

Review

Not peer-reviewed version

Functional Roles and Genomic Impact of Miniature Inverted-Repeat Transposable Elements (MITEs) in Prokaryotes

[Michael F Minnick](#)*

Posted Date: 14 February 2024

doi: 10.20944/preprints202402.0807.v1

Keywords: prokaryote; genomes; miniature inverted repeat transposable elements; MITEs; evolution; transposons; Tns; non-autonomous; mobile genetic elements; MGEs



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

Functional Roles and Genomic Impact of Miniature Inverted-Repeat Transposable Elements (MITEs) in Prokaryotes

Michael F. Minnick

Program in Cellular, Molecular, and Microbial Biology. Division of Biological Sciences, University of Montana, Missoula, Montana, USA; mike.minnick@mso.umt.edu

Abstract: Prokaryotic genomes are dynamic tapestries that are strongly influenced by mobile genetic elements (MGEs), including transposons (Tn's), plasmids, and bacteriophages. Of these, miniature inverted-repeat transposable elements (MITEs) are undoubtedly the least studied MGEs in bacteria and archaea. This review explores the diversity and distribution of MITEs in prokaryotes and describes what is known about their functional roles in the host and involvement in genomic plasticity and evolution.

Keywords: prokaryote; genomes; miniature inverted repeat transposable elements; MITEs; evolution; transposons; Tns; non-autonomous; mobile genetic elements; MGEs

Introduction

The discovery of “jumping genes” (transposable genetic elements or transposons) in maize by Barbara McClintock in the 1940's was monumental for its contribution to our understanding of the dynamic nature of the genome. From McClintock's work, it was clearly apparent that phenotypic traits were not dictated solely by Mendelian genetics, and that gene linkage was not fixed within a chromosome. Rather, phenotypes could be changed stochastically, gene linkages rearranged, and gene expression altered, as a result of insertion or loss of transposable elements (TEs) [1,2]. Although the tremendous significance of McClintock's research was not recognized for over thirty years, it eventually earned her a Nobel Prize in Physiology or Medicine in 1983. Today, the role of TEs in genomic plasticity and evolution is universally accepted dogma. Other mobile genetic elements have since been discovered, and together with TEs constitute a “mobilome” that is continually in flux and manipulating the genomic makeup of organisms in all three domains of life.

The basic structure of a TE includes a set of terminal inverted repeats (TIRs) that surround a transposase (*tnp*) gene whose encoded enzyme mobilizes the element. The simplest TE is the insertion sequence (IS), and it consists solely of this basic structure. As such, IS elements are relatively small (~800-2000 bp) with TIRs of ~10-50 bp (Figure 1A). If the TE includes the basic structure plus accessory genes (e.g., antibiotic-resistance markers, virulence genes, etc.), it is referred to as a simple TE. Such elements can reach several thousand base pairs in length, depending on their gene content (Figure 1B). Finally, when two IS elements flank a set of accessory genes, a composite TE is formed (Figure 1C). Here, the transposase is encoded by one or both of the flanking IS elements, rather than within the intervening core sequence, as observed in IS elements and simple TEs.

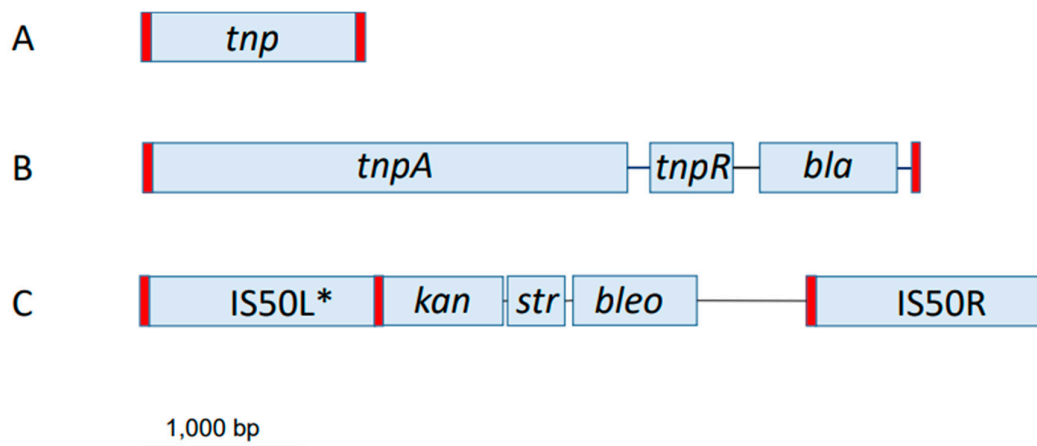


Figure 1. Examples of the three types of TEs. A) Insertion sequence IS2 from *Escherichia coli*. The IS2 *tnp* gene consists of two ORFs (*orfA* and *orfB*) that encode the active transposase, which is produced following a -1 translational frameshift and fusion. **B) Simple transposon Tn3 from *E. coli*** possessing *tnpA* (transposase), *tnpR* (resolvase), and *bla* (β -lactamase) genes. **C) Composite transposon Tn5** containing two flanking IS50 elements plus genes conferring resistance to kanamycin (*kan*), streptomycin (*str*), and bleomycin (*bleo*). *IS50L of Tn5 is defective for transposition due to a nonsense mutation in its transposase gene. Red rectangles indicate positions of TIRs. .

Transposition is an autonomous process catalyzed by enzymes encoded within the respective element, and TEs can be transposed by one of two methods. TEs that use a RNA intermediate as a template and reverse transcriptase to generate double-stranded DNA for re-integration into the genome undergo a “copy and paste” method of mobilization and are referred to as Class I TEs or retrotransposons. Nearly all known retrotransposons occur in eukaryotes. In contrast, TEs that employ a “cut and paste” method with a DNA intermediate are referred to as Class II TEs or DNA transposons. DNA transposons occur in both prokaryotes and eukaryotes, and are the type relevant to this review.

Mobilization of DNA transposons involves transposase-mediated: **1)** excision of the element at its TIRs, **2)** cutting of the target DNA sequence to yield sticky-end overhangs, and **3)** ligation of the TE into the target site. Following integration, host DNA polymerases are recruited to repair any gaps that surround the new insert. Target-site specificity of the mobilization process varies with the kind of TE. For instance, Tn7 targets a single specific site in the bacterial chromosome at *attTn7* [3], whereas many other TEs randomly target host DNA.

Two general types of transposition events are known. If a Tn is completely excised and mobilized to a different location in the genome, it is referred to as conservative transposition, since the process ensures maintenance of a single copy of the element in the cell. This kind of transposition is observed *in vitro* with Tn5 (Figure 1C) [4]. Alternatively, if the TE is replicated and the resulting copy mobilized, it is referred to as replicative transposition. The net effect of this mechanism is to maintain a copy of the element in its original locus plus at least one other integration site in the chromosome. Replicative transposition is observed with Tn3 (Figure 1B) [5].

While the ability to mobilize in a genome is a hallmark characteristic of TEs, not all elements can do so autonomously. For example, if mutations generate a defective transposase gene, the resulting element can only be mobilized *in trans* through transposase enzymes supplied by similar, or distantly-related, intact transposons. Such mutated versions of TEs are referred to as miniature inverted-repeat transposable elements (MITEs), in reference to their relatively smaller size (50-500 bp) due to sequence decay of the non-functional transposase and/or other core sequences. MITEs have been investigated to a much greater degree in eukaryotes than prokaryotes, especially in plants. Adaptive roles ascribed to MITEs of eukaryotes include serving as a source of microRNAs for gene

regulation and adaptation [6,7], modulation of gene expression in nearby genes [8], influencing alternative RNA splicing events (exonization) [9], and driving genome complexity and allelic diversity through transposition [10].

1. Prokaryotic MITEs

1.1. Conserved Structural Features

Prokaryotic MITEs share the same basic architecture as eukaryotic MITEs, with a total length typically < 500 bp, TIRs of ~4-30 bp, and a flanking set of direct repeats (DRs) in relatively “younger” MITEs, following duplication of the target sequence [11,12]. Of note is that maintenance and conservation of the DRs wanes over time [11]. MITEs are typically AT-rich, tend to integrate into AT-rich intergenic regions of the genome, and produce transcripts that form thermodynamically-stable, stem-loop secondary structures [13]. A representative MITE from *Coxiella burnetii* (QMITE2) and the predicted secondary structure of its transcript are shown in Figure 2 [14].

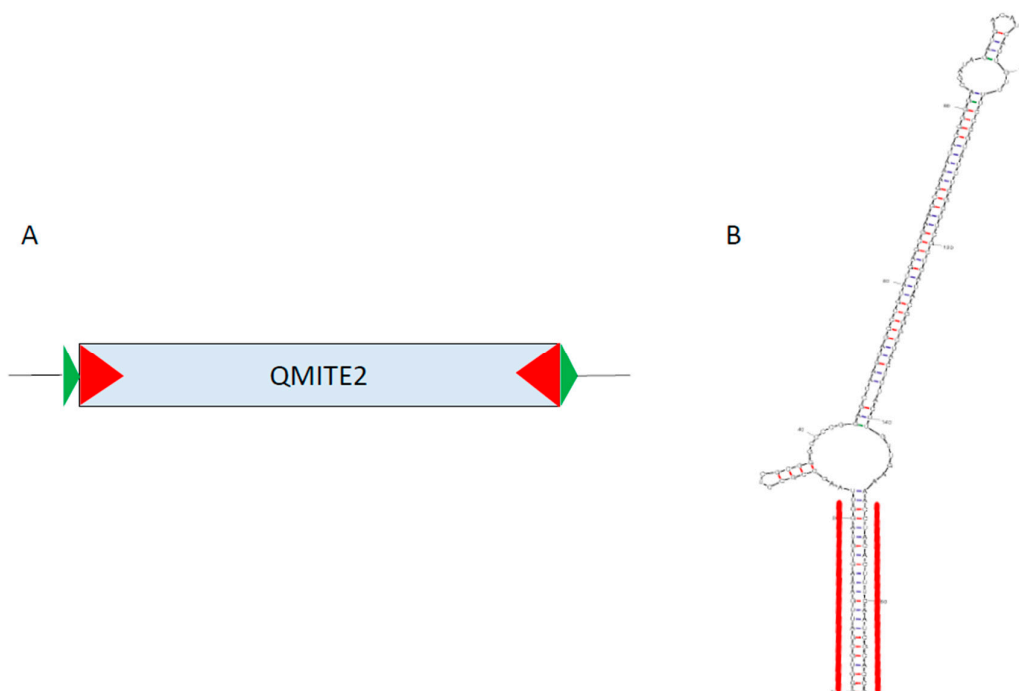


Figure 2. A representative prokaryotic MITE. QMITE2 of *C. burnetii* is depicted (adapted from [14]). **A)** Model of QMITE2 including its DRs (7-9 bp, green), TIRs (23-29 bp, red), and a core sequence (113-145 bp, blue) is shown. **B)** Secondary structure prediction of the full-length QMITE2 transcript by mFold [15] showing TIRs (red lines) and a ΔG value = -113.09 cal/mol.

IS elements are thought to be the main source of MITEs in prokaryotes. This is not surprising since IS's are the most common type of autonomous TE in these organisms, and they can catalyze mobilization of related MITEs *in trans* [16–18].

1.2. Size and Sequence Variation

There are two general categories of MITEs. Type I MITEs are those that share considerable sequence conservation with related, autonomous IS elements, while type II MITEs share sequence conservation that is strictly confined to their TIR regions [19]. In actuality, these designations are somewhat arbitrary, as a range of degeneracy from autonomous, recently-transposed IS elements to related, but “older”, type II TEs could co-exist at any point in time in an organism. Such a gradient is

clearly observed in the *C. burnetii* QMITE1 and QMITE2 families [14]. For instance, in the Nine-Mile type strain of *C. burnetii* (RSA493), it was found that the chromosome possessed 45 QMITE1 elements ranging in size from 39-400 bp. Roughly half of these MITEs were 39-321 bp long and shared < 75% sequence identity with the full-length (~400-bp) QMITE1, although their TIRs and DRs showed a relatively higher degree of conservation. A second family, QMITE2 (Figure 2), demonstrated even greater size and sequence variation. Of the 78 copies of QMITE2, only ten were “full-length” (~190 bp) MITEs. Of the remaining elements, 12 copies showed a loss of ~40 bases at their 5' ends. Notably, the loss of DRs and TIRs in these markedly truncated copies produced elements that more closely resembled repetitive extragenic palindromes (REPs) than typical MITEs [14].

The *Chungie* MITE of *Geobacter uraniireducens* also demonstrates considerable diversity with respect to size and sequence. For example, full-length *Chungie*-1 is a 235-bp element that disrupts one copy of the bacterium's two phosphate uptake operons. 37 additional copies of the element, with conserved TIRs and DRs but ranging in size from 178-235 bp, are located elsewhere in intergenic regions of the genome. Remarkably, BLAST searches revealed yet another 57 copies of partial *Chungie* elements, including presumably inactive elements lacking TIRs or DRs [11].

2. Distribution Across Prokaryotic Genomes

2.1. Taxonomic Distribution

The first prokaryotic MITE was discovered in *Neisseria gonorrhoeae* and *Neisseria meningitidis* in 1988. Named after a discoverer, the 204-bp Correia element was found in both species with an estimated 20 copies each [20]. This number was later revised upwards to >120 copies of the MITE based on the *N. gonorrhoeae* genome sequence [21]. Following the discovery of Correia in *Neisseria*, MITEs have been found in numerous other bacteria, including members of the *Enterobacteriaceae* family [22,23], cyanobacteria [12,24,25], *Staphylococcus aureus* [26], *Acinetobacter* [27,28], pseudomonads [29,30], *Porphyromonas* [31,32], *Polaramonas* [33], and an endosymbiont of spiders [34]. In fact, MITEs appear to be a universal feature of bacteria. Although relatively less is known about the MITEs of archaea, they have been identified in *Sulfolobus* spp. [35–37], *Halobacterium* [38]; *Haloquadratum walsbyi* [12], *Delsulfurococcus kamchatkensis* [39], and several other archaeal genomes [40], to date. Interestingly, the archaeon *Methanobacterium thermoautotrophicum* is unusual for its lack of any apparent MITEs [38].

2.2. Copy-Number Variation

While all MITEs tend to be multicopy in nature, the actual number in a genome varies considerably per given host and MITE family. This is perhaps best exemplified in the MITEs of plants, where an astounding 8-10% of the genome can be comprised of elements representing dozens of MITE families infecting a single plant species, and arising from single or numerous rounds of multiplication [9,41]. While nothing to this extreme has been observed in prokaryotes, MITE copy numbers can still be high and variable. For example, the MITE_{Aba12} and its orthologs vary from 1-22 elements per genome, depending on the host *Acinetobacter* species and strain examined [28]. In the archaeon *Sulfolobus solfataricus* P2, four MITE families (designated SM1-SM4) exist and range in copy numbers from 25 (SM2) to 44 (SM3) [42]. Finally, the copy-number of the enterobacterial repetitive intergenic consensus (ERIC) MITE found in the *Enterobacteriaceae* and *Vibrionaceae* families varies considerably, with copy numbers correlating with sub-family taxa [43]. ERIC genomic estimates of 30 copies in *E. coli* (strain K-12), 150 in *Salmonella enterica* (Typhimurium LT2), and 711 copies in *Photobacterium luminescens* have been reported [22,44,45]. As an example of strain-level variation, 14 additional ERIC elements have been identified in novel intergenic regions of genomes from *E. coli* strains other than K-12 [44].

2.3. Insertion Site Preferences

MITEs of prokaryotes, like their eukaryotic counterparts, are normally located in intergenic regions of the genome. For example, the BrickBuilt MITE of *P. gingivalis* and the CIR element of *Caulobacter crescentus* were reported to be strictly intergenic [32,46]. Similarly, MITEs derived from the ISSoc2 of *Synechococcus* were located intergenically or within genes of unknown function [47], and QMITE 2 of *C. burnetii* was primarily found in the 3' untranslated region of ORFs [14]. While these locations are typical, exceptions exist. For example, even in ERIC elements, where the “I” of the acronym actually denotes “intergenic”, examples of intragenic locations have been reported [48]. Presumably, any integration event that could confer a selective advantage on the host through the functional role(s) of a particular MITE would be favored and the element retained. Moreover, MITE insertion into intergenic regions would have no chance for deleterious gene inactivation, however, it could still play a role in genomic plasticity (see section 3.2, below). Interestingly, BOX, RUP and SPRITE elements of *Streptococcus pneumoniae* were frequently found bunched together in the chromosome, with elements even inserting into one another! These “clusters” were found together with MITEs in more typical intergenic regions, including regulatory sequences and pseudogenes [49].

3. Functional Roles and Genomic Impact

3.1. Influence on Gene Expression

Early work with Correia showed that a σ^{70} -like promoter sequence on the element could serve as a transcriptional promoter for the *N. gonorrhoeae* *uvrB* gene when expressed ectopically in *E. coli* *uvrB* mutants [50]. In fact, subsequent work showed that subtypes of the Correia element encoded two strong promoters that could enhance expression of proximal genes [51]. Clearly, the predominant mapped locations of Correia immediately upstream of genes would provide an optimal arrangement for promoter function [21]. Work has also shown that the Correia element can undergo inversion, especially under conditions where the host is under antibiotic stress [52]. Considering the endogenous promoters that occur in Correia, this kind of mobilization would certainly affect expression of nearby genes and might play a role in the reversible phase variation seen in *N. gonorrhoeae* [52]. Since the earlier work with Correia, σ^{70} -like promoter sequences have also been discovered in many other kinds of prokaryotic MITEs, including *Acinetobacter* MITEAba12 [28]; *Porphyromonas* BrickBuilt [32], *Coxiella* QMITE1 [14], and *Shewanella* SonMITE_1 [53].

Another means by which MITEs can regulate expression is by encoding an integration host factor (IHF)-binding site. IHF is a nucleoid-associated protein that can bind cognate IHF sites of 30-35 bp in length and subsequently bend the underlying DNA in order to optimize a variety of biological functions, including transcription [54]. A bona fide IHF-binding site was previously discovered in a Correia element inserted immediately downstream of the promoter for the *mtrCDE* efflux pump operon of *N. gonorrhoeae*. When the IHF-binding site was deleted, transcription of the operon was upregulated *in vivo*, clearly demonstrating a regulatory role for the MITE embedded in the sequence [55].

A third way that MITEs can post-transcriptionally regulate expression, is through RNase-mediated cleavage of a transcript containing the particular element. This was demonstrated with *mtrCDE* transcripts of *N. gonorrhoeae* containing a Correia element, where RNase III cleaved the mRNA at the TIRs [55]. Similar RNase III processing has also been reported for other kinds of transcripts containing Correia elements [56,57]. Interestingly, the insertion of ERIC elements at the 3' end of *Y. enterocolitica* genes can also affect levels of the respective ERIC-containing transcripts, by unmasking a RNase E site in the 3' TIR and enhancing the mRNA's degradation [58].

A fourth way that MITEs can regulate expression is through insertional inactivation of genes. Recent work with *Acinetobacter baumannii*, a highly adaptable, multidrug-resistant bacterial pathogen, has demonstrated the active mobilization of MITEAba12 into a histone-like nucleoid structuring (*hns*) global regulator gene, in response to desiccation stress [28]. Interestingly, the *hns* locus is apparently a “hotspot” for insertional inactivation by various TEs in *A. baumannii* during stress, and loss of H-NS

function following insertion generates high-level colistin resistance, hypermobility, and other phenotypic changes [59,60].

A fifth way that MITEs could conceivably regulate expression is through their production of non-coding small regulatory RNAs (sRNAs). Much like microRNAs originating from the MITEs of plants [reviewed in [61]], sRNAs encoded by prokaryotic MITEs could possibly play a role in prolonging or decreasing a target mRNA's longevity or by regulating translation following sRNA:mRNA complex formation. Of course, these activities would depend on the physical location of the sRNA:mRNA interaction(s). Production of sRNAs by MITEs has been reported in *C. burnetii*, where two sRNAs, Cbsr3 and Cbsr13, were found to originate from QMITE1, while Cbsr16 originated from QMITE2 [14]. The biological role and mRNA targets for these three sRNAs is unknown but intriguing.

Finally, it is conceivable that MITEs could also regulate expression through formation of riboswitches or Rho-independent transcriptional terminators in the mRNAs where they occur. As mentioned above, the characteristic formation of stable secondary structures in MITE transcripts, would clearly be amenable to the formation of such regulators. The occurrence of Rho-independent terminator-like sequences in SPRITE elements and potential T-box riboswitch motifs in BOX elements of *Streptococcus pneumoniae* underscores these possibilities [18].

3.2. Role in Genome Plasticity and Evolution

Genome plasticity is fundamental to successful evolution, and one of the most basic mechanisms is the acquisition of novel genes encoded by the MGEs themselves. Moreover, as these elements are mobilized and integrated, there is an increased likelihood for evolution through gene duplication and/or merodiploidy in the host. Research on prokaryotic MITEs has shown that they often encode novel ORFs, presumably derived from progenitor IS elements or acquired through recombinational events. For example, several MITEs of enteric bacteria, including ERIC, were found to possess ORFs [48]. In many instances, the ORFs within MITEs have been found to encode proteins or peptides of unknown functions or which lacked obvious homologs in the databases. Examples include the large BOX elements of pneumococcus [49], intragenic ERIC elements [48], QMITE1 of *C. burnetii* [14], and MITEb and MITEe of *Anabaena* sp. strain PCC 7120 [62].

While it is not unusual for ORFs in intragenic MITEs to generate in-frame fusions with genomic ORFs upon their integration, *Rickettsia* palindromic elements (RPEs) appear to have taken the situation to a unusually high level of frequency. As first reported, RPEs of ~150 bp, representing ten identified RPE families, integrated 44 times in unrelated metabolic genes of *Rickettsia conorii*, the agent of Mediterranean spotted fever [63]. Each RPE was found to encode an ORF with a conserved peptide. Nearly half the RPEs, from three of the ten families identified, inserted into genomic ORFs with the ability to modify the underlying gene product through a translational fusion. A dozen homologs of the RPEs were also found in genomes of *Rickettsia felis*, *Rickettsia prowazekii*, and *Rickettsia helvetica*, and eleven of the twelve elements were similarly inserted into genomic ORFs in an in-frame manner [63,64]. To what extent these insertions have contributed to the evolution of novel or modified proteins of *Rickettsia* spp. remains unexplored, but the predicted compatibility between 3-dimensional protein conformation and function following RPE modification is intriguing [64]. Moreover, it is conceivable that a similar MITE-mediated mechanism of protein evolution has occurred in other prokaryotes.

MITEs also play an important role in genome reduction through insertional inactivation of genes that are no longer needed by the host prokaryote. Together with autonomous TEs, this mechanism of purifying selection significantly contributes to genomic plasticity. Typically, there are many more MITEs per genome relative to the corresponding, autonomous precursor elements, suggesting that MITEs may play a significantly greater role in genome plasticity [65]. Insertional inactivation of genes by MITEs has been demonstrated in a variety of prokaryotes. For example, the cyanobacterial microcystin toxin gene cluster, *mcy*, was shown to be functionally inactivated by MITEs in two naturally-occurring strains of *Anabaena* collected from the Baltic Sea [66]. Interestingly, generation of non-toxic, microcystin-negative strains of cyanobacteria commonly occurs in both laboratory and

natural settings, and it is tempting to suggest that MITEs may be responsible for this widespread phenotypic variation [66]. In another example, comparative genomics work on *N. gonorrhoeae* has demonstrated that the Correia element differentially disrupts coding sequences in a strain-dependent manner, underscoring the MITE's mobility and utility in "shaping" the pathogen's genome [21]. In *Ralstonia solanacearum* and *Pseudomonas syringae*, two notorious plant pathogens, MITEs have been shown to insert and inactivate the type III secretion system effector genes *avrA* and *hopX1* (*avrPphE*), respectively [29,67]. These insertional inactivation events are thought to modulate a particular strain's virulence and expand its host range for certain plants. Finally, in the archaeon *Sulfolobus solfataricus*, a type II MITE has been shown to insertionally inactivate the *pyrE* gene (encoding orotate phosphoribosyl transferase) in certain naturally-occurring strains, demonstrating the element's ability to mobilize *in vivo*. The biological significance of this *pyrE* inactivation in natural populations is unknown but certainly intriguing [68].

MITEs have also been shown to mobilize antimicrobial resistance genes in bacteria, and this activity would clearly be adaptive to pathogens or soil microbes that are routinely exposed to antibiotics. For example, early work with E622, a 611-bp MITE involved in *P. syringae* virulence, demonstrated that it was able to mobilize a β -lactamase (*bla*) gene [30]. More recently, the *mcr-5* colistin resistance cassette was shown to be mobilized *in trans* by the IR148 MITE of *E. coli* using a transposase provided *in trans* by *TnAs1* [23]. Finally, intriguing work has implicated MITEs in the mobilization of integrons; a common bacterial genetic element that provides for acquisition and utilization of multiple gene cassettes encoding various gene products, including virulence factors, metabolic enzymes, and antimicrobial- or disinfectant-resistance determinants [69]. For example, a class I integron encoding a *bla* gene for imipenem resistance in a clinical isolate of *A. baumannii* was found to possess identical MITEs at both of its termini [70]. The exact same MITEs have also been found flanking a Class I integron of *Acinetobacter johnsonii* isolated from the gut of a prawn, with cassettes encoding streptomycin-resistance and methionine sulfoxide reductase [71]. Finally, additional work has discovered identical MITEs flanking different class I integrons in seven *Acinetobacter* spp., and one unique integron from *Acinetobacter bereziniae* possessed *aacC1* and *bla_{VIM-2}* genes for aminoglycoside and carbapenem resistance, respectively (Figure 3) [72,73].

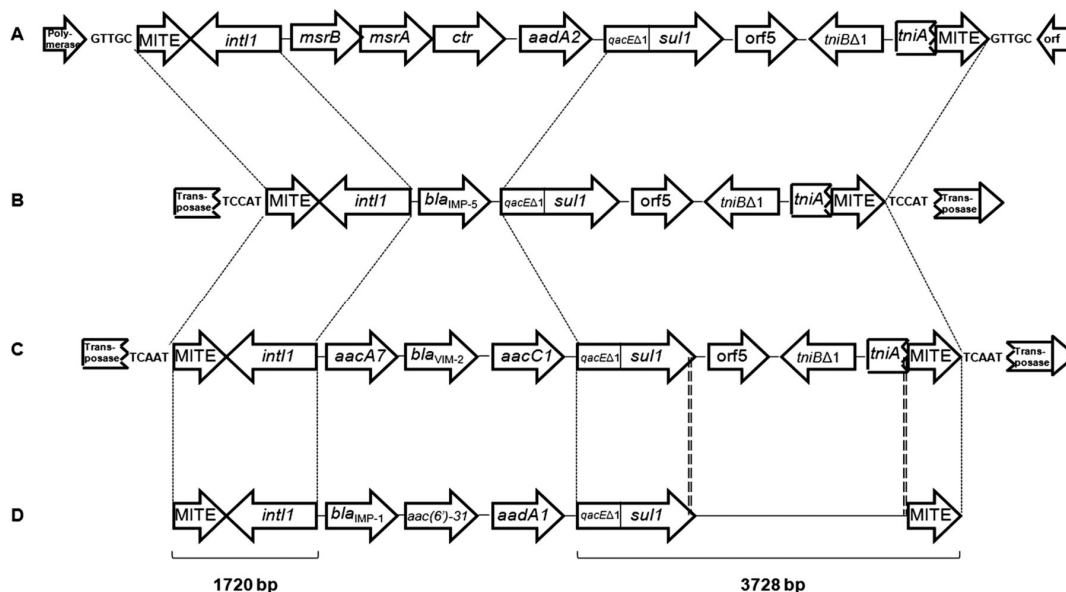


Figure 3. Depiction of various class I integrons of *Acinetobacter* spp. flanked by identical, MITEs (reprinted from [73] by permission). **A)** strain NFM2 of *A. johnsonii* isolated from prawns; **B)** strain 65FFC of *A. baumannii*; **C)** strain 118FFC of the nosocomial pathogen, *A. bereziniae*; and **D)** strains 694-696, 5227, 5248 and 9043 of *A. baumannii*. Nucleotide sequences were identical between dashed lines in A-C), whereas the region between double-dashed lines in **D)** was verified by size conformity

between PCR products compared to **A-C**) [73]. Arrows denote gene cassettes encoding identical MITEs (MITE); methionine sulfide reductase (*msrA/B*), integrase (*intI1*), transposase (*tniA/BΔ1*), and resistance genes (all others).

Taken as a whole, the existence of identical MITEs surrounding disparate integrons in various strains and species of *Acinetobacter* suggest that these elements have played a significant role in class I integron mobilization, thereby facilitating the acquisition of antimicrobial / disinfectant resistance and novel physiological traits within members of the genus. It is quite likely that MITE-mediated mobilization of integrons has also occurred in other bacteria where both MITEs and integrons are present.

Like IS elements, MITEs serve as important hotspots for homologous recombination with the chromosome during evolution. Likewise, stretches of genomic sequences flanked by homologous MITEs can be mobilized to new locations. A clear example is seen in *C. burnetii*, where QMITE2 once served as a target for transposition by IS1111, through the IS element's recognition of the MITE's GTAG repetitive extragenic palindrome [14]. Specifically, the two IS1111 TEs have integrated into the two QMITE2 copies that flank the chromosomal *icm/dot* genes for the type IV secretion apparatus [14]. Mobilization and acquisition of the *icm/dot* virulence determinant would clearly have been advantageous to *C. burnetii* over the course of its evolution to become an obligate intracellular parasite, and MITEs apparently played a role [74].

4. Bioinformatic Tools for MITE Analysis

At this writing, six open-source software programs were freely available online to identify putative MITEs in sequence datasets. From the most recent to earliest, these include MITEFinder II [75], MITE Tracker [76], MUSTv2 [77], detectMITE [78], MITE Digger [79], and MITE-Hunter [80]. Not surprisingly, these algorithms initially search for potential MITEs in genomic sequence data using their conserved structural features, including TIRs and DRs (i.e., target-site duplications; TSDs). Low-complexity TIRs are typically filtered out to decrease false-positives based on a set of selection criteria such as sequence identity and low similarity to MITEs in the Repbase database. Sequence alignments are also employed to determine sequence identity, allowing for mismatches, to determine the relatedness of the potential elements to known MITEs and to group potential MITEs into families. Major issues that have confounded software developers include the upper size-limit for the input sequence and the high rate of false-positives identified, although this problem has apparently been minimized with MITEFinder II, relative to other algorithms [75].

Conclusions and future directions

Prokaryotic MITEs are non-autonomous mobile genetic elements that shape and “flex” the genomes of host archaea and bacteria through their ability to be mobilized *in trans* by autonomous TEs. The resulting transposition event can introduce novel ORFs, insertionally inactivate genes no longer needed, and relocate sections of the genome, thereby facilitating evolution. Following integration, MITEs remain active participants in the host's biology through adaptive roles in transcriptional regulation, protein evolution, and post-transcriptional regulation by their encoded sRNAs. While several reports on the functional roles of MITEs in eukaryotic systems have been published, to date, relatively few studies exist for prokaryotic MITEs. As such, further investigation on prokaryotic MITEs should provide fertile ground for discovery.

Abbreviations

DRs (direct repeats); ERIC (enterobacterial repetitive intergenic consensus); IS (insertion sequences); MGEs (mobile genetic elements); ORF (open reading frame); REPs (repetitive extragenic palindromes); RUP (repeat units of pneumococcus); SPRITE (*Streptococcus pneumoniae* rho-independent terminator-like element); TEs (transposable elements); TIRs (terminal inverted repeats); Tn (transposon); TSDs (target site duplications); MITEs (miniature inverted-repeat transposable elements)

References

1. McClintock B. The origin and behavior of mutable loci in maize. *Proc Natl Acad Sci U S A*. 1950 Jun;36(6):344-55.
2. McClintock B. Induction of Instability at Selected Loci in Maize. *Genetics*. 1953 Nov;38(6):579-99.
3. Peters JE, Craig NL. Tn7: smarter than we thought. *Nat Rev Mol Cell Biol*. 2001 Nov;2(11):806-14.
4. Ahmed A. Alternative mechanisms for tn5 transposition. *PLoS Genet*. 2009 Aug;5(8):e1000619.
5. Nicolas E, Lambin M, Dandoy D, Galloy C, Nguyen N, Oger CA, Hallet B. The Tn3-family of Replicative Transposons. *Microbiol Spectr*. 2015 Aug;3(4).
6. Crescente JM, Zavallo D, Del Vas M, Asurmendi S, Helguera M, Fernandez E, Vanzetti LS. Genome-wide identification of MITE-derived microRNAs and their targets in bread wheat. *BMC Genomics*. 2022 Feb 22;23(1):154.
7. Guo Z, Kuang Z, Tao Y, Wang H, Wan M, Hao C, Shen F, Yang X, Li L. Miniature Inverted-repeat Transposable Elements Drive Rapid MicroRNA Diversification in Angiosperms. *Mol Biol Evol*. 2022 Nov 3;39(11):msac224.
8. Lu C, Chen J, Zhang Y, Hu Q, Su W, Kuang H. Miniature inverted-repeat transposable elements (MITEs) have been accumulated through amplification bursts and play important roles in gene expression and species diversity in *Oryza sativa*. *Mol Biol Evol*. 2012 Mar;29(3):1005-17.
9. Xin Y, Ma B, Xiang Z, He N. Amplification of miniature inverted-repeat transposable elements and the associated impact on gene regulation and alternative splicing in mulberry (*Morus notabilis*). *Mob DNA*. 2019 Jun 25;10:27.
10. Boutanaev AM, Osbourn AE. Multigenome analysis implicates miniature inverted-repeat transposable elements (MITEs) in metabolic diversification in eudicots. *Proc Natl Acad Sci U S A*. 2018 Jul 10;115(28):E6650-E6658.
11. Chen Y, Zhou F, Li G, Xu Y. A recently active miniature inverted-repeat transposable element, Chunjie, inserted into an operon without disturbing the operon structure in *Geobacter uraniireducens* Rf4. *Genetics*. 2008 Aug;179(4):2291-7.
12. Chen Y, Zhou F, Li G, Xu Y. MUST: a system for identification of miniature inverted-repeat transposable elements and applications to *Anabaena variabilis* and *Haloquadratum walsbyi*. *Gene*. 2009 May 1;436(1-2):1-7.
13. Tu Z. Three novel families of miniature inverted-repeat transposable elements are associated with genes of the yellow fever mosquito, *Aedes aegypti*. *Proc Natl Acad Sci U S A*. 1997 Jul 8;94(14):7475-80.
14. Wachter S, Raghavan R, Wachter J, Minnick MF. Identification of novel MITEs (miniature inverted-repeat transposable elements) in *Coxiella burnetii*: implications for protein and small RNA evolution. *BMC Genomics*. 2018 Apr 11;19(1):247.
15. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003 Jul 1;31(13):3406-15.
16. Siguier P, Filée J, Chandler M. Insertion sequences in prokaryotic genomes. *Curr Opin Microbiol*. 2006 Oct;9(5):526-31.
17. Siguier P, Gournayre E, Chandler M. Bacterial insertion sequences: their genomic impact and diversity. *FEMS Microbiol Rev*. 2014 Sep;38(5):865-91.
18. Delihias, N. Impact of small repeat sequences on bacterial genome evolution. *Genome Biol. Evol*. 2011. 3:959-973.
19. Oosumi T, Garlick B, Belknap WR. Identification of putative nonautonomous transposable elements associated with several transposon families in *Caenorhabditis elegans*. *J Mol Evol*. 1996 Jul;43(1):11-8.
20. Correia FF, Inouye S, Inouye M. A family of small repeated elements with some transposon-like properties in the genome of *Neisseria gonorrhoeae*. *J Biol Chem*. 1988 Sep 5;263(25):12194-8.
21. Snyder LA, Cole JA, Pallen MJ. Comparative analysis of two *Neisseria gonorrhoeae* genome sequences reveals evidence of mobilization of Correia Repeat Enclosed Elements and their role in regulation. *BMC Genomics*. 2009 Feb 9;10:70.
22. Hulton CS, Higgins CF, Sharp PM. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. *Mol Microbiol*. 1991 Apr;5(4):825-34.
23. Kieffer N, Nordmann P, Millemann Y, Poirel L. Functional Characterization of a Miniature Inverted Transposable Element at the Origin of *mcr-5* Gene Acquisition in *Escherichia coli*. *Antimicrob Agents Chemother*. 2019 Jun 24;63(7):e00559-19.
24. Zhou F, Tran T, Xu Y. Nezha, a novel active miniature inverted-repeat transposable element in cyanobacteria. *Biochem Biophys Res Commun*. 2008 Jan 25;365(4):790-4.
25. Kaneko T, Nakajima N, Okamoto S, Suzuki I, Tanabe Y, Tamaoki M, Nakamura Y, Kasai F, Watanabe A, Kawashima K, Kishida Y, Ono A, Shimizu Y, Takahashi C, Minami C, Fujishiro T, Kohara M, Katoh M, Nakazaki N, Nakayama S, Yamada M, Tabata S, Watanabe MM. Complete genomic structure of the bloom-forming toxic cyanobacterium *Microcystis aeruginosa* NIES-843. *DNA Res*. 2007 Dec 31;14(6):247-56.

26. Yui Eto K, Firth N, Davis AM, Kwong SM, Krysiak M, Lee YT, O'Brien FG, Grubb WB, Coombs GW, Bond CS, Ramsay JP. Evolution of a 72-Kilobase Cointegrand, Conjugative Multiresistance Plasmid in Community-Associated Methicillin-Resistant *Staphylococcus aureus* Isolates from the Early 1990s. *Antimicrob Agents Chemother*. 2019 Oct 22;63(11):e01560-19.
27. Zong Z. The complex genetic context of blaPER-1 flanked by miniature inverted-repeat transposable elements in *Acinetobacter johnsonii*. *PLoS One*. 2014 Feb 25;9(2):e90046.
28. Adams FG, Brown MH. MITE *Aba12*, a Novel Mobile Miniature Inverted-Repeat Transposable Element Identified in *Acinetobacter baumannii* ATCC 17978 and Its Prevalence across the *Moraxellaceae* Family. *mSphere*. 2019 Feb 20;4(1):e00028-19.
29. Robertson AE, Wechter WP, Denny TP, Fortnum BA, Kluepfel DA. Relationship between avirulence gene (*avrA*) diversity in *Ralstonia solanacearum* and bacterial wilt incidence. *Mol Plant Microbe Interact*. 2004 Dec;17(12):1376-84.
30. Stavrinides J, Kirzinger MW, Beasley FC, Guttman DS. E622, a miniature, virulence-associated mobile element. *J Bacteriol*. 2012 Jan;194(2):509-17.
31. Naito M, Hirakawa H, Yamashita A, Ohara N, Shoji M, Yukitake H, Nakayama K, Toh H, Yoshimura F, Kuhara S, Hattori M, Hayashi T, Nakayama K. Determination of the genome sequence of *Porphyromonas gingivalis* strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in *P. gingivalis*. *DNA Res*. 2008 Aug;15(4):215-25.
32. Klein BA, Chen T, Scott JC, Koenigsberg AL, Duncan MJ, Hu LT. Identification and characterization of a minisatellite contained within a novel miniature inverted-repeat transposable element (MITE) of *Porphyromonas gingivalis*. *Mob DNA*. 2015 Oct 6;6:18.
33. Mattes TE, Alexander AK, Richardson PM, Munk AC, Han CS, Stothard P, Coleman NV. The genome of *Polaromonas* sp. strain JS666: insights into the evolution of a hydrocarbon- and xenobiotic-degrading bacterium, and features of relevance to biotechnology. *Appl Environ Microbiol*. 2008 Oct;74(20):6405-16.
34. Halter T, Hendrickx F, Horn M, Manzano-Marín A. A Novel Widespread MITE Element in the Repeat-Rich Genome of the *Cardinium* Endosymbiont of the Spider *Oedothorax gibbosus*. *Microbiol Spectr*. 2022 Dec 21;10(6):e0262722.
35. Brügger K, Torarinsson E, Redder P, Chen L, Garrett RA. Shuffling of *Sulfolobus* genomes by autonomous and non-autonomous mobile elements. *Biochem Soc Trans*. 2004 Apr;32(Pt 2):179-83.
36. Guo L, Brügger K, Liu C, Shah SA, Zheng H, Zhu Y, Wang S, Lillestøl RK, Chen L, Frank J, Prangishvili D, Paulin L, She Q, Huang L, Garrett RA. Genome analyses of Icelandic strains of *Sulfolobus islandicus*, model organisms for genetic and virus-host interaction studies. *J Bacteriol*. 2011 Apr;193(7):1672-80.
37. Jaubert C, Danioux C, Oberto J, Cortez D, Bize A, Krupovic M, She Q, Forterre P, Prangishvili D, Sezonov G. Genomics and genetics of *Sulfolobus islandicus* LAL14/1, a model hyperthermophilic archaeon. *Open Biol*. 2013 Apr 17;3(4):130010.
38. Brügger K, Redder P, She Q, Confalonieri F, Zivanovic Y, Garrett RA. Mobile elements in archaeal genomes. *FEMS Microbiol Lett*. 2002 Jan 10;206(2):131-41.
39. Ravin NV, Mardanov AV, Beletsky AV, Kublanov IV, Kolganova TV, Lebedinsky AV, Chernyh NA, Bonch-Osmolovskaya EA, Skryabin KG. Complete genome sequence of the anaerobic, protein-degrading hyperthermophilic crenarchaeon *Desulfurococcus kamchatkensis*. *J Bacteriol*. 2009 Apr;191(7):2371-9.
40. Kamoun C, Payen T, Hua-Van A, Filée J. Improving prokaryotic transposable elements identification using a combination of de novo and profile HMM methods. *BMC Genomics*. 2013 Oct 11;14:700.
41. Chen J, Hu Q, Zhang Y, Lu C, Kuang H. P-MITE: a database for plant miniature inverted-repeat transposable elements. *Nucleic Acids Res*. 2014 Jan;42(Database issue):D1176-81.
42. Redder P, She Q, Garrett RA. Non-autonomous mobile elements in the crenarchaeon *Sulfolobus solfataricus*. *J Mol Biol*. 2001 Feb 9;306(1):1-6.
43. Zo YG. Reductive divergence of enterobacterial repetitive intergenic consensus sequences among Gammaproteobacteria genomes. *J Microbiol*. 2011 Feb;49(1):35-45.
44. Wilson LA, Sharp PM. Enterobacterial repetitive intergenic consensus (ERIC) sequences in *Escherichia coli*: Evolution and implications for ERIC-PCR. *Mol Biol Evol*. 2006 Jun;23(6):1156-68.
45. Duchaud E, Rusniok C, Frangeul L, Buchrieser C, Givaudan A, Taourit S, Bocs S, Boursaux-Eude C, Chandler M, Charles JF, Dassa E, Deroose R, Derzelle S, Freyssinet G, Gaudriault S, Médigue C, Lanois A, Powell K, Siguier P, Vincent R, Wingate V, Zouine M, Glaser P, Boemare N, Danchin A, Kunst F. The genome sequence of the entomopathogenic bacterium *Phototrhaddus luminescens*. *Nat Biotechnol*. 2003 Nov;21(11):1307-13.
46. Chen SL, Shapiro L. Identification of long intergenic repeat sequences associated with DNA methylation sites in *Caulobacter crescentus* and other alpha-proteobacteria. *J Bacteriol*. 2003 Aug;185(16):4997-5002.
47. Nelson WC, Bhaya D, Heidelberg JF. Novel miniature transposable elements in thermophilic *Synechococcus* strains and their impact on an environmental population. *J Bacteriol*. 2012 Jul;194(14):3636-42.

48. Delilhas N. Enterobacterial small mobile sequences carry open reading frames and are found intragenically-evolutionary implications for formation of new peptides. *Gene Regul Syst Bio*. 2007 Oct 16;1:191-205.
49. Croucher NJ, Vernikos GS, Parkhill J, Bentley SD. Identification, variation and transcription of pneumococcal repeat sequences. *BMC Genomics*. 2011 Feb 18;12:120.
50. Black CG, Fyfe JA, Davies JK. A promoter associated with the neisserial repeat can be used to transcribe the *uvrB* gene from *Neisseria gonorrhoeae*. *J Bacteriol*. 1995 Apr;177(8):1952-8.
51. Siddique A, Buisine N, Chalmers R. The transposon-like Correia elements encode numerous strong promoters and provide a potential new mechanism for phase variation in the meningococcus. *PLoS Genet*. 2011 Jan 20;7(1):e1001277.
52. Elbeyioglu F, Roberts SB, Spencer-Smith R, Pulijala M, Zelewska MA, Nebel JC, Snyder LAS. Inversion of Correia repeat enclosed elements in *Neisseria gonorrhoeae*. *Microbiology (Reading)*. 2017 Jan;163(1):31-36.
53. Romine MF, Carlson TS, Norbeck AD, McCue LA, Lipton MS. Identification of mobile elements and pseudogenes in the *Shewanella oneidensis* MR-1 genome. *Appl Environ Microbiol*. 2008 May;74(10):3257-65.
54. Goosen N, van de Putte P. The regulation of transcription initiation by integration host factor. *Mol Microbiol*. 1995 Apr;16(1):1-7.
55. Rouquette-Loughlin CE, Balthazar JT, Hill SA, Shafer WM. Modulation of the *mtrCDE*-encoded efflux pump gene complex of *Neisseria meningitidis* due to a Correia element insertion sequence. *Mol Microbiol*. 2004 Nov;54(3):731-41.
56. De Gregorio E, Abrescia C, Carlomagno MS, Di Nocera PP. The abundant class of nemis repeats provides RNA substrates for ribonuclease III in *Neisseriae*. *Biochim Biophys Acta*. 2002 Jun 7;1576(1-2):39-44.
57. Mazzone M, De Gregorio E, Lavitola A, Pagliarulo C, Alifano P, Di Nocera PP. Whole-genome organization and functional properties of miniature DNA insertion sequences conserved in pathogenic *Neisseriae*. *Gene*. 2001 Oct 31;278(1-2):211-22.
58. De Gregorio E, Silvestro G, Petrillo M, Carlomagno MS, Di Nocera PP. Enterobacterial repetitive intergenic consensus sequence repeats in *yersiniae*: genomic organization and functional properties. *J Bacteriol*. 2005 Dec;187(23):7945-54.
59. Wright MS, Jacobs MR, Bonomo RA, Adams MD. Transcriptome Remodeling of *Acinetobacter baumannii* during Infection and Treatment. *mBio*. 2017 Mar 7;8(2):e02193-16.
60. Deveson Lucas D, Crane B, Wright A, Han ML, Moffatt J, Bulach D, Gladman SL, Powell D, Aranda J, Seemann T, Machado D, Pacheco T, Marques T, Viveiros M, Nation R, Li J, Harper M, Boyce JD. Emergence of High-Level Colistin Resistance in an *Acinetobacter baumannii* Clinical Isolate Mediated by Inactivation of the Global Regulator H-NS. *Antimicrob Agents Chemother*. 2018 Jun 26;62(7):e02442-17.
61. Pegler JL, Oultram MJ, Mann CWG, Carroll BJ, Grof CPL, Eamens AL. Miniature Inverted-Repeat Transposable Elements: Small DNA Transposons That Have Contributed to Plant *MICRORNA* Gene Evolution. *Plants (Basel)*. 2023 Mar 1;12(5):1101.
62. Wolk CP, Lechno-Yossef S, Jäger KM. The insertion sequences of *Anabaena* sp. strain PCC 7120 and their effects on its open reading frames. *J Bacteriol*. 2010 Oct;192(20):5289-303.
63. Ogata H, Audic S, Barbe V, Artiguenave F, Fournier PE, Raoult D, Claverie JM. Selfish DNA in protein-coding genes of *Rickettsia*. *Science*. 2000 Oct 13;290(5490):347-50.
64. Ogata H, Audic S, Abergel C, Fournier PE, Claverie JM. Protein coding palindromes are a unique but recurrent feature in *Rickettsia*. *Genome Res*. 2002 May;12(5):808-16.
65. Federoff N, Wessler S, Shure M. Isolation of the transposable maize controlling elements *Ac* and *Ds*. *Cell*. 1983. 35:235-242.
66. Fewer DP, Halinen K, Sipari H, Bernardová K, Mänttari M, Eronen E, Sivonen K. Non-autonomous transposable elements associated with inactivation of microcystin gene clusters in strains of the genus *Anabaena* isolated from the Baltic Sea. *Environ Microbiol Rep*. 2011 Apr;3(2):189-94.
67. Bardaji L, Añorga M, Jackson RW, Martínez-Bilbao A, Yanguas-Casás N, Murillo J. Miniature transposable sequences are frequently mobilized in the bacterial plant pathogen *Pseudomonas syringae* pv. *phaseolicola*. *PLoS One*. 2011;6(10):e25773.
68. Blount ZD, Grogan DW. New insertion sequences of *Sulfolobus*: functional properties and implications for genome evolution in hyperthermophilic archaea. *Mol Microbiol*. 2005 Jan;55(1):312-25.
69. Gillings MR. Integrins: past, present, and future. *Microbiol Mol Biol Rev*. 2014 Jun;78(2):257-77.
70. Domingues S, Nielsen KM, da Silva GJ. The *blaIMP-5*-carrying integron in a clinical *Acinetobacter baumannii* strain is flanked by miniature inverted-repeat transposable elements (MITEs). *J Antimicrob Chemother*. 2011 Nov;66(11):2667-8.
71. Gillings MR, Labbate M, Sajjad A, Giguère NJ, Holley MP, Stokes HW. Mobilization of a Tn402-like class 1 integron with a novel cassette array via flanking miniature inverted-repeat transposable element-like structures. *Appl Environ Microbiol*. 2009 Sep;75(18):6002-4.
72. Mendes RE, Castanheira M, Toleman MA, Sader HS, Jones RN, Walsh TR. Characterization of an integron carrying *blaIMP-1* and a new aminoglycoside resistance gene, *aac(6)-31*, and its dissemination among

- genetically unrelated clinical isolates in a Brazilian hospital. *Antimicrob Agents Chemother.* 2007 Jul;51(7):2611-4.
73. Domingues S, Toleman MA, Nielsen KM, da Silva GJ. Identical miniature inverted repeat transposable elements flank class 1 integrons in clinical isolates of *Acinetobacter* spp. *J Clin Microbiol.* 2013 Jul;51(7):2382-4.
 74. Minnick MF, Raghavan R. Genetics of *Coxiella burnetii*: on the path of specialization. *Future Microbiol.* 2011 Nov;6(11):1297-314.
 75. Hu J, Zheng Y, Shang X. MiteFinderII: a novel tool to identify miniature inverted-repeat transposable elements hidden in eukaryotic genomes. *BMC Med Genomics.* 2018 Nov 20;11(Suppl 5):101.
 76. Crescente JM, Zavallo D, Helguera M, Vanzetti LS. MITE Tracker: an accurate approach to identify miniature inverted-repeat transposable elements in large genomes. *BMC Bioinformatics.* 2018 Oct 3;19(1):348.
 77. Ge R, Mai G, Zhang R, Wu X, Wu Q, Zhou F. MUSTv2: An Improved De Novo Detection Program for Recently Active Miniature Inverted Repeat Transposable Elements (MITEs). *J Integr Bioinform.* 2017 Aug 10;14(3):20170029.
 78. Ye C, Ji G, Liang C. detectMITE: A novel approach to detect miniature inverted repeat transposable elements in genomes. *Sci Rep.* 2016 Jan 22;6:19688.
 79. Yang G. MITE Digger, an efficient and accurate algorithm for genome wide discovery of miniature inverted repeat transposable elements. *BMC Bioinformatics.* 2013 Jun 7;14:186.
 80. Han Y, Wessler SR. MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. *Nucleic Acids Res.* 2010 Dec;38(22):e199.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.