lett.* **202**, 67-71 (2001).
30. Chen,Y.T. *et al.* Sequencing and analysis of the large virulence plasmid pLVPK of *Klebsiella pneumoniae* CG43. *Gene* **337**, 189-198 (2004).
31. Clarke,S.C., Haigh,R.D., Freestone,P.P., & Williams,P.H. Virulence of enteropathogenic *Escherichia coli*, a global pathogen. *Clin. Microbiol Rev.* **16**, 365-378 (2003).
32. Crouzet,J. *et al.* Nucleotide sequence and genetic analysis of a 13.1-kilobase-pair *Pseudomonas denitrificans* DNA fragment containing five cob genes and identification of structural genes encoding Cob(I)alamin adenosyltransferase, cobyrinic acid synthase, and bifunctional cobinamide kinase-cobinamide phosphate guanylyltransferase. *J. Bacteriol.* **173**, 6074-6087 (1991).
33. Davies,B.W. & Walker,G.C. Disruption of *sitA* compromises *Sinorhizobium meliloti* for manganese uptake required for protection against oxidative stress. *J. Bacteriol.* **189**, 2101-2109 (2007).
34. Davies,J. Origins and evolution of antibiotic resistance. *Microbiologia* **12**, 9-16 (1996).
35. de Almeida,A.M. *et al.* Chromosomal *irp2* gene in *Yersinia*: distribution, expression, deletion and impact on virulence. *Microb. Pathogen.* **14**, 9-21 (1993).
36. Deana,A., Ehrlich,R., & Reiss,C. Silent mutations in the *Escherichia coli* *ompA* leader peptide region strongly affect transcription and translation in vivo. *Nucleic Acids Res.* **26**, 4778-4782 (1998).
37. Demir,M. & Kaleli,I. Production by *Escherichia coli* isolates of siderophore and other virulence factors and their pathogenic role in a cutaneous infection model. *Clin. Microbiol. Infect.* **10**, 1011-1014 (2004).
38. Desrosiers,D.C. *et al.* The general transition metal (Tro) and Zn²⁺ (Znu) transporters in *Treponema pallidum*: analysis of metal specificities and expression profiles. *Mol. Microbiol.* **65**, 137-152 (2007).
39. Diaz-Mireles,E. *et al.* The Fur-like protein Mur of *Rhizobium leguminosarum* is a Mn(2+)-responsive transcriptional regulator. *Microbiology* **150**, 1447-1456 (2004).

40. Dobrindt,U. *et al.* Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J. Bacteriol.* **185**, 1831-1840 (2003).
41. Dobrindt,U. *et al.* Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *J. Bacteriol.* **185**, 1831-1840 (2003).
42. Dobrindt,U. *et al.* Genetic structure and distribution of four pathogenicity islands (PAI I(536) to PAI IV(536)) of uropathogenic *Escherichia coli* strain 536. *Infect. Immun.* **70**, 6365-6372 (2002).
43. Dobrindt,U. & Hacker,J. Whole genome plasticity in pathogenic bacteria. *Curr. Opin. Microbiol.* **4**, 550-557 (2001).
44. Dobrindt,U., Hochhut,B., Hentschel,U., & Hacker,J. Genomic islands in pathogenic and environmental microorganisms. *Nat. Rev. Microbiol.* **2**, 414-424 (2004).
45. Doublet,B., Boyd,D., Mulvey,M.R., & Cloeckert,A. The *Salmonella* genomic island 1 is an integrative mobilizable element. *Mol. Microbiol.* **55**, 1911-1924 (2005).
46. Dougan,G. *et al.* The *Escherichia coli* gene pool. *Curr. Opin. Microbiol.* **4**, 90-94 (2001).
47. Dubnau,D. DNA uptake in bacteria. *Annu. Rev. Microbiol.* **53**, 217-244 (1999).
48. Eaton,K.A., Brooks,C.L., Morgan,D.R., & Krakowka,S. Essential role of urease in pathogenesis of gastritis induced by *Helicobacter pylori* in gnotobiotic piglets. *Infect. Immun.* **59**, 2470-2475 (1991).
49. Emody,L., Kerenyi,M., & Nagy,G. Virulence factors of uropathogenic *Escherichia coli*. *Int. J. Antimicrob. Agents* **22 Suppl 2**, 29-33 (2003).
50. Eriani,G., Delarue,M., Poch,O., Gangloff,J., & Moras,D. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* **347**, 203-206 (1990).
51. Escolar,L., Perez-Martin,J., & de,L., V Opening the iron box: transcriptional metalloregulation by the Fur protein. *J. Bacteriol.* **181**, 6223-6229 (1999).
52. Farthing,M.J. Bugs and the gut: an unstable marriage
1. *Best. Pract. Res Clin Gastroenterol.* **18**, 233-239 (2004).
53. Finlay,B.B. & Falkow,S. Common themes in microbial pathogenicity. *Microbiol. Rev.* **53**, 210-230 (1989).
54. Foxman,B. & Brown,P. Epidemiology of urinary tract infections: transmission and risk factors, incidence, and costs. *Infect. Dis. Clin. North Am.* **17**, 227-241 (2003).

-
55. Groisman, E.A. & Ochman, H. Pathogenicity islands: bacterial evolution in quantum leaps. *Cell* **87**, 791-794 (1996).
 56. Guarner, F. Enteric flora in health and disease
1. *Digestion* **73 Suppl 1**, 5-12 (2006).
 57. Hacker, J., Blum-Oehler, G., Mühldorfer, I., & Tschäpe, H. Pathogenicity islands of virulent bacteria: structure, function and impact on microbial evolution. *Mol. Microbiol.* **23**, 1089-1097 (1997).
 58. Hacker, J. & Carniel, E. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Rep.* **2**, 376-381 (2001).
 59. Hacker, J. *et al.* Pathogenomics of mobile genetic elements of toxigenic bacteria. *Int. J. Med. Microbiol.* **293**, 453-461 (2004).
 60. Hacker, J. & Kaper, J.B. Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* **54**, 641-679 (2000).
 61. Hakenbeck, R. *et al.* Mosaic genes and mosaic chromosomes: intra- and interspecies genomic variation of *Streptococcus pneumoniae*. *Infect. Immun.* **69**, 2477-2486 (2001).
 62. Hallet, B. & Sherratt, D.J. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS Microbiol. Rev.* **21**, 157-178 (1997).
 63. Hanage, W.P., Fraser, C., & Spratt, B.G. The impact of homologous recombination on the generation of diversity in bacteria. *Journal of Theoretical Biology* **239**, 210-219 (2006).
 64. Hantke, K. Bacterial zinc transporters and regulators. *Biometals* **14**, 239-249 (2001).
 65. Hare, J.M., Wagner, A.K., & McDonough, K.A. Independent acquisition and insertion into different chromosomal locations of the same pathogenicity island in *Yersinia pestis* and *Yersinia pseudotuberculosis*. *Mol. Microbiol.* **31**, 291-303 (1999).
 66. Heesemann, J. Chromosomal-encoded siderophores are required for mouse virulence of enteropathogenic *Yersinia* species. *FEMS Microbiol. Lett.* **48**, 229-233 (1987).
 67. Heesemann, J. *et al.* Plasmids of human strains of *Yersinia enterocolitica*: molecular relatedness and possible importance for pathogenesis. *J. Infect. Dis.* **147**, 107-115 (1983).
 68. Higgins, C.F. ABC transporters: from microorganisms to man. *Annu. Rev. Cell. Biol.* **8**, 67-113 (1992).

69. Hoess,R.H. & Landy,A. Structure of the lambda att sites generated by int-dependent deletions. *Proc. Natl. Acad. Sci. U. S. A* **75**, 5437-5441 (1978).
70. Hooper,D.C. & Wolfson,J.S. Bacterial resistance to the quinolone antimicrobial agents. *Am. J. Med.* **87**, 17S-23S (1989).
71. Janakiraman,A. & Slauch,J.M. The putative iron transport system SitABCD encoded on SPI1 is required for full virulence of Salmonella typhimurium. *Mol. Microbiol.* **35**, 1146-1155 (2000).
72. Javaloyas,M., Garcia-Somoza,D., & Gudiol,F. Epidemiology and prognosis of bacteremia: a 10-y study in a community hospital. *Scand. J. Infect. Dis.* **34**, 436-441 (2002).
73. Johnson,J.R. Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev.* **4**, 80-128 (1991).
74. Johnson,J.R., Brown,J.J., Carlino,U.B., & Russo,T.A. Colonization with and acquisition of uropathogenic *Escherichia coli* as revealed by polymerase chain reaction-based detection. *J. Infect. Dis.* **177**, 1120-1124 (1998).
75. Johnson,J.R., Owens,K., Gajewski,A., & Kuskowski,M.A. Bacterial characteristics in relation to clinical source of *Escherichia coli* isolates from women with acute cystitis or pyelonephritis and uninfected women. *J. Clin. Microbiol.* **43**, 6064-6072 (2005).
76. Johnson,J.R. & Russo,T.A. Uropathogenic *Escherichia coli* as agents of diverse non-urinary tract extraintestinal infections. *J. Infect. Dis.* **186**, 859-864 (2002).
77. Johnson,J.R. & Stell,A.L. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J Infect Dis* **181**, 261-272 (2000).
78. Kammler,M., Schon,C., & Hantke,K. Characterization of the ferrous iron uptake system of *Escherichia coli*. *J. Bacteriol.* **175**, 6212-6219 (1993).
79. Kang,P.J. & Craig,E.A. Identification and characterization of a new *Escherichia coli* gene that is a dosage-dependent suppressor of a dnaK deletion mutation. *J. Bacteriol.* **172**, 2055-2064 (1990).
80. Karch,H. *et al.* A genomic island, termed high-pathogenicity island, is present in certain non-0157 Shiga Toxin-producing *Escherichia coli* lineages. *Infect. Immun.* **67**, 5994-6001 (1999).
81. Karlin,S. Detecting anomalous gene clusters and pathogenicity islands in diverse bacterial genomes. *Trends Microbiol.* **9**, 335-343 (2001).
82. Kayser,F.H., Bienz,K.A., Eckert,J., & Zinkernagel,R.M. *Medizinische Mikrobiologie*(Georg Thieme Verlag, Stuttgart, 1998).

83. Kehres,D.G., Janakiraman,A., Slauch,J.M., & Maguire,M.E. SitABCD is the alkaline Mn(2+) transporter of *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* **184**, 3159-3166 (2002).
84. Kehres,D.G. & Maguire,M.E. Emerging themes in manganese transport, biochemistry and pathogenesis in bacteria. *FEMS Microbiol. Rev.* **27**, 263-290 (2003).
85. Kelly,D. & Conway,S. Bacterial modulation of mucosal innate immunity. *Mol. Immunol.* **42**, 895-901 (2005).
86. Kerr,I.D. Structure and association of ATP-binding cassette transporter nucleotide-binding domains. *Biochim. Biophys. Acta* **1561**, 47-64 (2002).
87. Kim,G.J. & Kim,H.S. Identification of the structural similarity in the functionally related amidohydrolases acting on the cyclic amide ring. *Biochem. J.* **330 (Pt 1)**, 295-302 (1998).
88. Koczura,R. & Kaznowski,A. Occurrence of the Yersinia high-pathogenicity island and iron uptake systems in clinical isolates of *Klebsiella pneumoniae*. *Microb. Pathog.* **35**, 197-202 (2003).
89. Köster,W. ABC transporter-mediated uptake of iron, siderophores, heme and vitamin B12. *Res. Microbiol.* **152**, 291-301 (2001).
90. Lawrence,J.G. Gene transfer, speciation, and the evolution of bacterial genomes. *Curr. Opin. Microbiol.* **2**, 519-523 (1999).
91. Lederberg,J. & TATUM,E.L. Sex in bacteria; genetic studies, 1945-1952. *Science* **118**, 169-175 (1953).
92. Lerat,E. & Ochman,H. Recognizing the pseudogenes in bacterial genomes. *Nucleic Acids Res.* **33**, 3125-3132 (2005).
93. Lesic,B. *et al.* Excision of the high-pathogenicity island of *Yersinia pseudotuberculosis* requires the combined actions of its cognate integrase and Hef, a new recombination directionality factor. *Mol. Microbiol.* **52**, 1337-1348 (2004).
94. Lesic,B. & Carniel,E. Horizontal transfer of the high-pathogenicity island of *Yersinia pseudotuberculosis*. *J. Bacteriol.* **187**, 3352-3358 (2005).
95. Lewis,D.A. *et al.* Identification of the *znuA*-encoded periplasmic zinc transport protein of *Haemophilus ducreyi*. *Infect. Immun.* **67**, 5060-5068 (1999).
96. Linton,K.J. Structure and function of ABC transporters. *Physiology. (Bethesda.)* **22**, 122-130 (2007).
97. Luck,S.N., Turner,S.A., Rajakumar,K., Sakellaris,H., & Adler,B. Ferric dicitrate transport system (Fec) of *Shigella flexneri* 2a YSH6000 is encoded on a novel pathogenicity island carrying multiple antibiotic resistance genes. *Infect. Immun.* **69**, 6012-6021 (2001).

-
98. Lynch,M., Painter,J., Woodruff,R., & Braden,C. Surveillance for foodborne-disease outbreaks--United States, 1998-2002. *MMWR Surveill Summ.* **55**, 1-42 (2006).
 99. Manning,P.A. *et al.* L-pilin variants of Neisseria gonorrhoeae MS11. *Mol. Microbiol.* **5**, 917-926 (1991).
 100. Marcus,E.A., Moshfegh,A.P., Sachs,G., & Scott,D.R. The periplasmic alpha-carbonic anhydrase activity of Helicobacter pylori is essential for acid acclimation. *J. Bacteriol.* **187**, 729-738 (2005).
 101. Marlovits,T.C., Haase,W., Herrmann,C., Aller,S.G., & Unger,V.M. The membrane protein FeoB contains an intramolecular G protein essential for Fe(II) uptake in bacteria. *Proc. Natl. Acad. Sci. U. S. A* **99**, 16243-16248 (2002).
 102. Mau,B., Glasner,J.D., Darling,A.E., & Perna,N.T. Genome-wide detection and analysis of homologous recombination among sequenced strains of Escherichia coli. *Genome Biol.* **7**, R44 (2006).
 103. Maurelli,A.T., Fernandez,R.E., Bloch,C.A., Rode,C.K., & Fasano,A. "Black holes" and bacterial pathogenicity: a large genomic deletion that enhances the virulence of *Shigella* spp. and enteroinvasive *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A* **95**, 3943-3948 (1998).
 104. Mazel,D., Dychinco,B., Webb,V.A., & Davies,J. Antibiotic resistance in the ECOR collection: integrons and identification of a novel aad gene. *Antimicrob. Agents Chemother.* **44**, 1568-1574 (2000).
 105. McDaniel,T.K. & Kaper,J.B. A cloned pathogenicity island from enteropathogenic *Escherichia coli* confers the attaching and effacing phenotype on *E. coli* K-12. *Mol. Microbiol.* **23**, 399-407 (1997).
 106. Mehta,N., Olson,J.W., & Maier,R.J. Characterization of Helicobacter pylori nickel metabolism accessory proteins needed for maturation of both urease and hydrogenase. *J. Bacteriol.* **185**, 726-734 (2003).
 107. Mietzner,T.A. & Morse,S.A. The role of iron-binding proteins in the survival of pathogenic bacteria. *Annu. Rev. Nutr.* **14**, 471-493 (1994).
 108. Moncrief,J.S. *et al.* The enterotoxin of Bacteroides fragilis is a metalloprotease. *Infect. Immun.* **63**, 175-181 (1995).
 109. Morris,R.T. & Drouin,G. Ectopic gene conversions in four *Escherichia coli* genomes: increased recombination in pathogenic strains. *J. Mol. Evol.* **58**, 596-605 (2004).
 110. Morse,M.L., Lederberg,E.M., & Lederberg,J. Transduction in *Escherichia coli* K-12. *Genetics* **41**, 142-156 (1956).

-
111. Moss, J.E., Cardozo, T.J., Zychlinsky, A., & Groisman, E.A. The selC-associated SHI-2 pathogenicity island of *Shigella flexneri*. *Mol. Microbiol.* **33**, 74-83 (1999).
 112. Mushegian, A.R. & Koonin, E.V. A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc. Natl. Acad. Sci. U. S. A* **93**, 10268-10273 (1996).
 113. Muto, A. & Osawa, S. The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc. Natl. Acad. Sci. U. S. A* **84**, 166-169 (1987).
 114. Ochman, H. & Davalos, L.M. The nature and dynamics of bacterial genomes. *Science* **311**, 1730-1733 (2006).
 115. Ochman, H., Lawrence, J.G., & Groisman, E.A. Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**, 299-304 (2000).
 116. Ochman, H. & Moran, N.A. Genes lost and genes found: evolution of bacterial pathogenesis and symbiosis. *Science* **292**, 1096-1099 (2001).
 117. Ochman, H. & Selander, R.K. Standard reference strains of *Escherichia coli* from natural populations. *J. Bacteriol.* **157**, 690-693 (1984).
 118. Ochsner, U.A., Vasil, A.I., & Vasil, M.L. Role of the ferric uptake regulator of *Pseudomonas aeruginosa* in the regulation of siderophores and exotoxin A expression: purification and activity on iron-regulated promoters. *J. Bacteriol.* **177**, 7194-7201 (1995).
 119. Okeke, I.N., Scaletsky, I.C., Soars, E.H., Macfarlane, L.R., & Torres, A.G. Molecular epidemiology of the iron utilization genes of enteroaggregative *Escherichia coli*. *J. Clin. Microbiol.* **42**, 36-44 (2004).
 120. Pal, C., Papp, B., & Lercher, M.J. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat. Genet.* **37**, 1372-1375 (2005).
 121. Parkhill, J. *et al.* Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature* **413**, 848-852 (2001).
 122. Paterson, D.L. Resistance in gram-negative bacteria: Enterobacteriaceae. *Am. J. Infect. Control* **34**, S20-S28 (2006).
 123. Patzer, S.I. & Hantke, K. The zinc-responsive regulator Zur and its control of the *znu* gene cluster encoding the ZnuABC zinc uptake system in *Escherichia coli*. *J. Biol. Chem.* **275**, 24321-24332 (2000).
 124. Patzer, S.I. & Hantke, K. The ZnuABC high-affinity zinc uptake system and its regulator Zur in *Escherichia coli*. *Mol. Microbiol.* **28**, 1199-1210 (1998).
 125. Pelludat, C., Hogardt, M., & Heesemann, J. Transfer of the core region genes of the *Yersinia enterocolitica* WA-C serotype O:8 high-pathogenicity island to *Y. enterocolitica* MRS40, a strain with low levels of pathogenicity,

- confers a yersiniabactin biosynthesis phenotype and enhanced mouse virulence. *Infect. Immun.* **70**, 1832-1841 (2002).
126. Pembroke, J.T., MacMahon, C., & McGrath, B. The role of conjugative transposons in the *Enterobacteriaceae*. *Cell Mol. Life Sci.* **59**, 2055-2064 (2002).
127. Perna, N.T. *et al.* Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529-533 (2001).
128. Perry, R.D., Balbo, P.B., Jones, H.A., Fetherston, J.D., & DeMoll, E. Yersiniabactin from *Yersinia pestis*: biochemical characterization of the siderophore and its role in iron transport and regulation. *Microbiology* **145** (Pt 5), 1181-1190 (1999).
129. Perry, R.D. & San Clemente, C.L. Siderophore synthesis in *Klebsiella pneumoniae* and *Shigella sonnei* during iron deficiency. *J. Bacteriol.* **140**, 1129-1132 (1979).
130. Picard, B. *et al.* The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun.* **67**, 546-553 (1999).
131. Plunkett, G., III, Rose, D.J., Durfee, T.J., & Blattner, F.R. Sequence of Shiga toxin 2 phage 933W from *Escherichia coli* O157:H7: Shiga toxin as a phage late-gene product. *J. Bacteriol.* **181**, 1767-1778 (1999).
132. Portnoy, D.A. & Falkow, S. Virulence-associated plasmids from *Yersinia enterocolitica* and *Yersinia pestis*. *J. Bacteriol.* **148**, 877-883 (1981).
133. Purdy, G.E. & Payne, S.M. The SHI-3 iron transport island of *Shigella boydii* 0-1392 carries the genes for aerobactin synthesis and transport. *J. Bacteriol.* **183**, 4176-4182 (2001).
134. Qiu, X., Gurkar, A.U., & Lory, S. Interstrain transfer of the large pathogenicity island (PAPI-1) of *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. U. S. A* **103**, 19830-19835 (2006).
135. Rajadhyaksha, A.B. & Rao, S.S. The role of phage in the transduction of the toxinogenic factor in *Corynebacterium diphtheriae*. *J. Gen. Microbiol.* **40**, 421-429 (1965).
136. Rakin, A., Noelting, C., Schubert, S., & Heesemann, J. Common and specific characteristics of the high-pathogenicity island of *Yersinia enterocolitica*. *Infect. Immun.* **67**, 5265-5274 (1999).
137. Rakin, A., Saken, E., Harmsen, D., & Heesemann, J. The pesticin receptor of *Yersinia enterocolitica*: a novel virulence factor with dual function. *Mol. Microbiol.* **13**, 253-263 (1994).
138. Rakin, A., Schubert, S., Pelludat, C., Brem, D., & Heesemann, J. The high-pathogenicity island of *Yersinia* in *Pathogenicity Islands* (eds. Hacker, J. & Kaper, J.B.) (ASM-Press, Washington, 1999).

-
139. Rakin,A., Urbitsch,P., & Heesemann,J. Evidence for two evolutionary lineages of highly pathogenic *Yersinia* species. *J. Bacteriol.* **177**, 2292-2298 (1995).
 140. Ravatn,R., Zehnder,A.J., & van,d.M., Jr. Low-frequency horizontal transfer of an element containing the chlorocatechol degradation genes from *Pseudomonas* sp. strain B13 to *Pseudomonas putida* F1 and to indigenous bacteria in laboratory-scale activated-sludge microcosms. *Appl. Environ. Microbiol.* **64**, 2126-2132 (1998).
 141. Reiter,W.D., Palm,P., & Yeats,S. Transfer RNA genes frequently serve as integration sites for prokaryotic genetic elements. *Nucleic. Acids. Res.* **17**, 1907-1914 (1989).
 142. Roth,J.R., Lawrence,J.G., & Bobik,T.A. Cobalamin (coenzyme B12): synthesis and biological significance. *Annu. Rev. Microbiol.* **50**, 137-181 (1996).
 143. Russo,T.A. & Johnson,J.R. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes. Infect.* **5**, 449-456 (2003).
 144. Salama,N.R. & Falkow,S. Genomic clues for defining bacterial pathogenicity. *Microbes. Infect.* **1**, 615-619 (1999).
 145. Sanger,F., Nicklen,S., & Coulson,A.R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467 (1977).
 146. Schubert,S., Cuenca,S., Fischer,D., & Heesemann,J. High-pathogenicity island of *Yersinia pestis* in *Enterobacteriaceae* isolated from blood cultures and urine samples: prevalence and functional expression. *J. Infect. Dis.* **182**, 1268-1271 (2000).
 147. Schubert,S., Dufke,S., Sorsa,J., & Heesemann,J. A novel integrative and conjugative element (ICE) of *Escherichia coli*: the putative progenitor of the *Yersinia* high-pathogenicity island. *Mol. Microbiol.* **51**, 837-848 (2004).
 148. Schubert,S., Picard,B., Gouriou,S., Heesemann,J., & Denamur,E. *Yersinia* high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. *Infect Immun.* **70**, 5335-5337 (2002).
 149. Schubert,S., Rakin,A., Fischer,D., Sorsa,J., & Heesemann,J. Characterization of the integration site of *Yersinia* high-pathogenicity island in *Escherichia coli*. *FEMS Microbiol. Lett.* **179**, 409-414 (1999).
 150. Schubert,S., Rakin,A., & Heesemann,J. The *Yersinia* high-pathogenicity island (HPI): evolutionary and functional aspects. *Int. J Med Microbiol.* **294**, 83-94 (2004).
 151. Schubert,S., Rakin,A., Karch,H., Carniel,E., & Heesemann,J. Prevalence of the "high-pathogenicity island" of *Yersinia* species among

- Escherichia coli* strains that are pathogenic to humans. *Infect Immun.* **66**, 480-485 (1998).
152. Schubert, S. *et al.* HPI of high-virulent *Yersinia* is found in *E. coli* strains causing urinary tract infection. Structural, functional aspects, and distribution. *Adv. Exp. Med Biol.* **485**, 69-73 (2000).
153. Segelke, B., Knapp, M., Kadkhodayan, S., Balhorn, R., & Rupp, B. Crystal structure of Clostridium botulinum neurotoxin protease in a product-bound state: Evidence for noncanonical zinc protease activity. *Proc. Natl. Acad. Sci. U. S. A* **101**, 6888-6893 (2004).
154. Shin, J.H., Oh, S.Y., Kim, S.J., & Roe, J.H. The zinc-responsive regulator Zur controls a zinc uptake system and some ribosomal proteins in Streptomyces coelicolor A3(2). *J. Bacteriol.* **189**, 4070-4077 (2007).
155. Shinagawa, H. & Iwasaki, H. Processing the holliday junction in homologous recombination. *Trends Biochem. Sci.* **21**, 107-111 (1996).
156. Snyder, L. & Champness, W. *Molecular Genetics of Bacteria* (ASM Press, Washington, D.C., 1996).
157. Sokurenko, E.V. *et al.* Pathogenic adaptation of Escherichia coli by natural variation of the FimH adhesin. *Proc. Natl. Acad. Sci. U. S. A* **95**, 8922-8926 (1998).
158. Solaiman, D. & Wu, F.Y. Intrinsic zinc ion is essential for proper conformation of active Escherichia coli RNA polymerase. *Biochemistry* **23**, 6369-6377 (1984).
159. Stojiljkovic, I., Baumler, A.J., & Hantke, K. Fur regulon in gram-negative bacteria. Identification and characterization of new iron-regulated *Escherichia coli* genes by a fur titration assay [published erratum appears in J Mol Biol 1994 Jul 15;240(3):271]. *J. Mol. Biol.* **236**, 531-545 (1994).
160. Stojiljkovic, I., Cobeljic, M., & Hantke, K. Escherichia coli K-12 ferrous iron uptake mutants are impaired in their ability to colonize the mouse intestine. *FEMS Microbiol. Lett.* **108**, 111-115 (1993).
161. Stover, C.K. *et al.* Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen. *Nature* **406**, 959-964 (2000).
162. Sullivan, J.T. & Ronson, C.W. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc. Natl. Acad. Sci. U. S. A* **95**, 5145-5149 (1998).
163. TATUM, E.L. & Lederberg, J. Gene Recombination in the Bacterium Escherichia coli. *J. Bacteriol.* **53**, 673-684 (1947).
164. Thoden, J.B., Phillips, G.N., Jr., Neal, T.M., Raushel, F.M., & Holden, H.M. Molecular structure of dihydroorotase: a paradigm for catalysis through the use of a binuclear metal center. *Biochemistry* **40**, 6989-6997 (2001).

-
165. Tlaskalova-Hogenova, H. *et al.* Commensal bacteria (normal microflora), mucosal immunity and chronic inflammatory and autoimmune diseases. *Immunol. Lett.* **93**, 97-108 (2004).
 166. Tobe, T. *et al.* Complete DNA sequence and structural analysis of the enteropathogenic *Escherichia coli* adherence factor plasmid. *Infect. Immun.* **67**, 5455-5462 (1999).
 167. Torres, A.G., Redford, P., Welch, R.A., & Payne, S.M. TonB-dependent systems of uropathogenic *Escherichia coli*: aerobactin and heme transport and TonB are required for virulence in the mouse. *Infect Immun.* **69**, 6179-6185 (2001).
 168. Tsai, H.C. *et al.* Complete genome sequences of *Klebsiella pneumoniae* isolated from a liver abscess patient, unpublished, Accession AP006726. 2005.
Ref Type: Internet Communication
 169. Tsolis, R.M., Baumler, A.J., Heffron, F., & Stojiljkovic, I. Contribution of TonB- and Feo-mediated iron uptake to growth of *Salmonella typhimurium* in the mouse. *Infect. Immun.* **64**, 4549-4556 (1996).
 170. Velayudhan, J. *et al.* Iron acquisition and virulence in *Helicobacter pylori*: a major role for FeoB, a high-affinity ferrous iron transporter. *Mol. Microbiol.* **37**, 274-286 (2000).
 171. Vokes, S.A., Reeves, S.A., Torres, A.G., & Payne, S.M. The aerobactin iron transport system genes in *Shigella flexneri* are present within a pathogenicity island. *Mol. Microbiol.* **33**, 63-73 (1999).
 172. Wandersman, C. & Delepelaire, P. Bacterial iron sources: from siderophores to hemophores. *Annu. Rev. Microbiol.* **58**, 611-647 (2004).
 173. Welch, R.A. *et al.* Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A* **99**, 17020-17024 (2002).
 174. Wilkins, B.M. & Frost, L.S. Mechanism of Gene Exchange Between Bacteria in *Molecular Medical Microbiology* (ed. Sussmann, M.) 355-400 (Academic Press, 2002, 2007).
 175. Williams, A.M. *et al.* Effects of microflora on the neonatal development of gut mucosal T cells and myeloid cells in the mouse. *Immunology* **119**, 470-478 (2006).
 176. Wisplinghoff, H. *et al.* Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. *Clin. Infect. Dis.* **39**, 309-317 (2004).
 177. Wolf, G. Gut microbiota: a factor in energy regulation. *Nutr. Rev.* **64**, 47-50 (2006).

-
178. Wolf, Y.I., Aravind, L., Grishin, N.V., & Koonin, E.V. Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **9**, 689-710 (1999).
 179. Zheng, M., Doan, B., Schneider, T.D., & Storz, G. OxyR and SoxRS regulation of fur. *J. Bacteriol.* **181**, 4639-4643 (1999).
 180. Zhou, D., Hardt, W.D., & Galan, J.E. Salmonella typhimurium encodes a putative iron transport system within the centisome 63 pathogenicity island. *Infect. Immun.* **67**, 1974-1981 (1999).
 181. Ziebuhr, W., Ohlsen, K., Karch, H., Korhonen, T., & Hacker, J. Evolution of bacterial pathogenesis. *Cell Mol. Life Sci.* **56**, 719-728 (1999).
 182. ZINDER, N.D. & Lederberg, J. Genetic exchange in Salmonella. *J. Bacteriol.* **64**, 679-699 (1952).

Anhang

Tabelle 1: Offene Leseraster von RegX mit Positionsangabe, sowie Größe und Homologien der Translationsprodukte nach online Vergleichen mit der Dantenbank (NCBI Blast search Algorithmus, <http://www.ncbi.nlm.nih.gov/BLAST>).

ORF	ORF Position	Orf Größe	Homologien (BLAST)	Quelle	% Identität/ Ähnlichkeit	ident. aa	GenBank Accession-Nr
ORF0	1-331	105 aa	iron ABC transporter ATP-binding prot ein putative iron(III) ABC transporter, ATP-binding protein	<i>Klebsiella pneumoniae</i> <i>Erwinia carotovora</i>	96%/ 98% 55%/ 69%	99/103 57/102	AAR07733.1 YP_052341.1
ORF1	496-1308	271 aa	iron (III) ABC transporter periplasmic binding putative iron(III) ABC transporter, periplasmic-binding protein	<i>Klebsiella pneumoniae</i> <i>Erwinia carotovora</i> subsp. <i>atroseptica</i> SCRI1043	94%/ 97% 70%/ 80%	256/270 192/273	AAR07732.1 YP_052342.1
ORF2	1308-2327	339 aa	ABC transporter permease protein putative iron(III) ABC transporter, permease protein	<i>Klebsiella pneumoniae</i> <i>Erwinia carotovora</i>	98%/ 99% 65%/ 80%	329/333 217/329	AAR07731.1 YP_052343.1
ORF3	3036-2533	167 aa	hypothetical protein LV197 putative regulator	<i>Klebsiella pneumoniae</i> <i>Sodalis glossinidius</i>	54%/ 72% 32%/ 53%	90/165 44/137	AAR07788.1 YP_455819.1
ORF4	3605-4828	407 aa	CobW Cobalamin synthesis protein/P47K	<i>Klebsiella pneumoniae</i> <i>Burkholderia</i> sp. 383	94%/ 98% 73%/ 86%	383/406 290/392	AAR07790.1 YP_371076.1
ORF5	6744-5095	549 aa	ferrus ion transporter protein transporter 43 (probable ferrous iron transport protein)	<i>Klebsiella pneumoniae</i> <i>Cyanotheca</i> sp. CCY0110	88%/ 93% 39%/ 58%	488/549 214/547	AAR07792.1 ZP_01729856.1
ORF6	7886-6756	376 aa	hypothetical protein LV193 probable monooxygenase	<i>Klebsiella pneumoniae</i> <i>Crocospaera watsonii</i> WH 8501	98%/99% 44%/59%	360/366 159/359	AAR07793.1 ZP_00514781.1
ORF7	9339-7957	460 aa	cysteinyI-RNA cysteinyI-tRNA synthetase, class Ia	<i>Erwinia carotovora</i> <i>Shewanella</i> sp. PV-4	60%/ 74% 60%/ 74%	277/460 278/459	YP_051241.1 YP_001094636
ORF8	9873-9421	150 aa	hypothetical protein LV192 hypothetical protein LV119 transcriptional regulator, Fur family	<i>Klebsiella pneumoniae</i> <i>Klebsiella pneumoniae</i> <i>Myxococcus xanthus</i> DK 1622	98%/ 99% 65%/ 67% 31%/ 48%	147/150 92/141 45/145	AAR07794.1 AAR07725.1 YP_635082.1
ORF9	10369-11169	266 aa	hypothetical protein LV120 hypothetical protein LV191 hypothetical. protein	<i>Klebsiella pneumoniae</i> <i>Klebsiella pneumoniae</i> <i>E. coli</i> Nissle 1917	98%/ 98% 96%/ 97% 65%/ 78%	262/266 257/266 173/265	AAR07726.1 AAR07795.1 CAE55732.1
ORF10	11240-11662	140 aa	hypothetical protein LV190 transcriptional regulators, TraR/DksA family	<i>Klebsiella pneumoniae</i> <i>Serratia proteamaculans</i> 568	93%/ 97% 57%/73%	131/140 80/140	NP_943446.1 ZP_01537690.1
ORF11	11726-12661	311 aa	hypothetical protein LV189 protein of unknown function DUF198	<i>Klebsiella pneumoniae</i> <i>Serratia proteamaculans</i> 568	99%/ 99% 58%/ 71%	308/311 173/296	NP_943447.1 ZP_01534574.1
ORF12	12519-13211	230 aa	putative carbonic anhydrase-related protein carbonic anhydrase-related protein	<i>Klebsiella pneumoniae</i> <i>Serratia proteamaculans</i> 568	98%/ 99% 82%/ 90%	196/198 153/186	NP_943448.1 ZP_01534575.1
ORF13	13278-14207	309 aa	hypothetical protein LV187 [Klebsiella hypothetical protein Ping_2266	<i>Klebsiella pneumoniae</i> <i>Psychromonas ingrahamii</i> 37	97%/ 98% 42%/ 65%	300/309 123/289	NP_943449.1 YP_943606.1
ORF14	15033-14224	269 aa	putative inner membrane protein putative inner membrane protein	<i>Klebsiella pneumoniae</i> <i>Salmonella typhimurium</i> LT2	87%/ 92% 48%/ 69%	227/259 120/246	NP_943450.1 NP_461953.1
ORF15	15876-15049	275 aa	ABC-3 protein ABC-type Mn2+/Zn2+ transport systems, permease components	<i>Enterobacter</i> sp. 638 <i>Escherichia coli</i> F11	71%/ 89% 69%/ 86%	192/267 190/273	YP_001175846 ZP_00725758.1
ORF16	16730-15870	286 aa	ABC-3 protein Iron transport protein	<i>Enterobacter</i> sp. 638 <i>Shigella flexneri</i> 2a str. 301	81%/ 92% 81%/ 91%	229/280 229/280	YP_001175845 NP_707260.1
ORF17	17548-16727	273 aa	Iron transport prot., ATP-binding component Iron transport prot., ATP-binding component	<i>Shigella sonnei</i> Ss046 <i>Shigella flexneri</i> 2a str. 301	73%/ 86% 73%/ 86%	176/240 176/240	YP_310673.1 NP_707261.1
ORF18	18466-17552	304 aa	periplasmic solute binding protein Iron transport protein	<i>Enterobacter</i> sp. 638 <i>Shigella sonnei</i> Ss046	79%/ 87% 75%/ 86%	235/297 231/304	YP_001175843 YP_310672.1
ORF19	18719-20677	652 aa	threonyI-tRNA synthetase threonyI-tRNA synthetase	<i>Burkholderia pseudomallei</i> 305 <i>Pseudomonas fluorescens</i> Pf-5	61%/ 75% 56%/ 73%	390/639 355/623	ZP_01768717.1 YP_263242.1
ORF20	20547-21521	324 aa	hypothetical protein LV183 hypothetical protein Sden_2526	<i>Klebsiella pneumoniae</i> <i>Shewanella denitrificans</i> OS217	84%/ 87% 45%/ 64%	245/291 118/258	NP_943453.1 YP_563529.1
ORF21	21518-22864	448 aa	PyrC dihydroorotase	<i>Klebsiella pneumoniae</i> <i>Aeromonas hydrophila</i> subsp. <i>hydrophila</i> ATCC 7966	94%/ 96% 58%/ 72%	423/448 262/449	NP_943454.1 YP_857112.1
ORF22	23090-23947	285 aa	putative porphobilinogen synthase delta-aminolevulinic acid dehydratase	<i>Klebsiella pneumoniae</i> <i>Erwinia carotovora</i> subsp. <i>atroseptica</i> SCRI1043	93%/ 98% 67%/ 82%	266/285 191/282	NP_943455.1 YP_048330.1

Tabelle 2: Annotierte DNA-Sequenz der Region X (RegX_24196_bp.seq EU681267)

```

LOCUS      RegX_24196_bp.seq 24196 bp DNA linear 13-FEB-2008
DEFINITION Escherichia coli.
ACCESSION  RegX_24196_bp.seq EU681267
SOURCE     Escherichia coli
   ORGANISM Escherichia coli
            Unclassified.
REFERENCE  1 (bases 1 to 24196)
   AUTHORS Schubert,S. and Preisach,B.
   TITLE   Characterization of a novel DNA region in E. coli acquired by
            horizontal gene transfer
   JOURNAL Unpublished
REFERENCE  2 (bases 1 to 24196)
   AUTHORS Schubert,S. and Preisach,B.
   TITLE   Direct Submission
   JOURNAL Submitted (13-FEB-2008) Bacteriology, Max von Pettenkofer-
            Institut,
            Marchioninistr. 17, Munich 81377, Germany
FEATURES   Location/Qualifiers
   source   1..24196
            /organism="Escherichia coli"
            /mol_type="genomic DNA"
   gene     <1..331
            /locus_tag="ORF0"
            /note="putative iron(III) ABC transporter, ATP-binding
            protein"
   gene     496..1308
            /locus_tag="ORF1"
            /note="putative iron(III) ABC transporter,
            periplasmic-binding protein"
   gene     1308..2327
            /locus_tag="ORF2"
            /note="putative iron(III) ABC transporter, permease
            protein"
   gene     complement(2533..3036)
            /locus_tag="ORF3"
            /note="putative regulator"
   gene     3605..4828
            /locus_tag="ORF4"
            /note="putative cobalamin synthesis protein/P47K"
   gene     complement(5095..6744)
            /locus_tag="ORF5"
            /note="putative ferrous ion transport protein"
   gene     complement(6756..7886)
            /locus_tag="ORF6"
   gene     complement(7957..9339)
            /locus_tag="ORF7"
            /note="putative cysteinyl-tRNA synthetase, class Ia"
   gene     complement(9421..9873)
            /locus_tag="ORF8"
            /note="putative transcriptional regulator, Fur family"
   gene     10369..11169
            /locus_tag="ORF9"
   gene     11240..11662
            /locus_tag="ORF10"
            /note="putative transcriptional regulators, TraR/DksA
            family"
   gene     11726..12661
            /locus_tag="ORF11"
   gene     12519..13211

```

```

        /locus_tag="ORF12"
        /note="putative carbonic anhydrase-related protein"
gene    13278..14207
        /locus_tag="ORF13"
gene    complement(14224..15033)
        /locus_tag="ORF14"
        /note="putative inner membrane protein"
gene    complement(15049..15876)
        /locus_tag="ORF15"
        /note="putative ABC-type Mn2+/Zn2+ transport systems,
gene    complement(15870..16730)
        /locus_tag="ORF16"
        /note="putative ABC-3 protein"
gene    complement(16727..17548)
        /locus_tag="ORF17"
        /note="putative iron transport protein, ATP-binding
gene    complement(17552..18446)
        /locus_tag="ORF18"
        /note="putative periplasmic solute binding protein"
gene    18719..20677
        /locus_tag="ORF19"
        /note="putative threonyl-tRNA synthetase"
gene    20547..21521
        /locus_tag="ORF20"
gene    21518..22864
        /locus_tag="ORF21"
        /note="putative dihydroorotase"
gene    23090..23947
        /locus_tag="ORF22"
        /note="putative porphobilinogen synthase"
BASE COUNT    6299 a    5549 c    6048 g    6300 t
ORIGIN
1    ttggctcctc tgactggcac agacgcccga cctggatttg cttgatgaac ccaccaatca
61   tctcgatcct cccgggocgog aggaactact ttcccttgta aaaaccaagg gaattaccgt
121  cgtggcggta ttgcacgata tgacactgac agaatcattt gctgaccgtg ttctcctctt
181  gtcgcagggg caatcagtgga tctgtgacac accggagagg gtattggtat cagaatacct
241  ttatcccatc tttggcttga ccagttttac cgtaccgcat cccatacacg gaaagcgctt
301  gcgtatTTTT gaagttccgc attgtgcata aattactttc tgagagtcaG tgatgaaaaa
361  ttgtttatat ccctgctggc ggtactgtct gcctcagtag ccagtgctgc tgattttcct
421  gttactattg aaagttgtgg gactcctgtc acattcgccg gtccgccgaa aagggctggt
481  atcaacgacc ttaatatgtc agaaatggcg tttgactgac atttacagga tcgtattggt
541  gggctgactg gtattagtgg ctggtacaaa atgacgccgg aatttaaaaa ggcgatgggg
601  tcgatcccag aactggcacc caaatatccc actctggaaa cattgctggc agccgaacct
661  gatttctttt tcgctggctg gaactatggg atgaaggtgg gtggcgaagt gacgccgat
721  acgctaagca aatatggcat caagactttt gttctgagtG aaagctgtgt gttcaccact
781  gcgcacaaaa ataaagccac catggatctg ctgtataacg atgtcctcac gctgggtaag
841  atttttggca agcgcaatga cgcacagtcc ctggtcagcg gctggaaaaa gaggttgaat
901  gaactgccga agccagcggc aggtaccctg cttttaaagg tgttcgtcta cgactcaggc
961  gaggataaac cgtttaccag tggtaaatat gcgatgccta cgccataat tgaggcggct
1021 ggtggtaaaa atgtaatgga ggcgggtggat accagttggg gaacgacttc ctgggaaagc
1081 gtggccgcta ctgagccgga cttcattatt ttgttggatt accagacggg cagcgggtgcg
1141 gatgcactgc gtcattttct tgagaatcat ccgttaatga aactgactcc ggctgttcag
1201 catcatcgat acctgaaatt acaatatgct gaactgacac cggggcctgc caacgttgat
1261 gcgggtgaaa aactcgcgog ggcgatgtac ccgtcaacgg cgaagtgatG ctcatgtcgt
1321 accgttatca ctggggcgatg tggggagcca tggtagtGct gattgggttg atgtattca
1381 gtattaccag agggacagtt tcaactgtcg taatgcaggt gctgggagca ctgaagctgg
1441 tggatgagcc ggtatctgcc atggtcagca gaattgtgac ggatttacgc gtaccgagga
1501 tgttgctctc tgtgctgaca ggggcgggac tggcaatggg aggggcgttg ttacaaacta
1561 ccaccgtaaa cgatctggcc gatccttttt tgttcgggct atcctccggc gcgtcagcgg
1621 gggcgggtgct ggtgatcacc cgttttggtg atcgacttgg tgtattaacg ctccccgtat

```

1681 cggcggttgt tgggtggtgtg tgttccgcca tagcagtaat gctgttgttt cattttaaaa
1741 agcaacgcgg agctgagcat ctaattctct gcggtcttgc catctctttt ttgttcgggg
1801 cgctgaccag ttatttcatt ttttcggggc atcagcgagc tgccagctct gtgctgttct
1861 ggtcactggg tggctctggga ttagccacct ggaataattt gcctttcgca gtgttcagcc
1921 tgggtgttgc gtttgctttt gtctgtttac ggtggcgctc tctggatggt gtgctggcag
1981 gagaacagac agcgttgtcg ctggggatta acgtgagtcg tctgcgcacg gagatttttc
2041 tttgttgccg actggcaaca tcgttgatgg tggcattaac cggcggtgatt ggatttgttg
2101 gcttaatggt gcctcatctg tgtcggcatt ttgcaggagt taagcatctg ttgttgttac
2161 ctttgtgtgg cgtctggggg gcaatattac tgtgtggcgg cgatatagtc agccggacga
2221 tactggcacc gcaggaacta ccaattggca ttattacggc gggcattgga gggttgttta
2281 ttatcatggt actcgcaaga aattcgtcca gaaatacggc ctgttaagca catataagac
2341 agaaaaataa ggctctgatg tcttttattc attctgaagg ttgaggttgt tgacgccaat
2401 tatgaaagca accgacaatc cttgttgacg gttatatcta ttatagagtc aggtaatagt
2461 gtgatggctt ttagatattt taaaatttat ttatcgctcg actcgatttc aatgattatt
2521 cgtattggtg atttaaccct ttctcgcaca atcagggcaa attactttaa cttcaatatic
2581 ttttttttgt acagcacaac cataatctat tgcttttttc tcaatgtaag tatgaatgag
2641 gtggtcagta tgttggatta ttgttccgca gtttgagcaa ataaaaaaaa tagtcacaac
2701 atcagaggat aaattatttt ttttgataaa tttgctacag cattctattt ttacgataat
2761 ccctgcctgt tgcaaataat ccagagaacg ataaaccgta gccggttttg cattcggatt
2821 atatttcttc attaattcca gtatctggtg tgcgctaatt cctttttctt tagcatgact
2881 gacgaaatta taaactctct cccggagcgg tgtgattcgt attttcttac tctcaatcag
2941 ctctctccgc ccggaacaat cctcagacga taatgcagtg tccctgttac ttagagaata
3001 tgggctatta ctcaccgggc ctctgtaaat tttcaagcga tgttatcttt tcttgcctg
3061 aaacgatgga aatctactga tgttttgatc gtgccatgta aggaaagaag aactggcaaa
3121 ttttattggt acgttataac ataacacatg ataagtgagg tgccaatggt tctcttactg
3181 gaaaaagccc atgcaggtgc agtttttaaa cttgaagata tcctggcacc aataccctgg
3241 gattcccatg gattgatcgc tgcgattgca cagcaatatg aactgggaga ggtgttgatg
3301 cttgcctgga tgaatcaaca ggctcttcga tgaaatctta ttgactggca gagcctgta
3361 ctggctcgcga tcccgatcat gcctttggcg aaaagggtgag acctccggct gcgttcagca
3421 agtacatgat atccgtctgg attgtgacgg agatgcggta ctactgcttg tggatcaaaa
3481 aggtggagcc tgccacacag gtcggcggag ttgtttttac aacgccatta aaggcgatga
3541 gctaaccgtg ctcaacgacg ctgggttaatc cgaacacaac gaatatctga ggagatgtaa
3601 ggtaatgtcc gaagtcagtg caaattttca tggcgataaa cgctgcccgg ttaccgtact
3661 ttcgggattt ttgggagcag ggaaaacgac actgcttaac cataattctga acaaccggga
3721 aggccgtcgt gttgctgtca ttgttaatga tatgtctgag gtgaatattg atgctgccct
3781 tgtgagggaa ggtggtgcag aactctcccg gactgatgaa aaactggtag aaatgagcaa
3841 cggttgcatt tgctgcacat tgcgcgaaga tcttcttctc gaagttaatc gtctggcaaa
3901 agaagggcgt tttgatcagc ttgtcattga gtccaccggg atattctgagc ctctgctgt
3961 tttgaaaca tttaccttg cgggcgatga cggtgaaagc ctttccgagc ttgcccgcct
4021 tgacaccatg gtcactgtgg tggatgcgta caattttctg aaagactacg gttcggaaaga
4081 cagcattcag tcacgcggag aatcgttagg tgaaggatgat gaaaggagcg ttgttgactt
4141 gctgatcgat cagattgaa tctgtgatgt gctggtcctt aataaagtgt acctatcag
4201 tgaagcga aaagaaaaac tgatggcgat tttacgttcg cttaatccac gggccagaat
4261 cgttgtttca cagtttggac aggttccgct gggcaacata ctttaataccg ggctgtttga
4321 tttcagacag gccgctcagg caccggctg gcttaagaa ctccggggcg aacatacggc
4381 ggagacagaa gaatatggga tcaccagctt tgtctttcgc gcccgccggc cttttcatcc
4441 gaccctgttc tggcaggtga tggaaaatga gctggatgga gtcgtcagat caaaaggtta
4501 cttctggctg gccagccgct cggaatcgc aggttcatgg tcgcaggcag gtggcattgc
4561 gcgccagggg ctggggggca tgtggtgggc tagtgtgccg aaaggacggt ggccagagga
4621 tgctgagtca ctgaagtcca tcatgtctaa ctggatagat ggcatgggtg atgccctca
4681 ggaactcgtt tttattggaa tggatatgaa tgaatcagaa ctgcgcaaca gactggattc
4741 cgccctgtta acggatgcag aaatggcaga agggcctcaa aagtggcggc attattccga
4801 tcctgttgag ccctggttcg aagaatgaa cgaaggcacc cgcatttgtc ggggtgccttc
4861 gtatgaaagc gaaaaggggg cagatttaca acaattcttc ccagctttt ttgaaccact
4921 gtacaacact attcagatgt ggttgcccat atagcgagac aaaacgtctg gtttctgat
4981 gatcttggga ctcaaggact gccagagcgg atgcaataaa tgcctgcgct aagggaattc
5041 cacattcttc tgtgatgcac cggatcacc tctggtaagt cttgggagtc atacttatgc
5101 ctcatgaata aataactggg atatcaccag agccaccacc agagaactta atacctgtt
5161 tcccgcact gctgctgcaa accgccagtt tatttcccg gcgatggtaa ataccgtgac
5221 aaggcagggc atcaacgtag aggcgagcca tacaagtagc agtaactgag atgtgctgag
5281 agactgtatt aagctgccc catcctgatt gagcaccatc aaaccatctt tccttagcag
5341 tgaaaaaatg atcccgggca tcagttctgc tggcagttta aaaagatgaa gtaaaaggtgc

5401	cgtcgtttca	ctcagccagc	gagtgatccc	tgcgtaatca	agcatccccg	caacaatgca
5461	gataatcagg	aataacggca	tggcttgcgt	aatgaactga	cggaggatat	ttttcagcat
5521	ccatgtcaca	tttcgccagc	gtggccattg	tagccagggt	agctctgtca	gacgctggtt
5581	ttggccaggc	ttcaacgacc	agttccagag	gcgcgatgt	attgocgctg	tgaaaaacag
5641	cagactgaga	taggggataa	agagccatgg	ttgatgggcg	gcattaaaaa	gagaaagcgt
5701	cgcccccgtc	tgatagctgc	aagcggaccc	aaagctaata	atacttatac	aagcatggcg
5761	ggtgcatcgc	gaacacgaac	ggctctgaaa	aaccgccacg	acattacagc	cgaaaccgct
5821	taatacagga	ataaggtcct	gaccgctgag	gccgagtttt	ctcaaccagg	gatccagcgt
5881	cgctgtaatt	ctttccttaa	gtccactgtc	gtcagtcagc	gaaagagaca	accctatcag
5941	taccaccaca	ggaaatgccc	atacaaacga	ataaagtccc	agactaaaaa	gaccgtaatt
6001	gccgacaaat	agtgtcttca	gaacgtcggg	ccagctcgtc	gttatattct	taagcggctg
6061	aataaccgca	ttgtcaatga	caggctgaat	aaaatctgac	agcagccatg	ccccccatac
6121	cggaacggca	aacagcagga	aaaggcatac	tattgcagca	acttttccac	accaacgatg
6181	ttcaaatatc	gtttgttgag	gaaattcagc	cagcagatta	attaccggta	tcttttgacg
6241	gggtaaagag	acagcggagg	gtgcaggttt	atcctgcaac	agcaactgca	ctattccctg
6301	gcgaacattg	ctgccagctt	cccgggcatt	gactgccagt	accggcgcac	ctgatgtttc
6361	actcaggtat	tcagtcaact	ccagaaggcc	ctttctgatc	ttatcagcaa	aggtcaggag
6421	tataacaacg	tgttttgctt	gtggcgggat	gaggtgacag	agcgtttgcc	actcctgctg
6481	agcgtgctg	gcccgcagaa	ccaccagaat	gccatcatgc	tgatttagcg	catcaagggc
6541	gagttttgtg	gtttcactgt	cgctctgaaa	acgtatgcct	ggcgtatcaa	ccaggcagat
6601	ccctgcatca	ttaatgagac	cttcccagca	tacaacggtc	gatcccctga	aattacgctc
6661	atcttgcgca	gcactttccg	tcaggatatt	aaacagggta	gattttcctg	cactttctaa
6721	tcccattaaa	acgaatgatt	gcataaacac	ggacctcagc	acgcacagtc	ctcatcggca
6781	cagcagtcag	gctctgaagg	cagaaccac	cagtcccgtg	ctccttcaga	ggacaggccc
6841	agatgacgcg	tgatactttc	tgccacgacg	ggaaagcgtt	gacggaaact	atagataaag
6901	caataaatat	taccggcatg	ccgcacatga	gggccgacaa	ggtacaaaacc	aggaaaaagg
6961	gttgagcagt	cttcgtccgt	taagaccgga	tagccatcat	cattccattc	aaaaaaagca
7021	gccagttgcc	ttgctcccc	gccacactgg	aagccggtag	caagcacagg	cacctcgtca
7081	gtagcccagg	cgcgctccgt	tgtggtatgg	actttgtacg	aacttcccgg	gccgggcatt
7141	ctgattactt	cacagacatc	cgcacttcca	taaatttcga	aaagacggtg	attcatcact
7201	ctgttcaaac	gctcacgggt	ataaggggaa	agcgatagac	tgggatcgct	gatgtggtta
7261	gttgaccagg	gggctgagcg	ggttaacatt	tttaccgatg	agccattttc	cagcaggttt
7321	acagcggcat	ccaccgcgct	ttcatagccg	cctataacaa	tgtattcacc	tttctgaaa
7381	tccttccagc	tggtcacatc	accgtaatgg	caacaaatat	ctgctccagg	aaaaatctgc
7441	cggctctggga	actggaattc	gcctgtagcc	catattaaag	agcgtgtttc	cagcttctct
7501	cctgcttctg	ttgtcagtat	aaaattaccg	gaactgagca	acgcaacctt	tgcaatcctg
7561	gcgggagcca	tgacgggtat	ctgatactca	tccagtacga	ccttaaggta	actggcataa
7621	gttttcccc	caggatgttc	ttcaccgcag	aaaagcgcca	atgaactgtc	aggggtcaca
7681	gagttcaggt	ctggttgtcc	aaacggattg	gaaaaaaaag	aaggagtaat	aaaacggggt
7741	tcttccggcc	agcgcataaa	ggaggcacct	acttcgtgac	tgtcaacgac	caggatatcc
7801	tgtatggcac	tacgacgaag	aagcgcgccc	ataccccacc	cggccggacc	cgcaccgaca
7861	ataacgactt	cgcggtggtg	gttcataaaa	ttaccctttg	cattacgtta	taacataaca
7921	attgatattg	aaatTTTTTt	catccagagg	gacaacttat	cggatacggc	atgaggtatt
7981	ttcgccgaca	tcttcaagat	caatgccttg	ctccgctaat	tcatttctta	aacgatcggc
8041	cttgtcccat	ttgcccatac	gtcgtgcctc	actgcttggg	tgaattaact	tttcaataag
8101	ttgcagttct	gcaggatgaa	gatcactcgc	gtttttcaaa	aaggcatccg	ctgattgttc
8161	aagcaaacc	agcacattcc	ccaggctctt	cagcgtgggt	gcaagtgctg	cggcttcatc
8221	atcattcggt	ccgcaaagcc	tgtaaatctc	tcttgtcagt	tcaaataaaa	tggcaagcgc
8281	ctggggtgta	ttgaaatcat	catccattgc	ctcacagaat	ctggcccggc	atgcatcgtg
8341	ctgttcagat	cgagaccctg	cagcgaactg	gcatgcgggt	tagagccgtg	tcagcagggc
8401	gcgcgcgagg	tgaatattct	cctcgtctata	attcagcggc	ttacgataat	gtgcgcttag
8461	caggaagaag	cgcacactct	gggcatcgaa	ccgattcaga	atatcgcgca	gggtaaaaaa
8521	gttattcagc	gatttcgaca	tcttctcttt	attcaccgtc	accatgccgg	tgtgaagcca
8581	gtaatttacg	taagggccat	cattggcggg	ggtggattgg	gcaatctcat	tttcatgatg
8641	gggaaatata	aggtcagaac	caccgccatg	aatgctgaaa	tgggtaccaa	ggtaatggca
8701	gttcattgct	gaacattcaa	tatgccagcc	aggtctccct	tcaccccagg	gtgatggcca
8761	gacgggctcg	ccgggttttg	cccgttcca	gagaacaaaa	tccagcggat	cgcgtttata
8821	ctcattcact	tcgatgcggg	tcccggcaag	aagcatggta	atatacctgtc	ccgaaagctt
8881	tccatagcct	tttgccgatg	cagtggagaa	catgacgtca	ccattctctg	caatgtaagc
8941	aaattgtcga	tcgagaagtc	gttcgatcat	actgatgatt	tgtgggatgt	gctctgtggg
9001	acgaggttcg	atgtttggcg	gcaggatgcc	taactttgca	aaatcaatgt	gcatttcatg
9061	caccatccgc	tctgtcaggg	tgcgccagtc	ctctcccgtg	tcacggggcg	gctgaataat

9121	tttatcatct	atatcagtaa	tgttacgcac	gaaagtcaca	tcgatatccgc	gaaaacgcag
9181	atagcggaca	atcatatcaa	aagctacaaa	ggttctggcg	tgcccaagat	ggcaaaggtc
9241	ataaacccgtc	acgccgcata	catacatgcc	gattttgctc	tctttaaccg	gaactaatag
9301	ctccttttgc	ataccaagcg	tgtttggtag	tttaatcatt	acgtttacct	cattatTTTT
9361	agatagcagc	acacgaacac	gcctggcgct	aattaacagt	cgtccggaat	aagaaaaaaaa
9421	ttatTTGGTG	atagTACGAT	TACAGTCCGG	ACAGATACCA	TGACTGGAAA	GGGTAATTTT
9481	ccgttccggtc	ataccatgtc	cagtGAGATG	TGATTTTACC	AATTCAATGG	CCGCTTCCAG
9541	ggctggatcc	TGCTCAATAG	GAATAACATG	ATGGCAGTTA	TCGCATTCAA	AAAGTCTTAC
9601	ggcagcccgt	TGGGTcagag	CGAAACGCTG	AGTTCTTGCC	GAGTCGGTAT	AACGCTGGAC
9661	aatactgacc	TGCATTAGCC	GTTCCAGTAT	TCGGTAAAGG	GTCAcACGAT	CGGGAGCCTG
9721	ctcgcCCAGT	TGCTGAGTTA	AATCGTcATG	GTCcATGGCC	TGGCTGGCAT	GTTCAAGGCA
9781	ctgcagTgta	GCAAGGGTGG	ATGGCGTTGC	CCGCAAACCA	TGTGTGCGCA	ACGTTTGTCT
9841	ccagTcatac	GTTTTGGCTA	CTGGTGTGTG	CATATTTCCC	CCTCCGAGTC	ATGTcATTtG
9901	aatTTTGGCA	GGAAGCCGGT	ATTATTGCAA	CTTGATTGCA	TTAACGTTTA	GCGGCATAAG
9961	GCTGGATGCT	ATGAAAAGAA	CACCGCTCAT	TATCCTTAAT	GGTTTTCTGG	GGGcAGGGAA
10021	aactactctg	ctaaaaaacc	ttctTAcaca	ggcAcAcAAA	cggcGAATGA	ctgTATCCGT
10081	catcgTTAAT	GACATGAGTG	AaCTTgATGT	ggatGGCGTC	ctgATAGCAA	AcAcAGAAAT
10141	cgTtGATGCA	GCCAGTAACA	ATTTTGTAGT	ATTTcAGCAG	AcAGTATCAG	CAGCAGAAAG
10201	gggatccaga	agctggattc	cgCGCTGAAA	aacctgctcg	aaaaacggtc	TGCGGATTTT
10261	atTTTACTGG	aaacctcagG	TAGCAGCCAT	cctcttcccc	TGGTCCGCTA	TCTGCGAGAA
10321	cataccagG	taagtctcaa	agcattTTTTG	TCTCTTGTG	ATACGCTCAT	GTTAAATGAT
10381	GactacGagc	GTGGAAAGAA	GCTTATTCCG	GTTTTCCAGG	AGCACCTCAA	TAAAGTAACC
10441	cggggagTgg	aaagcctgct	ggcagaacag	ATTATGTTCT	GTAATAAGCT	CTTATTAACC
10501	aaaaacgacc	ggTtaccctt	ctatGTGGTT	actGAGGTAG	ccagagccat	AcATCCcCTT
10561	aatccccagG	TtGCAATTAT	ggcCGTgccc	TGGGGAAATT	TACAGCTGGA	TGAATTGCTT
10621	tccatgcccG	attacgattt	TcATCGGGTG	GCGTTACTTA	TcGATGAGCT	TcAGGATGCC
10681	attgacGcag	TTTTGCCGGA	TGAAACCGAT	AAAAAATACG	ATATTTCTGT	GAGGGTCATT
10741	gaggatgacc	gtccatttca	TCCGcAGCGG	CTGTGGGATA	CCTGCCATTG	CTTcATGGGC
10801	gcaggcgTtt	accgcagcaa	ggggTTTTTc	TGGCTACC GG	GAAGGGATGA	CCTGGCATAA
10861	TtGTGGAACC	aggcagcagG	GAGcATCAAT	CTCGAATTCA	TAAGCTACTG	GAAAGCCGGG
10921	gtcctgactc	acactgataa	TtCATtGACA	AAGGAAGAAC	GCGCAGCTAT	CCGTcAGCAA
10981	ctggcaaaaa	TGCCTGGCAG	ATTTGGTGAC	CGGCGCTGCC	GCCTGACTGT	TATTGGCGAG
11041	agcggcgaaa	TtGACGATTT	TGCCCTCGCA	TtGCGCCAGT	GTTTACTCAC	TGAAGAGGAA
11101	attctctggT	ggcagcaggg	ggggatTTTc	cacgatccct	ggccAcGAA	AGTGGCCCGG
11161	ctggcttaac	aggTtATTTT	GcAGCGGcAG	TtAATGTTAT	GTTATAACGT	TtCATTACGC
11221	aatTTTATCG	GTACAGGACA	TGACTAAACC	TACCCGCGCT	GAACGAGAAA	AAATCCTGGC
11281	gatctgcca	gaagattaca	TGAATCCTGT	GCAACAGGAC	TACTTCCGGC	GGCTTATCCG
11341	cccagaaagG	caggaggtgc	TtGATcATAT	agaccagatg	aagcaacaac	TGCAGGATAA
11401	Tcagaaaagc	ggTgacgaag	GTGACATCGC	CATTCTGTA	GAGCAGTTAC	TCTGTCTGTT
11461	cagacagata	GACAGAGAAA	GCCGCTGTGT	GCCGAAATAT	GATGcAGCGC	TtATGCGGCT
11521	Tgaaaaagga	gaatatggct	TctGccggga	TACAGGCGAA	CCTATCGGTC	TGCAACGCCT
11581	GtactcagG	cccaccgag	AATTAAGTAT	TGAAGCCAAA	ACAAAGGAGG	AAAAAGTTGA
11641	GattcagTAc	cGAAAGAAAT	AAAGCGGTGG	CTCAGGcAGG	GGGcAGGGTA	GGGATGAGAG
11701	Taaatatcag	ATCAGGcAGG	TATATATGAG	TATTGcAGCC	AAACAAACCT	TATGCCCCCGG
11761	AatgctaccG	GATATTcAGT	CCACTAAGGG	TGACGGTgAG	GGCgAGTcAT	TATCCTGGGT
11821	Tggcatggaa	CAAATTGACC	TGCCGATAGA	CATCGCAGGC	AGGCCGGTAT	CAGCAAAGGT
11881	caatgctggc	ATAAACCTTc	TcAGTtCACc	TGAAGCTGAA	AAGGGGATCC	ATATGTCTCG
11941	Tctttatctt	CTTCTGGACG	AaCTGAcACA	AGGGGAAATC	ACACCgGCAT	TATTGCAACA
12001	Tgtgttgaaa	GcATTTCTGG	TATCTCATCA	GGGACGTAGT	GATGAAGCCA	GtATTGAGAT
12061	Ttcaggggat	TtGTTACTGT	CgCGTAAGTc	GTTAAATTCT	AATcACTCCG	GCTGGAAAGC
12121	Ttatectctg	actctgagcg	ccGAACTcAG	GCAATCTTTT	ACGGTTACCC	TCAAAGTGGG
12181	Tatcccttat	TcGTCCACCT	GCCCGGCTTc	CgCTGcACTT	AGCCGCCACG	TcGCGGGTTT
12241	GcagTttagc	aaagactttg	GTAACAGGAT	CGACAGGCTc	CCGGCAGCAG	AAATAGCCGA
12301	ctggcttgtt	GAGAAAGGAA	TGCCGGCAAC	CCCGCACAGc	CAGCGAAGCT	GGGcATGGGT
12361	gagTatccgt	CTTAATCTGT	AAGCAAAATc	ACTGCCTGTG	ATTGAACTGA	TtGATTATGc
12421	TgaagTtggc	ctcgggacag	CCGTGCAAAC	GgTtGTCAAG	CGGAGCGAGC	AACAGCTTTT
12481	TgcggTggct	AACGGTcAGA	ATCTGATGTT	TtGCGAAGAT	GCTGCCCGTC	GAGCTCAATAA
12541	CgTtTTTcGT	TGTGACCOCT	TtTGTGAAGc	ATTTGATATc	AGGGTtGAGc	ATCAGGAAAG
12601	cctgcAtcct	cATAATGCCG	TtGCCCGTAT	TcACTGGAAA	GGAAGTAAAA	ATGTTACGTA
12661	aaaaccatc	AGGTcATAACC	CCAGTGGTGT	CAACAAAGGC	CTATATCGAT	CCCACCgCGG
12721	TaattTgtgg	acgtgtcatt	attcATGATT	atgTtTACGT	cggaccttat	GCGGTTATCC
12781	gggctgacga	GCTAAATGCA	GACGGGGATA	TGGATCCTAT	CATTATCCAT	TCCCACTCTA

12841 atattcagga cggggtggtt attcactcaa aaagcggcgc gccagtgaca ataggaagcg
12901 gaacgtctat tgcccaccgt gcaatagtgc atggcccctg tcaggttgac gagcgggtat
12961 ttattggatt caacagcgtg ttgtttaact gtcataattca gactggttgc gttattcgcct
13021 ataattgcagt tgtggatggc gttaccttgc ctgagaatac ctatattcca tcaacggaac
13081 gcgtggggcc tgattccgat cttaaaagtt attcgagggg agaccccgtc tctttacaat
13141 tttcagaaga agtggccagt accaacgtta aactggtcga aggttatcag aagttacgca
13201 atgaattctg agtatggctc ggctagtgcc gttaacgtaa caccgtatta ttaacgattt
13261 gttattaatg tagcaaaatg atacttgctc cttctttcct tttgcggaag aggcttttct
13321 gttttattaa agaggtagcg gtgaaacttt ttattaatga tctgacggtc atggatttta
13381 gtttccttga tgcggaatcc gggttgattg gcgatagcct tattggtgat attattcctg
13441 agggagatct gaatgctgag tcaatggtaa tggatttctc acatgctaaa aaaagtatca
13501 aacatgaaat tgataaactg gccgatcatg tcttgattgt gccggaacaa aatagtcata
13561 ttattgtag tagcagcggc actacgacgg aagtcgcat gctgcgtaag aatggggaga
13621 cccagtgttt tatttccggg ccgcaggaga gtttctggct ggtccaaaca gataatatta
13681 attcccgttg ccttgaatcc caaatagaaa aacatttact ggctgtttg ccgcaaggag
13741 taaaagacat aacgataacg ttacggccgg aaagtattaa tggggacagt tatcactatt
13801 cgcattgact gaaaaaacac cgcggaaatt gccagagaat agcgcattgg catcgctctg
13861 cgatcagaat ttttgtggat ggtgagcgaa gccatatgtg ggaacagaag tgggcaacgc
13921 gatggaataa cgcttatcta ttaagccggg aagatgtcgt tacagtgaca acactttcac
13981 cgcgtgccgt tgcatactgg cataaggggc tgacatgttc ttcttggcgg agttctcagg
14041 gatactttga aatcatgctt tgttctgag tagtagatat tctgcccgtg gataccagcg
14101 ttgagtcact ggcttttatt atcaggcagt caattgagca cggtttacc cggcggaaaa
14161 tagaagtcca tgcctttgag ggagtgggga aaggcgtat tgcctgacac actattctac
14221 ttttcagggg tattgatctg ttttgtctcg taaaactaca ataccgggtt ttgtagccat
14281 gtatctaata aaaggtgagt caactacctt tctatttttc gctaattgacg atgcatttct
14341 gaaccagatg ttattgtggg ctttgataat gatcccaaca tgtgaaacat ccagacctct
14401 ttttgttgaa taaataccca cataatcacc tgtttgcagt ttatttatta cctctttatc
14461 tattgcagaa gccggtatgt aacttatccg tctgtgagata atgccgagac ccttaacata
14521 ttcttgtttt ttattctttt ggttttagctg tttgttgaca gtaaatgat acggactgat
14581 gtcctgagta acatcttgtg catttcgagg agatatgacc gccagtcgg tgaaaaaatg
14641 ttttctgcca gtaaaatcaa tcattccggt tgtataacga gtcttaacaa gttcagatac
14701 ataagtttgc cagctactgg cgtgcgccat ggagcgaaca taatcaagaa aagtaaaaca
14761 atccacggta ctgggattag tgaccagcct ctcttgcac gaagatgatc caatcagtgt
14821 atgcccctgg taaggcgttc cgaggaaacg ggaggatata tcagcaacgc gttcattata
14881 gctgaggcca tttctgtaca acgaagtggg atgaaggatt ttgttaactt tgcctgact
14941 ttcttcaata gaatagtcgc ttctgtcctg ctgaggctca gcgaatcca cgcagctata
15001 aaatgaaagg agaaagataa ttgctttttt cataaaatat caacatcctc agagataatt
15061 ctattatat ttaaagcttg agtatataaa aataactata aatatcagg caaaaagcaa
15121 aacaaccggt ggcgagggcg cactatcaat gaaataagag aggtaaacgc cgtcgttga
15181 taccagagcc gataaagcca cagcgatgaa taatgcgttg ctaaatcgtt cgtcctaaaag
15241 aactgcaata gcaccgggcg cgataagcaa agagacggat aaaataatcc ccacggactt
15301 tagcgtggcg acaacagtaa gcgaaatcat gctcagtaat ccgtaataga gccatcccgt
15361 cgtagttcca ctgaccctgg cctgctgaat atcaaaacag taaagcagga aatcacgcca
15421 ttttaattccg gtaatcagca gtataaataa ggcaatggtg cccgtctgga gtatgtcggg
15481 ggcgtttata ccaagcatgt cgccaaagag gatatgttca agcggaaacat ccggacggac
15541 tgagacatac aaaatcagtc cggagggcga catacccgaa aaaatgatac ccagcagcgt
15601 atcctgcttt atgcccgtgt tatctttaat atatccggta agtaaagcac aggataatcc
15661 agccacaaaa gctccggtgc ctaatggcaa ccctacaatc caggccagta caaccccggg
15721 aaacacggca tggctcatgg catcccccat cagagcccaa cctttcagca cgagaaaaaac
15781 agacagcagg gcgcagggaa cagagaccag cagcgcgatt atcagggcat aattgatgaa
15841 gtcaaagcgg aaaggttcaa tgagcatatt cagcatttac gcggtccttt tgactgagaa
15901 accgatcggc ggttcgcccag taagccatgt accggggcga aaatgaacac cagcaagaaa
15961 atgagggctc gagcgaccac gataatgcca cccgttacgc catcaaggta ataacttgcc
16021 cacgcaccca gaaagctggt gacactgctc atggtcaccg caatacagac caggcggggg
16081 aacgcctcag taagcagcca ggccgttgcg ccaggcgtta ccacaaggca gatcacatg
16141 aagcgccta cggctgttaa tgccgaaacg gtggagaggg ttaacagaag gaaaaataat
16201 atcttaagcc tttccggatg taatcctacc gcccgcgcat gattttcatc gaaaaaggtc
16261 agtaataaat ctctccattt taataacaac gtaataaccg aaataatgcc aatgacggta
16321 agctgaagaa tatcttctgg cggaaaccgc aggatatttc ccagaatgat tgattgtata
16381 tttaccgaaa caggattcag cgagatgata aataatccaa gcgcgaaaaa ggatgaaaag
16441 atcaaaccba tgatgacatc ttcttttaga cgggtacgtt gattaagaag gaacatgctg
16501 cctgctgcca gcccgccaga aacaaacgca cccaccgaga agggcagtc cagcataaaa

16561	gcgccagcta	ctccgggtac	gatagaatgg	gacagcgc	cgccaattag	tgaccaaccc
16621	ttcaacatca	gataacagga	taaaaaagca	catattcctc	cgaccattgc	cgaaatccac
16681	atggcattga	gcatgtattg	ataaccaaaa	ggctccagaa	gccagctcat	gatgacctgc
16741	ctttatcttc	atattcagcg	gaaggttgat	aatttggctg	tgttgatcct	ccctgaaactg
16801	gcaatatatc	tgactcactc	tcaatgtttg	cactaaacgt	cttttttata	ttttccggcg
16861	taaatgtcgt	ggcgacagga	ccgcttgcca	gcacgggtgcc	tttcaccaga	acgggtataat
16921	cacaatatte	ggatacggag	gtaagggtat	gcggtgaaac	caccaggggtg	ctgccttctt
16981	ctttcagttc	gccaaagcagc	gtgatgattt	tcttttcggg	attgacatcc	acaccggtaa
17041	aaggctcatc	aagaaggatt	actttgcccc	gttgtgcaat	agaacgcgcg	aggaacaccc
17101	gcttcttttg	ccctccggat	aattcaccaa	tctgcccgtg	gcggtaaagga	agcatatcta
17161	cacgctcaag	cgcatcatta	acgatcctgc	gatcttcttt	acccgctcta	cgagcatac
17221	ccatatggcc	ataccggccc	atcatgacga	catcttccac	caataccggg	aacgaccagt
17281	caatttcctc	ggactgaggt	acataggcaa	taatattgcg	gcgaagcgcc	tgtgcaggag
17341	gcatcccag	aagtgaaatc	gttccgcgatg	aaggggatac	gaatcccatc	attgccctga
17401	atagtgtgga	ttttccggaa	ccatttacc	ccagcagggc	cgcgttcgaa	cctttagggga
17461	cgctgaatga	tgcatctctg	aatgccttat	ggccggttacg	ataaacaaca	gttaagccccg
17521	ctactgtgat	gcctgcccgc	tggtcatct	tttactcctt	cattccgctg	tgcaaaccct
17581	gtaccactgt	ccgggtagtg	acgttcagca	gatcaaggta	ggtgggaacc	gggccatttg
17641	ctgcgctcag	ggaatccaca	taaagtacgc	cccataatg	tgacactgtt	tgcggtgcga
17701	cctgacgggc	tggtttattg	gaaacggtag	tttcgctaaa	taccgcccga	attttattct
17761	gccgcacttt	atcgattacg	ttacgaacct	ggtgcggggg	tcctctgtga	tccggcattaa
17821	tcggccatag	atagagttct	ttcagtccca	gatcgcgctg	cagataagaa	aatgcccctt
17881	cgctcgatac	cagccagcgc	tgttcggcag	ggaggggttc	tatttgttta	cgagcgggtt
17941	caagcgtctg	acggatcttt	tgtttatagg	cgctccgcat	ggctttgtaa	gtttcagcat
18001	ttccctgatac	gtatttcacc	agcgcatac	gaatattatc	tacataaatc	agtgcattat
18061	caggcgacat	ccaggcatga	ggatttggtt	ttccatcata	cggaccctcg	gtaatgctta
18121	ttggagtgat	acctgtagtg	accaccacat	ccggtacggt	tttcaatcgt	tgataaaatt
18181	tttgaaccca	gagttcgagg	ttgaagccat	ttgaaaggat	tagctgtgcc	ccctgagcac
18241	gctttatatac	tccggttgta	ggctgatatt	catggatttc	tgcgccaggc	cggttaatcg
18301	aactgacttc	ggcggcgtcc	cctgccacgt	tctgtgccat	gtcagcaatt	accgtgaatg
18361	tcgtgatgac	cttaaacttt	tcagcagcat	aagcgcctatg	aatagctgtc	agtccctgtga
18421	taaagggtgct	taacaggatt	atctttaccg	tcttttagttg	aagcatgggt	atctcctgat
18481	gaaatcactg	tttgtttaata	attagtaatg	ttataacata	gcttatacta	tgcaacatag
18541	cacaggctat	atttaaatga	gaaaaattat	ctttatcatt	taagagctga	tgccacagta
18601	tttcatggct	tatgcaagcg	ctgatatactc	agacatgcca	cttttaatgc	aacatgggtg
18661	cgttaaaatg	tgcattaatg	caacatagtt	gctgttaatc	actgagctat	gttgggatag
18721	gaaaaatgca	gaagcgccta	taacagagtc	agaacctctc	acttttactc	ttcctgatgg
18781	ctctctgaag	cacgtcagaa	aaggggcaac	tttgagaaac	gttgagaaac	gcactcgggtc
18841	ttccgctctg	aaaagcgtcg	tgtacgcgga	aatgtatggt	aatgtatggt	aatctgctga
18901	gaaagcaaaag	aacagtggaa	agctcaatat	cattacagtt	tttgatgaag	aagcgttggg
18961	cccgatctgc	cgaggttgtc	ttctccttct	cgagcgtct	gtaaaacagc	tgtttcctc
19021	ggccaggatg	gttgaagggc	aactgactaa	ggaaggggtt	tattacgatt	tccggttaga
19081	taaacctttt	actcgcgatg	acctggggat	gattgagcag	cacatgacaa	gtctgatcgc
19141	tgaaaacccg	ccgattatta	aagagcaggt	cacgcgggca	caggctattg	actgttttga
19201	gcagcggggc	gagaaattca	aatcaacgga	gttgatggaa	atactggcag	ataagtcggg
19261	gacactatgc	catctcaggc	agttcacgga	cggttttaac	ggccatttgg	ttccggatct
19321	tcgttttcta	aaaaatatga	agttgcttaa	cgtatccggg	gcttactggc	gtggtgaagc
19381	caacaatgta	cagctacagc	ggatttatgg	cacggcatgg	gccagcaaaa	aacagcttga
19441	tggttggtct	gagaaaacgg	cagaagcggg	gaaacgtgat	catcgcgaagc	tggttcggga
19501	gctggatttg	tttcattttc	aggacaatgc	gccagggcgcg	gttttctggc	atcccagggg
19561	gtggacaatc	tttcagtcgc	ttattgccta	tatgcgtagc	cgccatgagg	tgccgggtta
19621	cgttgaggtc	aatacgcctg	atgtcatgga	ccgcagcctg	tgggaaattt	ccggtcactg
19681	ggataattat	cgtgatcata	tgtttacgac	acaaccgag	gatggccgga	atctcgtctc
19741	caagccgatg	aactgtcctg	ggcgggtatc	gctgtttaaa	tacgggatca	agagttaccg
19801	cgattttacc	gttcgcctgt	ctgaatttgg	taaagtccat	cgttatgaa	cctctggatc
19861	gctgatggc	ttattacggg	ttccgcaatt	tactcaggat	gacgcgcata	tcttttgac
19921	gctgcaacag	gtcagggggg	aatgtaaaag	tatccttcag	ttggtgcttg	atatttataa
19981	gcaatttgga	tttgaagagg	tcgcaatcaa	ggtatccact	cgctcctgaaa	agcggatggg
20041	aagcgacgca	gactgggata	ggctggaaaa	cgcgctctca	gcctctctgg	aagcgcaggg
20101	attacaatgg	tcggtaaatc	ccggtgaagg	tgcttctat	ggccctaaac	tggtgattgt
20161	gttacgcgac	gcgattggtc	gtgactggca	gtgcggaacg	ctacaggtgg	atatgaatct
20221	gcctgagcga	tttgatattg	gctatattgc	cgaggatggc	tcgacaaaac	gcccgggtgat

20281 gttgcacagg gcgctcttcg gctcccttga gcggttcacc ggtatccttc ttgaacatta
20341 cgctcggtaaa ctgcccgcct ggctgtcgcc ggttcaggcc gtggtcatga cgattaccga
20401 caagcagcat cactatgcgg agcaggtttt gaaggcgctt cgcaagaaag ggctgcgttg
20461 tgaggccgac ctccgcaatg aaaaaattgg atttaaaatc cgtgaacaga ctctggccccg
20521 cattccttac ctgctgatcg ttggcgatgc ggaggagaat gggcgcgctg tcacggtgcy
20581 cgatcgtacc gggagggtgca tggggacgct gttactcaat gaggcggcag acgctatcaa
20641 aatatgttgc cagccaccgg aggttcacct tgattaagca aagcagaagc gatatcccc
20701 gacgcgagag cggtcgctac gtctggccgg tgaaggagga tgatacccaa agcaacgata
20761 tcgattttta taaaacggcg ttcagagAAC agtccaggcg atgggggctt tccgagatca
20821 ttcttagtaa tactggtgat ggagaaatat cactctatca gtctgttaat cagggggcag
20881 atgttctca cgtactgatt gcggcagggt tcatggaga agaaccgcy ggatgttggg
20941 ggatgctgga tttttgcy gagggctccc ctgaattgtt tgataacgta gccgtgagtt
21001 ttttaccgct tgtaaacctc tctggctcgc gcaactgtag ccgccttaat agtcttggac
21061 agaatccaaa ccgaggcttt accaaagggtg cagaagtcga gccttctatt gaaggaaaag
21121 tactgctcaa ctatgaaacg cttctgaaaa acgcagcaag tcatggtgtg ctctgttgcc
21181 atgaagatat tcttcgctat aaggcgtatc tctacacttt cgaacatgct actcgttgg
21241 gtcatttcag cgtggcatta cgtgatgaac ttgaacgctt tttcccgtc atggaaaaag
21301 agcaggttga tggttgtgaa tgtgaggatg gcattatttt taaccatttc gattcgtcat
21361 ttgaatcatg gttgtttct tctgttccg acgtcgcagc gtgcaactgaa acgccgggat
21421 tgcaaccctt tgcaaaacga gcagaagcaa accgttattt aatcggcgcy tttatctcat
21481 cgatactgga gcgcaacagc atcgggagcy ggaatgcatg aagccttat tactgaccaa
21541 tgcattaata atcaatgaag acttgcgcta tccggccgac atcctcattg ataaaggacy
21601 gatccagaaa attgcctcgc ttattccctc ccgaacggaa tggcaggtta ttgatgtgcy
21661 agggaaatgg gtgatccccg gtatgattga cgatcaggtt catttccgcy aaccggtct
21721 gacgcataag ggaaccattg ccagtgaatc agctgccgct gtaatgggag gcattaccag
21781 ttttatggaa atgcctaacg ttacgccacc aacaacgacc cttcaggcy tgcgtgagaa
21841 atttcaacgt gcctcgcatt cttcgttggc gaactacagt ttttacttcg gcgcyacgaa
21901 cgataatctt gacgaactga aagcgttac ggcttcgcaa gcctgtggg tgaagggtt
21961 tatgggagct tctaccggaa atatgctggt ggacgatgaa cagatcctgg aaagtattt
22021 cgccaacgca ccctgcctgg tagcaacaca ttgcyagcat accccgacaa taaagcataa
22081 tgaggagaca tggcgggctc gtttggggga cgtattcct gctggcgaac atgccgctat
22141 tcgttctgtg gacgcctgoc tgacatcttc ccaccaggcc gtctctctgg cgaaaaaaca
22201 tcgtaccgca ttacatgtcc tgcataatcac caccgcagat gaactcgtc tctttgatgc
22261 tgctcccacg ctggaggcygc tgcgccagaa aaccattaca gctgaagcct gcgttcatca
22321 cctgtttttt aattatgacg attatgaaat gctggggcat aagctcaaat gtaatccatc
22381 cgtaaaatcc tcttttcatc aggaagccct ccggagaggc gttaatgaag gcattattga
22441 cgttatcgcg actgaccatg cgcgcattt gctggaggaa aagcagaatg actattttgc
22501 tgcgccttcc ggcttaccgc tgggtgcagca tgccttctc gcaactgctg atatgagcag
22561 cagggggatt tttaccocgg aaatgggtgg gagaaagacc agcctatctg tagccgagcy
22621 tttccagctc aaagatcgtg gctatatcag agaaggctat tggcgcgacc tgggtgttat
22681 cgatccgttt agtcatcaac agataatccg ggaggatgtg gcatacaaat gcgggtggtc
22741 cccctttgaa ggccggattt tgtcaggcgg ggccggttgat atgacgctgg taaatggcca
22801 cgttatctgg aatgggcyca ccatacagca gaagtacggc ctgccactgg aattctgtcy
22861 ctagccacgt atttattttg ctcaggcata acttctccc gagaaattcag aagaggcaaa
22921 tcttgtgagt attattatc ctggttcttt cccggcacga cgactcgcac gttctcgcac
22981 ccattctttc agtcgcctgc tggctcgtga aaatcagctg accgtagatg acctgatttt
23041 accggtatcc attattgagg ggcaaacct gcgacaggaa gtgccgtcga tgcgggcyt
23101 atttcgtctg tcggttgata tgttgcctca tgatgcccy caatttgtt aggcaggtat
23161 acctgctatt gcaactttc cgtgtataga cagttctgct aaatcgtcga tggcagacgc
23221 ttcgtgggat ccatccggtt tagtaccgcy agctgtaagg gcgctgaaaa aggcttttcc
23281 tgaactcggg gtgataactg acgttgcctc tgacgcctat acgacacatg gtcaggacgy
23341 catcactgac gaaaatggat atgtacttaa tgagccgaca aaagaaattc tcaccagaca
23401 ggcgcttacc cacgcogaag caggcgcaga tattgtcgt ccaagcgaca tgatggatgg
23461 ccgcatcggc catatcagac aagcgttga aacgaatgga ctgcttaata tccagataat
23521 ggcgatcca gcaaaatac cgtgcaacta ttatggtcct tttcgcgag cyacgaaatc
23581 tgctgtatca ctgggtaaac gggcaaaaa aagctatcag atggatccgy cgaatgcaat
23641 ggaagctatg ccttaacctg atattatcag ggaagccagg aaaactttt cagtaccctg
23701 gccgggcatg ccttaacctg atattatcag ggaagccagg aaaactttt cagtaccctg
23761 ctttgcctat caggtttcag gggaaatatac gatgcatatg gccgctttcc gaaatgggtg
23821 gctggacgaa gaaaaacgy tgcctgaaat acttatgtc ttcaaacgy ccggggcgy
23881 tggaaactg acttatttt ccaatacagt agcccagtt cttaatcaac ataaacggga
23941 atattgatcc cctgttgaca aaattcctta gtcgtaagc aggagcaccg ctgagagtaa

```
24001 acgtcataat cagcgagaat ttctatactt taattcttat ctttattgcg aattctgacg
24061 aatcgtttgt atcagcccaa taatctggag gaagcagtta tggtcgaaac cagtaaacag
24121 gaccgtatgc cttatctact tttcctgtgt gaaaaccgca ttcttgca ca gaatatcgat
24181 ggtcatgtca ttgatc
```

//

Lebenslauf

Bettina Preisach

5. September 1977 geboren in Arad, Rumänien

Schul- und Hochschulausbildung:

1984 – 1990	Grund- und Gesamtschule in Engelsbrunn, Rumänien
Juni 1997	Abitur am Maria-Theresia Gymnasium, Augsburg
WS 1997 – SS 2004	Studium der Humanmedizin, Ludwig Maximilians Universität, München
SS 2003	Studium der Humanmedizin, Humboldt Universität, Berlin
August 1999	Physikum
August 2001	1. Staatsexamen
April 2003	2. Staatsexamen
April 2004	3. Staatsexamen
Oktober 2004	Approbation

Berufliche Ausbildung

07/2004 – 07/2007	wissenschaftliche Assistentin an der Medizinischen Poliklinik (Innenstadt) der LMU München
seit 10/2007	Assistenzärztin an der Ilmtalklinik in Pfaffenhofen (Innere Medizin)

Der experimentelle Teil dieser Arbeit wurde zwischen September 2000 und August 2002 am Max von Pettenkofer-Institut für Hygiene und Medizinische Mikrobiologie durchgeführt. Der schriftliche Teil der Arbeit wurde ab Juli 2007 abgefasst.

Mein Dank gilt allen voran Sören Schubert und Prof. Heesemann für die Bereitstellung des Themas. Sören danke ich für die hervorragende Betreuung, für enthusiastische und geduldige Besprechungstermine zu christlichen und zu unchristlichen Uhrzeiten. Bei allen Mitgliedern der AG Schubert bedanke ich mich für die unvergleichlich gute Arbeitsatmosphäre, die Einarbeitung in die Methodik und die Hilfsbereitschaft.

Danke vor allem an Robert, Birgit, Katrin, Stephan und Clemens fürs Motivieren!