This is a postprint of an article published in

Kresse, A.U., Blöcker, H., Römling, U.
ISPa20 advances the individual evolution of Pseudomonas aeruginosa clone C subclone C13 strains
isolated from cystic fibrosis patients by insertional mutagenesis and genomic rearrangements
(2006) Archives of Microbiology, 185 (4), pp. 245-254.

ISPa20 advances the individual evolution of *Pseudomonas aeruginosa* clone C subclone C13 strains isolated from cystic fibrosis patients by insertional mutagenesis and genomic rearrangements

Andreas U. Kresse • Helmut Blöcker • Ute Römling

Andreas U. Kresse[1] • Ute Römling
Research Group 'Clonal Variability', Division of Cell- and Immune Biology, GBF – German Research Centre for Biotechnology, Mascheroder Weg 1, D-38124 Braunschweig, Germany

Helmut Blöcker
Department of Genome Analysis, GBF – German Research Centre for Biotechnology, Mascheroder Weg 1, D-38124 Braunschweig, Germany

Ute Römling
Microbiology and Tumor Biology Center, Karolinska Institutet, SE-17177 Stockholm, Sweden
E-mail: Ute.Romling@mtc.ki.se
Tel.: +46 8 524 87319

[1]Present addresss: FG11 - Bacterial Infections, Dept. of Infectious Diseases, Robert Koch Institute, Burgstrasse 37, D-38855 Wernigerode, Germany

Abstract *Pseudomonas aeruginosa* clone C strains, which chronically colonize the lungs of cystic fibrosis patients reorganize their genome structure. In this study, a novel member of the IS3 subfamily of IS elements, ISPa20, was detected which was specific for clone C subclone C13 strains. ISPa20, which was present in high copy number, mediated events of genomic reorganization. ISPa20 was inserted into *P. aeruginosa* backbone genes leading to adaptation to the CF lung habitat and into DNA acquired through horizontal gene transfer. Furtheron, large chromosomal inversions were mediated by ISPa20. In contrast to strains of other subclonal linages high rates of genomic rearrangements of subclone C13 strains were observed in vitro. The acquisition of mobile elements by *P. aeruginosa* clone C strains in the lungs of cystic fibrosis patients supports the chronic colonization by insertional mutagenesis and chromosome restructuring leading to microevolution within clone C that reflects macroevolution observed on the species level.

Keywords clone C • cystic fibrosis • IS element • large chromosomal inversion • *Pseudomonas aeruginosa*

Introduction

The ability to create diversity by re-organization of the vast genomic information contributes to the adaptability of *Pseudomonas aeruginosa* to various habitats (Ernst et al. 2003; Kresse et al. 2003; Head and Yu 2004). Those mechanisms also work in the chronic colonization of cystic fibrosis (CF) patients by *P. aeruginosa* which is the major cause for lung deterioration and pulmonary failure (Hoiby et al. 1998). The underlying genetic modifications are reflected by phenotypic changes leading to a CF phenotype of the isolates which is characterized by, for example, alginate overproduction, lipopolysaccharide (LPS) modifications, loss of motility and the mutator phenotype (Kresse et al. 2003).

P. aeruginosa clone C is a major clone spread throughout Europe and worldwide (Römling et al. 2005). It was found associated mainly with CF but also prevalent in other habitats such as water biotopes (Römling et al. 1994). Clone C subclonal classification is based on distinct pulsed field gel electrophoresis (PFGE) patterns, which reflect the acquisition and loss of genetic information (Römling et al. 1997b; Larbig et al. 2002; Ernst et al. 2003; Klockgether et al. 2004). Genome analysis of subclone C13 strains, which were exclusively isolated from one patient suffering from CF have shown distinct differences such as several small insertions of 1-3 kbp (Römling et al. 1997a).

While analysis of the biological function of newly acquired genes is pending, the acquisition of mobile elements has been shown to play an important role in bacterial adaptation to the cystic fibrosis lung habitat (Kresse et al. 2003). Recently, we demonstrated individual genomic evolution in subgroup C strains of clone C. Subclone C strains were the predominant strains exclusively colonizing patients of the Medical School of Hannover, Germany during an observation period from 1986 to 1992 (Römling et al. 1997b). Those strains had acquired the 21 kbp composite transposon TNCP23 that was flanked by IS6100 elements (Kresse et al. 2003; Klockgether et al. 2004). The transposition activity of IS6100 was coupled with large chromosomal inversions (LCIs) thereby leading to gross chromosome reorganisation (Kresse et al. 2003). Significantly, IS6100 transposition caused adaptive mutations, shifting bacteria into the direction of the common phenotype of *P. aeruginosa* observed in CF.

IS6100 in *P. aeruginosa* was exclusively found in the subclonal lineage C, but not in the other subclonal linages C2, C13, or SG17M. Since LCIs were also found in CF strains from other subgroups, the LCIs of subclone C13 strains isolated from an individual CF patient at the MHH from 1985 to 1987 were analyzed. A novel IS element of the IS3 family exclusively detected in subclone C13 strains was found to cause inversions. Furthermore, the generation of insertional mutations was another important contribution of ISPa20 to the subclonal evolution of the C13 linage.

Materials and methods

Bacterial strains and growth conditions

P. aeruginosa strains used in this study are described in Table 1. If not otherwise indicated, *P. aeruginosa* was cultured in tryptone soy broth (TSB) or on tryptone soy agar (TSA).

General molecular methods

For DNA preparations the Nucleospin Kit (Machery and Nagel) or GFX miniprep Kit (Amersham Pharmacia) was used. Quantitative real time PCR was performed using a GeneAmp PCR System and the SYBR Green PCR Master Mix (Applied Biosystems). Primers for real time PCR for the detection of ISPa20 were IS3-Arev (5'–CCGAGCGAATTGGCGTTAGTA–3') and IS3-B (5'–CAAGATCAAAGTA-CGCCGCGG–3'), and for the detection of single copy ORF PA5170 were PA5170-1 (5'–CTACTTCGTCCTGTTGTTCAGC–3') and PA5170-2 (5'–ACGGTGAC-CAGCATCATGTTG–3'). Inverse PCR was carried out with religated genomic DNA, which had been digested with either of the restriction enzymes *Apa*I, *Acc*I, *Nco*I or *Sma*I and subsequent PCR using the outwarded primers orfB-for3 (TTGAAACCGACTATAACCGCCAGCG) and IS3-C (CAGTTCTTCGGCTTGC-TCGGC) specific for ISPa20.

Analysis of LCIs by PCR

For the analysis of inversion breakpoint by PCR, the following primers specific for PA5160, *rmlB* (PA5161) and the intergenic region between both ORFs were used: 5160-Afor (5'–ATCGCCCTGGCCTTCTTCATCATC-3'), 5160-Brev (5'–AGCCGCCGCCAGGAGTAAACC-3'), 5160-Crev (5'–GCGGGGCTTTTCGGG-ATTCA-3'), 5161-1 (5'–TTCATCGGCGCCAATTTCGTG-3'), 5161-Brev (5'–GGCCGTAGTTGTTCGAGCAGTTGG-3').

Pulsed-field gel electrophoresis

Preparation of high molecular-weight DNA and pulsed field gel electrophoresis (PFGE) was carried out as described before (Kresse et al. 2003). In brief, 2.5 x 10^9 cells/ml were embedded in low-melting-point agarose (type VII; Sigma). The DNA was digested with *Spe*I. PFGE was performed in a CHEF MAPPER (Bio-Rad). Experimental conditions: 0.5x Tris-borate-EDTA (TBE) buffer (45 mM Tris, 45 mM boric acid, 1 mM EDTA [8.3]), a 120º reorientation angle, and 1.5% agarose gel. Running conditions: 6 V/cm; pulse times: 8-50 s for 24 h, 12-25 s for 22 h, and 1-14 s for 14 h. Ethidium bromide stained gels were visualized using a CCD camera system (Fujifilm Image Reader LAS-1000).

Southern blotting

Primers for the amplification of the IS3 specific hybridization probe were IS3-Arev and IS3-B (see Molecular methods). The PCR product was purified by the QIAquick Nucleotide Removal kit (Qiagen) and DIG labelled using the DIG High Prime kit (Roche). Blotting of DNA to nylon membrane, the hybridisation procedure, and immunological detection of probe signals were performed according to previous protocols (Römling et al. 1994).

Sequence analysis

Sequence analysis was performed using the DNASTAR software package (DNASTAR, Inc., Madison, WI 53715, USA) and the Pseudomonas database (Stover et al. 2000). Nucleotide positions in ORFs refer to the designation in the Pseudomonas database (Stover et al. 2000). The sequence of ISPa20 is available at the EMBL database under accession number AJ849357 and in the IS finder database (http://www-is.biotoul.fr/).

Lipopolysaccharide analysis

LPS was isolated by the proteinase K method (Hitchcock and Brown 1983) and preparations were analysed by SES-PAGE using 15% discontinuous gels (Tsai and Frasch 1982). LPS bands were transferred to polyvinylidenefluoride membranes (Millipore) and blocked with 5% skimmed milk. Blotted membranes were incubated with mAbs 177 (specific for lipidA), 7-4 (inner core), monovalent P1 antiserum (Biorad; B-band, variable O-antigen with O-chain repeating units) and N1F10 (A-band, common conserved antigen).

Results

Genomic analysis of *P. aeruginosa* clone C subgroup C13

The genetic basis for large chromosomal inversions in subgroup C13 strains was analyzed. While strain C13 shared the same global chromosome structure as strain C, the clonal members of clone C subgroup C13, C14 and C15 showed LCIs of 5664 and 5886 kb (Fig. 1). As shown before IS6100 which caused genomic rearrangements in clone C, subgroup C strains, is not present in subgroup C13 strains (Kresse et al. 2003).

Previous detailed restriction fragment pattern analysis indicated that the inversion points in genotypes C14 and C15 are located in the same region of the *Spe*I restriction fragment SpF (Fig. 2a, (Römling et al. 1997b)). We therefore hypothesized that the same mechanism could be responsible for the LCIs in both genotypes. The recombination point was expected between PA5159 and PA5171 (Fig. 2a). Southern hybridization and combinatorial PCR narrowed down the region of inversion to 1761 bp between PA5160 and *rmlB* on the PAO1 chromosome (Fig. 2b). PCRs optimized for short range amplifications covering these 1761 bp were negative for strains C13, C14 and C15 while PCRs only covering sequences up- or down-stream of that region were positive. As described below, novel DNA harbouring an IS element of the IS3 family was inserted upstream of *rmlB*, at position 5810024 with reference to the PAO1 chromosome, in C13, C14 and C15 strains, but not in PAO1 or other clone C strains as determined by PCR (data not shown).

Inversion points in strains C14 and C15

In strain C15 the other inversion point derived from parental fragment SpAL which is invariable in size in all clone C strains (Römling et al. 1997b). Fragment SpAL had been cloned before from strain C9 (data not shown, (Kresse et al. 2003)) consequently strain C9 derived SpAL was sequenced to target the inverted region in strain C15 (Fig. 3). The 6.65 kb SpAL fragment is identical to the chromosomal region 56415 to 63067 of PAO1, which contains the gene encoding type III secreted

bacterial GTPase-activating protein ExoT as well as open reading frames (ORFs) of unknown functions.

To that end, with combinatorial PCRs using a primer for *rmlB* and various primers located in the SpAL fragment the inversion point of strain C15 was amplified. A PCR product approximately 1.6 kb larger in size than expected from the genomic backbone of PAO1 was obtained and sequenced. Novel DNA was inserted at position 54 of the $tRNA^{Thr}$ locus (PA5160.1) in the intergenic region between PA5160 and *rmlB* on fragment SpF (pos. 5810024 on the PAO1 chromosome). Upon this novel DNA an IS element of the IS3 family, ISPa20 (described below) was inserted. On fragment SpAL ISPa20 was inserted into PA0042 (pos. 56803 on the PAO1 chromosome, Fig. 4), which was confirmed by the sequence of PCR products specific for PA0042 and the IS3 element.

Efforts to amplify the DNA region covering the insertion point of ISPa20 downstream of PA5160 by using primers specific for PA0042' and PA5160 failed. We conclude that the DNA sequence, which is inserted into the $tRNA^{Thr}$ locus and partially inverted has a size that does not allow standard PCR amplification.

In strain C14 the insertion of ISPa20 upstream of the *rml* operon was confirmed by PCR (data not shown). The other inversion point in strain C14 is located on fragment SpAC, which, though present in all clone C strains, is absent from PAO1. Thus the DNA sequence is not available to check the recombination breakpoint.

LPS pattern in subgroup C13 strains

The gene *rmlB* encodes for a dTDP-D-glucose 4,6-dehydratase and is the first gene of a polycistronic operon involved in the core oligosaccharide and O polysaccharide assembly (Rahim et al. 2000). A $Sigma^{70}$ like promoter was suggested to be located 35 bp downstream of PA5160 (Fig. 2b, (Rahim et al. 2000)) indicating that insertion of the novel DNA into the $tRNA^{Thr}$ between PA5160 and *rmlB* could have separated the natural promoter from the *rmlB* operon. However, inner core specific antibodies did not reveal reduced size of the molecule subgroup C13 strains compared to an environmental isolate of clone C (Fig. 5) as it has been shown for a *rmlC* mutant (Rahim et al. 2000) suggesting that the *rml* operon is not affected in its functionality by insertion of the novel DNA into $tRNA^{Thr}$. However, the amount of variable O-

antigen (B-band) and conserved antigen (A-band) was reduced or not present in subgroup C13 strains.

A novel IS element of the IS*3*-family

The IS element found to cause LCIs in subgroup C13 reveals its highest homologies to members of the IS3-family, whereby IS3H from *Shigella dysenteriae* ATCC13313 showed the highest nucleotide sequence identity (acc. no. Z11609). BLASTX search at the *Pseudomonas* website revealed PA1938 as the closest homologue in PAO1 (51% similarity over 141 aa). ISPa20 is 1246 bp in length and classically composed of two ORFs framed by inverted repeats of 26 bp length. OrfA, which is proposed to be essential for sequence-specific binding to the terminal inverted repeats of the IS element as well as multimerization, exhibits the highest homology to a transposase from *Shigella flexneri* serotype 2a strain 2457T (acc. no. Q83J10). OrfB, which is of unknown function (Mahillon and Chandler 1998), exhibits the closest matches with *insF* (acc. no. P05822) from *Escherichia coli* and *S. flexneri* IS3 elements (74% similarity) and IS3G (73% similarity, acc. no. Z11608). An adenine rich sequence stretch at the C-terminus of OrfA presumably allows by programmed translational frameshifting the generation of a transframe protein which acts as the actual transposase (Mahillon and Chandler 1998). This frameshift also would complete a region resembling a leucine zipper motif (LZ) that starts at the C-terminus of OrfA. Many but not all members of the IS3 family have a LZ being completed by translational frameshifting (Mahillon and Chandler 1998). As another characteristic of the IS3-family the new IS element ISPa20 also shows a helix-turn-helix motif at the N-terminus of OrfA which is typical for regulatory factors. Based on sequence alignments of OrfB using IS finder (http://www-is.biotoul.fr/) and exhibiting a conserved $DD_{(35)}E$ consensus motif from the IS3 subfamily (Mahillon and Chandler 1998) the new IS element is identified as a member of the IS3-subgroup of the IS3-family (Fig. 6) and was designated ISPa20 according to the IS element nomenclature standard (http://www-is.biotoul.fr/).

At positions 799-808 an IHF binding motif (AAnnnnTTGAT) is found. Promoter sites recognized by $Sigma^{70}$ of *P. aeruginosa* have the consensus sequence YSTTGRn(17-18)YRTAAT (Ronald 1992). Scanning the ISPa20 sequence for this motif elucidated the slightly altered sequence TGTTTGAn(16)TATAAC at

nucleotide positions 1119-1147, where at position 3 of the motif T is inserted, the spacer region is 16 bp long and the last position holds C instead of T. Prediction of this putative promoter site was supported applying BPROM promoter prediction for prokaryotes (http://www.softberry.com). Interestingly, a reduced spacer region of the $Sigma^{70}$ consensus sequence leads to an enhanced activity of the promoter (McLean et al. 1997), and a 16 bp-spacer region can be found in $Sigma^{70}$ dependent promoters for *P. aeruginosa* genes *oprI*, *algP*, *algR2*, *iaaM* and *nylA* (Ronald 1992). A member of the IS3 family has been shown before to carry a strong promoter which induced the expression of *argE* in *E. coli* (Zafarullah et al. 1981; Charlier et al. 1982). Our observation suggests an almost conserved strong promoter region at the proximal right end of ISPa20 oriented outwards. An insertion of the IS element upstream of a genetic element might thus lead to activation of gene expression.

ISPa20 is found in high copy numbers in subgroup C13 strains

The insertion of an IS element coupled with an inversion event resembled what had been found in isolates of clone C subgroup C from CF patients (Kresse et al. 2003). In order to independently confirm the presence of ISPa20 on both inversion fragments Southern blot analysis was carried out using a 210 bp fragment of *orfA* as a probe using the the *Spe*I macrorestriction fragment pattern separated by PFGE. As expected, ISPa20 was found on fragment SpF in strain C13, as well as in C14 and C15 on fragments SpU' and SpK', and SpW'" and SpL', respectively (Fig. 6B, indicated by arrows). ISPa20 was exclusively found in subclone C13 strains (C13-C16), but was not present in the other clone C subgroups C, C2 and SG17M nor in PAO1 (Fig. 6B).

Surprisingly, the distribution of ISPa20 copies in subclone C13 strains was not restricted to the pair of inverted fragments. The hybridisation pattern of *orfA* revealed that strain C13 has at least 14 copies of ISPa20, which are distributed over the whole genome (Fig. 6B; Table 2). In strains C14 and C15 similar numbers and fragment distribution profiles were mapped for ISPa20, when using the *orfA* sequence as a probe.

Southern hybridisation applied to the *Spe*I macrorestriction fragment pattern separated by PFGE will give an underestimation of the copy number of ISPa20, because of the low resolution of the method. As an alternative estimation of the copy number of ISPa20 in subgroup C13 strains, quantitative real time PCR was applied whereby primers for *orfA* of ISPa20 were used for copy number estimation and primers for PA5170 as a reference. Amplification values for ISPa20 were calculated according to the directions for relative quantification of loci by the ABI Prism manual (Applied Biosystems) yielding the estimated copy numbers for ISPa20 in each strain. The result by quantitative real time PCR showed that subclone C13 strains might have up to 25 copies of ISPa20 (Table 2).

Targets of ISPa20 insertion

To identify the targets of ISPa20 insertion, inverse PCR with primers specific for ISPa20 in outward direction using C13, C14 and C15 genomic DNA as template was carried out. PCR products of different sizes were generated which most likely represented different insertion points of ISPa20. Sequencing identified 13 insertion breakpoints of ISPa20 in strain C13, C14 and C15. ISPa20 was inserted into genomic sequences homologous to PAO1 DNA, but also into novel DNA sequences specific for clone C strains (Table 3).

Most strikingly, several insertion sites of ISPa20 are located within genes involved in the adaptation of *P. aeruginosa* to the CF lung. In strain C13 a copy of ISPa20 was found inserted into *pilM*. *pilM* is part of the *pilMNOP* gene cluster involved in the biosynthesis of Type IV pili as judged by the lack of twitching motility and insensitivity to pili-specific phage (Martin et al. 1995). Most likely, the twitching motility negative phenotype of all subclone C13 strains (Table 1, (Römling et al. 2005)) is due to the insertional inactivation of *pilM* as manifested in the first isolate of this patient, strain C13, although we cannot exclude that additional genes involved in twitching motility in clone C13 strains are non-functional. We and others have shown that many CF isolates show a reduced capability to twitch (Mahenthiralingam et al. 1994; Kresse et al. 2003). Interestingly, in strain C8, subgroup C, the lack of twitching motility is caused by another IS element, IS6100, inserted into *pilB* coupled with an inversion (Kresse et al. 2003).

Another copy of ISPa20 was found between ORF13 (glycosyl transferase-like protein) and ORF14 (NAD dependent epimerase/dehydratase-like protein) of the gene cluster responsible for the biosynthesis of the O1-antigen of LPS (AF540990). Strain C13 agglutinated only with antiserum against O1-antigen, but Western blot analysis showed that reduced amounts of O1-antigen (B-band) of LPS were produced, indicating that insertion of ISPa20 affected O1-antigen biosynthesis. However, during the infection process strains C14, C15 and C16 subsequently lost the O1-antigen completely (Fig. 5). Loss of O-antigen of *P. aeruginosa* CF isolates is frequently observed, and IS elements play a role (Spencer et al. 2003). Altered O1-antigen profile in clone C strains could also be caused by IS element insertion. ISPa20 mediated the alteration of the O1-antigen profile in subgroup C13 strains, while in CF strain C9 IS6100 is inserted into *wbpM*, the last gene of the O1-antigen cluster (Kresse et al. 2003).

Other insertions were less obviously associated with the adaptation process of the strains to the CF lung. Another copy of ISPa20 was found inserted into *ribD* coding for an enzyme involved in de novo riboflavin biosynthesis from GTP. However, integration of ISPa20 into *ribD* was generated by homologous recombination between identical stretches of 10 nts (positions 1148-1157 of ISPa20, positions 1026-1017 of ribD), which leads to a slightly truncated copy of ISPa20.

Inserted into the very distal end of a putative two-component sensor (PA4036 at position 4518869) another copy of ISPa20 was found. Since only the last coding triplet is hit the insertion of ISPa20 might not affect the function of PA4036.

34 bp upstream of the start codon of PA1169 but in opposite orientation another copy of IS3 is located. PA1169 is a secreted lipoxygenase involved in the conversion of eukaryotic arachidonic acid into 15-hydroxyeicosatetraenoic acid thereby potentially modulating the local inflammatory responses during *P. aeruginosa* infection (Vance et al. 2004). Interference with activity of primary immune cells (polymorphonuclear neutrophils) plays an important role in the pathogenesis of *P. aeruginosa* in the CF lung (Sorrell et al. 1992; Ernst et al. 1999). Thus, the copy of ISPa20 in this genetic location might give an adaptation advantage to clone C13 by alteration of PA1169 expression and the manipulation of inflammatory responses.

Interestingly, another copy of ISPa20 was inserted into ORF C42, which is located on PAGI-2(C), a gene island described for *P. aeruginosa* clone C strains (Larbig et al. 2002). C42 encodes for a conserved integral membrane protein. Homologues of

C42 are found on several DNA elements acquired by clone C and other *P. aeruginosa* strains (He et al. 2004; Klockgether et al. 2004). The C42 homologue RL016 showed differential virulence in *P. aeruginosa* PA-14. An in-frame deletion mutant of RL016 was significantly less virulent in an Arabidopsis plant model, but equally virulent than wild type in the mice burn wound model (He et al. 2004).

Other insertions of ISPa20 into DNA sequences of subgroup C13 strains with no homologies in sequence databases show that novel DNA is affected by insertional mutagenesis. It remains a future task to demonstrate how the newly acquired DNA of clone C strains is involved in adaptation processes in the CF lung and how ISPa20 integration modulates these processes.

Discussion

The *P. aeruginosa* population in environmental and patient habitats is composed of diverse clonal lineages but some clones, like clone C, are dominant and spread across state borders thereby prevailing for decades. In those clones the individual evolution dependent on macro- and micro-ecological conditions can be studied.

P. aeruginosa from the lungs of cystic fibrosis (CF) patients develops high genotypic and phenotypic diversity (Römling et al. 1997b; Kresse et al. 2003; Spencer et al. 2003). The process of diversification founds mainly on the loss of genetic functions by mutation or insertional gene inactivation (Oliver et al. 2000; Kresse et al. 2003) or deletions (Römling et al. 1997b) whereas high frequency of recombination leading to the random association of alleles (Kiewitz and Tummler 2000) and the integration of additional genetic material clustered in gene islands (Larbig et al. 2002; Ernst et al. 2003; Klockgether et al. 2004) are features of the entire *P. aeruginosa* population.

A feature that seems to be unique to *P. aeruginosa* isolates from the CF lung habitat is large chromosomal inversions. Recent studies on LCIs in *P. aeruginosa* clone C isolates from CF patients indicated an active IS6100 element in subclone C strains, whereby transposition lead to disruption of genes essential for the adaptation to CF lungs in conjunction with global rearrangements of the chromosome (Kresse et al. 2003). In this study we show that in clone C subgroup C13 which did not harbour IS6100, an IS element of the IS3 family, ISPa20, is involved in the LCIs. This substantiates the important role and potential of acquired mobile elements for the shaping of the gross chromosome organisation of individual isolates of pathogens in new and challenging environments. However, unlike with IS6100 in *P. aeruginosa* clone C subclone C strains, the transpositional activity of ISPa20 in subclone C13 strains is not instantly coupled to an insertion-inversion mechanism. Several copies of ISPa20 are found in regions where no chromosomal rearrangements have been registered.

However, the transposition of ISPa20 did not randomly target genes on the chromosome, but preferentially targeted CF adaptive traits as well as genes on DNA islands. Already strain C13 exhibited a high degree of pathoadaptive traits, as for instance LPS modulation, lack of twitching motility and stable expression of alginate, some of which could be associated with ISPa20 insertions in respective

genes. Later isolates of subclone C13 show further progression towards the common CF phenotype, for instance the reduction of pyocyanin biosynthesis and reduced levels of total protease activities (Table 1). During colonization and persistence in the CF lung many regions of the large chromosome of *P. aeruginosa* might become redundant, if not disadvantageous when expressed, ISPa20 not only supports reorganization of the chromosomal structure, but caused mutational attrition by insertional activity.

Insertional gene inactivation, however, is only one mean of modulating gene expression essential for the bacterial adaptation to CF lungs (Kresse et al. 2003). ISPa20 comprises a putative strong promoter at the distal end of the element, which can switch on adjacent genes. IS elements have been noted to function as "mobile promoters" (Zafarullah et al. 1981; Charlier et al. 1982) provided that the downstream gene is in the proper orientation and at a permissive distance. Thus we conclude that the role of ISPa20 in the adaptation process not only lies in the disruption of gene expression but vice versa also might activate factors, which favour the persistence of *P. aeruginosa* in the CF lung.

Hypermutators are often found in CF patients and underline the importance of fast adaptation by mutation in the lung environment (Oliver et al. 2000). Frequently, the *mutS* gene is disrupted which leads to point mutations and small insertions/deletions that would be otherwise corrected by the mismatch repair system. In subclone C13 strains ISPa20 conducted the role of the mutagen. Thus IS elements constitute a variation of the hypermutator phenotype, which contributes to the adaptation of *P. aeruginosa* to the CF lung. Intriguingly, IS element insertion has also been found to be coupled with insertional inactivation of *mutS* (Kresse et al. 2003), although this might be assumed as a rare yet highly selective event.

Fast progression of the individual evolution of isolates within one CF patient has lead to an expansion of ISPa20 in subclone C13 strains. Already the parental strain C13 contained many copies of ISPa20 spread over the chromosome. This dissemination of potential recombination sites due to high copy numbers of ISPa20 not only implies the chance for gross adaptation by chromosomal rearrangements but also results in genomic lability and potentially in the impairment of fitness. Strain C14 grew much slower on agar plates reflecting a decrease in the overall fitness of the population at least under laboratory conditions. Indeed, when screening for a reversion event of the LCI in strain C14 by in vitro passaging, larger colonies, which showed chromosomal

deletions were often encountered. Interestingly, many deletions were associated with fragments containing ISPa20. In addition, it was not possible to recover the original C16 strain, since PFGE analysis showed changes of the macrorestriction fragment pattern after storage (data not shown). Little is known how transposition of IS elements is activated but nutrient stress and host factors play a role (Coros et al. 2005; Twiss et al. 2005).

Genome sequencing has shown that subspecies or species which develop fast on the evolutionary scale utilize IS element expansion to adapt to the environmental context rapidly. Prominent examples are *Shigella* (Jin et al. 2002), *Mycobacteria* (Gordon et al. 1999), *Helicobacter pylori* (Kalia et al. 2004), *Salmonella typhi* (Alokam et al. 2002) and *Yersinia pestis* (Parkhill et al. 2001). Thus, the inactivation of genes by insertion of IS elements as it is observed in clone C subclone C13 isolates from a CF patient can be regarded as the first step in a specialisation process, which occurs on a microscale within a short time frame.

Acknowledgements

We gratefully acknowledge Maren Scharfe for sequencing of SpAL, J. Lam for providing the antisera against LPS components and Yilmaz Mashid for performing the Western blot analysis. Michael Templin and Patricia Gatzlaff for assistance with real time PCR and Juergen Wehland for the provision of excellent working conditions.

Figure legends

Fig. 1 Graphical display of LCI. A. Macrorestriction X-alignments of the genomic arrangements of two inversion genotypes compared to the reference genotype C13. Chromosome sizes were normalized to 100% and chromosomes were plotted pairwise comparing positions of *Spe*I sites, while considering *oriC* as fixed. C14, genotype C14 versus genotype C13 (LCI 5664 kb); C15, genotype C15 versus genotype C13 (LCI 5886 kb). Grey box, proposed Ter region; open circle, arithmetical midpoint of the chromosome. Arrowheads give the position and orientation of the four rrn operons: filled*, rrnA*; grey, *rrnB*; dotted*, rrnC*; open, *rrnD*.

Fig. 2 Analysis of LCI in C14.

A. *Spe*I map of genotypes C13 and C14 displaying the calculated region of inversion in genotype C14 (not proportional). The open reading frames are withdrawn from the Pseudomonas database (27). Designations of *Spe*I fragments, capital letters; S = *Spe*I; inverted regions of *Spe*I fragments, grey or striked bar, respectively; position of a *Spe*I site in PAO1 as a reference, underlined.

B. Identification of the inversion breakpoint by combinatorial PCR. Chromosomal DNA homologous to the PAO1 backbone is shown as a striked line; disrupted $tRNA^{Thr}$, black bar; position and orientation of primers, arrows; the 1761 bp region from PAO1 amplified with primers 5160-Afor and 5160-Brev, grey bars; inserted novel DNA is indicated by a grey dotted bar with indication of the position with respect to PAO1; the position of an IS element is given. The $Sigma^{70}$ like promoter of the *rmlB* operon is shown by an arrowhead.

Fig. 3 *Spe*I map of genotypes C13 and C15 displaying fragment involved in the inversion in genotype C15 (not proportional) and the sequenced SpAL fragment from genotype C (not proportional). The open reading frames are withdrawn from the Pseudomonas database (Stover et al. 2000). Designations of *Spe*I fragments, capital letters; S = *Spe*I; inverted regions of *Spe*I fragments, grey or striked bar, respectively; position of a *Spe*I site in PAO1 as a reference, underlined.

Fig. 4 Involvement of ISPa20 in the LCI in strain C15. The position of ISPa20 insertions and inversion are indicated by vertical numbers with reference to the Pseudomonas database (Stover et al. 2000). Numbers in brackets give nucleotide regions of PA0042 fragments. A duplication of 5 bp of PA0042 is shown above the sequence ruler. Black bars indicate inverted repeats of ISPa20. The sequence right hand of ISPa20 located on SpW"' could not be resolved due to acquired unknown DNA sequence. However, the most likely arrangement is drawn in grey (not to scale). Ruler unit, 100 bp.

Fig. 5 Western blot analysis of lipopolysaccharide of C13 strains. The signals for lipid A, the inner core, O1-antigen and A-band are shown. After detection of the O1-antigen using P1 antiserum, the same blot was used to detect the inner core.

Fig. 6 Characterization and distribution of ISPa20 in C13 strains.

A. Schematic view on ISPa20 open reading frames and motifs. Regulatory binding sites are shown. IRL and IRR, inverted repeats; SD, Shine Dalgarno sequence; TFS window, translational frameshift window; HTH, helix-turn-helix motif; LZL, leucine zipper like motif; DD(35)E, $DD_{(35)}E$ motif as grey bar with black lines indicating positions of the aspartate and glutamate residues. Positions of primers for inverse PCR are indicated by arrows. a) Sequence alignment of inverted repeats. b) Sequence motif of the leucine zipper like motif. c) Sequence from ISPa20 of the highly conserved DD(35)E motif of the IS3 subgroup of IS3 elements.

B. Chromosomal distribution of ISPa20 in subgroup C13. *Spe*I digested chromosomal DNA from *P. aeruginosa* strains was separated by PFGE and hybridized with an PCR-amplified and DIG-labelled 210 bp probe specific for ISPa20. I, Ethidiumbromide stained PFGE gel; II, Southern blot. *P. aeruginosa* strains: 1, PAO1; 2, subclone C strain C; 3, subclone C2 strain C2; 4, subclone C13 strain C13; 5, subclone SG17M, strain SG17M; 6, subclone C13, strain C14; 7, subclone C13, strain C15; 8, subclone C13, strain C16.

Table 1 Characteristics of *P. aeruginosa* clone C subgroup C13 strains used in this study

Strain	**Source**	**Colony**	**Twitching motility**	**Swim**	**Mucoidy**	**Pellicle**	**Haemolysin**	**Pyocyanin**	**Protease**	**Mutator**	**Prototrophy**
C13	CF, Patient 13	mucoid	-	+	+	++	+	+++	+++	no	+
C14	CF, Patient 13	inhomogenous	-	-	-/+	++	(+)	++	(+)	n.d.	+
C15	CF, Patient 13	inhomogenous	-	-	+	+	-	(+)	(+)	n.d.	+
C16	CF. Patient 13	inhomogenous	-	+	-/+	+++	-	+	-	n.d.	+
SG31M	River sediment	homogenous	+	+	-	+	n.d.	n.d.	n.d.	n.d.	+

Table 2 Estimation of ISPa20 copy numbers in clone C13 strains.

strain	ISPa20 copy number as estimated by	
	Southern blot	quantitative real time PCR
C13	14	16
C14	11	24
C15	12	14
C16	14	25

Table 3 Insertion points of ISPa20 into clone C13 strains as revealed by inverse PCR.

affected gene	function	PAO1-position	strain
between PA1168 and PA1169	unknown; lipoxygenase	1267646	C13
C42 (SG50) on PAGI-2(C) (PAGI-3(SG))	integral membrane protein	*	C13
pilM	required for pilin biosynthesis	5680048	C13
PA4036	putative histidine kinase	4518869	C13
between ORF13 and ORF14 (AF540990)	glycosyl transferase-like protein and NAD dependent epimerase/dehydratase-like protein of O1-antigen biosynthesis cluster	(13914 of AF540990)*	C13
	unknown	*	C14
	unknown	*	C14
BZ55693.1	unknown	*	C14
	unknown	*	C14
Homologous to sequence on *Klebsiella oxytoca* plasmid pTKH11 (Y17716.1)	unknown	*	C14
ribD	Riboflavin biosynthetic protein	4537016	C14
between ORF41 and ORF42 *Yersinia enterocolitica* cryptic plasmid (AJ519722.2)	unknown	*	C15
Novel sequence inserted into $tRNA^{Gly}$	unknown	797641-797738 (404-501 of C15)	C15

		sequence)	

* additional DNA not existent in PAO1

References

Alokam S, Liu SL, Said K, Sanderson KE (2002) Inversions over the terminus region in Salmonella and Escherichia coli: IS200s as the sites of homologous recombination inverting the chromosome of Salmonella enterica serovar typhi. J Bacteriol 184:6190-6197

Charlier D, Piette J, Glansdorff N (1982) IS3 can function as a mobile promoter in E. coli. Nucleic Acids Res 10:5935-5948

Coros AM, Twiss E, Tavakoli NP, Derbyshire KM (2005) Genetic evidence that GTP is required for transposition of IS903 and Tn552 in Escherichia coli. J Bacteriol 187:4598-4606

Ernst RK et al. (2003) Genome mosaicism is conserved but not unique in Pseudomonas aeruginosa isolates from the airways of young children with cystic fibrosis. Environ Microbiol 5:1341-1349

Ernst RK et al. (1999) Specific lipopolysaccharide found in cystic fibrosis airway Pseudomonas aeruginosa. Science 286:1561-1565

Gordon SV, Heym B, Parkhill J, Barrell B, Cole ST (1999) New insertion sequences and a novel repeated sequence in the genome of Mycobacterium tuberculosis H37Rv. Microbiology 145 (Pt 4):881-892

He J et al. (2004) The broad host range pathogen Pseudomonas aeruginosa strain PA14 carries two pathogenicity islands harboring plant and animal virulence genes. Proc Natl Acad Sci U S A 101:2530-2535

Head NE, Yu H (2004) Cross-sectional analysis of clinical and environmental isolates of Pseudomonas aeruginosa: biofilm formation, virulence, and genome diversity. Infect Immun 72:133-144

Hitchcock PJ, Brown TM (1983) Morphological heterogeneity among Salmonella lipopolysaccharide chemotypes in silver-stained polyacrylamide gels. J Bacteriol 154:269-277

Hoiby N, Koch C, Frederiksen B (1998) [Cystic fibrosis]. Nord Med 113:328-330

Jin Q et al. (2002) Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157. Nucleic Acids Res 30:4432-4441

Kalia A et al. (2004) Evolutionary dynamics of insertion sequences in Helicobacter pylori. J Bacteriol 186:7508-7520

Kiewitz C, Tummler B (2000) Sequence diversity of Pseudomonas aeruginosa: impact on population structure and genome evolution. J Bacteriol 182:3125-3135

Klockgether J, Reva O, Larbig K, Tummler B (2004) Sequence analysis of the mobile genome island pKLC102 of Pseudomonas aeruginosa C. J Bacteriol 186:518-534

Kresse AU, Dinesh SD, Larbig K, Römling U (2003) Impact of large chromosomal inversions on the adaptation and evolution of Pseudomonas aeruginosa chronically colonizing cystic fibrosis lungs. Mol Microbiol 47:145-158

Larbig KD et al. (2002) Gene islands integrated into tRNA(Gly) genes confer genome diversity on a Pseudomonas aeruginosa clone. J Bacteriol 184:6665-6680

Mahenthiralingam E, Campbell ME, Speert DP (1994) Nonmotility and phagocytic resistance of Pseudomonas aeruginosa isolates from chronically colonized patients with cystic fibrosis. Infect Immun 62:596-605

Mahillon J, Chandler M (1998) Insertion sequences. Microbiol Mol Biol Rev 62:725-774

Martin PR, Watson AA, McCaul TF, Mattick JS (1995) Characterization of a five-gene cluster required for the biogenesis of type 4 fimbriae in Pseudomonas aeruginosa. Mol Microbiol 16:497-508

McLean BW, Wiseman SL, Kropinski AM (1997) Functional analysis of sigma-70 consensus promoters in Pseudomonas aeruginosa and Escherichia coli. Can J Microbiol 43:981-985

Oliver A, Canton R, Campo P, Baquero F, Blazquez J (2000) High frequency of hypermutable Pseudomonas aeruginosa in cystic fibrosis lung infection. Science 288:1251-1254

Parkhill J et al. (2001) Genome sequence of Yersinia pestis, the causative agent of plague. Nature 413:523-527

Rahim R, Burrows LL, Monteiro MA, Perry MB, Lam JS (2000) Involvement of the rml locus in core oligosaccharide and O polysaccharide assembly in Pseudomonas aeruginosa. Microbiology 146 (Pt 11):2803-2814

Ronald S, Farinha, M.A., Ilan, B.J., Kropinski, A.M. (1992) Cloning and physical mapping of transcriptional regulator (sigma) factors from Pseudomonas aeruginosa. In: Galli E, Silver, S., Withold, B. (ed) Pseudomonas, Molecular Biology and Biotechnology. American Society for Microbiology, Washington, DC

Römling U, Kader A, Sriramulu DD, Simm R, Kronvall G (2005) Worldwide distribution of Pseudomonas aeruginosa clone C strains in the aquatic environment and cystic fibrosis patients. Environ Microbiol 7:1029-1038

Römling U, Schmidt KD, Tummler B (1997a) Large chromosomal inversions occur in Pseudomonas aeruginosa clone C strains isolated from cystic fibrosis patients. FEMS Microbiol Lett 150:149-156

Römling U, Schmidt KD, Tummler B (1997b) Large genome rearrangements discovered by the detailed analysis of 21 Pseudomonas aeruginosa clone C isolates found in environment and disease habitats. J Mol Biol 271:386-404

Römling U, Wingender J, Muller H, Tummler B (1994) A major Pseudomonas aeruginosa clone common to patients and aquatic habitats. Appl Environ Microbiol 60:1734-1738

Sorrell TC, Muller M, Sztelma K (1992) Bacterial metabolism of human polymorphonuclear leukocyte-derived arachidonic acid. Infect Immun 60:1779-1785

Spencer DH et al. (2003) Whole-genome sequence variation among multiple isolates of Pseudomonas aeruginosa. J Bacteriol 185:1316-1325

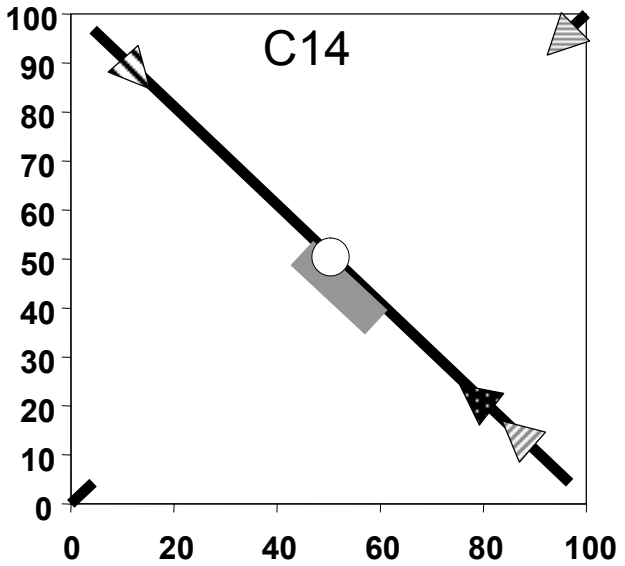
Stover CK et al. (2000) Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen. Nature 406:959-964

Tsai CM, Frasch CE (1982) A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. Anal Biochem 119:115-119

Twiss E, Coros AM, Tavakoli NP, Derbyshire KM (2005) Transposition is modulated by a diverse set of host factors in Escherichia coli and is stimulated by nutritional stress. Mol Microbiol 57:1593-1607

Vance RE, Hong S, Gronert K, Serhan CN, Mekalanos JJ (2004) The opportunistic pathogen Pseudomonas aeruginosa carries a secretable arachidonate 15-lipoxygenase. Proc Natl Acad Sci U S A 101:2135-2139

Zafarullah M, Charlier D, Glansdorff N (1981) Insertion of IS3 can "turn-on" a silent gene in Escherichia coli. J Bacteriol 146:415-417


C14
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100

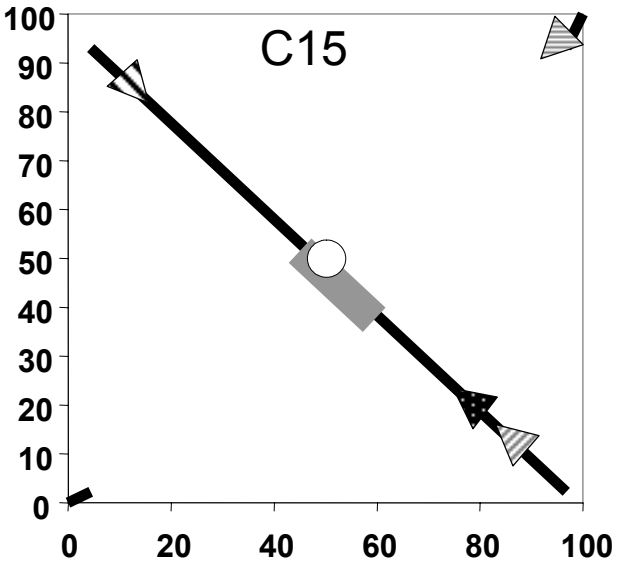

C15
100
90
80
70
60
50
40
30
20
10
0
0
20
40
60
80
100

Fig. 1

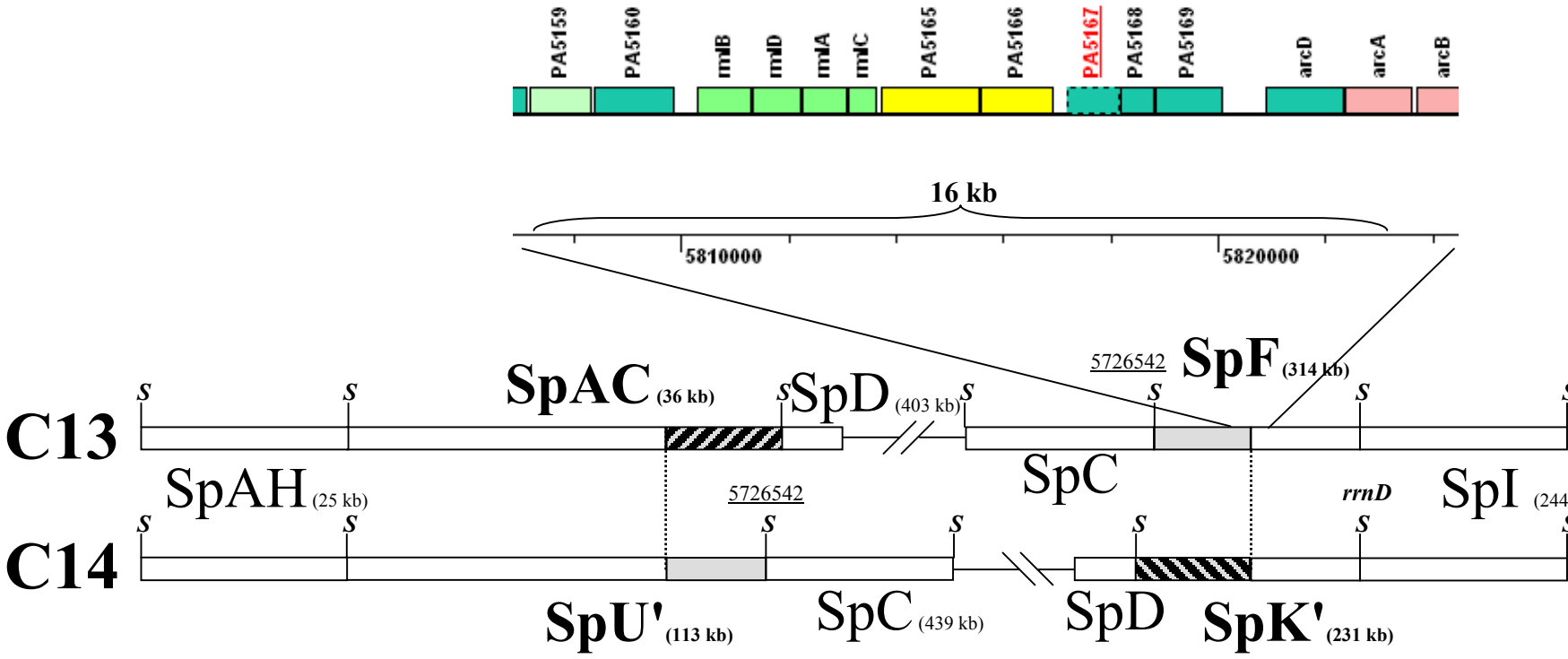

PA5159
PA5160
mlaB
mlaD
mlaA
mlaC
PA5165
PA5166
PA5167
PA5168
PA5169
arcD
arcA
arcB
16 kb
5810000
5820000
5726542
SpF (314 kb)
C13
SpAC (36 kb)
SpD (403 kb)
SpC
SpAH (25 kb)
5726542
rrnD
SpI (244
C14
SpU' (113 kb)
SpC (439 kb)
SpD
SpK' (231 kb)

Fig. 2A

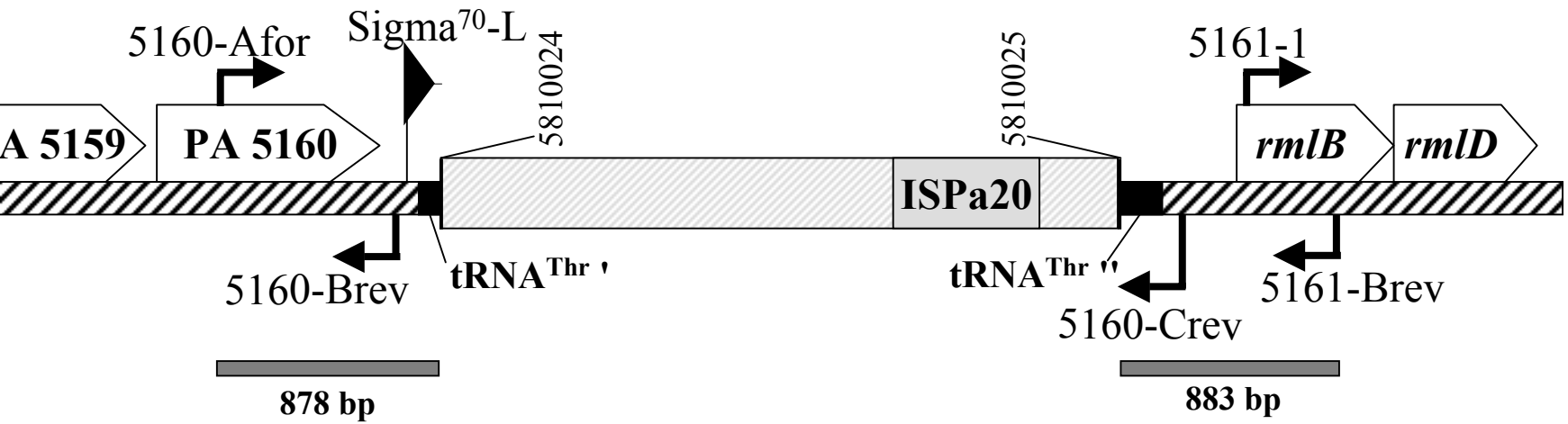

5160-Afor
Sigma70-L
5810024
5810025
5161-1
A 5159
PA 5160
rmlB
rmlD
ISPa20
5160-Brev
tRNAThr '
tRNAThr ''
5160-Crev
5161-Brev
878 bp
883 bp

Fig. 2

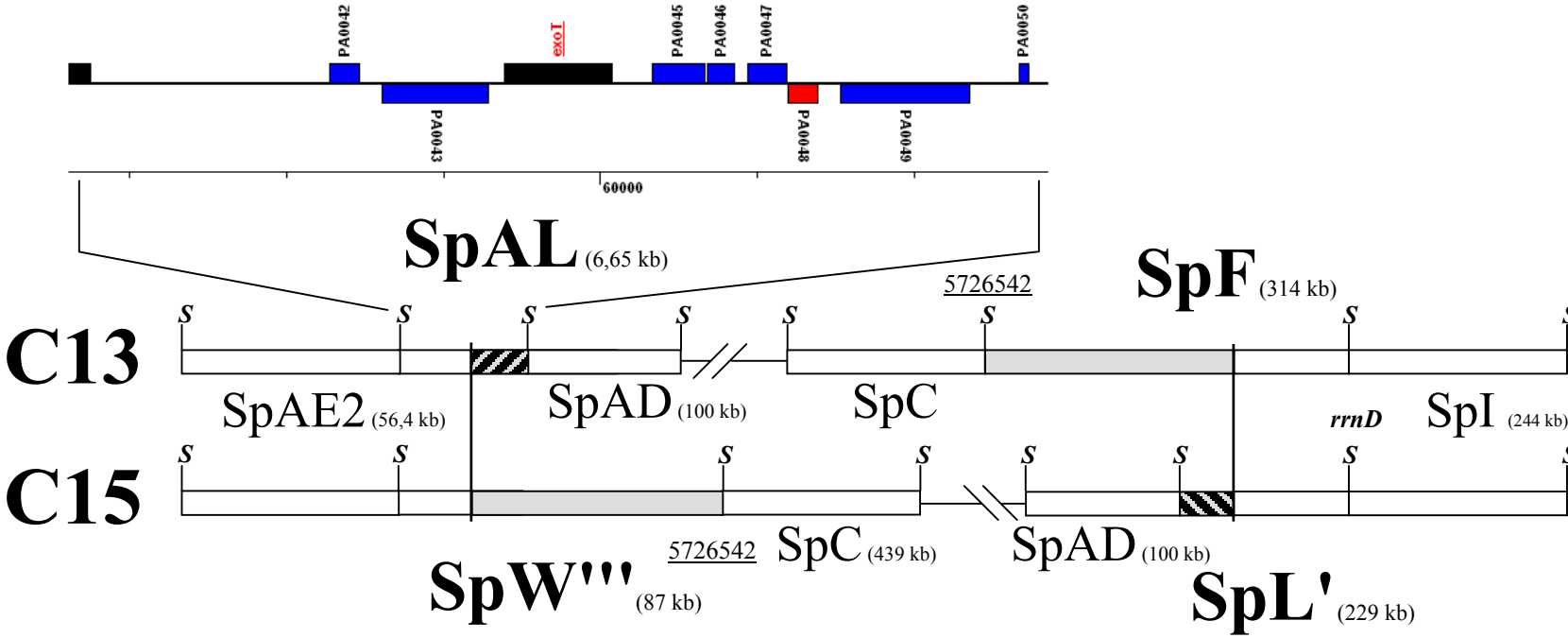
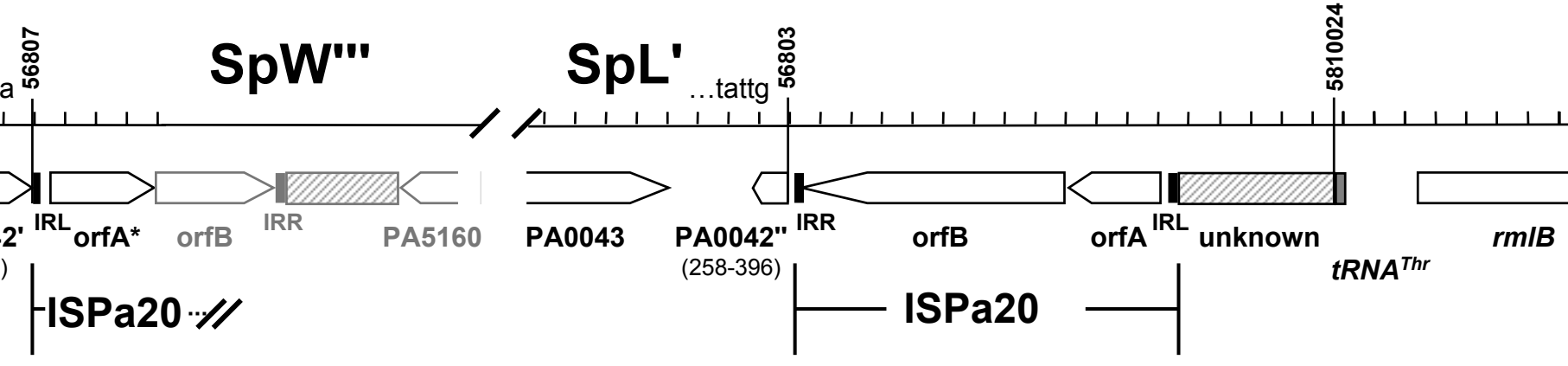
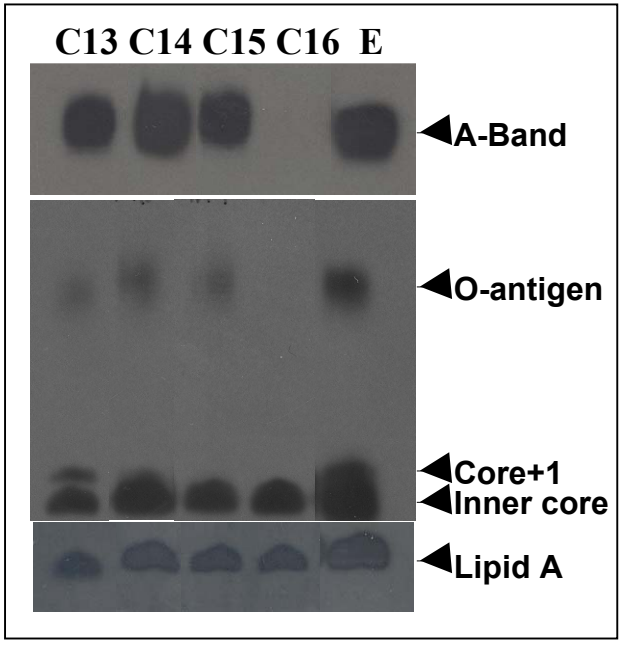
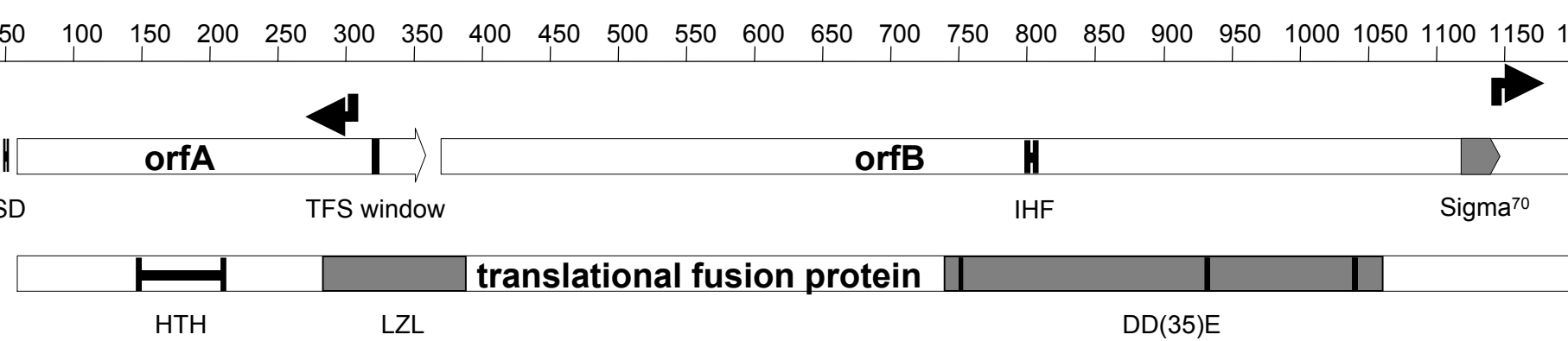

PA0042
exoT
PA0045
PA0046
PA0047
PA0050
PA0043
PA0048
PA0049
60000
SpAL (6,65 kb)
5726542
SpF (314 kb)
C13
SpAE2 (56,4 kb)
SpAD (100 kb)
SpC
rrnD
SpI (244 kb)
C15
5726542
SpC (439 kb)
SpAD (100 kb)
SpW''' (87 kb)
SpL' (229 kb)

Fig. 3


SpW'''
SpL'
…tattg
56807
56803
5810024
IRL
orfA*
orfB
IRR
PA5160
PA0043
PA0042"
(258-396)
IRR
orfB
orfA
IRL
unknown
rmlB
tRNA Thr
ISPa20
ISPa20

Fig


C13 C14 C15 C16 E
A-Band
O-antigen
Core+1
Inner core
Lipid A




```
          1                15          26
IRL       TGATCTTGCCCAGCAACCGTGGACAC
          ||||||||||||||||...||||||||
IRR       TGATCTTGCCCAGCACGAGTGGACAC
          1246             1232        1221

LZL             L(6)L(6)Q(6)V(6)L(6)

DD(35)E         W..DITY(53)HSDRGS.Y.S(16)S..G...DN...ESF...LK
```

Fig.

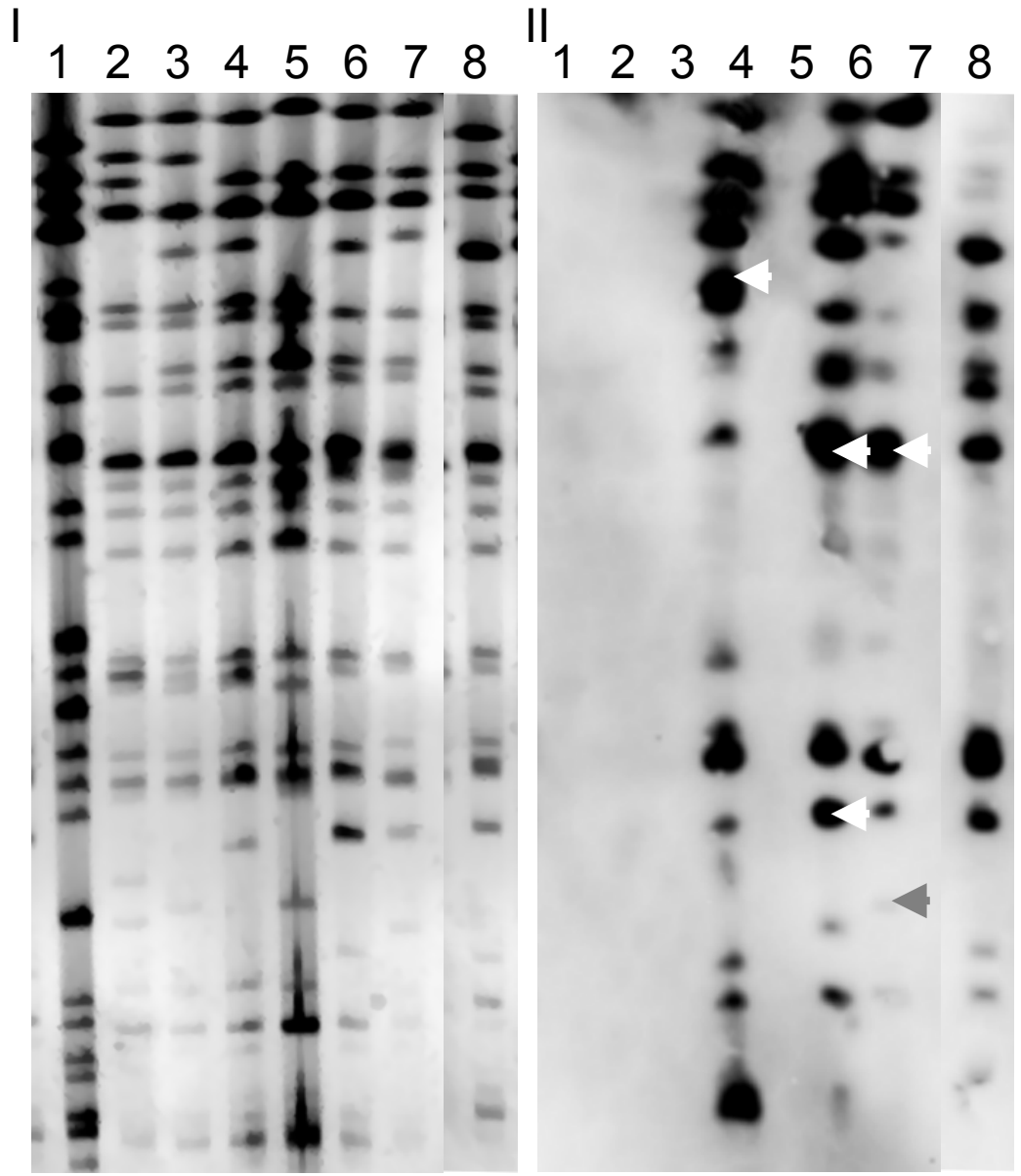


Fig.