Miniature transposable sequences are frequently mobilized in the bacterial plant pathogen Pseudomonas syringae pv. phaseolicola

Article

Published Version

Bardaji, L., Añorga, M., Jackson, R. W., Martínez-Bilbao, A., Yanguas-Casás, N. and Murillo, J. (2011) Miniature transposable sequences are frequently mobilized in the bacterial plant pathogen Pseudomonas syringae pv. phaseolicola. PLoS ONE, 6 (10). e25773. ISSN 1932-6203 doi: 10.1371/journal.pone.0025773 Available at http://centaur.reading.ac.uk/23879/

It is advisable to refer to the publisher's version if you intend to cite from the work.
Published version at: http://dx.doi.org/10.1371/journal.pone.0025773
To link to this article DOI: http://dx.doi.org/10.1371/journal.pone.0025773

Publisher: Public Library of Science

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.
www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online
Miniature Transposable Sequences Are Frequently Mobilized in the Bacterial Plant Pathogen *Pseudomonas syringae* pv. phaseolicola

Leire Bardaji¹, Maite Añorga¹, Robert W. Jackson², Alejandro Martínez-Bilbao¹, Natalia Yanguas-Casás¹, Jesús Murillo¹*

¹Laboratorio de Patología Vegetal, Departamento de Producción Agraria, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Pública de Navarra, Pamplona, Spain, ²School of Biological Sciences, University of Reading, Reading, United Kingdom

**Abstract**

Mobile genetic elements are widespread in *Pseudomonas syringae*, and often associate with virulence genes. Genome reannotation of the model bean pathogen *P. syringae* pv. phaseolicola 1448A identified seventeen types of insertion sequences and two miniature inverted-repeat transposable elements (MITEs) with a biased distribution, representing 2.8% of the chromosome, 25.8% of the 132-kb virulence plasmid and 2.7% of the 52-kb plasmid. Employing an entrapment vector containing *sacB*, we estimated that transposition frequency oscillated between 2.6 × 10⁻⁵ and 1.1 × 10⁻⁶, depending on the clone, although it was stable for each clone after consecutive transfers in culture media. Transposition frequency was similar for bacteria grown in rich or minimal media, and from cells recovered from compatible and incompatible plant hosts, indicating that growth conditions do not influence transposition in strain 1448A. Most of the entrapped insertions contained a full-length IS801 element, with the remaining insertions corresponding to sequences smaller than any transposable element identified in strain 1448A, and collectively identified as miniature sequences. From these, fragments of 229, 360 and 679-nt of the right end of IS801 ended in a consensus tetranucleotide and likely resulted from one-ended transposition of IS801. An average 0.7% of the insertions analyzed consisted of IS801 carrying a fragment of variable size from gene PSPPH_0008/PSPPH_0017, showing that IS801 can mobilize DNA *in vivo*. Retrospective analysis of complete plasmids and genomes of *P. syringae* suggests, however, that most fragments of IS801 are likely the result of reorganizations rather than one-ended transpositions, and that this element might preferentially contribute to genome flexibility by generating homologous regions of recombination. A further miniature sequence previously found to affect host range specificity and virulence, designated MITEPsy1 (100-nt), represented an average 2.4% of the total number of insertions entrapped in *sacB*, demonstrating for the first time the mobilization of a MITE in bacteria.

**Introduction**

Insertion sequences (IS) are probably the simplest autonomous mobile DNA elements and generally consist of a transposase coding gene, responsible for their transposition, which is bound by terminal inverted repeats [1,2]. ISs can generate significant variability in bacteria and contribute to their evolution [3], in part because they are usually present in more than one copy per genome and thus represent mobile regions of recombination. Their mobility, together with their capacity to mobilize unrelated DNA in their proximity, can lead to a panoply of mutations and rearrangements in the host bacteria, which include insertions, deletions, duplications, translocations, cointegrations, inversions and gene activation [4]. From these activities, it easily follows that they have an enormous potential to alter the genome and influence bacterial evolution. They can also shuffle DNA among different genetic replicons such as chromosomes and plasmids sustaining a gene trading activity that widely contributes to the horizontal spread of genetic information [e.g. 5,6]. ISs are widespread among bacteria and archaea, present in nearly all of the sequenced genomes and often in high numbers [1]. There is also a large variety of ISs, with more than 2500 types included in the IS repository database and grouped in 25 families [7]. Additional types of small, non-autonomous mobile sequences are the REPINs (repetitive extragenic palindromic (REP) doublets forming hairpins) [8] and MITEs (miniature inverted-repeat transposable elements) [1]. MITEs are generally less than 300 bp long and usually contain terminal inverted repeat sequences; MITEs are thought to derive from ISs by internal deletions and to be mobilized *in trans* by the transposase of their parental IS [1,9]. The impact of MITE activity in the prokaryotic genome is potentially very high and their small size allows them to contribute to phenotypic variation in many different and creative ways, such as generating new gene alleles and functions, or new regulatory signals for preexisting genes [9].

*Pseudomonas syringae* is a plant pathogenic gamma proteobacterium that is being extensively used as a model to study the molecular bases of plant-microbe interactions and the evolution of
pathogenicity [10]. In this respect, *P. syringae* is a major study subject because of the large pathogenic variation within the species. Indeed, *P. syringae* can be divided into at least 60 groups, or pathovars, that are characterized by their host range [11]. While the different pathovars show a highly conserved core genome, there is a large variation in their virulence gene complement [12,19]. The pathogenicity of *P. syringae* depends on the activity of a type III secretion system that delivers specialized proteins, known as effectors, into the plant cell, where they contribute to the suppression of the plant defence responses and the establishment of an infection [10]. Occasionally, effectors are recognized by the plant machinery, leading to the activation of a general defense response, the hypersensitive response, which ultimately leads to a complete plant resistance phenotype. Therefore, effectors might have a dual effect, either promoting pathogenicity and virulence or restricting host range in specific plants. Remarkably, effectors and other virulence genes are commonly bordered by ISs and other types of repeated elements in diverse strains of *P. syringae* [14,15,16], which would likely favor their frequent horizontal transfer and exchange. This is supported by the fact that mobile genetic elements are often associated with regions that are not syntenic among different *P. syringae* genomes [17,18,19]. Sequencing of native plasmids and genomes has revealed a diverse collection of ISs in pathovars of *P. syringae* [17,19,20,21,22], with partial or complete sequences of 32 newly described *P. syringae* ISs included in the IS Finder repository. However, it is not possible to estimate the actual diversity due to the lack of a complete IS inventory and the use of different names for the same element. Additionally, there is a large variation in the types and frequency of ISs in the only three complete genomes currently available, of *P. syringae* pv. tomato DC3000 [19], *P. syringae* pv. phaseolicola 1448A [17] and *P. syringae* pv. syringae B728a [21]. Nevertheless, the ability to transpose within *P. syringae* has been demonstrated for only a few elements, including IS443 (syn. ISPsy2) and IS52 [23], IS801 [24], ISPsy2 and ISPsy3 [25], and the hopXI effector transposon [26].

*P. syringae* pv. phaseolicola is the causal agent of halo blight of beans (*Phaseolus vulgaris* L) and is used as a model for the study of the molecular basis of pathogenicity and virulence [27]. We are interested in the characterization of transposable elements in this bacterium in order to explore their impact on the evolution of virulence and in the frequent genome changes that were reported to occur during the interaction with plant hosts [28]. Additionally, and due to the current lack of other suitable markers, we want to expand the use of repeated DNA for typing populations [29,30]. Repeated DNA has been directly involved in shaping host range in *P. syringae* pv. phaseolicola by the inactivation or alteration of effector genes [31,32]. A chimeric transposable element promoted large deletions eliminating effector gene *avrPphF* (syn. *hopF1*), allowing the emergence of populations that overcame the resistance conferred by gene *Rl*, which was widely used for the control of halo blight [32]. Likewise, the generation of new alleles of effector gene *avrPphE* (syn. *hopXI*) after the insertion of a small repeated sequence [31], which has been determined in this work to be a MITE (see below), lead to the emergence of new pathogenic races. Genome rearrangements have also been linked to the activity of repeated DNA. Originally isolated from a *P. syringae* pv. phaseolicola strain, IS801 was shown to participate in the transmission and excision of a native plasmid mediated by recombination between copies of the IS, which resulted in a dynamic exchange of DNA between the plasmid and the chromosome [24,33]. IS801 belongs to the IS91 family of rolling circle transposable elements, which seem to be preferentially involved in the dissemination of pathogenicity-related genes [34,35,36]. Nevertheless, there is still limited information on the content and functionality of transposable elements in this pathogen. Five percent of the ORFs in the closed genome of *P. syringae* pv. *phaseolicola* 1448A were identified as mobile genetic elements [17], but this figure includes integrases and phage-related DNA sequences, making it difficult to estimate the relevance of transposable elements. Although an inventory of transposable elements in strain 1448A is not available, the annotation of the genome (accession no. CP000050) includes the coding regions for transposases, but not other non-coding sequences related to transposable elements, limiting the utility of this information.

To characterize and quantify the impact of transposable elements on the life cycle of *P. syringae*, we analyzed and reannotated the genome of the model bacterium *P. syringae* pv. *phaseolicola* 1448A. We also used an IS trapping vector to identify functionally active elements and assess their impact in the genome flexibility. This study provides a foundation for a larger scale analysis of transposable elements in other *P. syringae* genomes.

**Results**

*P. syringae* pv. *phaseolicola* 1448A contains at least seventeen insertion sequences and two MITEs

A comparison of the genome of *P. syringae* pv. *phaseolicola* 1448A with the databases indicates that it harbors at least seventeen different types of insertion sequences and two MITEs (Table 1 and Annotation files S1, S2 and S3). The transposases of these insertion sequences have all been previously annotated in the genome of strain 1448A [17], and here we contributed the definition and annotation of the complete elements, as well as the identification of 21 truncated copies and several MITEs that were not annotated previously.

The copy number of transposable elements was highly variable, from one to 48 complete copies, and many of them were fragmented, suggesting the occurrence of DNA reorganizations. In total, the mobile DNA amounted to at least 199,018 nt, representing 3.3% of the nearly 6 Mb genome (Table 2); however, the density of ISs varied for the chromosome (2.8%), the 132 kb plasmid (25.8%), and the 52 kb plasmid (2.7%). Coding sequences (CDSs) for transposases were reported to clearly disrupt 8 reading frames in the 1448A genome (Table S1) [17]. Our annotation of the complete mobile elements reveals that they interrupt at least five other CDSs and form chimeras with at least 40 CDSs (Table S1). The majority of these CDSs would likely have an altered functionality or not be functional at all, as occurs with the interrupted CDSs and 11 chimeric CDSs whose 5' end correspond to mobile DNA sequences.

Nine copies of IS801 (copies IS801:T1 to T6 and T10 to T12) and three copies of ISPsy19 (ISPsy19:T6 to T8), were of wild type length but contained a premature stop in the transposase gene. Since the IS801 transposase can efficiently act in trans [36], it is conceivable that the interrupted derivatives of IS801 could also transpose in the cell. Element ISPsy17:T5 is also full-length but interrupted by a tandem insertion of ISPsy19 and of a new mobile element identified in this work and designated MITE/Psy1 (see below). Additionally, element ISPsy17:T8 contains an independent insertion of MITE/Psy1 14 nt upstream of the transposase start codon, although this would predictably not affect transposition of the element because it does not affect the terminal repeats or the CDS.

There is a very limited description of insertion sequences ISPsy17, ISPsy19 and ISPsy24, for which nearly only the transposase genes were identified [17,32]. Therefore, to estimate the size of the complete element, we aligned the coding sequences
The right border defined in this way coincided with that previously bordered by 28 nt imperfect repeats (Tables 1 and 3).

Length insertion sequence.

resulting from insertion, was considered to be that of the full-contained terminal repeats, but not any duplicated sequence.

of the corresponding transposable elements found in the genome of 1448A plus up to 1 kb on either side. The longest consensus sequence that contained terminal repeats, but not any duplicated sequence resulting from insertion, was considered to be that of the full-length insertion sequence.

For ISPsy17, we determined a consensus sequence of 1374 nt that was bordered by 28 nt imperfect repeats (Tables 1 and 3). The right border defined in this way coincided with that previously

The following insertion elements, previously found in P. syringae, were not present in 1448A: ISPsy1, ISPsy1, ISPsy1, ISPsy5, ISPsy6, ISPsy7, ISPsy8, ISPsy9, ISPsy10, ISPsy11, ISPsy12, ISPsy13, ISPsy14, ISPsy15, ISPsy27, ISPsy28, ISPsy30, IS52, and IS1240. The size of ISPsy25 and ISPsy26 was estimated to be at least the size of the transposase plus the surrounding DNA that included the typical bordering inverted repeats.

Correspond to loci PSPPH_0182, PSPPH_0183. Loci PSPPH_0182 belongs to family Pfam PF05621, of NTP-binding proteins involved in transposition. PSPPH_0183 is reorganized, but contains a C-terminal Mu transposase domain (PF09299) found in various prokaryotic integrases and transposases.

Correspond to loci PSPPH_3494, PSPPHA0085, and PSPPHA0131.

Corresponds to locus PSPPH_3497.

Corresponds to locus PSPPH_4298.

doi:10.1371/journal.pone.0025773.t001

doi:10.1371/journal.pone.0025773.t002

defined from deletion variants generated by a chimeric element derived from ISPsy17 [32]. The element includes a 1260 nt ORF that could code for a putative 419 aa transposase, which was found in 48 intact copies and 62 in total, in the genome of P. syringae pv. phaseolicola 1448A (Table 1). The predicted start codon of the transposase is located only 53 nt downstream of the 5’ end of the element, which may indicate that transcription of the transposase gene could be exclusively under the control of promoters upstream of the insertion point (i.e. promoters outside of the IS). We analyzed a comparison of the 1448A genome with those of P. syringae pv. syringae B728a and P. syringae pv. tomato DC3000, which do not harbor ISPsy17. Only 13 of the 47 chromosomal insertions of ISPsy17 (insertions 1, 2, 4, 8, 10, 12, 16, 19, 27, 37–39, 45) in 1448A were located in DNA regions that were otherwise complete in at least one of the two other bacteria. These insertions, that we will call here “genuine insertions”, were surrounded in 10 out of 13 cases by a perfect target duplication of 6 or 8 nt. In all, twenty four insertions were surrounded by an 8 nt target duplication, whereas others were surrounded by either a 6 nt duplication, an imperfect 4 to 5 nt direct repeat or by no discernible target repeat (Table 4). Although in all cases the sequence was apparently unique for each of the target duplications, we evaluated the possible occurrence of a target preference of DNA corresponding to putative insertion sequences in the genome of P. syringae pv. phaseolicola 1448A.

Table 1. Type and number of mobile elements found in the genome of P. syringae pv. phaseolicola 1448A.

| Mobile element | Synonym | Family | Size (nt) | No. of CDSs | Inverted repeats | Chromosome | pA | pB |
|---------------|---------|--------|-----------|-------------|----------------|------------|----|----|
| IS51         | ISPsy21 | IS3    | 1312      | 2           | 26             | 2          | 1  | (1) |
| IS53         | ISPsy20 | IS21   | 2572      | 2           | 27             | 2          | (3) |    |
| IS801        | IS91    | IS91   | 1512      | 1           | no             | 3          | (6) | 1  | (10) |
| ISPsy2       | IS5     | IS5    | 1194      | 1           | 12             | 5          | (6) | (2) |    |
| ISPsy3       | IS91    | IS91   | 1507      | 1           | no             |            | (3) |    |
| ISPsy4       | ISPsy23 | IS21   | 1962      | 2           | 23             | (3)        | (1) |    |
| ISPsy16      | IS110   | IS110  | 1461      | 1           | 12             | 2          | (1) | 3  | (3) |
| ISPsy17      | ISPsy18 | IS256  | 1374      | 1           | 28             | 47         | (10)| (4) | 1  |
| ISPsy19      | IS5     | IS5    | 1178      | 1           | 17             | 28         | (10)| 3  | |
| ISPsy22      | IS5     | unk    | unk       | unk         | unk            |            | (2) |    |    |
| ISPsy24      | IS5     | 1235   | 2         | 26           | 2              | 4          | (1) |    |    |
| ISPsy25      | IS630   | IS630  | 1177      | 1           | 19             | 1          | (1) |    |    |
| ISPsy26      | ISPsy29 | IS63  | 1231      | 2           | 28             |            | 1   | (1) |    |
| unnameda     | unk     | unk    | unk       | unk         | unk            |            | (1) |    |    |
| unnameda     | Tn3     | unk    | unk       | unk         | unk            |            | (1) | (2) |    |
| unnameda     | IS66    | unk    | unk       | unk         | unk            |            | (1) |    |    |
| unnameda     | IS66    | unk    | unk       | unk         | unk            |            | (1) |    |    |
| MITEPsy1     | IS57    | 100    | 0         | 18           | 3              | 5          | 1  |    |    |
| MITEPsy2     | unk     | 228    | 0         | 26           | 1              | 1          | 1  |    |    |

Table 2. Amount of DNA corresponding to putative insertion sequences in the genome of P. syringae pv. phaseolicola 1448A.

| Size (kb) | Insertion sequence DNA (kb) | % of ISs |
|----------|----------------------------|----------|
| Chromosome | 5928  | 163.6  | 2.8   |
| p1448A-A  | 132   | 34.0   | 25.8  |
| p1448A-B  | 52    | 1.4    | 2.7   |
| Total     | 6112  | 199.0  | 3.3   |

doi:10.1371/journal.pone.0025773.t0002

doi:10.1371/journal.pone.0025773.t0001

doi:10.1371/journal.pone.0025773.t0000

Miniature Transposable Sequences in P. syringae
for ISPsy17. For this, we constructed alignments of all the insertions with a perfect 8 nt duplication, because members of the IS256 family, such as ISPsy17, generate an 8 or 9 bp direct target repeat [2]. From these alignments, we could not discern any obvious consensus sequence that could serve as the insertion target, except that there was always a C or a T 24 nt downstream of the ISPsy17, suggesting that this element might insert randomly.

ISPsy19 was defined as a 1178 nt element containing a perfect 17 nt inverted repeat (Tables 1 and 3) and coding for a putative 367 aa transposase. There were 31 intact copies and it was often associated with other mobile elements, with 9 complete or degenerate insertions adjacent to ISPsy2 and three adjacent to MITEPsy1. Six chromosomal insertions of ISPsy19 (insertions 7, 10, 14, 18, 27, and T7) could be considered genuine insertions in a genome comparison with strains B728a and DC3000. Fifteen of the ISPsy19 copies were surrounded by a 3 nt perfect duplication of the target DNA; alignment of the DNA surrounding these elements suggests that the insertion target for ISPsy19 is the sequence CHHD.

ISPsy24 is a 1235 nt element surrounded by an imperfect 26 nt repeat (Tables 1 and 3) and coding for two putative proteins of 102 and 273 aa, of which the first belongs to the ISJ family of transposases (InterPro IPR002514) whereas the second contains a ribonuclease H-like domain (InterPro IPR012337). ISPsy24 is present only in two intact copies, and only one of them generated a 4 nt target repeat; therefore, it is not possible to deduce if this element has any target preference or if it leads to target duplication.

A 104 nt sequence with terminal inverted repeats was previously found inserted into effector gene avrPphE (syn. hopX1) in race 8 strains of P. syringae pv. phaseolicola, leading to a change in host range [31]. We found that this sequence is present in several copies in the 1448A genome (Table 1) and was able to transpose (see below); we therefore considered it to be a functional MITE and designated it as MITEPsy1. Comparison of the MITEPsy1 insertions in 1448A and in other P. syringae pv. phaseolicola strains (accession no. AB023077 and AY147025), plus the analysis of the insertions entrapped with pGEN500 (Table 5), indicate that this element is 100 nt long, has 18 nt terminal inverted repeats and that it inserts after the sequence CTAG or YTAA, which is duplicated upon insertion (Table 3). The high similarity between their repeated sequences suggests that MITEPsy1 might have generated from ISPsy2 (Figure 1). BlastN comparisons with microbial genomes indicated that well conserved (>97% identity) copies of MITEPsy1 are present only in P. syringae pv. aesculi 0893-23 and NCPPB 3681, glycinea B076, lachrymans M301315, mori 301020 and savastanoi NCPPB 3335, all belonging to genomospecies 2. Additional sequences homologous to MITEPsy1 were only found in the genomes of P. syringae pv. actinidiae M302091 (91 nt, 77% identity) and tomato DC3000 (52 nt, 91% identity) (Figure S1), indicating a somewhat restricted distribution of this element in the P. syringae complex. Additionally, the genomes of P. stutzeri ATCC 17588 and of the denitrifying and alkane degrading Gammaproteobacteria Hdn1 (accession no. FP929140) contain an identical 99 nt sequence that showed a global 71% identity with MITEPsy1 (Figure S1). A further Blast

**Table 3.** Characteristics of selected mobile elements in P. syringae pv. phaseolicola 1448A.

| Insertion element | Terminal inverted repeats* | Sequence duplicationb | Target sequenceb |
|-------------------|---------------------------|----------------------|-----------------|
| ISPsy17           | GAGACTGTCAGAAAAATTGTGTTGGGC | Probably 8 nt         | Random          |
| ISPsy19           | GAGGGGTTGATAGACAAAAT       | 3 nt                 | CHHD            |
| ISPsy24           | TGTTAGGTTGCTATTGAAACCCGACAC | 4 nt ?               | unk             |
| MITEPsy1          | GAGGGGTTGATAGACAAAAT       | 4 nt                 | CTAG or YTAA    |
| MITEPsy2          | GGGGGTTGATAGACAAAAT       | unk                  | unk             |

*Terminal inverted repeats are shown 5’-3’, with the left repeat in the top and the right repeat at the bottom.

bunk, unknown.
doi:10.1371/journal.pone.0025773.t003

**Table 4.** Nucleotide direct repeats found bordering insertions of ISPsy17 in P. syringae pv. phaseolicola 1448Aa.

| Type of insertion | Direct repeat | Genuine | Other |
|------------------|---------------|---------|-------|
| Perfect 8 nt     | 7             | 13      |
| Imperfect 8 ntb  | 3             | 1       |
| 7 nt             | -             | 3       |
| 6 nt             | 3             | 1       |
| Degeneratec      | -             | 5       |
| None             | -             | 12      |

aOnly those insertions containing a complete copy of ISPsy17 were analyzed.

bGenuine insertions are those that disrupt a sequence in 1448A that is present only in P. syringae pv. phaseolicola 1448A.

cDegenerate indicates direct repeats of 4-5 nt, which might be surrounded by mismatches.
doi:10.1371/journal.pone.0025773.t004
search with the homolog from \textit{P. stutzeri} found similar sequences in plasmids or chromosomes of diverse \textit{Pseudomonas} species, \textit{Marinobacter aquaeolei}, \textit{Klebsiella pneumoniae} and plasmid pRSB105 from an uncultured bacterium, suggesting that sequences similar to MITE\textsubscript{Psy1} might be widespread among bacteria.

A putative 228 nt MITE, designated MITE\textsubscript{Psy2}, is present in the native plasmids of \textit{P. syringae pv. phaseolicola} strain NCPPB 3335 \cite{37}. This element contains terminal repeats highly similar to those of IS\textit{Psy30}, an insertion sequence of the \textit{Tn3} family also present in strain NCPPB 3335 and in several other \textit{P. syringae} strains (Figure 1). MITE\textsubscript{Psy2} is also present in two copies in the 1448A genome (Table 1), and shows 26 nt imperfect inverted repeats (Table 3). Homologs of this MITE are widespread in the \textit{P. syringae} group, although it shows high sequence variability with identity levels in pair comparisons as low as 78%.

Importantly, neither MITE\textsubscript{Psy1} nor MITE\textsubscript{Psy2} were identified using the programs MUST \cite{38} and MITE-Hunter [Han, personal communication; 39], specifically designed for the detection of MITEs, indicating the need to design enhanced prediction programs.

| Mobile element | Size (kb) | PT1 | PT2 | PT3 | T1 | T2 | T3 | T4 | Total no. | percentage of suc\textsuperscript{b} | percentage of insertions\textsuperscript{a} |
|----------------|----------|-----|-----|-----|----|----|----|----|-----------|----------------|---------------------------|
| Transposition frequency \((\times 10^{-4})\) | 6.0±3.0 | 3.7±1.2 | 3.8±0.6 | 13.0±2.4 | 1.1±0.5 | 9.8±0.8 | 9.6±0.8 | | | | |
| IS807 | 1.5 | 1 | 0 | 0 | 1 | 0 | 1 | 3 | 0.7 | 0.7 | |
| | 1.5 | 69 | 62 | 56 | 25 | 30 | 31 | 27 | 300 | 65.2 | 71.2 | |
| | 0.679 | 19 | 13 | 25 | 6 | 4 | 1 | 3 | 71 | 15.4 | 16.9 | |
| | 0.360 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 2 | 0.4 | 0.5 | |
| | 0.229 | 6 | 15 | 8 | 2 | 0 | 3 | 1 | 35 | 7.6 | 8.3 | |
| MITE\textsubscript{Psy1} | 0.1 | 1 | 0 | 4 | 1 | 1 | 1 | 2 | 10 | 2.2 | 2.4 | |
| None | - | 4 | 9 | 8 | 5 | 4 | 1 | 5 | 39 | 8.5 | | |
| Total no. | 100 | 99 | 101 | 40 | 40 | 40 | 40 | 40 | 460 | | |

\textsuperscript{a}Average percentage of each insertion type over the total number of insertions on suc\textsubscript{R}.

The transposition frequency is similar in different growing conditions

Our analysis suggested that there have been unique, independent transposition events within 1448A. To test the hypothesis that some of these transposons are transposing within the genome, we estimated transposition frequency of mobile DNA in \textit{P. syringae pv. phaseolicola} using the entrapment vector pGEN500 \cite{40}. This plasmid contains gene \textit{sacB} from \textit{Bacillus subtilis}, which confers sucrose-dependent lethality to many Gram-negative bacteria, thus providing a positive selection for insertions of mobile elements into \textit{sacB} \cite{41}. We examined transposition using individual transformants of \textit{P. syringae pv. phaseolicola} 1448A as well as heterogeneous bacterial populations originated from single transformation experiments, which we designated as “pool of transformants”.

The overall transposition frequency of different pools of transformants or clones of \textit{P. syringae pv. phaseolicola} 1448A(pGEN500), originating from independent electroporation events, ranged from \(2.6\times10^{-5}\) to \(1.1\times10^{-6}\), depending on the transformant (Table 5). Similar frequencies were observed after six consecutive rounds of growth in liquid KMB plus tetracycline (not shown), suggesting that suc\textsuperscript{R} clones did not accumulate in the bacterial populations. We found comparable transposition frequencies using \textit{P. syringae pv. phaseolicola} strain 1449B (race 7; not shown), suggesting that the data obtained with strain 1448A might be representative of this pathovar. Since there is a similar proportion of the different types of insertion among the different transformants analyzed (Table 5), this variation must be due to phenomena that have a general effect. An obvious explanation is that, for each transformant, there is a differential level of toxicity of the entrapment vector in the absence of sucrose. This would predictably cause the premature death of cells, even though they are not exposed to sucrose, and would result in an artificial increase of the apparent rate of transposition. Indeed, in some clones we observed a lack of correspondence between the optical density of the culture and the expected number of colony forming units (not shown). If this would be the case, then the use of entrapment vectors based on \textit{sacB} would not allow for the estimation of accurate absolute transposition frequencies, at least in \textit{P. syringae}, although it could be used for comparative analyses.

Transposition activity has been reported to be influenced by the growth environment \cite{42}, and particularly by diverse stressful conditions.
conditions [40,43,44], although the effect of host and growth factors on transposition are considered to be specific for each mobile element [2,42]. We therefore evaluated the transposition frequency in bacteria subjected to both favorable and stressful conditions using five independent transformants of strain 1448A(pGEN500). For all the conditions tested, there were no significant differences among the frequencies estimated independently for each of the transformants and, in consequence, further analyses were done using the combined data of all the transformants for each condition. The average transformation frequency was not significantly different for any of the conditions tested, which included bacteria grown in liquid rich medium (KMB; transposition frequency 7.72±6.85x10⁻⁶) or minimal medium (MG; 6.56±5.75x10⁻⁶), and bacteria recovered from artificially inoculated bean cv. Tendegreen (compatible host; 6.77±3.51x10⁻⁶) and tobacco leaves (incompatible host; 3.87±2.31x10⁻⁶). These results suggest that growth conditions do not have a significant influence on the transposition frequency of mobile DNA in P. syringae pv. phaseolicola 1448A.

Only two mobile elements were identified in a functional assay using an entrapment vector

We were able to use the strains recovered from the previous experiments to identify the mobile DNA disrupting sacB in pGEN500. A preliminary sequence analysis of over 40 clones indicated that pGEN500 contained insertions of IS801 in nearly all the sucR colonies. Therefore, native plasmid profiles from 460 sucR colonies obtained from three independent pools of transformants and four individual transformants were analyzed by Southern hybridization using a specific IS801 probe. Only 10.7% of the plasmids in sucR clones did not hybridize with IS801 (Table 5). Sequencing of the sacB gene in these strains revealed that they either contained insertions of MITEPsy1 (2.2% of the plasmids) or had deletions or point mutations in sacB (8.5%). All the insertions of MITEPsy1 produced, as expected, a 4 nt duplication of the target sequence (Table 3).

The remaining pGEN500 derivatives, representing 89.3%, contained DNA cross-hybridizing to the IS801 probe (Table 5), although the size of the insertions was variable. Only 65% of these plasmids contained inserts of a size compatible with the full-length IS801 (1512 nt), whereas the remaining plasmids contained inserts derived from IS801 of 229 nt, 360 nt, 679 nt, and over 1.5 Kb (Figure 2), as shown by DNA sequencing. The 229, 360 and 679 nt derivatives, representing 26.3% of the total number of IS801 insertions, are the result of a one-ended transposition and consist of continuous fragments flanked by the right terminal repeat and a tetranucleotide with homology to the wild type left terminus (Figure 2), as shown by DNA sequencing. The 229, 360 and 679 nt derivatives, representing 26.3% of the total number of IS801 insertions, are the result of a one-ended transposition and consist of continuous fragments flanked by the right terminal repeat and a tetranucleotide with homology to the wild type left terminus. The mobilized fragments of gene PSPPH_008/PSPPH_0017 that were mobilized by IS801 in three different experiments, with an indication of their sizes. Drawings are to scale.
doi:10.1371/journal.pone.0025773.g002

right repeat and lacked the left repeat. However, only two of the five fragments could be considered as the result of a one-ended transposition because their 5’ end was homologous to the left repeat of IS801; both fragments were identical to the 229 nt functional fragment identified in this work. The analysis of partial sequences from diverse P. syringae strains previously described to contain IS801 fragments [34] also showed that none of them had the typical characteristics of miniature fragments originating from one-ended transposition (containing an intact right end repeat and having as a left end a tetranucleotide with homology to the wild type left end of IS801). This suggests that the large majority of IS801 fragments in the P. syringae genomes, and associated to virulence genes, is probably the result of DNA reorganizations.

Transpositions do not occur in bursts

Transposition could occur in the cell either as a single event or as multiple simultaneous events. To investigate this in P. syringae, we examined the hybridization pattern of five clones of P. syringae pv. phaseolicola 1448A(pGEN500), each containing a different type of insertion in sacB (see Table 5), using insertion sequences IS33 (syn. ISPsy20), IS801, ISPsy2, and ISPsy24 as probes. All the clones showed identical patterns of hybridization with all the probes, with the exception of the bands corresponding to the insertion in the entrapment plasmid, suggesting that a single insertion has occurred in each of them (Figure S2 and data not shown).

Discussion

The very large variability in the occurrence of ISs in prokaryotic genomes cannot be easily explained and, unexpectedly, the analysis of over 200 genomes indicated that the variability is apparently not correlated with pathogenicity or rates of horizontal gene transfer, with genome size being the only significant predictor of abundance and diversity [45]. The genome of P. syringae pv. phaseolicola 1448A contains 102 complete IS elements, which is nearly twice the number that would be expected given its genome size [45]. In this case, it is likely that the complement of virulence genes present in strain 1448A is partly responsible for this...
abundance, given the close association between virulence genes and mobile elements in \textit{P. syringae}, given the close association between virulence genes and mobile elements in \textit{P. syringae} \cite{14,34,46}. Elements in the genome of strain 1448A is relative low, 4 and 6 types of mobile elements. The number of intact copies of these two \textit{pv. phaseolicola} 1448A harbors at least nineteen putative insertions in p1448A-B and pPsv48B, at least in part, because they \cite{47,48}. It is therefore possible that evolution selects against and that might be implicated in plasmid survival or virulence \cite{47,48}, suggesting that insertions of these elements do not accumulate to high numbers. In particular, the number of insertions of IS801 do not appear to be very different in comparisons between and within the two genetic lineages previously defined in \textit{P. syringae pv. phaseolicola} \cite{29}. Remarkably, we observed the transposition of miniature sequences corresponding to partial fragments of IS801 (see below) that, together with MITEspy3 represented around 28% of the total mobilized elements in strain 1448A. We did not observe the transposition of any of the other mobile elements found in strain 1448A among the ca. 500 clones analyzed, indicating that they might transpose at a frequency lower than $10^{-7}$, that they have become fixed in the chromosome or that other conditions stimulate excision. Given their apparent stability, these mobile elements might therefore be suitable as epidemiological markers. The transposition frequency was stable for each clone in all growing conditions examined, including repeated transfers in culture, suggesting that the frequency of transposition in \textit{P. syringae} pv. phaseolicola is not influenced by the type of interaction with the plant host or by the growing medium. Nevertheless, we cannot rule out that the different growing conditions are affecting the relative proportion of each type of mobile element that are transposing, or their insertion specificity.

| Genotype | Molecule | accession no. | 5′ | Internal | 3′ | Both ends | Total no. | Possible mobile fragment$^{b}$ |
|---------|----------|---------------|----|----------|----|-----------|----------|-------------------|
| \textit{P. syringae pv. phaseolicola} 1448A | Chromosome | CP000058 | - | - | 1 | - | 1 | - |
| \textit{P. syringae pv. phaseolicola} 854326 | p1448A-A | CP000059 | 3 | 1 | 2 | 1 | 7 | 2 (229 nt) |
| \textit{P. syringae pv. phaseolicola} 854326 | pPMA4326B | AY603980 | - | 2 | - | 1 | 3 | - |
| \textit{P. syringae pv. tomato} DC3000$^{c}$ | Chromosome | AE016853 | - | 1 | - | 1 | 2 | - |
| \textit{P. syringae pv. phaseolicola} 8728A | pDC3000A | AE016855 | 1 | - | - | - | 1 | - |
| \textit{P. syringae pv. phaseolicola} 8728A | pDC3000B | AE016854 | 1 | 1 | 1 | - | 3 | - |

$^{a}$Only truncated fragments larger than 100 nt were taken into account, but not the complete element with premature stops; (-) indicates that no fragments were found in the category. Plasmid pFKN (accession no. AF359557) contains a copy of IS801, but not truncated forms. The following sequences (accession no.) did not contain any IS801 sequences: p1448A-B (CP000060), pPMA4326A (AY603979), pPMA4326C (AY603982), pPMA4326D (AY603981), pPMA4326E (AY603983), pPSR1 (AY342395), and \textit{P. syringae pv. syringae} B728A (CP000075).

$^{b}$We assumed that, as it happens with IS91 \cite{34}, the right end (3′ end) of IS801 is essential for transposition; therefore, partial fragments of IS801 were considered to be able to transpose when: (1) they contained an intact right end, and (2) the fragment started with any of the tetranucleotides described as an IS801 insertion target (GAAC, GGAC, CAAG and CGAC) \cite{36}; these partial fragments could be found only among those with a 5′ end deletion because fragments with internal deletions were specifically excluded.

$^{c}$IS801 has not been described in strain DC3000 \cite{19}, and the fragments found in this strain probably correspond to insertion sequence ISPsy3, a close relative of IS801 that also belongs to the IS91 family.

doi:10.1371/journal.pone.0025773.t006

Table 6. Type of truncated fragments derived from IS801 found in closed genomes of \textit{Pseudomonas syringae}.$^{a}$
IS801 transposition events (Table 5). The left end of members of the IS91 family shows a highly conserved 17 nt G-C rich sequence that includes a potential stem and loop structure [35]. However, we did not find this sequence in the 5’ end of any of the three miniature fragments identified here, suggesting that it might be irrelevant for transposition. Likewise, we did not identify in these ends any obvious repeat, palindrome or consensus sequence, except for the conserved tetranucleotide marking the left end of the fragment (see Figure 2). This suggests that any IS801 fragment could move in one-ended transposition event as long as it contains the right IS801 end and any of the tetranucleotides with homology to the left end of IS801 (GAAC). However, in the original sequence of IS801 (accession no. X57269) there are 7 occurrences of the tetranucleotides GAAC, GGAC, CAAG and CGAC, plus 11 of GCAC; these sequences border the IS801 miniature fragments (Figure 2) and were previously identified as target sequences for IS801 insertion [36], except GCAC. In spite of this, we observed one-ended transposition of only three of these 18 potential transposing fragments. Additionally, although the tetranucleotide GCAC found at the beginning of the 679 nt fragment has not been described as a target sequence for IS801, this fragment was mobilized with a much higher frequency than the others (Table 5). Therefore, it is likely that other sequences around the left end might be important for recognition of the IS801 left terminus and complete the transposition process. Importantly, the miniature IS801 derivatives identified here must be the result of the partial transposition of full-length elements, rather than the mobilization of pre-existing submolecular fragments, because there are no copies of the 679 or 360 nt fragments in the 1448A genome. This clearly indicates that transposition of a complete IS801 copy is normally inefficient, producing the mobilization of miniature fragments with a high frequency.

As occurs with IS91 [34], IS801 was also capable of mobilizing adjacent DNA, albeit at a very low frequency. In three independent events, we observed the mobilization of different partial fragments of the same gene, PSSPH_0008/PSSPH_0017 coding for the alpha subunit of a molybdopterin oxidoreductase, which in all cases started in the same tetranucleotide than the IS801 left end, GAAC (Figure 1). We do not have a satisfactory explanation for this preference; moreover, both PSSPH_0008 and PSSPH_0017 are probably non-functional because the IS801 insertion appears to have eliminated most of the 3’ end of the gene (Table S1), while at least 153 nt are missing from the 5’ end of the PSSPH_0008 reading frame in comparison with its closest homolog (not shown). Nevertheless, our results indicate that IS801 could potentially mobilize by transposition the virulence genes with which it is often associated. This could be relevant, because elements related to IS91 are involved in the mobilization of virtually every class of antibiotic resistance genes [50]. Although the mobilization of adjacent DNA by IS801 occurred at low frequency in our experimental conditions, it is possible that the interaction with plant hosts provides a highly selective environment favoring the exchange of effector and other virulence genes mobilized by IS801.

IS801 is limited in distribution within P. syringae, and is tightly associated to virulence genes [34], very often appearing as a partial element of various sizes and whose origins are unclear. Our analysis of partial fragments in complete plasmid and chromosome genomes of P. syringae suggests that most of them originated by recombination, rather than by one-ended transposition (Table 6), although a 229 nt fragment in the large plasmid of strain 1448A is interrupting a copy of ISPsy3 (PSSPH_A0015), suggesting that it originated by a true transposition event. Nevertheless, it is of course possible that any of these miniature fragments could mobilize accompanying DNA, even if this happens at a low frequency. One-ended transposition was proposed to be the main mechanism for gene propagation mediated by IS91 [34]; conversely, our results suggest that IS801, and partial fragments thereof, might preferentially contribute to the generation of recombination regions around virulence genes instead of serving as carrier elements. Indeed, insertion sequences are known to play a major role in genome flexibility, and variation between isolates of P. syringae appears to be due more to recombination than to mutation [31,32].

A remarkable outcome is the identification of a functionally active MITE in strain 1448A, that we have designated MITEPsy1; a second putative MITE, designated MITEPsy2, was also discovered in the genome at lower copy number, but this was not trapped in the transposition assay. MITEs are non-autonomous elements, in that they lack a transposase and rely entirely on transposases encoded within other elements existing within the genome, which act in trans [53]. The lower copy number of MITEPsy2 and the fact that it was not captured in the transposition assay may indicate that it is a recent addition to the 1448A genome, which lacks a transposase to mobilise it. MITE elements are widespread in bacteria and there is indirect evidence of their mobility [54,55]. For example, a MITE-like sequence was shown to transpose as a composite element and mobilized an antibiotic resistance gene [56]. However, this is the first time that a MITE is reported to actively transpose in vivo, and at a high frequency, opening the way to functionally test the requirements and molecular mechanisms for their transposition. MITEPsy1 has probably originated from ISPsy2, because their terminal repeats are nearly identical; in any case, it is likely that the ISPsy2 transposase is responsible for the mobility of MITEPsy1, because transposase specificity is generally determined by the sequence of the mobile element terminal repeats [2,4]. However, a potentially contradicting result is that we observed the frequent mobilization of MITEPsy1, reaching nearly 2.5% of the total number of insertions (Table 5), although in no case did we observe the mobilization of ISPsy2 despite strain 1448A encoding five complete copies of this element. Two of the six copies of MITEPsy1 present in the genome of 1448A are interrupting two loci, PSSPH_0770 and PSSPH_A0017 (Table S1), whereas another two are intergenic and a further two are inside other transposons. The insertion of MITEPsy1 into a CDS implies a change in the reading frame, because the element is 100 nt long and produces a 4 nt duplication, potentially leading to gene inactivation or to the generation of new alleles. Indeed, insertion of a 104 nt sequence (identified here as MITEPsy1 plus the 4 nt duplication) in the 3’ end of effector gene avrPphE (syn. hopX1) in a strain of P. syringae pv. phaseolicola race 8 lead to the generation of a new allele that no longer induced the hypersensitive response in resistant bean cultivars, causing the expansion of its host range [31]. Therefore, it shall be important to screen effector genes in different phytopathogenic bacteria for the presence of repeated sequences that might be surrounding them and could either participate in their mobility or alter their coding sequences, because they could represent new miniature mobile elements. Examples of this type of small repeated sequences have already been found in the chromosome and plasmids of different strains of P. syringae [17,57], whereas two different MITEs were found to alter the host range of Ralstonia solanacearum and contribute to the generation of epidemic genotypes of the pathogen [55].

Materials and Methods

Bacterial strains and growth conditions

Escherichia coli DH5a was routinely grown using LB medium [58] at 37°C and was used for cloning purposes. Pseudomonas
syringae pv. phaseolicola 1448A [race 6; 59] and 1449B [race 7; 60] were routinely propagated at 25°C using King’s medium B (KMB) [61] and the frequency of transposition was generally estimated using nutrient agar (NA; Oxoid, Basingstoke, UK) and NA supplemented with 5% (w/v) sucrose (SNA). Medium MG [62] was used as a minimal medium to evaluate nutrient limitation on transposition. When required, media were supplemented with tetracycline (Tc) at a final concentration of 12.5 μg ml⁻¹.

DNA hybridization and sequence analyses

For hybridization probes, we amplified a complete copy of IS901 or internal fragments of ISPsy2, IS53 and ISPsy24 from strain 1448A using specific primers (Table S2); these fragments were cloned in pGEM-T Easy and used as template for DNA amplification and labeling. Preparation of labeled probes with digoxigenin by PCR, Southern hybridization, and detection of hybridization signals were carried out with a DIG DNA labeling and detection kit (Roche Diagnostics) following the manufacturer's instructions.

To complete and update the genome annotation, the sequence of all the insertion sequences already found in P. syringae IS Finder Database; www-is.biotoul.fr [7], except ISPsy27 whose sequence was not available, were individually compared to the genome of P. syringae pv. phaseolicola 1448A using the genomic blastn program [63] at the NCBI website. Likewise, the genome of strain 1448A was compared to the IS Finder Database using the blast programs to examine for the existence of new, not registered, insertion sequences. Following the criteria of the IS Finder repository, an insertion sequence was considered new when its deduced product was less than 98% similar and/or its DNA sequence less than 95% identical to any other previously identified insertion sequence in the databases. During annotation, an element was considered to be truncated when it lacked part or all of at least one of the terminal repeats, or when it contained internal deletions or insertions that affected the transposase promoter or caused premature stops or deletions larger than 100 nt in its reading frame.

Multalin [64] and Blast 2 Sequences [65] were used to search for homology among sequences and to produce sequence alignments. Annotation was done with Artemis [66] and genome comparisons with ACT [67] using WebACT [68]. All the transposable elements, and truncated fragments larger than 200 nt, were annotated in the genome of 1448A and numbered consecutively; truncated copies, or those having a premature stop codon or deletion in the transposase coding region, were numbered separately and identified with the letter T and a codon or deletion in the transposase coding region, were transposable elements, and truncated fragments larger than 500 analyzed (see Table 5, and data not shown) were found to originate by the insertion of a mobile element, whereas the remaining 8.5% was due to small deletions or individual nt changes in sacB. The frequency of transposition in rich and minimal media was estimated for five independent transformants of strain 1448A(pGEN500) grown overnight in KMB or MG, both supplemented with tetracycline; each experiment was repeated at least seven times. To estimate transposition from cells growing in planta, those five transformants were independently inoculated on leaves of bean (Phaseolus vulgaris L.) cv. Tendergreen, or fully expanded leaves of tobacco (Nicotiana tabacum cv. xanthi) as previously described [71], using plants maintained in growth chambers at 20°C, a photoperiod of 16/8 h day/night and 80% of humidity. Six hours after inoculation with bacterial suspensions adjusted to an OD₆₀₀ of 0.5 (approximately 5×10⁸ cfu ml⁻¹), excised leaf disks were ground in sterile 10 mM MgCl₂ and appropriate dilutions plated on NA and SNA, both supplemented with tetracycline. The experiment was repeated at least three times.

The analysis of transposition frequency data in different growing conditions was done using a two-way ANOVA test (p<0.05).

Supporting Information

Figure S1 Alignment of MITEPsy1 with homologs in other bacteria. Global alignments were done with Blast, and curated manually, between [upper sequence] a copy of MITEPsy1 from P. syringae pv. phaseolicola 1448A (accession no. CP000058, positions 705671–705770) and [lower sequence] A contig 32.3 (accession no. AEAL01099292.1) from the draft genome of P. syringae pv. actinidiae M302091; B the genome of P. syringae pv. tomato DC3000 (accession no. AE016853); and C the genome of P. stutzeri ATCC 17586 (accession no. CP002881; an identical transformants of P. syringae pv. phaseolicola containing pGEN500 were selected on KMB plus tetracycline after electroporation, grown in the same conditions in liquid medium and stored at −80°C in 20% glycerol. To avoid the accumulation of insertions during routine transfer of cultures, the frequency of sacB inactivation was estimated, in general, using cultures started from cryopreserved transformants. To obtain pools of transformants, cells were incubated at 20°C for 2–4 h in KMB immediately after electroporation and a 100 μl aliquot was transferred to 5 ml of KMB+Tc. After overnight growth, aliquots of this culture were cryopreserved whereas other aliquots were directly used for the estimation of the transposition frequency. To isolate derivatives containing mobile elements inserted into sacB, cultures were diluted and spread onto SNA+Tc and the occurrence of insertions was monitored by examining changes in the mobility of pGEN500 in uncult plasmid preparations separated in 0.8% agarose gels [70]. The location and size of the corresponding insertions was analyzed by restriction digestion, PCR of the 5’ or the 3’ ends of the sacB gene and, in some cases, by DNA sequencing. PCR was carried out using primer pairs sacB3L1 (5’CCCCGTAGTCTAGGAAATCTT3’) - sacB5R1 (5’GGCGTAATGTTTACCGGAGA3’) and sacB3L1 (5’GTCGTAGTTTACCGGAGA3’) - sacB3R1 (5’GGATTTCCCCACATTTT3’) designed from the published sacB sequence (accession no. X02730); these primer pairs allow for the amplification of the complete sacB CDS (1422 nt) plus its promoter in two overlapping fragments.

A working transposition frequency was estimated using the formula: [(number of sucrose and tetracycline resistant cfu per millilitre×0.92)/total number of tetracycline resistant cfu per millilitre]; in this formula, we multiplied by 0.92 because an average of 91.5% of the sucrose-resistant (suc₆) clones among more than 500 analyzed (see Table 5, and data not shown) were found to originate by the insertion of a mobile element, whereas the remaining 8.5% was due to small deletions or individual nt changes in sacB. The frequency of transposition was estimated for five independent transformants of strain 1448A(pGEN500) grown overnight in KMB or MG, both supplemented with tetracycline; each experiment was repeated at least seven times. To estimate transposition from cells growing in planta, those five transformants were independently inoculated on leaves of bean (Phaseolus vulgaris L.) cv. Tendergreen, or fully expanded leaves of tobacco (Nicotiana tabacum cv. xanthi) as previously described [71], using plants maintained in growth chambers at 20°C, a photoperiod of 16/8 h day/night and 80% of humidity. Six hours after inoculation with bacterial suspensions adjusted to an OD₆₀₀ of 0.5 (approximately 5×10⁸ cfu ml⁻¹), excised leaf disks were ground in sterile 10 mM MgCl₂ and appropriate dilutions plated on NA and SNA, both supplemented with tetracycline. The experiment was repeated at least three times.

The analysis of transposition frequency data in different growing conditions was done using a two-way ANOVA test (p<0.05).
alignment was obtained with the genome of the *Gammaproteobacterium* HdN1, accession no. FP929140, positions 593503–593601.

**D** Terminal inverted repeats of the MITEPsy homolog present in *P. stutzeri* ATGC 17508 and the *Gammaproteobacterium* HdN1.

(DOC)

**Figure S2** Each transposition involves the movement of a single element. Southern hybridization of DNA digested with PstI using a complete copy of IS801 as a probe. Lanes contain genomic (Lane 1) or total plasmid DNA (Lane 7) from 1448A(pGEN500) or DNA isolated from clones containing independent insertions of mobile elements in gene *sab* of pGEN500 as follows: two independent MITEPsy1 insertions (Lanes 2 and 3); two independent insertions of IS801 (Lanes 4 and 5), and IS801 that recruited 1431 nt of adjacent DNA (Lane 6). Asterisks to the left of lanes indicate hybridization bands corresponding to the elements inserted in the vector. M, molecular weight marker (Kb Ladder, Agilent Technologies).

(PPT)

**Table S1** Coding sequences annotated in the genome of *P. syringae* pv. phaseolicola (Pph) 1448A that are chimeras with or that are interrupted by mobile elements.

(DOC)

**Table S2** List of primers used for the amplification of insertion sequences.

(DOC)

**Annotation File S1** Annotation of insertion sequences, and fragments thereof larger than 200 nt, and MITEs in the chromosome of *P. syringae* pv. phaseolicola 1448A (accession no. CP000058). The file is in the format of a feature table (tab file) to be read as an entry with the Artemis browser.

(TAB)

**Annotation File S2** Annotation of insertion sequences, and fragments thereof larger than 200 nt, and MITEs in plasmid p1448A-A of *P. syringae* pv. phaseolicola 1448A (accession no. CP000059). The file is in the format of a feature table (tab file) to be read as an entry with the Artemis browser.

(TAB)

**Annotation File S3** Annotation of insertion sequences, and fragments thereof larger than 200 nt, in plasmid p1448A-B of *P. syringae* pv. phaseolicola 1448A (accession no. CP000060). The file is in the format of a feature table (tab file) to be read as an entry with the Artemis browser.

(TAB)

**Acknowledgments**

We are very grateful to Magdalena Linderberg for her help with the annotation updates, to Yoshiyuki Ohtsubo for supplying us with pGEN500 and to Patricia Siguer (CNRS, Toulouse) for her help in the analysis of MITEPsy1 and the identification of its relationship to ISPsy2. We thank Frederic Bertels and Paul Rainey for sharing information before publication.

**Author Contributions**

Conceived and designed the experiments: LB MA RJ JM. Performed the experiments: LB MA AM-B NY-C. Analyzed the data: LB MA RJ JM. Contributed reagents/materials/analysis tools: JM. Wrote the paper: LB RJ JM.

**References**

1. Siguer P, File J, Chandler M (2006) Insertion sequences in prokaryotic genomes. Current Opinion in Microbiology 9: 526–531.
2. Mahillon J, Chandler M (1998) Insertion sequences. Microbiology and Molecular Biology Reviews 62: 725–774.
3. Jackson RW, Vinatzer B, Arnold DL, Dorus S, Murillo J (2011) The influence of the accessory genome on bacterial pathogen evolution. Mobile Genetic Elements 1: 55–65.
4. Craig NL, Craigie R, Gellert M, Lambowitz AM, eds (2002) Mobile DNA II.
5. McEvoy CRE, Falmer AA, van Putten NCC, VGC TG, van Helden PD, et al. (2007) The role of IS6110 in the evolution of *Mycobacterium tuberculosis*. Tuberculosis 87: 393–404.
6. Sundin GW, Momin J (2009) Gene traders: characteristics of native plasmids from plant pathogenic bacteria. In: Jackson RW, ed. Plant pathogenic bacteria: genomics and molecular biology. Cams: Caister Academic Press. pp 295–310.
7. Siguer P, Perochon J, Lestrade I, Mahillon J, Chandler M (2006) Sfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Research 34: D32–D36.
8. Bertels F, Rainey PB (2011) Within-genome evolution of REPINs: a new family of miniature mobile DNA in bacteria. PLoS Genetics 7: e1002132.
9. Delihas N (2008) Small mobile sequences in bacteria display diverse structure/function motifs. Molecular Microbiology 67: 475–481.
10. Mansfield JW (2009) From bacterial avirulence genes to effector functions via the avrB gene. Molecular Plant-Microbe Interactions 22: S5–S14.
11. Joardar V, Linderberg M, Jackson RW, Selengut J, Dodson R, et al. (2005) A transposable element from *Pseudomonas syringae* pv. *phaseolicola* 1448A reveals divergence among pathogens in genes involved in virulence and transposition. Journal of Bacteriology 187: 6488–6498.
12. Rodriguez-Palenzuela P, Matos I, Murillo J, Lopez-Solala E, Bardaji J, et al. (2010) Annotation and overview of the *Pseudomonas savastanoi* genes. NCBi: 3335 draft genome reveals the virulence gene complement of a tumour-inducing pathogen of woody hosts. Environmental Microbiology 12: 1604–1620.
13. Buell CR, Joardar V, Linderberg M, Selengut J, Paulsen IT, et al. (2003) The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. *tomato* DC3000. Proceedings of the National Academy of Sciences USA 100: 10181–10186.
14. Vivian A, Murillo J, Jackson RW (2001) The role of plasmids in phytopathogenic bacteria: mobile archaeoids? Microbiology 147: 763–780.
15. Feil H, Feil WS, Chain P, Larimer F, DiBartolo G, et al. (2005) Comparison of the complete genome sequences of *Pseudomonas syringae* pv. *phaseolicola* B728a and pv. *tomato* DC3000. Proceedings of the National Academy of Sciences USA 102: 11064–11069.
16. Sundin GW (2007) Genomic insights into the contribution of phytopathogenic bacterial plasmids to the evolutionary history of their hosts. Annual Review of Phytopathology 45: 129–151.
17. Yamaoka T, Lee PD, Kosuge T (1986) Insertion sequence elements of *Pseudomonas* and *Agrobacterium* plasmids. Plant Pathology 35: 1106–1109.
18. Romantschuk M, Richter GY, Mukhopadhyay P, Mills D (1991) IS100 of *Pseudomonas syringae* pv. *phaseolicola*. Molecular Plant Pathology 7: 355–364.
19. Landgraf A, Weingart H, Tsiamis G, Boch J, et al. (2006) Different versions of the ISPsy3 element in *Pseudomonas syringae* phaseolicola ATCC 17588 and the *Gammaproteobacterium* HdN1. Molecular Plant Pathology 7: 355–364.
29. Oquiña JA, Rico A, Rivas LA, Sotra I, Vivian A, et al. (2004) Pseudomonas syringae pv. phaseolicola can be separated into two genetic lineages distinguished by the possession of the phaseolotoxin biosynthetic cluster. Microbiology 150: 473–482.

30. Finhure ME, Navarro de la Fuente I, Rivas I, Hernández-Flores JL, Garcedua-Pita R, et al. (2008) Genetic relatedness among the different lineages of Pseudomonas syringae pv. phaseolicola. In: Fattou, Collmer A, Iacobellis NS, Mansfield JW, Murillo J, et al., eds. Pseudomonas syringae pathovars and related pathogens Identification, epidemiology and genomics. Dordrecht: Springer. pp 173–172.

31. Stevens G, Bennett MA, Athanasopoulou E, Tsiamis G, Taylor JD, et al. (1998) Sequence variations in alleles of the avirulence gene AvrPphE R2 from Pseudomonas syringae pv. phaseolicola lead to loss of recognition of the AvrPphE protein within bean cells and a gain in cultivar-specific virulence. Molecular Microbiology 29: 553–577.

32. Rivas LA, Mansfield J, Tsiamis G, Jackson RW, Murillo J (2005) Changes in race-specific virulence in Pseudomonas syringae pv. phaseolicola are associated with a chimeric transposable element and rare deletion events in a plasmid-born pathogenicity island. Applied and Environmental Microbiology 71: 3778–3785.

33. Szabo LJ, Mills D (1984) Integration and excision of pMC7105 in Pseudomonas syringae pv. phaseolicola: involvement of repetitive sequences. Journal of Bacteriology 157: 821–827.

34. García-Bárcia MP, de la Cruz F (2002) Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. FEMS Microbiology Ecology 42: 303–313.

35. García-Bárcia MP, Bernales I, Mendolia MV, de la Cruz F (2002) IS91 rolling-circle transposition. In: Craig NL, Craigie R, Gellert M, Lambowitz AM, eds. Mobile DNA II. Washington, D.C.: ASM Press. pp 891–904.

36. Richter GY, Björklof K, Romantschuk M, Mills D (1998) Insertion specificity. In: Drevinek P, Baldwin A, Lindenburg L, Joshi LT, Marchbank A, et al. (2010) Oxidative stress of bacteria. Journal of Bacteriology 164: 918–921.

37. Bardaji L, Pérez-Martínez I, Rodríguez-Moreno L, Rodríguez-Palenzuela P, Sánchez-Diezpérez A, et al. (2009) Sequence and role in virulence of the three plasmid complement of the model tumor-inducing bacterium Pseudomonas syringae pv. savastanoi NCPPB 3335. PLoS ONE, in press.

38. Chen Y, Zhou F, Li G, Xu Y (2009) MUST: A system for identification of miniature inverted-repeat transposable elements and applications to A. thaliana varities and Halosphaera violacea. Gene 436: 1–7.

39. Han Y, Wessler SR (2010) MITE-Hunter: a program for discovering miniature inverted-repeat transposable elements from genomic sequences. Nucleic Acids Research 38: 4259–4266.

40. Ohnabe Y, Grunke H, Komatsu H, Nagata Y, Tsuda M (2005) High-temperature-induced transposition of insertion elements in Burkholderia multivorans ATCC 17616. Applied and Environmental Microbiology 71: 1822–1828.

41. Gay P, Le Coq D, Steinmetz M, Belkeman T, Kado CI (1985) Positive selection procedure for the entrapment of insertion sequences elements in gram-negative bacteria. Journal of Bacteriology 164: 918–921.

42. Nagy Z, Chandler M (2004) Regulation of transposition in bacteria. Research in Microbiology 155: 387–398.

43. Drevinek P, Baldwin A, Lindenburg L, Joshi LT, Marchbank A, et al. (2010) Oxidative stress of Burkholderia cenocepacia induces insertion sequence-mediated genomic rearrangements that interfere with macrorestriction-based genotyping. Journal of Clinical Microbiology 48: 34–40.

44. Valle J, Vargara-Brugay M, Merino N, Penades JR, Lasa I (2007) eBE regulates IS256-mediated Staphylococcus aureus biofilm phenotypic variation. Journal of Bacteriology 189: 2086–2096.

45. Touchon M, Rocha EPC (2007) Causes of insertion sequences abundance in prokaryotic genomes. Molecular Biology and Evolution 24: 969–981.

46. Lindenberg M, Myers CR, Collmer A, Schneider DJ (2008) Roadmap to new virulence determinants in Pseudomonas syringae: insights from comparative genomics and genome organization. Molecular Plant-Microbe Interactions 21: 605–700.

47. Walden K, Rivera-Calzada A, Waksman G (2010) Type IV secretion systems: versatility and diversity in function. Cellular Microbiology 12: 1201–1212.

48. Norman A, Hansen LH, Sørensen SJ (2009) Conjugative plasmids: vessels of the communal gene pool. Philosophical Transactions of the Royal Society B: Biological Sciences 364: 2273–2289.

49. Mendolia MV, Bernales I, de la Cruz F (1994) Differential roles of the transposon terminus in IS91 transposition. Proceedings of the National Academy of Sciences USA 91: 1922–1926.

50. Toleman MA, Bennet PM, Walsh TR (2006) ISCR elements: novel gene-capturing systems of the 21st century? Microbiology and Molecular Biology Reviews 70: 296–316.

51. Stavrinides J, Ma W, Gutman DS (2006) Terminal reassortment drives the quantum evolution of type III effectors in bacterial pathogens. PLoS Pathogens 2: e1104.

52. Yuan S, Liu H, Mohr TJ, Jeufrette J, Chioudin R, et al. (2008) Role of recombination in the evolution of the model plant pathogen Pseudomonas syringae pv. tomato DC3000, a very atypical tomato strain. Applied and Environmental Microbiology 74: 3171–3181.

53. Burt A, Trivers R (2006) Genes in conflict: the biology of selfish genetic elements. CambridgeMassachusetts: Belknap Press of Harvard University Press.

54. Zhou F, Tran T, Xu Y (2008) Nezha, a novel active miniature inverted-repeat transposable element in cyanobacteria. Biochemical and Biophysical Research Communications 365: 790–794.

55. Robertson AE, Wechter WP, Denny TF, Fortman BA, Klupfel DL (2004) Relationship between avirulence gene (avr) diversity in Ralstonia solanacearum and bacterial virulence and incidence. Molecular Plant-Microbe Interactions 17: 1376–1384.

56. Poirel L, Carrer A, Pitout JD, Nordmann P (2009) Integron mobilization unit as a source of mobility of antibiotic resistance genes. Antimicrobial Agents and Chemotherapy 53: 2492–2498.

57. Stavrinides J, Gutman D (2004) Nucleotide sequence and evolution of the five-plasmid complement of the phytopathogen Pseudomonas syringae pv. maculicola ES4320. Journal of Bacteriology 186: 5101–5115.

58. Sambrook J, Frisch EF, Maniatis T (1989) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

59. Teverson DM (1993) Genetics of pathogenicity and resistance in the halo-blight disease of beans in Africa [Ph. D. Thesis]. Birmingham, UK: University of Birmingham.

60. Taylor JD, Teverson DM, Allen DJ, Pastor-Correllas MA (1996) Identification and origin of races of Pseudomonas syringae pv. phaseolicola from Africa and other bean growing areas. Plant Pathology 45: 469–478.

61. King EO, Ward NK, Kaney DE (1954) Two simplex media for the demonstration of pyocyanin and fluorescein. Journal of Laboratory and Clinical Medicine 44: 301–307.

62. Keane PJ, Kerr A, New, PB (1976) Crown gall of stone fruit. II. Identification and nomenclature of Agrobacterium isolates. Australian Journal of Biological Sciences 29: 585–595.

63. Cummings L, Riley L, Black L, Sourovar A, Resenuch S, et al. (2002) Genomic BLAST: custom-defined virtual databases for complete and unfinished genomes. FEMS Microbiology Letters 216: 133–138.

64. Corpet F (1988) Multiple sequence alignment with hierarchical clustering. Nucleic Acids Research 16: 10811–10890.

65. Tatsumo TA, Madder TL (1999) BLAST 2 SEQUENCES, a new tool for comparing protein and nucleotide sequences. FEMS Microbiology Letters 174: 247–250.

66. Rutherford K, Parkhill J, Crook J, Hornsell T, Rice P, et al. (2000) Artemis: sequence visualisation and annotation. Bioinformatics 16: 944–945.

67. Carver TJ, Rutherford KM, Berriman M, Rajandream MA-A, Barrett BG, et al. (2005) ACT: the Artemis Comparison Tool. Bioinformatics 21: 3422–3423.

68. Abbott JC, Aanensen DM, Rutherford K, Butcher S, Spratt BG (2005) WebACT: an online companion for the Artemis Comparison Tool. Bioinformatics 21: 3665–3666.

69. Michele L, Faure D, Blot M, Cabannes-Giulì E, Balanderoux J (2001) Detection and activity of insertion sequences in environmental strains of Burkholderia. Environmental Microbiology 3: 766–773.

70. Seoana E, Sundin GW, Murillo J (1998) Closely related replicons coexisting in closely related replicons coexisting in Burkholderia multivorans ATCC 17616. Applied and Environmental Microbiology 71: 1822–1828.

71. Abbott JC, Aanensen DM, Rutherford K, Butcher S, Spratt BG (2005) WebACT: an online companion for the Artemis Comparison Tool. Bioinformatics 21: 3422–3423.