Transposase interaction with the β sliding clamp: effects on insertion sequence proliferation and transposition rate

Héctor Díaz-Maldonado1, Manuel J. Gómez2,†, Mercedes Moreno-Paz1, Patxi San Martín-Uriz1, Ricardo Amils1,2, Víctor Parro1 & Francisco J. López de Saro1

Insertion sequences (ISs) are ubiquitous and abundant mobile genetic elements in prokaryotic genomes. ISs often encode only one protein, the transposase, which catalyzes their transposition. Recent studies have shown that transposases of many different IS families interact with the β sliding clamp, a DNA replication factor of the host. However, it was unclear to what extent this interaction limits or favors the ability of ISs to colonize a chromosome from a phylogenetically-distant organism, or if the strength of this interaction affects the transposition rate. Here we describe the proliferation of a member of the IS1634 family in *Acidiphilium* over ~600 generations of cultured growth. We demonstrate that the purified transposase binds to the β sliding clamp of *Acidiphilium*, *Leptospirillum* and *E. coli*. Further, we also demonstrate that the *Acidiphilium* IS1634 transposase binds to the archaeal sliding clamp (PCNA) from *Methanosarcina*, and that the transposase encoded by *Methanosarcina* IS1634 binds to *Acidiphilium* β. Finally, we demonstrate that increasing the strength of the interaction between β and transposase results in a higher transposition rate in vivo. Our results suggest that the interaction could determine the potential of ISs to be mobilized in bacterial populations and also their ability to proliferate within chromosomes.

Insertion sequences (ISs) are the simplest mobile genetic elements as they often contain just a single gene encoding the transposase required for transposition. ISs are ubiquitous in bacterial genomes, where they play a key role in genome evolution by promoting chromosomal rearrangements and gene inactivation and mobilization1,2. A characteristic feature of ISs is their ability to cross species barriers and transpose actively in new hosts, which makes ISs essential players in the process of horizontal gene transfer (HGT). IS-mediates HGT has been documented extensively among habitat-sharing organisms for Bacteria and Archaea3,4. ISs have often been observed flanking recently-incorporated foreign DNA in genomic islands5, or in smaller gene clusters in, for example, pathogens6, extremophiles7, or endosymbionts8. However, the invasion of a new host relies on the compatibility of the IS with the new molecular environment: the transposase gene has to be transcribed and translated, and any interactions of the transposase with host functions ought to be maintained9. Indeed, since most ISs do not confer a direct selective advantage to the host, IS persistence in populations implies that they must be constantly imported into genomes by HGT2,3,10. It could be predicted that an active IS has a short time-window to proliferate in a chromosome before inactivating mutations accumulate11–14. ISs have been classified into a number of families depending on their organization, terminal inverted repeats and mechanism of transposition1. While some IS families are widely distributed in Bacteria and Archaea, others seem restricted

1Centro de Astrobiología (INTA-CSIC), Ctra. de Ajalvir, Km. 4, Torrejón de Ardoz, 28850 Madrid, Spain. 2Centro de Biología Molecular Severo Ochoa (UAM-CSIC), 28049 Madrid, Spain. †Current address: Manuel J. Gómez, Bioinformatics Unit, Centro Nacional de Investigaciones Cardiovasculares (CNIC), 28029 Madrid, Spain. Correspondence and requests for materials should be addressed to F.J.L.S. (email: lopezsfj@cab.inta-csic.es)
to specific groups. Similarly, some IS families can proliferate explosively in a given bacterial lineage and account for a large fraction of the genome, while others are always found in one or few copies in the chromosome\textsuperscript{13}. IS abundance has been correlated with lifestyle or with episodes of rapid evolution, such as the colonization of new ecological niches\textsuperscript{16–18}.

Although ISs are highly autonomous, there are several examples of interaction with host enzymes and processes, especially chromosomal replication (for a review, see Ref. 1). The association of transposition to replication was first observed for IS10, which is activated by the passing of the replication fork\textsuperscript{19}. Recently it has been demonstrated that transposases of various IS families interact with the host replication machinery via the $\beta$ sliding clamp, an essential factor of DNA replication and repair\textsuperscript{20–22}. $\beta$ is a homodimeric protein that encircles DNA and serves as an anchoring platform for diverse enzymes. In addition to conferring processivity to DNA polymerases, $\beta$ targets enzymes to the chromosome and the replication fork. These enzymes, including the transposases, often contain a short $\beta$-binding motif at their C-terminal end, consisting of the consensus sequence Q-L-x-L-F\textsuperscript{21,23,24}. The fact that transposases with very different transposition pathways and mechanisms bind to $\beta$ suggests convergent evolution for a critical contact with the host\textsuperscript{21}. The ability of a transposase to bind to $\beta$ could possibly determine its fate during the invasion of a new host. If binding is weak, the transposase may not have a chance to compete with the many other host factors (or other transposases) that also bind to $\beta$, transposition could fail and the IS would become inactivated by mutation. Alternatively, if the transposase binds strongly to $\beta$, competition with host enzymes would likely interfere with essential host functions and be detrimental to the cell. Depending on the role of $\beta$ in transposition, a strong interaction could either favor or block transposase activity.

The goal of this study was to determine whether the interaction of transposases with $\beta$ is a possible requirement that limits IS movement across species, and whether the strength of the interaction could influence the frequency of transposition events. We analyzed the dynamics of IS populations during 600 generations in a recently-isolated acidophilic bacterium, Acidiphilium sp. PM, and identified those ISs that had changed in copy number during that period of time. We detected proliferation of a member of the IS1634 family, a commonly observed IS element in the genome of acidophiles. We demonstrated that the IS1634 transposase can physically interact with the Acidiphilium $\beta$ sliding clamp and with $\beta$ from other acidophiles. We then tested the ability of the Acidiphilium IS1634 family member to proliferate in E. coli, an organism in which this IS family has never been detected, and obtained direct evidence suggesting that the affinity of the transposase for $\beta$ could directly determine transposition rate. Finally, we also found that peptides containing the $\beta$ binding motif of the transposase encoded by the Acidiphilium IS1634 can bind to an archaeal processivity factor, PCNA, from Methanosarcina, and that peptides derived from a Methanosarcina IS1634 transposase can bind to the Acidiphilium $\beta$ sliding clamp.

Results

A member of the IS1634 family is active in Acidiphilium sp. PM, and its transposase contains a $\beta$-binding motif. The sequencing of the genome of Acidiphilium sp. PM, an acidophilic $\alpha$-proteobacterium isolated from the Tinto river in the south of Spain, revealed that up to 2.4% of its chromosome is composed of repeated regions\textsuperscript{25}. 85 ORFs related to insertion sequences (transposases and associated factors) were detected and classified by comparative analysis against the Pfam and ISfinder databases (See Ref. 21 for details). Although some of these ORFs corresponded to full-size transposases or associated proteins, many of them were likely inactivated remnants and gene fragments that are no longer capable of transposition. In order to identify active transposable elements, we performed a long-term cultivation of Acidiphilium sp. PM, and analyzed changes in IS abundance that had become fixed in the population at the end of the experiment. The sequenced strain was used to start a culture which was then grown for ~600 generations over four years (See Methods). To identify changes in the number of ISs in the population, we used an oligonucleotide microarray consisting of probes representing 1,358 genes associated to ISs, phage and plasmid mobility functions present in the genomes of selected acidophilic organisms, including Acidiphilium sp. PM (see Methods). Although this methodology would not detect transposition events that did not change IS copy number in the chromosome (i.e., conservative transposition) or chromosomal rearrangements, we expected that it would detect IS proliferation or loss.

DNA obtained from the founder strain (‘2007’) and from the 4-year-old population (‘2011’) were differentially labelled with fluorescent dyes Cy3 and Cy5 and hybridized together against the microarray (Fig. 1a). We observed changes $\log_2(2011/2007)>0.5$ in nine elements (Fig. 1b and Table S1). Changes implied the apparent loss of the IS ORF in six cases (IS5-4, IS110-4, IS110-12, IS21-1, IS3-11, IS5a-2), and a small gain in copy number in three cases (IS1634-1, IS5b-8 and IS66-4).

A member of the IS1634 family, IS1634-1, present in single copy in the ‘2007’ culture, showed an increase to three copies in the ‘2011’ culture, and we focused on this element for the rest of the study. IS1634-1 in the sequenced ‘2007’ culture (Fig. S1, Accession number: WP_007423974) could be a recent insertion in the chromosome, as suggested by its anomalous %GC and GC skew (Fig. S2). The copy number increase was confirmed by qPCR, inverse PCR, and sequencing (Fig. 1c,d). Our analysis of the insertion sites of the new copies (IS1634-2 and IS1634-3) revealed the characteristic inverted repeats generated by true transposition events, but did not detect obvious adaptive benefits for the evolving culture of Acidiphilium.
The IS1634 family is a poorly characterized family of insertion sequences related to IS4, ISH3 and IS701\textsuperscript{20}. A similarity tree for the transposases encoded by IS1634 members (Fig. 2) shows that this IS family can be found in Euryarchaea and in numerous groups of Bacteria. Interestingly, IS1634 appears in organisms that share acidic, heavy-metal rich environments, and is present, among others, in the chromosomes of most species of *Leptospirillum* (Nitrospira), *Acidithiobacillus* (\(\gamma\)-proteobacteria) and *Acidiphilium* (\(\alpha\)-proteobacteria), as well as in *Paracoccus ferrooxidans* (\(\alpha\)-proteobacteria), *Ferroplasma acidarmanus* (Euryarchaeae), *Alicyclobacillus acidocaldarius* (Firmicutes), *Thiimonas* sp. (\(\beta\)-proteobacteria), and *Desulfovibacca acetoxidans* (\(\delta\)-proteobacteria) (Fig. 2). IS-sharing has been observed before for co-habiting organisms\textsuperscript{3}.

The interaction of transposases with the host \(\beta\) sliding clamp occurs almost always at the C-terminus of the protein via a short sequence with a weak consensus sequence (Q\(_1\)-L\(_2\)-x-L\(_4\)-F\(_5\))\textsuperscript{21}. A search of the \(\beta\)-binding motif among IS1634 transposases shows that a putative motif can be found at the C-terminus of these proteins (Fig. 2). Despite high variation in sequence context, the alignment of this region shows conservation of the Gln at position 1 of the putative \(\beta\)-binding motif plus conservation of hydrophobic or aromatic residues at positions 4 and 5, suggesting a possible interaction of IS1634 transposases with the host \(\beta\) sliding clamp via this region.
The IS1634 transposase from Acidiphilium can bind \( \beta \) from diverse organisms. To determine whether Acidiphilium IS1634 transposase (Tnp) could bind Acidiphilium \( \beta \), we purified both proteins from overexpressing E. coli strains (See Methods) and probed their interaction biochemically. We purified the full-size, wild-type transposase and, for comparative purposes, two engineered mutant proteins, one in which the Q1 and the F5 of the \( \beta \)-binding motif had been mutated to Alanine (named 5A), and one containing a full consensus sequence (QLSLF, named CN). Also, in addition to Acidiphilium \( \beta \) protein (Ac\( \beta \)), we purified \( \beta \) from Leptospirillum ferrooxidans (Lf\( \beta \)) and Escherichia coli (Ec\( \beta \)) (See Methods). We chose Leptospirillum (phylum Nitrospira) because, although distant from Acidiphilium phylogenetically, it shares the same acidic habitats, where it is usually the most common organism. E. coli was chosen because its sliding clamp and its ligands have been well described, facilitating the realization of comparative and genetic experiments as described below, and because no IS1634 member has been yet detected in Enterobacteria. For each transposase the sequence of the C-terminus of the protein is presented and residues involved in the putative \( \beta \)-binding motif are aligned and in bold. Underlined sequences are those of peptides used for biochemical analysis in this work.

Figure 2. Unrooted similarity tree of IS1634 transposases and alignment of C-terminal region. IS1634 transposase sequences were obtained from a previous study\textsuperscript{21}, in which 2,216 genomes downloaded from the National Centre for Biotechnology Information (NCBI) Genome database on October 2012, were systematically scanned to identify and classify ISs. The collection of IS1634 transposase sequences was made non-redundant at a sequence similarity level of 90% and by choosing one representative sequence per genus. The tree was computed with PhyML using a JTT model and a bootstrap of 500 replicates, and the results visualized with Seaview\textsuperscript{61}. Bootstrap values are shown for the main branches. The analysis reveals five distinct groups of IS1634 transposases. The code preceding each organism name is as follows: aP, bP, gP and dP stand for \( \alpha \), \( \beta \), \( \gamma \) and \( \delta \) Proteobacteria, respectively; F, Firmicutes; Te, Tenericutes; N, Nitrospira; C, Cyanobacteria; T, Thermodesulfolobus; D, Deinococcus-Thermus; A, Actinobacteria; EA, Euryarchaeota. For each transposase the sequence of the C-terminus of the protein is presented and residues involved in the putative \( \beta \)-binding motif are aligned and in bold. Underlined sequences are those of peptides used for biochemical analysis in this work.
Since experiments in Fig. 3a,b are mostly qualitative, we used a competition assay to analyze the relative strength of the interaction between WT and CN peptides and Ec\(β\). This assay makes use of the gel-shift generated by DNA polymerase IV (LF domain, PolIV LF) bound to \(β\) in native gel electrophoresis. These two proteins bind strongly and their interaction has been described in atomic detail. PolIV LF binds to Ec\(β\) at the canonical hydrophobic pocket on the C-side of the ring which is the binding site for all other proteins studied to date. Therefore a displacement of this complex by an excess of peptide, observable in the native gel by a change in the mobility of \(β\), would indicate that the peptide binds strongly to the same pocket on \(β\). A titration of the WT and CN peptides on the preformed complex (See Methods for details) indicate that the CN peptide binds to Ec\(β\) with higher affinity than the WT peptide, as expected (Fig. 3c,d).

The alignment of Fig. 2 showed that in some species (Aromatoleum, Thiomonas, Desulfovibrio, and Acidithiobacillus) a potential second motif can be found at the C-terminal end of some IS1634 family transposases. To determine whether these sequences could bind to \(β\), we synthesized a peptide containing the C-terminal sequence of Acidiphilium IS1634 Tnp. As shown in Fig. 4a, this peptide (AfTnp-C WT) can bind and retain Acidiphilium, Leptospirillum and E. coli \(β\), while a double mutant in which the Q and L in first and fourth positions of the motif were mutated could not...
Further, AfTnp-C WT can displace PolIV LF from a complex with Ecβ, revealing that the Acidithiobacillus peptide binds to the same region of Ecβ as other peptides (Fig. 4b). Finally, we have also detected a sliding clamp-binding motif in the transposases of two IS1634 members from Euryarchaea (Methanosarcina and Methanosaeta) compatible with the consensus PCNA-binding motif (Q-x-x-I-x-x-F-F)²⁸. PCNA (Proliferating Cell Nuclear Antigen) is the Archaeal and Eukaryotic sliding clamp, functionally homologous to bacterial β²⁹. We purified Methanosarcina barkeri PCNA and demonstrated binding to a peptide derived from the IS1634 transposase present in this organism and containing the putative PCNA binding motif (Fig. 4c). Given the relatedness and overlap between the β and PCNA binding motifs, we also tested whether interactions could be maintained across phylogenetic domains. Certainly, we could observe the interaction between the peptide derived from the Methanosarcina transposase and Acidiphilium β (Fig. 4c). Then, we tested whether peptides derived from Acidiphilium Tnp could interact with Methanosarcina PCNA. As shown in Fig. 4d, these proteins interact, showing that the interaction could potentially be maintained, bidirectionally, across the Bacteria-Archaea boundary.

A stronger β-binding motif increases transposition in vivo. Although the β-binding motif is universal, the established consensus sequence (QLxLF) was experimentally validated using Escherichia coli β²³,²⁶,³¹. A comparison of several β-binding motifs found in E. coli and Acidiphilium DNA replication
and DNA repair enzymes shows that those of E. coli better conform to the consensus (Fig. S4). On the other hand, the identity between Acidiphilium and E. coli β is only 35.4%. Although IS1634 Tnp binds E. coli β (Fig. 3 and crosslinking, Fig. S3), especially the variant containing the β-binding consensus (CN), it was uncertain whether this higher affinity (Fig. 3c,d) would result in more efficient transposition. In order to determine whether the interaction detected in vitro with β has an effect in the ability of transposases to be functional in other organisms, we performed an in vivo transposition assay for the three variants of the Acidiphilium IS1634 transposase in E. coli. We used a recently-developed vector that generates genomic insertions of the lacZ gene flanked by the IS inverted repeats32. The Acidiphilium IS1634 transposase gene was placed under the transcriptional control of the BAD promoter, allowing for the modulation of its expression by addition of arabinose (Fig. 5a). We transformed E. coli DH5α cells with the three plasmid variants containing wild-type Tnp and the mutants 5A and CN, plus a control plasmid with no transposase cloned, and incubated for 15 days at 30°C. We observed papillae reflecting transposition events in WT colonies, although clearly occurring at a low frequency in our experimental conditions. Mutation 5A showed lower generation of papillae than WT but further experimentation will be required to investigate whether this mutation decreases or eliminates transposition. However, cells harbouring the plasmid with the CN mutation clearly had a significant increase of transposition with respect to WT, generating ~4–6-fold more papillae (Fig. 5b,c). Our results therefore suggest that stronger binding of the transposase to β increases the frequency of transposition events in the cell.

Discussion

The objective of this study was to identify an active insertion sequence from a natural environment and explore its interaction with the host β sliding clamp regarding its ability to proliferate within the chromosome and its potential for dispersal to other species. We searched for changes in IS abundance in the genome of Acidiphilium sp PM that had become fixed in the population during 4 years of laboratory cultivation, and identified one IS that had increased its copy number during this period. We demonstrated that the transposase encoded by this IS interacts with β and mapped the interaction motif to the C-terminus of the protein. Next, we showed that the transposase could bind to β from various species, some of them distant phylogenetically, and identified additional examples of how the interaction between transposases and processivity factors (β and PCNA) can occur even across large phylogenetic distances. Finally, we demonstrated that optimizing the binding site on the transposase to fit the consensus β-binding motif increases transposition in a new host.

Our 600-generation laboratory evolution experiment with Acidiphilium sp. PM, aimed at detecting changes in IS copy number that became fixed in the culture by comparative microarray hybridization, found three cases of proliferation and six deletions. Other IS transposition events could have occurred and remained undetected because they were detrimental to the host and were selected against, because they were lost by genetic drift, or because they involved IS relocation in the chromosome (in the case of cut and paste transposition). As it is the case with other studies that have analyzed spontaneous global transposition activity in genomes, our study suggests that only a small fraction of the ISs detected by sequencing are active33. The detectable transposition activity can vary greatly with IS element and strain (from 10⁻³ to 10⁻⁷ per generation, according to some estimates)34–36. Unlike the well-defined and relatively predictable point mutation rates resulting from the combined action of DNA polymerases and repair genes, transposition rates are likely the result of the combination of diverse factors such as IS sequence variation, transposase expression and activity, chromosomal location effects, or various host regulatory mechanisms. Additionally, it is likely that the process of isolation and adaptation to laboratory conditions change transposition rates. For example, the observed transposition events in Ferroplasma were substantially more frequent in culture than in environmental samples of the same organism37. Although transposase gene expression in laboratory conditions has not been explored, metatranscriptomic and metaproteomic analysis of bacterial populations have also detected relatively quick changes in transposase gene expression in response to environmental stimuli38,39, although the underlying mechanisms are unknown.

The generally low rates of transposition observed in nature suggest that they are determined by regulatory mechanisms that are intrinsic to IS elements, perhaps as a strategy to limit any negative effects on the host genome4. However, the interaction of ISs with the essential replication and repair factor β, first demonstrated for transposon Tn7,12, has now been generalized to nine IS families with very different transposition pathways21. This extreme case of functional evolutionary convergence for a highly conserved and universal element suggests that β binding could be an integral aspect of transposition regulation. The β binding motif is present in a large number of proteins involved in DNA synthesis and repair. Since β is limiting in the cell14, the concentration of transposases and the strength of their interaction could determine the efficiency of transposition. All proteins so far studied in detail interact with β on the same face of the ring, binding competitively to the same conserved ‘hydrophobic pocket’. Using peptides, it was shown that different types of transposases also bind competitively to the same site on β2 and we show here that the same principle applies to the transposase of several members of the IS1634 family. The affinity with which enzymes bind β is likely to be finely-tuned in the cell to accommodate the various processes for which β is an essential component24,40. Since β-binding by a transposase could be a burden on the host, it would be unlikely to find strong motifs in transposases. Also, the relatively simplicity of
Figure 5. In vivo assay to study the effect of different $\beta$ binding motifs in IS1634 transposition. (a) Plasmid design of vector pSKT1-IS1634 (See Methods). The sequences for the left (IR-L) or right (IR-R) inverted repeats is shown. The sequences of the wild type and mutants 5A and CN is shown (amino acids 562–566 of IS1634 Tnp). (b) Papillation assay of transposition. The pictures show representative examples of E. coli colonies for the three versions of the transposase (WT and 5A or CN mutants), and a negative control C (pSKT1 containing the inverted repeats but no transposase gene). (c) Quantification of the papillation assay for 8 colonies (C = 1.51 ± 0.6; WT = 3.15 ± 0.5; 5A = 2.05 ± 0.85; CN = 13.0 ± 2.95).
the binding motif, which often resides in unstructured sequences, could be created de novo by random mutation and selection. This could explain the appearance of motifs at distinct locations in homologous transposases, as it is the case here with the secondary motif identified at the C-terminal sequence of Acidithiobacillus ferrivorans IS1634 (Fig. 4a). Likewise, the fact that the motif appears consistently in non-homologous transposases with very different transposition pathways seems to favour the idea that it has appeared independently many times in the evolutionary history of this diverse group of proteins.

Our results show that interaction with β should not be a strong barrier for IS exchange among phylogenetically-distant but habitat-sharing organisms. Because ISs are subject to inactivation by mutation, horizontal transfer is an essential aspect of their lifestyle and critical for the persistence of ISs in bacterial populations. Acidiphilium IS1634 Tnp can bind to Leptospirillum or Escherichia β, and even to an archaeal sliding clamp, PCNA from Methanosarcina. Although β and PCNA share no detectable sequence similarity, these proteins are structurally and functionally very similar, and the binding motifs of the proteins interacting with them are highly related. Our results suggest that transposases could interact relatively easily and bidirectionally with the replication machinery of bacterial or archaeal hosts. With a few exceptions (e.g., ISH6, found only in Archaea), most IS families can be found in bacterial and archaeal genomes, and genome sequencing suggest that movement of these mobile elements between the two domains is fluid. However, a recent survey of prokaryotic IS elements in eukaryotic genomes detected few events of IS transfer in recent eukaryotic evolutionary history, with the possible exception of cyanobacterial IS607. Although HGT events from prokaryotic to eukaryotes have been documented extensively, it remains to be investigated if mobile elements are involved. In eukaryotic organisms, PCNA-binding has been detected for Pogo, a Drosophila transposase and, recently, for the endonuclease/reverse transcriptase of LINE-1, where it has been shown to be critical for retrotransposition.

By interacting with β, transposases are targeted to the replication fork, thus associating chromosomal replication with transposition. Coupling of transposition to replication would ensure, first, the possibility of recombination and repair with the sister chromosome, which would limit the potential damage to the host caused by IS excision and the generation of double-strand breaks. Second, IS elements with a conservative (cut and paste) mechanism of transposition would have a chance of jumping to the sister chromosome and therefore increase their number in that replication. Third, transposition would co-localize with repair factors required in the later stages of transposition to fill in the gaps and ligate (e.g., both DNA polymerase I and DNA ligase interact with β). Yet another possibility is that transposases couple β-binding with allosteric changes that initiate transposition, such as monomer-dimer transitions or binding to the end sequences of the IS. Transposases could require β to be released from a state of self-inhibition which has often been observed in these enzymes. Formation of the transpososome involves dimerization and conformational changes which often involve C-terminal regions of the protein, which are, almost always, also the sites of interaction with β.

Although transposase binding to β assures integration with the host's chromosomal replication, it remains to be investigated if an excess of transposase, or a transposase with strong affinity for β, would be disruptive for DNA synthesis. It is remarkable that all enzymes involved in mutagenic processes, namely the DNA polymerases, the mismatch repair system, and transposition, directly interact with the β sliding clamp. Indeed, a recent experimental study demonstrated a direct 'conflict' between these processes, as mismatch repair mutator alleles, often present in bacterial populations, limit insertion sequence proliferation in the early stages of invasion in a new host.

Further, different IS families could directly compete among them for access to a limiting host resource, the DNA polymerase, and that can be explained just with two parameters: the duplication-deletion ratio and the LGT-deletion ratio. We propose that the ability of transposases to interact with β is one of the molecular processes that contribute to the “duplication” component of the first parameter and, as such, it is subject to natural selection. Variations in the binding site that increase the affinity for β may be selected upon colonization of new hosts and cause explosive proliferation of IS. On the contrary, variations that decrease the affinity for β would reduce transposition rates, contributing to a reduction in IS populations. Future studies will be required to understand how these various processes contribute to modulate genetic variability and to the adaptive potential of bacterial populations.

The in vivo assay reported here (Fig. 5) shows that transferring IS1634 from Acidiphilium sp. PM to Escherichia coli is greatly favoured by point mutations in the motif that make the interaction stronger (Fig. 3c,d). Our IS1634 transposase CN mutant defines a new category of hyperactive transposases: one with improved interface with the host. The search for hyperactive transposases has been based typically in genetic screens of random point mutants which are then combined in a single transposase to greatly increase its activity. For example, commercial Tn5 transposase contains three amino acid changes which increase activity by >1000-fold, PiggyBac mutations in seven amino acids generate a 17-fold increase in activity, and a combination of multiple mutations in Sleeping Beauty can generate up to a 100-fold increase in transposition. More recent studies have shown that mariner transposases could be easily made hyperactive by mutation of the dimerization interface, thus disrupting the autoregulatory allosteric mechanism between the subunits that limits transposition. The principles derived from our designed hyperactive IS1634 mutant could potentially open the door for development of novel hyperactive transposases for use in biotechnology, or for the design of transposases with an expanded range of hosts.
Methods

Bacterial strains and long-term cultivation. Acidiphilium sp. PM (DSM 24941) was originally isolated from the Tinto river in Huelva (Spain) and its genome sequenced. The sequenced strain was used to initiate a 100 ml culture in GYE medium, which is composed of a mineral salt solution (0.2% (NH₄)₂SO₄, 0.01% KCl, 0.03% KH₂PO₄·H₂O, 0.025% MgSO₄·7H₂O, 0.0014% Ca(NO₃)₂·4H₂O) supplemented with 0.2% (v/v) glucose and 0.01% (w/v) yeast extract. The pH was adjusted to 2.5 with 1 N H₂SO₄ prior to autoclaving (111 °C, 0.5 atm, 30 min). Cultivation took place at 30 °C with vigorous shaking. Serial transfers (1:50 dilution) took place in periods of 14 days. Under these conditions the cultures reached stationary phase in 5 days and the culture experienced 9-day–periods in nutrient-deprived medium. The 1:50 dilution allowed ~5.6 generations (log 250) per serial dilution, or 584 generations in 4 years. Samples from populations were taken periodically and stored at −80 °C.

Oligonucleotide microarray of IS-related genes. To design an oligonucleotide microarray that could detect changes in copy number and gene expression of mobile elements in acidophilic organisms, we selected 12 fully sequenced acidophilic genomes (including Acidiphilium sp. PM) (manuscript in preparation). We then used RPS Blast (NCBI) to scan 40,231 proteins identified in these genomes with a set of 149 position-specific scoring matrix (PSSM) profiles, selected from Pfam and corresponding to mobile element proteins. This procedure identified a set of 1,358 transposases and associated proteins. The corresponding gene sequences were progressively clustered by means of a customized pipeline that used UCLUST and Cons to generate a set of 769 consensus sequences characterized by having at least two conserved blocks of minimal length equal to 50 bp and zero ambiguities. The program Array Designer 4 (Premier Biosoft International) was then used to design two oligonucleotides for each consensus sequence (1,538 oligonucleotides) with an average length of 40 nt and a constant estimated melting temperature (72°C). See Table S1 for sequences corresponding to Acidiphilium IS1634 transposase gene; and oligo #3 and #4 for the dnaX gene (See Table S2 for oligonucleotide sequences). PCR reactions (25 μL) were performed with 0.4 μM primers (each), Acidiphilium sp. PM ‘2011’ target genomic DNA at two different final concentrations (40 pg/μL and 8 pg/μL) and 12.5 μL of IQ SYBR Green Supermix (Life Technologies) according to the manufacturer’s instructions. Two biological and three technical replicates were prepared for each gene and each DNA concentration. Reactions were carried out with a MyiQ Single-Colour Real-Time PCR Detection System (Bio-Rad). Cycling parameters comprised an initial cycle of 3 minutes at 95°C followed by 40 cycles of 30 seconds at 95°C, 30 seconds at 60°C and 30 seconds at 72°C. A melting curve analysis was performed for each reaction to rule out non specific reaction products or primer dimers. ‘2007’ genomic DNA was serially diluted in 5-fold increments from 5 ng to 1, 6 pg. These serial dilutions with three technical replicates, were used to create a standard curve for the reference dnaX gene (r = 0.992) and another curve for IS1634 transposase gene (r = 0.997). The genomic copy numbers of Acidiphilium IS1634 and dnaX in ‘2011’ time-point was determined from their corresponding ‘2007’ standard curves where both genes are present in a single copy per chromosome, using the comparative threshold cycle method (CT). Results in Fig. 1D represent the average of ‘2011’ dnaX and Acidiphilium IS1634 chromosomal gene copy number and the standard deviation (SD) of the replicates.

Inverse PCR. Inverse PCR was used to amplify the flanking regions of the Acidiphilium IS1634 insertion sites detected in the ‘2011’ culture. For this purpose 1 μg of ‘2007’ and ‘2011’ Acidiphilium sp. genomic DNA, extracted as previously described, were digested with EcoRI (NEB), an enzyme that does not cut within the transposase sequence, for 2 h at 37°C in 20 μL. Digestions were ligated in 200 μL reaction volume with T4 ligase (NEB) according to the manufacturer’s recommendation. A PCR was performed using 1 μL of ligation products and 1 μM of each divergent primers #5 and #6 (See Table S2). Inverse PCR reaction products of ‘2007’ and ‘2011’ time-points were resolved and visualized in a 1% agarose gel. DNA fragments were extracted of the gel with QIAquick Gel Extraction Kit (Qiagen) and sequenced.

Protein purifications. E. coli β, Methanosarcina barkeri PCNA and GST-PoIVL were cloned, overexpressed and purified as described. Acidiphilium sp. PM β was amplified by PCR using oligonucleotides
β Reactions were stopped with 1% SDS, subjected to SDS-PAGE electrophoresis and of 0.8 mg mL⁻¹ of them containing transposase were dialyzed against 2 L of Buffer J. Tnp was obtained at concentrations
utes at 25 °C in 5 mL of Buffer J
incubated (15 m, 25 °C) with the Tnp-coated beads, and washed three times with Binding Buffer A to
Dynabeads M-280 Tosyl-activated (Invitrogen). The reactions (50 μL) contained 1.2 mg of magnetic beads in Binding Buffer A (phosphate-buffered saline, 0.1% Tween-20, mercaptoethanol, 0.1% BSA) with 1 mg of Dynabeads M-270 Streptavidin (Invitrogen), incubated (30 m, 25 °C) and then centrifuged (15,000 g, 30 m, 4 °C). The inclusion body pellet was then washed three times with 20 mL of buffer G to remove residual Triton X-100. In a subsequent step, inclusion bodies were solubilized in 5 mL of buffer H (100 mM Tris-HCl, 100 mM NaCl, 2 mM EDTA, 10% glycerine, 1 mM β-mercaptoethanol, 1 mM PMSF, pH 8) and centrifuged (30 m, 15,000 g, 4 °C). Insoluble material was removed by centrifugation 15 minutes at 15000 g. The solubilised inclusion bodies were slowly drop by drop diluted in 400 mL of buffer H with constant stirring at 4 °C and allowed to refold for 1 hour. The solution was centrifuged 20 minutes at 15000 g to remove insoluble material. This clarified solution was applied on a 30 mL Q sepharose FF (GE Healthcare) ion-exchange chromatography column equilibrated in Buffer H. Protein was eluted with Buffer H over a NaCl gradient (120–L, 0.1–1.0 M). Fractions containing β were pooled and dialyzed against 2 L of Buffer I (50 mM Tris–HCl, 50 mM NaCl, 1 mM EDTA, 10% glycerine, 1 mM DTT, pH 8). This fraction (1.3 mg mL⁻¹) was aliquoted and stored at −80°C.

To amplify the β gene from Leptospirillum ferrooxidans by PCR we used oligos #9 and #10 (see Table S2). These oligos introduced a sequence encoding a FLAG epitope at the N-terminus of the gene. The PCR product was cloned as a BglII/NheI fragment into pET11a (Novagen). L. ferrooxidans β formed inclusion bodies and was purified following the same protocol as for Acidiphilium sp. PM β. The final yield was 1.4 mg mL⁻¹.

The gene for IS1634 transposase from Acidiphilium sp PM (Tnp) was amplified by PCR from genomic DNA with oligonucleotides #11 and #12 (see Table S2), cloned in vector pET16b, and sequenced. The 5A and CN Tnp mutants were created by site-directed mutagenesis of this plasmid genomic DNA with oligonucleotides #13 and #14 (Table S2); cloning of the transposase as a NcoI/EcoRI PCR product were #11 and #17 (Table S2).

Acidiphilium IS1634 Tnp pull-down assay. The binding assay to test the interaction of IS1634 Tnp (WT, 5A and CN) with β clamp or PCNA. 400 μM of each β labeled with Alexa Fluor 350 C5-maleimide (Life Technologies) were incubated (25 °C, 30 m) with 12 μM of Tnp covalently coupled to 1.2 mg of magnetic beads in Binding Buffer A (phosphate-buffered saline, 0.1% Tween-20, 0.1% BSA, pH 7.2). 5 μM of each β labeled with Fluor 350 C5-maleimide (Life Technologies) were incubated (15 m, 25 °C) with the Tnp-coated beads, and washed three times with Binding Buffer A to remove unbound β. Reactions were stopped with 1% SDS, subjected to SDS-PAGE electrophoresis and visualized on a UV transilluminator.

IS1634 transposase peptides pull-down assay. Peptides were obtained from ProteoGenix SAS (Schiltigheim, France). IS1634 transposase derived peptides from Acidiphilium sp. PM, Acidithiobacillus ferrooxidans and Methanosarcina barkeri were assayed for the interaction with β clamp or PCNA. 400 μM of biotinylated peptides were mixed in 50 μL in Binding Buffer B (50 mM Tris, 50 mM NaCl, 5% glycerine, 0.1% BSA) with 1 mg of Dynabeads M-270 Streptavidin (Invitrogen), incubated (30 m, 25 °C) and washed three times with the same buffer. 4 μM of labeled Ecoβ, Lfβ, Acβ or 3 μM labeled MbaPCNA were added to the beads in a reaction volume of 50 μL with Binding Buffer B, incubated (15 m, 25 °C), and washed three times with the same buffer. Reactions were stopped with 1% SDS, loaded on a SDS-PAGE and analyzed on a UV transilluminator.

Cloning of IS1634 into pSKT1. The pSKT1 plasmid was generously provided by H. Savilahti (University of Turku, Finland) and used as described (32). Oligos for cloning of right inverted repeat (RIR) of IS1634 were #13 and #14 (Table S2); cloning of the left inverted repeat (LIR) were #15 and #16 (Table S2); and cloning of the transposase as a NcoI/EcoRI PCR product were #11 and #17 (Table S2).
References

1. Siguer, P., Gourbeyre, E. & Chandler, M. Bacterial insertion sequences: their genomic impact and diversity. FEMS Microbiol. Rev. 38, 863–891 (2014).
2. Cerveau, N., Leclercq, S., Bouchon, D. & Cordaux, R. in Evolutionary Biology – Concepts, Biodiversity, Macroevolution and Genome Evolution (ed. Pontarotti, P.) 291–312 (Springer-Verlag, Berlin Heidelberg, 2011).
3. Hooper, S. D., Mavromatis, K. & Kyrpides, N. C. Microbial co-habitation and lateral gene transfer: what transposases can tell us. Genome Biol. 10, R45 (2009).
4. Toussaint, A. & Chandler, M. in Bacterial Molecular Networks: Methods and Protocols (ed. van Helden, J.) (Springer Science, 2012).
5. Juhás, M. et al. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol. Rev. 33, 376–93 (2009).
6. Cunneen, M. M. & Reeves, P. R. The Yersinia kristensenii O11 O-antigen gene cluster was acquired by lateral gene transfer and incorporated at a novel chromosomal locus. Mol. Biol. Cell. 24, 1355–1365 (2007).
7. Dillenburg, J. et al. Evidence of recent lateral gene transfer among hyperthermophilic Archaea. Mol. Microbiol. 38, 684–693 (2000).
8. Degnan, P. H., Yuh, Y., Sisneros, N., Wing, R. A. & Moran, N. A. Hamiltonella defensa, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. Proc. Natl. Acad. Sci. USA 106, 9063–9068 (2009).
9. Nagy, Z. & Chandler, M. Regulation of transposition in bacteria. Res. Microbiol. 155, 387–98 (2004).
10. Iranzo, J., Gomez, M. J., Lopez de Saro, F. J. & Mannubia, S. Large-scale genomic analysis suggests a neutral punctuated dynamics of transposable elements in bacterial genomes. PLoS Comput. Biol. 10, e1003680 (2014).
11. Le Rouzic, A. & Capy, P. The first steps of transposable elements invasion: parasitic strategy vs. genetic drift. Genetics 169, 1033–1043 (2005).
12. Bichsel, M., Barbour, A. D. & Wagner, A. The early phase of a bacterial insertion sequence infection. Theor. Popul. Biol. 78, 278–88 (2010).
13. Lohse, A. R., Moriyama, E. N., Lidholm, D. & Hartl, D. L. Horizontal transmission, vertical inactivation, and stochastic loss of mariner-like transposable elements. Mol. Biol. Evol. 12, 62–72 (1995).
14. Feher, T. et al. Competition between transposable elements and mutator genes in bacteria. Mol. Biol. Evol. 29, 3135–3159 (2012).
15. Treangen, T. J., Abraham, A. L., Touchon, M. & Rocha, E. P. Genesis, effects and fates of repeats in prokaryotic genomes. FEMS Microbiol. Rev. 33, 539–71 (2009).
16. Newton, I. L. G. & Bordenstein, S. R. Correlations between bacterial ecology and mobile DNA. Curr. Microbiol. 62, 198–208 (2010).
17. Moran, N. A. & Plague, G. R. Genomic changes following host restriction in bacteria. Curr. Opin. Genet. Dev. 14, 627–33 (2004).
18. Parkhill, J. et al. Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiiseptica. Nat. Genet. 35, 32–40 (2003).
19. Roberts, D., Hoopes, B. C., McClure, W. F. & Kleckner, N. IS10 transposition Is regulated by DNA adenine methylation. Cell 43, 117–130 (1985).
20. Chandler, M. Clamping down on transposon targeting. Cell 138, 621–3 (2009).
21. Gómez, M. J., Díaz-Maldonado, H., González-Tortuero, E. & López de Saro, F. J. Chromosomal replication dynamics and interaction with the β sliding clamp determine orientation of bacterial transposable elements Genome Biol. Evol. 6, 727–740 (2014).
22. Parks, A. R. et al. Transposition into replicating DNA occurs through interaction with the processivity factor. Cell 138, 685–95 (2009).
23. Dalrymple, B. P., Kongswan, K., Wijffels, G., Dixon, N. E. & Jennings, P. A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. Proc. Natl. Acad. Sci. USA 98, 11627–11632 (2001).
24. López de Saro, F. J. Regulation of interactions with sliding clamps during DNA replication and repair. Current Genomics 10, 206–215 (2009).
25. San Martin Úrtiz, P., Gómez, M. J., Arcas, A., Bargiela, R. & Amils, R. Draft genome sequence of the electrician Acidiphilium sp. strain PM (DSM 24941). J. Bacteriol. 193, 5585–5586 (2011).
26. De Palmenaer, D., Siguer, P. & Mahillon, J. IS4 family gives genomic. BMC Evol. Biol. 8, 18 (2008).
27. Bunting, K. A., Roe, S. M. & Pearl, L. H. Structural basis for recruitment of translesion DNA polymerase Pol IV/DinB to the β-clamp. EMBO J. 22, 5883–5892 (2003).
28. Warbrick, E. The puzzle of PCNA ‘s many partners. Cold Spring Harb. Perspect. Biol. 5, a010165 (2013).
29. Georgescu, R. E. et al. Structure of a small-molecule inhibitor of a DNA polymerase sliding clamp. Proc. Natl. Acad. Sci. USA 105, 11116–21 (2008).
30. López de Saro, F. J., Georgescu, R. E., Goodman, M. F. & O’Donnell, M. Competitive processivity-clamp usage by DNA polymerases during DNA replication and repair. Embryo. J. 22, 6408–6418 (2003).
31. Pajunen, M. I. et al. Universal platform for quantitative analysis of DNA transposition. Mobile DNA 1, 24 (2010).
32. Flartszewisch, E., Sensen, C. W. & Schleper, C. High spontaneous mutation rate in the hyperthermophilic archaeon Sulfolobus solfataricus is mediated by transposable elements. J. Bacteriol. 182, 2574–81 (2000).
33. Kleckner, N. Regulating Tn10 and IS10 transposition. Genetics 124, 449–454 (1990).
34. Reznikoff, W. S. Transposon Tn5. Annual Review of Genetics 42, 269–286 (2008).
35. Sousa, A., Bourgard, C., Wahl, L. M. & Gordo, I. Rates of transposition in Escherichia coli. Biol. Lett. 9, 20130838 (2013).
36. Allen, E. E. et al. Genome dynamics in a natural archaeal population. Proc. Natl. Acad. Sci. USA 104, 1883–8 (2007).
37. Muller, R. S. et al. Ecological distribution and population physiology defined by proteomics in a natural microbial community. Mol. Syst. Biol. 6, 374 (2010).
38. Hewson, I. et al. In situ transcriptomic analysis of the globally important keystone N2-fixing taxon Crocosphaera watsonii. ISME J. 3, 618–631 (2009).
39. Mau, R. W., Ponticelli, S. K., Duzen, J. M. & Sutton, M. D. Differential binding of Escherichia coli DNA polymerases to the beta-sliding clamp. Mol. Microbiol. 65, 811–827 (2007).
40. Wagner, A. & de la Chaux, N. Distant horizontal gene transfer is rare for multiple families of prokaryotic insertion sequences. Mol. Genet. Genomics 280, 397–408 (2008).
41. Filel, J., Siguer, P. & Chandler, M. Insertion sequence diversity in Archaea. Microbiol. Mol. Biol. Rev. 71, 121–157 (2007).
42. Gilbert, C. & Cordaux, R. Horizontal transfer and evolution of prokaryote transposable elements in eukaryotes. Genome Biol. Evol. 5, 822–832 (2013).
43. Schonknecht, G. et al. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339, 1207–1210 (2013).
design all research, and performed some experimental work. H.D.M., M.J.G. and F.J.L.S. wrote the
experiment with *Acidiphilium* and V.P. constructed the oligonucleotide microarray. P.S.M.U. and R.A. performed the laboratory evolution
H.D.M. performed most experimental work. M.J.G. performed most of the bioinformatics work. M.M.P.

**Author Contributions**

This work was funded by the Subdirección General de Proyectos de Investigación of the Spanish
Ministry for Economy and Competitiveness Grants No. CGL2010-17384 and AYA2011-24803, and by
the European Research Council (ERC) Advanced Grant No. 250350 (IPBSL). HDM has a FPI fellowship
from the Spanish Government.

The authors gratefully thank Harri Savilahti (University of Turku, Finland) for vector pSKT1, and to
Marina Postigo (Centre for Astrobiology, Spain) for sequencing all genetic constructs. We also thank
Enrique González-Tortuero (Centre for Astrobiology, Spain) for discussion and bioinformatics support.

**Acknowledgments**

The authors gratefully thank Harri Savilahti (University of Turku, Finland) for vector pSKT1, and to
Marina Postigo (Centre for Astrobiology, Spain) for sequencing all genetic constructs. We also thank
Enrique González-Tortuero (Centre for Astrobiology, Spain) for discussion and bioinformatics support.
This work was funded by the Subdirección General de Proyectos de Investigación of the Spanish
Ministry for Economy and Competitiveness Grants No. CGL2010-17384 and AYA2011-24803, and by
the European Research Council (ERC) Advanced Grant No. 250350 (IPBSL). HDM has a FPI fellowship
from the Spanish Government.

The authors gratefully thank Harri Savilahti (University of Turku, Finland) for vector pSKT1, and to
Marina Postigo (Centre for Astrobiology, Spain) for sequencing all genetic constructs. We also thank
Enrique González-Tortuero (Centre for Astrobiology, Spain) for discussion and bioinformatics support.
This work was funded by the Subdirección General de Proyectos de Investigación of the Spanish
Ministry for Economy and Competitiveness Grants No. CGL2010-17384 and AYA2011-24803, and by
the European Research Council (ERC) Advanced Grant No. 250350 (IPBSL). HDM has a FPI fellowship
from the Spanish Government.

**Author Contributions**

H.D.M. performed most experimental work. M.J.G. performed most of the bioinformatics work. M.M.P.
and V.P. constructed the oligonucleotide microarray. P.S.M.U. and R.A. performed the laboratory evolution
experiment with *Acidiphilium* and provided expertise with acidophilic organisms. F.J.L.S. conceived and
designed all research, and performed some experimental work. H.D.M., M.J.G. and F.J.L.S. wrote the
manuscript. All authors read and approved the final version of this manuscript.

**Additional Information**

**Supplementary information** accompanies this paper at http://www.nature.com/srep

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Diaz-Maldonado, H. et al. Transposase interaction with the β sliding clamp:
effects on insertion sequence proliferation and transposition rate. *Sci. Rep.* 5, 13329; doi: 10.1038/
srep13329 (2015).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or
other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/