RESEARCH REVIEW

INTERNATIONAL MICROBIOLOGY (2011) 14:187-199
DOI: 10.2436/20.1501.01.148 ISSN: 1139-6709 **www.im.microbios.org**

INTERNATIONAL MICROBIOLOGY

Unconventional lateral gene transfer in extreme thermophilic bacteria

Carolina E. César[1], Laura Álvarez,[1] Carlos Bricio,[1] Esta van Heerden,[2] Dereck Littauer,[2] José Berenguer[1]*

[1]Center of Molecular Biology 'Severo Ochoa', Autonomous University of Madrid-CSIC, Madrid, Spain.
[2]Department of Microbial, Biochemical and Food Biotechnology, University of the Free State, Bloemfontein, South Africa

Received 30 October 2011 · Accepted 17 November 2011

Summary. Conjugation and natural competence are two major mechanisms that explain the acquisition of foreign genes throughout bacterial evolution. In recent decades, several studies in model organisms have revealed in great detail the steps involved in such processes. The findings support the idea that the major basis of these mechanisms is essentially similar in all bacteria. However, recent work has pinpointed the existence of new, evolutionarily different processes underlying lateral gene transfer. In *Thermus thermophilus* HB27, at least 16 proteins are required for the activity of one of the most efficient natural competence systems known so far. Many of those proteins have no similarities to proteins involved in natural competence in other well-known models. This unusual competence system is conserved, in association with the chromosome, in all other *Thermus* spp. genomes so far available, it being functional even in strains from isolated environments, such as deep mines. Conjugation is also possible among *Thermus* spp. Homologues to proteins implicated in conjugation in model bacteria are encoded in the genome of a recently sequenced strain of *Thermus thermophilus* and shared by other members of the genus. Nevertheless, processive DNA transfer in the absence of a functional natural competence system in strains in which no conjugation homologous genes can be found hints at the existence of an additional and unconventional conjugation mechanism in these bacteria. [**Int Microbiol** 2011; 14(4):187-199]

Keywords: *Thermus* · thermophiles · lateral gene transfer (LGT) · conjugation · transformation

Introduction

Comparative whole-genome analyses have revealed that prokaryotic genomes are extraordinarily plastic. The enormous internal variation found within closely related species and strains is most likely produced by the rapid gain (and loss) of genetic material through lateral gene transfer (LGT) [26,29]. Indeed, lateral gene transfer is a leading force driving the accelerated evolution of prokaryotes, conferring on bacteria the unique ability to rapidly adapt to various environmental changes. Of particular relevance is the capability of prokaryotes to colonize novel, previously restricted ecological niches. This characteristic has come about by the lateral inheritance of settlement modules or operons that can deliver an entire new physiological ability in a single evolutionary event [32,39].

There are three classical pathways described for LGT: transduction, transformation, and conjugation. While transduction involves the accidental packaging of bacterial DNA into prophage capsids and its ulterior propagation through defective bacteriophage infection, transformation and conjugation involve specialized DNA transport machineries able to actively acquire foreign DNA or to transfer it to recipient cells. Natural transformation is the least encountered of the three common LGT processes, with about 70 species of bac-

***Corresponding author:** J. Berenguer
Centro de Biología Molecular Severo Ochoa (UAM-CSIC)
Universidad Autónoma de Madrid
28049 Cantoblanco (Madrid), Spain
Tel. +34-911964498. Fax +34-911964420
E-mail: jberenguer@cbm.uam.es

teria described as able to introduce external DNA into the cell through an energy-requiring mechanism [28]. In most cases this capability is not constitutive but is induced under specific, normally adverse, growth conditions during which cells become 'competent' for DNA uptake. This state implies the synthesis of a complex protein machinery that forces the entrance of external DNA through the various cell envelopes of gram-negative bacteria [17]. The induction of competence and the physiology of natural transformation models have been intensively studied in gram-positive and gram-negative bacteria [16,17,30,33,41]. These models are remarkably different among species; in fact, it is likely that many naturally competent bacteria have not been detected yet because of our inability to define the appropriate inducing conditions [15]. In contrast to these highly regulated pathways in well-studied naturally competent bacteria, there are at least two species, *Helicobacter pylori* and *Thermus thermophilus*, for which a high efficiency of DNA incorporation is found along the whole growth curve, such that these species are considered as constitutively competent [23,24]. The nature of the competence apparatus of *T. thermophilus* is discussed in this article.

An important issue regarding natural transformation is the fate of the DNA incorporated into the cells, which depends very much on the nature of the DNA itself. In most cases DNA that enters the cell as a single-stranded molecule is used as a nutrient, degraded by nucleases to enter catabolic cycles or to be used as building blocks for new cell material [45]. In some cases, single-stranded DNA (ssDNA) sequences that bear homology to genome sequences can undergo RecA-mediated homologous recombination, allowing gene alleles and even novel genes to be incorporated into the genome of the transformant strain, hence favoring the horizontal spread of new traits.

Contrary to natural transformation, which depends on the ability of the recipient cell to take up DNA, conjugation relies on the direct contact between a donor cell harboring all the essential genetic determinants for conjugation [42] and a susceptible recipient strain. Unlike transformation and transduction, conjugation is a highly specialized process in which the nature of the transferred DNA involves, in most cases, an episomal double-stranded DNA (dsDNA) molecule—rarely, chromosomal DNA—that either carries all the genetic material needed for conjugative transfer, i.e., self-transmissible mobile genetic elements, or requires an accessory, generally plasmid-borne, conjugative apparatus in order to be mobilized from one cell to another. However, in some cases, the integration of one of these conjugative episomes into the chromosome, either by specific integrases or through homologous recombination, results in the mobilization of the whole chromosome, as in the case of Hfr strains of *Escherichia coli* [25,47]. The existence of the so-called conjugative transposons, or integrative and conjugative elements (ICEs) that can integrate into the genome and be excised from it as a plasmid via a prophage-like mechanism, is also common. Such ICEs encode the conjugative apparatus required for their transfer to a recipient strain [53].

There is little doubt about the contribution of LGT in the exceeding dynamism of prokaryotic genomes. However, LGT is not random but instead has been shown to be greatly influenced by internal and external environmental variables that delimit particular gene-exchange communities on the basis of shared factors such as genome size, GC contents, and carbon and oxygen sources [26]. High-temperature environments could, in principle, be perceived as a limiting factor for LGT, as DNA from mesophiles rarely encodes thermostable proteins. Nonetheless, it has been shown that temperature tolerance is a factor heavily associated with elevated rates of LGT exchange [26]. Systematic genome-scale comparison analyses revealed a complex history of LGT events between and within thermophilic bacteria and archaea. Evidence of ample LGT from *Archaea* to *Bacteria* comes from the Aquificae class hyperthermophile *Aquifex aeolicus*, in which 16 % of the predicted coding sequences are likely derived from *Archaea* [6]. In another hyperthermophile, *Thermotoga maritima*, 24 % of the predicted coding regions have probably arisen as a consequence of extensive LTG from thermophilic archaea [37]. By contrast, in mesophilic bacteria the percentage of genes that have probably been acquired from *Archaea* is much lower, suggesting that thermal environments are especially prone to select for LGT events as a means to accelerate adaptation to such extreme conditions.

The *Thermus-Deinoccocus* clade provides a good example of specialization and innovation through LGT, as evidenced by the thermophile *Thermus thermophilus* and the radiation-resistant mesophile *Deinococcus radiadurans*. Remarkably, both organisms are archetypical members of two very distinct lifestyles. Parsimonial evolutionary reconstruction methods predict that the two species share a mesophilic or moderately thermophilic ancestor, from which they evolved through LGT events that facilitated the gradual acquisition of either thermophile adaptations from other thermophilic bacteria and archaea or radiation resistance determinants [40]. This scenario supports the notion that the common ancestor of the two genomes was prone to receive genes by LGT, a property that nowadays is preserved in *Thermus* spp. but is apparently absent in the genus *Deinococcus*.

Moreover, a particularly interesting characteristic of *T. thermophilus* strain HB27 that makes it the thermophilic organism of choice for genetic modification is the constitutive expression of an extremely efficient natural competence system [7,23], which allows strain transformation at high frequencies under laboratory conditions [14]. Indeed, transformability seems to be a widespread characteristic in the genus *Thermus* [31]. In addition to the natural transformation phenotype of *Thermus*, DNAse-resistant DNA transfer by means of conjugation has been demonstrated in *T. thermophilus*, namely, by the ability of aerobic strain HB27 to acquire from the facultative NAR1 strain a respiratory nitrate reductase gene cluster (*nar*) [43,44]. This conjugative process also has been shown to allow the transfer of a complete set of denitrification genes from strain PRQ25 to the same aerobic HB27 recipient strain [5,9].

In the following sections of this review, we aim to provide the reader with a general overview of the mechanisms and promiscuity of LGT in the genus *Thermus*, one of the most widespread genera of thermophilic bacteria. We focus in particular on *T. thermophilus*, a reference organism for the genetic study of thermophiles and for the structural characterization of proteins, as well as an exceptional source of enzymes of great biotechnological potential.

Natural competence systems. DNA uptake during competence is mediated by a complex macromolecular assembly comprising a large number of proteins. In order to achieve transfer of the incoming DNA into the cytoplasm, the natural barrier imposed by the cell envelope must be overcome. For gram-positive bacteria, this process means the breaching of a robust peptidoglycan (PG) layer and the cytoplasmic membrane (CM). In gram-negative bacteria, the presence of the outer membrane (OM) is an added barrier to this process. Despite the marked differences in the cell-envelope structures and the distinct characteristics of the DNA-uptake pathways, gram-positive and gram-negative bacteria use related proteins in the construction of the DNA transporter apparatus. Regarding the cell-envelope structure, and due to its unique nature which does not conform to the Gram staining method for classification, the genus *Thermus* is a fairly singular case. *Thermus thermophilus* cells are covered by an EDTA-extractable material with high sugar content that hides an array of hexagonal symmetry built up by the SlpA protein, a structure known as the S-layer but which more closely resembles a regular OM protein [8]. This structure constitutes the scaffold for an OM and is tightly bound to an underlying layer of peptidoglycan-associated secondary cell wall polymers (SCWP) [13].

These special features involving the anatomy of the *Thermus* cell envelope together with its thermophilic nature hint at an unusual model of DNA translocation in *T. thermophilus* strain HB27 [49]. Whole-genome analysis and directed knockout mutagenesis have led to the identification of 16 genes involved in natural competence that are arranged in seven transcriptional units. These competence proteins have been split into three distinct groups: (i) homologues to DNA-translocation-specific proteins from *Bacillus subtilis* and *Neisseria gonorrhoeae* (ComEC, ComEA, DrpA), (ii) homologues to components of the type IV pili (PilA1-4, PilD, F, C, Q), and (iii) homologues that are not related to any other natural transformation systems (ComZ, PilM, N, O, W) [7].

Localization and structural analyses performed on these proteins have yielded a model in which the core of the DNA translocator apparatus is a pseudopilus that originates from the cytoplasmatic membrane (CM) and spans the periplasm (Fig. 1). This type IV-like pilus (T4P) is formed by the polymerization of small pilin proteins into long dynamic fibers, which in gram-negatives emerge through the OM via a ring formed by the oligomerization of a secretin-like protein (for a comprehensive review, see [7,17]). *T. thermophilus* HB27 encodes four pilin-like proteins, PilA1, PilA2, PilA3, and PilA4. Mutations in any of these lead to a transformation-negative phenotype but only *pilA4* mutants lack pili, as established by electron microscopy [20]. The multimeric secretin-like PilQ is also essential for transformation and its deletion results in a complete loss of DNA binding ability. PilQ localizes to the OM and forms an oligomeric structure of 15 × 34 nm that spans the periplasmic space, forming a channel-like structure wide enough to accommodate a pilus [11]. Localization of PilQ in the OM depends on the presence of the competence protein PilW, which is unique to the genus *Thermus* [46].

Similarly to other systems, prepilins must be processed into a mature form by a prepilin peptidase, probably encoded by *pilD* [20,38]. In addition, T4P systems require the presence of an AAA^+ ATPase to drive the polymerization of pilins, which are required for both piliation and transformation in gram-negative bacteria [34]. In the *Thermus* system, the putative traffic NTPase function is covered by PilF; however, in contrast to other traffic NTPases, mutations in this gene that affect transformation ability result in a piliated phenotype, as seen under the electron microscope [21].

In the *Thermus* "system", the well-conserved proteins ComEA and ComEC are thought to play roles analogous to those proposed for other species. ComEA is a periplasmic protein anchored to the membrane. Its function may be to

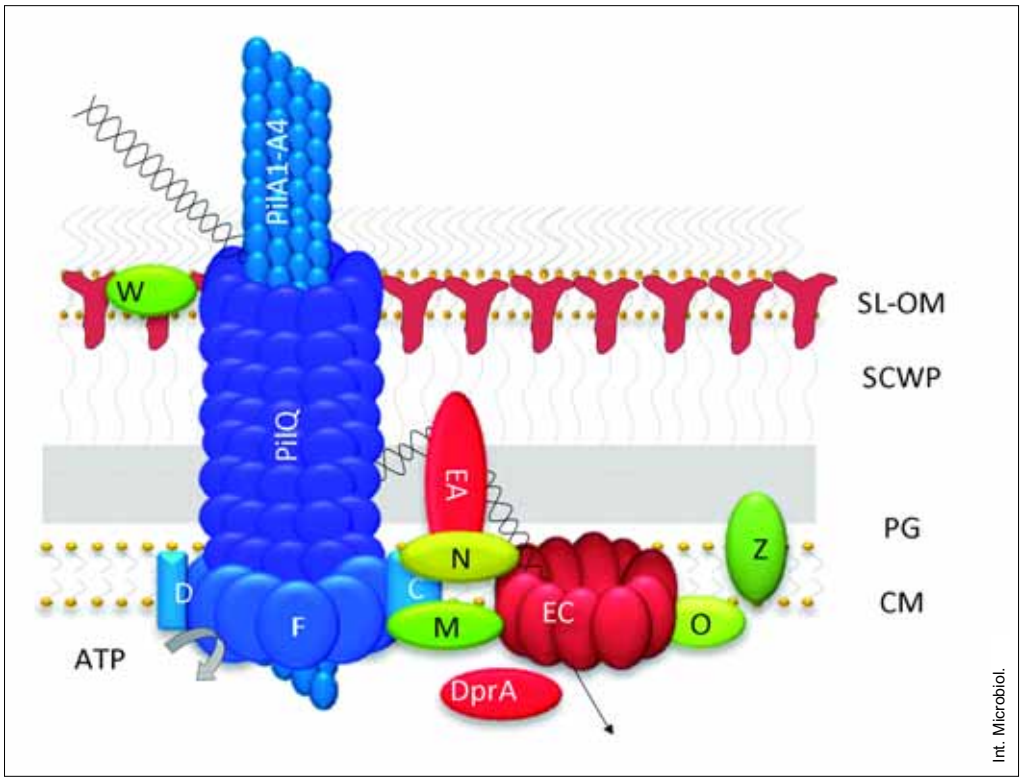

Fig. 1. Model of DNA translocation in *Thermus thermophilus* HB27, as described in [7]. Pre-pilins A1–A4 are processed to mature pilins by the cytoplasmic membrane peptidase PilD and secreted to the periplasm, where they are assembled into a fiber-like pilus, powered by the ATPase PilF. The pilin fiber emerges through the outer membrane (OM) via a secretin ring, formed by PilQ units. Localization of the secretin ring to the OM is strictly dependent on PilW. Retraction of the pilus (likely powered by PilF) facilitates DNA transport through the cell wall, where it binds ComEA and is then presented to ComEC. The latter transfers one strand of the DNA molecule to the cytoplasm, while the other strand is degraded in the cytoplasm by an unknown nuclease. PilC, PilN, and PilO are part of the multiprotein inner membrane complex. Type IV pilin proteins are shown in blue, homologous proteins found in *B. subtilis* and *N. gonorrhoeae* systems in red, and proteins present only in the *Thermus* transport system in green. SCWP: secondary cell wall polymers.

bind DNA in the periplasm and then deliver it to the cytoplasmic membrane channel formed by the polytopic ComEC [49]. PilC is another conserved competence protein essential for transformation. It lies on the CM, where it is believed to link the periplasmic and cytoplasmic components of the T2SS and T4P systems [20]. DprA is a cytosolic protein that is also present in other transformation systems. Although not essential for transformation, it has an active role in processing incoming DNA in the cytoplasm [35]. Very little is known about the remaining components of the *Thermus* competence system. Conservation of ComZ, PilM, PilN, or PilO in other studied systems is not evident. ComZ appears to be anchored to the CM as part of the transporter assembly scaffold [21]. PilM, PilN, and PilO are all located in the inner membrane [46]. Mutations in the *pilM*, *pilN*, or *pilO* genes render the cell incapable of being transformed, although no effect has been demonstrated on DNA binding or transport to the periplasm in any case [48,49].

These remarkable structural differences in the DNA translocator apparatus of *T. thermophilus* HB27 are conserved in other *Thermus* spp. isolates, as revealed by genome sequence analyses. pBLAST searches against the GenBank *Thermus* taxon (including whole-genome projects for *T. thermophilus* HB8, *T. thermophilus* SG0.5JP17-16, *T. aquaticus* Y51MC23, and *T. scotoductus* SA-01 strains) retrieved high-scoring homology matches at the protein level with all the components of six of the seven competence loci mentioned above for *T. thermophilus* HB27. These well-conserved loci comprise the *pilM-Q* and *comEA-comEC* operons, *pilD*, *pilC*, *pilF*, and *dprA*. Genetic organization of the competence genes is also preserved (Fig. 2A) (in our analyses we have included unpublished data from our laboratory involving the *T. thermophilus* strain PRQ25). Amino acid sequence alignments of each individual set of orthologous genes yielded overall identity scores ranging from 41 % (PilW) to 83 % (PilF).

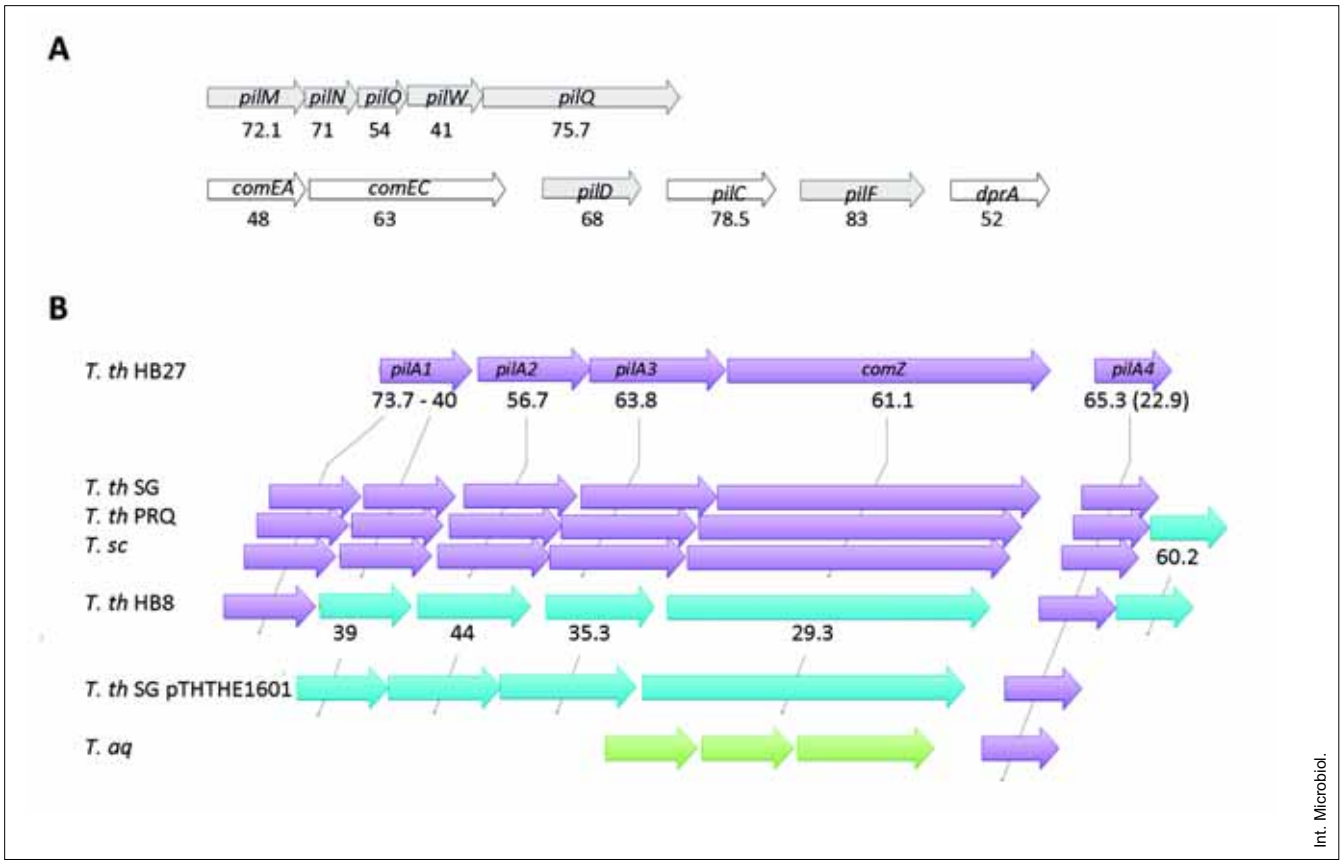

Fig. 2. Conservation of competence proteins and loci amongst *Thermus* spp., based on the transformation system defined for *T. thermophilus* HB27. Species include: *T. thermophilus* HB8, SG0.16-17, and PRQ25, *T. aquaticus* Y51MC23, and *T. scotoductus* SA-01 spp. and their corresponding megaplasmids, where applicable. (**A**) Competence loci presenting no deviation from those described in strain HB27. Numbers below the transcription units denote the percentage of overall identity amongst homologous proteins. (**B**) Conservation of the *pilA1-4* operon within *Thermus* spp. Blocks of proteins presenting significantly similar identity are represented in the same colors (purple for strain HB27-like proteins, blue for those of strain HB8, and green for those of *T. aquaticus*) and drawn together by broken lines. Overall identities are given below each transcription unit block.

The organization and protein array of the *pilA1-4* operon, on the other hand, are somehow less conserved (Fig. 2B). BLAST searches on the four prepilin-like proteins of strain HB27 retrieved positive matches for all the proteins in *T. thermophilus* strains SG0.5JP17-16 and PRQ25. In *T. scotoductus* SA-01, no results were obtained for PilA2-4; however, a detailed inspection of the genome context of *pilA1* and *comZ* in the genome of this strain revealed the presence of three putative genes coding for homologues to PilA2, PilA3, and PilA4. Overall identities range between 56.7 % for PilA2 and 72.7 % for PilA1 (also present in the *T. thermophilus* HB8 strain), with the exception of PilA4, which is conserved in all the analyzed strains with a low overall identity (22.9 %), except for the N-terminal 50 amino acids (65.3 %). This N-terminal part of the protein includes the conserved cleavage/methylation signal that is characteristic of all prepilin-like proteins [38]. Interestingly, all the strains that conserve the genetic structure of the HB27 *pilA1-4* operon present a gene duplication involving *pilA1*. The protein products of these duplicates are more loosely conserved and have an identity of 40 % with HB27 PilA1, in contrast with the 73.7 % identity of the other PilA1 proteins.

In *T. thermophilus* HB8, a pBLAST approach did not identify any homologue to either PilA2, PilA3, or ComZ. Nevertheless, inspection of the genomic context of the conserved *pilA1* and *pilA4* genes revealed a genetic organization similar to that of the *pilA1-4* operon but comprising four prepilin-like proteins that bear no homology at the sequence level with any of HB27 competence pilins, the exceptions being the conserved PilA1 protein; a larger protein of similar size and operon position to ComZ, which is well conserved at the N-terminal leader region; and a PilA4 homologue and a PilinV like protein. A homologue (60.2 % identity) to this pilinV-like protein was also found downstream from *pilA4*, in

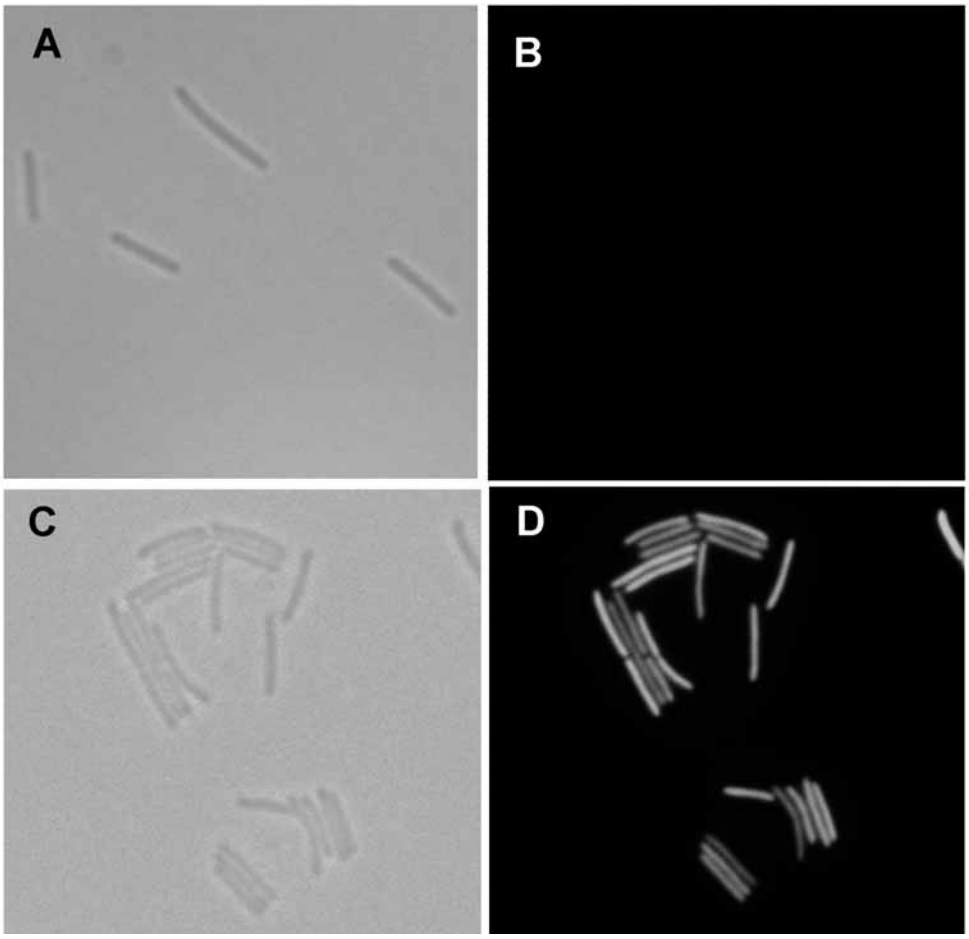

Fig. 3. Transformation in *Thermus scotoductus* SA1. Cells of exponential cultures *T. scotoductus* SA1 were observed by phase contrast (**A**,**C**) and fluorescence (**B**,**D**) microscopy. Panels **C** and **D** correspond to kanamycin-resistant cells of *T. scotoductus* SA1 obtained by transformation with a *Thermus-E. coli* bifunctional vector expressing a thermostable variant of GFP (green fluorescent protein). Panels **A** and **B** correspond to the untransformed controls.

the PRQ25 strain. We performed pBLAST searches with each individual protein sequence within HB8 *pilA1-4 (ttha1217-ttha1220)* and found homologous proteins (29.3–44 % identity) in the *T. thermophilus* SG0.5JP17-16 plasmid pTHTHE1601 (*ththe16_2408-2411*), preceding the PilA4 gene sequence. Therefore, two types of *pil-comZ* clusters (HB27- and HB8-like operons) are present within *Thermus* strains and they probably play similar roles in transformation. In this sense, it is worthy to note the presence of both types of clusters in *T. thermophilus* SG0.5JP17-16, one HB27-like in the chromosome and the other HB8-like in the megaplasmid, supporting the idea that their activities are compatible within the cell. Nevertheless, in natural competence assays parallel to those used with transformable strains, we were not able to obtain any transformants with this *Thermus* strain (C.E. César, unpublished results).

Thermus aquaticus Y51MC23 offers a very different picture. The pilin operon consists of only four coding sequences (TaqDRAFT_4097-4094) that code for four putative prepilin-like proteins, including a homologue to PilA4 (TaqDRAFT_4094). The first three prepilin-like proteins of the operon show no identity to any of the previously mentioned pilins. No trace of a putative *comZ* gene was found by manual inspection, and pBLAST searches did not retrieve any homologous protein in this organism. Assuming that the available genome draft is complete, the absence of this gene suggests that this strain cannot take up external DNA. Unfortunately, there are no experimental data to confirm these results.

The presence in the genome of homologues to all of the *T. thermophilus* HB27 competence proteins does not guarantee a high efficiency of transformation. Actually, there is a great difference in the transformation efficiencies of strain HB27 and any other strain so far studied. The extraordinary DNA transformation efficiency conspicuous to strain HB27 (10^{-2} transformants/viable cells [31]) is not shared by other *T. thermophilus* strains, such as HB8, which shows transformation frequencies one or two orders of magnitude lower than those of HB27, and PRQ25, which in spite of the high degree of homology and perfect conservation of the compe-

tence protein components shows very low transformability (four orders of magnitude lower) when transformed either with plasmids or with its own chromosomal DNA (our own results). Nevertheless, it is still possible to transfer genetic markers to other strains for genetic analysis if appropriate conditions for transformation and selection are chosen. An example is strain *T. scotoductus SA-01*, whose competence genes are remarkably similar to those of *T. thermophilus* HB27 and which can be transformed at low efficiencies with plasmids expressing fluorescent proteins, when a long expression period is used (Fig. 3).

Conjugation proteins in *Thermus* spp. Undoubtedly, LGT accounts for most of the plasticity found within the restricted thermophile environment, as demonstrated by the widespread presence of homologous clusters within thermophilic bacteria and archaea. Natural transformation in the genus *Thermus* is an example of the tremendous promiscuity shown by these bacteria, as inferred not only by the high frequency of transformation observed by its members but also by the broad range of DNA species they are able to accommodate, i.e., those from *Archaea*, *Bacteria*, and *Eukarya* [40]. In this sense, note that several of the genes that can be considered as having been acquired by LGT in the aerobic strains HB27 and HB8 are located on their respective pTT27 megaplasmid. These include several genes encoding a DNA repair system found in thermophilic archaea and even a reverse gyrase in strain BH8 that could contribute to the thermophilic lifestyle of *Thermus* spp. [10]. Recently, genes for denitrification present in facultative strains of this species were identified in the megaplasmid [5,9]. However, whether this extraordinary natural competence ability is sufficient to explain the plasticity observed by *Thermus* spp. remains to be determined.

It is worth considering that conjugation, like DNA transformation, involves a DNA transport system that has to adapt to different environments and cell types in order to transport nucleoprotein complexes across cell membranes. The type 4 secretion system (T4SS) is a well-conserved and versatile system that mediates the transfer of DNA, toxins, and nucleoprotein complexes in both *Bacteria* and *Archaea* [3,4]. T4SSs have been recently classified in three functional types: a first type, involved in conjugation processes (cT4SS), which mediates the cell to cell transfer of proteins and DNA; a second type, involved in the translocation of virulence factors to the cytosol of eukaryotic cells by pathogenic species; and a third type, responsible for the specialized uptake and release of DNA [4].

The cT4SSs are large macromolecular assemblies that vary substantially in their subunit number and composition, they present a number of mating-pair-formation proteins and two or three dedicated ATPases that provide energy for DNA and protein transfer and for pilus assembly. The prototypical models for cT4SS (Fig. 4A) are derived from *E. coli* pKM101 plasmid and from *Agrobacterium tumefaciens* T4SS [51,52]; the latter does not participate in conjugation but is specialized in the delivery of oncogenic nucleoprotein complexes into plant cells. These systems are characterized by: three ATPases (VirB4, VirB11, and VirD4 or coupling protein, CP), which are postulated to pump ssDNA during conjugative transfer [12]; a series of scaffolding proteins (VirB6-VirB10, and possibly VirB3), that cross the entire cell envelope; and T4Ps similar to those involved in the competence DNA translocator system.

While the presence of VirB-like proteins is not apparent in those *T. thermophilus* strains in which conjugation has been addressed (*T. thermophilus* HB27, *T. thermophilus* HB8, *T. thermophilus* NAR1, and *T. thermophilus* PRQ25), the recently available *T. thermophilus* SG0.5JP17-16 megaplasmid pTHTHE1601 sequence contains a putative and structurally novel VirB operon consisting of 20 open reading frames (ORFs; conjugative transfer region, *ctr*, *1-20*, *Tthe16_2024-2005*, Fig. 4B). Three protein products display clear homologies in the amino acid sequence with VirB proteins: the pilin precursor TrbC/VirB2 and the VirB4 and VirD4 ATPases. pBLAST searches involving VirB4 and VirD4 retrieved matches in many different phyla, which is not surprising since VirB4 and T4CPs are extremely well conserved and systematically found in all T4SSs described to date [19]. VirB2 searches, on the other hand, produced significant alignments ($E < 0.01$) only with *T. aquaticus* Y51MC23, *T. scotoductus* SA-01, *Meiothermus ruber* DSM1279, and *M. silvanus* DSM9946 proteins. Routine pBLAST searches with each *orf* product revealed that the conservation of the *Thermus* VirB-like operon is restricted to these species (Fig. 5).

Assignment of VirB functions. Aside from the already mentioned VirB2, VirB4, and VirD4 coded by *ctr1*, *3,* and *7*, respectively, the predicted protein products of *ctr10*, *12*, *13*, *18,* and *20* have annotated functions, while the remaining *ctr2*, *4–6*, *8*, *9*, *11*, *14–17,* and *19* code for proteins of unknown function. Ctr10 is a putative lytic transglycosidase (LT). These murein-degrading enzymes are ubiquitous to gram-negative T4SS, where they participate in the degradation of the PG to allow pilus assembly [55,56]. In the

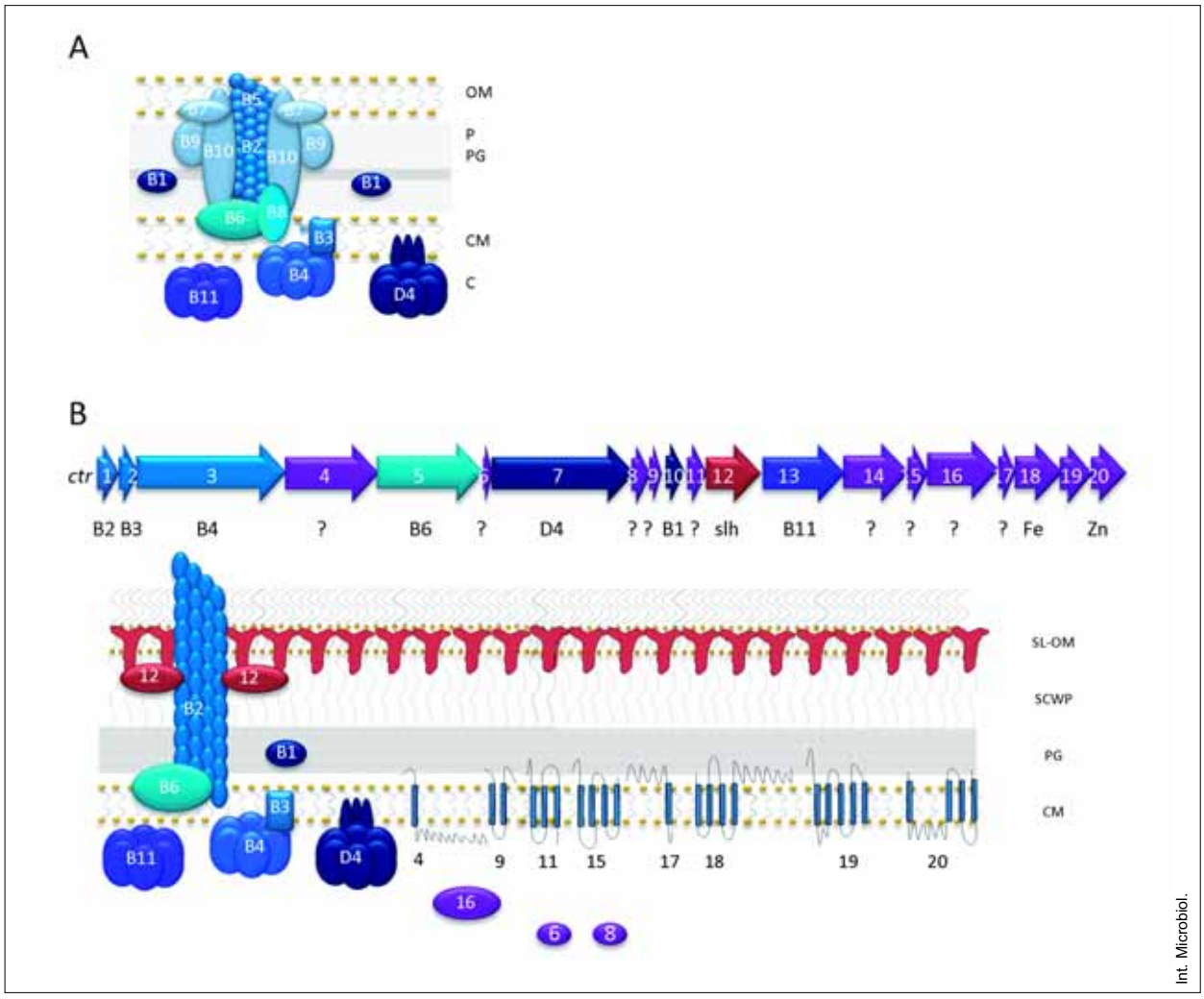

Fig. 4. Type 4 secretion system model structures. (**A**) Experimentally predicted conjugative VirB4/VirD4 transfer system, based on [51]. OM, outer membrane; CM, cytoplasmic membrane; P, periplasm; C, cytoplasm; PG, peptidoglycan. (**B**) Predicted subunits for the *Thermus termophilus* SG0.Jl VirB-like system. Membrane disposition as well as the transmembrane (TM) domains of T4SS subunits are represented schematically, according to SOSUI predictions and the presence of putative peptide signals. The localizations of Ctr12, VIrB1, VirB2, VirB3, VirB4, VirB6, VirB11, and VirD4 are based on the VirB4/VirD4 *Agrobacterium tumefaciens* model [52]. Proteins with homologous functions are depicted by the same color, proteins of unknown function are represented in purple. SL-OM, S-layer outer-membrane; SCWP, secondary cell wall polymers; PG, peptidoglycan; CM, cytoplasmic membrane.

A. tumefaciens system, the LT role is performed by VirB1. Based on the functional similarities and widespread presence of LTs in T4SS, we have defined Ctr10 as VirB1. *ctr12* codes for a 466-residue polypeptide that contains an N-terminal S-layer homology (SLH) domain, involved in the anchoring of secreted protein to the cell surface, and a spectrin-repeat domain typical of cytoskeletal structural proteins. Two other genes, *ctr4* and *ctr16,* code for proteins with 3D structures similar to actinin (E = 0.65) and contractile/cell adhesion proteins (E = 2.5×10^{-05}). The conserved presence of cytoskeletal elements in the cluster is unforeseen. The prokaryotic cytoskeleton is postulated to be a central organizer for the accurate positioning of proteins and nucleoprotein complexes, thereby controlling macromolecular trafficking in cells. We can only speculate that these three proteins are members of a novel cytoskeleton-like scaffold that provides a dynamic brace to the T4SS, with Ctr4 anchored to the inner membrane (topological predictions, SOSUI server, reveal the presence of a transmembrane helix (TM) at the N-terminal end of the protein), Ctr12 anchored to the cell surface via SLH, and the predicted soluble Ctr16 located in the cytoplasm (Fig. 4B).

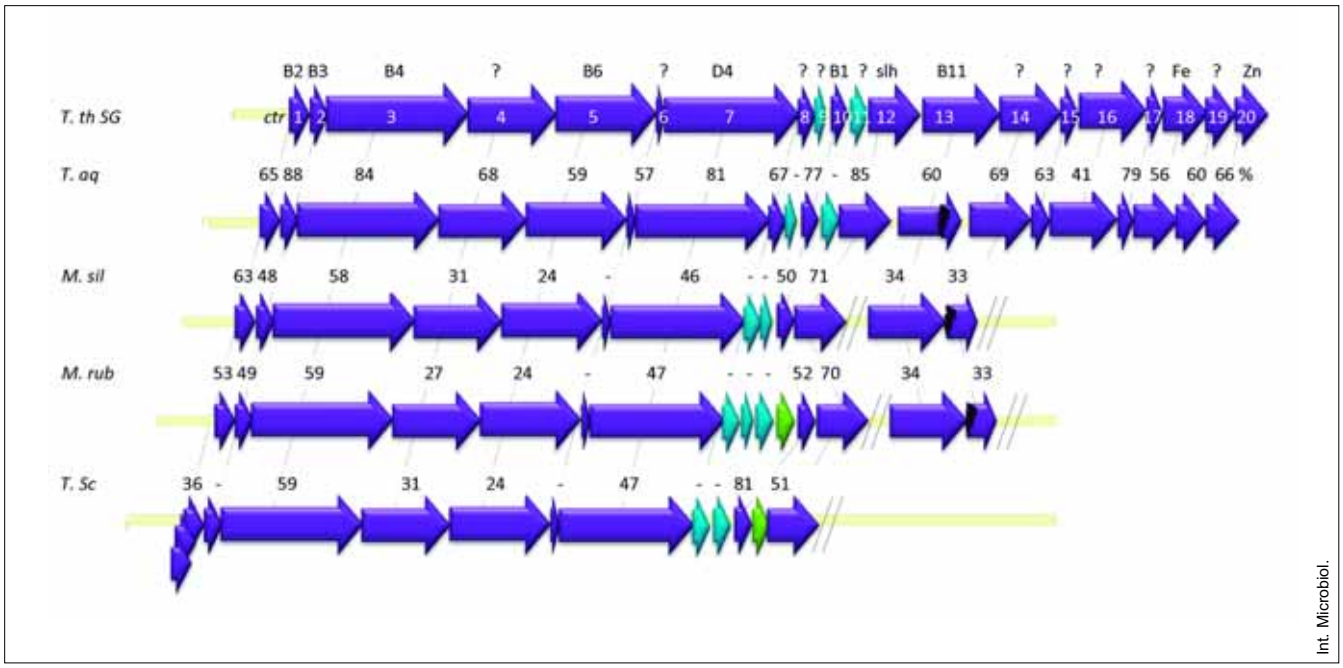

Fig. 5. Conservation of the putative conjugative *virB* operon in *Thermus* spp. Overall identities (as obtained by BLASTp searches) with respect to the *T. thermophilus* SG0.15-16 *virB* operon are shown above each open reading frame (ORF). The (–) symbol denotes homologous proteins whose score value is lower than the threshold for BLASTp retrieval. A hypothetical/predicted role for each gene product is given, whereas uncertain roles are denoted by question marks. Purple arrows represent ORFs with significant homology. Blue ORFs code for predicted homologous transmembrane proteins of reduced overall identity. Light green ORFs code for proteins that are unique for a particular spp. A black-filled lightning bolt represents either an interruption of the coding frame or a truncated product.

Ctr13 is an AAA^+ type ATPase, as is VirB11. No sequence similarity is found between Ctr13 and VirB11-like proteins, and VirB11 is not part of all T4SS. T4SS, as it is absent from most gram-positive cT4SSs [22]. We have designated this protein as VirB11, since the protein structure prediction (Phyre version 0.2) of Ctr11 ATPase revealed a 100 % match with *Brucella suis* VirB11 ATPase ($E = 3.1 \times 10^{-20}$).

The two other annotated proteins are putative iron-dependent transcriptional repressors, Ctr18, of unknown function within the system, and a zinc-dependent metallopeptidase, Ctr20. While no peptidases are commonly found in VirB operons, many VirB-like proteins contain a signal peptide for translocation into the periplasm and that peptide is processed from the mature protein once the latter has been delivered. Hence, the presence of a membrane peptidase within the VirB-like operon could represent a novel feature of the *Thermus* T4SS, whose role is covered by an unknown accessory function not preserved in other well-characterized systems. The fact that *ctr20* is not conserved within other *Thermus* VirB-like operons, or anywhere in the genome, points to the existence of an alternative system, perhaps reminiscent of other, traditional VirB-like systems, for the processing of T4SS subunits. Note that the protein product of *ctr19*, which shows no significant homology to any known function either at the amino acid sequence or structural level but is topologically predicted to be an integral membrane protein, yields multiple low-score pBLAST hits with eukaryotic zinc transporters. We can therefore hypothesize a putative role for this transporter, in which it would act in concert with the zinc-dependent metallopeptidase.

With the exception of Ctr6 and Ctr8, which failed to retrieve any sequence or structural similarities with any known protein and were not predicted to contain any signal peptide or TM helices, protein products from *ctr2*, *5*, *9*, *11*, *15*, and *17* contain one or more TM helices (Fig. 4B). Ctr2 is a small protein, 105 residues, that contains two TM helices; its gene is located immediately upstream from that of VirB4. These characteristics are reminiscent of those of VirB3, which in many systems is coded as a chimeric protein fused to the N-terminal portion of VirB4. The presence of VirB4, together with VirB7 and VirB8, is essential in stabilizing VirB3 during T-pilus biogenesis [36]. While we could infer the role of Ctr2 to be that of VirB3, we were unable to extend our analysis to predict putative VirB7 and VirB8 functions.

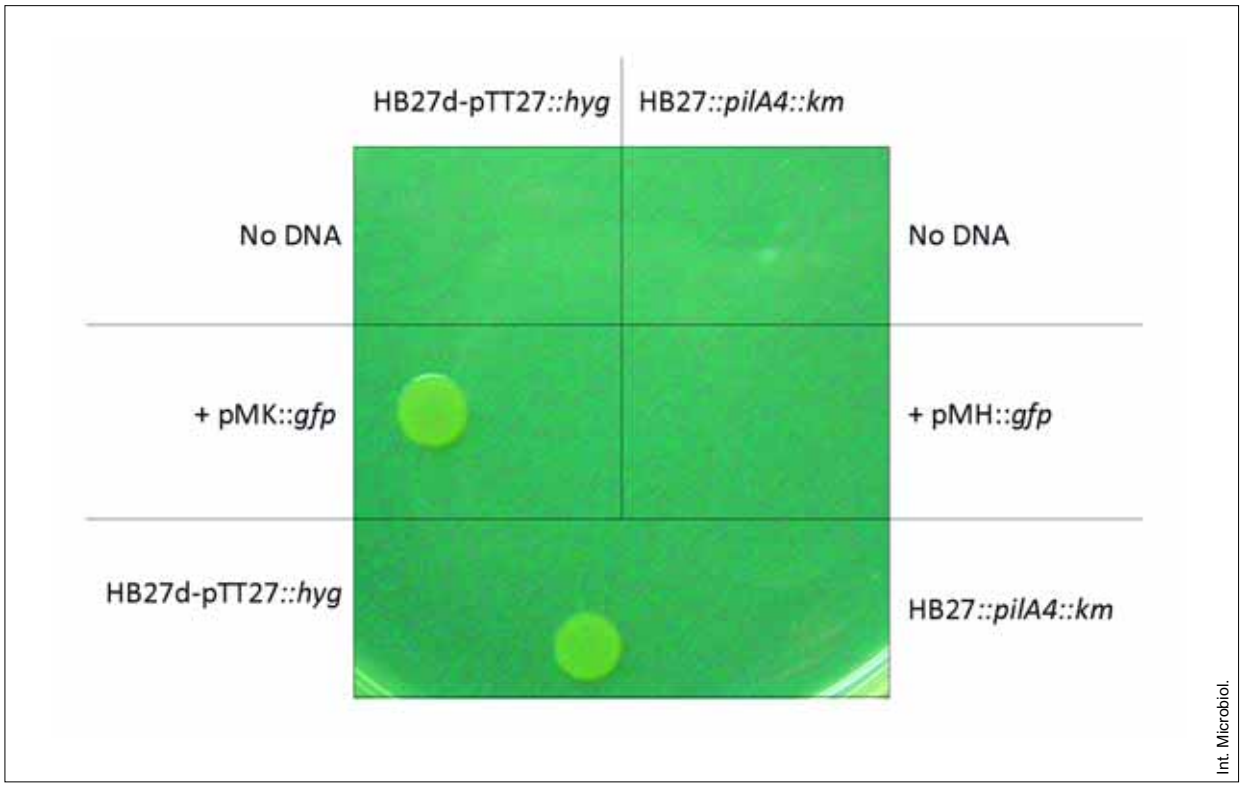

Fig. 6. Ten µl of saturated cultures of HB27d-pTT27::*hyg* strain (containing a plasmid-encoded hygromycin resistance marker, hyg), and HB27::*pilA4*::*km* (containing a chromosomally encoded kanamycin marker, km, and transformation-deficient) were plated on *hyg*-*km*10 ng/µl. Growth was observed only for the transformation-proficient HB27d-pTT27::*hyg* strain. When 10 µl of each strain were mixed together, growth was observed (bottom image), due to the conjugation-like transfer of marker genes.

Similarly, we could not assign identities to VirB9- or VirB10-like functions. VirB7-10, together with the already mentioned VirB3, and VirB6 are probably members of the theoretical scaffold that spans the cell envelope and supports substrate translocation [51,52]. VirB6-like proteins are large polytopic proteins (5 ≤ TM segments) characterized by the presence of a central periplasmic domain and a C-terminal hydrophilic domain [4, 27]. Like VirB6, Ctr5 is a polytopic protein containing nine TM segments as well as a central and a C-terminal hydrophilic domain. VirB6-like proteins display a low overall similarity at the sequence level but are ubiquitous to all T4SS described to date. They are proposed to be essential for conjugation and the absolute minimal requirement for substrate translocation across membranes together with the ATPases T4CP and VirB4 [4]. Because of the predicted structural similarities between Ctr5 and VirB6 proteins and corroborated by its position within the *Thermus* VirB-like operon in addition to the low conservation displayed between homologues present in the genus *Thermus* (Fig. 5), we have tentatively assigned Ctr5 to the large family of VirB6 proteins. It is likely that this scaffolding function is provided by the above-mentioned cytoskeletal elements Ctr4, 12, and 16, together with some or all of the non-assigned proteins, Ctr8, 9, 11, 15–17. The presence of genes coding for small, non-conserved, TM-containing proteins overlapping *virB*-like genes is common to gram-positive VirB operons and very probably represents different strategies, with a common denominator, the T4SS, used by bacterial cells to translocate substances across the cell envelope, in cell surface recognition, and in the formation of mating pairs [4].

Thus, we propose that these unique features of the *Thermus* T4SS represent an imposed requirement due to the thermophile niche of the genus and the unusual structure and nature of the cell envelope of *Thermus* species. Our model is corroborated by the strict preservation of this unusual T4SS system within the *Thermus* group, including not only *Thermus* but also *Meiothermus* spp. It is feasible that with the gathering and sequencing of new *Thermus* strains other structurally similar *virB*-like operons will come to light.

Experimental evidence is still needed to address the participation of these T4SS in conjugation. All functioning cT4SSs are always associated, by CPs, with DNA transfer

replication systems, determined by the presence of a relaxase protein, a site-specific endonuclease that recognizes transferred DNA at a specific site, *origin of transfer* (OriT), cleaves it, and brings it to the recipient cell [18]. Remarkably, no relaxase-like sequence has ever been described for any *Thermus* organism. Whereas relaxases have been thoroughly characterized and categorized within six distinct families, there is a strong belief that new families of relaxases could emerge with the sequencing and characterization of novel conjugation systems, particularly in bacteria other than proteobacteria [50]. We believe the genus *Thermus,* and by extension thermophilic bacteria, are examples of this novel, yet ancient, family of relaxases and as such, part of an unusual conjugation system.

A new conjugation mechanism in *Thermus thermophilus*? Although no conjugation-like homologous proteins have been found in *T. thermophilus* NAR1, HB27, or PRQ25, a conjugative-like mechanism has been described that involves these strains. This uncharacterized LGT mechanism is responsible for the transfer of a pTT27-plasmid-encoded *nar* operon from *T. thermophilus* NAR1 to strain HB27 [43,44], and from *T. thermophilus* PRQ25 to strain HB27 [5]. Further proof of the existence of this atypical LGT between *Thermus* strains comes from experiments involving the transfer of an antibiotic resistance marker, coded within the pTT27 megaplasmid, from HB27-pTT27::*hyg* (hygromycin resistant) to HB27*pilA4*::*Km*-pTT27 (transformation-deficient due to the insertion of kanamycin resistance cassette into *pilA4*, a gift from B. Averhoff) strain (C.E. César, unpublished results, Fig. 6) at frequencies up to 10^{-2} transconjugants per recipient cell. In all cases, this conjugation-like mechanism involves transfer of (part of) the pTT27 megaplasmid—which is ubiquitously found as part of the *T. thermophilus* genome, although chromosomal markers have also been shown to be transferred in what was reportedly an HFR-like process [44]—at much lower frequencies than those obtained with plasmid markers (C.E. César, unpublished results).

Transduction. Literature reports on *Thermus* bacteriophages are scarce, and despite the identification of 113 viruses in a single article [54], only a pair of them have been thoroughly studied. This small number of known strains probably correspond to the very tip of an unknown repertoire of thermophilic bacteriophages. A good indication of this misrepresentation is the presence in the HB8 strain of 12 cluster regulatory interspaced short palindromic repeat (CRISPR) systems [1]. CRISPR systems are host genetic modules encoded by several bacteria and most archaea to act against invading genetic elements, such as plasmids, transposons and, particularly, bacteriophages (for a comprehensive review see [2]). The unusually high number of CRISPR systems coded by HB8 supports a scenario of frequent foreign-DNA invasion events. Notably, the transcription profiles of most of the CRISPRS-associated proteins are up-regulated after phage infection [1], supporting phage infection and probably transduction as a fairly active event in *Thermus* spp. strains. It is difficult to acknowledge and quantify LGT by transduction in the genus *Thermus*; nonetheless, transduction is yet another mechanism that likely participates in the extensive LGT shown by *Thermus* spp.

Concluding remarks. The ability of microorganisms to exploit and inhabit different environments relies on the plasticity of prokaryotes in acquiring those genetic determinants that facilitate the colonization of diverse, often extreme, niches. In extremophilic bacteria, LGT has been recognized as a major force in adaptation and diversification, and inter-domain gene flow between *Archaea* and *Bacteria* has been widely reported. The high-temperature habitat is a good example of this phenomenon, and *T. thermophilus*, due to its unique repertoire of LGT strategies, provides an excellent model to study horizontal gene flow in this harsh environment. The enormous LGT promiscuity observed in *Thermus* strains is manifested in the presence of a large megaplasmid as part of the genome, where most of the plasticity observed between *T. thermophilus* strains resides [11]. We believe that *T. thermophilus* megaplasmid acts as a reservoir for non-essential to life adaptation traits (e.g., denitrification and carotenoid biosynthesis clusters, CRISPR antiviral systems, etc.) and that the constant gene flow of these traits as needed is a common characteristic in the genus *Thermus*.

Acknowledgements. This work was supported by grant BIO2010-18875 from the Spanish Ministry of Science and Innovation. An institutional grant from Ramón Areces Foundation to CBMSO is acknowledged. CEC holds a Juan de la Cierva postdoctoral contract from the Spanish Ministry of Science and Innovation. CB and LA are founded by FPI and JAE fellowships from the Ministry of Education and the Spanish National Research Council (CSIC), respectively. We thank Dr. B. Averhoff for sending us the *pilA4* mutant.

Competing interests: None declared.

References

1. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A (2010) Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. J Mol Biol 395:270-281

2. Al-Attar S, Westra ER, van der Oost J, Brouns SJ (2011) Clustered regularly interspaced short palindromic repeats (CRISPRs): the hallmark of an ingenious antiviral defense mechanism in prokaryotes. Biol Chem 392:277-289
3. Albers SV, Pohlschroder M (2009) Diversity of archaeal type IV pilin-like structures. Extremophiles 13:403-410
4. Álvarez-Martínez CE, Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. Microbiol Mol Biol Rev 73:775-808
5. Álvarez L, Bricio C, Gómez MJ, Berenguer J (2011) Lateral transfer of the denitrification pathway genes among *Thermus thermophilus* strains. Appl Environ Microbiol 77:1352-1358
6. Aravind L, Tatusov RL, Wolf YI, Walker DR, Koonin EV (1998) Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. Trends Genet 14:442-444
7. Averhoff B (2009) Shuffling genes around in hot environments: the unique DNA transporter of *Thermus thermophilus*. FEMS Microbiol Rev 33:611-626
8. Bahl H, Scholz H, Bayan N, Chami M, Leblon G, Gulik-Krzywicki T, Shechter E, Fouet A, Mesnage S, Tosi-Couture E, Gounon P, Mock M, Conway de Macario E, Macario AJ, Fernandez-Herrero LA, Olabarria G, Berenguer J, Blaser MJ, Kuen B, Lubitz W, Sara M, Pouwels PH, Kolen CP, Boot HJ, Resch S (1997) Molecular biology of S-layers. FEMS Microbiol Rev 20:47-98
9. Bricio C, Álvarez L, Gómez MJ, Berenguer J (2011) Partial and complete denitrification in *Thermus thermophilus*: lessons from genome drafts. Biochem Soc Trans 39:249-253
10. Bruggemann H, Chen C (2006) Comparative genomics of *Thermus thermophilus*: Plasticity of the megaplasmid and its contribution to a thermophilic lifestyle. J Biotechnol 124:654-661
11. Burkhardt J, Vonck J, Averhoff B (2011) Structure and function of PilQ, a secretin of the DNA transporter from the thermophilic bacterium *Thermus thermophilus* HB27. J Biol Chem 286:9977-9984
12. Cabezón E, de la Cruz F (2006) TrwB: an F(1)-ATPase-like molecular motor involved in DNA transport during bacterial conjugation. Res Microbiol 157:299-305
13. Cava F, de Pedro MA, Schwarz H, Henne A, Berenguer J (2004) Binding to pyruvylated compounds as an ancestral mechanism to anchor the outer envelope in primitive bacteria. Mol Microbiol 52:677-690
14. Cava F, Hidalgo A, Berenguer J (2009) *Thermus thermophilus* as biological model. Extremophiles 13:213-231
15. Claverys JP, Martin B (2003) Bacterial "competence" genes: signatures of active transformation, or only remnants? Trends Microbiol 11:161-165
16. Claverys JP, Prudhomme M, Martin B (2006) Induction of competence regulons as a general response to stress in gram-positive bacteria. Annu Rev Microbiol 60:451-475
17. Chen I, Dubnau D (2004) DNA uptake during bacterial transformation. Nat Rev Microbiol 2:241-249
18. Draper O, César CE, Machón C, de la Cruz F, Llosa M (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. Proc Natl Acad Sci USA 102:16385-16390
19. Fernández-López R, Garcillán-Barcia MP, Revilla C, Lázaro M, Vielva L, de la Cruz F (2006) Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. FEMS Microbiol Rev 30:942-966
20. Friedrich A, Prust C, Hartsch T, Henne A, Averhoff B (2002) Molecular analyses of the natural transformation machinery and identification of pilus structures in the extremely thermophilic bacterium *Thermus thermophilus* strain HB27. Appl Environ Microbiol 68:745-755
21. Friedrich A, Rumszauer J, Henne A, Averhoff B (2003) Pilin-like proteins in the extremely thermophilic bacterium *Thermus thermophilus* HB27: implication in competence for natural transformation and links to type IV pilus biogenesis. Appl Environ Microbiol 69:3695-3700
22. Grohmann E, Muth G, Espinosa M (2003) Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev 67:277-301
23. Hidaka Y, Hasegawa M, Nakahara T, Hoshino T (1994) The entire population of *Thermus thermophilus* cells is always competent at any growth phase. Biosci Biotechnol Biochem 58:1338-1339
24. Israel DA, Lou AS, Blaser MJ (2000) Characteristics of *Helicobacter pylori* natural transformation. FEMS Microbiol Lett 186:275-80
25. Jacob F, Wollman EL (1961) Viruses and genes. Sci Am 204:93-107
26. Jain R, Rivera MC, Moore JE, Lake JA (2003) Horizontal gene transfer accelerates genome innovation and evolution. Mol Biol Evol 20:1598-1602
27. Jakubowski SJ, Krishnamoorthy V, Cascales E, Christie PJ (2004) *Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion system. J Mol Biol 341:961-977
28. Johnsborg O, Eldholm V, Havarstein LS (2007) Natural genetic transformation: prevalence, mechanisms and function. Res Microbiol 158: 767-778
29. Koonin EV, Wolf YI (2008) Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. Nucleic Acids Res 36:6688-6719
30. Kovacs AT, Smits WK, Mironczuk AM, Kuipers OP (2009) Ubiquitous late competence genes in *Bacillus* species indicate the presence of functional DNA uptake machineries. Environ Microbiol 11:1911-1922
31. Koyama Y, Hoshino T, Tomizuka N, Furukawa K (1986) Genetic transformation of the extreme thermophile *Thermus thermophilus* and of other *Thermus* spp. J Bacteriol 166:338-340
32. Lawrence JG (1997) Selfish operons and speciation by gene transfer. Trends Microbiol 5:355-359
33. Lorenz MG, Wackernagel W (1994) Bacterial gene transfer by natural genetic transformation in the environment. Microbiol Rev 58:563-602
34. Maier B, Potter L, So M, Long CD, Seifert HS, Sheetz MP (2002) Single pilus motor forces exceed 100 pN. Proc Natl Acad Sci USA 99:16012-16017
35. Mortier-Barriere I, Velten M, Dupaigne P, Mirouze N, Pietrement O, McGovern S, Fichant G, Martin B, Noirot P, Le Cam E, Polard P, Claverys JP (2007) A key presynaptic role in transformation for a widespread bacterial protein: DprA conveys incoming ssDNA to RecA. Cell 130:824-836
36. Mossey P, Hudacek A, Das A (2010) *Agrobacterium tumefaciens* type IV secretion protein VirB3 is an inner membrane protein and requires VirB4, VirB7, and VirB8 for stabilization. J Bacteriol 192:2830-2838
37. Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Nelson WC, Ketchum KA, McDonald L, Utterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg J, Sutton GG, Fleischmann RD, Eisen JA, White O, Salzberg SL, Smith HO, Venter JC, Fraser CM (1999) Evidence for lateral gene transfer between *Archaea* and *Bacteria* from genome sequence of *Thermotoga maritima*. Nature 399:323-329
38. Nunn DN, Lory S (1991) Product of the *Pseudomonas aeruginosa* gene pilD is a prepilin leader peptidase. Proc Natl Acad Sci USA 88: 3281-3285
39. Ochman H, Davalos LM (2006) The nature and dynamics of bacterial genomes. Science 311:1730-1733
40. Omelchenko MV, Wolf YI, Gaidamakova EK, Matrosova VY, Vasilenko A, Zhai M, Daly MJ, Koonin EV, Makarova KS (2005) Comparative genomics of *Thermus thermophilus* and *Deinococcus radiodurans*: divergent routes of adaptation to thermophily and radiation resistance. BMC Evol Biol 5:57

41. Palmen R,Hellingwerf KJ (1997) Uptake and processing of DNA by *Acinetobacter calcoaceticus*—a review. Gene 192:179-190
42. Pérez-Mendoza D, de la Cruz F (2009) *Escherichia coli* genes affecting recipient ability in plasmid conjugation: are there any? BMC Genomics 10:71
43. Ramírez-Arcos S, Fernández-Herrero LA, Berenguer J (1998) A thermophilic nitrate reductase is responsible for the strain specific anaerobic growth of *Thermus thermophilus* HB8. Biochim Biophys Acta 1396:215-227
44. Ramírez-Arcos S, Fernández-Herrero LA, Marín I, Berenguer J (1998) Anaerobic growth, a property horizontally transferred by an Hfr-like mechanism among extreme thermophiles. J Bacteriol 180:3137-3143
45. Redfield RJ (1993) Genes for breakfast: the have-your-cake-and-eat-it-too of bacterial transformation. J Hered 84:400-404
46. Rumszauer J, Schwarzenlander C, Averhoff B (2006) Identification, subcellular localization and functional interactions of PilMNOWQ and PilA4 involved in transformation competency and pilus biogenesis in the thermophilic bacterium *Thermus thermophilus* HB27. FEBS J 273:3261-3272
47. Ryan FJ, Lederberg J (1946) Reverse-mutation and adaptation in leucineless *Neurospora*. Proc Natl Acad Sci USA 32:163-173
48. Schwarzenlander C, Averhoff B (2006) Characterization of DNA transport in the thermophilic bacterium *Thermus thermophilus* HB27. FEBS J 273: 4210-4218
49. Schwarzenlander C, Haase W, Averhoff B (2009) The role of single subunits of the DNA transport machinery of *Thermus thermophilus* HB27 in DNA binding and transport. Environ Microbiol 11:801-808
50. Smillie C, Garcillán-Barcia MP, Francia MV, Rocha EP, de la Cruz F (2010) Mobility of plasmids. Microbiol Mol Biol Rev 74:434-452
51. Waksman G, Fronzes R (2010) Molecular architecture of bacterial type IV secretion systems. Trends Biochem Sci 35:691-698
52. Wallden K, Rivera-Calzada A, Waksman G (2010) Type IV secretion systems: versatility and diversity in function. Cell Microbiol 12:1203-1212
53. Wozniak RA. Waldor MK (2010) Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat Rev Microbiol 8:552-863
54. Yu MX, Slater MR, Ackermann HW (2006) Isolation and characterization of *Thermus* bacteriophages. Arch Virol 151:663-679
55. Zahrl D, Wagner M, Bischof K, Bayer M, Zavecz B, Beranek A, Ruckenstuhl C, Zarfel GE, Koraimann G (2005) Peptidoglycan degradation by specialized lytic transglycosylases associated with type III and type IV secretion systems. Microbiology 151:3455-3467
56. Zupan J, Hackworth CA, Aguilar J, Ward D, Zambryski P (2007) VirB1* promotes T-pilus formation in the vir-Type IV secretion system of *Agrobacterium tumefaciens*. J Bacteriol 189:6551-6563