ORIGINAL PAPER

Actinomycete integrative and conjugative elements

**Evelien M. te Poele · Henk Bolhuis ·
Lubbert Dijkhuizen**

Abstract This paper reviews current knowledge on actinomycete integrative and conjugative elements (AICEs). The best characterised AICEs, pSAM2 of *Streptomyces ambofaciens* (10.9 kb), SLP1 (17.3 kb) of *Streptomyces coelicolor* and pMEA300 of *Amycolatopsis methanolica* (13.3 kb), are present as integrative elements in specific tRNA genes, and are capable of conjugative transfer. These AICEs have a highly conserved structural organisation, with functional modules for excision/integration, replication, conjugative transfer, and regulation. Recently, it has been shown that pMEA300 and the related elements pMEA100 of *Amycolatopsis mediterranei* and pSE211 of *Saccharopolyspora erythraea* form a novel group of AICEs, the pMEA-elements, based on the unique characteristics of their replication initiator protein RepAM. Evaluation of a large collection of *Amycolatopsis* isolates has allowed identification of multiple pMEA-like elements. Our data show that, as AICEs, they mainly coevolved with their natural host in an integrated form, rather than being dispersed via horizontal gene transfer. The pMEA-like elements could be separated into two distinct populations from different geographical origins. One group was most closely related to pMEA300 and was found in isolates from Australia and Asia and pMEA100-related sequences were present in European isolates. Genome sequence data have enormously contributed to the recent insight that AICEs are present in many actinomycete genera. The sequence data also provide more insight into their evolutionary relationships, revealing their modular composition and their likely combined descent from bacterial plasmids and bacteriophages. Evidence is accumulating that AICEs act as modulators of host genome diversity and are also involved in the acquisition of secondary metabolite clusters and foreign DNA via horizontal gene transfer. Although still speculative, these AICEs may play a role in the spread of antibiotic resistance factors into pathogenic bacteria. The novel insights on AICE characteristics presented in this review may be used for the effective construction of new vectors that allows us to engineer and optimise strains for the production of commercially and medically interesting secondary metabolites, and bioactive proteins.

E. M. te Poele · L. Dijkhuizen (✉)
Department of Microbiology, Groningen Biomolecular
Sciences and Biotechnology Institute (GBB), University
of Groningen, Kerklaan 30, 9751 NN Haren,
The Netherlands
e-mail: l.dijkhuizen@rug.nl

H. Bolhuis
Department of Marine Microbiology, Centre for Estuarine
and Marine Ecology, Netherlands Institute of Ecology
(NIOO-KNAW), P.O. Box 140, 4400 AC Yerseke,
The Netherlands

Integrative and conjugative elements

Actinomycete integrative and conjugative elements (AICEs) belong to a broad class of integrative and conjugative elements (ICEs) (Burrus et al. 2002). The name ICE was initially proposed by Burrus et al. (2002) for a diverse group of mobile genetic elements (MGEs), which have both plasmid-and bacteriophage-like features. ICEs are present in all major divisions of bacteria and include mobile genomic islands and conjugative transposons. They are characterised by their prophage-like mode of maintenance, i.e. replication along with the host chromosome, and their ability to excise, conjugate to a new host and integrate in the host chromosome by site-specific recombination, irrespective of the specificity and mechanism of integration and conjugation (Burrus et al. 2002).

The ICE backbone is composed of three modules, which are involved in maintenance, conjugation and regulation (Burrus and Waldor 2004). ICEs often contain genetic elements, such as transposons and insertion sequences as well as genes encoding specific recombinases. These genetic elements and recombinases mediate the acquisition of additional modules encoding functions, such as resistance and metabolic traits, which confer a selective advantage to the host under certain environmental conditions (Burrus and Waldor 2004). For example, the 100 kb ICE SXT of *Vibrio cholerae*, the agent of cholera, encodes resistance to multiple antibiotics (Waldor et al. 1996).

Integration of ICEs is mediated by site-specific recombinases, mainly tyrosine recombinases, which function in many different hosts and have a minimal requirement for specific host factors. These features may contribute to the extremely broad host range of many ICEs and hence their important role in the spreading of antibiotic resistance genes (Mullany et al. 2002).

Prior to conjugative transfer, ICEs excise from the host chromosome to form a circular, mostly non-replicative molecule that is nicked at the origin of transfer and is transferred to recipient cells as an ssDNA intermediate. Excision and transfer of ICEs can be triggered by different factors, such as sub-inhibitory concentrations of tetracycline in the case of Tn916, the conjugative transposon from *Enterococcus faecalis* (Salyers et al. 1995), and the SOS response triggering transfer of SXT (Burrus and Waldor 2004). In the latter case, the SOS response causes release from repression by a phage lambda repressor orthologue SetR encoded by SXT and subsequent derepression of the excision and transfer functions.

Actinomycete integrative and conjugative elements

AICEs represent a special class of ICEs, because unlike other ICEs, they have the ability to replicate autonomously like a plasmid. Several AICEs have been described in literature (Table 1; Fig. 1a), but only a few of these have been characterised in detail. AICEs are maintained integrated in a specific tRNA gene in the host chromosome. The majority are self-transmissible and a number of them can mobilise chromosomal markers (Moretti et al. 1985; Vrijbloed 1996; Hopwood et al. 1984; Bibb et al. 1981; Brown et al. 1988b; Smokvina et al. 1988). However, genes encoding clear beneficial functions for the physiology of the hosts were not, or very rarely, found on these elements and only little is known about their distribution in actinomycetes and their roles in the evolution of their hosts. The elements pMEA300 of *Amycolatopsis methanolica*, pMEA100 of *Amycolatopsis mediterranei*, and pSE211 of *Saccharopolyspora erythraea* form a separate group within the AICEs, the pMEA-elements (Fig. 1a). A core characteristic of the pMEA-elements is their novel replication initiator protein RepAM, which is unrelated to other Rep proteins and has several unique DNA-binding properties (te Poele et al. 2006) (see below).

AICEs are present in diverse actinomycete genera

We recently found that AICEs are present in diverse actinomycete genera. Database searches of bacterial genome sequences revealed various putative AICEs in several actinomycete genomes (Fig. 1). AICEs were discovered by screening the genomes for integrase genes flanking tRNA genes with known integrase genes of previously known AICEs by batch Blast. Analysis of the *Sac. erythraea* genome sequence revealed the presence of two putative AICEs in addition to the previously identified pSE211 and pSE101 elements (te Poele et al. 2008) (Fig. 1b). The first of these new elements (∼20.4 kb) was designated pSE222 and encodes proteins related to those of the

Table 1 Overview of previously characterised AICEs and several element-specific functions

AICE	Strain	Reference[a]	Size (kb)	Element specific functions[b]				
				tRNA	Replication	Conjugation	Pock formation	CMA[c]
pMEA100	A. mediterranei	Moretti et al. (1985)	23.3	Phe	+	+	+	+
pMEA300	A. methanolica	Vrijbloed et al. (1994)	13.3	Ile	+	+	+	+
pSE211	Sac. erythraea	Brown et al. (1988b)	17.2	Phe	+	+	+	+
pSE101	Sac. erythraea	Brown et al. (1988a)	10.9	Thr	+	−	−	−
pMR2	M. rosaria	Hosted Jr et al. (2005)	11.2	Phe	+	nd[d]	nd	nd
pSAM2	S. ambofaciens	Pernodet et al. (1984)	10.9	Pro	+	+	+	+
SLP1	S. coelicolor	Bibb et al. (1981)	17.3	Tyr	+	+	+	+
pSG1	S. griseus	Cohen et al. (1985)	16.9	Ser	+	nd	nd	nd
pIJ110	S. parvulus	Hopwood et al. (1984)	13.6	nd	+	+	+	+
pIJ408	S. glaucescens	Hopwood et al. (1984)	15.05	nd	+	+	+	+

[a] Reference to paper first describing the element

[b] See text for references to element-specific functions

[c] CMA, chromosome mobilising ability

[d] nd, not determined

pMEA-like elements and of other AICEs. The second integrated element ($\sim$11.7 kb) is remarkably similar to pSE101, and was therefore designated pSE102. Our analysis also identified AICEs in the genomes of *Frankia alni* strain ACN14a (AICE*Fraal*5456; 14.3 kb), and *Frankia* sp. strain EAN1pec (AICE*Franean*5323 (19.5 kb) and AICE*Franean*6303 (20.1 kb)), in the obligate marine species *Salinispora arenicola* (AICE*Sare*1922 (14.4 kb) and AICE-*Sare*1562 (13.3 kb)) and *Salinispora tropica* (AICE*Strop*0058; 14.9 kb), in *Streptomyces avermitilis* (AICE*Sav*3708 (24.3 kb) and AICE*Sav*3728 (22.5 kb)) and in *Mycobacterium gilvum* (AICE*Mflv*3036; 17.9 kb) (Fig. 1b). Interestingly, the two AICEs of *S. avermitilis* are located in tandem, with AICE*Sav*3708 directly following AICE*Sav*3728. This is most likely the result of tandem integration of AICE*Sav*3728 and AICE*Sav*3708 into the "same" Arg-tRNA gene. We propose AICE*Sav*3728 integrated first into the Arg-tRNA gene, restoring the gene after integration, in which AICE*Sav*3708 subsequently inserted, leaving a 44 bp *att* site between the two elements.

Moreover, the genome of *Streptomyces coelicolor* A3(2) contains six pSAM2-like insertions, three of which are integrated into a tRNA gene (Bentley et al. 2002; this study). Analysis of these three insertions showed that two of them are putative functional AICEs, AICE*Sco*3250 (14.3 kb) and AICE*Sco*5349 (21.2 kb). The third insertion appears to be a remnant of

an AICE (*Sco*3997, 11.3 kb, Fig. 1c), encoding a few typical AICE proteins namely integrase, excisionase and Kor proteins, but lacking the N-terminal half of the RepSA coding sequence. Another pSAM2-like insertion (*Sco*3937, 10.4 kb, Fig. 1c) is integrated into a thiamine biosynthesis gene *thiC*, instead of a tRNA gene. This insertion encodes a serine integrase instead of a tyrosine integrase typical for AICEs (see below) that is related (26% identity) to a putative integrase of *Mycobacterium tuberculosis* bacteriophage ΦRv1, but also encodes homologues of several pSAM2 proteins. An AICE remnant is also present on the *Sal. arenicola* genome (*Sare*1208, Fig. 1c). Its size could not be determined, since clear attachment sites were missing.

Modular organisation of the AICEs

An eye-catching feature of the AICEs is their highly conserved structural organisation, with functional modules for excision/integration, replication, conjugative transfer, and regulation (Fig. 1). These modules are described in more detail below.

Site-specific integration

The AICE integration system resembles that of several temperate bacteriophages (Boccard et al. 1989b). AICEs integrate into a specific tRNA gene

◄**Fig. 1** Structural organisation of (**a**) previously characterised and sequenced AICEs: pMEA100 of *A. mediterranei*; pMEA300 of *A. methanolica*; pSE211 of *Sac. erythraea* NRRL23338; pSAM2 of *S. ambofaciens*; pMR2 of *M. rosaria*; SLP1 of *S. coelicolor* A3(2); pSE101 of *Sac. erythraea* NRRL23338 (**b**) newly found AICEs: pSE101 and pSE222 of *Sac. erythraea* NRRL23338; AICE*Sav*3728 and AICE*Sav*3708 of *S. avermitilis* MA-4680; AICE*Sco*3250 and AICE*Sco*5349 of *S. coelicolor* A3(2); AICE*Mflv*3036 of *Myc. gilvum* PYR-GCK; AICE*Franean*5323 and AICE*Franean*6303 of *Frankia* sp. strain EAN1pec; AICE*Fraal*5456 of *F. alni* ACN14a; AICE*Sare*1562 and AICE*Sare*1922 of *Sal. arenicola* CNS205; AICE*Strop*0058 of *Sal. tropica* CNB-440. (**c**) AICE-remnants: *Sare*1208 of *Sal. arenicola* CNS205; *Sco*3997 of *S. coelicolor* A3(2); insertion *Sco*3937 of *S. coelicolor* A3(2). All newly found elements were named after the locus tag of their integrase gene. The prefixes of locus tags of the AICEs of *Sac. erythraea* NRRL23338 (SACE), *S. coelicolor* A3(2) (SCO), *S. avermitilis* MA-4680 (SAV), *Frankia* sp. strain EAN1pec (Franean), *F. alni* ACN14a (FRAAL), *Sal. arenicola* CNS205 (Sare), *Sal. tropica* CNB-440 (Strop), and *Myc. gilvum* PYR-GCK (Mflv) were left out for clarity. The size of the elements and the tRNA gene in which the elements are inserted are indicated below the element name at the right. Colour coding: orange, genes and sites involved in excision/integration; dark yellow, genes most likely involved in replication and its control; dark yellow with vertical black lines, *repAM* genes; dark yellow with vertical red lines, *repSA* genes; red bar, pMEA-specific hairpin structure; blue, putative conjugation genes; dark blue, putative main transfer genes; lime, putative regulatory genes; dark green, Nudix hydrolase genes; white, *orfs* with unknown functions; arrows with diagonal black lines, *orfs* with G + C content <55%; pink, transposons (*tn*), IS-elements (IS), and pseudogenes (*ps*); lavender with white diagonal lines, genes encoding DNA primase/polymerase (Prim-pol) proteins; red, genes with annotated function: *gltA*, glycosyltransferase; *hnh*, HNH-endonuclease signature; *aph*, aminoglycoside phosphotransferase; *flav*, flavoprotein; *re*, predicted restriction endonuclease; *phd*, encoding prevent-host-death family protein; *thio*, thioesterase superfamily protein; *exp*, predicted RND superfamily drug exporter; *pept*, peptidase C14 caspase catalytic subunit p20; HAD, HAD-superfamily hydrolase subfamily IA; *ich*, isochorismatase hydrolase; *dprA*; encoding DNA-protecting protein; *rni*, ribonuclease inhibitor; *thiC*, thiamine biosynthetic gene. The following colours correspond to genes shared by two or more elements, lavender, pMEA100, pSE211, SLP1, AICE*Sco*5349, AICE*Sare*1922, AICE*Strop*0058, *Sare*1208; grey, pSE222, AICE*Sare*1562, AICE*Sare*1922, AICE*Strop*0058; bright yellow (GGDEF-domain), pSE211, pSAM2, AICE*Sco*5349, AICE*Sare*1562; light blue, pSE101, pSE102, pSE222; bright green (single-stranded binding protein), pMEA100 and pSE222; plum, pSE211, pSE222; black, AICE*Sare*1562, AICE*Sare*1922, *Sare*1208; arrows with black confetti; highly similar protein clusters on AICE*Sco*3250 and *Sco*3997. Grey band between pSE101 and pSE102 indicates a highly similar DNA region. Black arrow head indicates partial protein coding sequence and back diagonal bar indicates a frameshift. The shared *att* site located between AICE*Sco*3708 and AICE*Sco*3728 is highlighted in both elements. Figure 1 was taken from te Poele et al. (2008)

in the host chromosome by site-specific recombination, which is mediated by a tyrosine integrase and occurs between two identical short sequences, the *att* identity segments, in the attachment site on the element (*attP*) and on the chromosome (*attB*). The *attB* overlaps the 3′ end of the tRNA, keeping the gene functional after integration and excision. In most cases, an integrase and excisionase are both required for excision whereas the integrase alone can mediate integration of the element. The *Streptomyces lividans* SLP1 encoded integrase carries out both integration and excision (Brasch and Cohen 1993).

Some elements can integrate site-specifically in tRNA genes of genomes other than their own host. Integration of pSAM2 was observed in many other *Streptomyces* species (Kuhstoss et al. 1989; Boccard et al. 1989a) and vectors based on the pSAM2 integration system were shown to integrate in a Pro-tRNA gene in *Mycobacterium smegmatis* (Martin et al. 1991; Seoane et al. 1997). The pSE101 element integrates at multiple chromosomal locations in *S. lividans*, whereas it is integrated site-specifically in a Thr-tRNA of its original host *Sac. erythraea* (Brown et al. 1988a, b).

Replication

AICEs can replicate like genuine plasmids, but replication is not required for maintenance. The elements are maintained integrated in the host chromosome, in which they are propagated along with the host during cell division. Mutations in the replication gene or origin of replication (*ori*) of pSAM2 lead to loss of the transfer function, showing that autonomous replication is required for conjugal transfer (Smokvina et al. 1991; Hagege et al. 1994).

Replication of plasmids is initiated by a chromosomal or plasmid-encoded replication initiator protein (Rep) that binds to a specific DNA sequence at the *ori* and guides the assembly of the replication initiation complex. This complex largely consists of proteins of the host replication machinery and in the subsequent steps of elongation and termination of replication the plasmid often relies extensively on host proteins as well.

After binding to the *ori*, replication proceeds by one of two basic mechanisms. The first mechanism involves opening of the parental strands at an AT-rich region followed by RNA-priming and is used in theta

replication and in the related strand displacement. The second mechanism, which is used by rolling circle replication (RCR) plasmids, introduces a nick to relax the DNA and to generate a 3′ OH that is used as a primer for initiation of replication.

The *ori* of theta-replicating plasmids typically consist of several short DNA repeats, called iterons, to which Rep specially binds. Opening of the double-stranded plasmid DNA occurs at an AT-rich region adjacent to these iterons. Unwinding of the strands is performed by a helicase and an RNA primer is synthesised by either RNA polymerase or by a host- or plasmid-encoded primase. DNA synthesis is performed by DNA polymerase III holoenzyme, and is continuous on the leading strand and discontinuous on the lagging strand (del Solar et al. 1998). For more details on the mechanism of replication and proteins involved therein see the review of del Solar et al. (1998).

The *ori* of RCR plasmids is called the double strand origin (DSO) and consists of a binding region and a nicking site. The DSO regions of many RCR plasmids have secondary structures, such as hairpins and cruciforms. Also binding of Rep may cause bending of the DNA and the generation of a hairpin, in which the nicking site is located in the single-stranded loop (Khan 2005). This nicking site is highly conserved among plasmids belonging to the same family, whereas the binding site is less well conserved. Hence the replication specificity of RCR plasmids is determined by the specific binding of Rep to the DSO (Khan 2005). Rep nicks the DNA and becomes covalently attached to the 5′ phosphate end, leaving a free 3′ OH end which is used as primer for leading strand synthesis. Rep recruits other proteins, such as a DNA helicase which unwinds the DNA, single-stranded DNA-binding proteins (SSBs) which coat the displaced ssDNA, and DNA polymerase III which initiates leading strand replication (Khan 2005). After the leading strand has been fully displaced, Rep introduces a second single-stranded break at the DSO and ligates the single-stranded DNA ends, generating a double-stranded plasmid and a circular single-stranded plasmid. Generation of ssDNA replicative intermediates is the hallmark for RCR plasmids (te Riele et al. 1986). The single-stranded plasmid is converted into dsDNA using the single strand origin (SSO) and the host replication machinery.

pSAM2 of *Streptomyces ambofaciens*, the only AICE for which the replication mechanism has been elucidated thus far, replicates via the RCR mechanism (Hagege et al. 1993a) and homologues of its replication initiator protein RepSA were present on newly found AICEs of *Streptomyces*, *Mycobacterium*, *Salinispora*, and *Frankia* (Fig. 2). MGEs replicating by RCR are divided into five families: pT181, pC194, pMV158 and pSN2 (del Solar et al. 1998), and a family based on *Corynebacterium* replicons (Osborn et al. 2000). Based on the presence of conserved motifs and a conserved nicking site, pSAM2 belongs to the pC194 family together with most of the characterised actino-mycete RCR plasmids, like pSG5 of *Streptomyces ghanaensis* and pSVH1 of *Streptomyces venezuelae* (Muth et al. 1995; Reuther et al. 2006b). However, the amino acid sequence of RepSA of pSAM2 appears not to be related to those of the Rep proteins of pSG5 and pSVH1 (Fig. 2).

On SLP1, two regions are required for autonomous replication (Omer and Cohen 1989). The first region contains the transcriptional regulator *impA*, and the second region encodes the two proteins SCO4617 and SCO4618, which are highly similar to SAV3711 (89% identity) and SAV3712 (71% identity) of AICE*Sav3*708 (Figs. 1a, b, 2). On both elements, the genes encoding these proteins are directly upstream of *xis*. SCO4618 and SAV3712 are weakly related to a DNA primase/polymerase (Prim-pol) domain protein of pSE211 (SACE_7028). SCO4617 and SAV3711 may therefore encode the actual replication initiator proteins of the elements. The replication mechanism of SLP1 is still unknown. The elements pMEA100, pSE211, pSE222, AICE-*Strop*0058, AICE*Sare*1922, and an AICE remnant of *Sal. arenicola Sare*1208 all encode a Prim-pol domain protein (Fig. 1). These proteins show similarity to the N-terminal Prim-Pol domain of a 955 aa long phage/plasmid primase P4-like protein of *Sphingomonas wittichii* RW1. This protein has an additional C-terminal helicase domain, which is missing from the much smaller pMEA-encoded Prim-pols. It has been suggested that the Prim-pol proteins with the associated helicases could form a replication initiation complex (Lipps 2004). Whether AICE-encoded Prim-pol proteins are involved in replication is not known.

pMEA300 also requires two regions for replication. Similar to SLP1, one region encodes a
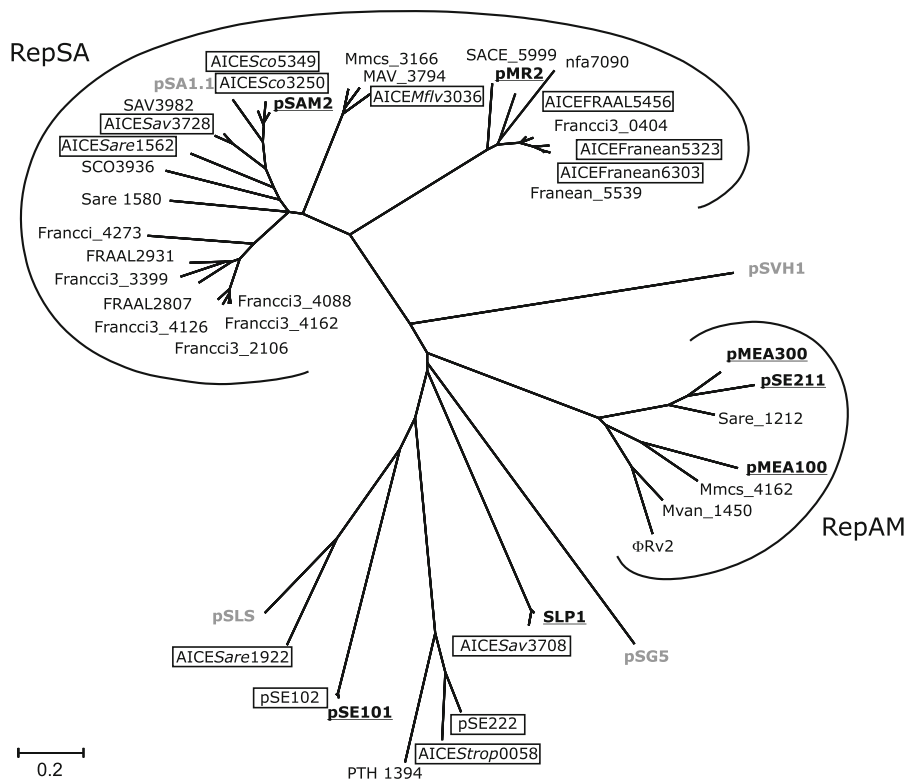
Fig. 2 Evolutionary relationship between (putative) replication proteins of previously characterised AICEs (bold, underlined), newly found AICEs (boxed), actinomycete plasmids (bold grey) and other chromosome encoded homologues with more than 35% identity to AICE-encoded Rep proteins. Abbreviations: SACE, *Sac. erythraea* NRRL23338; SCO/*Sco*, *S. coelicolor* A3(2); SAV/*Sav*, *S. avermitilis* MA-4680; Sare, *Sal. arenicola* CNS205; Strop, *Sal. tropica* CNB-440; FRAAL, *Frankia alni* strain ACN14a; Franean, *Frankia* sp. strain EAN1pec; Francci3, *Frankia* sp. strain CcI3; *Mflv*, *Myc.* *gilvum*; Mmcs, *Mycobacterium* sp. MCS; ΦRv2, prophage of *Mycobacterium tuberculosis* H37Rv; Mvan, *Mycobacterium vanbaalenii* PYR-1; MAV, *Mycobacterium avium* 104; nfa, *Nocardia farcinica* IFM 10152; PTH, *Pelotomaculum thermopropionicum* SI. Plasmids: pSLS of *Streptomyces laurentii*; pSG5 of *S. ghanaensis*; pSVH1 of *Streptomyces venezuelae*; pSA1.1, *Streptomyces cyaneus*. The phylogenetic tree was constructed using the neighbour-joining algorithm of Mega version 4.0 (Tamura et al. 2007). The scale bar represents 0.2 substitutions per site

transcriptional regulator, KorA, related to ImpA, and the second region contains two genes, *orfA* and *repAM*, which are transcriptionally coupled (Vrijbloed et al. 1995a). The *repAM* gene encodes the replication protein of pMEA300. Since *orfA* and *repAM* are transcriptionally coupled a role for OrfA in replication is anticipated. However, its disruption did not significantly affect replication, therefore its function remains unknown (Vrijbloed et al. 1995a; te Poele et al. 2006). Closely related homologues of RepAM can be found on pMEA100 and pSE211 (Fig. 2). We have shown that RepAM of pMEA300 and its homologues on pMEA100 and pSE211 belong to a novel class of replication initiator proteins (te Poele et al. 2006). The RepAM proteins do not

resemble any previously known replication proteins, but a highly similar protein Sare_1212 with ∼59% identity to the RepAM of pMEA300 (Fig. 2) is encoded on the AICE remnant *Sare*1208 (Fig. 1c). The RepAM proteins are also related to hypothetical proteins of *Mycobacterium* sp. MCS and *Mycobacterium vanbaalenii* PYR-1, and to a possible ΦRv2 prophage protein of *M. tuberculosis* (Fig. 2). RepAM has unique DNA-binding properties. Purified pMEA300 RepAM protein binds specifically to multiple identical 8 bp repeats within its own coding sequence (te Poele et al. 2006). The repeat sequences within this putative *ori* can form a stable secondary hairpin structure. Similar hairpin structures with multiple identical 8 bp repeats are also present on

Fig. 3 (**a**) Alignment of the
putative nicking sites and
the flanking 8 bp repeats
(underlined) of pMEA300,
pMEA100, pSE211,
pSE101, pSE102, and the
AICE remnant *Sare*1208
(**b**) Secondary structures of
the hairpins of pMEA300,
pMEA100, pSE211,
pSE101, pSE102, and the
AICE remnant *Sare*1208, as
predicted by mfold (Zuker
2003). The conserved 8 bp
repeats are outlined by bars.
The black arrows indicate
the putative nicking sites,
which are located in ssDNA
regions. Black star indicates
1 bp mismatch in 8 bp
repeat of *Sare*1208.
Figure 3 was adapted from
te Poele et al. (2006)
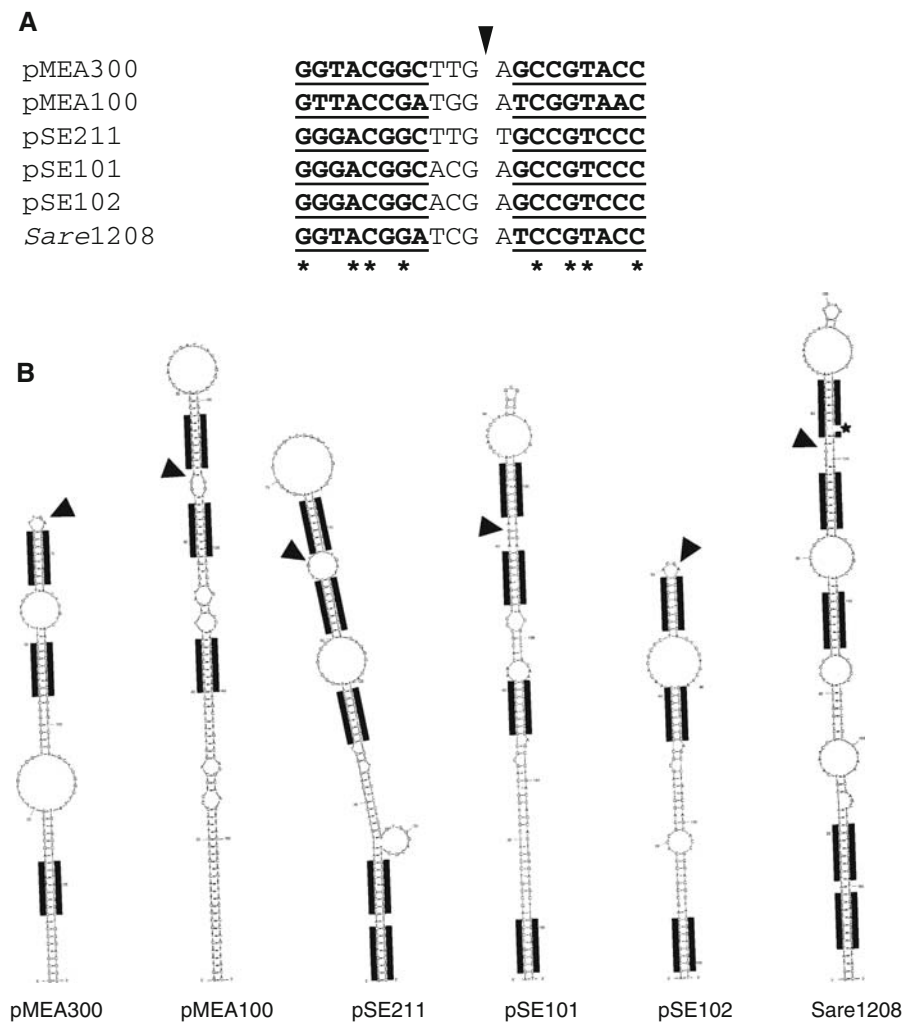
pMEA100, pSE211 and on the AICE remnant
*Sare*1208 of *Sal. arenicola*, and on pSE101 and
pSE102 (Fig. 3b). However, since the replication
proteins of pSE101 and pSE102 are clearly distinct
from the pMEA-encoded RepAM proteins, it remains
unclear whether they also use the conserved hairpin
structure as *ori* or whether they use an alternative *ori*.
Besides functioning as *ori*, the unique location of the
RepAM binding site also hints to a regulatory role for
the expression of RepAM, or of Xis that is located
directly downstream of RepAM in case of the pMEA-
elements.

All six hairpin structures contain a consensus
sequence that is similar to the nicking site (5′-
CTTGAT-3′) of the pC194 family of RCR plasmids
(Fig. 3a). Binding of RCR Rep proteins to the

double stranded origin of replication (DSO) may
cause bending of the DNA and/or the generation of
a hairpin structure in which the nicking site is
located in the single-stranded loop (Khan 2000).
Similarly, the putative nicking sites of the AICEs
are also located within unpaired regions of the
hairpin structures, and are directly flanked by the 8
bp inverted repeats (Fig. 3a), although one mismatch
can be observed in one of the 8 bp repeats of the
hairpin structure of AICE remnant *Sare*1208
(Fig. 3b).

The presence of a hairpin structure and pC194-like
nicking site in the putative DSO, suggests that the
RepAM-based elements replicate via the RCR mech-
anism. However, the presence of multiple identical
RepAM binding sites and the location of the DSO at

the 3′-end of *repAM* are atypical for RCR plasmids. Furthermore, the sizes of pMEA300 (13.3 kb), pSE211 (17.2 kb) and pMEA100 (23.3 kb) are believed to be too large for the RCR mechanism due to structural instability of large ssDNA intermediates (Helinski et al. 1996). Interestingly, the largest element pMEA100 (23.3 kb) encodes an SSB protein. Host-encoded SSB proteins coat ssDNA intermediates during RCR (Khan 2005), protecting them against nuclease attack and formation of undesired secondary structures (Greipel et al. 1987). The pMEA100-encoded SSB may prevent a shortage of host SSB proteins in protecting its large ssDNA intermediates during autonomous replication. The other, smaller, ICEs do not encode SSB proteins, suggesting that their host SSB pool is sufficient for stabilising the ssDNA replication intermediates of these elements.

Besides the presence of a hairpin structure and the pC194-like nicking site in the RepAM binding region, RepAM and its homologues on pMEA100 and pSE11 are not related to known RCR proteins. They lack consensus sequences typical for these RCR proteins, such as motifs of the catalytic domain, or of a putative metal-binding domain (del Solar et al. 1998; Ilyina et al. 1992). The elements are also devoid of Rep proteins and structural features like iterons as found in theta-replicating plasmids (del Solar et al. 1998).

Conjugation

The conjugation process of AICEs appears to be similar to that of *Streptomyces* plasmids, but differs greatly from that of other bacteria. Only one protein is essential for intermycelial transfer from donor to recipient. Streptomycetes grow as mycelia and it has been suggested that the hyphal tips of a plasmid-carrying donor and a recipient can grow together when their mycelia intertwine, making special aggregation systems for cell-to-cell contact, as seen in Gram-negative and unicellular Gram-positive bacteria, unnecessary (Grohmann et al. 2003; Wu et al. 1995). The transfer protein TraB of pSG5 was found to be localised at the hyphal tips of *S. lividans*, indicating that conjugation takes place at the tips of the mating mycelium (Reuther et al. 2006a). The main transfer proteins of AICEs and *Streptomyces*

plasmids all contain a FtsK-SpoIIIE domain that is also present in proteins involved in chromosome partitioning during sporulation and cell division, translocating double-stranded chromosomal DNA (Begg et al. 1995; Wu et al. 1995). Mating experiments with pSAM2 indicated that it was transferred as dsDNA (Possoz et al. 2001). In fact, the TraB proteins of plasmids pSG5 and pSVH1 were shown to be DNA translocators mediating the translocation of unprocessed dsDNA molecules to recipient strains (Reuther et al. 2006a).

A few small and often hydrophobic spread proteins (Spd) are involved in spread of the element through the septal crosswalls of the compartments of the recipient mycelium. This intramycelial spread is accompanied by pock formation, a phenotype that appears to be restricted to *Streptomyces* plasmids and AICEs (Moretti et al. 1985; Vrijbloed et al. 1995c; Smokvina et al. 1991; Brown et al. 1988a, b). Pocks are macroscopically visible inhibition zones that reflect temporary growth delay of plasmid-acquiring recipient cells when these are grown in a confluent lawn of plasmid-free recipients (Bibb et al. 1977). The proposed function is to delay growth until the copy number of the element is sufficiently high for efficient spread in the recipient mycelium (Hagege et al. 1993b; Grohmann et al. 2003).

Regulation

AICEs have a prophage-like mode of maintenance as they are mainly integrated in the host chromosome. The copy-number of the freely replicating form is normally low, in case of pMEA300 less than 1 freely replicating copy per 5–10 chromosomes (Vrijbloed 1996). For pSAM2 it has been observed that only under conditions favourable for conjugation, the element is excised and freely replicating molecules appear prior to transfer to recipients (Possoz et al. 2001). Once transferred, the element replicates in the recipient, spreads in its mycelium and integrates into its chromosome.

Studies on pSAM2 showed that two element-encoded proteins, the transcriptional regulator KorSA and the hydrolase Pif, keep pSAM2 integrated in the absence of potential recipient cells (Sezonov et al. 2000; Possoz et al. 2003). KorSA belongs to the GntR-family of repressor proteins and has homologues on all

the other AICEs. KorSA represses *pra*, the activator of the *rep*, *xis*, *int* operon (Sezonov et al. 1998, 2000). Initiation of transfer is established by temporary inactivation of KorSA. This leads to derepression of Pra and subsequent replication and conjugal transfer.

Pif, the pSAM2 immunity factor, contains a Nudix hydrolase motif that is required for the immunity activity (Possoz et al. 2003). Nudix proteins hydrolyse the pyrophosphate bond in a Nucleoside diphosphate linked to some other moiety, X. The Nudix protein MutT of *E. coli* degrades potentially mutagenic nucleotides; other Nudix proteins control the levels of metabolic intermediates and signalling compounds (McLennan 2006). Pif of pSAM2 is the first Nudix protein shown to be involved in bacterial conjugation. It prevents redundant transfer between pSAM2 harbouring cells by rendering its host unable to induce DNA transfer into neighbouring donor cells. It has been suggested that Pif prevents recognition between donor cells by modifying a host component in donor cells (Possoz et al. 2003). However, the mechanism triggering conjugal transfer is unknown.

When *A. methanolica* is grown in medium containing autoclaved sucrose or fructose, a drastic increase in freely replicating pMEA300 molecules is observed (Vrijbloed et al. 1995a). Conceivably, a specific degradation product of the sugars somehow triggers the onset of conjugation thus leading to excision and autonomous replication of pMEA300. pMEA300 and several other AICEs also encode Nudix hydrolase proteins, but it is not known whether these are involved in immunity as well (Fig. 1).

SLP1 is one of the elements (7 out of 20) lacking a Nudix hydrolase homologue. For this element another mechanism for preventing redundant exchange was proposed, in which the *imp* (inhibition of plasmid maintenance) locus works as a master regulator of the replication, integration and transfer functions (Shiffman and Cohen 1993; Hagege et al. 1999). It encodes two proteins, ImpA and ImpC that both contain a helix-turn-helix (HTH)-motif, suggesting that they function as DNA-binding proteins. ImpA is related to the Kor transcriptional regulators found on other AICEs. Cell-to-cell contact between donor and recipient may cause derepression of the *imp* controlled genes resulting in replication and conjugation. Derepression may be caused by dilution of the Imp proteins or by signal transduction

mechanisms altering *imp* expression or activity (Hagege et al. 1999).

Several AICE encoded *kor* genes are also part of a *kil-kor* system. The expression of certain *kil* genes is lethal in the absence of a *kor* (*kil* override) gene that controls expression of the Kil phenotype. The *kil-kor* systems of AICEs, like the *kil-kor* systems found on conjugative and pock forming *Streptomyces* plasmids, are associated with conjugation, in which Kor transcriptionally represses transfer genes responsible for the Kil phenotype. The proposed function is to retard growth of recipient cells, resulting in pock formation.

On pMEA300, *traA* and *traB* are most likely involved in the Kil phenotype, since disruption of the two genes strongly affected pock size (Vrijbloed et al. 1995c). KorA binds to the *korA-traA* intergenic region, and may therefore regulate its own expression and that of the transfer genes (Vrijbloed et al. 1995c). The binding region contains a 14 bp inverted repeat, that is also present directly upstream of *orfA*, one of the genes involved in replication. Furthermore, *korA* cannot be deleted from autonomous replicating derivatives of pMEA300, suggesting a role for KorA in replication of pMEA300 as well (Vrijbloed et al. 1995a).

ImpA, the Kor-like protein encoded by the *imp* locus of SLP1, also regulates its own expression by binding to a 16 bp inverted repeat (Shiffman and Cohen 1993) that is highly similar to the 14 bp repeat found on pMEA300, and regulates transfer by binding to a second promoter upstream of the transfer genes (Hagege et al. 1999). However, in contrast to the situation on pMEA300, the *imp* locus represses replication of SLP1 (Shiffman and Cohen 1993).

KorSA of the pSAM2 encoded *kil-kor* system does not bind to the transfer genes. Instead, it negatively regulates transfer indirectly by repressing *pra*, which encodes the activator of replication and transfer, to maintain pSAM2 integrated in the chromosome (Sezonov et al. 2000).

Deletion of the pMEA300 encoded *stf* gene was shown to result in reduced transformation frequencies with pMEA300 DNA in *A. methanolica* (Vrijbloed et al. 1995b). The Stf protein contains a partial C-terminal DUF921 domain that is found in several putative regulatory proteins in *Streptomyces*, one of which also has a putative N-terminal HTH-motif and is thought to be involved in sporulation regulation (Babcock and Kendrick 1990). Whether Stf is the activator protein of

replication and transfer of pMEA300, similar as the function of Pra on pSAM2, is not known. A highly similar protein is present on pSE101 (SACE_0474; Fig. 1a) showing 70% identity to Stf.

Interestingly, on the pMEA-elements and the related elements pSE101, pSE102, and pSE222, a small ORF (59–69 aa) was found directly upstream and convergently transcribed from Kor- or, in case of pSE222, a XRE-family transcriptional regulator. TMHMM analysis of its amino acid sequence predicts a signal peptide suggesting that this protein is either secreted or associated with the cell membrane. No homologues of these ORFS were found on the other AICEs or in databases, suggesting that they are pMEA-specific. Its function is currently unknown.

pMEA-elements show a clear geographical distribution

Screening of a collection of more than 100 *Amycolatopsis* strains that had been isolated from different geographical locations around the world (Tan et al. 2006) showed that pMEA-specific *repAM* and *traJ* sequences are more widely distributed among *Amycolatopsis* and revealed a very interesting geographical distribution of pMEA-like elements (te Poele et al. 2007).

Some strains only had a *repAM* sequence and several strains only had a *traJ* sequence, suggesting that these genes are not necessarily linked. It is likely that *traJ* has become associated with a different type of replication initiation protein than *repAM*, and vice versa, emphasising the mosaic structure of mobile genetic elements (MGEs) (Osborn et al. 2000) that is also found in AICEs.

Two geographically distinct populations of pMEA-like elements were identified. Phylogenetic analysis of their deduced RepAM and TraJ protein sequences revealed clustering with the protein sequences of either pMEA300 or pMEA100 (Fig. 4). The sequences clustering with pMEA300 consisted of Australasian strains whereas the pMEA100 cluster originated from European strains. These *repAM* and *traJ* genes were phylogenetically clearly linked with the 16S rRNA gene of the investigated host strains (Fig. 4). Apparently, the pMEA-elements mainly coevolved with their host in an integrated form, rather than being dispersed via horizontal gene transfer (HGT). So far, *repAM* and

traJ sequences were not detected in the few *Amycolatopsis* strains isolated from other locations, notably Egypt (Henssen et al. 1987) and the USA (Lee and Hah 2001; Labeda 1995; Mertz and Yao 1993; Stapley et al. 1972; Lechevalier et al. 1986). Clearly, more *Amycolatopsis* strains from various regions around the world need to be isolated and screened for pMEA-like sequences to draw firm conclusions about the geographical distribution and spread of pMEA-elements, or to assess whether their diversity is restricted to the two described populations.

The origin and evolution of AICEs

Several proteins are encoded on all AICEs, such as an integrase and GntR- or XRE-family transcriptional regulator, as well as an attachment site *attP*, directly downstream of the integrase gene. Most AICEs also encode an excisionase, a protein involved in conjugal transfer, and a Nudix hydrolase. One or two proteins involved in replication are encoded by all AICEs, although most of the replication proteins are not related to each other. The novel replication initiator protein RepAM is only encoded by pMEA300, pMEA100, pSE211, and on an AICE remnant of *Sal. arenicola*. The other three elements of *Sac. erythraea*, pSE101, pSE102 and pSE222, lack a *repAM*-related gene and apparently encode other replication genes instead. In that sense, they are not genuine pMEA-elements, but are clearly pMEA-related, with a similar overall organisation, and several other common genes and structural features. The small *orf* directly upstream and convergently transcribed from *korA* (an XRE-family regulator gene in case of pSE222) is a feature that all the pMEA-related elements have in common. No homologues of this small protein were found on other AICEs or in the databases, suggesting that these proteins are specific for the pMEA-elements and therefore restricted to the closely related *Amycolatopsis* and *Saccharopolyspora* strains in which these elements reside.

Some genes are present only on one or a few AICEs and mostly encode hypothetical proteins with unknown function. Interestingly, pSE102, one of the newly found elements, encodes a putative aminoglycoside phosphotransferase protein (APH). Such proteins are known to confer resistance to various aminoglycoside antibiotics. The pSE102-encoded

Fig. 4 Comparison of the phylogenetic relationship, as shown by connecting lines, of the *Amycolatopsis* 16S rDNA sequences (~1,250 bps) with that of the RepAM and TraJ sequences reveals co-evolution of the pMEA-elements with their *Amycolatopsis* host strains. Phylogenetic trees were reconstructed by neighbour-joining with Mega version 4.0 using CLUSTALW alignment and by calculating evolutionary distances by the Kimura-2 parameter method. Bootstrap values were calculated from 1,000 replicate trees. Bootstrap values over 50% are shown. The scale bar represents 0.1, 0.01, and 0.05 substitutions per nucleotide position for RepAM, 16S rDNA and TraJ, respectively. Figure 4 was taken from te Poele et al. (2007)
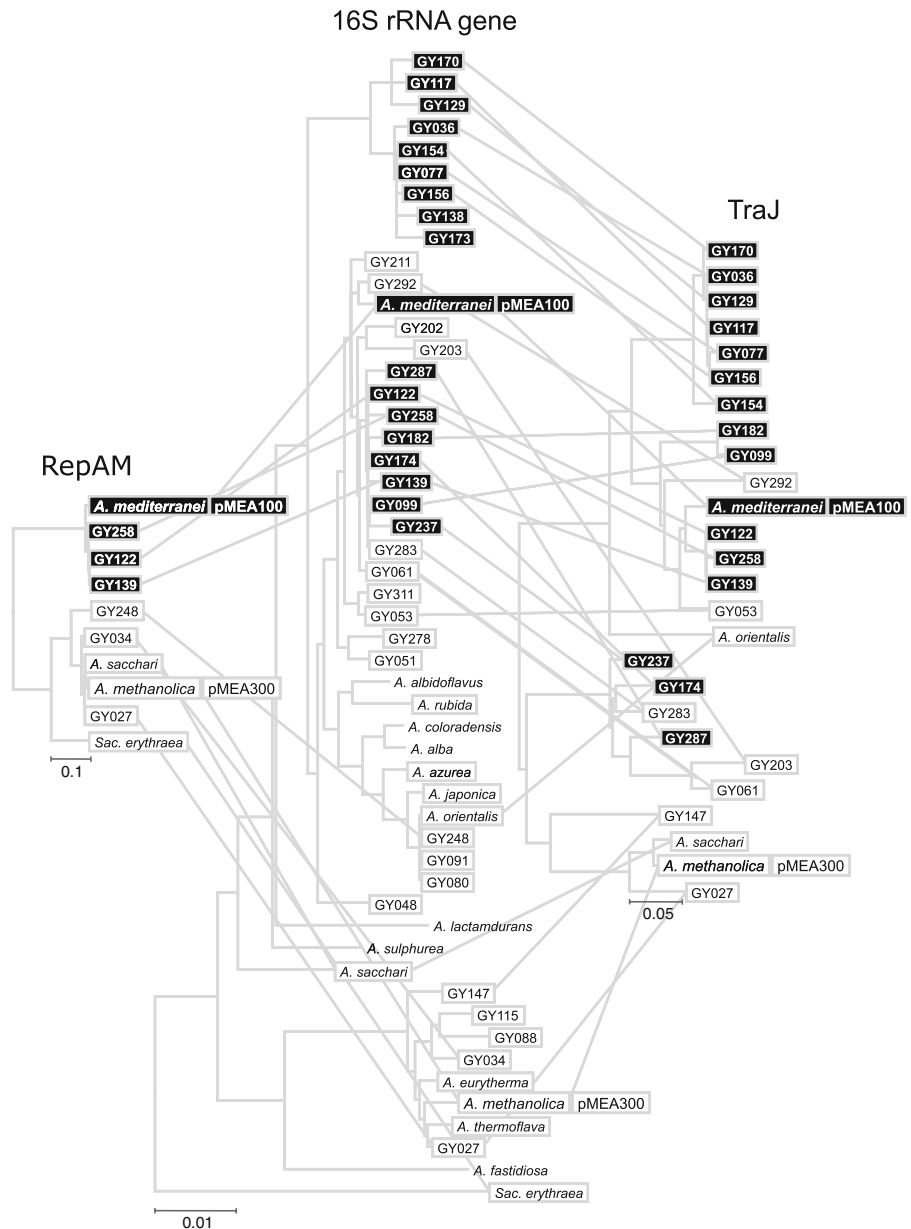
APH may therefore give the host a predictable beneficial phenotype, i.e. antibiotic resistance.

AICE replication and excision/integration modules most likely originate from bacteriophages. RepAM of pMEA100, pSE211 and pMEA300 are 34–41% identical to a ΦRv2 prophage protein (Rv2655c) of *Myc. tuberculosis*. Similar to the pMEA-elements, the gene encoding Rv2655c can be found directly upstream of an excisionase and an integrase. ΦRv2 also integrates in one of the tRNA genes (Bibb and Hatfull 2002). RepSA of pSAM2 resembles the Rep

proteins of some bacteriophages that replicate by the RCR mechanism (Hagege et al. 1994). The excision and integration system of AICEs is very similar to that of temperate bacteriophages (Boccard et al. 1989b). The regulatory and conjugation modules of the pMEA-like elements most likely have a bacterial/plasmid origin: homologues can be found in many actinomycetes and other bacterial species and do not appear to be phage-related.

Despite the highly conserved modular structure of these AICEs, within each module a large number of

rearrangements can be observed with genes conserved in function but of different phylogenetic descent, especially those involved in autonomous replication. The elements either have a different evolutionary origin and converged to functionally similar elements, or have a common ancestor and diverged later in time.

Within this very complex organisation several highly similar genes and gene clusters were identified on a number of elements, such as the replication genes of pMEA300, pSE211 and pMEA100 (Fig. 2), and the excisionases and integrases of pMEA100, pSE211, pSE101 and pSE102 (data not shown). Furthermore, pMEA100 and pSE211 share a highly similar DNA region (62% identity) that consists of three genes, encoding a metal-dependent phosphohydrolase (Mdp) (65% identity), a Nudix hydrolase (Nud) (75% identity) and a HTH-XRE motif containing regulatory protein (58% identity). Highly similar gene clusters are also present in AICE*Sco*5349 of *S. coelicolor*, AICE*Franean*5323 of *Frankia* EAN1pec, and AICE*Mflv*3036 of *Myc. gilvum* PYR-GCK (Fig. 1b), in several actinomycete genomes, such as *Nocardia farcinica* IFM 10152 and *Frankia* sp. ClI3, and in a number of *Mycobacterium* genomes (data not shown). Further analysis of the regions flanking these *Mycobacterium* genomic gene clusters showed that they are often part of a larger pSAM2-related gene cluster consisting of *repSA*, *traSA*, *pra*, *mdp*, *nud*, and *xre* homologues, and could therefore be remnants of AICEs. The high similarity between the gene clusters on several actinomycete genomes and AICE genes indicates that these genes have a common origin. Conceivably, exchange of DNA has occurred between the different actinomycete genera, possibly mediated by AICEs.

AICEs are widely distributed amongst actinomycetes, but the RepAM-based pMEA-elements were only found in *Amycolatopsis* and the closely related genus *Saccharopolyspora*, except for the RepAM-based AICE remnant in *Sal. arenicola*. The RepAM homologue of *Myc. tuberculosis* is part of a prophage ΦRv2, and analysis of the genes flanking the RepAM homologues in the genomes of *Mycobacterium* sp. MCS and *Myc. vanbaalenii* suggests that they are prophage-related as well. Furthermore, the observation that the two isolated integrative elements pMEA100 and pMEA300 can only be transferred into their own (pMEA-free) host strains and cannot be exchanged between *A. methanolica* and

A. mediterranei (H. Kloosterman and E.M. te Poele, unpublished results) supports the small host range of pMEA-elements.

The host range of AICEs may depend on both their replication and integration machinery. Autonomous replication is required for conjugal transfer of AICEs, as shown for pSAM2 (Smokvina et al. 1991; Hagege et al. 1994). The excised element replicates in the donor and after conjugal transfer replicates and spreads within the recipient mycelium (Possoz et al. 2001). Integration of the element in the recipient chromosome requires the presence of a suitable integration site. Although RepAM-based pMEA-elements may be able to transfer to other species outside the host range, once inside the recipient strain these elements may not be able to establish themselves properly because of the inability to replicate and subsequently spread within the mycelium.

AICEs encoding a Rep protein related to RepSA of pSAM2 of *S. ambofaciens* on the other hand, were found to be widely distributed in actinomycetes. These newly found AICEs were, in addition to other species of *Streptomyces*, also present in several less closely related actinomycete genera from other suborders, i.e. *Mycobacterium*, *Frankia* and *Salinispora*, indicating that RepSA-based AICEs have a broader host range than the pMEA-elements and may disperse to other genera in several suborders.

The host range of AICEs is also dependent on the integration system, although it may be broader than the host range for autonomous replication (Martin et al. 1991). AICEs integrate site-specifically into tRNA genes, which are highly conserved in actinomycetes. The pSAM2 integration system for instance mediates site-specific integration in a Pro-tRNA gene in the *Myc. smegmatis* chromosome (Martin et al. 1991). Furthermore, unlike phage lambda, which requires an integration host factor (IHF) for integration (Gardner and Nash 1986), AICEs appear not to require host factors for integration (Raynal et al. 1998; Katz et al. 1991), making them less dependent on the host for integration.

AICEs as modulators of host genome diversity and their role in HGT

Despite the new insights into the distribution, origin and evolution of AICEs as described above, the eco-physiological role of AICEs is not clear.

Nevertheless, the rearrangements and apparent exchange of genes between the different AICEs suggest that these elements are continuously subject to HGT between different actinomycete genera and are likely to play an important role in genome plasticity of the host in which they reside.

A strong indication that AICEs play a role in host genome plasticity comes from the observation that several AICEs contain more genes than strictly required for maintenance and transfer of the elements. The majority of these genes are located on a highly variable region that flanks the conserved modules. For most of these additional genes it is not directly clear whether, and how, they can be beneficial to the host strain. Rather, these additional genes may be remnants of previous chromosomal reorganisation events. At this moment we cannot reconstruct the actual events leading to the rearrangements, which by themselves may be beneficial to the host.

Furthermore, the highly variable regions often have a significantly lower G + C content, and frequently encompass transposases and encode non-actinomycete proteins. For example, pSE222 contains a gene (SACE_1142a) with a considerable lower G + C content of 51% than the average 66.3%, and its protein product (349 aa) shows 27 and 30% identity to conserved hypothetical proteins (ZP_01980106; ZP_01705385) of the γ-proteobacteria *Vibrio cholerae* MZO-2 and *Shewanella putrefaciens* 200. The gene SCO5329 (52% G + C) of AICE*Sco*5349 (65% G + C) is 28% identical to a hypothetical protein (NP_463839) of the food-borne pathogen *Listeria monocytogenes* EGD-e. This indicates that AICEs may also mediate genetic exchange outside of the actinomycetes. However, database searches fail to identify clear AICE homologues in non-actinomycetes. Alternatively, the low G + C DNA may also have been acquired from an actinophage with a low G + C content.

Additional indications for the involvement of AICEs in genome rearrangements and HGT events come from the analysis of several related actinomycete genomes. The *S. coelicolor* genome contains a number of potentially horizontally acquired insertions (Bentley et al. 2002). Several of the insertions present in the central core region are absent from the highly syntenic core regions of *S. ambofaciens* and *S. lividans* genomes, showing that these regions have a recent origin in *S. coelicolor* (Choulet et al. 2006; Jayapal et al. 2007). The elements AICE*Sco*3250 and

AICE*Sco*5349 represent two such horizontally acquired insertions in the *S. coelicolor* genome. Interestingly, AICE*Sco*3250 directly flanks the calcium-dependent antibiotic (CDA) biosynthetic gene cluster and AICE*Sco*5349 is adjacent to the *whiE* cluster (Bentley et al. 2002) encoding the biosynthetic enzymes for the grey polyketide spore pigment (Davis and Chater 1990). These two AICEs are absent in *S. lividans* (Jayapal et al. 2007) and the *S. ambofaciens* genome lacks the CDA cluster and also the complete region containing the *whiE* cluster and AICE*Sco*5349 (Choulet et al. 2006). Possibly, these AICEs were associated with the acquisition of these secondary metabolite clusters in *S. coelicolor*. The suggestion that the *whiE* cluster has a recent origin in *S. coelicolor* is substantiated by findings of Metsa-Ketela et al. (2002). They observed that the *whiE* sequences of several *Streptomyces* species did not share the same phylogeny as the 16S rRNA genes, which is suggestive of horizontal transfer.

Moreover, the *S. coelicolor* variable arm region contains a large 153 kb genomic island (SCO6806-SCO6953) consisting of multiple insertions. This fitness island contains IS-elements and numerous genes involved in various modes of microbial defence, such as several polyketide synthase genes, two DNA methylase genes, genes similar to several multi-drug efflux transporters and arsenic resistance genes (Bentley et al. 2002; Jayapal et al. 2007). It is flanked by a 43 bp direct repeat one of which is located within the 3′ end of the Pro-tRNA. The tRNA gene is directly downstream of an integrase (SCO6806) that is highly similar (69% identity) to the integrase of pSE222. Conceivably, an AICE was involved in the acquisition of this island and may have been subject to multiple additional insertions.

Actinomycetes are the most important bacterial producers of bioactive secondary metabolites and notably antibiotics. In addition to genes involved in antibiotic biosynthesis, these strains often carry resistance genes such as APHs, to protect themselves against the antibiotics they produce. Interestingly, pSE102 encodes a putative APH. Related proteins (31–37% identity) are also encoded by the large 153 kb genomic island of *S. coelicolor* (SCO6951) and by the RepAM-based AICE remnant of *Sal. arenicola* (Sare_1225). The observation that putative antibiotic resistance genes are associated with AICEs is one of the first examples of the possible role of AICEs in

spread of antibiotic resistance via HGT, at least between actinomycetes.

It is clear from the above that our knowledge about AICEs is still far from complete. Novel elements are being identified continually, especially thanks to the large number of genome sequences that are now being resolved. Particularly the isolation of a large number of strains of one species, or several closely related species, and analysis of their integrative elements, will be of great value to understand the processes that have lead to the current diversity in integrative elements, and the effects that these elements may have on host genome plasticity and therefore on evolution of actinomycetes. We have initiated such a comparative analysis with the *Amycolatopsis* pMEA-elements (Tan et al. 2006; te Poele et al. 2007, 2008).

The continued development and discovery of novel antibiotics is essential, especially in light of the alarming recurrence of pathogenic micro-organisms that are resistant to a broad range of unrelated antibiotics (Wright 2007). Therefore, particularly the rich diversity of actinomycetes needs to be further explored. In order to do so, there will also be a great need for genetic tools and cloning vectors that allow us to manipulate and control expression of genes involved in secondary metabolite production, and for construction of overproducing strains. AICEs offer an impressive range of elements for the development of such genetic tools. Further detailed information about these integrative and conjugative elements, and their physiological and evolutionary roles, is urgently needed.

References

Babcock MJ, Kendrick KE (1990) Transcriptional and translational features of a sporulation gene of *Streptomyces griseus*. Gene 95:57–63

Begg KJ, Dewar SJ, Donachie WD (1995) A new *Escherichia coli* cell division gene, *ftsK*. J Bacteriol 177:6211–6222

Bentley SD, Chater KF, Cerdeno-Tarraga AM et al (2002) Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature 417:141–147

Bibb LA, Hatfull GF (2002) Integration and excision of the *Mycobacterium tuberculosis* prophage-like element, phiRv1. Mol Microbiol 45:1515–1526

Bibb MJ, Freeman RF, Hopwood DA (1977) Physical and genetic characterization of a second sex factor, SCP2, for *Streptomyces coelicolor* A3(2). Mol Gen Genet 154:155–166

Bibb MJ, Ward JM, Kieser T, Cohen SN, Hopwood DA (1981) Excision of chromosomal DNA sequences from *Streptomyces coelicolor* forms a novel family of plasmids detectable in *Streptomyces lividans*. Mol Gen Genet 184:230–240

Boccard F, Smokvina T, Pernodet JL, Friedmann A, Guerineau M (1989a) Structural analysis of loci involved in pSAM2 site-specific integration in *Streptomyces*. Plasmid 21:59–70

Boccard F, Smokvina T, Pernodet JL, Friedmann A, Guerineau M (1989b) The integrated conjugative plasmid pSAM2 of *Streptomyces ambofaciens* is related to temperate bacteriophages. EMBO J 8:973–980

Brasch MA, Cohen SN (1993) Excisive recombination of the SLP1 element in *Streptomyces lividans* is mediated by Int and enhanced by Xis. J Bacteriol 175:3075–3082

Brown DP, Chiang SJ, Tuan JS, Katz L (1988a) Site-specific integration in *Saccharopolyspora erythraea* and multisite integration in *Streptomyces lividans* of actinomycete plasmid pSE101. J Bacteriol 170:2287–2295

Brown DP, Tuan JS, Boris KA, DeWitt JP, Idler KB, Chiang SJ, Katz L (1988b) Plasmid-chromosome interactions in *Saccharopolyspora erythraea* and *Streptomyces lividans*. Dev Ind Microbiol 29:97–105

Burrus V, Waldor MK (2004) Shaping bacterial genomes with integrative and conjugative elements. Res Microbiol 155:376–386

Burrus V, Pavlovic G, Decaris B, Guedon G (2002) Conjugative transposons: the tip of the iceberg. Mol Microbiol 46:601–610

Choulet F, Aigle B, Gallois A et al (2006) Evolution of the terminal regions of the *Streptomyces* linear chromosome. Mol Biol Evol 23:2361–2369

Cohen A, Bar-Nir D, Goedeke ME, Parag Y (1985) The integrated and free states of *Streptomyces griseus* plasmid pSG1. Plasmid 13:41–50

Davis NK, Chater KF (1990) Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Mol Microbiol 4:1679–1691

del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R (1998) Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev 62:434–464

Gardner JF, Nash HA (1986) Role of *Escherichia coli* IHF protein in lambda site-specific recombination. A mutational analysis of binding sites. J Mol Biol 191:181–189

Greipel J, Maass G, Mayer F (1987) Complexes of the single-stranded DNA-binding protein from *Escherichia coli* (Eco SSB) with poly(dT). An investigation of their structure and internal dynamics by means of electron microscopy and NMR. Biophys Chem 26:149–161

Grohmann E, Muth G, Espinosa M (2003) Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev 67:277–301

Hagege J, Pernodet JL, Friedmann A, Guerineau M (1993a) Mode and origin of replication of pSAM2, a conjugative

integrating element of *Streptomyces ambofaciens*. Mol Microbiol 10:799–812

Hagege J, Pernodet JL, Sezonov G, Gerbaud C, Friedmann A, Guerineau M (1993b) Transfer functions of the conjugative integrating element pSAM2 from *Streptomyces ambofaciens*: characterization of a *kil-kor* system associated with transfer. J Bacteriol 175:5529–5538

Hagege J, Boccard F, Smokvina T, Pernodet JL, Friedmann A, Guerineau M (1994) Identification of a gene encoding the replication initiator protein of the *Streptomyces* integrating element, pSAM2. Plasmid 31:166–183

Hagege JM, Brasch MA, Cohen SN (1999) Regulation of transfer functions by the *imp* locus of the *Streptomyces coelicolor* plasmidogenic element SLP1. J Bacteriol 181:5976–5983

Helinski DR, Toukdarian AE, Novick RP (1996) Replication control and other stable maintenance mechanisms of plasmids. In: Neidhardt FC, Curtiss RIII, Ingraham JL, Lin ECC, Low KB, Magasanik B, Reznikoff WS, Riley M, Schaechter M, Umbarger HE (eds) *Escherichia coli* and *Salmonella*: cellular and molecular biology, 2nd edn. ASM Press, Washington, DC, pp 2295–2324

Henssen A, Kothe HW, Kroppenstedt RM (1987) Transfer of *Pseudonocardia azurea* and *Pseudonocardia fastidiosa* to the genus *Amycolatopsis*, with emended species description. Int J Syst Bacteriol 37:292–295

Hopwood DA, Hintermann G, Kieser T, Wright HM (1984) Integrated DNA sequences in three streptomycetes form related autonomous plasmids after transfer to *Streptomyces lividans*. Plasmid 11:1–16

Hosted TJ Jr, Wang T, Horan AC (2005) Characterization of the *Micromonospora rosaria* pMR2 plasmid and development of a high G + C codon optimized integrase for site-specific integration. Plasmid 54:249–258

Ilyina TV, Gorbalenya AE, Koonin EV (1992) Organization and evolution of bacterial and bacteriophage primase-helicase systems. J Mol Evol 34:351–357

Jayapal KP, Lian W, Glod F, Sherman DH, Hu WS (2007) Comparative genomic hybridizations reveal absence of large *Streptomyces coelicolor* genomic islands in *Streptomyces lividans*. BMC Genomics 8:229

Katz L, Brown DP, Donadio S (1991) Site-specific recombination in *Escherichia coli* between the att sites of plasmid pSE211 from *Saccharopolyspora erythraea*. Mol Gen Genet 227:155–159

Khan SA (2000) Plasmid rolling-circle replication: recent developments. Mol Microbiol 37:477–484

Khan SA (2005) Plasmid rolling-circle replication: highlights of two decades of research. Plasmid 53:126–136

Kuhstoss S, Richardson MA, Rao RN (1989) Site-specific integration in *Streptomyces ambofaciens*: localization of integration functions in *S. ambofaciens* plasmid pSAM2. J Bacteriol 171:16–23

Labeda DP (1995) *Amycolatopsis coloradensis* sp. nov., the avoparcin (LL-AV290)-producing strain. Int J Syst Bacteriol 45:124–127

Lechevalier MP, Prauser H, Labeda DP, Ruan JS (1986) Two new genera of nocardioform actinomycetes: *Amycolata* gen. nov. and *Amycolatopsis* gen. nov. Int J Syst Bacteriol 36:29–37

Lee SD, Hah YC (2001) *Amycolatopsis albidoflavus* sp. nov. Int J Syst Evol Microbiol 51:645–650

Lipps G (2004) The replication protein of the *Sulfolobus islandicus* plasmid pRN1. Biochem Soc Trans 32:240–244

Martin C, Mazodier P, Mediola MV, Gicquel B, Smokvina T, Thompson CJ, Davies J (1991) Site-specific integration of the *Streptomyces* plasmid pSAM2 in *Mycobacterium smegmatis*. Mol Microbiol 5:2499–2502

McLennan AG (2006) The Nudix hydrolase superfamily. Cell Mol Life Sci 63:123–143

Mertz FP, Yao RC (1993) *Amycolatopsis alba* sp. nov., isolated from soil. Int J Syst Bacteriol 43:715–720

Metsa-Ketela M, Halo L, Munukka E, Hakala J, Mantsala P, Ylihonko K (2002) Molecular evolution of aromatic polyketides and comparative sequence analysis of polyketide ketosynthase and 16S ribosomal DNA genes from various *Streptomyces* species. Appl Environ Microbiol 68:4472–4479

Moretti P, Hintermann G, Hutter R (1985) Isolation and characterization of an extrachromosomal element from *Nocardia mediterranei*. Plasmid 14:126–133

Mullany P, Roberts AP, Wang H (2002) Mechanism of integration and excision in conjugative transposons. Cell Mol Life Sci 59:2017–2022

Muth G, Farr M, Hartmann V, Wohlleben W (1995) *Streptomyces ghanaensis* plasmid pSG5: nucleotide sequence analysis of the self-transmissible minimal replicon and characterization of the replication mode. Plasmid 33:113–126

Omer CA, Cohen SN (1989) SLP1: a paradigm for plasmids that site-specifically integrate in the actinomycetes. In: Berg DE, Howe MM (eds) Mobile DNA. American Society for Microbiology, Washington, DC, pp 289–296

Osborn M, Bron S, Firth N, Holsappel S, Huddleston A, Kiewiet R, Meijer W, Seegers J, Skurray R, Terpstra P, Thomas CM, Thorsted P, Tietze E, Turner SL (2000) The evolution of bacterial plasmids. In: Thomas CM (ed) The horizontal gene pool: bacterial plasmids and gene spread. Harwood Academic Publishers, Amsterdam, The Netherlands, pp 301–361

Pernodet JL, Simonet JM, Guerineau M (1984) Plasmids in different strains of *Streptomyces ambofaciens*: free and integrated form of plasmid pSAM2. Mol Gen Genet 198:35–41

Possoz C, Ribard C, Gagnat J, Pernodet JL, Guerineau M (2001) The integrative element pSAM2 from *Streptomyces*: kinetics and mode of conjugal transfer. Mol Microbiol 42:159–166

Possoz C, Gagnat J, Sezonov G, Guerineau M, Pernodet JL (2003) Conjugal immunity of *Streptomyces* strains carrying the integrative element pSAM2 is due to the *pif* gene (pSAM2 immunity factor). Mol Microbiol 47:1385–1393

Raynal A, Tuphile K, Gerbaud C, Luther T, Guerineau M, Pernodet JL (1998) Structure of the chromosomal insertion site for pSAM2: functional analysis in *Escherichia coli*. Mol Microbiol 28:333–342

Reuther J, Gekeler C, Tiffert Y, Wohlleben W, Muth G (2006a) Unique conjugation mechanism in mycelial streptomycetes: a DNA-binding ATPase translocates unprocessed plasmid DNA at the hyphal tip. Mol Microbiol 61:436–446

Reuther J, Wohlleben W, Muth G (2006b) Modular architecture of the conjugative plasmid pSVH1 from *Streptomyces venezuelae*. Plasmid 55:201–209

Salyers AA, Shoemaker NB, Stevens AM, Li LY (1995) Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiol Rev 59: 579–590

Seoane A, Navas J, Garcia Lobo JM (1997) Targets for pSAM2 integrase-mediated site-specific integration in the *Mycobacterium smegmatis* chromosome. Microbiology 143:3375–3380

Sezonov G, Duchene AM, Friedmann A, Guerineau M, Pernodet JL (1998) Replicase, excisionase, and integrase genes of the *Streptomyces* element pSAM2 constitute an operon positively regulated by the *pra* gene. J Bacteriol 180:3056–3061

Sezonov G, Possoz C, Friedmann A, Pernodet JL, Guerineau M (2000) KorSA from the *Streptomyces* integrative element pSAM2 is a central transcriptional repressor: target genes and binding sites. J Bacteriol 182:1243–1250

Shiffman D, Cohen SN (1993) Role of the *imp* operon of the *Streptomyces coelicolor* genetic element SLP1: two *imp*-encoded proteins interact to autoregulate *imp* expression and control plasmid maintenance. J Bacteriol 175: 6767–6774

Smokvina T, Francou F, Luzzati M (1988) Genetic analysis in *Streptomyces ambofaciens*. J Gen Microbiol 134:395–402

Smokvina T, Boccard F, Pernodet J-L, Friedmann A, Guerineau M (1991) Functional analysis of the *Streptomyces ambofaciens* element pSAM2. Plasmid 25:40–52

Stapley EO, Jackson M, Hernandez S, Zimmerman SB, Currie SA, Mochales S, Mata JM, Woodruff HB, Hendlin D (1972) Cephamycins, a new family of beta-lactam antibiotics. I. Production by actinomycetes, including *Streptomyces lactamdurans* sp. nov. Antimicrob Agents Chemother 2:122–131

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) Software Version 4.0. Mol Biol Evol 24:1596–1599

Tan GYA, Ward AC, Goodfellow M (2006) Exploration of *Amycolatopsis* diversity in soil using genus-specific primers and novel selective media. Syst Appl Microbiol 29:557–569

te Poele EM, Kloosterman H, Hessels GI, Bolhuis H, Dijkhuizen L (2006) RepAM of the *Amycolatopsis methanolica* integrative element pMEA300 belongs to a novel class of replication initiator proteins. Microbiology 152: 2943–2950

te Poele EM, Habets MN, Tan GY, Ward AC, Goodfellow M, Bolhuis H, Dijkhuizen L (2007) Prevalence and distribution of nucleotide sequences typical for pMEA-like accessory genetic elements in the genus *Amycolatopsis*. FEMS Microbiol Ecol 61:285–294

te Poele EM, Samborskyy M, Oliynyk M, Leadlay PF, Bolhuis H, Dijkhuizen L (2008) Actinomycete integrative and conjugative pMEA-like elements of *Amycolatopsis* and *Saccharopolyspora* decoded. Plasmid 59:202–216

te Riele H, Michel B, Ehrlich SD (1986) Are single-stranded circles intermediates in plasmid DNA replication? EMBO J 5:631–637

Vrijbloed JW (1996) Functional analysis of the integrative plasmid pMEA300 of the actinomycete *Amycolatopsis methanolica*. PhD Thesis, University of Groningen

Vrijbloed JW, Madon J, Dijkhuizen L (1994) A plasmid from the methylotrophic actinomycete *Amycolatopsis methanolica* capable of site-specific integration. J Bacteriol 176:7087–7090

Vrijbloed JW, Jelinkova M, Hessels GI, Dijkhuizen L (1995a) Identification of the minimal replicon of plasmid pMEA300 of the methylotrophic actinomycete *Amycolatopsis methanolica*. Mol Microbiol 18:21–31

Vrijbloed JW, Madon J, Dijkhuizen L (1995b) Transformation of the methylotrophic actinomycete *Amycolatopsis methanolica* with plasmid DNA: stimulatory effect of a pMEA300-encoded gene. Plasmid 34:96–104

Vrijbloed JW, van der Put NM, Dijkhuizen L (1995c) Identification and functional analysis of the transfer region of plasmid pMEA300 of the methylotrophic actinomycete *Amycolatopsis methanolica*. J Bacteriol 177:6499–6505

Waldor MK, Tschape H, Mekalanos JJ (1996) A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. J Bacteriol 178:4157–4165

Wright GD (2007) The antibiotic resistome: the nexus of chemical and genetic diversity. Nat Rev Microbiol 5: 175–186

Wu LJ, Lewis PJ, Allmansberger R, Hauser PM, Errington J (1995) A conjugation-like mechanism for prespore chromosome partitioning during sporulation in *Bacillus subtilis*. Genes Dev 9:1316–1326

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31: 3406–3415