Phage-inducible islands in the Gram-positive cocci

Roser Martínez-Rubio[1,5], Nuria Quiles-Puchalt[2,5], Miguel Martí[3,5], Suzanne Humphrey[2], Geeta Ram[4], Davida Smyth[4], John Chen[4], Richard P Novick[4,*], José R Penadés[2,*]

[1]Departamento de Ciencias Biomédicas, Universidad CEU Cardenal Herrera, 46113 Moncada, Valencia, Spain; [2]Institute of Infection, Immunity and Inflammation, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, G12 8TA, UK; [3]Centro de Investigación y Tecnología Animal, Instituto Valenciano de Investigaciones Agrarias (CITA-IVIA), Apdo. 187, 12.400 Segorbe, Castellón, Spain; [4]Skirball Institute Program in Molecular Pathogenesis and Departments of Microbiology and Medicine, New York University Medical Center, 540 First Avenue, New York, NY 10016, USA.

[5]These authors contributed equally to this work.

Running title: Phage-inducible chromosomal islands

Keywords: *Staphylococcus aureus*, pathogenicity island, SaPI, PICIs, gene transfer, mobile genetic elements, helper phage, TerS

[*]Corresponding authors: José R. Penadés
Institute of Infection, Immunity and Inflammation
College of Medical, Veterinary and Life Sciences
University of Glasgow
e-mail: JoseR.Penades@glasgow.ac.uk

Richard P. Novick
Departments of Microbiology and Medicine
Skirball Institute,
New York University Medical Center
e-mail: Richard.Novick@med.nyu.edu

Abstract

The SaPIs are a cohesive sub-family of extremely common phage-inducible chromosomal islands (PICIs) that reside quiescently at specific *att* sites in the staphylococcal chromosome and are induced by helper phages to excise and replicate. They are usually packaged in small capsids composed of phage virion proteins, giving rise to very high transfer frequencies, which they enhance by interfering with helper phage reproduction. As the SaPIs represent a highly successful biological strategy, with many natural *Staphylococcus aureus* strains containing two or more, we assumed that similar elements would be widespread in the Gram-positive cocci. On the basis of resemblance to the paradigmatic SaPI genome, we have readily identified large cohesive families of similar elements in the lactococci and pneumococci/streptococci plus a few such elements in *Enterococcus faecalis*. Based on extensive ortholog analyses, we find that the PICI elements in the four different genera all represent distinct but parallel lineages, suggesting that they represent convergent evolution towards a highly successful life style. We have characterized in depth the enterococcal element, EfCIV583, and have shown that it very closely resembles the SaPIs in functionality as well as in genome organization, setting the stage for expansion of the study of elements of this type. In summary, our findings greatly broaden the PICI family to include elements from at least three genera of cocci.

Introduction

During the past several years, we have discovered and extensively characterized the SaPIs, 15 kb phage-inducible pathogenicity islands in staphylococci (reviews: Novick *et al.*, 2010; Penadés and Christie, 2015; Novick and Ram, 2015). The SaPIs are highly evolved mobile genetic elements (MGE's) that are descended from an ancestral prophage or protophage, of which they have retained a number of key features that define their functionality and lifestyle. These include prophage-like transcriptional organization, SOS-insensitive repressors that are countered by helper phage-encoded de-repressors, integration/excision, autonomous replication when de-repressed, specific packaging in phage-derived proheads, high frequency transfer of unlinked host genes as well as of their own genomes, and interference with phage reproduction. Their genome organization is prophage-like and consists of a small set of genes transcribed in one direction, starting with an integrase (*int*) gene, a somewhat larger set of genes transcribed in the opposite direction, and a pair of divergent regulatory genes flanking the transcriptional switch (Fig. 1). The larger transcriptional region encodes an excision function (*xis*), a primase homolog (*pri*) and a replication initiator (*rep*), which are sometimes fused, followed by a replication origin, the genes involved in phage interference, and, usually, a terminase small subunit homolog (*ter*S). The SaPIs have diverged radically from the putative ancestral prophage (or protophage) and occupy a unique evolutionary niche in their host bacteria. It is likely that they have maintained their evolutionary independence owing to their development of phage interference and the acquisition of hypothetical and accessory genes, none of which are phage derived. The SaPIs form a large cohesive family with many natural *S. aureus* strains containing two or more. The cohesiveness of the SaPI family is underlined by the fact that their ORFs belong to large sets of orthologs, within which the first 8-10 are almost always SaPI genes, most of which never appear in other genetic elements (Novick and Ram, 2015). SaPIs are found at 5 different attachment sites; those in the same site are often more closely related to one another than to those in other sites. The helper phages and SaPIs undergo rapid coevolution, which is likely

to play a key role in the evolution and ecology of the bacteriophages as well as of their prokaryotic hosts (Frígols *et al.*, 2015).

Given their unique, fascinating, and highly successful biological strategy, plus their importance for the biological activites of *S. aureus* (Penadés and Christie, 2015; Novick and Ram, 2015), it seemed highly likely that elements with similar genome organization and functionality would be found in other bacteria. The first hint of SaPI-like elements in other bacteria was provided by the genomic analysis of *E. faecalis* strain V583, which was sequenced by Paulsen *et al.* (Paulsen *et al.*, 2003) and shown to contain 7 prophage-like elements. We observed that one of these, p7, has the same genome organization and is about the same size as a SaPI (Fig. 1). We mentioned it in a review paper on the SaPIs and suggested that it be re-designated EfCI1 (for *E. faecalis* Chromosomal Island 1) (Novick *et al.*, 2010).

Later, Duerkop *et al*. showed that EfCI1 was co-transferred at very high frequency with phage p1. Since EfCI1 was only 12 kb in length and would clearly be unable to generate infective particles on its own, the authors proposed that it and p1 formed “composite” particles (Duerkop *et al.*, 2012). Soon thereafter, Matos, *et al.*, unable to demonstrate “composite” particles, suggested instead that EfCI1 and p1 might be similar to a SaPI-helper phage pair (Matos *et al.*, 2013). They reported that p1, a fully functional temperate phage, is uniquely required for EfCI1 packaging and transfer, but unlike the SaPI helper phages, not for induction. The authors proposed, at our suggestion, that EfCI1 be re-designated EfCIV583, replacing the “1” in the generic designation, EfCI1, with “V583”, to indicate the strain of origin, in keeping with the nomenclature that we had earlier proposed for elements of this type (Novick *et al.*, 2010; Penadés and Christie, 2015). Since p1 did not seem to be required for EfCIV583 induction, we wondered whether EfCIV583 was, in fact, fully comparable to a SaPI or, perhaps, was a new type of phage-related element. In this report, we demonstrate that EfCIV583 closely resembles a typical SaPI, requiring a helper phage for induction, packaging and transfer; however, it differs critically from the SaPIs in lacking the

terminase small subunit homolog that determines packaging specificity. We demonstrate also that other low G+C Gram-positive cocci, especially the lactococci and streptococci (particularly the pneumococci), harbor large families of SaPI-like elements and that there are a few in the enterococci, very similar to EfCIV583. We designate the entire class of these elements phage-inducible chromosomal islands (PICIs), of which the SaPIs are a subset.

Materials and Methods

Bacterial strains and growth conditions.

Bacterial strains used in these studies are listed in Table S1. Phage and PICI analyses were performed as described (Tormo-Más *et al.*, 2013; 2010; Chopin *et al.*, 2001).

Identification of PICIs.

The analysis of orthologies points to elements that might correspond to PICIs. Examination of the corresponding KEGG genome maps (http://www.genome.jp/kegg; release May 1 2016) was used to confirm the identifications. In these analyses we examined only the 13 *L. lactis* and the 26 *S. pneumonaie* genomes that have been coded for KEGG because the KEGG genome maps enable PICI-like elements to be readily identified. Since each of these genomes contains at least 1 PICI-like element, it is certain that such elements are very common among the lactococci and streptococci; additional details are given in Supplementary Materials and Methods.

DNA methods.

General DNA manipulations were performed as previously described (Ubeda *et al.*, 2008). DNA probes were generated by PCR using primers listed in Table S2.

Plasmid construction.

Plasmid constructs (Table S3) were prepared by cloning PCR products obtained with oligonucleotide primers as listed in Table S2.

Enzyme assays.

β-Lactamase assays, using nitrocefin as substrate, were performed with cells in exponential growth phase as described (O'Callaghan *et al.*, 1972).

In-gel enzymatic digestion and mass fingerprinting.

Protein bands of interest were analyzed as previously described (Tormo-Más *et al.*, 2010).

Results

The results presented in this report are in two major subsections. The first contains an analysis of the resemblance between EfCIV583 and a typical SaPI. The second presents an extensive analysis of lactococcal and streptococcal genomes for SaPI-like elements.

Enterococcus faecalis EfCIV583

We tested the following key life cycle genes in this element: *int*/xis (integration/excision), *rep/ori* (replication initiator, replication origin) and *rpr* (repressor) for functionality, and noted, incidentally, that the element lacked a recognizable homolog of *ter*S*,* a gene present in most SaPIs and, in those SaPIs, essential for SaPI-specific transfer (Ubeda *et al.*, 2007b). We also identified the p1 gene that is responsible for EfCIV583 induction.

Integration/excision

To confirm the functionality of the *int*/*xis* system of EfCIV583, we performed PCR analysis with inward- and outward-directed primers as shown in Figs. 2A and B. In both cases, amplicons were obtained, suggesting *int* functionality. To confirm this, we analyzed *int* activity ectopically in *E. coli*. We prepared derivatives of the thermosensitive plasmid pMAK700 (Hamilton *et al.*, 1989) containing the chromosomal attachment site (att_C) for EfCIV583 and derivatives of pCN51 (Charpentier *et al.*, 2004), containing the cognate PICl attachment site (att_{PI}) plus *int* (Fig. 2C). Plasmid pairs were tested for cointegrate formation by overnight growth (at 30ºC) followed by plating on doubly selective medium at 43ºC, the restrictive temperature for pMAK700. Colonies were obtained only with plasmids containing the cognate *att* sites and *int* gene; no colonies were obtained when the *att* sites and *int* genes were from different elements. Cointegrate formation was confirmed by restriction

155 analysis (Fig. 2D), and by plasmid sequencing.

156 *Replication*

157 The Pri-Rep-*ori* segment of SaPIbov1, SaPI1 or SaPIn1 can drive replication of a suicide
158 plasmid in *S. aureus* (Ubeda *et al.*, 2008; 2007a). Since the EfCIV583 replicon seemed to
159 have the same overall organization as the SaPIs (see Fig. 1 and S1A), we tested for
160 functionality of its replicon by cloning the corresponding segment of EfCIV583 (plasmid
161 pJP782) into an erythromycin resistance (Em^r) plasmid containing the *E. coli* ColE1 replicon
162 (incapable of replication in *S. aureus* or other Gram-positive bacteria). We also constructed
163 plasmids carrying mutations in the *rep* gene or in the *ori* site (plasmids pJP1306 and pJP781)
164 of EfCIV583. As controls, we used similar plasmids carrying the replication module from
165 SaPI1 or SaPIbov1 (plasmids pRN9217 and pRN9211 respectively). Plasmids were
166 transferred to *S. aureus* RN4220, *E. faecalis* JH2-2 and *Bacillus subtilis* RL-3 by
167 transformation with selection for the Em^R marker of the plasmid. Additionally, since
168 staphylococcal phages can transfer SaPIs and plasmids to *Listeria monocytogenes* (Chen
169 and Novick, 2009; Chen *et al.*, 2015a), we transduced the functional plasmids from *S. aureus*
170 to this organism, also with erythromycin selection. Plasmids carrying the *pri-rep-ori* segment,
171 but not those with mutations in these loci produced colonies in the recipient strains (Table 1).
172 DNA extraction from the different hosts confirmed the presence of the plasmids containing
173 the replication module (Fig. S1B). These results confirm that EfCIV583, SaPI1 and SaPIbov1
174 can replicate in different Gram-positive bacteria, and that this replication depends on the
175 presence of a functional replicon.

176 *Functionality of EfCV583 master repressor*

177 SaPI gene expression is controlled by a master repressor, Stl, analogous to λ-c1 (Ubeda *et*
178 *al.*, 2008). To test for a comparable repressor in EfCIV583, we cloned into a β-lactamase
179 reporter vector, pCN41 (Charpentier *et al.*, 2004), the region of EfCIV583 corresponding to
180 the SaPI regulatory region, which contains *stl*, the promoter that it represses, and the site of
181 transcriptional divergence (Ubeda *et al.*, 2008; Tormo-Más *et al.*, 2010). Clones were

constructed in *E. coli* either with or without the *stl*-like gene, which is here designated *rpr* (repressor, see Fig. 3A), the reporter constructs were introduced by transformation in *S. aureus* RN4220, and β-lactamase activities were measured. The clone containing *rpr* showed sharply lower β-lactamase activities than that lacking it, consistent with repressor function for Rpr (Fig. 3B).

Identification of the phage p1 coded EfCIV583 inducer.

Since p1 did not seem to be required for EfCIV583 induction, we wondered whether EfCIV583 was, in fact, fully comparable to a SaPI or, perhaps, was a new type of phage-related element. Accordingly, we subjected the element to the two tests by which the original SaPIs had been defined. First, we SOS-induced strain VE14089 with mitomycin C (MC), removing aliquots of the culture at different times during induction. VE14089 is a plasmid-cured derivative of strain V583 (Table S1), which facilitates the interpretation of agarose gels used for analysis of EfCIV583. The cells were lysed with lysozyme and the lysates analyzed by agarose gel electrophoresis. In the SaPI system, in an experiment of this type, late in the lytic cycle, a new band appears in the electropherogram, migrating more rapidly than the sheared chromosomal (bulk) DNA (Lindsay *et al.*, 1998; Ubeda *et al.*, 2005). This band has the mobility of the SaPI monomeric DNA and represents SaPI DNA released from intracellular mature small particles or filled small proheads during preparation of the DNA sample (Ubeda *et al.*, 2007b; 2008). As shown in Fig. 4A, the gel pattern observed with VE14089 following phage induction was indistinguishable from that seen with the classical SaPIs (Lindsay *et al.*, 1998; Ubeda *et al.*, 2005), and the identity of the SaPI-like band was confirmed as EfCIV583 by Southern blotting (Fig. 4B).

The second test involved the pelleting of the particles in a MC-induced VE14089 lysate and isolation of the particle DNA followed by agarose gel electrophoresis. With a SaPI-helper phage combination, two bands are seen with such DNA – one of phage monomer size, the other of SaPI monomer size (Ubeda *et al.*, 2007b). Again, as shown in Fig. 4C, the same gel pattern was seen with EfCIV583 and confirmed by Southern blotting. These results imply that

EfCIV583 is packaged in small capsids as are many of the SaPIs (Ruzin *et al.*, 2001; Ubeda *et al.*, 2005); indeed, such small particles have been observed by Matos *et al.* (Matos *et al.*, 2013). The sequence of EfCIV583, however, does not contain any identifiable homologs of the SaPI *cpm* genes, which are responsible for small particle formation (Ubeda *et al.*, 2007b), and we are presently attempting to identify such genes by mutational analysis.

To test further the implication that EfCIV583 and p1 represent a SaPI-like element and its helper phage, we introduced a selectable marker (*tet*M) into a putatively non-essential region of EfCIV583 (Fig. S2), then prepared a mitomycin C (MC)-induced lysate of JP11028, containing this derivative and phage p1, and tested for transfer of the *tet*M marker to two different non-lysogenic *E. faecalis* strains. Transfer of the *tet*M marker was observed at a frequency of up to ~10^5/ml (Table 2), 3 orders of magnitude greater than would be expected for generalized transduction. Incidentally, sequencing of the DNA obtained from phage particles in these experiments revealed that the individual phage or PICI elements are packaged independently, confirming the absence of any fused p1/EfCIV583 genome.

To confirm that p1 is the helper phage for EfCIV583, using pMAD derivative plasmids (Table S3), we cured VE14089 of prophages 1-6. We found that the elimination of p1 but not any of the others eliminated EfCIV583 transfer (in accordance with (Matos *et al.*, 2013)), as well as the appearance of the PICI band following MC induction (Fig. 4B and Table 2). Moreover, MC treatment of a non-lysogenic EfCIV583-positive strain did not showed either induction or any extra chromosomal species of the EfCIV583 element (Fig. S3A). Although these results clearly confirm that p1 is required for the induction and packaging of EfCIV583, they raise a further question of the packaging mechanism, since, as noted above, EfCIV583 does not encode a recognizable *terS* gene. We therefore considered the possibility that, like SaPIbov5 (Quiles-Puchalt *et al.*, 2014), EfCIV583 contains a packaging site that is recognized by the p1 terminase. Since the terminase packaging (*pac*) site for most phages is embedded in the *terS* coding sequence, we performed a blast search with the p1 *terS* coding sequence looking for a matching subsequence within EfCIV583. As shown in Fig. 4D we found such a

sequence and suggest that it represents the requisite packaging site for both phage and PICI, and therefore that EfCIV583 utilizes the p1 terminase for its packaging. This sequence, as occurs with the SaPIs (Quiles-Puchalt *et al.*, 2014), is located in an intergenic region of EfCIV583 (Fig. 4D). A BLAST search showed that this sequence was present in the *ter*S gene of two additional *E. faecalis* prophages and, remarkably, in a pneumococcal PICI, SpnCI6706B-1.

The existence of a functional Stl-like repressor of EfCIV583 replication, plus the failure of MC to induce the appearance of replicative forms of EfCIV583 in a non-lysogen suggested that, like the SaPI *stl,* the EfCIV583 repressor *rpr* is SOS-insensitive. We confirmed the SOS insensitivity of the EfCIV583 repressor by means of the β-lactamase reporter fusion, as shown in Fig. 3. As the most likely source of the putative de-repressor protein was the helper phage p1, we subcloned several segments of the phage, tested them for induction of EfCIV583 with the β-lactamase reporter fusion (Fig. S3B), and found that a 4.5 kb phage segment could relieve Rpr-mediated repression of the β-lactamase reporter fusion (Fig. S3C-D). We were then able to demonstrate that a single gene in this segment (EF0309) (Fig. S3E) was responsible, and that a p1 derivative with a mutation in this gene did not induce transfer of the element (Table 2). We then used affinity chromatography with a histidine tagged derivative of Rpr to test for the formation of a complex between the phage-encoded inducer and the EfCIV583-encoded repressor (Fig. 5A), as previously reported for the SaPIs (Tormo-Más *et al.*, 2010). These results suggested that the product of EF0309 acts directly on the repressor, as previously reported for the well characterized *dut*-specific induction of the SaPIs (Tormo-Más *et al.*, 2010; 2013). Interestingly, an orthology analysis of EF0309 revealed that it is highly likely to be the phage's *xis* gene, and it is referred to as *xis* hereafter.

Finally, the role of Xis in EfCIV583 induction was confirmed by the introduction of an in frame *xis* deletion in p1, which also eliminated EfCIV583 induction, mobilization and EfCIV583-mediated phage interference (Fig. 5B, 5C and Table 2).

In summary, EfCIV583 appears to closely resemble a typical SaPI, with p1 serving as its

helper phage. Our results however, are inconsistent with those of Matos, *et al.* (Matos *et al.*, 2013), who reported that, unlike the SaPIs, EfCIV583 can be SOS induced and that EfCIV583 and p1 are each packaged exclusively in small and large particles, respectively. In both cases, however, our data do not agree with those of Matos *et al.* (Matos *et al.*, 2013). First, we could not demonstrate SOS inducibility and we show, rather, that EfCIV583 is induced by the phage coded Xis protein (see Fig. 3 and 5). Second, since in the Southern blot analyses the p1 and EfCIV583 probes hybridized both with the EfCIV583-sized (small) and the phage-sized (large) DNA bands (Fig. 4C), our results suggest that phage DNA can be packaged in the EfCIV583-sized particles, and conversely the island DNA can also be packaged in full-sized phage particles. As with the SaPIs, these results indicate that packaging is not dependent on prohead size (Ubeda *et al.*, 2007b; Maiques *et al.*, 2007). Packaging of a significant proportion of phage DNA in the small particles, which would generate defective phages, could be responsible for the observed phage interference (Matos *et al.*, 2013; Frígols *et al.*, 2015).

Extension to other Gram-positive cocci

Genomic analyses

A key feature of the present study is that we considered it important to focus on species in which cohesive families of SaPI-like elements could be identified and in which these were the predominant form of phage related elements in the genus. This would be, first of all, parallel to the *S. aureus* situation and secondly would reinforce the concept that the biological success of such elements would be underlined by the existence of large – possibly exclusive – intrageneric families. Although the first PICI element was identified in *E. faecalis*, it subsequently became clear that *E. faecalis* does not contain a significant family of such elements. Thus, only four others have been identified and all are very closely related (Fig. S4, Table S4), three of them being in the same site as EfCIV583. Moreover, the orthology analysis of the EfCIV583 ORFs does not reveal membership in any family (see Table S5). Nevertheless, EfCIV583 shares not only most of the functional features of a typical SaPI, but

also the typical genome organization, indicating that that it is clearly a member of the overall PICI family, as described above, despite the atypical pattern of its orthologs.

By contrast, in the lactococci and the streptococci (especially *S. pneumoniae*), there does appear to be a series of elements that fits the genomic pattern described for the SaPIs. These elements, which form cohesive families on the basis of ortholog analysis (see below), could readily be separated from other types of inserted elements, occupying specific and unique chromosomal sites. In this report, we characterize these cohesive families of SaPI-like elements, the lactococcal and streptococcal PICIs.

Lactococcal PICIs.

Identification and genomic characterization. For the lactococci, we started with two reports describing six "prophage" genomes in *L. lactis* strain IL1403 (Chopin *et al.*, 2001; Bolotin *et al.*, 2001). Three of these, bIL309, bIL285 and bIL286, are typical prophage genomes, 35-44 kb in length and 3, bIL310, bIL311, and bIL312, are much smaller, 11-15 kb, and lack virion protein genes. Two of the three, bIL310 and bIL312 (but not bIL311) are apparently inducible since their DNA could be detected in lysates after mitomycin C-induction of the resident prophages present in strain IL1403 (Bolotin *et al.*, 2001). The genomic patterns of these three were highly similar and were also highly similar to the genomic patterns of the SaPIs, as noted above (see Fig. 1).

To identify similar elements in other lactococci, we used the KEGG orthology tool (Kanehisa and Goto, 2000). Here, we started with the phage-related element present in the strain CV56 (LlCIV56-1), which is similar to bIL310 in IL1403 (henceforth, LlCI-lL1403(I)), and generated orthology lists for all 22 ORFs in the island (Table S6). The orthologs are listed in decreasing order of similarity to the index gene, and the length of the list depends on how well the gene is conserved. Each gene in the list is linked to a KEGG map of the corresponding region of the organism's genome, and the KEGG map patterns are often highly informative with respect to the local genetic context. Indeed, visual inspection of the KEGG genome map usually enables the identification of such inserted elements. For example, as shown in Fig.

S5A and described above for the SaPIs, a PICI consists of a short set of genes transcribed in one direction and a longer set transcribed in the opposite direction, with the overall size being ~12-16 kb . The *int* gene is at or near the end of the shorter set, the *ter*S gene near the end of the longer set, within which are one or two large genes corresponding to *rep* and *pri*, and the divergence is flanked by regulatory genes corresponding to SaPI s*tl* and *str.* A typical example is illustrated in Fig. S5A. Visual scanning of the entire genome often reveals one or more other elements with this pattern or with the prophage pattern (Fig. S5B); in the staphylococci, streptoocci and lactococci, other types of inserted elements, aside from transposons and IS's, are very rare (although sets of genes with this overall pattern can be found, BLAST searches readily determine whether these are inserted elements such as PICIs or not). PICIs and prophages have small numbers of specific *att* sites and once these are identified, BLAST searches with the flanking genes reveal the unoccupied sites, as shown in Fig. S5A.

In the ortholog Tables, generally at least the first 10 orthologs are listed. Occasionally, orthologs are found in the absence of other PICI-related or prophage related genes. These are listed as "no Insert". Where there are fewer than 10, either all the matches are listed or prophage genes have begun to appear (and are listed), at which point the list is terminated. The left-hand column in the chart records the locus tag of the gene that has been identified by the KEGG orthology search. In the next 3 columns are the length of the hypothetical protein, its % similarity with the index protein, and the length of the overlap region between the two that was used for the similarity calculation. For each of the orthologs listed, the corresponding genome region was inspected to determine the type of insert, if any, in which the gene was located. This result is listed in the next column. If it resembles the genomic pattern of a typical PICI or prophage (see below), this is so indicated. "NI" (no insert) means that the flanking regions do not have the pattern of genes typical of PICIs, prophages, or other mobile elements, nor do the flanking genes resemble genes of mobile elements. The next 2 columns indicate the coordinates of the ends of the putative inserted element, and the

next, its size. Any relevant comments are in the next column. Several problematical elements, listed as “hybrid” probably represent hybrids between PICIs and prophages. They represent additional examples of possible recombinants and are included for completeness.

By this means, we identified a set of 26 SaPI-like elements in the lactococci, and observed that they occupy 7 different *att* sites (Table S7 and Fig. S6). These elements, like those in IL1403, have a number of features in common with the SaPIs: *i*) unique attachment (*att*) sites that are not also occupied by prophages; *ii*) the above-mentioned major point of transcriptional divergence flanked by regulatory genes (also a feature of temperate phages); *iii*) absence of bacteriophage morphogenetic and lytic genes; *iv*) size around 15 kb; and *v*) presence of primase (*pri*) and initiator (*rep*) protein genes, plus location and organization of the replication origin. Although the Pri-Rep proteins are often annotated, remarkably, as virulence related proteins (VapE) or as phage resistance proteins, the genes encoding these proteins are always homologous to the SaPI replication initiation genes (Ubeda *et al.*, 2007a) and are indicated as such in our maps. Of note is that fact that, as with the SaPIs, distinct families of PICI Pri-Rep proteins are encoded in these elements. Thus, in some cases the *pri-rep* genes are fused as is the case with some SaPIs. A subset of these newly identified lactococcal PICIs is illustrated in Fig. 6A. Their ORFs are colour coded to indicate putative functions, and, as can be seen, their organization corresponds closely to that seen in Figs. 1 and S5A.

A problem with the assembly and characterization of these element families is that the overall recombination frequency in lactoccoci is high and consequently, genomic rearrangements are common; in order to pinpoint possible rearrangements, it is especially important to have a family of closely related elements for comparison. An example of genotypic modifications is provided by the elements at the *mtl*D site in *L. lactis* (site I). Among the currently available sequenced genomes, four have PICIs at this site, NZ9000, MG1363, IL1403, and CV56. The first two (LlCINZ9000-1 and LlCIMG1363-1) are virtually identical in sequence as are the latter two (bIL312 and LlCICV56-1), and a comparison of the

first (LICINZ9000-1) and third (bIL310) elements reveals several areas of virtual identity separated by 3 major unshared segments, which are most likely to be insertions. At the extreme 5' end is a 4 kb unshared region that is also present at the insertion site in strain KW2 and, in a modified form including a transposon, in strain SK11, and is absent in strain KLDS.

As can be seen in the Table S6, for most of the 22 ORFs examined the top 10 orthologs belong to other putative PICIs. Prophages appear toward the end of the list, but mostly for those ORFs that are obviously phage-related (*int*, *pri, rep or terS*), and in all cases the phage ORFs have low similarity with the PICI gene. The other types of ORFs in the list include: *i*) accessory genes that were presumably inserted by some unknown type of recombination event; and *ii*) ORFs encoding hypothetical proteins (HP's).

HP analysis. As with most genetic units, the newly identified PICIs always contain many ORFs with unknown functions, whose putative products are annotated as "hypothetical proteins" (HPs). There are several orthology patterns among the HPs: some HPs are highly conserved and present in many different PICIs, but never elsewhere, whereas others are specific to one or two PICIs only. Prophage orthologs are rarely found, usually very far down when there is a long list. In these cases, the gene may have originated in a prophage or have been acquired by one. We suggest that HPs matching only conspecific PICIs have either been acquired since the divergence of PICIs from their ancestral element or have evolved *de novo*. Either underscores the long evolutionary independence of these elements. Although the HP ORFs found only in PICIs may be free to diverge unrestrictedly, we have not encountered any that have recently been fragmented by mutation and they nevertheless retain their exclusive membership in the PICI family, supporting the concept that the family is coherent and that its members have evolved independently and in concert.

Packaging of the lactococal elements. We have recently reported that some SaPIs have *cos* sites. These variant SaPIs, of which SaPIbov5 is the prototype (Viana *et al.*, 2010), are induced by *cos* phages, that share the same *cos* site, and are efficiently packaged by these

phages, leading to high frequency intra- and inter-generic transfer (Chen *et al.*, 2015a; Quiles-Puchalt *et al.*, 2014). Since all three phages present in *L. lactis* IL1403 are *cos* phages, and since the bIL310 and bIL312 elements present in this strain can be packaged after induction of the resident prophages (Bolotin *et al.*, 2001), we hypothesised that these elements utilise for packaging the same strategy used by SaPIbov5, namely carrying the *cos* sequence present in one of the prophages. To test this, we analysed the sequence contained between the *hnh* and the *ter*S phage genes. In many phages from Gram-positive bacteria, this region contains the phage *cos* site (Quiles-Puchalt *et al.*, 2014). Examination of the bIL310 sequence revealed a putative *cos* site identical to the putative phage bIL286 *cos* site (Fig. S7). Related with the other two phages present in the *L. lactis* IL1403 strain, bIL285 and bIL309, other PICIs also share the putative *cos* sites present in these phages (Fig. S7). To test these phage *cos* sites for function, we cloned the bIL286 and bIL310 putative *cos* sites to a plasmid, pAGEnt (Table S3), which was not transferrable after induction of the resident prophages present in strain IL1403, and found that the cloned *cos* sites, enabled transfer of the plasmids by the bIL286 phages (Table S8). This result confirmed the identity of these sequences as *cos* sites.

Pneumococcal and other streptococcal PICIs

A similar search in 8 of the now very large number of pneumococcal genomes revealed essentially the same pattern of SaPI-like elements as in the lactococci. Diagrams of the genomes of ten of these are presented in Fig. 6B and Table S7. We have also done an orthology analysis for all of the ORFS in one of these, in strain 670-6B, at Mb 0.01 (*ych*F site). The results of this analysis, shown in Table S9, are similar to those obtained in the analysis of the lactococcal elements (Table S6), confirming the idea that the PICI genes belong exclusively to the PICI elements. The difference with the lactococcal elements is that in the pneumococcal PICIs some orthologs are found frequently among other streptococci and orthologs that do not belong to any inserted element are sometimes found in a wide variety of other genera – which may be a result of the high level of transformation

competence in pneumococci and related streptococci (Straume *et al.*, 2015). As with the lactococci, with one exception, prophages do not appear or only appear far down the list. The exception is gene SP670_0020, an ORF encoding a 55 aa hypothetical protein, for which the first 3 orthologs are 3 different prophages. These prophages are at different sites in 3 different host strains, and none of these sites contains a PICI. This ORF presumably represents a very rare episode of gene exchange between PICIs and prophages.

The appearance of other streptococci in this list suggests either that the PICI lineage split from an ancestral element before the differentiation of streptococcal species, or reflects horizontal transfer of these elements. Transfer, however, need not have been mediated by a helper phage, since these streptococci are transformation competent, and pneumococci have a habit of extruding their DNA under certain conditions (Claverys *et al.*, 2007). It is especially notable that a site 3' to the enolase gene is occupied by closely related PICIs in *S. suis* and *S. oligofermentans*, as well as in *S. pneumoniae* (see Fig. 6B and Table S7).

Examination of KEGG genome maps obtained with orthologs of the *rep* gene of LlCI1403(I) revealed similar elements in many species. In the pneumococci, only elements of the PICI type were identified in this screen. We thus considered it likely that these elements would belong to a pneumococcus-specific family analogous to those in *S. aureus* and in the lactococci. We have identified eleven different *att* sites for the pneumococcal PICIs. As our analysis of the pneumococcal genomes has not been as extensive as that of the lactococcal genomes, there may well be other sites. Again, however, we have not found any intact or defective prophages at any of these sites. Other streptococcal species contain PICIs that are closely related to those in pneumococci (Fig. 6B; Table S7).

The streptococcal PICIs have been studied in some detail by McShan and coworkers (Scott *et al.*, 2012) who have demonstrated that one frequently occupied insertion site in *S. pyogenes* is between the DNA repair genes, *mutS* and *mutL* at ~1.75 MB in the *S. pyogenes* genome (Scott *et al.*, 2008). PICI-like elements at this site block transcription of the downstream *mut*L but excise during exponential growth, restoring transcription, and re-insert

when stationary phase is reached (Scott *et al.*, 2008). Unlike the typical *S. pyogenes* PICIs, the PICI-like elements at the *mut*S/L site appear to have SOS-sensitive repressors and to lack a *terS* homolog, and therefore, their primary function may be gene regulation rather than transfer. Nguyen and McShan (Nguyen and McShan, 2014) have suggested that other streptococcal PICIs may also have regulatory roles.

From an evolutionary perspective, as can be seen from the ortholog analysis of SpnCI670-6B (table S9), genes from *S. pyogenes* PICIs located between *mutS* and *mutL* are in the same list as genes from *S. pyogenes* PICIs located elsewhere, suggesting that the streptococcal PICIs belong to a co-ancestral family that has branched into at least two subfamilies – one involved in regulation, encoding a SOS inducible repressor, lacking a TerS homolog, and integrating between the DNA repair genes *mutS* and *mutL*, and the other in gene transfer, encoding an SOS-insensitive repressor, a TerS homolog, and integrating elsewhere.

In summary, the PICIs within each of these two genera are closely related, suggesting that they are a coherent family that does not contain genetic units of any other type. Very importantly, elements containing additional modules or other complicating segments have not been found in these species – which is why the families are referred to as cohesive.

Accessory genes.

Many of the PICIs carry identifiable genes or other elements that do not appear to be involved in the PICI lifecycle (accessory genes – see Table S7), of which there seem to be at least 4 types – *i*) transposons, IS sequences and other obviously inserted elements, even including, in at least one case, SpnCI-TCH8431-2, a possibly intact plasmid; *ii*) genes involved in helper phage interference – well-characterized in *S. aureus* but as yet defined in only one PICI, EfCIV583; *iii*) genes that contribute in possibly important ways to the host organism. In the SaPIs, most of these encode superantigens or other virulence genes (Lindsay *et al.*, 1998; Viana *et al.*, 2010; Ubeda *et al.*, 2003), many of which are carried exclusively by the SaPIs. In the lactococci and streptococci, this class includes genes for

diverse metabolic activities and resistances to antibiotics, bacteriocins and bacteriophages; superantigen genes have not been observed; and *iv*) phage-related genes that are not standard and seem accidental – including occasional capsid, phage protease, and regulatory genes. One phage-related regulatory gene, an *ltr*C homolog, is of interest since *ltr*C is an essential phage gene that controls late phage gene transcription and always occurs just 5' to *terS* in prophage genomes (Quiles-Puchalt *et al.*, 2013; Ferrer *et al.*, 2011). However, among the PICIs, the *ltr*C homolog is always 3' to *terS*, is sometimes duplicated, and sometimes occurs without any *terS* homolog. In the SaPIs, *terS* is adjacent to the phage interference module, one of whose targets is *ltr*C (Ram *et al.*, 2014). Perhaps the PICI-carried *ltrC* homologs, which are <50% similar their prophage counterparts, are involved in phage interference.

In summary, the PICIs, like other non-essential genomic elements, suffer diverse types of rearrangements, including plasmids, insertions of transposons, IS sequences, etc. Such adventitious insertions occur, of course, in any non-essential region, and are not unique to the PICIs.

Discussion

During the course of this study it has gradually become apparent that prophages and PICIs have evolved in much more interesting ways than has generally been realized. Remarkably, the genomes of at least the two genera highlighted here, as well as of the staphylococci, contain one or more highly conserved and highly functional lineages of a novel family of mobile genetic elements, the PICIs, which form special archipelagos in the coccal sea. Their special role, which has been defined in *S. aureus,* is to connect with functional prophages which induce them to reproduce and spread to other individual cells. The prototype of these lineages is the SaPIs, which lead a highly productive existence within the chromosomes of the staphylococci, enabling the phage-mediated spread of superantigens and other important bacterial products (Chen *et al.*, 2015a; 2015b; Chen and Novick, 2009). Additionally, they are of considerable benefit to their bacterial host cells as they interfere with the reproduction of

infecting phages and increase the survival of host cells attacked by phages (Ram *et al.*, 2012). The success of this evolutionary strategy is evidenced by the widespread existence of elements of the same type in related species.

Since they contain genes that are recognizably phage-related, these elements have been universally annotated as defective prophages. What distinguishes them is that they form cohesive families whose members are closely related both genetically and structurally and are only distantly related to other prophage-derived genetic segments. This is most clearly demonstrated by two of their common features: firstly, they occupy unique *att* sites that are not occupied by either intact or defective prophages, and vice versa. Secondly, many of their genes are conserved within their family but are not detected in other genomes. For the SaPIs, probably the most important of these genes are *ppi* plus those genes located in operon I, of which only t*erS* is phage-related (Ubeda *et al.*, 2007b). The other genes have key functions in SaPI biology and lack prophage homologs (Novick and Ram, 2015). They were either remodelled from genes of unknown origin, or evolved de novo within the SaPIs. Along with *terS*, they serve to define the SaPI-helper phage interaction and, along with the relatively low frequency of general recombination, have served to maintain the long-term separation between SaPIs and prophages. This separation, applied to the PICIs as well as the SaPIs, is also clearly demonstrated by analyzing the orthology patterns of ORFs encoding hypothetical proteins (HPs). Not only are most of the HPs not detected outside of the family, but as it is not known whether they are functional or even whether their genes are translated, their presence, with rare exceptions, as noted, can simply not be explained as acquisition by horizontal transfer – *i.e.,* they must have evolved *in situ*. Consequently, their relatedness serves as an index of the relatedness of the elements carrying them. The same is true of prophages, which also occur as families and also encode many HPs, which follow the same patterns as those of the PICIs.

In summary, the KEGG analyses indicate that the lactococcal and streptococcal PICIs plus EfCIV583 are at most very distantly related to one another or to the SaPIs, yet they share a

common genome organization and content. This suggests first that the PICIs in these different genera are probably not co-ancestral and therefore must have originated independently, after the diversification of the genera, and thus represent a remarkable example of convergent evolution. This result also indicates that the PICI genomic organization has a powerful selective value, since the PICIs are far more common than any other type of phage-related element in the 3 genera analyzed in detail, staphylococci, streptococci and lactococci. Moreover, it appears that two types of phage-related elements are in the vast majority – the prophages themselves and the SaPIs/PICIs, of which there may be more than a single lineage in some genera. Other types of phage-related elements can be identified, but these are few and far between.

Acknowledgements

This work was supported by grants BIO2011-30503-C02-01, Eranet-pathogenomics PIM2010EPA-00606 and Consolider-Ingenio CSD2009-00006 from the Ministerio de Ciencia e Innovación (MICINN), and grant MR/M003876/1 (Medical Research Council, UK) to J.R.P., and by National Institutes of Health Grant R01AI022159 to R.P.N. and J.R.P.

Conflict of Interest.

The authors declare no conflict of interest.

References

Bolotin A, Wincker P, Mauger S, Jaillon O, Malarme K, Weissenbach J, *et al.* (2001). The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. lactis IL1403. *Genome Res* **11**: 731–753.

Charpentier E, Anton AI, Barry P, Alfonso B, Fang Y, Novick RP. (2004). Novel cassette-based shuttle vector system for gram-positive bacteria. *Appl Environ Microbiol* **70**: 6076–6085.

Chen J, Carpena N, Quiles-Puchalt N, Ram G, Novick RP, Penadés JR. (2015a). Intra- and inter-generic transfer of pathogenicity island-encoded virulence genes by cos phages. *ISME J* **9**: 1260–1263.

Chen J, Novick RP. (2009). Phage-mediated intergeneric transfer of toxin genes. *Science* **323**: 139–141.

Chen J, Ram G, Penadés JR, Brown S, Novick RP. (2015b). Pathogenicity island-directed

563 transfer of unlinked chromosomal virulence genes. *Mol Cell* **57**: 138–149.

564 Chopin A, Bolotin A, Sorokin A, Ehrlich SD, Chopin M. (2001). Analysis of six prophages in
565 *Lactococcus lactis* IL1403: different genetic structure of temperate and virulent phage
566 populations. *Nucleic Acids Res* **29**: 644–651.

567 Claverys J-P, Martin B, Håvarstein LS. (2007). Competence-induced fratricide in streptococci.
568 *Mol Microbiol* **64**: 1423–1433.

569 Duerkop BA, Clements CV, Rollins D, Rodrigues JLM, Hooper LV. (2012). A composite
570 bacteriophage alters colonization by an intestinal commensal bacterium. *Proc Natl Acad Sci*
571 *USA* **109**: 17621–17626.

572 Ferrer MD, Quiles-Puchalt N, Harwich MD, Tormo-Más MÁ, Campoy S, Barbé J, *et al.* (2011).
573 RinA controls phage-mediated packaging and transfer of virulence genes in Gram-positive
574 bacteria. *Nucleic Acids Res* **39**: 5866–5878.

575 Frígols B, Quiles-Puchalt N, Mir-Sanchis I, Donderis J, Elena SF, Buckling A, *et al.* (2015).
576 Virus Satellites Drive Viral Evolution and Ecology. *PLoS Genet* **11**: e1005609.

577 Hamilton CM, Aldea M, Washburn BK, Babitzke P, Kushner SR. (1989). New method for
578 generating deletions and gene replacements in *Escherichia coli*. *J Bacteriol* **171**: 4617–4622.

579 Kanehisa M, Goto S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic*
580 *Acids Res* **28**: 27–30.

581 Lindsay JA, Ruzin A, Ross HF, Kurepina N, Novick RP. (1998). The gene for toxic shock
582 toxin is carried by a family of mobile pathogenicity islands in *Staphylococcus aureus*. *Mol*
583 *Microbiol* **29**: 527–543.

584 Maiques E, Ubeda C, Tormo MA, Ferrer MD, Lasa I, Novick RP, *et al.* (2007). Role of
585 staphylococcal phage and SaPI integrase in intra- and interspecies SaPI transfer. *J Bacteriol*
586 **189**: 5608–5616.

587 Matos RC, Lapaque N, Rigottier-Gois L, Debarbieux L, Meylheuc T, Gonzalez-Zorn B, *et al.*
588 (2013). *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits.
589 *PLoS Genet* **9**: e1003539.

590 Nguyen SV, McShan WM. (2014). Chromosomal islands of *Streptococcus pyogenes* and
591 related streptococci: molecular switches for survival and virulence. *Front Cell Infect Microbiol*
592 **4**: 109.

593 Novick RP, Christie GE, Penadés JR. (2010). The phage-related chromosomal islands of
594 Gram-positive bacteria. *Nat Rev Microbiol* **8**: 541–551.

595 Novick RP, Ram G. (2015). The Floating (Pathogenicity) Island: A Genomic Dessert. *Trends*
596 *in Genetics* 1–13.

597 O'Callaghan CH, Morris A, Kirby SM, Shingler AH. (1972). Novel method for detection of
598 beta-lactamases by using a chromogenic cephalosporin substrate. *Antimicrob Agents*
599 *Chemother* **1**: 283–288.

600 Paulsen IT, Banerjei L, Myers GSA, Nelson KE, Seshadri R, Read TD, *et al.* (2003). Role of
601 mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* **299**:
602 2071–2074.

603 Penadés JR, Christie GE. (2015). The Phage-Inducible Chromosomal Islands: A Family of
604 Highly Evolved Molecular Parasites. *Annual Review of Virology* **2**: 181–201.

605 Quiles-Puchalt N, Carpena N, Alonso JC, Novick RP, Marina A, Penadés JR. (2014).
606 Staphylococcal pathogenicity island DNA packaging system involving cos-site packaging and
607 phage-encoded HNH endonucleases. *Proc Natl Acad Sci USA* **111**: 6016–6021.

608 Quiles-Puchalt N, Tormo-Más MÁ, Campoy S, Toledo-Arana A, Monedero V, Lasa I, *et al.*
609 (2013). A super-family of transcriptional activators regulates bacteriophage packaging and
610 lysis in Gram-positive bacteria. *Nucleic Acids Res* **41**: 7260–7275.

611 Ram G, Chen J, Kumar K, Ross HF, Ubeda C, Damle PK, *et al.* (2012). Staphylococcal
612 pathogenicity island interference with helper phage reproduction is a paradigm of molecular
613 parasitism. *Proc Natl Acad Sci USA* **109**: 16300–16305.

614 Ram G, Chen J, Ross HF, Novick RP. (2014). Precisely modulated pathogenicity island
615 interference with late phage gene transcription. *Proc Natl Acad Sci USA* **111**: 14536–14541.

616 Ruzin A, Lindsay J, Novick RP. (2001). Molecular genetics of SaPI1--a mobile pathogenicity
617 island in Staphylococcus aureus. *Mol Microbiol* **41**: 365–377.

618 Scott J, Nguyen SV, King CJ, Hendrickson C, McShan WM. (2012). Phage-like
619 *Streptococcus pyogenes* chromosomal islands (SpyCI) and mutator phenotypes: control by
620 growth state and rescue by a SpyCI-encoded promoter. *Front Microbiol* **3**: 317.

621 Scott J, Thompson-Mayberry P, Lahmamsi S, King CJ, McShan WM. (2008). Phage-
622 associated mutator phenotype in group A streptococcus. *J Bacteriol* **190**: 6290–6301.

623 Straume D, Stamsås GA, Håvarstein LS. (2015). Natural transformation and genome
624 evolution in *Streptococcus pneumoniae*. *Infect Genet Evol* **33**: 371–380.

625 Subedi A, Ubeda C, Adhikari RP, Penadés JR, Novick RP. (2007). Sequence analysis
626 reveals genetic exchanges and intraspecific spread of SaPI2, a pathogenicity island involved
627 in menstrual toxic shock. *Microbiology* **153**: 3235–3245.

628 Tormo MA, Ferrer MD, Maiques E, Ubeda C, Selva L, Lasa I, *et al.* (2008). *Staphylococcus*
629 *aureus* pathogenicity island DNA is packaged in particles composed of phage proteins. *J*
630 *Bacteriol* **190**: 2434–2440.

631 Tormo-Más MÁ, Donderis J, García-Caballer M, Alt A, Mir-Sanchis I, Marina A, *et al.* (2013).
632 Phage dUTPases control transfer of virulence genes by a proto-oncogenic G protein-like
633 mechanism. *Mol Cell* **49**: 947–958.

634 Tormo-Más MÁ, Mir I, Shrestha A, Tallent SM, Campoy S, Lasa I, *et al.* (2010). Moonlighting
635 bacteriophage proteins derepress staphylococcal pathogenicity islands. *Nature* **465**: 779–
636 782.

637 Ubeda C, Barry P, Penadés JR, Novick RP. (2007a). A pathogenicity island replicon in
638 *Staphylococcus aureus* replicates as an unstable plasmid. *Proc Natl Acad Sci USA* **104**:
639 14182–14188.

640 Ubeda C, Maiques E, Barry P, Matthews A, Tormo MA, Lasa I, *et al.* (2008). SaPI mutations
641 affecting replication and transfer and enabling autonomous replication in the absence of
642 helper phage. *Mol Microbiol* **67**: 493–503.

643 Ubeda C, Maiques E, Knecht E, Lasa I, Novick RP, Penadés JR. (2005). Antibiotic-induced
644 SOS response promotes horizontal dissemination of pathogenicity island-encoded virulence
645 factors in staphylococci. *Mol Microbiol* **56**: 836–844.

646 Ubeda C, Maiques E, Tormo MA, Campoy S, Lasa I, Barbé J, *et al.* (2007b). SaPI operon I is
647 required for SaPI packaging and is controlled by LexA. *Mol Microbiol* **65**: 41–50.

648 Ubeda C, Tormo MA, Cucarella C, Trotonda P, Foster TJ, Lasa I, *et al.* (2003). Sip, an
649 integrase protein with excision, circularization and integration activities, defines a new family
650 of mobile *Staphylococcus aureus* pathogenicity islands. *Mol Microbiol* **49**: 193–210.

651 Viana D, Blanco J, Tormo-Más MÁ, Selva L, Guinane CM, Baselga R, *et al.* (2010).
652 Adaptation of *Staphylococcus aureus* to ruminant and equine hosts involves SaPI-carried
653 variants of von Willebrand factor-binding protein. *Mol Microbiol* **77**: 1583–1594.

654

655

Figure legends

Figure 1. Genome maps for PICIs and related elements. The originally identified PICIs from *E. faecalis* and *L. lactis* compared with SaPI1 and SaPIbov1. Additional details for these elements can be found in Table S7.

Figure 2. Characterization of EfCIV583-encoded Int protein. (A) Schematic representation of the EfCIV583-dependent excision and circularization processes. The relevant genes, genetic markers and PCR primer binding sites are shown. (B) Detection of EfCIV583 excision and circularization. DNA from *E. faecalis* VE14089 was extracted and PCR-amplified using specific primers (see scheme in A) recognizing the external and internal sequence of the element (integration: I), primers recognizing the flanking sequences (excision: E) or PCR-amplified using a pair of primers set divergently at both termini of the island (circularization: C). M: molecular weight marker. (C) Constructs used for test site-specific integration in *E. coli*. Top: pCN51 derivatives containing the EfCIV583 att_{PI}-*int* gene. Bottom: thermosensitive derivatives of pMAK700 carrying the att_C from the *E. faecalis* chromosome. The relevant genetic markers and restriction enzyme sites are shown. (D) Plasmid DNA was isolated from overnight cultures grown at 37ºC (for strains carrying uniquely the pCN51 or pMAK700 derivatives) or at 43ºC (for strain carrying both plasmids), in presence of ampicillin (pCN51) or chloramphenicol (pMAK700 derivative and cointegrative plasmid). Plasmids were digested with *Bam*HI (pMAK700 derivatives) or with *Sal*I (pCN51 derivatives and cointegrative plasmids).

Figure 3. Characterization of the EfCIV583-encoded StI repressor. (A) Schematic representation of the different *blaZ* transcriptional fusions. The relevant genes are shown. (B) *S. aureus* RN4220 strains containing the plasmids represented in panel A were assayed for β-lactamase activity under standard conditions, or after MC induction. Samples were normalized for total cell mass. Values presented are the averages (±SD) of three independent assays.

Figure 4. Induction of EfCIV583 by mitomycin C (MC). MC (1 μg/ml) was added to a

culture of *E. faecalis* VE14089 (EfCIV583-positive) or *E. faecalis* VE14089 ∆p1, followed by incubation at 32ºC. Samples were removed at the indicated time points and used to prepare minilysates, which were resolved on a 0.7% agarose gel (A), and Southern blotted with an EfCIV583 probe (B). In panel C is a stained gel and a Southern blot of DNA extracted from phage particles in a lysate of an MC-treated culture of VE14089. (D) The putative p1 *pac* site (coloured in red) is embedded in the *ter*S gene (coloured in green). Its homolog in EfCIV583 is located between two genes (coloured in blue). See text for explanation. Bottom: the predicted EfCIV583 and p1 *pac* sites are aligned using ClustalW2.

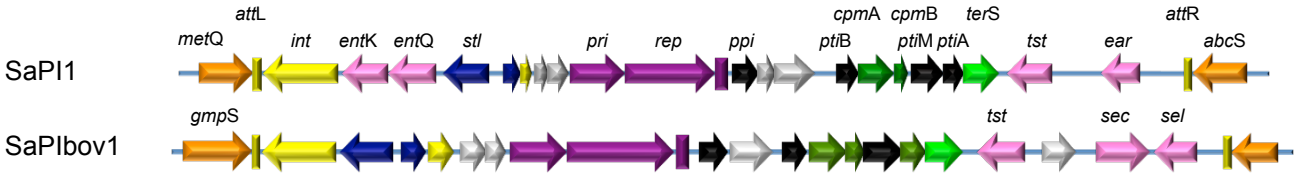
Figure 5. Identification of the EfCIV583 inducer. (A). Affinity chromatography of p1 EF0309 using His6-$Rpr_{EfCIV583}$. *E. coli* strain expressing the EF0309 / His6-$Rpr_{EfCIV583}$ pair was IPTG-induced and, after disruption of the cells, the expressed proteins were applied to a Ni2+ agarose column and eluted. The presence of the different proteins was monitored in the load (lanes E), flow-through, wash and elute fractions (P) by Coomassie staining. The identity of the purified proteins was determined by in-gel enzymatic digestion and mass fingerprinting. It is assumed that the presence of Xis in the eluate represents an Rpr-Xis complex. (B). MC (1 μg/ml) was added to a culture of *E. faecalis* JP11028 (EfCIV583/p1-positive) or *E. faecalis* JP13142 (JP11028 p1∆*xis*), followed by incubation at 32ºC. Samples were removed at the indicated time points and used to prepare minilysates, which were resolved on a 0.7% agarose gel, and Southern blotted with an EfCIV583 probe. CCC: closed circular form. (C) EfCIV583 interference with phage reproduction. Approximately 10^8 bacteria (carrying or not the EfCIV583 element) were infected with 400 plaque forming units (p.f.u.) of phage ϕ1 or phage p1 ∆*xis*, plated on phage bottom agar, and incubated 24h at 32 °C. Plates were stained with 0.1% triphenyl tetrazolium chloride in TSB media and photographed.

Figure 6. PICI genomes. (A) *Lactococcus lactis* PICI genomes identified by searching with b3IL10 genes, arranged by *att* sites. Also shown is a highly unusual element from *L. lactis* KF147, which may be related to ICE elements and plasmids, since it has an integrase plus putative plasmid replication and segregation genes. It could be confused with a PICI, except

710 that its transcriptional organization does not fit. (B) Genomes of PICIs of *S. pneumoniae* and
711 other streptococci. The colour coding of the PICI genes is the same as in Fig. 1.

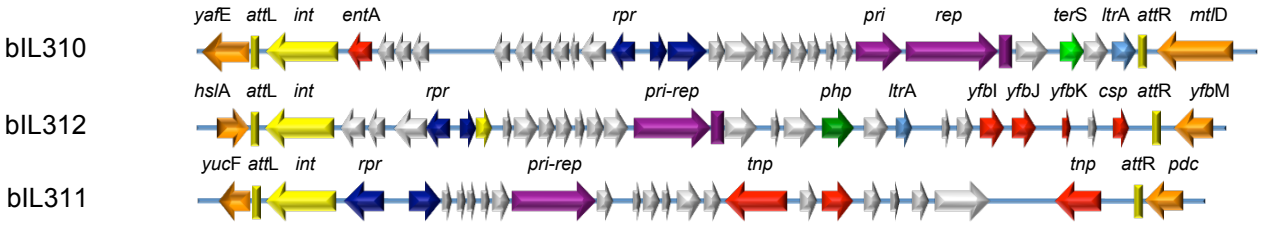
712

Staphylococcus aureus



Enterococcus faecalis

Lactococcus lactis



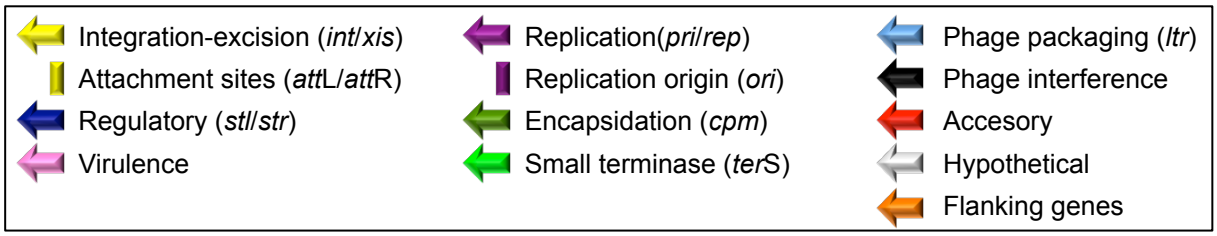


Figure 2

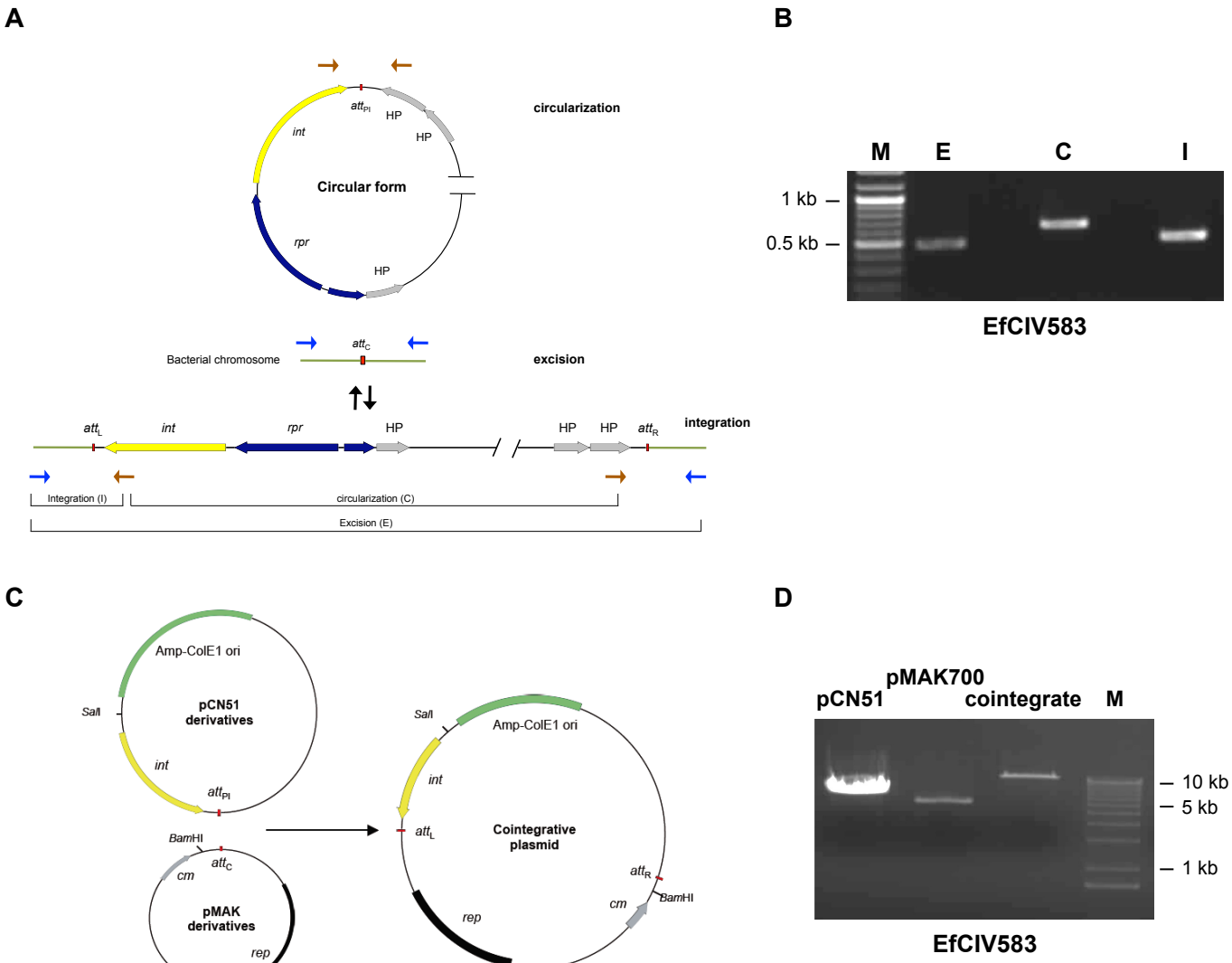
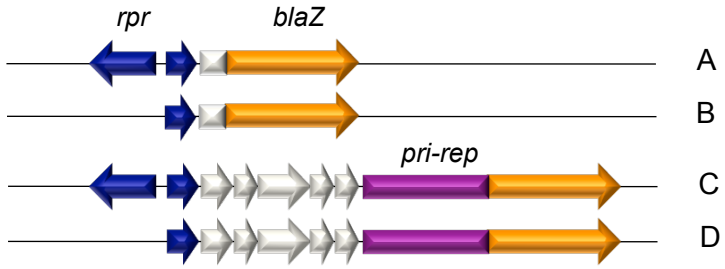
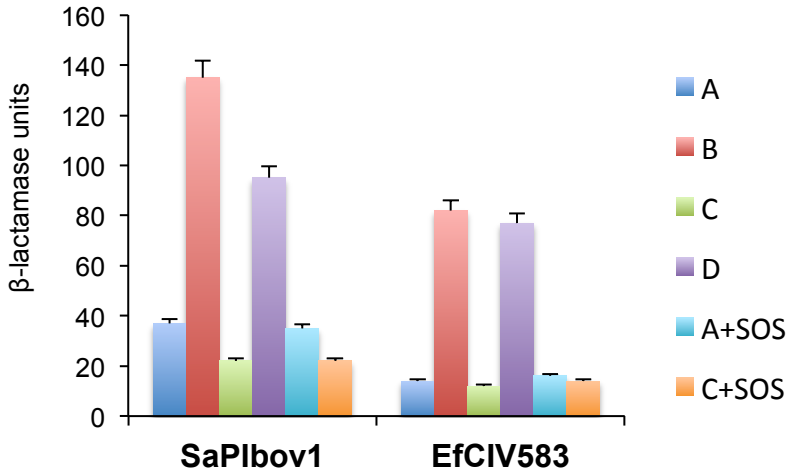

A
circularization
Circular form
int
rpr
HP
Bacterial chromosome
excision
integration
Integration (I)
circularization (C)
Excision (E)
B
M
E
C
I
1 kb
0.5 kb
EfCIV583
C
Amp-ColE1 ori
pCN51 derivatives
SalI
int
BamHI
cm
pMAK derivatives
rep
Cointegrative plasmid
D
pMAK700
pCN51
cointegrate
M
10 kb
5 kb
1 kb
EfCIV583

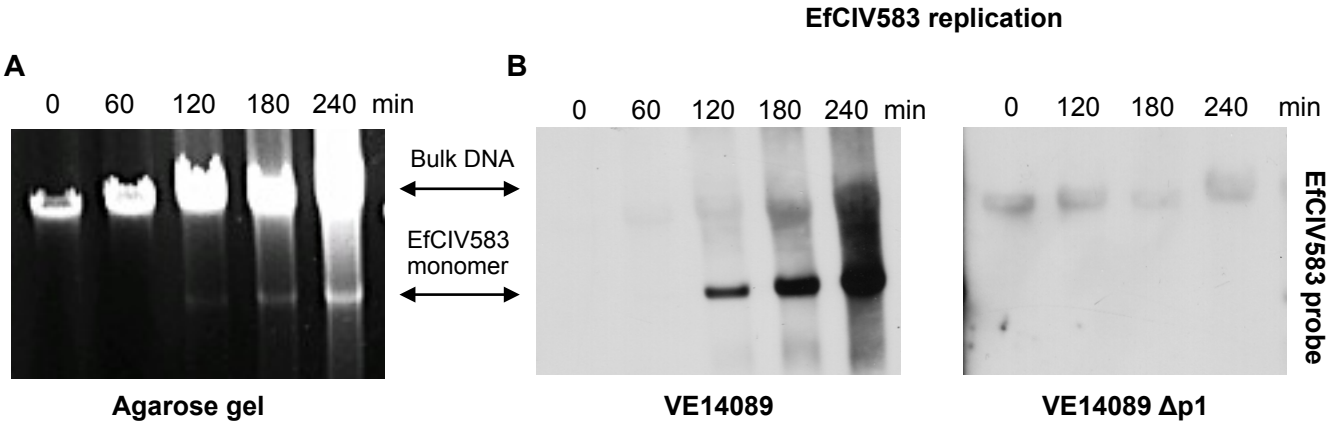
Figure 3

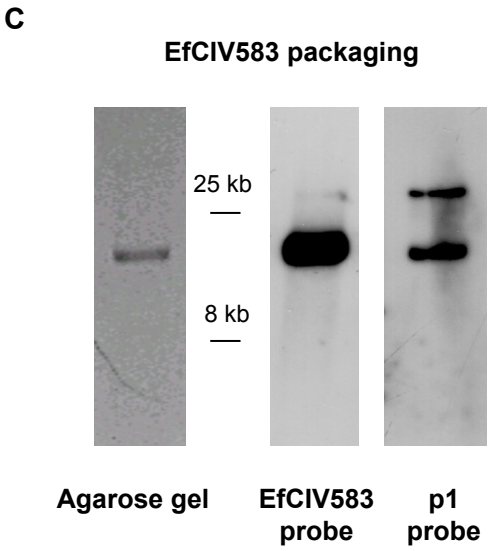
A


rpr
blaZ
pri-rep
A
B
C
D

B


β-lactamase units
160
140
120
100
80
60
40
20
0
SaPIbov1
EfCIV583
A
B
C
D
A+SOS
C+SOS





D

Phage p1 *terS*

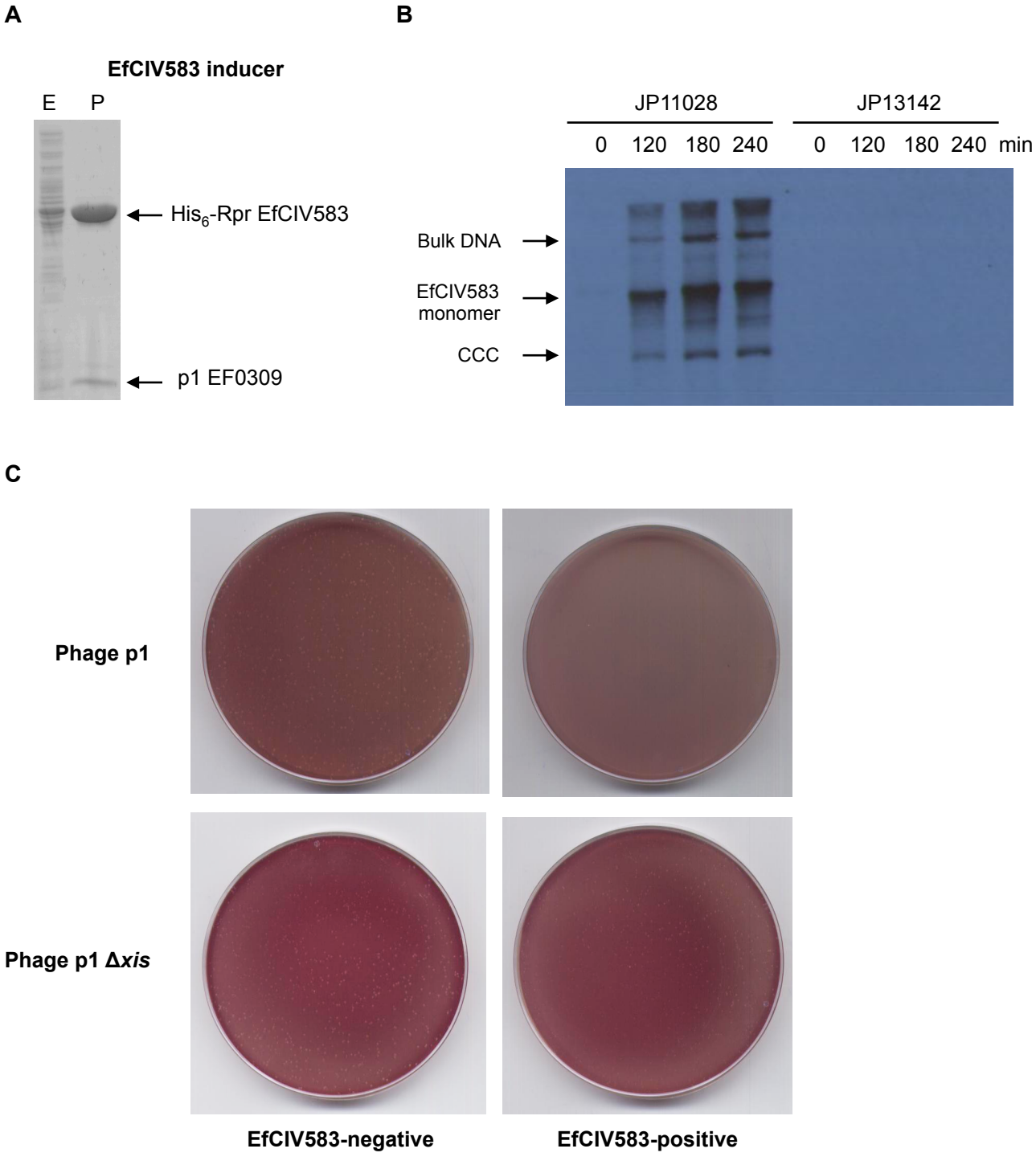
ATGGCGAAGTACACAGAGTGGCTAACCGAGGAAGGGTTAATCAAAATCGAAGGAT
GGGCACGAGATGGCCTCATCGATAAGCAGATAGCCCAGAATATCGGGGTAGCTTA
TTCAACTTTCAGAGAATGGGT**AAAAAAATTTCCGGCACTTTCGGCATCCTTAAAA**
AAGGGAAAAGAAGTCGTTGATAGACAAGTTGAAAATGCTTTATTTAAGAGTGCTA
CAGGCTACGAATATACCGAAGTTACAGAAGAACTAACAGAGAACGGTATGGAAAT
TACAAAAAAGGTTAAAAAACAAGTAGCTCCTAATCCAACTGCAGCTATTTTCTGG
TTGAAAAATAGAAAGCCGGACGAATGGCGAGATCGAAAAGAAACTGAAGTTTCAG
GCGTGCTTAATATCTCCGATGCAGCTGTCGAAATCGAGCAATTTTTCGAGGATGA
TTCTACATGA

EfCIV583

ATGGATAAAGTATACATTGATAACAGTAAAAAAACGGAAGTGGTAGAACTACCA
AAATTTGGGGAAGTTAAGTTGATTGTAAAAGATGGGAAGGTAGTAAAATACGAT
ACAATTACATCGCATGTTTTACCTAAGAATTAGCAGTCTAAACTATCTATTAGG
CGAAAAAAGTGCTATAATAAAGACAGTTAAATAAGTATATCTTGCTATTCTAGC
GCTATGTATCTTACTAGGAAGAACCAGAGGATATGAGACAGCTTACTCTTAATT
GAGTGAGGTCTTGTGTCCTCTTTTTTGTTTAAAATCTGACAAATCAGACACCT**A**
AAAAAGTCGGCACTTTCGGCAGCCTAAAAAAGACAATTCAGGAAAGAGAGGAAC
GTTGTCAATGGATCCATTTGATGAAATTGTTATAAAAAACTATCTCTCATTTCC
GTTAGTGTTGCAAAAGTTGCTTTCAATAAAAAAATTAAGGCGTATTGAATTTTT
CAGCCAAAATATGGCAACGCATATTGTGTTTGATGAGTTAGGAATGACTACAAA
CAAAAACATTTTGAGCGGTTTTGGGCTACTCTTAAGCCTTTAGAGCAACAATTG
CTCATTAGGCGATTCAAGTATAAAGAAGAGGTTAATTGCCCTCACAGGCTTATA
GAGAGCGTATTAGATGAGATTGAAGAGATAGAAACAGCTATTTGTTTGATGGAA
AATATAGAATTAGAAGAAAACGAGCTTTCCGACGATGTAGAAGAAAACTTAGAG
AGGATGTGTGACTTCTTTGCTTTATGA

Alignment of the putative *pac* sites

```
p1          TAAAAAAATTTCCGGCACTTTCGGCATCCTTAAAAAAG
EfCIV583    ---TAAAAAAGTCGGCACTTTCGGCAGCCTAAAAAAG-
               :****::  ********************* ***:*****.
```


A
EfCIV583 inducer
E
P
His6-Rpr EfCIV583
p1 EF0309
B
JP11028
JP13142
0 120 180 240
0 120 180 240 min
Bulk DNA
EfCIV583 monomer
CCC
C
Phage p1
Phage p1 Δxis
EfCIV583-negative
EfCIV583-positive

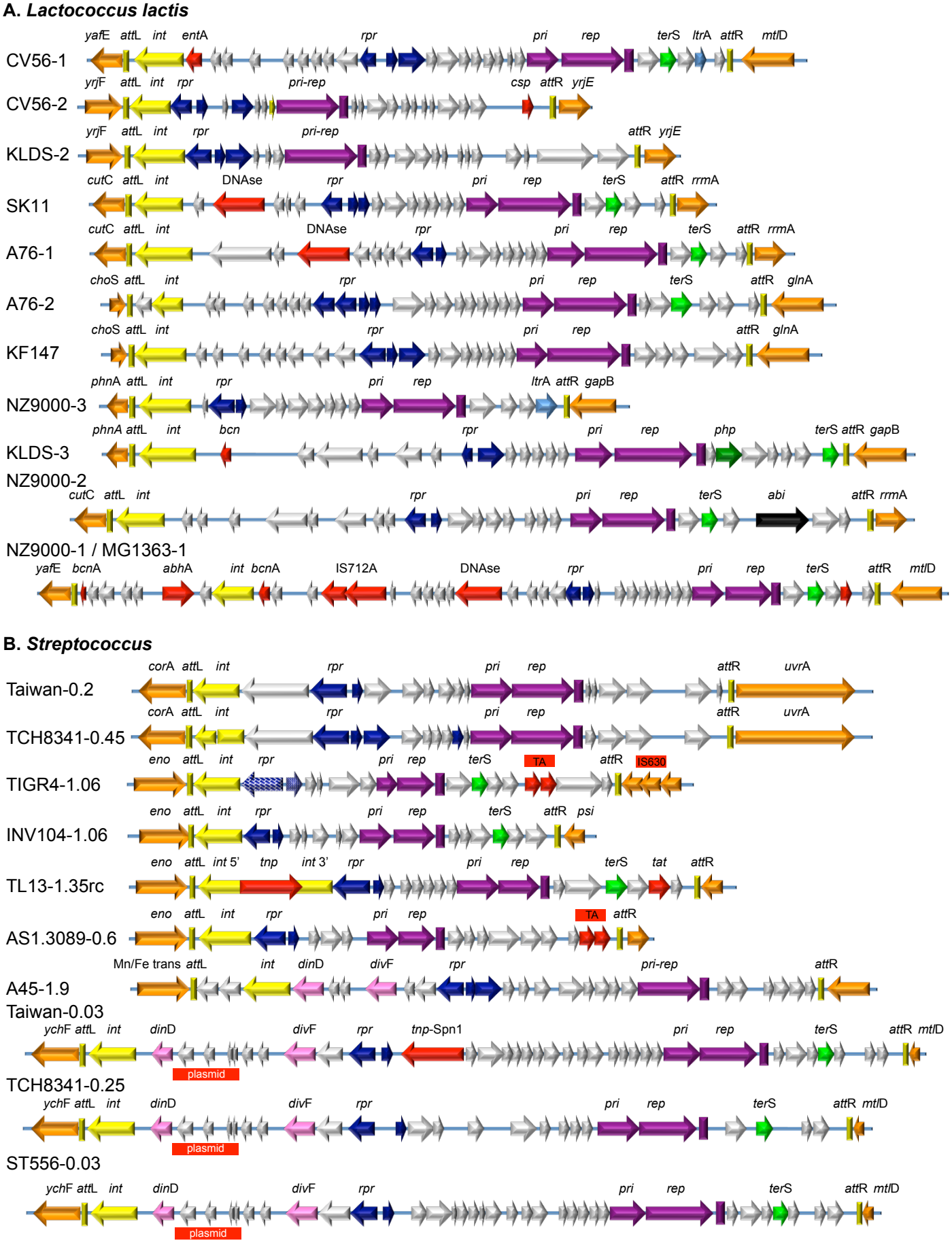

A. Lactococcus lactis
CV56-1
yafE attL int entA rpr pri rep terS ltrA attR mtlD
CV56-2
yrjF attL int rpr pri-rep csp attR yrjE
KLDS-2
yrjF attL int rpr pri-rep attR yrjE
SK11
cutC attL int DNAse rpr pri rep terS attR rrmA
A76-1
cutC attL int DNAse rpr pri rep terS attR rrmA
A76-2
choS attL int rpr pri rep terS attR glnA
KF147
choS attL int rpr pri rep attR glnA
NZ9000-3
phnA attL int rpr pri rep ltrA attR gapB
KLDS-3
phnA attL int bcn rpr pri rep php terS attR gapB
NZ9000-2
cutC attL int rpr pri rep terS abi attR rrmA
NZ9000-1 / MG1363-1
yafE bcnA abhA int bcnA IS712A DNAse rpr pri rep terS attR mtlD
B. Streptococcus
Taiwan-0.2
corA attL int rpr pri rep attR uvrA
TCH8341-0.45
corA attL int rpr pri rep attR uvrA
TIGR4-1.06
eno attL int rpr pri rep terS TA attR IS630
INV104-1.06
eno attL int rpr pri rep terS attR psi
TL13-1.35rc
eno attL int 5' tnp int 3' rpr pri rep terS tat attR
AS1.3089-0.6
eno attL int rpr pri rep TA attR
A45-1.9
Mn/Fe trans attL int dinD divF rpr pri-rep attR
Taiwan-0.03
ychF attL int dinD divF rpr tnp-Spn1 pri rep terS attR mtlD
plasmid
TCH8341-0.25
ychF attL int dinD divF rpr pri rep terS attR mtlD
plasmid
ST556-0.03
ychF attL int dinD divF rpr pri rep terS attR mtlD
plasmid

Table 1. Characterization of the PICI-encoded replicon.

	Recipient strain			
PICI replicon	*S. aureus*	*E. faecalis JH2-2*	*L. monocytogenes*	*B. subtilis*
SaPI1	+++	+++	+++	+++
SaPIbov1	+++	+++	+++	+++
EfCIV583	+++	+++	+++	-

Table 2. Effects of phage mutants on EfCIV583 transfer[a].

Donor strain	Recipient strain[b]	
E. faecalis	**VE18590**	**JH2-2**
V583 wt	8.9×10^4	1.9×10^2
V583 $\Delta\phi1$[c]	<1	<1
JP11028[d]	5.6×10^4	2.4×10^2
JP13142[d] (p1 Δ*xis*)	<1	<1

[a]The means of results from three independent experiments are presented. Variation was within ± 5% in all cases.

[b]Transductants/ml of lysate, using the non-lysogenic VE18590 or JH2-2 as recipient strains.

[c]Phage deleted in strain V583.

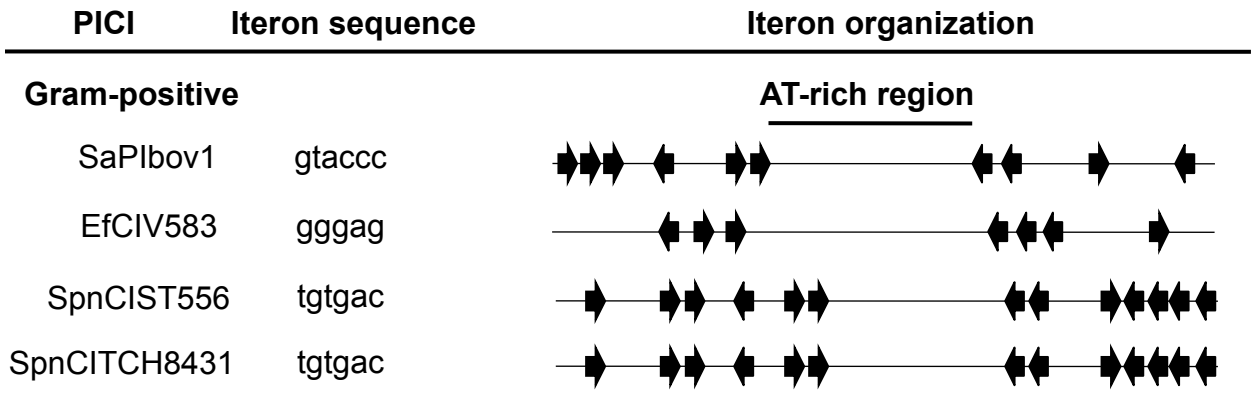
[d]JP11028: lysogenic for phage p1; EfCIV583-positive.

[d]JP13142: Derivative of JP11028, p1 Δ*xis* mutant.

***Identification of PICIs*.**

The analysis of orthologies points to elements that might correspond to PICIs. Examination of the corresponding KEGG genome maps was used to confirm the identifications. In Fig. S5A, the lower diagram shows a genome with an inserted PICI, and the smaller diagram above shows the corresponding syntenic region of a different strain lacking any insert. Comparison of such paired genome regions usually enables identification of the ends of the inserted element and reveals a short direct repeat that represents the core *att* site. A typical example is shown in Fig. S5A, where the BLAST pattern was obtained by searching with the putative flanking genes, *rrm*A and *cut*C, which are located at about 2.1 Mb. The BLAST lines with an interruption represent strains containing an insert between these two genes. The uninterrupted ones represent strains with no insert at that location. The nucleotides immediately flanking the interruption are at the ends of the insert and usually represent the *att* site core. This can be confirmed by performing a BLAST search with either of the flanking genes plus 3-500 nucleotides corresponding to the end of the PICI. In some cases, no *att* site core can be identified, presumably owing to sequence modifications that have occurred following insertion of the element. In these cases, we have assumed that the element is defined by the genes flanking the insertion, as revealed by comparing strains as in Fig. S5A. In Fig. S5B, the diagram shows a genome with a prophage integrated on the *suf*B gene.

A. PICI replication origins



B. PICI autonomous replication

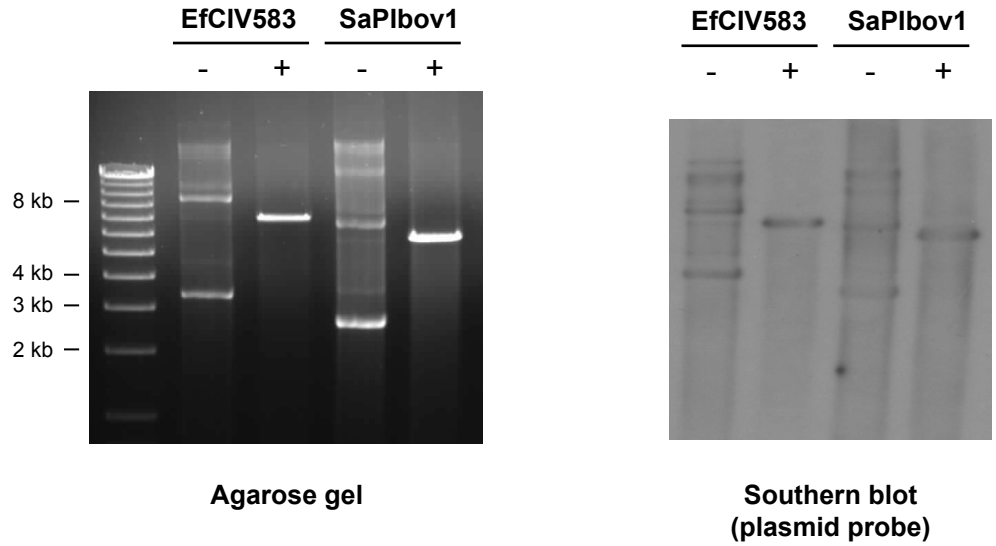


Figure S1. Characterization of the PICI replication origins. (A) Comparative map of the replication origins of several PICIs. The iterons are represented by arrows, and their sequences are shown at left. Note that there are always two sets of iterons flanking an AT-rich region, which could be the melting site. (B) Plasmid DNA carrying the Pri-Rep-*ori* segment of EfCIV583 or SaPIbov1 was isolated from overnight cultures of *S. aureus* RN4220, in presence of erythromycin. Plasmids were analyzed in agarose gels (left) or transferred for Southern blot studies using a probe specific for the plasmid (right). (-): non-digested plasmid; (+) plasmids digested with *Bam*HI, which cuts one in the plasmids.

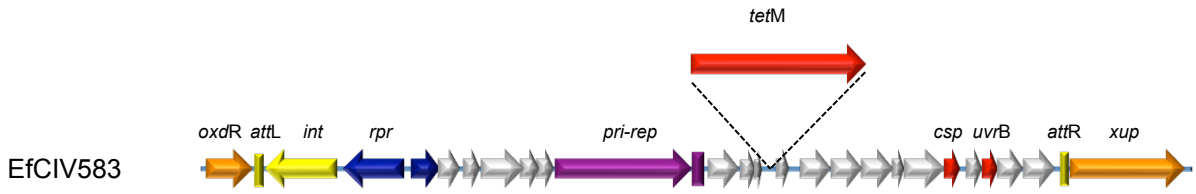
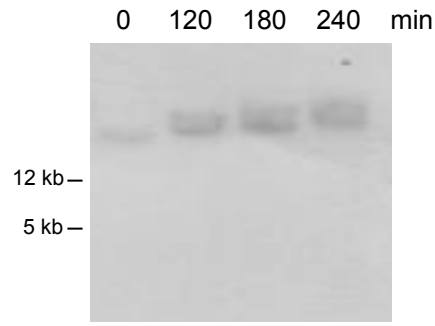


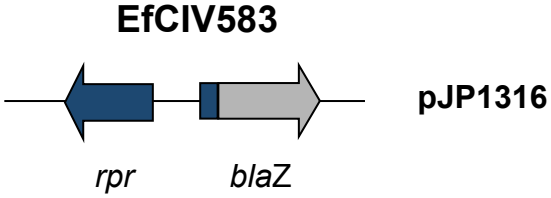
Figure S2. Location of the *tet*M marker in the EfCIV583 genome.

A

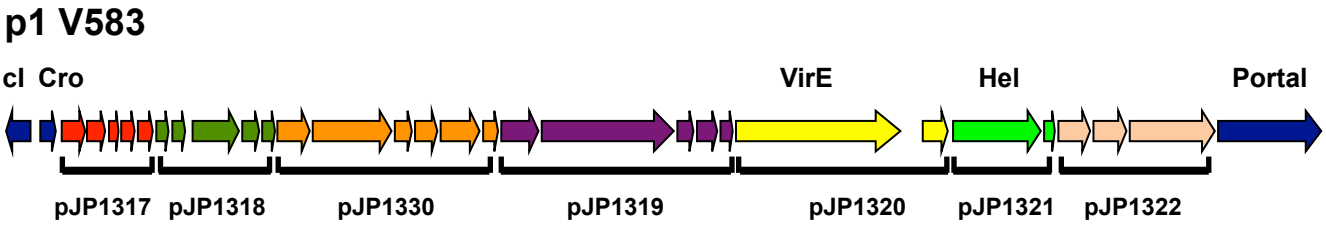


VE18589
EfCIV583, non lysogenic

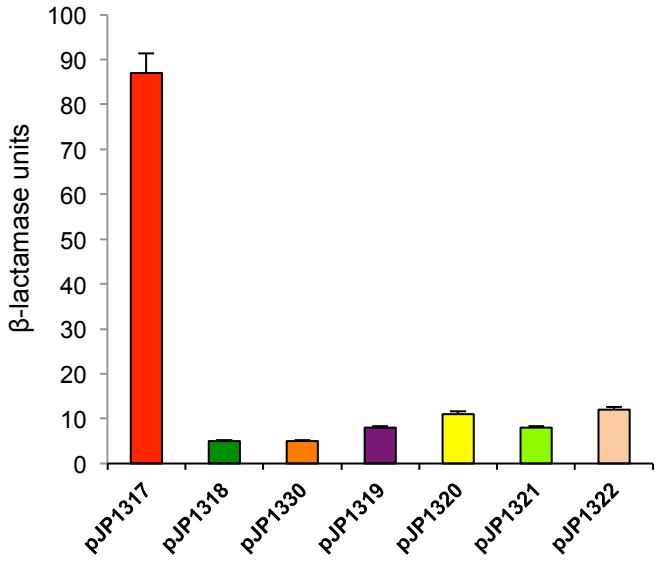
B



C



D



E

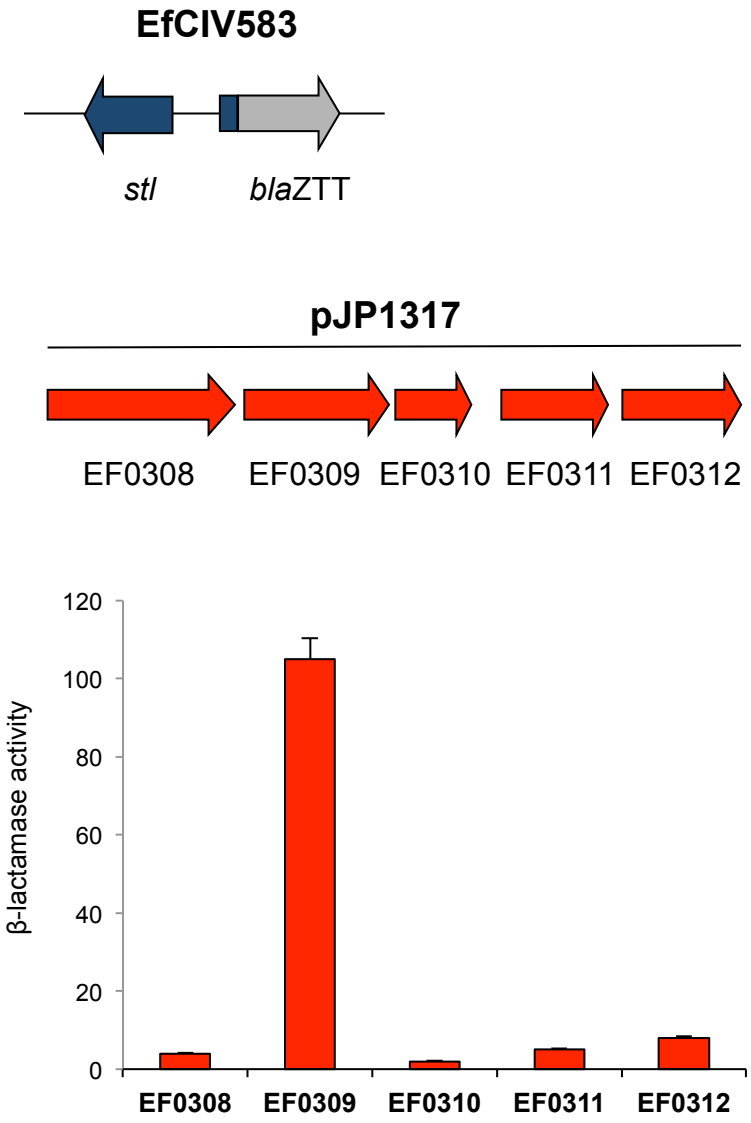


Figure S3. Identification of p1 region containing the derepressor of EfCIV583. (A) Failure of MC to induce excision/replication of EfCIV583. A culture of VE18589 (EfCIV538 positive) was induced with MC and samples taken during the subsequent incubation were used to prepare minilysates that were separated on agarose and Southern blotted with a EfCIV538 probe. (B) Construct used to test for de-repression. (C) Map of p1 showing regions that were cloned and tested for de-repression. *S. aureus* RN4220 derivative strains containing pJP1316 and pCN51 derivative plasmids (P*cad* promoter) containing the different regions of p1 were assayed for β-lactamase activity after induction with 5 μM $CdCl_2$. Samples were normalized for total cell mass. (D) β-lactamase activity generated by the above reporter construct in the presence of each of the cloned p1 regions, following induction with 5 um $CdCl_2$. (E) Demonstration of the de-repression activity of a subclone of pJP1317. Tests were performed with the construct shown in B. Values presented are the averages (±SD) of three independent assays.

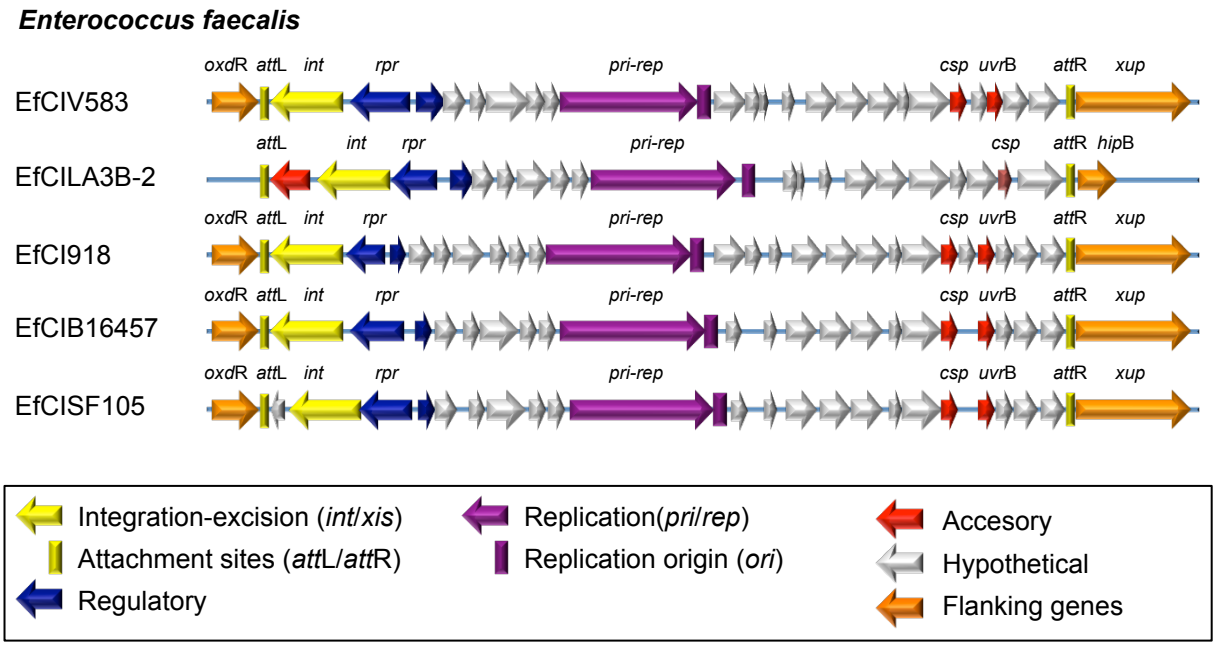


Figure S4. *Enterococcus faecalis* PICI genomes.

Figure S5A.
LICI-SK11 at *cut C* site, plus empty *cutC* site in strain NCDO

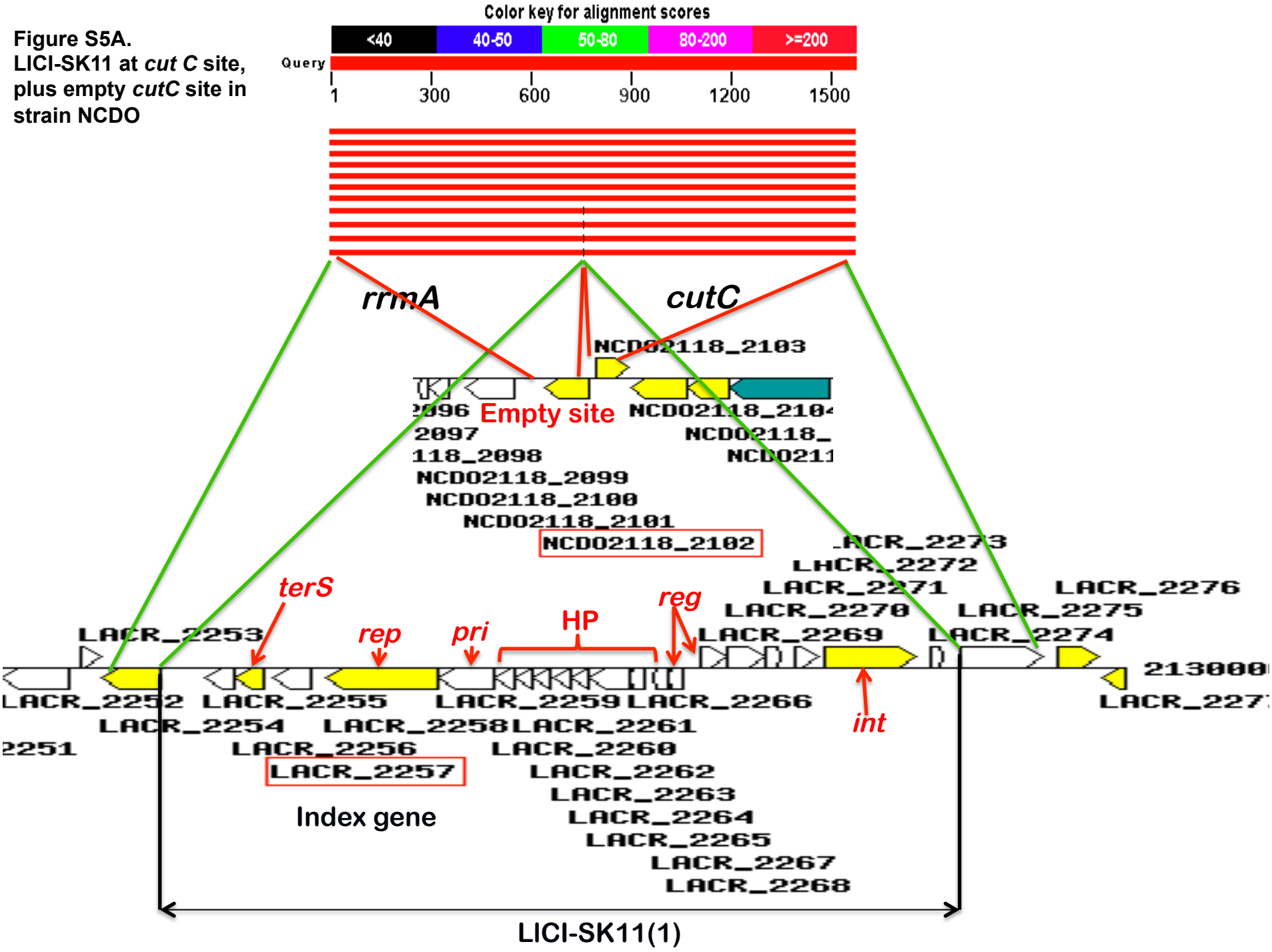
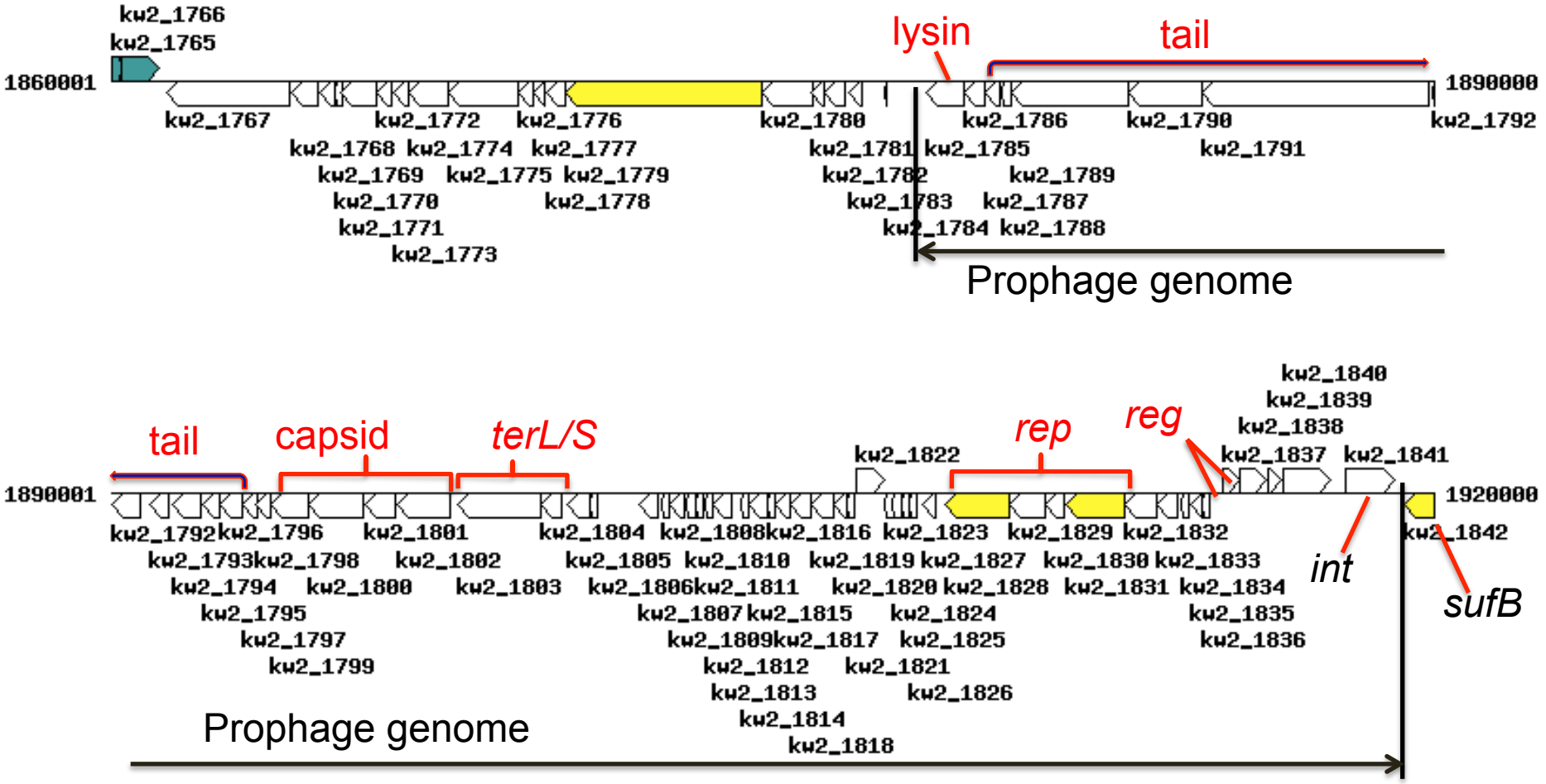


Figure S5B. LIPH-KW2 (Prophage) at *suf*B site



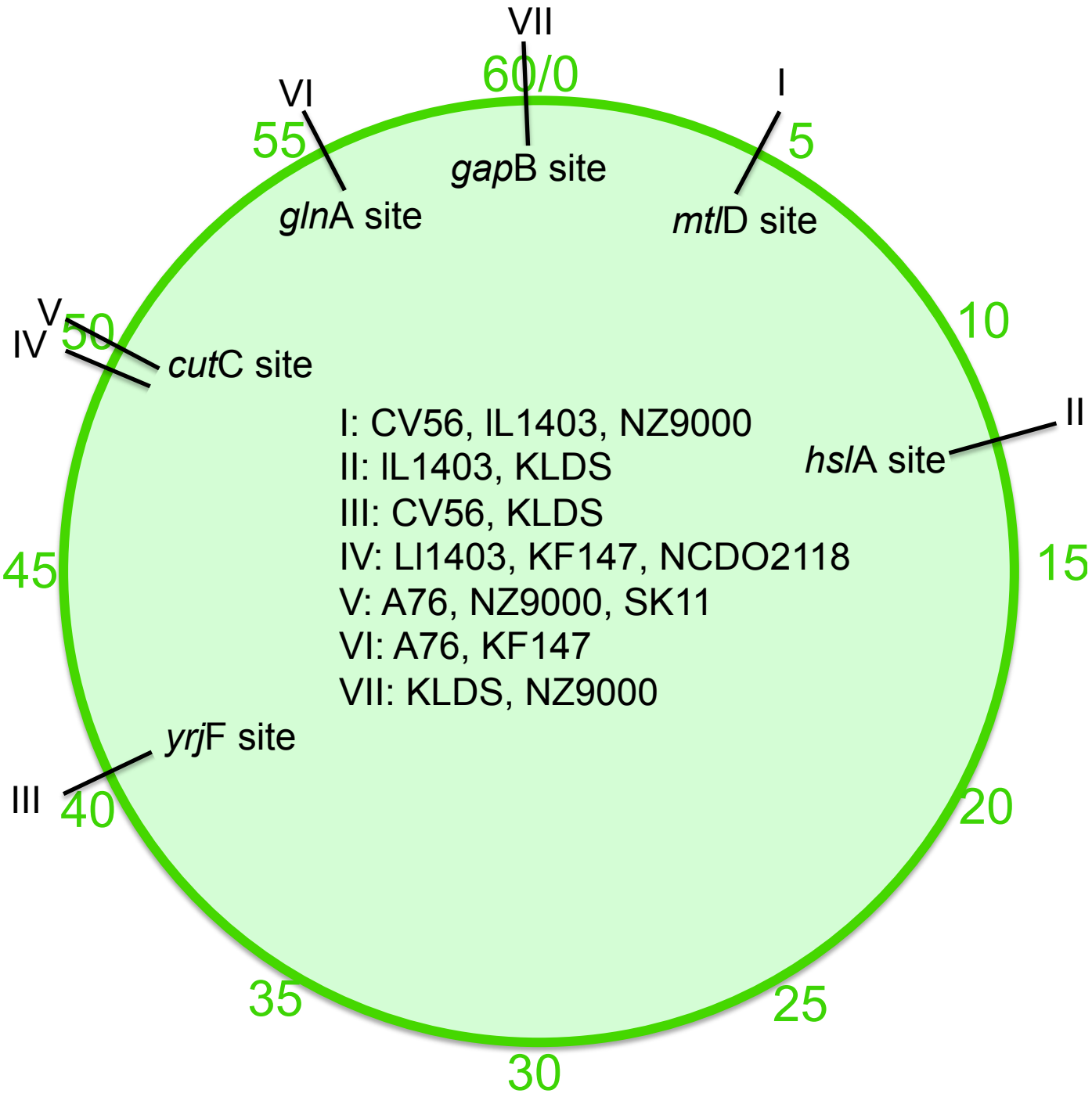


Figure S6. Locations of PICI *att* sites in lactococci.

```
Phage bIL286    TTTTAATAACCCTCCCCCCGTATCTTTTCACAGGGAAACCACACACATAGGTATCGTCTT
LlCI-bIL310     TTTTTAAACACCCCCGCCCGTATCTTTTCACAGGGAAACCACACACATAGGTATCATCTT
                ****:*:*..** ** ****************************************.****

Phage bIL286    GCGTGAAAACCCATTTTTCAAAATTTTTATATAGGGGGGGTCAAAACACTAAAA
LlCI-bIL310     GCGTGAAAGTCCATTTTTGAAAATTTTTATATAGG-GGGGTCAAAAATCTAAAA
                ********. ******** *************** **********.:******

Phage bIL285    AAAGAACCGAGTGAGTTTAGCTTTTTCCAAGTGTGAGGAAATTTGAAAATATTTTTTTAC
LlCINZ9000-2    AAAGAACCGAGTGAGTTTAGCTTTTTCCAAGTGTACAAAGTCCTGAATCTATTTTTTTAC
                **********************************....*.:  ****:.***********

Phage bIL309    CCCCCCGTCATCGCTTTTAGGAATACCGTATACCAATGGTGGCTTCCTGAGTAAAAAAGT
LlCI-A76-2      CCCCCCGTCATTGCTTTTAGGAATACCGTATACCAATAGTGGCTCCCTGAGTAAAAAAGT
                *********** *************************.****** ***************

Phage bIL309    GATTTTTTAAAATTTTTGCATAGGGGGGGT
LlCI-A76-2      GGTTTTTTAAAATTTTTGCATAGGGGGGGT
                *.****************************
```

Figure S7. Cos alignment. The predicted latococcal PICI and phage *cos* sites and their flanking sequences are aligned using ClustalW2.

Table S1. Strains used in this work.

Strain	Description	Reference
Enterococcus faecalis		
VE14089	*E. faecalis* V583 cured of its plasmids. EfCIV583 positive	(Rigottier-Gois *et al.*, 2011)
E. faecalis JH2-2	Laboratory strain	(Jacob and Hobbs, 1974)
VE18590	*E. faecalis* V583 deriviative. Non-lysogenic, EfCIV583-negative.	(Matos *et al*. 2013)
VE18589	VE18590 (EfCIV583)	(Matos *et al*. 2013)
JP10318	VE14089 EfCIV583 *tet*M	This work
JP10416	JP10318 Δp1	This work
JP10860	JP10318 Δp2	This work
JP10861	JP10318 Δp3	This work
JP10862	JP10318 Δp4	This work
JP10863	JP10318 Δp5	This work
JP10864	JP10318 Δp6	This work
JP10984	VE 18562 (pp1+)	(Matos *et al*. 2013)
JP11211	JP10984 ΔEF0309	This work
JP11028	JP10984 EfCIV583 *tet*M	This work
JP13142	JP11028 ΔEF0309	This work
Bacillus subtilis		
B. subtilis RL-3	Natural competent	Richard Losick lab
Listeria monocytogenes		
L. monocytogenes RN10983	Natural lysogenic strain	Richard Novick lab
Staphylococcus aureus		
S. aureus RN4220	Restriction-defective strain	(Kreiswirth *et al.*, 1983)
JP8546	RN4220 pJP1097	(Mir-Sanchis *et al.*, 2012)
JP8545	RN4220 pJP1096	(Mir-Sanchis *et al.*, 2012)
JP10837	RN4220 pJP1350	This work
JP10838	RN4220 pJP449	This work
JP5267	RN4220 pJP788	This work
JP5266	RN4220 pJP787	This work
JP9991	RN4220 pJP1277	This work
JP9992	RN4220 pJP1278	This work
JP5118	RN4220 pJP782	This work
RN10733	RN4220 pRN9211	(Ubeda *et al.*, 2007)
RN10734	RN4220 pRN9217	(Ubeda *et al.*, 2007)
JP10399	RN4220 pJP1316	This work
JP10739	JP10399 pJP1317	This work
JP10740	JP10399 pJP1318	This work

Strain	Description	Reference
Staphylococcus aureus		
JP10822	JP10399 pJP1330	This work
JP10741	JP10399 pJP1319	This work
JP10742	JP10399 pJP1320	This work
JP10743	JP10399 pJP1321	This work
JP10744	JP10399 pJP1322	This work
JP10745	JP10399 pJP1323	This work
JP10746	JP10399 pJP1324	This work
JP10747	JP10399 pJP1325	This work
JP10748	JP10399 pJP1326	This work
JP10749	JP10399 pJP1327	This work
Escherichia coli		
JP5630	DH5α pJP795/pJP796	This work
JP5631	DH5α pJP795/pJP797	This work
JP5632	DH5α pJP795/pJP798	This work
JP5039	DH5α pJP782	This work
JP4994	DH5α pJP781	This work
JP9821	DH5α pJP1306	This work
JP10488	BL21 (DE3) pJP1328	This work
Lactococcus lactis		
IL1403	Laboratory strain	(Chopin *et al.*, 1984)
JP14198	IL1403 pAGEnt	This work
JP14199	IL1403 pJP1868	This work
JP14203	IL1403 pJP1869	This work

Table S2. Oligonucleotides used in this work.

Plasmid	Oligonucleotides	Sequence (5'-3')
Excision, circularisation and integration		
$EfCI_{V583}$		
Excision	EfPI-1m	TAAAAACAGCGCCTTCGTCC
	EfPI-2c	AATCGAGTAGTAGCTGAAACG
Circularization	EfPI-8m	CTTCTTCAATCAGGAGTGCC
	EfPI-12c	TATGGTTGGTACTGATAGGCG
In/out	EfPI-8m	CTTCTTCAATCAGGAGTGCC
	EfPI-2c	AATCGAGTAGTAGCTGAAACG
pJP795	EfCI-int-5mH	CCCAAGCTTTTGGCTTAAACCAAGAAAAGC
	EfCI-int-4cB	CGCGGATCCATTATGGGTGTTTTAAATGGC
pJP796	EfCI-int-1mS	ACGCGTCGACCAGTTATAGAAACATCTCTCC
	EfCI-int-3cB	CGCGGATCCAGTGATAATCAGTCAGTTGGC
pJP797	EfCI-int-1mS	ACGCGTCGACCAGTTATAGAAACATCTCTCC
	EfCI-int-4cB	CGCGGATCCATTATGGGTGTTTTAAATGGC
pJP798	EfCI-int-2mS	ACGCGTCGACTGAAACACTTCAAATTATGGC
	EfCI-int-3cB	CGCGGATCCAGTGATAATCAGTCAGTTGGC
pJP1350	SaPIbov-149cB	CGCGGATCCGATCAGTACCTAAATATGCG
	SaPIbov1-243mK	CGGGGTACCTACGACATTAACGTCATTGCG
pJP788	EfPI-16mE	CCGGAATTCGCTTTTTATCAAGCGTATGGC
	EfPI-15cB	CGCGGATCCACATATAGGCGGTTGTACCG
pJP787	EfPI-16mE	CCGGAATTCGCTTTTTATCAAGCGTATGGC
	EfPI-17cB	CGCGGATCCCGTGGAATACCTAACTCCTC
pJP1277	EfPI-56mE	CCGGAATTCGGAAACGCCCTCTACTATCTTC
	EfPI-15cB	CGCGGATCCACATATAGGCGGTTGTACCG
pJP1278	EfPI-56mE	CCGGAATTCGGAAACGCCCTCTACTATCTTC
	EfPI-17cB	CGCGGATCCCGTGGAATACCTAACTCCTC
pJP782	EfPI-7mB	CGCGGATCCAATGACCTCGTGTAAAAGCC
	EfPI-9cS	ACGCGTCGACCAATAGAGAATCCGAGATAGC
pJP781	EfPI-6mB	CGCGGATCCAAGGGATTTGGTCGGTTACC
	EfPI-9cS	ACGCGTCGACCAATAGAGAATCCGAGATAGC
pJP1306	EfPI-7mB	CGCGGATCCAATGACCTCGTGTAAAAGCC
	EfPI-49c	GCATTGGGAGATTTTTTCAGC
	EfPI-50m	GCTGAAAAAATCTCCCAATGCCTCTGGTAGGTACTCC ACAAG
	EfPI-9cS	ACGCGTCGACCAATAGAGAATCCGAGATAGC
pJP1312	EfPI-27mXS	GCTCTAGAGCACGCGTCGACTGACAACGTTCCTCTCT TTCC
	EfPI-28cB	CGCGGATCCTCTTAAGGAGTGCTAAAGAGC
	EfPI-29mP	AAAACTGCAGGAAGCGGAAGATTTCATGCCG
	EfPI-30cE	CCGGAATTCCTTACTGAGAATCAGGAGAGC
pJP1313	EFV583phi1-3mBglII	GAAGATCTTTAGGAACACCGCCAGAAACC
	EFV583phi1-4cS	ACGCGTCGACTTTACGACCAGACGAAGAGCC

Plasmid	Oligonucleotides	Sequence (5'-3')
pJP1351	EFV583phi2-3mB	CGCGGATCCTTAGCCGCAGCAAGTAATGCG
	EFV583phi2-6m	ACAGAATAATCCCTAAATTCCCAAACGATGGCAACGCACAG
	EFV583phi2-5c	GAATTTAGGGATTATTCTGTG
	EFV583phi2-4cS	ACGCGTCGACTAAATCCGACATATGGGCAGG
pJP1352	EFV583phi3-3mB	CGCGGATCCAGTTGAAGCTGATGCGGAAGG
	EFV583phi3-4cS	ACGCGTCGACGAAATTTCGGAAAATTCTCCG
pJP1353	EFV583phi4-3mB	CGCGGATCCTTTATGGCAATATGGAAGGAG
	EFV583phi4-4cS	ACGCGTCGACAATTAACAGCGGTTGATAGCC
pJP1354	EFV583phi5-3mB	CGCGGATCCAATACTCAATGCCATATAGGG
	EFV583phi5-4cS	ACGCGTCGACAGCGTTTTGCTAGTAAAGGGC
pJP1355	EFV583phi6-3mB	CGCGGATCCACAGTACGTTTCCACTGTCGC
	EFV583phi6-4cS	ACGCGTCGACTCCAATACCTTTCCCGATACG
pJP1552	EFV583phi1-5mS	AGCGTCGACTTTGAAACTTTGTGGGAATACG
	EF0309-5c	AACCGGTTTTGGCATACCCC
	EF0309-6m	GGTATGCCAAAACCGGTTAAGAAAAGAAAGGGCGGATAG
	EFV583phi1-6cB	CGCGGATCCTGCCGCTATACGTCTTAATTG
pJP1316	EfPI-15cB	CGCGGATCCACATATAGGCGGTTGTACCG
	TT-1cSp	ACATGCATGCTGTCACTTTGCTTGATATATGAG
pJP1317	EFV583phi1-5mS	ACGCGTCGACTTTGAAACTTTGTGGGAATACG
	EFV583phi1-6cB	CGCGGATCCTGCCGCTATACGTCTTAATTG
pJP1318	EFV583phi1-7mS	ACGCGTCGACAATATCTCAATTTATGAGGTGTAC
	EFV583phi1-8cB	CGCGGATCCTGAATCTGCTTCAATATTTAAATAG
pJP1330	EFV583phi1-9mS	ACGCGTCGACAAAGCGATTTTCAATGTAACAGATG
	EFV583phi1-10cB	CGCGGATCCTGAATTTTTTAAAGTAATCACATGG
pJP1319	EFV583phi1-11mS	ACGCGTCGACGAATTCGGAGATACTAATTTTTATG
	EFV583phi1-12cB	CGCGGATCCTAATTTGATGTTTTTTTCTGGCTG
pJP1320	EFV583phi1-13mS	ACGCGTCGACTATCAAGAAGGATGGCTTGAC
	EFV583phi1-14cB	CGCGGATCCTCTTAGAGTATTCCTGATAGGG
pJP1321	EFV583phi1-15mS	ACGCGTCGACAGCAACGTTACATCCCTATCAG
	EFV583phi1-16cB	CGCGGATCCTAATTTTTGATAGCTGACTAACC
pJP1322	EFV583phi1-17mS	ACGCGTCGACATACCGTATGATGTTAGATTTATATTG
	EFV583phi1-18cB	CGCGGATCCGTCTTCACTAAGTAAAGCTTCC
pJP1323	EFV583phi1-5mS	ACGCGTCGACTTTGAAACTTTGTGGGAATACG
	EFV583phi1-35cB	CGCGGATCCTTCGACAGCTTCCAGATCAAC
pJP1324	EFV583phi1-36mS	ACGCGTCGACGTTACCGATTATGTGGCTGTG
	EFV583phi1-37cB	CGCGGATCCGCATAATATCCACGCTTCTTG
pJP1325	EFV583phi1-38mS	ACGCGTCGACGTGTTTGATGAGTTTGTTCGC
	EFV583phi1-39cB	CGCGGATCCCCTCCTACGGAATTAATCTGT
pJP1326	EFV583phi1-40mS	ACGCGTCGACGTCGTTTAGAAGATAAGAACCG
	EFV583phi1-41cB	CGCGGATCCAAATTCCACGCTAGCCTTTTG
pJP1327	EFV583phi1-42mS	ACGCGTCGACGATGAACGTTTGCCAGAAGCA
	EFV583phi1-6cB	CGCGGATCCTGCCGCTATACGTCTTAATTG

Plasmid	Oligonucleotides	Sequence (5'-3')
pJP1328	EfPI-45mB	CGCGGATCCGATGAGAAAGGAGTTTCCTCTGATG
	EfPI-46cE	CCGGAATTCTCATTCTTTAGCTTTTGATTTACG
	EF0309-1mS	ACGCGTCGACTTCACACAGGAAACAGACCATGCCAAAACCGGTTAAGGTT
	EF0309-2cP	AACTGCAGCTATCCGCCCTTTCTTTTCTT
pJP1868	bIL286-cos_mP	TCAGTTCTGCAGATTTTTAATAACCCTCCCCCCGTATCTT
	bIL286-cos_cS	TCTCTTACTAGTTTTTTCTCCTTTCTTTTAGTGTTTTGAC
pJP1869	bIL310-cos_mP	TCAGTTCTGCAGTTTTTAAACACCCCCGCCCGTAT
	bIL310-cos_cS	CTTACTAGTTTTTAGATTTTTGACCCCCCTATATAAAAAT

Southern blot	Oligonucleotides	Sequence (5'-3')
EfCIV583 probe	EfPI-29mP	AAAACTGCAGGAAGCGGAAGATTTCATGCCG
	EfPI-30cE	CCGGAATTCCTTACTGAGAATCAGGAGAGC
Phage p1 probe	EFV583phi1V-1m	GTGCCTAAATCATAAGGACGG
	EFV583phi1V-2c	AAAGATTCCGTGCGATTATCC

Table S3. Plasmids used in this work.

Plasmid	Description	Reference
pCN41	Ap^r. Used in transcriptional fusions to the staphylococcal β–lactamase *blaZ*	(Charpentier *et al.*, 2004)
pCN33	Ap^r. Cloning vector	(Charpentier *et al.*, 2004)
pCN51	Ap^r. Expression vector	(Charpentier *et al.*, 2004)
pMAK700	Cm^r. Plasmid thermosensitive in *E. coli*	(Hamilton *et al.*, 1989)
pMAD	Vector for efficient allelic replacement	(Arnaud *et al.*, 2004)
pCU1	Cm^r. Cloning vector	(Augustin *et al.*, 1992)
pPROEX HTa	Expression vector	Invitrogen
pAGEnt	Cm^r. Expression vector	(Linares *et al.*, 2014)
pJP795	pMAK700 att_C EfCIV583	This work
pJP796	pCN51 *int*-att_{PI} EfCIV583	This work
pJP797	pCN51 *int*-att_L EfCIV583	This work
pJP798	pCN51 Δ*int*-att_{PI} EfCIV583	This work
pJP1097	pCN41 *stl-xis-blaZ* SaPIbov1	(Mir-Sanchis *et al.*, 2012)
pJP1096	pCN41 Δ*stl-xis-blaZ* SaPIbov1	(Mir-Sanchis *et al.*, 2012)
pJP1350	pCN41 *stl-hel-blaZ* SaPIbov1	This work
pJP449	pCN41 Δ*stl-hel-blaZ* SaPIbov1	(Ubeda *et al.*, 2008)
pJP788	pCN41 *stl-xis-blaZ* EfCIV583	This work
pJP787	pCN41 Δ*stl-xis-blaZ* EfCIV583	This work
pJP1277	pCN41 *stl-hel-blaZ* EfCIV583	This work
pJP1278	pCN41 Δ*stl-hel-blaZ* EfCIV583	This work
pJP782	Suicide plasmid. Em^R Δ*stl-pri-hel-ori* EfCIV583	This work
pJP781	Suicide plasmid. Em^R Δ*stl-pri-hel-*Δ*ori* EfCIV583	This work
pJP1306	Suicide plasmid. Em^R Δ*stl-pri-*Δ*hel-ori* EfCIV583	This work
pRN9211	Suicide plasmid. Em^R *pri-hel-ori* SaPIbov1	(Ubeda *et al.*, 2007)
pRN9217	Suicide plasmid. Em^R *pri-hel-ori* SaPI1	(Ubeda *et al.*, 2007)
pJP1312	pMAD derivative. Insertion of *tet*M cassette in EfCIV583	This work
pJP1313	pMAD derivative. Deletion of phage p1 from VE14089	This work
pJP1351	pMAD derivative. Deletion of phage p2 from VE14089	This work
pJP1352	pMAD derivative. Deletion of phage p3 from VE14089	This work
pJP1353	pMAD derivative. Deletion of phage p4 from VE14089	This work
pJP1354	pMAD derivative. Deletion of phage p5 from VE14089	This work
pJP1355	pMAD derivative. Deletion of phage p6 from VE14089	This work
pJP1552	pMAD derivative. Deletion of EF0309 from V583 p1.	This work
pJP1316	Transcriptional analysis of stl EfCIV583, pCU1 β-lactamase *blaZ*	This work
pJP1317	Expression of EF0308 to EF0312, pCN51 derivative	This work
pJP1318	Expression of EF0313 to EF0317, pCN51 derivative	This work
pJP1330	Expression of EF0318 to EF0323, pCN51 derivative	This work
pJP1319	Expression of EF0324 to EF0327, pCN51 derivative	This work
pJP1320	Expression of EF0328 to EF0329, pCN51 derivative	This work
pJP1321	Expression of EF0330, pCN51 derivative	This work

Plasmid	Description	Reference
pJP1322	Expression of EF0331 to EF0333, pCN51 derivative	This work
pJP1323	Expression of EF0308, pCN51 derivative	This work
pJP1324	Expression of EF0309, pCN51 derivative	This work
pJP1325	Expression of EF0310, pCN51 derivative	This work
pJP1326	Expression of EF0311, pCN51 derivative	This work
pJP1327	Expression of EF0312, pCN51 derivative	This work
pJP1328	Expression in *E. coli* of His-Stl EfCIV583 + EF0309, pPROEX HTa derivative	This work
pJP1868	pAGEnt containing phage bIL286 *cos* site	This work
pJP1869	pAGEnt containing LICIbIL310 putative *cos* site	This work

References

Arnaud M, Chastanet A, Débarbouillé M. (2004). New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl Environ Microbiol* **70**: 6887–6891.

Augustin J, Rosenstein R, Wieland B, Schneider U, Schnell N, Engelke G, *et al.* (1992). Genetic analysis of epidermin biosynthetic genes and epidermin-negative mutants of *Staphylococcus epidermidis*. *Eur J Biochem* **204**: 1149–1154.

Charpentier E, Anton AI, Barry P, Alfonso B, Fang Y, Novick RP. (2004). Novel cassette-based shuttle vector system for gram-positive bacteria. *Appl Environ Microbiol* **70**: 6076–6085.

Chopin A, Chopin MC, Moillo-Batt A, Langella P. (1984). Two plasmid-determined restriction and modification systems in *Streptococcus lactis*. *Plasmid* **11**: 260–263.

Hamilton CM, Aldea M, Washburn BK, Babitzke P, Kushner SR. (1989). New method for generating deletions and gene replacements in *Escherichia coli*. *J Bacteriol* **171**: 4617–4622.

Jacob AE, Hobbs SJ. (1974). Conjugal transfer of plasmid-borne multiple antibiotic resistance in *Streptococcus faecalis* var. zymogenes. *J Bacteriol* **117**: 360–372.

Kreiswirth BN, Löfdahl S, Betley MJ, O'Reilly M, Schlievert PM, Bergdoll MS, *et al.* (1983). The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* **305**: 709–712.

Linares DM, Perez M, Ladero V, del Rio B, Redruello B, Martin MC, *et al.* (2014). An agmatine-inducible system for the expression of recombinant proteins in *Enterococcus faecalis*. *Microb Cell Fact* **13**: 169.

Mir-Sanchis I, Martínez-Rubio R, Martí M, Chen J, Lasa I, Novick RP, *et al.* (2012). Control of *Staphylococcus aureus* pathogenicity island excision. *Mol Microbiol* **85**: 833–845.

Rigottier-Gois L, Alberti A, Houel A, Taly J-F, Palcy P, Manson J, *et al.* (2011). Large-scale screening of a targeted *Enterococcus faecalis* mutant library identifies envelope fitness factors. *PLoS ONE* **6**: e29023.

Ubeda C, Barry P, Penadés JR, Novick RP. (2007). A pathogenicity island replicon in *Staphylococcus aureus* replicates as an unstable plasmid. *Proc Natl Acad Sci USA* **104**: 14182–14188.

Ubeda C, Maiques E, Barry P, Matthews A, Tormo MA, Lasa I, *et al.* (2008). SaPI mutations affecting replication and transfer and enabling autonomous replication in the absence of helper phage. *Mol Microbiol* **67**: 493–503.

Table S4. Description and relationships between the *E. faecalis* PICI elements[a].

EfCIV583[b]	GenBank acces		EfCI918	EfCISF105	EfCIB16457	EfCILA3B-2
EfCIV583_1	EF2955	*int*	95	93	100	85
EfCIV583_2	EF2954	*rpr*			100	99
EfCIV583_3	not annotated		85		100	97
EfCIV583_4	EF2953		90	76	100	83
EfCIV583_5	EF2952		90		100	99
EfCIV583_6	EF2951		72	85	99	84
EfCIV583_7	EF2950		80	80	94	95
EfCIV583_8	EF2949		80	98	82	82
EfCIV583_9	EF2948	*pri-rep*				
EfCIV583_10	EF2947					
EfCIV583_11	not annotated		83	80	78	77
EfCIV583_12	EF2946		65	65	65	65
EfCIV583_13	EF2945		100	100	100	100
EfCIV583_14	EF2944		94	100	94	91
EfCIV583_15	EF2943		93	100	93	96
EfCIV583_16	EF2942		94	99	94	92
EfCIV583_17	EF2941		88	100	88	100
EfCIV583_18	EF2940		97	100	97	97
EfCIV583_19	EF2939	*csp*	96	100	96	90
EfCIV583_20	EF2938		60	100	60	
EfCIV583_21	EF2937	*uvr*B	100	100	100	
EfCIV583_22	EF2936		97	90	97	
EfCIV583_23	I574_00041		98	98	98	

EfCILA3B-2[b]	GenBank acces		EfCIV583	EfCI918	EfCISF105	EfCIB16457
EfCILA3B-2_1	D347_01789	*fic*				
EfCILA3B-2_2	D347_01790	*int*	85	84	87	85
EfCILA3B-2_3	D347_01791	*rpr*	99			99
EfCILA3B-2_4	D347_01792		97	81		97
EfCILA3B-2_5	D347_01793		83	83	70	83
EfCILA3B-2_6	D347_01794		99	90		99
EfCILA3B-2_7	D347_01795		84	84	84	85
EfCILA3B-2_8	D347_01796		95	73	75	94
EfCILA3B-2_9	D347_01797		83	90	81	100
EfCILA3B-2_10	D347_01798	*pri-rep*		98	97	96
EfCILA3B-2_11	D347_01799		80	94	97	95
EfCILA3B-2_12	D347_01800		64	64	64	64
EfCILA3B-2_13	D347_01801		100	100	100	100
EfCILA3B-2_14	D347_01802		91	93	91	93
EfCILA3B-2_15	D347_01803		96	92	95	92
EfCILA3B-2_16	D347_01804		92	94	91	94
EfCILA3B-2_17	D347_01805		100	86	100	86
EfCILA3B-2_18	D347_01806		97	97	97	97
EfCILA3B-2_19	D347_01807	*csp*	88	93	88	93
EfCILA3B-2_20	D347_01808					

EfCISF105[b]	GenBank acces		EfCIV583	EfCI918	EfCILA3B-2	EfCIB16457
EfCISF105_1	UM9_00916					
EfCISF105_2	UM9_00917	*int*	93	92	87	93
EfCISF105_3	UM9_00918	*rpr*				
EfCISF105_4	UM9_00919					
EfCISF105_5	UM9_00920		80	81	71	80
EfCISF105_6	UM9_00921					
EfCISF105_7	UM9_00922		85	83	83	84
EfCISF105_8	UM9_00923		100	98	95	96
EfCISF105_9	UM9_00924		99	79	81	81
EfCISF105_10	UM9_00925	*pri-rep*		97	97	96
EfCISF105_11	UM9_00926		74	92	88	98
EfCISF105_12	UM9_00927		100	100	100	100
EfCISF105_13	UM9_00928		100	94	91	94
EfCISF105_14	UM9_00929		99	93	95	93
EfCISF105_15	UM9_00930		99	93	91	93
EfCISF105_16	UM9_00931		100	86	100	86
EfCISF105_17	UM9_00932		100	97	97	97
EfCISF105_18	UM9_00933	*csp*	99	94	94	94
EfCISF105_19	UM9_00934		100	60		60
EfCISF105_20	UM9_00935	*uvr*B	100	100		100
EfCISF105_21	UM9_00936		99	98		98
EfCISF105_22	UM9_00937		98	97		97

EfCIB16457[b]	GenBank acces		EfCIV583	EfCI918	EfCILA3B-2	EfCISF105
EfCIB16457_1	Q95_00339	*int*	100	95	85	93
EfCIB16457_2	Q95_00340	*rpr*	100		99	
EfCIB16457_3	Q95_00341		100	100	99	
EfCIB16457_4	Q95_00342		100	94	83	80
EfCIB16457_5	Q95_00343		100	90	99	
EfCIB16457_6	Q95_00344		99	84	84	84
EfCIB16457_7	Q95_00345		94	76	94	78
EfCIB16457_8	Q95_00346		83	90	100	81
EfCIB16457_9	Q95_00347	*pri-rep*		96	96	96
EfCIB16457_10	Q95_00348		82	92	84	98
EfCIB16457_11	Q95_00349		100	100	98	100
EfCIB16457_12	Q95_00350		94	100	93	94
EfCIB16457_13	Q95_00351		93	99	92	93
EfCIB16457_14	Q95_00352		94	100	94	93
EfCIB16457_15	Q95_00353		91	100	93	91
EfCIB16457_16	Q95_00354		97	100	97	97
EfCIB16457_17	Q95_00355	*csp*	94	99	93	94
EfCIB16457_18	Q95_00356		73	99		73
EfCIB16457_19	Q95_00357	*uvr*B	100	100		100
EfCIB16457_20	Q95_00358		90	100		90
EfCIB16457_21	Q95_00359		98	100		97

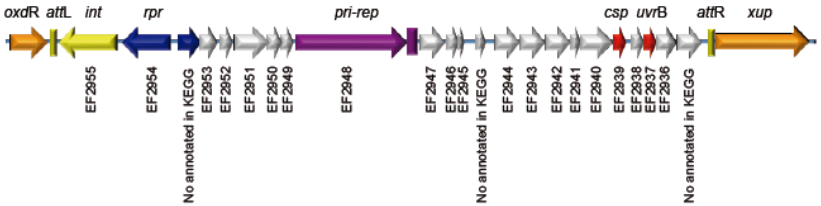
EfCI918[b]	GenBank acces		EfCIV583	EfCIB16457	EfCILA3B-2	EfCISF105
EfCI918_1	HMPREF2097_00609	*int*	95	95	84	92
EfCI918_2	HMPREF2097_00610	*rpr*				
EfCI918_3	HMPREF2097_00611		67	67	64	
EfCI918_4	HMPREF2097_00612		99	95	85	82
EfCI918_5	HMPREF2097_00613		88	88	90	
EfCI918_6	HMPREF2097_00614		69	69	94	70
EfCI918_7	HMPREF2097_00615		93	93	93	91
EfCI918_8	HMPREF2097_00616		98	95	100	98
EfCI918_9	HMPREF2097_00617		80	90	90	79
EfCI918_10	HMPREF2097_00618	*pri-rep*		96	98	97
EfCI918_11	HMPREF2097_00619		58	78	57	78
EfCI918_12	HMPREF2097_00620		64	64	64	60
EfCI918_13	HMPREF2097_00621		100	100	98	100
EfCI918_14	HMPREF2097_00622		94	100	93	94
EfCI918_15	HMPREF2097_00623		93	99	92	94
EfCI918_16	HMPREF2097_00624		94	100	94	93
EfCI918_17	HMPREF2097_00625		91	100	91	91
EfCI918_18	HMPREF2097_00626		97	100	97	97
EfCI918_19	HMPREF2097_00627	*csp*	94	99	93	94
EfCI918_20	HMPREF2097_00628			100		
EfCI918_21	HMPREF2097_00629		55	100		55
EfCI918_22	HMPREF2097_00630	*uvr*B	100	100		100
EfCI918_23	HMPREF2097_00631		91	100		91
EfCI918_24	HMPREF2097_00632		98	100		97

[a]PICI similarities were determined with BLASTX. Shading: light grey, >90 % similarity to corresponding gene in the PICI element described in the left column; dark grey, 50–90 %; black, <50 %; white, no corresponding gene. Abbreviations: *int*, integrase; *rpr*, PICI represor; *pri-rep*, primase-helicase homologues; *csp*, cold-shock protein.

[b]We have annotated genes in the PICIs according to the following nomenclature: *PICI_ORF number.*

Table S5. EfCIV583 orthologies.

EfCIV583



EfCIV583 gene	Function	Orthologs	Species	Length	Similarity (aa)	Overlap	Element	Start	End	Size (kb)	Comments
EF2936		efa:EF2936	*Enterococcus faecalis*	112			EfCIV583				No matches in database (DB)
EF2937		efa:EF2937	*Enterococcus faecalis*	69	1.000	69	EfCIV583				
		efq:DR75_2900	*Enterococcus faecalis*	69	0.492	59	No insert				
		efl:EF62_pB0056	*Enterococcus faecalis*	69	0.475	59	PICI	39611	46683	17.0	
		lpk:LACPI_0388	*Lactococcus piscium*	59	0.475	59	No insert[a]				
		lrg:LRHM_0140	*Lactobacillus rhamnosus*	69	0.491	57	No insert				
		lrh:LGG_00140	*Lactobacillus rhamnosus*	68	0.491	57	No insert				
		lcl:LOCK919_0751	*Lactobacillus casei*	66	0.439	57	No insert				
		lcz:LCAZH_0582	*Lactobacillus casei*	66	0.439	57	No insert				
		lpb:SH83_03640	*Lactobacillus plantarum*	66	0.368	57	No insert				
		lpj:JDM1_0749	*Lactobacillus plantarum*	66	0.368	57	No insert				
		lpl:lp_0899	*Lactobacillus plantarum*	66	0.368	57	No insert				
EF2938		efa:EF2938	*Enterococcus faecalis*	92			EfCIV583				No match in enterococci; no match in DB with less than 2x MW & over 35% similarity
EF2939		efa:EF2939	*Enterococcus faecalis*	67			EfCIV583				Cold-shock protein, has many matches in different species, none in any inserted element
EF2940		efa:EF2940	*Enterococcus faecalis*	190			EfCIV583				No significant match in DB
EF2941		efa:EF2941	*Enterococcus faecalis*	65	1.000	65	EfCIV583				
		spu:576079	*Strongylocentrotus purpuratus*	297	0.358	67	No insert				
		wse:WALSEDRAFT_57668	*Wallemia sebi*	491	0.391	64	No insert				
		bdi:100822641	*Brachypodium distachyon*	1101	0.381	63	No insert				
		ath:AT1G10490	*Arabidopsis thaliana*	1028	0.413	63	No insert				

EfCIV583 gene	Function	Orthologs	Species	Lenght	Similarity	Overlap	Element	Start	End	Size (kb)	Comments
		…									
		wci:WS105_0618	*Weissella ceti*	74	0.429	49	Prophage				
		wct:WS74_0819	*Weissella ceti*	74	0.429	49	Prophage				
EF2942		efa:EF2942	*Enterococcus faecalis*	185			EfCIV583				No match in enterococci; no significant match in DB
EF2943		efa:EF2943	*Enterococcus faecalis*	177			EfCIV583				No match in enterococci; no significant match in DB
EF2944		efa:EF2944	*Enterococcus faecalis*	160			EfCIV583				No match in enterococci; no significant match in DB
EF2945		efa:EF2945	*Enterococcus faecalis*	46			EfCIV583				No match in DB
EF2946		efa:EF2946	*Enterococcus faecalis*	47			EfCIV583				No match in DB
EF2947		efa:EF2947	*Enterococcus faecalis*	136	1.000	136	EfCIV583				
		efc:EFAU004_02137	*Enterococcus faecium*	135	0.406	128	PICI	2159624	2173033	13.4	Has full capsid module
		efu:HMPREF0351_12135	*Enterococcus faecium*	135	0.406	128	PICI				Has full capsid module
		pper:PRUPE_ppa005581mg	*Prunus persica*	453	0.302	96	No insert				
		cts:Ctha_2453	*Chloroherpeton thalassium*	755	0.333	66	No insert				
EF2948	*pri-rep*	efa:EF2948	*Enterococcus faecalis*	794	1.000	794	EfCIV583				
		lmoc:LMOSLCC5850_1262	*Listeria monocytogenes*	780	0.585	482	Prophage				
		lmod:LMON_1266	*Listeria monocytogenes*	780	0.585	482	Prophage				
		lmow:AX10_00090	*Listeria monocytogenes*	780	0.585	482	Prophage				
		lwe:lwe1216	*Listeria welshimeri*	780	0.585	482	Prophage				
		saa:SAUSA300_0809	*Staphylococcus aureus*	790	0.449	637	SaPI	881996	895129	13.1	
		saui:AZ30_04285	*Staphylococcus aureus*	790	0.449	637	SaPI				
		sax:USA300HOU_0861	*Staphylococcus aureus*	790	0.449	637	SaPI				
		bthu:YBT1518_01285 pu	*Bacillus thuringiensis*	797	0.532	477	Prophage				
EF2949		efa:EF2949	*Enterococcus faecalis*	82			EfCIV583				No significant match in DB

EfCIV583 gene	Function	Orthologs	Species	Length	Similarity	Overlap	Element	Start	End	Size (kb)	Comments
EF2950		efa:EF2950	*Enterococcus faecalis*	97	1.000	97	EfCIV583				
		efc:EFAU004_02141	*Enterococcus faecium*	52	0.396	48	PICI	2159624	2174178	14.5	Has full capsid module
		efu:HMPREF0351_12139	*Enterococcus faecium*	52	0.396	48	PICI				Has full capsid module
		efd:EFD32_1868	*Enterococcus faecalis*	84	0.333	72	defective phage?			3.2	
		cpae:CPAST_c11890	*Clostridium pasteurianum*	398	0.457	46	No insert				
		gmx:100776868	*Glycine max (soybean)*	922	0.322	90	No insert				
EF2951		efa:EF2951	*Enterococcus faecalis*	221	1.000	221	EfCIV583				No match in enterococci
		ppen:T256_00520	*Pediococcus pentosaceus*	234	0.471	227	PICI	99518	111337	11.8	
		lbh:Lbuc_0024	*Lactobacillus buchneri*	232	0.416	219	PICI	24658	40308	15.6	
		cpas:Clopa_0076	*Clostridium pasteurianum*	225	0.420	207	PICI	76442	85886	9.4	Defective
		std:SPPN_01200	*Streptococcus pseudopneumoniae*	209	0.421	202	PICI	190041	202462	12.4	Defective
		lbk:LVISKB_0740	*Lactobacillus brevis*	225	0.392	222	PICI	765893	774063	8.1	Defective
		smb:smi_2013	*Streptococcus mitis*	208	0.408	201	PICI	2057587	2068441	10.9	
		snc:HMPREF0837_10280	*Streptococcus pneumoniae*	201	0.399	203	PICI	249288	267373	18.1	
		snd:MYY_0022	*Streptococcus pneumoniae*	201	0.399	203	PICI				Defective
		snt:SPT_0025	*Streptococcus pneumoniae*	201	0.399	203	PICI				Defective
		ssut:TL13_0174	*Streptococcus suis*	206	0.391	202	PICI	131344	143429	12.1	
EF2952		efa:EF2952	*Enterococcus faecalis*	82			EfCIV583				No matches in DB
EF2953		efa:EF2953	*Enterococcus faecalis*	105	1.000	105	EfCIV583				
		xne:XNC1_0454	*Xenorhabdus nematophila*	313	0.300	100	No insert				
		xnm:XNC2_0444	*Xenorhabdus nematophila*	313	0.300	100	No insert				
		rca:Rcas_1163	*Roseiflexus castenholzii*	325	0.323	93	No insert				
		gca:Galf_2037	*Gallionella capsiferriformans*	79	0.316	76	PICI	2192302	2211152	18.8	Integrase backwards, no transcriptional divergence
		…									
		efd:EFD32_2452	*Enterococcus faecalis*	89	0.306	72	Prophage	2470933	2506883	36.0	
		elo:EC042_2423	*Escherichia coli*	63	0.333	54	Prophage	2552120	2559893	7.8	Defective phage

EfCIV583 gene	Function	Orthologs	Species	Length	Similarity	Overlap	Element	Start	End	Size (kb)	Comments
EF2954	*rpr*	efa:EF2954	*Enterococcus faecalis*	316	1.000	316	EfCIV583				No close match in enterococci
		scp:HMPREF0833_11008	*Streptococcus parasanguinis*	175	0.398	118	No insert				
		rto:RTO_22090	*Ruminococcus torques*	138	0.400	110	No insert				
		ssui:T15_0899	*Streptococcus suis*	182	0.412	102	PICI				
		csc:Csac_2097	*Caldicellulosiruptor saccharolyticus*	211	0.318	198	No insert				
		lam:LA2_03865	*Lactobacillus amylovorus*	147	0.540	63	No insert				
		lay:LAB52_03720	*Lactobacillus amylovorus*	241	0.507	71	No insert				
EF2955	*int*	efa:EF2955	*Enterococcus faecalis*	381	1.000	381	EfCIV583				
		efau:EFAU085_00503	*Enterococcus faecium*	393	0.467	379	Genomic island				
		efc:EFAU004_00565	*Enterococcus faecium*	393	0.467	379	Genomic island				
		efl:EF62_2611	*Enterococcus faecalis*	380	0.467	379	Genomic island				
		efu:HMPREF0351_10594	*Enterococcus faecium*	393	0.467	379	Genomic island				
		ehr:EHR_01850	*Enterococcus hirae*	380	0.467	379					
		efq:DR75_1138	*Enterococcus faecalis*	380	0.464	379	Genomic island				
		ecas:ECBG_01612	*Enterococcus casseliflavus*	380	0.488	381					
		crn:CAR_c08850	*Carnobacterium* sp	382	0.448	384	Genomic island				

[a]No insert: The ortholog seems to be chromsomally encoded, not been part of a defined mobile element.

Table S6. Orthology analysis of the 22 ORFs of LICICV56-1

Abbreviations

ll — *Lactococcus lactis*
lla: IL1403
llc: SK11
lld: KLDS
lli: UC509.9
llm:MG1363
lln: NZ9000
llr: A76
lls: IO-1
llk: KF147
llt: CV56
llw: KW2
llx:NCDO2118

cbr - *Cenorabditis elegans*
cat - *Croceibacter atlanticus*
cdf - *Peptoclostridium difficile* 630
ckl - *Clostridium kluyveri* DSM 555
cso - *Clostridium cf. saccharolyticum* K10
efq - *E. faecalis* ATCC 29212
ehr - *E. hirae*
efm - *Enterococcus faecium* NRRL B-2354
ere - Eubacterium rectale ATCC 33656
evi - *Echinicola vietnamensis*
lpl - *Lactobacillus plantarum* WCFS1
lcb - *Lactobacillus casei*
lpg - Lactobacillus paracasei N1115
lca - *Lactobacillus casei* ATCC 334

lga - *Lactobacillus gasseri* ATCC 33323
lmc - *Listeria monocytogenes* Clip81459
lcn - *Leuconostoc carnosum*
loa - Loa loa (eye worm)
lgr - *Lactococcus garveae* ATCC 49156
lgv - *Lactococcus garveae* Lg2
mar - *Microcystis aeruginosa*
mtr - *Meleagris gallopavo* (turkey): 100539070
myr - *Myroides sp.* A21
mer - *Methanomassiliicoccus sp. Mx1-Issoire*
smb - *S. mitis* B
stk - *S. parauberis*
snu - *S. pneumoniae* A45
stx - *S. pyogenes* MGAS1882

Orthologies

Gene		**length**	**sim**	**OL**	**insert**	**start**	**end**	**size (kb)**	**comment**
llt:CVCAS_0027a*	*ltr*C	107	1.000	107	PICI	36265	50942	14.7	
lla:L35519		107	1.000	107	PICI	35516	50948	15.4	
llm:llmg_2538		107	0.897	107	PICI	2478636	2491949	13.3	
lln:LLNZ_13110		107	0.897	107	PICI	2479452	2492765	13.3	
lld:P620_13395		106	0.796	98	PICI	2530163	2544289	14.2	
lls:lilo_1806		111	0.698	96	PICI	1963199	1976945	13.7	
myr:MYRA21_3441		179	0.344	64	NI**				
llt:CVCAS_0027b*	HP	136	1.000	136	PICI	36265	50942	14.7	
lla:L35867		136	1.000	136	PICI	35516	50948	15.4	
llm:llmg_2537		136	0.949	136	PICI	2478636	2491949	13.3	
lln:LLNZ_13105		136	0.949	136	PICI	2479452	2492765	13.3	
llc:LACR_2255		136	0.904	136	PICI	2111920	2125925	14.0	
:llh_11410		136	0.904	136	PICI	2111920	2125925	14.0	
lgr:LCGT_1117		140	0.530	132	Prophage	1109676	1146622	36.9	
lgv :LCGL_1137		140	0.530	132	Prophage				
lld:P620_10790		140	0.534	133	Prophage	2078353	2118530	40.2	
llt:CVCAS_0027	*terS*	147	1.000	147	PICI	36265	50942	14.7	
lla:L36274		147	1.000	147	PICI	35516	50948	15.4	
llc:LACR_2256		147	0.959	147	PICI	2111920	2125925	14.0	
llr:llh_11415		147	0.959	147	PICI	2111920	2125925	14.0	
llm:llmg_2250		146	0.938	146	PICI	2211673	2228382	16.7	Has phage resistance
lln:llnZ_11610		146	0.938	146	PICI	2211673	2228382	16.7	Has phage resistance
lga:LGAS_0603		173	0.550	120	Prophage	600763	680457	79.7	Probably 2 prophages in tandem
lpl:lp_2423		169	0.530	115	Prophage	2163938	2203818	39.9	
llt:CVCAS_0028	HP	193	1.000	193	PICI	36265	50942	14.7	
lla:L36850		193	0.995	193	PICI	35516	50948	15.4	
llm:llmg_0029		193	0.902	193	PICI	32974	51647	18.7	
lln:llnZ_00140		193	0.902	193	PICI	32974	51647	18.7	
llc:LACR_2257		180	0.631	179	PICI	2116658	2128806	12.1	
llr:llh_11420		180	0.631	179	PICI	2111920	2125925	14.0	

Gene		length	sim	OL	insert	start	end	size	comment
llk:llkF_2464		179	0.609	179	PICI	2505047	2520783	15.7	
lld:P620_02915		133	0.568	139	PICI	527748	537492	9.7	Defective; has prohead protease
lgr:LCGT_1791		287	0.328	125	PICI	175594	178604	12.7	Has phage protease
lgv:LCGL_1812		287	0.328	125	PICI	175594	178604	12.7	Has phage protease
Mtr:mgr		1527	0.272	173	NI				
llt:CVCAS_0029	*rep*	542	1.000	542	PICI	36265	50942	14.7	
lla:L37667		542	1.000	542	PICI	35516	50948	15.4	
llm:llmg_0030		542	0.985	542	PICI	32974	51647	18.7	
lln:llnZ_00145		542	0.985	542	PICI	32974	51647	18.7	
lld:P620_13370		542	0.954	542	PICI	2530163	2544289	14.2	
llk:llkF_2463		542	0.941	542	PICI	2505047	2520783	15.7	
llr:llh_12835		544	0.930	542	PICI	2352357	2368410	16.0	
llc:LACR_2258		542	0.917	542	PICI	2111920	2125925	14.0	
llw:kw2_1828		487	0.495	489	Prophage	1878501	1919139	40.6	
smb:smi_0425		534	0.458	509	Prophage	398225	440170	41.9	
stx:MGAS1882_1149		491	0.452	489	Prophage	1099359	1143106	46.7	
llt:CVCAS_0030	*pri*	264	1.000	264	PICI	36265	50942	14.7	
lla:L39306		264	1.000	264	PICI	35516	50948	15.4	
llc:LACR_2259		264	0.962	264	PICI	2111920	2125925	14.0	
llm:llmg_2253		264	0.958	264	PICI	2211673	2228382	16.7	Has phage resistance
lln:llnZ_11625		264	0.958	264	PICI	2211673	2228382	16.7	Has phage resistance
llr:llh_11430		264	0.951	264	PICI	2111920	2125925	14.0	
llk:llkF_2462		264	0.920	264	PICI	2505047	2520783	15.7	
lld:P620_11530		264	0.917	264	PICI	2217535	2230575	13.0	
lcb:LCABL_30870		273	0.323	266	Hybrid	3030347	3044256	13.9	PICI-prophage hybrid, with entire capsid module
llt:CVCAS_0031	HP	109	1.000	109	PICI	36265	50942	14.7	
lla:L40104		109	1.000	109	PICI	35516	50948	14.4	
llm:llmg_2532		109	0.954	109	PICI	2478636	2491949	13.3	
lln:llnZ_13075		109	0.954	109	PICI	2479452	2492765	13.3	
llr:llh_11435		111	0.963	108	PICI	2111920	2125925	14.0	
lld:P620_11535		109	0.944	108	PICI	2217535	2230575	13.0	
llk:llkF_2461		109	0.917	109	PICI	2505047	2520783	15.7	
llc:LACR_2260		111	0.907	108	PICI	2116658	2128806	12.1	
lgr:LCGT_1789		109	0.562	105	PICI	175594	178604	12.7	Has phage protease
lgv:LCGL_1810		109	0.562	105	PICI	75594	178604	12.7	Has phage protease
stk:STP_1275		115	0.321	81	Prophage	1375102	1412591	37.5	
llt:CVCAS_0032	HP	64	1.000	64	PICI	36265	50942	14.7	
lla:L200001		64	1.000	64	PICI	35516	50948	14.4	
llk:llkF_2460		64	0.969	64	PICI	2505047	2520583	15.0	
lld:P620_02875		64	0.953	64	Hybrid	521256	537492	16.2	Has 2 integrases
llr:llh_12820		104	0.344	61	PICI	2352357	2368410	16.0	
llt:CVCAS_0033	HP	79	1.000	79	PICI	36265	50942	14.7	
lla:L40862		79	0.987	79	PICI	35516	50948	14.4	
lld:P620_11550		79	0.949	79	PICI	2217535	2230575	13.0	
llk:llkF_2458		79	0.924	79	PICI	2505047	2520583	15.0	
llc:LACR_2262		79	0.924	79	PICI	2116658	2128806	12.2	
llm:llmg_0034		79	0.911	79	PICI	32974	51647	18.6	
lln:llnZ_00165		79	0.911	79	PICI	32974	51647	18.6	
llr:llh_11440		79	0.848	79	PICI	2111920	2125925	14.0	
lgr:LCGT_1792		59	0.426	61	PICI	1755947	1768604	12.7	
lgv:LCGL_1813		59	0.426	61	PICI	1769776	1782433	12.7	

Gene		length	sim	OL	insert	start	end	size	comment
evi:Echvi_3649		1172	0.387	75	NI				
llt:CVCAS_0034	HP	173	1.000	173	PICI	36265	50942	14.7	
lla:L41670		173	1.000	173	PICI	35516	50948	14.4	
llm:llmg_2260		173	0.789	171	PICI	2211673	2228382	16.7	Has phage resistance
lln:llnZ_11660		173	0.789	171	PICI	2211673	2228382	16.7	Has phage resistance
lld:P620_01795		173	0.789	171	PICI	321043	333060	12.0	3' end uncertain
llk:llkF_2456		174	0.759	174	PICI	2505047	2520583	15.0	
llc:LACR_2265		201	0.724	170	PICI	2116658	2128806	12.2	
llr:llh_11460		201	0.679	168	PICI	2111920	2125925	14.0	
mar:MAE_57910		265	0.313	134	NI				
cbr:CBG10226		806	0.312	112	No map				Caenorhabditis
llt:CVCAS_0035	HP	64	1.000	64	PICI	36265	50942	14.7	
lla:L42195		64	1.000	64	PICI	35516	50948	14.4	
lld:P620_02515		64	0.969	64	PICI	453940	463397	9.3	Defective
llr:llh_11465		64	0.953	64	PICI	2111920	2125925	14.0	
lls:lilo_1816		73	0.922	64	PICI	1963199	1976945	13.7	
llk:llkF_2455		64	0.938	64	PICI	2505047	2520583	15.0	
llc:LACR_2266		64	0.938	64	PICI	2116658	2128806	12.2	
llm:llmg_0037		64	0.922	64	PICI	32974	51647	18.6	
lln:llnZ_00180		64	0.922	64	PICI	32974	51647	18.6	
llt:CVCAS_0036	*reg*	246	1.000	246	PICI	36265	50942	14.7	
lla:L42465		246	1.000	245	PICI	35516	50948	14.4	
llm:llmg_2527		230	0.327	196	PICI	2478636	2491949	13.3	
lln:llnZ_13050		230	0.327	196	PICI	2479452	2492765	13.3	
lld:P620_01785		230	0.321	196	PICI	321043	330910	9.9	
llk:llkF_2454		230	0.306	196	PICI	2505047	2520583	15.0	
lgr:LCGT_1785		232	0.298	188	PICI	1755947	1768604	12.7	
lgv:LCGL_1806		232	0.298	188	PICI	1769776	1782433	12.7	
efm:M7W_2070		241	0.367	120	Prophage	1889045	1924864	39.5	
ehr:EHR_09450		241	0.358	120	Prophage	1847859	1882330	34.5	
llt:CVCAS_0037	HP	246	1.000	246	PICI	36265	50942	14.7	
lla:L42465		246	1.000	245	PICI	35516	50948	14.4	
llm:llmg_2527		230	0.327	196	PICI	2478636	2491949	13.3	
lln:llnZ_13050		230	0.327	196	PICI	2479452	2492765	13.3	
lld:P620_01785		230	0.321	196	PICI	321043	330910	9.9	
llk:llkF_2454		230	0.306	196	PICI	2505047	2520583	15.0	
lgr:LCGT_1785		232	0.298	188	PICI	1755947	1768604	12.7	
lgv:LCGL_1806		232	0.298	188	PICI	1769776	1782433	12.7	
efm:M7W_2070		241	0.367	120	Prophage	1889045	1924864	39.5	
ehr:EHR_09450		241	0.358	120	Prophage	1847859	1882330	34.5	
llt:CVCAS_0038	HP	128	1.000	246	PICI	36265	50942	14.7	
lla:L43680		128	1.000	128	PICI	35516	50948	14.4	
llr:llh_12770		128	0.781	128	PICI	2352357	2368410	16.0	
llc:LACR_2269		127	0.701	127	PICI	2116658	2128806	12.2	
llm:llmg_0041		140	0.550	129	PICI	32974	51647	18.6	
lln:llnZ_00200		146	0.550	129	PICI	32974	51647	18.6	
llw:kw2_0886		118	0.505	111	PICI	903732	908883	5.1	Defective PICI? Has lysin gene
lld:P620_12480		141	0.454	130	Prophage	2348850	2374861	26.0	Defective prophage, 2 int genes, no tail
lgr:LCGT_1146		114	0.470	115	Prophage	1110376	1146622	36.2	
lgv:LCGL_1166		114	0.470	115	Prophage				
lli:lli_0637		115	0.461	115	NI				

Gene		length	sim	OL	insert	start	end	size	comment
efq:DR75_1672		110	0.456	114	Prophage	1660409	1695776	35.4	
llt:CVCAS_0039	reg	184	1.000	184	PICI	36265	50942	14.7	
lla:L44085		184	1.000	184	PICI	35516	49727	14.2	
llm:llmg_0042		184	0.924	184	PICI	32974	51647	18.6	
lln:llnZ_00205		184	0.924	184	PICI	32974	51647	18.6	
llr:llh_12765		184	0.913	184	PICI	2352357	2368410	16.0	
llc:LACR_2270		183	0.799	184	PICI	2116658	2128806	12.2	
lld:P620_12940		98	0.887	97	PICI	2464618	2479987	15.4	
lgr:LCGT_0311		196	0.395	195	NI				
lgv:LCGL_0311		196	0.395	195	NI				
llk:llkF_1033		187	0.343	181	Prophage	1066254	1100809	34.6	
llt:CVCAS_0040	reg	64	1.000	64	PICI	36265	50942	14.7	
lla:L45035		64	1.000	64	PICI	35516	49727	14.2	
llm:llmg_0044		64	0.984	64	PICI	32974	51647	18.6	
lln:llnZ_00215		64	0.984	64	PICI	32974	51647	18.6	
llr:llh_12760		64	0.984	64	PICI	2352357	2368410	16.0	
llc:LACR_2271		64	0.969	64	PICI	2116658	2128806	12.2	
lld:P620_12935		64	0.969	64	PICI	2464618	2479987	15.4	
llk:llkF_2448		64	0.938	64	PICI	2505047	2520583	15.0	
loa:LOAG_08304		788	0.372	43					eye worm gene
llt:CVCAS_0041	reg	53	1.000	53	PICI	36265	50942	14.7	
lla:L45351		80	0.981	53	PICI	35516	49727	14.2	
llm:llmg_0047		80	0.962	53	PICI	32974	51647	18.6	
lln:llnZ_00240		80	0.962	53	PICI	32974	51647	18.6	
cat:CA2559_02645		294	0.368	38	NI				
llt:CVCAS_0042	HP	74	1.000	74	PICI	36265	50942	14.7	
lla:L45702		74	1.000	74	PICI	35516	49727	14.2	
llm:llmg_0048		74	1.000	74	PICI	32974	51647	18.6	
lln:llnZ_00245		70	1.000	70	PICI	32974	51647	18.6	
cso:CLS_07510		55	0.491	53	NI				
ckl:CKL_1132		56	0.446	56	NI				
ere:EUBREC_3588		55	0.462	52					Tiny defective prophage remnant?
mer:H729_05640		69	0.441	59	NI				
cdf:CD630_05860		59	0.453	53	NI				
lmc:Lm4b_00359		62	0.463	54	NI				
llt:CVCAS_0043	HP	108	1.000	108	PICI	36265	50942	14.7	
lld:P620_13295		108	1.000	108	PICI	2530163	2544289	14.1	
llm:llmg_0051		108	1.000	108	PICI	32974	51647	18.6	
lln:llnZ_00260		108	1.000	108	PICI	32974	51647	18.6	
llk:llkF_2449		108	0.963	108	PICI	2505047	2520583	15.0	
llc:LACR_C39		120	0.627	102					Transposon on plasmid
lli:lli_p6024		120	0.627	102					Transposon on plasmid
llr:llh_13780		117	0.618	102					Transposon on plasmid
lpq:AF91_13470		105	0.519	106	NI				
lca:LSEI_2757		105	0.519	106	NI				
llt:CVCAS_0044	HP	54	1.000	54	PICI	36265	50942	14.7	
lla:L200004		54	1.000	54	PICI	35516	50948	14.4	
lld:P620_12920		54	1.000	54	PICI	2463882	2479987	16.1	
llm:llmg_0053		54	1.000	54	PICI	32974	51647	18.6	
lln:llnZ_00265		54	1.000	54	PICI	32974	51647	18.6	

Gene		length	sim	OL	insert	start	end	size	comment
llk:llkF_2446		54	0.926	54	PICI	2505047	2520583	15.0	
llw:kw2_2347		53	0.731	52	NI				
llc:LACR_2598		53	0.712	52	NI				
lli:lli_2241		53	0.712	52	NI				
llr:llh_13190		53	0.712	52	NI				
lls:lilo_0576		66	0.558	52	NI				
lgr:LCGT_1418		64	0.560	50	NI				
lgv:LCGL_1439		64	0.560	50	NI				
llt:CVCAS_0045	HP	101	1.000	101	PICI	36265	50942	14.7	
lla:L47979		101	1.000	101	PICI	35516	50948	14.4	
llm:llmg_0054		101	1.000	101	PICI	32974	51647	18.6	
lln:llnZ_00270		101	1.000	101	PICI	32974	51647	18.6	
llk:llkF_2445		101	0.990	101	PICI	2505047	2520583	15.0	
lgr:LCGT_1323		99	0.494	87	NI				
lgv:LCGL_1344		99	0.494	87	NI				
lcn:C270_07595		99	0.354	99	NI				
llt:CVCAS_0046	*int*	394	1.000	394	PICI	36265	50942	14.7	
lla:L48477		394	0.997	394	PICI	35516	50948	14.4	
llm:llmg_0055		394	0.997	394	PICI	32974	51647	18.6	
lln:llnZ_00275		394	0.997	394	PICI	32974	51647	18.6	
lld:P620_01770		398	0.515	396	PICI	321043	333060	12.0	3’end uncertain
llc:LACR_0301		398	0.513	396	PICI	278489	289370	10.9	3’end uncertain
llk:llkF_2008		399	0.477	396	PICI	2063071	2074196	11.1	
lls:lilo_1819		410	0.470	396	PICI	1963199	1976945	13.7	
llr:llh_10885		393	0.471	397	Prophage	1973375	2011192	37.8	
lgr:LCGT_1777		396	0.415	393	PICI	1755947	1768604	12.5	
lgv:LCGL_1798		396	0.415	393	PICI	1769776	1782433	12.6	
lli:lli_1862		343	0.444	347	Prophage	1846819	1857113	10.9	Defective. Has *dut*
snu:SPNA45_01857		388	0.414	391	PICI	1886061	1902018	15.9	
smb:smi_2017		388	0.376	391	PICI	2057587	2070717	13.1	

*ORFs that are not annotated in LlCICV56-1 but are annotated in LlCIIL1403-1, which has the identical sequence in that region.

Table S7. Putative phage-inducible chromosomal islands of Gram-positive cocci.

PICI	Strain	Accession number (Genomic location)	Size (kb)	*att* site core	Accessory genes[b]
EfCIV583	*E. faecalis* V583	AE016830 (2816732-2829670)	12.9	TATTAATGAAACAACGTG	UvrB protein; Cold-shock protein
EfCILA3B-2	*E. faecalis* LA3B-2	ATJC01000082 (5397-18340)	12.9	TAAACTGTAAGTTTAGT	Cold-shock protein
EfCI918	*E. faecalis* 918	AVNY01000040 (1-12614)	12.6	TATTAATGAAACAACGTG	UvrB protein; Cold-shock protein
EfCIB16457	*E. faecalis* B16457	AIIL01000003 (236044-248787)	12.7	TATTAATGAAACAACGTG	UvrB protein; Cold-shock protein
EfCISF105	*E. faecalis* SF105	AJEE01000013 (53058- 65413)	12.3	TATTAATGAAACAACGTG	UvrB protein; Cold-shock protein
LICIbIL310	*L. lactis* IL1403	AE005176 (34907-49863)	14.9	CAAAAAAACACTGATTGAATGCCGTATG	Enterocin immunity (EntA); LtrA
LICIbIL312	*L. lactis* IL1403	AE005176 (502595-517773)	15.1	GAAAGACGCAGTTAAATAATTATAGCTAT	Peptidase_M48; Cold shock protein
LICINZ9000-1	*L lactis* - cremoris NZ9000	CP002094 (32370-51783)	19.4	CAAAAAAACACTGATTGAATGCCGT	*bcn*A; IS712A; Non-specific endonuclease
LICINZ9000-2	*L. lactis* - cremoris NZ9000	CP002094 (2210232-2228483)	18.3	TAGAACTATGTTAAAA	Abortive phage resistance
LICI-NZ9000-3	*L lactis* - cremoris NZ9000	CP002094 (2482743-2492765)	10.3	ATTCACTTGAGCAATGAATATA	LtrA
LICISK11	*L. lactis* - cremoris SK11	CP000425 (2115981-2128907)	12.9	TAGAACTATGTTAAAA	DNA/RNA non-specific endonuclease
LICICV56-1	*L. lactis* - lactis CV56	CP002365 (36265-50942)	14.7	CAAAAAAACACTGATTGAATGCCGTATG	Enterocin immunity (EntA); LtrA
LICI-CV56-2	*L. lactis* - lactis CV56	CP002365 (1723782-1733895)	10.1	J_LTAAAAAATAGGACCTAAGACTGATGA J_RTAAAAAATCAGACCTAAGACTTATGA	Cold shock protein
LICIKLDS-2	*L. lactis* - *lactis* KLDS 4.0325	CP006766 (1903120-1915186)	12.1	TCAGACCTAAGACTGATGATATAAAG	
LICI-KLDS-3	*L. lactis* – *lactis* KLDS 4.0325	CP006766 2464508-2479769	15.2	GCTATAATAAAACTATAT	Prohead protease
LICIA76-1	*L. lactis* - cremoris A76	CP003132 (2111244-2126036)	14.8	TTTTAACATAGTTCTATTTTATCACA	
LICIA76-2	*L. lactis* - cremoris A76	CP003132 (2352993- 2368188)	15,2	TAAAACTATA	
LICIKF147	*L. lactis* - lactis KF147	CP001834 (2505680- 2520561)	14.9	TAAAACTATA	BcnA-imm; Pyrimidine dimer DNA glycosylases
MG1363-1	*L. lactis* - lactis MG1363	AM406671 (32370-51783)	19.4	CAAAAAAACACTGATTGAATGCCGT	*bcn*A; IS712A; Non-specific endonuclease

PICI	Strain	Accession number (Genomic location)	Size (kb)	*att* site core	Accessory genes[b]
SpnCI-Taiwan-0.03	*S. pneumoniae* Taiwan	NC_012469 3563-23357	19.8	CCCTTTTTGTGTTA	
SpnCI-ST556-0.03	*S. pneumoniae* ST556	CP003357 3563-21646	18.1	CCCTTTTTGTGTTA	
SpnCI-Taiwan-0.2	*S. pneumoniae* Taiwan	NC_012469 197987-210888	12.9	TACAAAATCGGCTTTTTT	
SpnCI-Tigr4-1.06	*S. pneumoniae* Tigr4	NC_003028 1063231-1073321	10.1	CCTAACAAAAC	TA system
SpnCI-INV104-1.06	*S. pneumoniae* INV104	FQ312030 1070841-1079518	8.7	CCTTAAAAAATAA	
SpnCI-A45-1.9	*S. pneumoniae* A45	NC_018594.1 1887286-1902018	14.7	GCCCATACAAACCCCATA	DNA-damage-inducible protein D
SsuCI-TL13	*S. suis* TL-13	CP003993 1339840-1351861	12.0	CTTGAAAAAATAA	
SolCI-AS1.3089-0.6	*S. oligofermentans* AS1.3089	CP004409.1 602331`- 612649	10.3	CTTGAAAAAATAA	TA system
SpnCI-TCH8341-0.45	*S. pneumoniae* TCH8431/19A	CP001993 (440193-453094)	12.9	ATTATACTACAAAATCGGC	
SpnCI-TCH8341-0.25	*S. pneumoniae* TCH8431/19A	CP001993 (248868-267453)	18.6	TAACACAAAAAGGG	DNA-damage-inducible protein D Inserted plasmid

[a]NI: Not identified.
[b]The identities of the accessory genes are based on annotations; none has been tested experimentally.

Table S8. Role of the cloned *cos* site in pAGEnt transfer[a].

Donor strain	Cloned site *cos*	Plasmid titre[b]
JP14198	Empty vector	< 10
JP14199	Phage bIL286	4.5×10^2
JP14203	LICI-bIL310	1.2×10^2

[a]The means of results from three independent experiments are shown. Variation was within ±5% in all cases.
[b]No. of transductants/ml induced culture, using IL1403 as recipient strain.

Table S9. Orthology analysis of SpnCl6706B

Abbreviations

SPN = *Streptococcus pneumoniae*

SPNN = *S. pseudopneumoniae*

SMB = *S. mitisB*

SPS = *S. parasanguinis*

STH = *S. thermophilus*

SPU = *S. parauberis*

SGL = *S. galactiae*

SUB = *S. uberis*

SSU = S. suis

SLT = *S lutetiensis*

SAG = *S. agalactiae*

NCE = *Nocema seranae*

SPY = *S. pyogenes*

LMO = *Listeria monocytogenes*

FSU = Fibrinobacter succinogenes

PPH = Pelodictyon phaeoclathratiforme

spn: SPN TIGR4
std: SPPN
snb: SPN 670-6B
snc: SPN TCH8431/19A
snd: SPN ST556
sni: SPN INV104
snt: SPN Taiwan19f-14
snu: SPN A45
spi: SPY MGAS10750
smb: SMB
scp: SPS ATCC 15912
stc: STH CNRZ1066
ste: STH LMD-9
stn: STH ND03
stw: STH MN-ZLW-002
spv: SPN Hungary19A 6
stk: SPU
*sthe:*STH ASCC 1725
sga: SGL UCN34
sgg: SGL ATCC BAA-2069
sgt: SGL ATCC 43143
sjj: SPN JJA
sagm:SAG 09mas018883
sub: SUB
sui: SSU T15
ssut: SSU TL-13
sst: SSU ST3
ssuy: SSU YB51
slu: SLT KE3
sak: SAG A909
sagt:SAG COH1
nce: NCE
stx: SPY MGAS1882
spnn:SPN A026
lmn: LMO 08-5778
fsc: FSU
pph: PPH
drs: =Dehalobacter restrictus
dec: =Dehalobacter sp. CF
ded: =Dehalobacter sp. DCA
rbr: =Ruminococcus bromii

Orthologs

Gene	length	sim	OL	insert	site	size	site
snb:SP670_0026 (388 a.a.)	*int*						
snc:HMPREF0837_10291	388	0.874	388	PICI	0.25	14.9	*dna*A
snd:MYY_0032	398	0.874	388	PICI	0.01	16.1	*dna*A
snt:SPT_0037	388	0.874	388	PICI	0.01	14.9	*dna*A
smb:smi_2017	388	0.546	388	PICI	2.05	12.2	sugar hydrolase
scp:S. parasangui:HMPREF0833_11010	388	0.549	388	PICI	1.02	10.6	*cna*
ste:STER_0829	388	0.531	388	PICI	0.74	9.4	*pab*B
stc:str0783	388	0.531	388	PICI	0.73	8.1	*pab*B
sthe:T303_05120	388	0.531	388	PICI	0.93	7.7	*dlt*D
stn:STND_0774	388	0.531	388	PICI	0.75	10.2	*uvr*A
stw:Y1U_C0751	388	0.531	388	PICI	0.73	6.5	*pab*B
spv:SPH_0289	388	0.536	388	PICI	0.26	12.4	*uvr*A
stkpara:STP_1346	386	0.525	387	PICI	1.48	11	SAM
sga:GALLO_2149	388	0.525	387	PICI	2.2	11.9	*gsh*A
sgg:SGGBAA2069_c21460	388	0.525	387	PICI	2.2	11.8	*gsh*A
sgt:S GGB_2132	388	0.525	387	PICI	2.2	11.9	*gsh*A
sagm:BSA_21480	388	0.503	388	PICI	2.1	14.6	*rps*D
sub:SUB1840	388	0.505	388	PICI	1.8	12.6	*rps*D
slu:KE3_0026	381	0.518	388	PICI	0.03	11.0	*tyr*S
sak:SAK_2094	388	0.508	388	PICI	2.06	15.6	*rps*D
snb:SP670_0025 (284 a.a.)	*din*D						
snu:SPNA45_01858	274	0.993	271	PICI	1.9	14.7	*yes*MN
drs:DEHRE_03560	278	0.598	276	NI*			
dec:DCF50_p2453	278	0.583	276	NI			
ded:DHBDCA_p2442	278	0.583	276	NI			
rbr:RBR_05470	280	0.561	278	NI			
snc:HMPREF0837_10290	186	0.987	155	PICI	0.25	14.9	*dna*A
snd:MYY_0031	186	0.987	155	PICI	0.01	14.9	*dna*A
snt:SPT_0036	186	0.987	155	PICI	0.01	14.9	*dna*A
fsc:FSU_1649	279	0.570	270	NI			
fsu:Fisuc_1187	279	0.570	270	NI			
pph:Ppha_0973	362	0.525	276	NI			
xne:XNC1_0195	271	0.549	266	defective and rearranged PICI-like fragment xne:XNE =Xenorabdus nematophila			
snb:SP670_0024 (238 a.a.)	*rpr*						
snc:HMPREF0837_10283	246	0.983	237	PICI	0.25	18.1	*dna*N
snd:MYY_0025	246	0.983	237	PICI	0.01	16.1	*dna*A
snt:SPT_0029	238	0.983	237	PICI	0.01	14.9	*dna*A
snu:SPNA45_01863	255	0.745	239	PICI	1.9	15.4	*yes*MN
sjj:SPJ_1901	250	0.626	246	prophage			

Gene		length	sim	OL	insert	site	size	site
snb:SP670_0023 (50 a.a.)	*reg*							
snc:HMPREF0837_10282		50	1.000	50	PICI	0.25	18.1	*dna*N
snd:MYY_0024		50	1.000	50	PICI	0.01	16.1	*dna*A
snt:SPT_0028		50	1.000	50	PICI	0.01	14.9	*dna*A
saua:SAAG_02514		73	0.460	50	Proph	2.0	43.4	*hlb*
snb:SP670_0022 (194 a.a.)	*ant*R							
ssq:SSUD9_2185		248	0.568	192	PICI	2.1		*rec*F
sst:SSUST3_2012		248	0.568	192	PICI	2.0	10.7	*rec*F
ssuy:YB51_9970		248	0.568	192	PICI	2.0	10.7	*rec*F
spy:SPy_2127		255	0.457	184	PICI	1.8	12.9	*mut*S/L
stz:SPYALAB49_001792		193	0.457	184	PICI	1.7	13.2	*mut*S/L
sdq:SDSE167_2198		255	0.446	184	PICI	2.0	19.6	*mut*S/L
sdg:SDE12394_10830		187	0.415	183	PICI	2.1	14.8	*rps*D
sds:SDEG_2138		187	0.415	183	PICI	2.1	11.2	*rps*D
ssui:T15_1236		261	0.425	153	proph	1.2	33.0	*rps*A
spb:M28_Spy1861		188	0.416	185	PICI	1.8	10.8	*rps*D
sph:MGAS10270_Spy1948		188	0.416	185	PICI	1.8	10.7	*rps*d
stg:MGAS15252_1718		188	0.416	185	PICI	1.7	12.2	*rps*D
clb:Clo1100_2443		239	0.453	148	proph	2.7	26.3	defective
ctc:CTC01071		270	0.484	126	proph	1.3	36.2	
snb:SP670_0021 (208 a.a.)								
snc:HMPREF0837_10490		208	0.952	208	PICI	0.44	15.4	*uvr*A
snt:SPT_0226		208	0.952	208	PICI	0.2	12.4	*uvr*A
spnn:T308_00870		208	0.952	208	PICI	1.6	12.4	*uvr*A
spv:SPH_0293		208	0.952	208	PICI	0.26	12.4	*uvr*A
snd:MYY_0262		203	0.951	203	PICI	0.24	12.4	*uvr*A
smb:smi_2013		208	0.841	208	PICI	2.1	10.9	sugar hydrolase
std:SPPN_01200		209	0.833	209	PICI	0.2	12.4	*mnm*A
scp:HMPREF0833_11006		208	0.812	208	PICI	1.02	10.6	*cna*
snu:SPNA45_01864		200	0.779	199	PICI	1.9	15.4	*yes*MN
ssut:TL13_1976		208	0.721	208	PICI	2.0	15.1	*rec*F
spn:SP_1134		205	0.744	203	PICI	1.1	10.1	enolase
ssui:T15_1587		206	0.677	201	PICI	1.60	10.0	*fab*G
snb:SP670_0020 (55 a.a.)								
ssui:T15_1694		54	0.698	53	proph	1.7	34.5	*hup*B
spy:SPy_0978		56	0.472	53	proph	0.8	40.2	*cps*FQ/*mut*X
stz:SPYALAB49_001012		56	0.472	53	proph	1.0	47.3	*glg*P
lld:P620_11580		61	0.412	51	PICI	2.2	13.0	*cut*C
spb:M28_Spy1797		64	0.491	55	PICI	1.8	13.3	*mut*S/L
sph:MGAS10270_Spy1881		64	0.491	55	PICI	1.8	13.9	*mut*S/L
sagm:BSA_21410		57	0.439	57	PICI	2.1	14.6	*rps*D
snc:HMPREF0837_10278		77	0.473	55	PICI	0.25	18.1	*dna*N
snd:MYY_0020		57	0.473	55	PICI	0.01	16.1	*dna*A
snt:SPT_0023		57	0.473	55	PICI	0.01	14.9	*dna*A
snu:SPNA45_01866		57	0.473	55	PICI	1.9	15.4	*yes*MN
aoe:Clos_1887		60	0.357	56	NI			

Gene	length	sim	OL	insert	site	size	site
snb:SP670_0019 (142 a.a.)							
snc:HMPREF0837_10277	142	0.993	142	PICI	0.25	18.1	*dna*N
snd:MYY_0019	142	0.993	142	PICI	0.01	16.1	*dna*A site
snt:SPT_0022	142	0.993	142	PICI	0.01	14.9	*dna*A
spb:M28_Spy1799	189	0.384	138	PICI	1.80	14.0	*mut*S/L
snb:SP670_0018 (48 a.a.) HP (4 hits only)							
snc:HMPREF0837_10276	48	1.000	48	PICI	0.25	18.1	*dna*N
snd:MYY_0018 hypothet	48	1.000	48	PICI	0.01	16.1	*dna*A
snt:SPT_0021 hypothet	48	1.000	48	PICI	0.01	14.9	*dna*A
rdn:HMPREF0733_10312	1301	0.396	48	NI			*rpo*C
snb:SP670_0017							
snc:HMPREF0837_10275	49	0.980	49	PICI	0.25	18.1	*dna*N
snb:SP670_0016 (38 a.a.) (4 hits only)							
snc:HMPREF0837_10274	46	1.000	29	PICI	0.25	18.1	*dna*N
snd:MYY_0017	46	1.000	29	PICI	0.01	16.1	*dna*A
snt:SPT_0020	46	1.000	29	PICI	0.01	14.9	*dna*A
ssk:SSUD12_0495	155	0.862	29	PICI	0.5	11.6	*lys*S
snb:SP670_0015 (166 a.a.)							
snc:HMPREF0837_10273	170	0.910	166	PICI	0.25	18.1	*dna*N
snd:MYY_0016	170	0.910	166	PICI	0.03	13.2	*dna*A
snt:SPT_0019	166	0.910	166	PICI	0.01	14.9	*dna*A
std:SPPN_01215	169	0.800	165	PICI	0.2	12.4	*mnm*A
eol:Emtol_1450	514	0.342	73	NI			
snb:SP670_0014 (46 a.a.)							
snc:HMPREF0837_10272	46	0.913	46	PICI	0.25	18.1	*dna*N
snt:SPT_0018	46	0.913	46	PICI	0.01	14.9	*dna*A
smb:smi_2011	66	0.826	46	PICI	2.1	10.9	sugar hydrolase
snd:MYY_0015	39	0.923	39	PICI	0.01	16.1	*dna*A
spnn:T308_00880	66	0.674	46	PICI	1.6	12.4	*uvr*A
spv:SPH_0295	66	0.674	46	PICI	0.26	12.4	*uvr*A
spn:SP_1136	148	0.630	46	PICI	1.1	11.5	enolase
scp:HMPREF0833_11005	66	0.630	46	PICI	1.02	10.6	*cna*
sni:INV104_09830	59	0.590	39	PICI	1.1	8.2	enloase
ssui:T15_1589	63	0.571	42	PICI	1.60	10.0	*fab*G
snb:SP670_0013 (71 a.a.)							
snc:HMPREF0837_10270	71	0.986	71	PICI	0.25	18.1	*dna*N
snt:SPT_0016	71	0.986	71	PICI	0.01	14.9	*dna*N
std:SPPN_01220	71	0.901	71	PICI	0.2	13.2	*dna*N
scp:HMPREF0833_11004	71	0.859	71	PICI	1.02	12.4	*mnm*A
smb:smi_2010	71	0.845	71	PICI	2.1	10.9	sugar hydrolase
spnn:T308_00885	71	0.789	71	PICI	1.6	12.4	*mnm*A
spv:SPH_0296	71	0.789	71	PICI	0.26	12.4	*uvr*A
dav:DESACE_04385	580	0.319	69	NI			
scg:SCI_1442	64	0.349	63	PICI		10.5	*gpm*A

Gene	length	sim	OL	insert	site	size	site
scon:SCRE_1399	64	0.349	63	PICI		10.5	*gpm*A
scos:SCR2_1399	64	0.349	63	PICI		0.5	*gpm*A
bvu:BVU_2093	583	0.432	37	NI			
snb:SP670_0012 (97 a.a.)							
snc:HMPREF0837_10269	97	0.959	97	PICI	0.25	18.1	*dna*N
std:SPPN_01225	98	0.928	97	PICI	0.2	12.4	*mnm*A
snu:SPNA45_01869	94	0.957	93	PICI	1.9	14.7	*yes*MN has dinD
smb:smi_2009	97	0.856	97	PICI	2.1	10.9	sugar hydrolase
snd:MYY_0013	68	0.956	68	PICI	0.01	16.1	*dna*A
snt:SPT_0015	58	0.948	58	PICI	0.01	14.9	*dna*A
nce:NCER_100853	221	0.302	96	NI			
snb:SP670_0011 (113 a.a.)							
snc:HMPREF0837_10268	113	1.000	113	PICI	0.25	18.1	*dna*N
snd:MYY_0012	113	1.000	113	PICI	0.01	16.1	*dna*A
snt:SPT_0014	113	1.000	113	PICI	0.01	14.9	*dna*N
snu:SPNA45_01870	113	1.000	113	PICI	1.9	15.4	*yes*MN
std:SPPN_01230	113	1.000	113	PICI	0.2	12.4	*mnm*A
smb:smi_2008	113	0.841	113	PICI	2.1	10.9	sugar hydrolase
spnn:T308_00890	113	0.841	113	PICI	1.6	12.4	*uvr*A
spv:SPH_0297	113	0.841	113	PICI	0.26	12.4	*uvr*A
scp:HMPREF0833_11003	113	0.735	113	PICI	1.02	10.6	*cna*
sdq:SDSE167_2216	109	0.667	111	PICI	2.0	19.6	*mut*L/S
spa:M6_Spy1814	109	0.667	111	PICI	1.8	13.3	*mut*L/S
stg:MGAS15252_1660	109	0.667	111	PICI	1.7	12.2	*rps*D
snb:SP670_0010 (90 a.a.)							
snc:HMPREF0837_10267	90	1.000	90	PICI	0.25	18.1	*dna*N
snt:SPT_0013	90	1.000	90	PICI	0.01	14.9	*dna*N
std:SPPN_01235	90	0.967	90	PICI	0.2	12.4	*mnm*A
snd:MYY_0266	93	0.708	89	PICI	0.24	12.4	*uvr*A
spnn:T308_00895	93	0.708	89	PICI	1.6	12.4	*uvr*A
spv:SPH_0298	93	0.708	89	PICI	0.26	12.4	*uvr*A
smb:smi_2007	91	0.674	89	PICI	2.1	10.9	sugar hydrolase site
scp:HMPREF0833_11002	91	0.607	89	PICI	1.02	10.6	*cna*
ssut:TL13_0170	92	0.539	89	PICI	0.14	13.1	*ack*A
sagt:GBSCOH1_1946	90	0.556	90	PICI	2.0	14.8	*rps*D
sst:SSUST3_2006	94	0.570	86	PICI	2.0	13.5	*rec*F
snb:SP670_0009 (286 a.a.) *pri*							
snc:HMPREF0837_10266	286	0.965	286	PICI	0.25	18.1	*dna*N
snd:MYY_0011	286	0.965	286	PICI	0.01	16.1	*dna*A
snt:SPT_0012	286	0.965	286	PICI	0.01	14.9	*dna*N
smb:smi_2006	286	0.930	286	PICI	2.1	10.9	sugar hydrolase
spnn:T308_00905	288	0.907	289	PICI	1.6	12.4	*uvr*A
spv:SPH_0299	288	0.907	289	PICI	0.26	12.4	*uvr*A
sagi:MSA_22030	289	0.810	289	PICI	2.1	17.4	*rps*D
spf:SpyM51773	288	0.741	290	PICI	1.8	11.9	*mut*S/L
spy:SPy_2135	285	0.718	287	PICI	1.8	12.9	*mut*S/L

Gene	length	sim	OL	insert	site	size	site
stz:SPYALAB49_001801	285	0.718	287	PICI	1.8	14.3	*mut*S/L
spa:M6_Spy1816	285	0.718	287	PICI	1.8	13.3	*mut*S/L
sak:SAK_2084	285	0.725	287	PICI	2.1	15.6	*rps*D
snb:SP670_0008 (492 a.a.) *rep*							
spf:SpyM51774	500	0.915	492	PICI	1.8	11.9	*mut*S/L
spv:SPH_0300	489	0.918	488	PICI	0.26	12.4	*uvr*A
snc:HMPREF0837_10498	489	0.916	488	PICI	0.44	15.4	*uvr*A
snd:MYY_0268	489	0.916	488	PICI	0.26	12.4	*uvr*A
snt:SPT_0233	489	0.916	488	PICI	0.2	12.4	*uvr*A
spnn:T308_00910	489	0.916	488	PICI	1.6	12.4	*uvr*A
sagi:MSA_22020	498	0.897	485	PICI	2.1	17.4	*rps*D
sak:SAK_2083	480	0.900	480	PICI	2.1	15.6	*rps*D
sthe:T303_05075	501	0.758	483	PICI	0.93	7.7	*dlt*D
stc:str0775	500	0.754	483	PICI	0.73	8.1	*pab*B
ste:STER_0819	501	0.754	483	PICI	0.74	9.4	*pab*B
stn:STND_0765	501	0.754	483	PICI	0.75	10.2	*uvr*A
ssui:T15_1593	507	0.484	494	PICI	1.60	10	*fab*G
spi:MGAS10750_Spy1910	498	0.474	485	PICI	1.87	13.5	*mut*S/L
Snb:SP670_0007 (167aa)							
snc:HMPREF0837_10500	176	0.889	162	PICI	0.44	15.4	*uvr*A
snd:MYY_0270	176	0.889	162	PICI	0.26	12.4	*uvr*A
snt:SPT_0235	176	0.889	162	PICI	0.2	12.4	*uvr*A
spnn:T308_00920	176	0.889	162	PICI	1.6	12.4	*uvr*A
spv:SPH_0302	176	0.889	162	PICI	0.25		*uvr*A
snu:SPNA45_01874	142	0.923	142	PICI	1.88		*yes*MN
smb:smi_2003	153	0.865	141	PICI	2.05		fucosidase
ssq:SSUD9_2176	204	0.529	172	PICI	2.16		*rec*F
sn:SSUST3_2003	182	0.529	172	PICI	2.00		*rec*F
sagi:MSA_21990	231	0.537	164	PICI	2.05		*rps*D
ssuy:YB51_9925	150	0.500	148	PICI	2.02		*rec*F
sub:SUB1829	205	0.426	162	PICI	2.01		*rps*D
Snb:SP670_0006 (167aa)							
snu:SPNA45_01875	167	1.000	167	PICI	1.88		*yes*MN
smb:smi_2002	167	0.934	167	PICI	2.05		fucosidase
snc:HMPREF0837_10501	169	0.879	165	PICI	0.44		*uvr*A
spnn:T308_00925	169	0.879	165	PICI	1.6		*uvr*A
spv:SPH_0303	169	0.879	165	PICI	0.25		*uvr*A
sagi:MSA_21980	163	0.722	162	PICI	2.05		*rps*D
ssui:T15_1595	166	0.548	166	PICI	1.60	10	*fab*G
sagm:BSA_21300	162	0.525	162	PICI	2.08	14.6	*rps*D
sak:SAK_2081	162	0.519	162	PICI	2.07	15.7	*rps*D
spa:M6_Spy1821	162	0.512	162	NI			

Gene	length	sim	OL	insert	site	size	site
Snb:SP670_0005 (130aa)							
smb:smi_2000	130	0.946	130	PICI	0.19	12.4	*mnm*N
std:SPPN_01285	140	0.311	132	PICI?	1.94	12.6	very poor annot
slu:KE3_2016	140	0.311	132	PICI	1.94	14.1	pgp
sga:GALLO_2136	140	0.303	132	PICI	2.23	11.4	
sgg:SGGBAA2069_c21340	140	0.303	132	PICI	2.18	11.8	
sgt:SGGB_2119	140	0.303	132	PICI	2.18	11.2	
lsa:LSA0600	113	0.350	80	PICI	0.60	8.1	

snb:SP670_0004 (52 a.a.) no matching protein in DB