A self-transmissible plasmid from a hyperthermophile that facilitates genetic modification of diverse Archaea

Ryan J Catchpole, Valérie Barbe, Ghislaine Magdelenat, Evelyne Marguet, Michael Terns, Jacques Oberto, Patrick Forterre, Violette da Cunha

A self-transmissible plasmid from a hyperthermophile that facilitates genetic modification of diverse Archaea

Ryan J. Catchpole[1,2,3]; Valérie Barbe[4]; Ghislaine Magdelenat[4], Evelyne Marguet[2]; Michael Terns[3]; Jacques Oberto[2]; Patrick Forterre[1,2]; Violette Da Cunha[1,2,4.]

[1]Institut Pasteur, Unité de Biologie Moléculaire du Gène chez les Extrêmophiles (BMGE), Département de Microbiologie F-75015 Paris, France

[2]Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Univ. Paris-Saclay, 91198, Gif-sur-Yvette Cedex, France

[3]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

[4]Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France

Corresponding authors
Ryan J. Catchpole: ryan.catchpole@uga.edu
Violette Da Cunha: vdacunha@genoscope.cns.fr

ABSTRACT

Conjugative plasmids are self-transmissible mobile genetic elements which transfer DNA between host cells via Type IV Secretion Systems (T4SS). While T4SS-mediated conjugation has been well-studied in bacteria, information is sparse in Archaea and known representatives exist only in the Sulfolobales order of Crenarchaeota. Here we present the first self-transmissible plasmid identified in a Euryarchaeon, *Thermococcus sp. 33-3*. The 103 kbp plasmid, pT33-3, is seen in CRISPR-spacers throughout the Thermococcales order. We demonstrate that pT33-3 is a bona fide conjugative plasmid that requires cell-to-cell contact, and is dependent upon canonical, plasmid-encoded T4SS-like genes. Under laboratory conditions pT33-3 transfers to various Thermococcales, and transconjugants propagate at 100°C. Using pT33-3, we developed a genetic toolkit which allows modification of phylogenetically diverse Archaeal genomes. We demonstrate pT33-3–mediated plasmid mobilization and subsequent targeted genome modification in previously untransformable Thermococcales species, and extend this process to interphylum transfer to a Crenarchaeon.

MAIN

Conjugative plasmids (CPs) are self-transmissible mobile genetic elements which encode Type IV Secretion Systems (T4SS) to facilitate their transfer between host cells. The paradigmatic plasmid "F" of *E. coli* has been well-studied for over 70 years[1], and other CPs have been identified throughout both gram-positive and gram-negative bacteria[2,3]. Such plasmids are important epidemiological markers, often carrying virulence factors and drug resistance genes, contributing to pathogenesis[4,5]. Even among non-pathogenic organisms, CPs are important contributors not only to host genome evolution[6], but also to the population dynamics of the mobilome[7,8]. Despite their prevalence in diverse bacteria, CPs have only been observed in one family of Archaea, the Sulfolobales, where it is estimated that 3% of strains harbour such plasmids[9]. These archaeal CPs encode genes homologous to T4SS of bacterial CPs and phylogenetic analysis suggests they arose via inter-domain transfer from a bacterial ancestor (possibly multiple times)[10]. While no CPs have been observed outside of the Sulfolobales family of Crenarchaetoa, bioinformatic analyses have identified homologues of bacterial conjugation machinery (e.g. v*irB4* and *virD4*) in other archaeal genome and plasmid sequences[10], suggesting they may be more widespread than previously appreciated. Indeed, in sequencing the genomes of isolated Thermococcales strains[11], we identified an extrachromosomal element with both *virB4* and *virD4* homologues reminiscent of a bacterial

CP. Here, we show that this plasmid, named pT33-3 after the strain in which it was identified, *Thermococcus sp. 33.3*, is a true CP encoding genes effecting its own transfer. This plasmid represents the first archaeal CP outside of Sulfolobales, here in the phylum Euryarchaeota. As DNA uptake is a major limiting factor in establishing genetic techniques in Archaea, we use the conjugative ability of pT33-3 to develop pT33-3–mediated transformation and genome modification techniques for a variety of hyperthermophilic archaea, including those of highly divergent phyla.

RESULTS

Identification and characterization of pT33-3

We sequenced the complete genome of the archaeal strain *Thermococcus sp.33-3,* a hyperthermophilic euryarchaeon isolated from the East Pacific Ridge (depth 2,330m)[11] and assembled a complete closed circular genome of 2,013,327 bp (ENA accession GCA_946300405). The 16S rRNA of *T. sp.33-3* is identical to that of the published species *T. nautili*[12] and *T. sp.26-2*[13]*,* suggesting they are strains of the same species; however, other common phylogenetic markers are non-identical (*rpoB* at 97% nucleotide identity, *pheS* at 92%), and whole genome alignments reveal significant large-scale rearrangements between these isolates (Extended Data Figure 1), as observed in other Thermococcales species[6]. Phylogenetic reconstruction using RpoB places *T. sp.33-3* in a clade with *T. nautili, T. henrietii*[14] and the as yet unnamed *T. sp.21S9*[15] (Figure 1B). In addition to the circular chromosome, we identified an extra-chromosomal circular plasmid of 103,230 bp (hereafter referred to as pT33-3) (Figure 1A), representing the largest known plasmid in the Thermococcales order. Illumina read depth for the plasmid was approximately equal to that of the chromosome, suggesting pT33-3 exists at single copy number (1.47±0.33). Marker frequency analysis did not reveal the presence of an active replication origin (unsurprising in Thermococcales[16]), though indicators of replication initiation and termination regions (GC skew and keto excess[17]) give clear inflection points at 3.6 kbp and 19.0 kbp (Extended Data Figure 2).

While annotation of archaeal mobile genetic elements (MGEs) remains difficult, we were able to ascribe potential functions to 36/166 genes by a combination of BLAST sequence searching and HHblits homology searching (Supplementary Table). This revealed the presence of many genes which are cogent for a large replicon e.g. toxin-antitoxin pairs (RelBE- and MazEF-family) promoting plasmid stabilisation and infectivity[18], cell

division/septum binding proteins (FtsK- and MinD-like) implicated in plasmid segregation/partitioning[19], and an MCM-family helicase likely involved in plasmid replication[20]. No integrase homologues were identified, and no evidence of integration in the host chromosome was found in long-read sequencing, suggesting pT33-3 is exclusively extra-chromosomal in nature. Analyses also reveal that pT33-3 encodes proteins homologous to VirB4/TrbE and VirD4/TraG (genes *p0019* and *p0132*, respectively) required for plasmid transfer through Type IV Secretion Systems (T4SS) in bacteria[21]. Combined with the detection of many other predicted integral membrane proteins (e.g. *p0012*, *p0071*, *p0101*, *p0150*), we sought to investigate whether pT33-3 is capable of conjugative transfer. If present, this would represent the first self-transmissible plasmid in Archaea outside of the Sulfolobales.

Determination of pT33-3 transmissibility

If pT33-3 is mobile between Thermococcales species/strains (or has been in the past), we predict a record of these encounters to be found in the CRISPR arrays of such cells[22]. Indeed, of the 70 available Thermococcales genome assemblies, 12 strains encode a total of 45 CRISPR spacers potentially targeting pT33-3 at >85% identity (Figure 1A). Additionally, the 45 spacers target different sites on pT33-3 (Figure 1B), indicating multiple spacer acquisition events and thus multiple infections by pT33-3, rather than a single ancestral infection[23].

The potential T4SS-like machinery of pT33-3, combined with its apparent historical promiscuity prompted us to investigate its ability to transfer under laboratory conditions. Although genetic tools have been developed for a few select Thermococcales species[24], *T. sp. 33-3* is not naturally competent for DNA uptake, so we instead developed the plasmid in a *T. kodakarensis* host. However, the presence of a potential anti-pT33-3 CRISPR spacer in *T. kodakarensis* (CRISPR1 spacer 12 with 87% match to *p0132*) required generation of a strain devoid of CRISPR-Cas activity. To achieve this, we disrupted the entire 20 kb cas gene locus and the upstream CRISPR array (Δ*TK0445-TK0464*,ΔCR1) in *T. kodakarensis*[25]. The resulting strain indeed lacked CRISPR-Cas activity against incoming plasmid DNA (Extended Data Figure 3) and was thus used as a host for pT33-3 in subsequent experiments.

Using a combination of triparental mating and homologous recombination (see materials and methods), we successfully generated pT33-3 variants encoding an HMG-CoA over-expression cassette (conferring resistance to mevinolin[26]). Illumina sequencing confirmed the

entire, circular pT33-3 plasmid was present in the *T. kodakarensis* recipient, and sequencing read depth confirmed the approximately single copy number per chromosome (0.86±0.25) observed in the native *T. sp. 33-3* host. In addition, the marker-encoding pT33-3 readily transferred to the naturally competent *P. furiosus* COM1, confirming inter-genus host compatibility of this plasmid. To test transfer of pT33-3, auxotrophic strains of the *T. kodakarensis* host were generated which allowed selection against the donor strain, and selection of recipients containing pT33-3 by selection for prototrophy in combination with mevinolin resistance (Figure 2A). Using this methodology, efficient transfer of pT33-3 was achieved across the Thermococcales order, not only between divergent *Thermococcus* species, but also from *Thermococcus* to *Pyrococcus* and *Palaeococcus* species (Figure 1B). Of the tested species, only *T. kodakarensis* and *P. furiosus* are naturally competent for DNA uptake, thus this transfer of pT33-3 to the 10 other non-competent species represents the first genetic transformation of these isolates. Average pT33-3 transfer efficiencies were approximately 10^{-3}-10^{-4} per donor cell (with recipients provided in ~10-fold excess) (Figure 2). Despite the low efficiency of transfer, when (and only when) appropriate donors and recipients were combined, transfer was always detected at least two orders of magnitude above the detection limit (10^{-7} per donor in most cases). In fact, of the 14 species tested as transfer recipients, only two proved inaccessible – *T. peptonophilus* which is naturally resistant to mevinolin and thus not experimentally suitable; and *T. sp. 26-2* which encodes 11 CRISPR spacers perfectly matching pT33-3, and 17 at 85% identity (Figure 1).

The ability of pT33-3 to transfer from a non-native host (*T. kodakarensis*) to various non-competent recipients, and the inability of shuttle vectors to transfer under identical conditions (Figure 2C) suggests that this mobility is plasmid-encoded, rather than host-encoded (as in other Euryarchaeota where plasmid transfer results from cell fusion[27]). Following identification of virB/D homologues (*p0019* and *p0132*), above, we hypothesised that the transfer of pT33-3 occurs via T4SS-like machinery, relying on this virB/D combination[21]. We successfully deleted either *p0019* (virB4 homologue) or *p0132* (virD4 homologue) from pT33-3 (Extended Data figure 4), indicating that they are not required for plasmid replication or maintenance in hosts. Removal of either completely abolished the ability of pT33-3 to transfer to *T. nautili* or *T. gammatolerans* recipients (Figure 2D). In contrast, the knockout plasmids readily transferred between *T. kodakarensis* strains (likely due to natural competency[28] – Extended Data Figure 5) indicating their ability to replicate in naive hosts is

unperturbed. The requirement of both VirB4 and VirD4 homologues for transmission confirms a conjugative transfer mechanism for pT33-3 using a T4SS-like machinery.

True T4SS-mediated conjugation requires cell-to-cell contact with the formation of a pore or protein complex bridging cells (Figure 2A). To test whether pT33-3 transfer occurs through conjugation or through an alternative plasmid-encoded mechanism (such as virions or plasmidions[29,30]), we used a membrane filter to separate donor and recipient cells. Here, transfer of pT33-3 from *T. kodakarensis* to *T. nautili* and *T. gammatolerans* was completely blocked by the membrane filter (Figure 2). In contrast, plasmid transfer was still observed between strains of the naturally competent *T. kodakarensis*, reaffirming competence-mediated transformation. This confirms that cell-to-cell contact is required for the plasmid-mediated transfer of pT33-3, identical to that of bacterial conjugative plasmids utilising T4SS.

Mobilization of shuttle vectors by pT33-3

A further defining feature of conjugative elements is the use of an *oriT* (origin of transfer) sequence as the initiation point for transfer[31]. In bacterial CPs, *oriT*s are often encoded adjacent to the transfer genes[32]. We identified two candidate *oriT* sequences (hereafter named *oriT1* and *oriT2*, Figure 1) of ~2 kbp due to their proximity to potential transfer genes, being relatively long intergenic spaces, encoding significant potential for stem-loop formation, and being at inflection points of GC-skew (points 3.6 kb and 19.0 kb (Extended Data Figure 2)). We cloned each region into a small pTN1-based replicative vector encoding the mevinolin resistance marker (p*oriT*1/2) (Extended Data Figure 6). We then tested the ability of *T. kodakarensis* donors with and without pT33-3 to mobilise these plasmids to other Thermococcales. Strikingly, *oriT2* facilitated transfer to *T. nautili* and *T. gammatolerans* at a higher frequency than pT33-3 itself (Figure 3, Extended Data Figure 6). The elevated transfer frequency may be due to the smaller size of mobilised plasmids, and the higher copy number relative to pT33-3. This transfer was reliant upon the presence of pT33-3, suggesting mobilization of shuttle vectors occurs through the pT33-3 encoded conjugation apparatus. We were able to recover intact plasmid DNA from cultures of *T. litoralis, T. gammatolerans, T. pacificus,* and *T. sp. 5-4* (Figure 3). These species are not amenable to genetic transformation with purified DNA, demonstrating the utility of this conjugation/mobilization-based approach. We further narrowed this *oriT*-encoding region to a 300 bp sequence which facilitated mobilization (Extended Data Figure 6). Deletion of this 300bp *oriT* region from

pT33-3 decreased transfer efficiency, but did not completely abolish transfer, suggesting some level of *oriT* redundancy.

Evolutionary origin of pT33-3

With the knowledge that pT33-3 encodes true conjugation apparatus including a virB4 homologue, we wished to investigate the evolutionary source of this conjugation machinery. We reconstructed phylogenies for virB4 sequences similar to that described previously[33]. VirB4 is the only ubiquitous protein in all 8 mating-pair formation (MPF) classes of T4SS, and thus can be used to generate phylogenetic trees of T4SS. In such phylogenies, all archaeal VirB4 sequences (almost all encoded chromosomally on integrated conjugative elements) cluster in the MPF_{FATA} class, having been transferred from bacteria[33]. In our tree, p0019 of pT33-3 grouped with other archaeal virB4 sequences; however, p0019 clustered within Crenarchaeotal sequences (Extended Data Figure 7), rather than virB4 homologues of integrated mobile elements in Euryarchaeota. This phylogenetic position suggests transfer of an ancestral pT33-3-like element from Crenarchaeota (or an ancestral lineage) to *T. sp.33-3* (or an ancestor thereof). These analyses reveal a complex evolutionary history of pT33-3 - phylogenies suggest T4SS emerged in ancestral diderm bacteria and were transferred later to ancestral monoderm bacteria[10], then by our analyses were transferred to an ancestor of Crenarchaeota and Thaumarchaeota, with pT33-3 emerging from a further transfer from Crenarchaeota to Euryarchaeota. This suggests that conjugative plasmids such as pT33-3 are involved horizontal gene transfer between distant taxonomic groups of Archaea.

pT33-3 for genomic manipulation of diverse Archaea

We next tested the ability of pT33-3 to transfer non-replicative DNAs with the potential to integrate into recipient genomes. We designed plasmids with a mevinolin resistance cassette flanked by homology arms for allelic exchange with the recipient genome, cloned between tandem *oriT2* sequences. As such, transfer will initiate at the first *oriT2* sequence, passing the allelic exchange cassette to the recipient cell, and terminate at the second *oriT2*[34] (Extended Data Figure 8). The resulting fragment does not encode an origin of replication, and thus will only confer mevinolin resistance to hosts upon integration into the host genome (and simultaneous allelic replacement). Using this approach, we replaced the *pyrF* gene from both *T. litoralis* and *P. abyssi* with a mevinolin-resistance marker. Allelic exchange was confirmed in the resulting colonies by PCR, and these cultures displayed resistance to both mevinolin and 5-fluoroorotic acid (*pyrF* confers sensitivity to 5-FOA) (Figure 4). Hence, pT33-3

mediated mobilization can also be harnessed to carry out genetic modification in non-competent strains.

Additionally, we were able to extend this plasmid delivery/genetic manipulation approach beyond Euryarchaeota to the Crenarchaeal species *Staphylothermus marinus*[35]. Using *T. kodakarensis* as a mobilization donor, we replaced the putative adenosine phosphoribosyltransferase (*apt*) gene (*hpt1*/Smar_0750/ABN69851.1) with a mevinolin-resistance marker from *Sulfolobus islandicus*[36]. Transformants were selected on a combination of 6-methylpurine and mevinolin, and colonies screened by PCR. In this case, spontaneous resistance was more frequent than true allelic exchange, but knockouts were detectable by PCR (Figure 4, Extended Data Figure 9). To our knowledge, this is the first example of inter-phylum conjugation in Archaea, and the first example of genetic modification of *Staphylothermus marinus.*

Conjugation at high temperatures

Finally, we were intrigued by the ability of pT33-3 to transfer to *Pyrococcus abyssi* (maximum growth temperature 102°C[37]) in our experiments, as well as the presence of pT33-3 CRISPR spacers in various Pyrococcus species, suggesting an impressive ability to transfer at very high temperature. We thus sought to test the maximum and minimum temperatures at which we can detect pT33-3 transfer between species. Transconjugants of both *P. abyssi* and *T. litoralis* were isolated following co-culturing at 100°C but not at 105°C; however, this is outside the growth temperature range of the *T. kodakarensis* donor, thus host survival may be the limiting factor rather than thermostability of the pT33-3 transfer machinery. It is difficult to say whether conjugation occurs during the 100°C incubation period, or at some intermediate temperature while cultures warm to 100°C. Regardless, isolation of transconjugants at 100°C sets new upper bounds on the function of known conjugative plasmids. At lower temperatures, transfer was successful at 55°C, but not detected at 50°C. Again, this is outside the growth temperatures of both donor and recipient cells. Such high thermostability makes pT33-3 an ideal candidate for studying the structure of archaeal conjugative apparatus.

DISCUSSION

We demonstrate that pT33-3 is the first identified self-transmissible plasmid in Archaea outside of the Sulfolobales, here in the phylum Euryarchaeota, family Thermococcales.

Transfer of pT33-3 requires cell-to-cell contact and is dependent upon plasmid-encoded homologues of bacterial conjugation machinery, *virB4* and *virD4,* found in Type IV Secretion Systems. Taken together, these results suggest pT33-3 is a true conjugative plasmid transferring between archaeal cells via a plasmid-encoded T4SS-like apparatus.

The abundance of CPs in Sulfolobales, and their absence outside of this family has led to the hypothesis that archaeal CPs are restricted to Sulfolobales[38], likely resulting from an exceedingly rare bacteria-to-archaea transfer event involving an ancestor of modern Sulfolobales[10]. The discovery of pT33-3 shows CPs do exist outside of the Sulfolobales, and in highly divergent archaeal phyla (Sulfolobales being Crenarchaea; Thermococcales being Euryarchaea). Phylogenetic reconstruction of VirB4 evolution suggests that pT33-3 was transferred from an ancestor of Sulfolobales (Extended Data Figure 7), perhaps suggesting pT33-3 reflects a later diversification of archaeal CPs following an original establishment in Sulfolobales. Such interdomain transfers are likely rare - the cell membranes and proteinaceous cell walls of archaea differ greatly from those of bacteria[39] thus it is difficult to imagine bacteria-to-archaea CP transfer through a classical T4SS. These differences in the cell envelopes of bacteria and archaea also indicate archaeal CP transfer apparatus must necessarily differ from bacterial T4SS, despite sharing some homologous proteins. Determining the nature of this archaeal transfer system will be important as our understanding of gene flow in Archaea advances.

Recent studies of the genetics and cell biology of Archaea have revealed important insights into not only this enigmatic domain of life, but into of bacterial and eukaryotic biology. The Thermococcales have proven to be invaluable model organisms for the study of Archaea, as well as the study of hyperthermophiles. There are now 70 Thermococcales genome assemblies available, with most representing culturable isolates. Despite the abundance of data and relative ease of cultivation, genetic tractability of this family has been limited to the few species which are naturally competent for DNA uptake (*T. kodakarensis* KOD1[40]*, P. furiosus* COM1[41]*,* and to a lesser extent *T. onnurineus* NA1[42] *and P. yayanosii* A1[43]). Our development of a conjugation-based DNA transfer system for Thermococcales breaks this transformation barrier, giving the potential to modify *any* Thermococcales isolate, without reliance upon natural-transformation. Furthermore, the apparent promiscuity of the pT33-3 transfer machinery makes this technique extendable to other archaeal families, such as Desulfurococcales, shown here (Figure 4D). This work opens the door for genetic dissection

of previously unexplored archaeal species, and the development of a broad-range archaeal genetic toolkit.

Methods

Culturing

All Thermococcales species, except *T. kodakarensis* and *P. furiosus*, were cultured anaerobically in Thermococcales Rich Medium (TRM)[44] with sulfur at 85°C. *T. kodakarensis* was cultured anaerobically in ASW medium[28] at 85°C. *P. furiosus* was cultured in defined medium described previously[41]. *S. marinus* was cultured anaerobically in 0.65x TRM with sulfur at 85°C. Where appropriate, agmatine was added at a final concentration of 1 mM; mevinolin at 10 µM; 6-methylpurine at 100 µM; 5-FOA at 1.5% w/v. Solid media was prepared by adding Gelrite to a final concentration of 1% w/v, polysulfide solution at 0.2% v/v[26], and starch azure added to a final concentration of 0.15% w/v to aid in colony visualisation (not required for *P. furiosus*)[45]. All manipulations of Thermococcales were carried out under anaerobic conditions.

DNA sequencing and analyses

Total DNA was isolated from cultures by phenol-chloroform extraction as described previously[11]. Sequencing of the *Thermococcus sp. 33-3* strain (ENA accession GCA_946300405) was carried out at Genoscope (Centre National de Sequençage, France), using Illumina MiSeq and Oxford Nanopore MinION. Reads were assembled with NaS (https://github.com/institut-de-genomique/NaS) and corrected with Canu[46].

Mapping of Illumina reads to assembled genomes (as an estimate of copy number) was performed and visualised using Geneious (BioMatters). 16S, RpoB and PheS sequences were each aligned with MAFFT[47]. Whole genomes were aligned with Gepard v2.1[48].

Functional annotations of pT33-3 were predicted using each ORF as a query sequence for BLASTp[49] against the nr database (cutoff E=0.01); HHblits (HHSuite v3)[50] against the UniRef30_2020_06 database (n=2 iterations, cutoff E=0.01).

To analyse CRISPR spacer matches, CRISPR arrays were predicted in all Thermococcales genomes available through NCBI (November 2021) using CRT[51] in Geneious. Spacers were extracted and used as query sequences for BLASTn search against pT33-3 (cutoff parameters of 85% identity over the entire spacer length).

Strain generation

To generate a *T. kodakarensis* strain devoid of CRISPR activity, we deleted the region of the genome comprising *TK0445-TK0464* and adjacent the CRISPR1 array[25] from strain TS559 using techniques described previously[52]. Activity was tested by cloning spacers from each of the 3 *T. kodakarensis* CRISPR arrays[25] in to pLC70[53], and used to transform CRISPR-knockout strains (and parental strains) of *T. kodakarensis* as described previously[40].

To generate a *T. kodakarensis* strain harboring pT33-3, we generated a pTP2-derived plasmid[54] encoding TK0664 (conferring 6-methylpurine sensitivity), and a recombination cassette encoding 1kb homology arms identical to pT33-3 flanking a *pdaD* agmatine marker gene. A triparental mating was carried out by mixing cell pellets from 50mL of early stationary phase *T. sp. 33-3, T. kodakarensis* (trp-) harboring the recombination substrate plasmid, and the *T. kodakarensis* (trp+) CRISPR-null strain. The mixed suspension was spotted on solid TRM media (supplemented with agmatine) and incubated anaerobically (N_2 atmosphere) at 85°C overnight. The following day, plates were chilled to 4°C before suspending the plated cell mass in 200 µL 0.8x ASW, and spreading on solid ASW-CH media containing 6-methylpurine (media devoid of tryptophan to select for recipient cells over recombination donors; containing 6-methylpurine to select against the full recombination substrate plasmid and against *T. sp. 33-3*; devoid of agmatine to select for the presence of the agmatine marker cassette). Colonies were screened by PCR for the presence of pT33-3 with the agmatine marker cassette, grown in liquid culture, and total DNA sequenced as described above.

Once established in *T. kodakarensis,* the marker cassette of pT33-3 was swapped by allelic exchange. Deletion of virB/D homologues, and *oriT* was carried out as described previously for chromosomal genes[52].

Mating assays

25 mL cultures of donor and recipient cells were grown to early stationary phase in appropriate media (see above). Cultures were chilled on ice before pelleting cell mass at 3,500x*g* for 20 minutes at 4°C. Media was decanted and pellets resuspended in 100 µL 0.8x ASW. Appropriate cell suspensions were mixed and spotted on the centre of a solid TRM plate containing agmatine. Where donors and recipients were separated by a membrane filter, donor suspension was first spotted directly on the media then overlaid with a 0.2 µm nitrocellulose filter, atop which the recipient suspension was spotted. Plates were incubated at 85°C (except when testing temperature limits) for 16 hours to allow mating, before chilling to 4°C. Cell mass was scraped from the plate and suspended in 200 µL 0.8x ASW. Serial

dilutions were then plated on media selective for donors, recipients, or transconjugants, and incubated at 85°C to allow colony formation. Colonies were enumerated, screened by PCR and amplicon sequencing by Sanger sequencing (Eurofins Genomics), and/or transferred to liquid media where required.

Mobilization assays

Candidate *oriT* sequences were cloned into a pLC70-derived plasmid encoding only a mevinolin resistance cassette by Gibson Assembly (sequences and plasmid maps of all plasmids available in supplementary data on Dryad). These plasmids were used to transform CRISPR-null *T. kodakarensis* harboring a trp-selectable pT33-3 plasmid. The resulting strains were used as donors in mating assays (as above), with transconjugants selected by resistance to mevinolin (mobilised plasmid marker) and agmatine prototrophy (donors are agmatine auxotrophs).

When mobilised DNA was designed to integrate into recipient genomes, a transfer cassette was cloned as above. The transfer cassette consists of a mevinolin-resistance marker surrounded by 1kb homology arms for pop-in/pop-out allelic exchange with the recipient chromosome, all surrounded by tandem *oriT* sequences. Modified recipients were selected by resistance to mevinolin, screened by PCR, and confirmed by resistance to 5-FOA or 6-methylpurine resulting from gene knockouts. Growth curves generated using MicrobeMeter HighT (Humane Technologies).

Phylogeny

For RpoB phylogenetic reconstruction, sequences aligned with MAFFT[47], alignment gaps removed with BMGE[55], evolutionary model selected with ModelFinder[56], and phylogenetic trees generated with IQ-TREE v1.6.12[57]. Trees were visualised with iTol[58].

For VirB4 phylogenetic reconstruction, sequences of VirB4 proteins from MPF_{FATA} , and MPF_{FA} used previously[10,33] were retrieved from Genbank. Additional archaeal sequences were retrieved by BLASTp search. Sequences of the MPF_{F} class were used as an outgroup. Phylogenetic trees were generated as for RpoB as above. The maximum likelihood tree was constructed using IQ-TREE with the MFP option for model selection. Branch supports were either computed through SH-like aLRT[59] (10,000 replicates) and UFBoot[60] (10,000 replicates).

Data Availability

The genome sequence for *Thermococcus sp. 33-3* has been deposited with ENA number GCA_946300405. Sequences of all plasmids, genomes, and primers used in this study are available on Dryad with doi 10.5061/dryad.2jm63xst9.

Acknowledgements

We thank Nicolas Soler for advice while writing this manuscript. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (grant agreement no.772725, project HiPhore to PF), under the European Union's Seventh Framework Program (grant agreement no. 340440, project EVOMOBIL to PF) and by the National Institutes of Health (R35GM118160 to MT).

Author Contributions

RJC and EM performed the experiments, VB and GM performed the sequencing, RJC, JO, and VDC performed bioinformatic analyses. MT, PF, JO, and VDC supervised aspects of the project and provided essential expert analysis. RJC wrote the manuscript, with input from MT, PF, JO, and VDC.

Competing interests

The authors declare no competing interests.

References

1. Lederberg, J., Cavalli, L. L. & Lederberg, E. M. Sex Compatibility in Escherichia Coli. *Genetics* **37**, 720–730 (1952).
2. Grohmann Elisabeth, Muth Günther & Espinosa Manuel. Conjugative Plasmid Transfer in Gram-Positive Bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 277–301 (2003).
3. De La Cruz, F., Frost, L. S., Meyer, R. J. & Zechner, E. L. Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol. Rev.* **34**, 18–40 (2010).
4. Ramirez, M. S., Traglia, G. M., Lin, D. L., Tran, T. & Tolmasky, M. E. Plasmid-Mediated Antibiotic Resistance and Virulence in Gram-negatives: the Klebsiella pneumoniae Paradigm. *Microbiol Spectr* **2**, 1–15 (2014).
5. Montero, I., Herrero-Fresno, A., Rodicio, R. & Rodicio, M. R. Efficient mobilization of a resistance derivative of pSLT, the virulence plasmid specific of Salmonella enterica serovar Typhimurium, by an IncI1 plasmid. *Plasmid* **70**, 104–109 (2013).
6. Cossu, M. *et al.* Flipping chromosomes in deep-sea archaea. *PLoS Genet.* **13**, e1006847 (2017).
7. Tamminen, M., Virta, M., Fani, R. & Fondi, M. Large-scale analysis of plasmid relationships through gene-sharing networks. *Mol. Biol. Evol.* **29**, 1225–1240 (2012).
8. Basta, T., Smyth, J., Forterre, P., Prangishvili, D. & Peng, X. Novel archaeal plasmid pAH1 and its interactions with the lipothrixvirus AFV1. *Mol. Microbiol.* **71**, 23–34 (2009).
9. Prangishvili, D. *et al.* Conjugation in Archaea: Frequent Occurrence of Conjugative Plasmids inSulfolobus. *Plasmid* **40**, 190–202 (1998).
10. Guglielmini, J., de la Cruz, F. & Rocha, E. P. C. Evolution of conjugation and type IV secretion systems. *Mol. Biol. Evol.* **30**, 315–331 (2013).
11. Lepage, E. *et al.* Molecular diversity of new Thermococcales isolates from a single area

of hydrothermal deep-sea vents as revealed by randomly amplified polymorphic DNA fingerprinting and 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.* **70**, 1277–1286 (2004).

12. Gorlas, A. *et al.* Thermococcus nautili sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal deep-sea vent. *Int. J. Syst. Evol. Microbiol.* **64**, 1802–1810 (2014).
13. Soler, N. *et al.* Two novel families of plasmids from hyperthermophilic archaea encoding new families of replication proteins. *Nucleic Acids Res.* **38**, 5088–5104 (2010).
14. Alain, K. *et al.* Thermococcus henrietii sp. nov., a novel extreme thermophilic and piezophilic sulfur-reducing archaeon isolated from a deep-sea hydrothermal chimney. *Int. J. Syst. Evol. Microbiol.* **71**, (2021).
15. Strang, L. C. Genomic Insights and Ecological Adaptations of Deep-Subsurface and Near Subsurface Thermococcus Isolates. (Western Washington University, 2020).
16. Gehring, A. M. *et al.* Genome Replication in Thermococcus kodakarensis Independent of Cdc6 and an Origin of Replication. *Front. Microbiol.* **8**, 2084 (2017).
17. Freeman, J. M., Plasterer, T. N., Smith, T. F. & Mohr, S. C. Patterns of Genome Organization in Bacteria. *Science* (1998) doi:10.1126/science.279.5358.1827a.
18. Díaz-Orejas, R., Espinosa, M. & Yeo, C. C. The Importance of the Expendable: Toxin-Antitoxin Genes in Plasmids and Chromosomes. *Front. Microbiol.* **8**, 1479 (2017).
19. Guynet, C. & de la Cruz, F. Plasmid segregation without partition. *Mob. Genet. Elements* **1**, 236–241 (2011).
20. Krupovic, M., Gribaldo, S., Bamford, D. H. & Forterre, P. The evolutionary history of archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of mobile genetic elements. *Mol. Biol. Evol.* **27**, 2716–2732 (2010).
21. Grohmann, E., Christie, P. J., Waksman, G. & Backert, S. Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* **107**, 455–471 (2018).

22. Terns, M. P. & Terns, R. M. CRISPR-based adaptive immune systems. *Curr. Opin. Microbiol.* **14**, 321–327 (2011).
23. Pinilla-Redondo, R. *et al.* Type IV CRISPR-Cas systems are highly diverse and involved in competition between plasmids. *Nucleic Acids Res.* (2019) doi:10.1093/nar/gkz1197.
24. Atomi, H., Imanaka, T. & Fukui, T. Overview of the genetic tools in the Archaea. *Front. Microbiol.* **3**, 337 (2012).
25. Elmore, J. R. *et al.* Programmable plasmid interference by the CRISPR-Cas system in Thermococcus kodakarensis. *RNA Biol.* **10**, 828–840 (2013).
26. Matsumi, R., Manabe, K., Fukui, T., Atomi, H. & Imanaka, T. Disruption of a sugar transporter gene cluster in a hyperthermophilic archaeon using a host-marker system based on antibiotic resistance. *J. Bacteriol.* **189**, 2683–2691 (2007).
27. Naor, A. & Gophna, U. Cell fusion and hybrids in Archaea: prospects for genome shuffling and accelerated strain development for biotechnology. *Bioengineered* **4**, 126–129 (2013).
28. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1. *J. Bacteriol.* **185**, 210–220 (2003).
29. Forterre, P., Da Cunha, V. & Catchpole, R. Plasmid vesicles mimicking virions. *Nat Microbiol* **2**, 1340–1341 (2017).
30. Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M. J. & Cavicchioli, R. A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. *Nat Microbiol* **2**, 1446–1455 (2017).
31. Willetts, N. & Wilkins, B. Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48**, 24–41 (1984).
32. Zatyka, M. & Thomas, C. M. Control of genes for conjugative transfer of plasmids and

other mobile elements. *FEMS Microbiol. Rev.* **21**, 291–319 (1998).

33. Guglielmini, J. *et al.* Key components of the eight classes of type IV secretion systems involved in bacterial conjugation or protein secretion. *Nucleic Acids Res.* **42**, 5715–5727 (2014).
34. Draper, O., César, C. E., Machón, C., de la Cruz, F. & Llosa, M. Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 16385–16390 (2005).
35. Fiala, G., Stetter, K. O., Jannasch, H. W., Langworthy, T. A. & Madon, J. Staphylothermus marinus sp. nov. Represents a Novel Genus of Extremely Thermophilic Submarine Heterotrophic Archaebacteria Growing up to 98 °C. *Syst. Appl. Microbiol.* **8**, 106–113 (1986).
36. Zhang, C., Cooper, T. E., Krause, D. J. & Whitaker, R. J. Augmenting the genetic toolbox for Sulfolobus islandicus with a stringent positive selectable marker for agmatine prototrophy. *Appl. Environ. Microbiol.* **79**, 5539–5549 (2013).
37. Erauso, G. *et al.* Pyrococcus abyssi sp. nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. *Arch. Microbiol.* **160**, 338–349 (1993).
38. Wagner, A. *et al.* Mechanisms of gene flow in archaea. *Nat. Rev. Microbiol.* **15**, 492–501 (2017).
39. Albers, S.-V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–426 (2011).
40. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon Thermococcus kodakaraensis. *Appl. Environ. Microbiol.* **71**, 3889–3899 (2005).
41. Lipscomb, G. L. *et al.* Natural competence in the hyperthermophilic archaeon Pyrococcus furiosus facilitates genetic manipulation: construction of markerless deletions

of genes encoding the two cytoplasmic hydrogenases. *Appl. Environ. Microbiol.* **77**, 2232–2238 (2011).

42. Kim, M.-S. *et al.* CO-dependent H2 production by genetically engineered Thermococcus onnurineus NA1. *Appl. Environ. Microbiol.* **79**, 2048–2053 (2013).
43. Li, X. *et al.* Genetic tools for the piezophilic hyperthermophilic archaeon Pyrococcus yayanosii. *Extremophiles* **19**, 59–67 (2015).
44. Zeng, X. *et al.* Pyrococcus CH1, an obligate piezophilic hyperthermophile: extending the upper pressure-temperature limits for life. *ISME J.* **3**, 873–876 (2009).
45. Akpan, I., Bankole, M. O. & Adesemowo, A. M. A rapid plate culture method for screening of amylase producing micro-organisms. *Biotechnol. Tech.* **13**, 411–413 (1999).
46. Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).
47. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
48. Krumsiek, J., Arnold, R. & Rattei, T. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* **23**, 1026–1028 (2007).
49. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
50. Steinegger, M. *et al.* HH-suite3 for fast remote homology detection and deep protein annotation. *bioRxiv* 560029 (2019) doi:10.1101/560029.
51. Bland, C. *et al.* CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* **8**, 209 (2007).
52. Santangelo, T. J., Cubonová, L. & Reeve, J. N. Thermococcus kodakarensis genetics: TK1827-encoded beta-glycosidase, new positive-selection protocol, and targeted and repetitive deletion technology. *Appl. Environ. Microbiol.* **76**, 1044–1052 (2010).

53. Santangelo, T. J., Cubonová, L. & Reeve, J. N. Shuttle vector expression in Thermococcus kodakaraensis: contributions of cis elements to protein synthesis in a hyperthermophilic archaeon. *Appl. Environ. Microbiol.* **74**, 3099–3104 (2008).
54. Catchpole, R., Gorlas, A., Oberto, J. & Forterre, P. A series of new E. coli-Thermococcus shuttle vectors compatible with previously existing vectors. *Extremophiles* **22**, 591–598 (2018).
55. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol. Biol.* **10**, 210 (2010).
56. Kalyaanamoorthy, S., Minh, B. Q. & Wong, T. K. F. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature* (2017).
57. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2014).
58. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242-5 (2016).
59. Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
60. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).

FIGURE LEGENDS

Figure 1

Overview of pT33-3 and evidence for CRISPR spacer targeting. A) Plasmid map of pT33-3, with predicted ORFs indicated by grey block arrows. *virB4* and *virD4* homologues highlighted in blue; potential origins of transfer highlighted in orange. CRISPR spacers of Thermococcales are indicated by black arrowheads on the interior.

B) Phylogenetic tree of Thermococcales species used/described in this work. Phylogeny based on RpoB with Methanococcales as an outgroup. Number of CRISPR spacers targeting pT33-3 and ability to receive pT33-3 under laboratory conditions is indicated (ND indicates not determined).

Figure 2

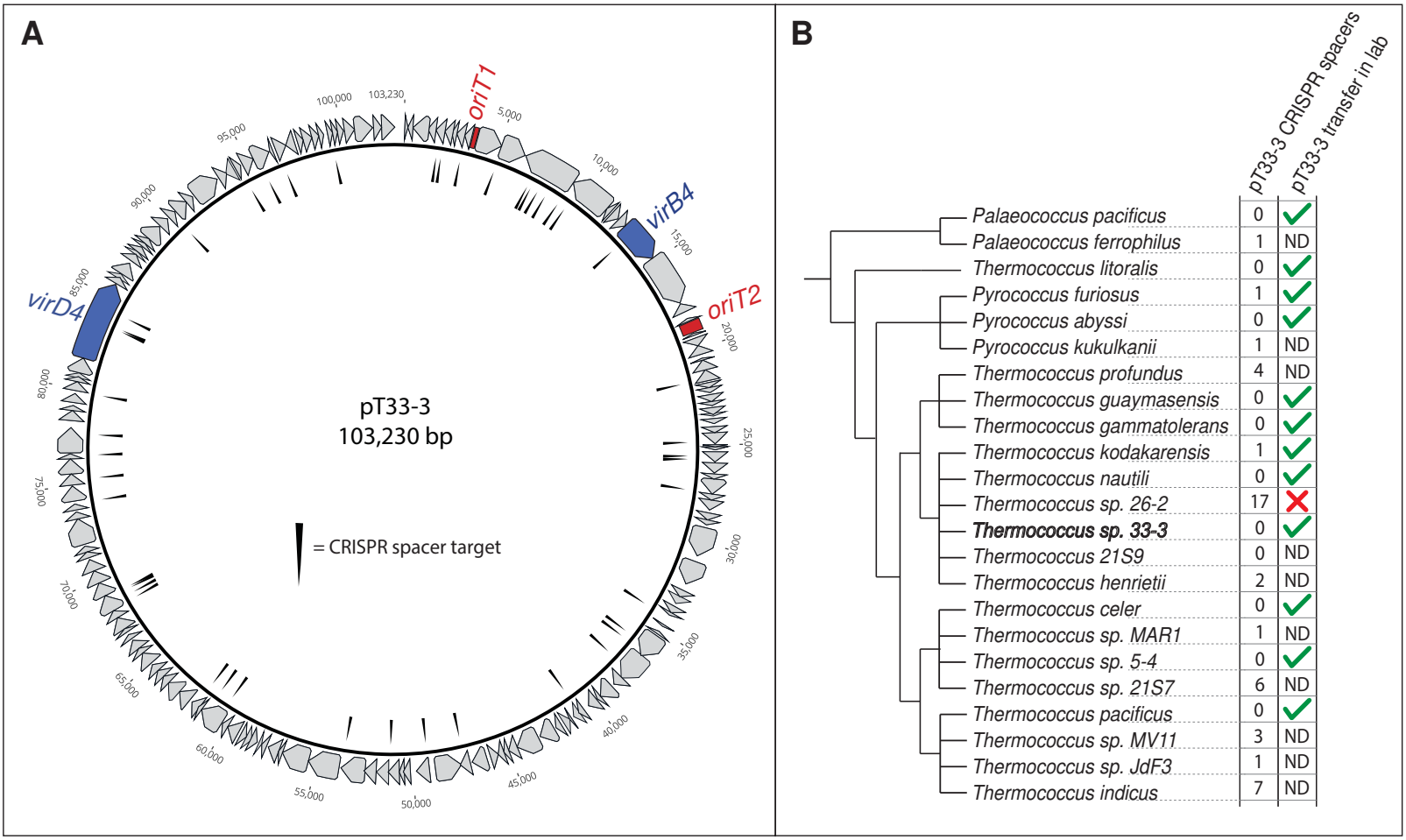
Interspecies transfer of pT33-3. A) In bacterial CPs, the plasmid-encoded T4SS forms a pore bridging the cytoplasm of donor and recipient cells through which the conjugative plasmid is actively transferred. Phenotypes selected in this work are indicated. B) pT33-3 readily transfers from *T. kodakarensis* donor to species not naturally competent for DNA uptake. Individual points indicate conjugation efficiency in a single biological replicate. C) pT33-3 transfers from *T. kodakarensis* to *T. nautili*; non-conjugative shuttle vector (pLC70) does not, indicating transfer ability is plasmid-encoded (not host-encoded). D) Transfer of pT33-3 to non-competent species requires homologues of known T4SS proteins, p0019 (VirB4 homologue), and p0132 (VirD4 homologue). Transfer also requires cell-to-cell contact, being blocked by a 0.2 μm filter.

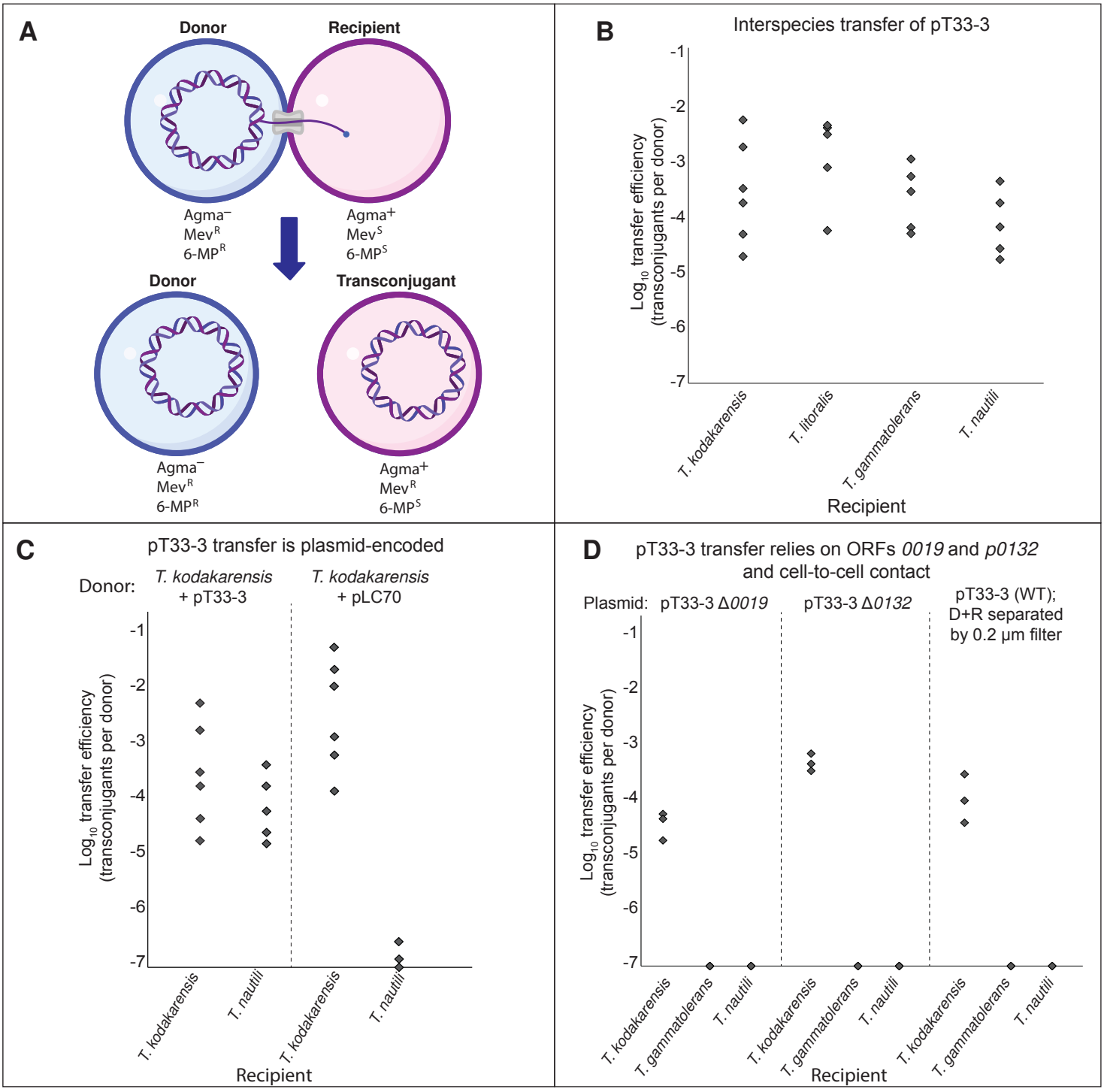
Figure 3

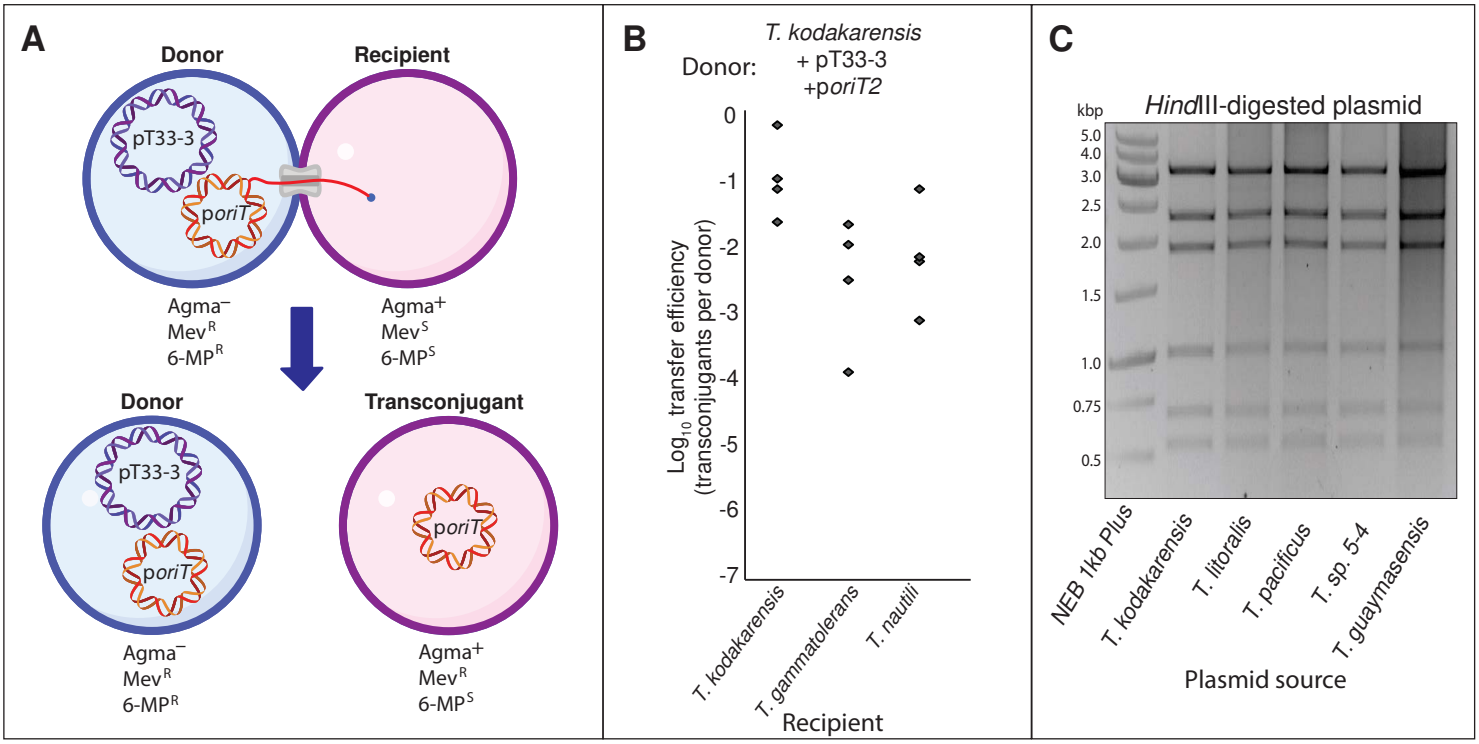
pT33-3 mediated mobilization of shuttle vectors. A) T4SSs encoded by conjugative plasmids mobilize other DNAs encoding their cognate origin of transfer (*oriT*). Phenotypes selected in this work are indicated. B) A shuttle vector encoding pT33-3 *oriT* (p*oriT2*) is mobilized to non-competent donor species in the presence of pT33-3. C) Plasmid DNA isolated from transconjugants has identical *Hind*III digestion patterns to the that of the donor strain (*T. kodakarensis*). Representative data from a single experiment shown, but identical results were obtained in 4 replicate experiments.

Figure 4

pT33-3 mediated genome modification. A) pT33-3 *oriT* allows mobilisation of non-replicative DNA to recipient organisms. Here, homology arms allow allelic exchange with the chromosome, facilitating genetic modification of recipients. Phenotypes selected in this work are indicated. B) Deletion of *pyrF* from *T. litoralis* via pT33-3 mediated mobilisation. Top: diagnostic PCR for deletion shows loss of WT gene and replacement with a pop-in/pop-out MevR cassette; Lower: Resulting strains form colonies on 5-FOA, as expected for a Δ*pyrF* genotype. C) As per B, but with a *P. abyssi* recipient. Growth curves show resistance to 5-FOA. D) *Staphylothermus marinus*, a Crenarchaeon, receives DNA designed to replace the *apt* gene with a MevR cassette. Top: Diagnostic PCR of knockout. Lower: Phylogenetic tree based on 16S rDNA, indicating phylogenetic distance between *T. kodakarensis* donor (Thermococci, labelled red), and *S. marinus* recipient (Desulfurococcales, labelled blue). Representative data from a single experiment shown for B-D, but identical knockouts obtained in 3 biological replicates.


A
oriT1
virB4
oriT2
virD4
pT33-3
103,230 bp
= CRISPR spacer target
B
pT33-3 CRISPR spacers
pT33-3 transfer in lab
Palaeococcus pacificus 0
Palaeococcus ferrophilus 1 ND
Thermococcus litoralis 0
Pyrococcus furiosus 1
Pyrococcus abyssi 0
Pyrococcus kukulkanii 1 ND
Thermococcus profundus 4 ND
Thermococcus guaymasensis 0
Thermococcus gammatolerans 0
Thermococcus kodakarensis 1
Thermococcus nautili 0
Thermococcus sp. 26-2 17
Thermococcus sp. 33-3 0
Thermococcus 21S9 0 ND
Thermococcus henrietii 2 ND
Thermococcus celer 0
Thermococcus sp. MAR1 1 ND
Thermococcus sp. 5-4 0
Thermococcus sp. 21S7 6 ND
Thermococcus pacificus 0
Thermococcus sp. MV11 3 ND
Thermococcus sp. JdF3 1 ND
Thermococcus indicus 7 ND


A
Donor
Recipient
Agma⁻
Mev^R
6-MP^R
Agma+
Mev^S
6-MP^S
Donor
Transconjugant
Agma⁻
Mev^R
6-MP^R
Agma+
Mev^R
6-MP^S
B
Interspecies transfer of pT33-3
Log10 transfer efficiency (transconjugants per donor)
-1
-2
-3
-4
-5
-6
-7
T. kodakarensis
T. litoralis
T. gammatolerans
T. nautili
Recipient
C
pT33-3 transfer is plasmid-encoded
Donor:
T. kodakarensis + pT33-3
T. kodakarensis + pLC70
Log10 transfer efficiency (transconjugants per donor)
-1
-2
-3
-4
-5
-6
-7
T. kodakarensis
T. nautili
T. kodakarensis
T. nautili
Recipient
D
pT33-3 transfer relies on ORFs 0019 and p0132 and cell-to-cell contact
Plasmid:
pT33-3 Δ0019
pT33-3 Δ0132
pT33-3 (WT); D+R separated by 0.2 µm filter
Log10 transfer efficiency (transconjugants per donor)
-1
-2
-3
-4
-5
-6
-7
T. kodakarensis
T. gammatolerans
T. nautili
T. kodakarensis
T. gammatolerans
T. nautili
T. kodakarensis
T. gammatolerans
T. nautili
Recipient


A
Donor
Recipient
pT33-3
oriT
Agma−
Mev^R
6-MP^R
Agma+
Mev^S
6-MP^S
Donor
Transconjugant
pT33-3
oriT
oriT
Agma−
Mev^R
6-MP^R
Agma+
Mev^R
6-MP^S
B
T. kodakarensis
Donor: + pT33-3
+oriT2
Log10 transfer efficiency
(transconjugants per donor)
0
-1
-2
-3
-4
-5
-6
-7
T. kodakarensis
T. gammatolerans
T. nautili
Recipient
C
kbp
HindIII-digested plasmid
5.0
4.0
3.0
2.5
2.0
1.5
1.0
0.75
0.5
NEB 1kb Plus
T. kodakarensis
T. litoralis
T. pacificus
T. sp. 5-4
T. guaymasensis
Plasmid source

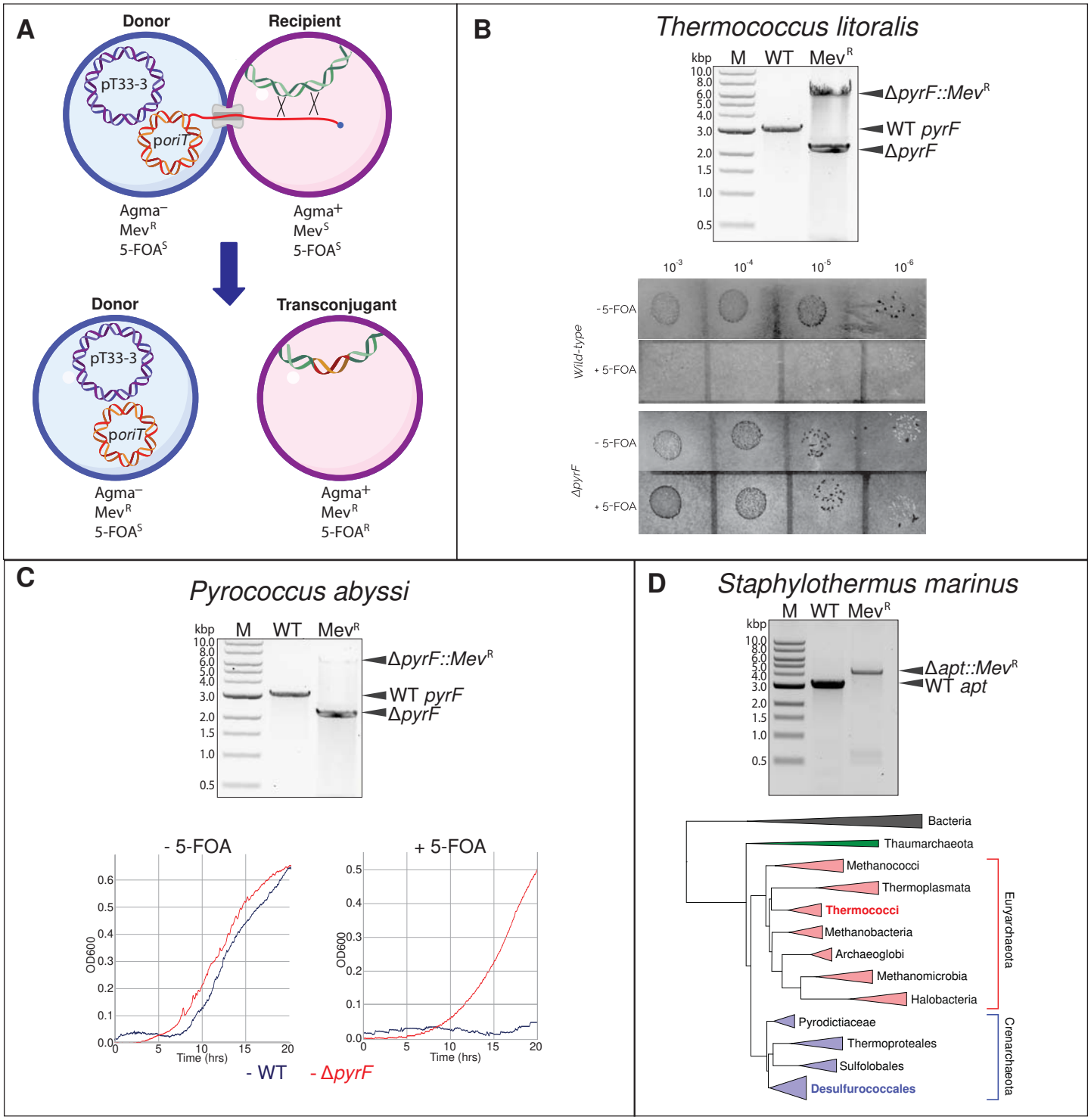

A
Donor
Recipient
pT33-3
oriT
Agma⁻
Mev^R
5-FOA^S
Agma+
Mev^S
5-FOA^S
Donor
Transconjugant
pT33-3
oriT
Agma⁻
Mev^R
5-FOA^S
Agma+
Mev^R
5-FOA^R
B
Thermococcus litoralis
kbp
M
WT
Mev^R
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.0
0.5
ΔpyrF::Mev^R
WT pyrF
ΔpyrF
10^-3
10^-4
10^-5
10^-6
- 5-FOA
+ 5-FOA
Wild-type
ΔpyrF
C
Pyrococcus abyssi
kbp
M
WT
Mev^R
ΔpyrF::Mev^R
WT pyrF
ΔpyrF
- 5-FOA
+ 5-FOA
OD600
Time (hrs)
- WT
- ΔpyrF
D
Staphylothermus marinus
M
WT
Mev^R
kbp
ΔapT::Mev^R
WT apt
Bacteria
Thaumarchaeota
Methanococci
Thermoplasmata
Thermococci
Methanobacteria
Archaeoglobi
Methanomicrobia
Halobacteria
Pyrodictiaceae
Thermoproteales
Sulfolobales
Desulfurococcales
Euryarchaeota
Crenarchaeota

Figure #	Figure title One sentence only	Filename This should be the name the file is saved as when it is uploaded to our system. Please include the file extension. i.e.: *Smith_ED_Fig1.jpg*	Figure Legend If you are citing a reference for the first time in these legends, please include all new references in the main text Methods References section, and carry on the numbering from the main References section of the paper. If your paper does not have a Methods section, include all new references at the end of the main Reference list.
Extended Data Fig. 1	Similarity of *T. sp. 33-3* to other species.	Ex_data_fig_1.eps	Genome alignment between *Thermococcus. sp. 33-3* and closely related strains or species, *T. nautili* (A); *T. henrietii* EXT12c (B); and *T. sp. 26-2* (C). Dots (which form into lines) indicate sequence identity between corresponding regions of the chromosome for the two indicated species. Discontinuities indicate indels; direction changes e.g. change from pointing north-east to south-east, indicate large-scale inversions.
Extended Data Fig. 2	Bioinformatic identification of pT33-3 origin of replication and transfer	Ex_data_fig_2.eps	Plot of cumulative GC skew [(G-C)/(G+C)] and keto excess [(G+T)-(C+A)] for pT33-3. Sharp inflection points are often indicative of origins of replication, origins of transfer, or replication termination sites. These sites are indicated by vertical dashed lines.
Extended Data Fig. 3	Interference assays with CRISPR WT and null strains of *T. kodakarensis*	Ex_data_fig_3.eps	Diamonds indicate data from a single biological replicate (for WT with target plasmid, overlapping points obscure data, n=3). Wild-type (TS559) *T. kodakarensis* is unable to be transformed by plasmids encoding a sequence complementary to spacers in its CRISPR array (target plasmid). In contrast, non-target plasmids transform at ~100cfu/µg DNA. Deletion of genes encoding all CRISPR-associated (Cas) proteins abolishes this targeting activity, restoring transformation with target-encoding plasmids.
Extended Data Fig. 4	Deletion of genes homologous to bacterial T4SS	Ex_data_fig_4.eps	PCR of knockouts for predicted transfer genes, *p0019* and *p0132*. PCR was carried out using primers binding to pT33-3 outside the homology arms used in pop-in/pop-out recombination
Extended Data Fig. 5	Transfer rates of genetic markers between *T. kodakarensis* strains does not require conjugation	Ex_data_fig_5.eps	A) *T. kodakarensis* incubated with purified plasmid DNA (presented as transformants per fg DNA for comparable scale) - *T. kodakarensis* is naturally competent for DNA uptake. B) Transfer of a non-conjugative plasmid from a *T. kodakarensis* donor to a plasmid-free recipient – plasmids transfer between *T. kodakarensis* strains occurs

			by simple co-culturing. C) Transfer of a chromosomal prototrophic marker from a *T. kodakarensis* donor to an auxotrophic recipient - chromosomal markers transfer between *T. kodakarensis* strains. D) Transfer of a chromosomal prototrophic marker from a *T. gammatolerans* donor to an auxotrophic *T. kodakarensis* recipient - chromosomal markers are unable to transfer between *T. gammatolerans* and *T. kodakarensis,* suggesting allelic exchange is mediated by homologous recombination. E) Transfer of a chromosomal prototrophic marker from a *T. kodakarensis* donor to an auxotrophic recipient where the marker is encoded at a locus encoding a second prototrophic marker – transfer chromosomal markers requires a suitable receptive genomic site (non-essential). Individual biological replicates (-, n=3) are presented with average and standard deviation indicated by error bars. G-I) In contrast to *T. kodakarensis, T. nautili* is unable to receive a shuttle vector by co-culturing, whereas pT33-3 readily transfers (see main text and Figure 2C).
Extended Data Fig. 6	Identification of *oriT*-encoding region of pT33-3.	Ex_data_fig_6.eps	1.5kb regions surrounding GC-skew/keto excess minima/maxima were cloned into a shuttle vector and transfer observed in the presence of pT33-3 from *T. kodakarensis* donors. A) *oriT*1 candidate region. While plasmids encoding this region readily transferred between *T. kodakarensis* strains, minimal transfer was observed to *T. nautili* and *T. gammatolerans* recipients. B) *oriT*2 candidate region. Plasmids encoding this region readily transferred from *T. kodakarensis* to *T. kodakarensis, T. nautili,* and *T. gammatolerans,* indicating this region encodes the *oriT* of pT33.3. C) *oriT*2 region was split into four overlapping fragments. Plasmids encoding both *oriT*2.1 and *oriT*2.2 regions transferred to the non-competent recipients, whereas *oriT*2.3 and *oriT*2.4 did not, indicating the *oriT* is encoded by the overlap region between *oriT*2.1 and *oriT*2.2. D) The 300bp overlap between *oriT*2.1 and *oriT*2.2 (named *oriT*300) also confers mobilization ability to plasmids, indicating that this region encodes the *oriT* of pT33.3.

Extended Data Fig. 7	Phylogeny of VirB4	Ex_data_fig_7.eps	A) Figure adapted from Figure 6 of Guglielmini et al. 2014 (doi: 10.1093/nar/gku194) where it was proposed that all archaeal VirB4 sequences arise from a transfer from bacteria within the MPF_{FATA} group. B) Re-creation of MPF_{FATA} and MPF_{FA} phylogeny including VirB4 sequences from integrated elements in other archaeal genomes, and the VirB4 homologue from pT33-3. Phylogeny rooted with MPF_{F} outgroup. pT33-3 VirB4 groups within a clade of Crenarchaeal sequences, suggesting pT33-3 arose from an inter-phylum transfer. Supported branches, computed by aLRT >80 or UFBoot >95, are indicated by dots at nodes (the full tree is provided as an extended data file). The scale bar indicates the average number of substitutions per site.
Extended Data Fig. 8	Strategy employed for allelic exchange in non-competent recipient cells	Ex_data_fig_8.eps	A) Schematic diagram of mobilizable plasmid (pMob) encoding pT33-3 *oriT* sequences and homologous recombination + selectable marker cassette for allelic exchange. B) Transfer of pMob initiating at *oriT*1 and terminating at *oriT*2 results in transfer of a non-replicative DNA encoding recombination/marker cassette. This can be a substrate for homologous recombination and allelic exchange with the recipient chromosome. C) Transfer of pMob initiating at *oriT*2 and terminating at *oriT*1 results in transfer of a replicative DNA without any selectable marker. Transformant selection on drug-containing media renders these cells non-viable.
Extended Data Fig. 9	PCR screen of *S. marinus* colonies following conjugation mediated mobilisation of a recombination substrate	Ex_data_fig_9.eps	Eight colonies were screened using primers binding to the *S. marinus* chromosome, outside the homology arms used in allelic exchange. Clean allelic exchange (Δ*apt*::SimR) was observed in 2/8 colonies. Data from a representative replicate is shown, but similar data obtained on 3 occasions.

11

Item	Present?	Filename This should be the name the file is saved as when it is uploaded to our system,	A brief, numerical description of file contents.

		and should include the file extension. The extension must be .pdf	
Supplementary Information	No		
Reporting Summary	Yes	Catchpole Flat RS.pdf	
Peer Review Information	No	*OFFICE USE ONLY*	

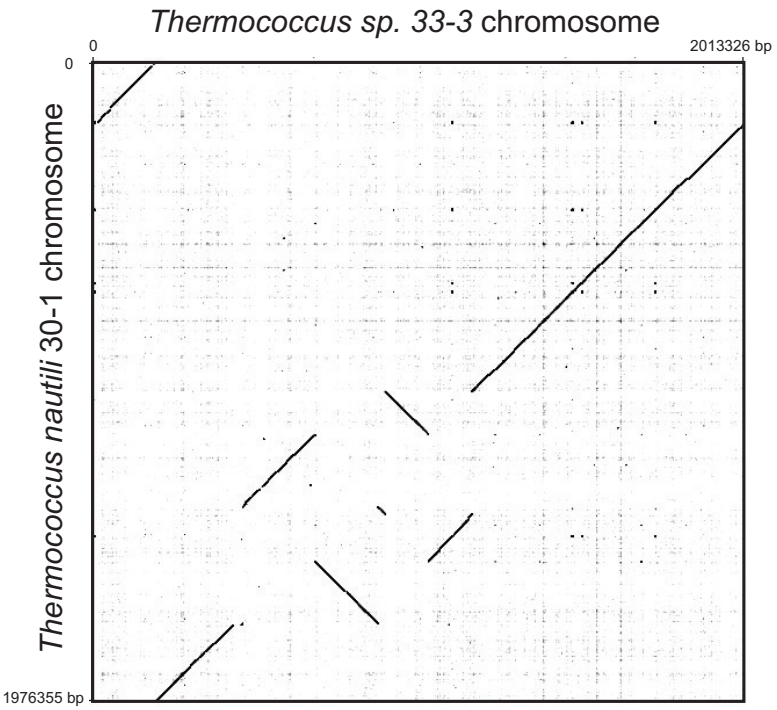
12

Type	Number If there are multiple files of the same type this should be the numerical indicator. i.e. "1" for Video 1, "2" for Video 2, etc.	Filename This should be the name the file is saved as when it is uploaded to our system, and should include the file extension. i.e.: *Smith_ Supplementary_Video_1.mov*	Legend or Descriptive Caption Describe the contents of the file
Supplementary Table		Supplementary Table 1.xlsx	Bioinformatic prediction of pT33-3 ORF function. A combination of BLAST and HHblits was used to predict functions for each ORF in pT33-3. BLAST was performed against the non-redundant protein database. Hits with E-value<0.01 are highlighted in yellow. HHblits was performed against the Uniprot database; however, the majority of hits returned 'hypothetical protein' or 'uncharacterised protein'. Thus, hits were further limited to those with functional annotations.

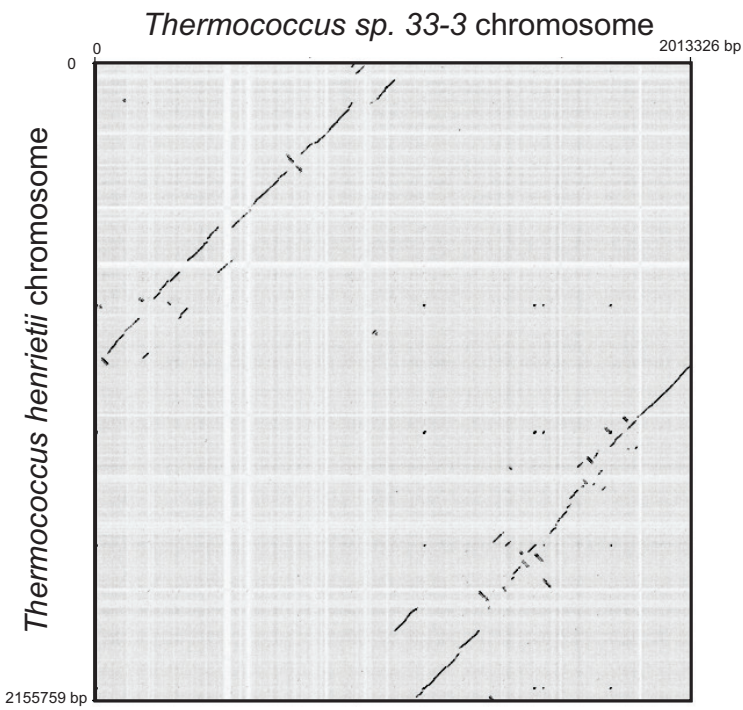
13

Parent Figure or Table	Filename This should be the name the file is saved as when it is uploaded to our system, and should include the file extension. i.e.:	Data description i.e.: Unprocessed Western Blots and/or gels, Statistical Source Data, etc.

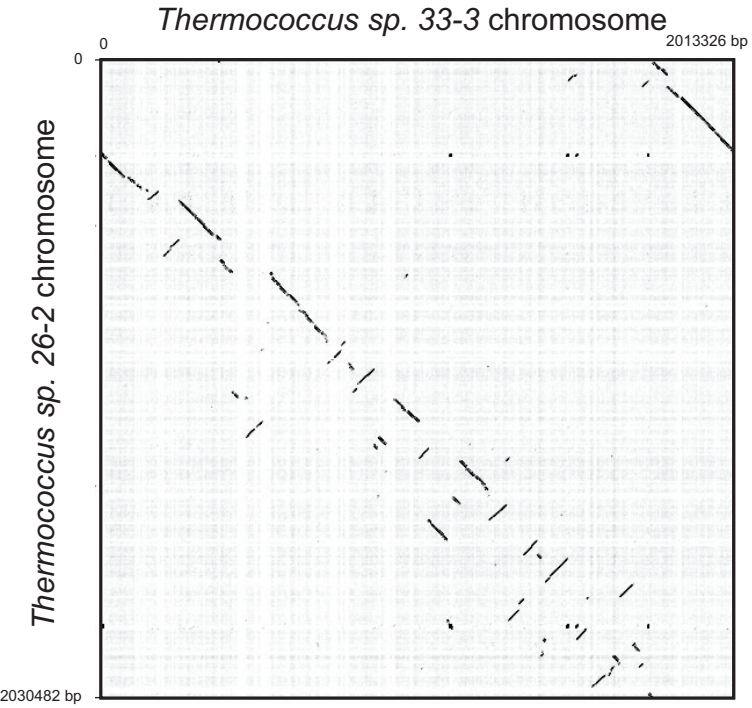
A

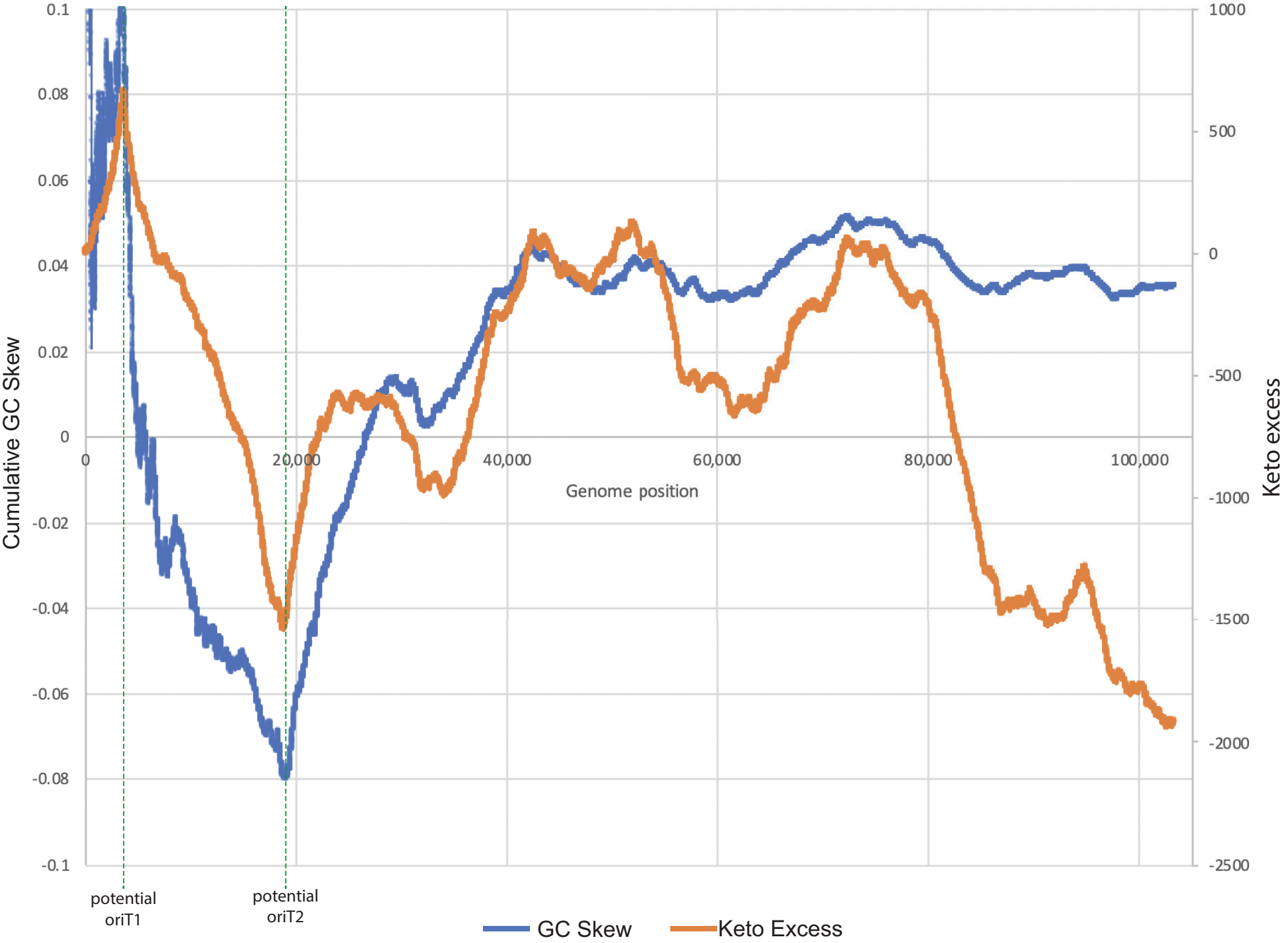

Thermococcus sp. 33-3 chromosome
0
2013326 bp
0
Thermococcus nautili 30-1 chromosome
1976355 bp

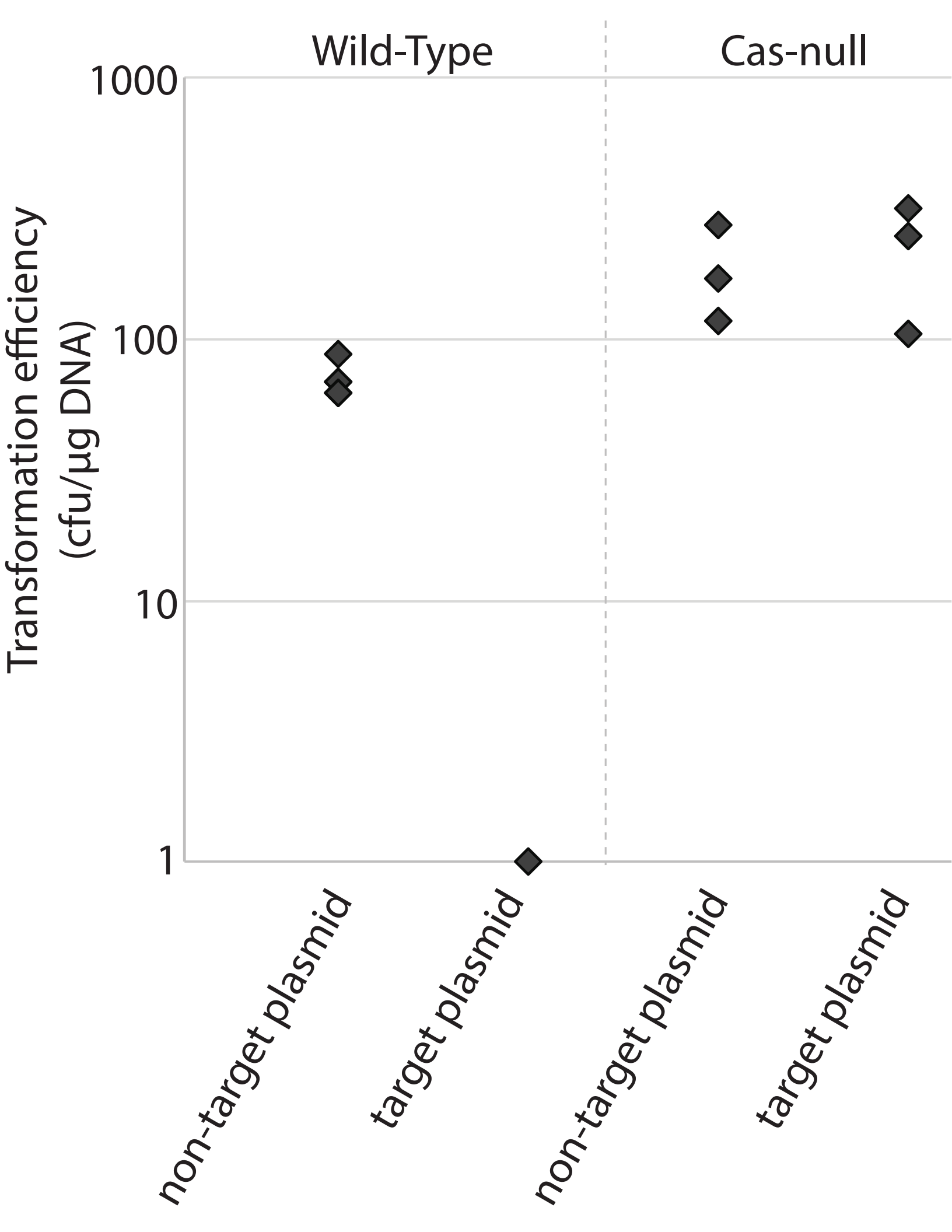
B

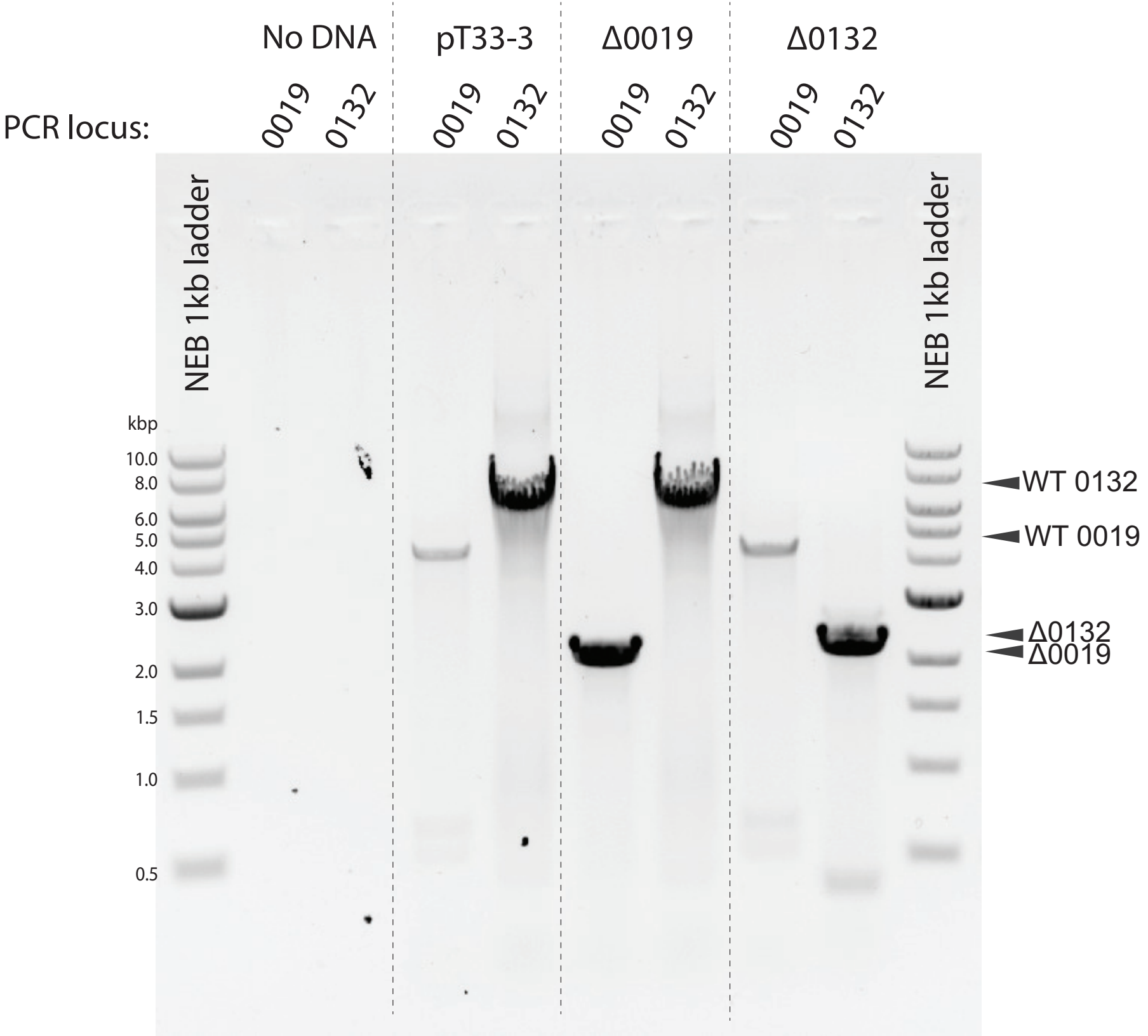

Thermococcus sp. 33-3 chromosome
0
2013326 bp
0
Thermococcus henrietii chromosome
2155759 bp

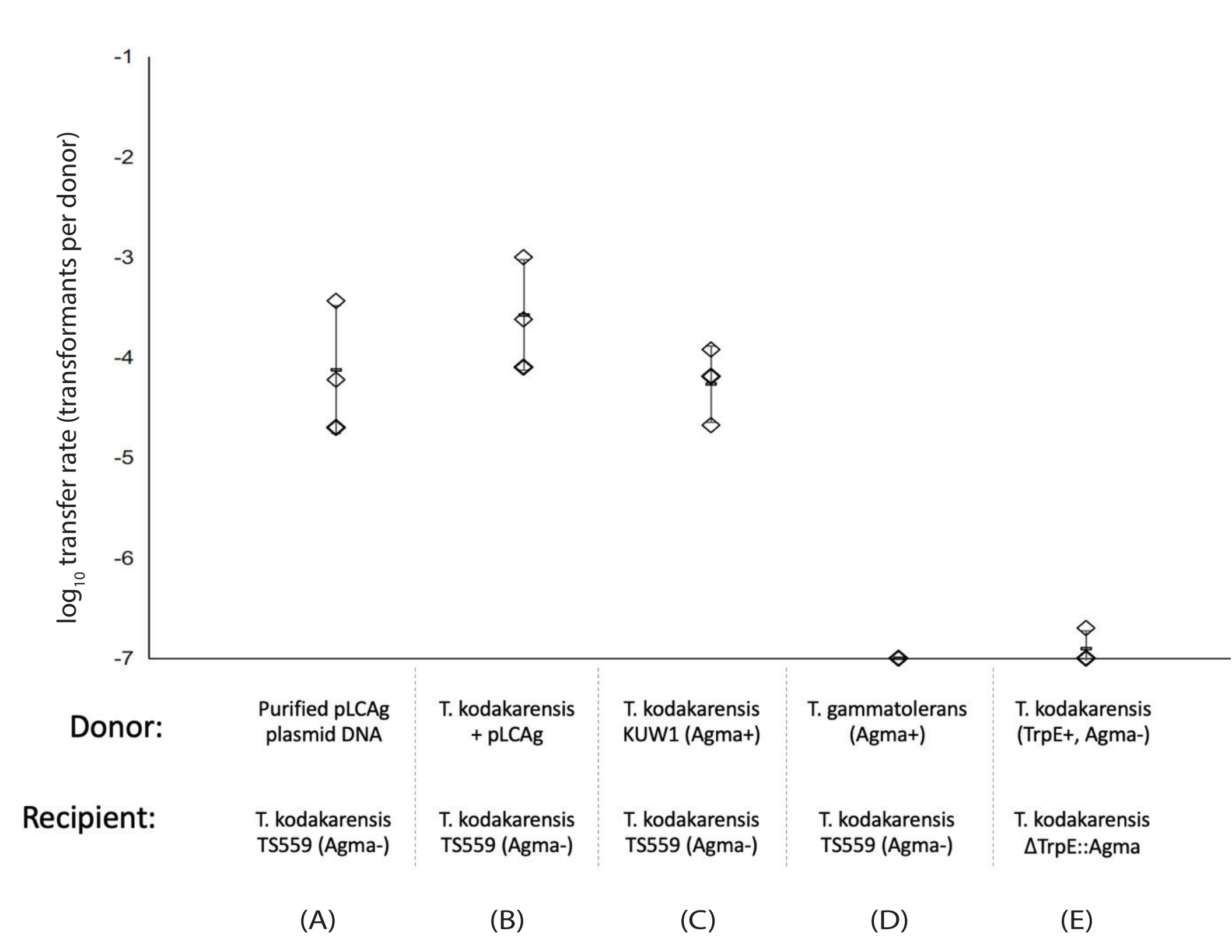
C

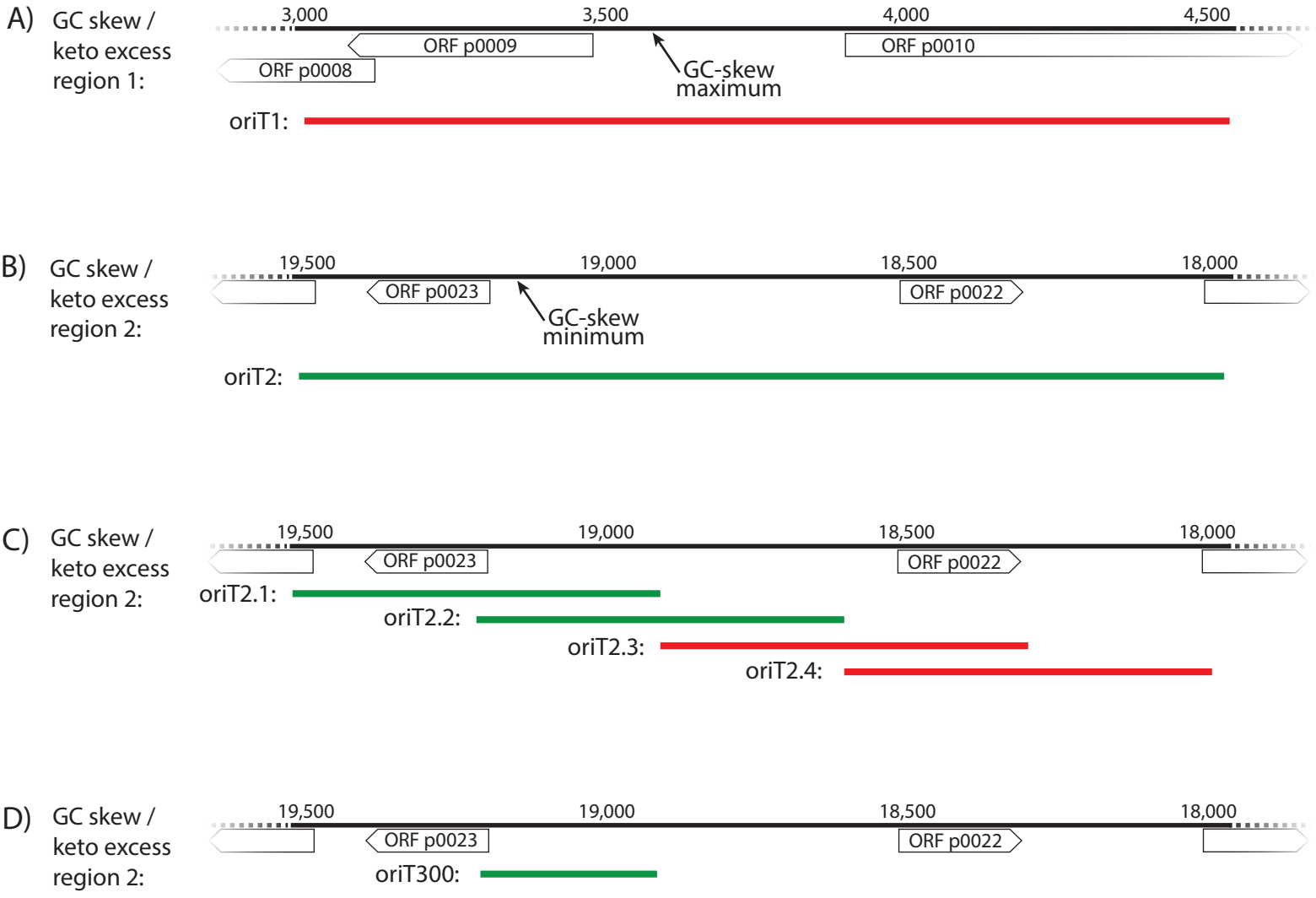

Thermococcus sp. 33-3 chromosome
0
2013326 bp
0
Thermococcus sp. 26-2 chromosome
2030482 bp

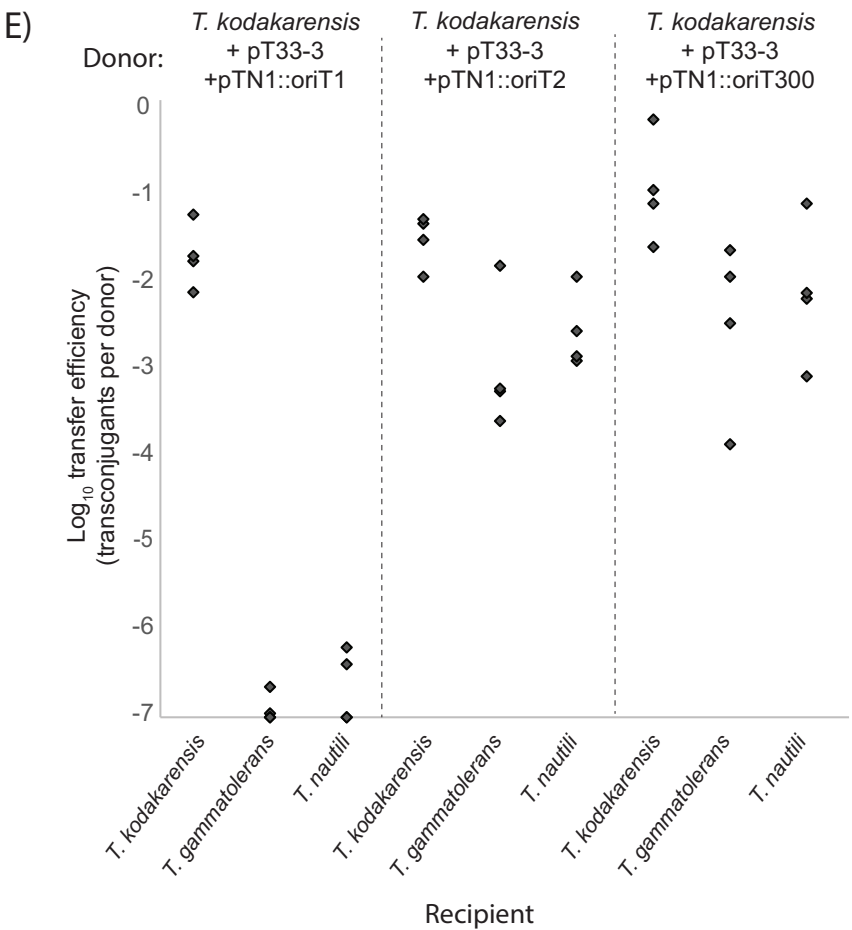

0.1
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
Cumulative GC Skew
1000
500
0
-500
-1000
-1500
-2000
-2500
Keto excess
0
20,000
40,000
60,000
80,000
100,000
Genome position
potential oriT1
potential oriT2
GC Skew
Keto Excess

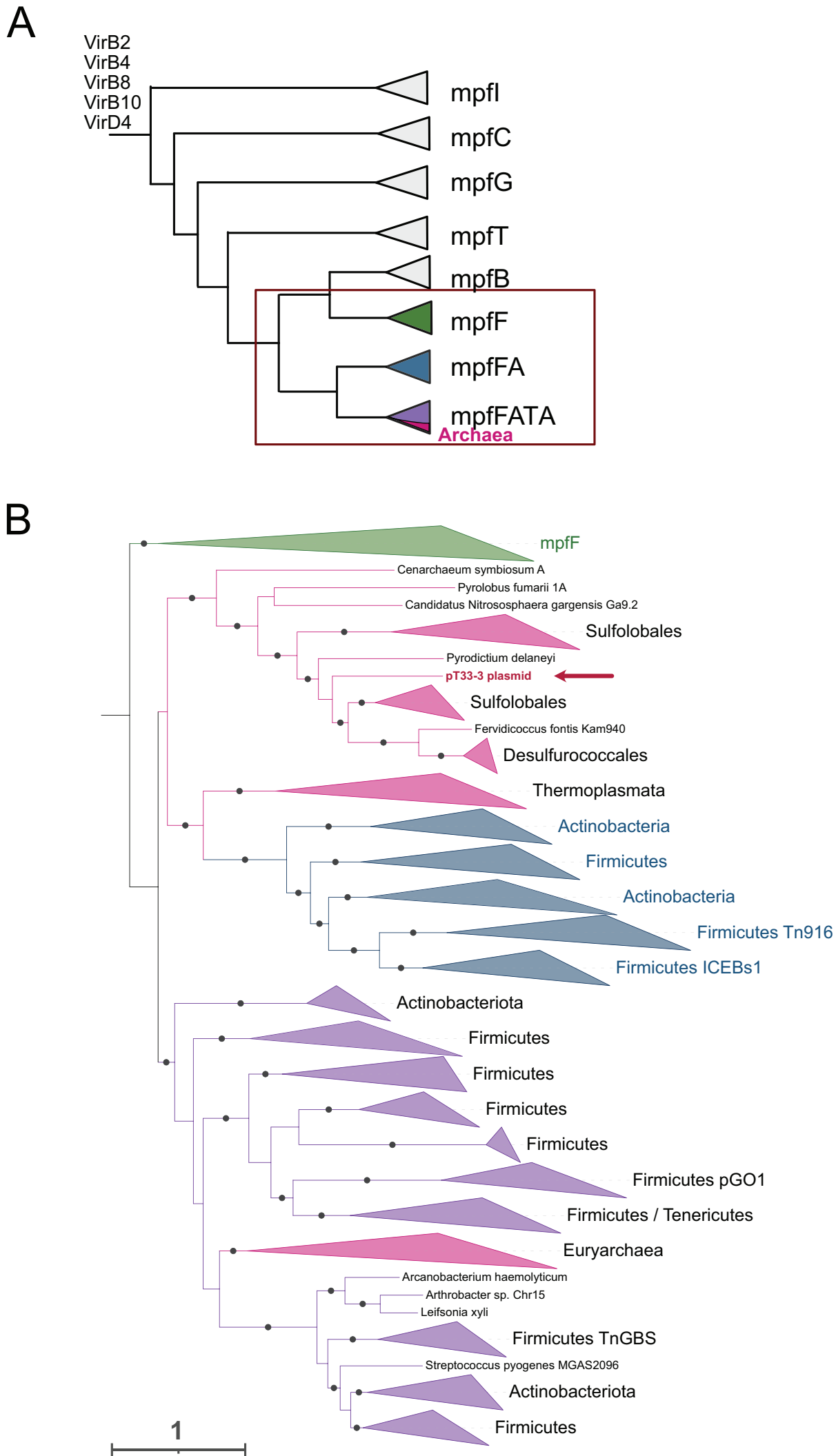

Wild-Type
Cas-null
1000
100
10
1
Transformation efficiency
(cfu/µg DNA)
non-target plasmid
target plasmid
non-target plasmid
target plasmid


No DNA
pT33-3
Δ0019
Δ0132
PCR locus:
0019
0132
0019
0132
0019
0132
0019
0132
NEB 1kb ladder
NEB 1kb ladder
kbp
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.0
0.5
WT 0132
WT 0019
Δ0132
Δ0019

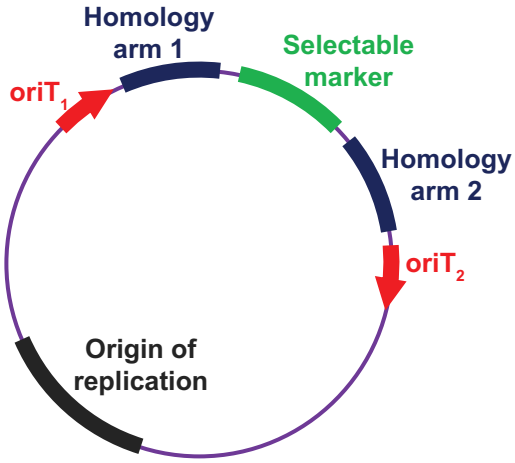

-1
-2
-3
-4
-5
-6
-7
log10 transfer rate (transformants per donor)
Donor:
Purified pLCAg plasmid DNA
T. kodakarensis + pLCAg
T. kodakarensis KUW1 (Agma+)
T. gammatolerans (Agma+)
T. kodakarensis (TrpE+, Agma-)
Recipient:
T. kodakarensis TS559 (Agma-)
T. kodakarensis TS559 (Agma-)
T. kodakarensis TS559 (Agma-)
T. kodakarensis TS559 (Agma-)
T. kodakarensis ΔTrpE::Agma
(A)
(B)
(C)
(D)
(E)


A)
GC skew / keto excess region 1:
3,000
3,500
4,000
4,500
ORF p0009
ORF p0010
ORF p0008
GC-skew maximum
oriT1:
B)
GC skew / keto excess region 2:
19,500
19,000
18,500
18,000
ORF p0023
ORF p0022
GC-skew minimum
oriT2:
C)
GC skew / keto excess region 2:
19,500
19,000
18,500
18,000
ORF p0023
ORF p0022
oriT2.1:
oriT2.2:
oriT2.3:
oriT2.4:
D)
GC skew / keto excess region 2:
19,500
19,000
18,500
18,000
ORF p0023
ORF p0022
oriT300:

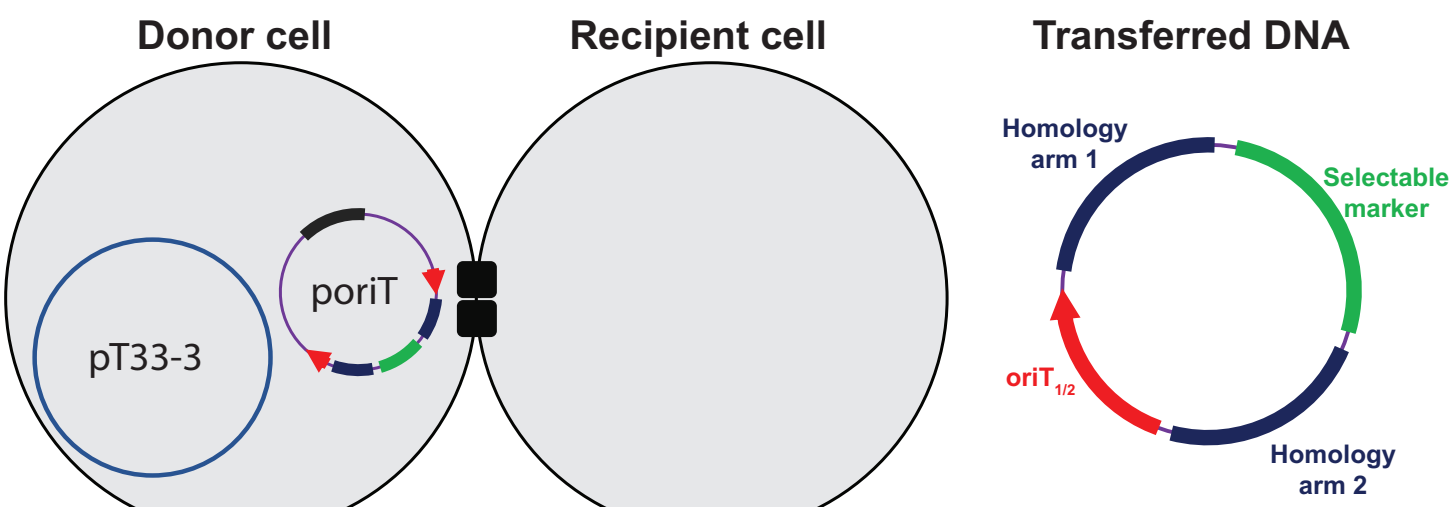

E)
Donor:
T. kodakarensis + pT33-3 +pTN1::oriT1
T. kodakarensis + pT33-3 +pTN1::oriT2
T. kodakarensis + pT33-3 +pTN1::oriT300
Log10 transfer efficiency (transconjugants per donor)
0
-1
-2
-3
-4
-5
-6
-7
T. kodakarensis
T. gammatolerans
T. nautili
Recipient


A
VirB2
VirB4
VirB8
VirB10
VirD4
mpfI
mpfC
mpfG
mpfT
mpfB
mpfF
mpfFA
mpfFATA
Archaea
B
mpfF
Cenarchaeum symbiosum A
Pyrolobus fumarii 1A
Candidatus Nitrososphaera gargensis Ga9.2
Sulfolobales
Pyrodictium delaneyi
pT33-3 plasmid
Sulfolobales
Fervidicoccus fontis Kam940
Desulfurococcales
Thermoplasmata
Actinobacteria
Firmicutes
Actinobacteria
Firmicutes Tn916
Firmicutes ICEBs1
Actinobacteriota
Firmicutes
Firmicutes
Firmicutes
Firmicutes
Firmicutes pGO1
Firmicutes / Tenericutes
Euryarchaea
Arcanobacterium haemolyticum
Arthrobacter sp. Chr15
Leifsonia xyli
Firmicutes TnGBS
Streptococcus pyogenes MGAS2096
Actinobacteriota
Firmicutes
1

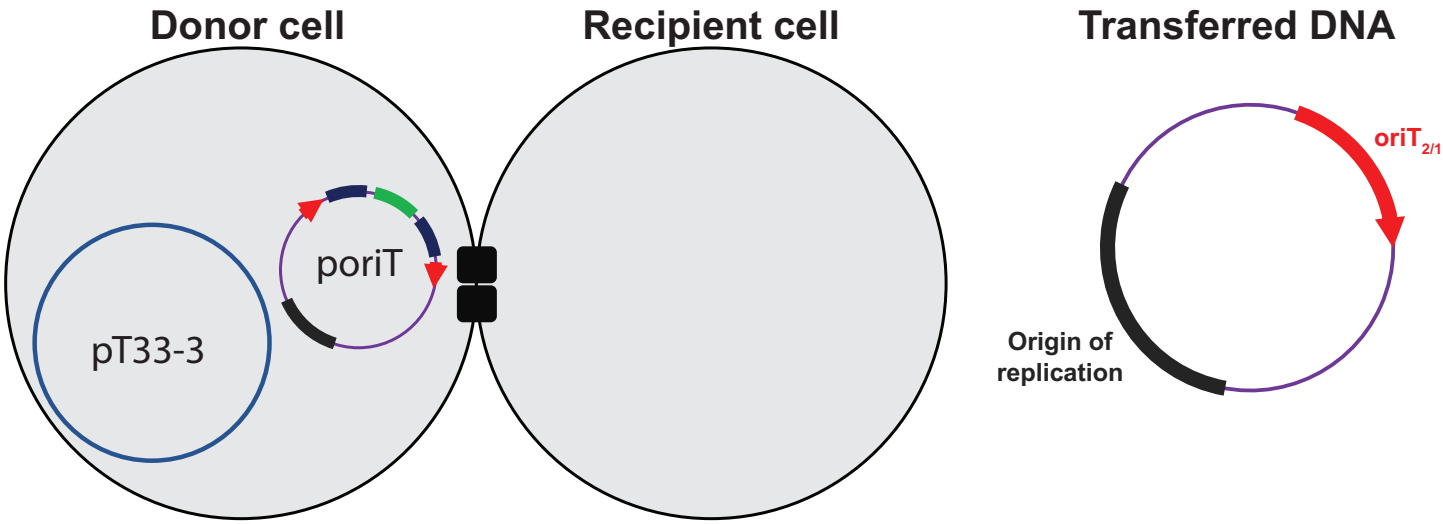
A)

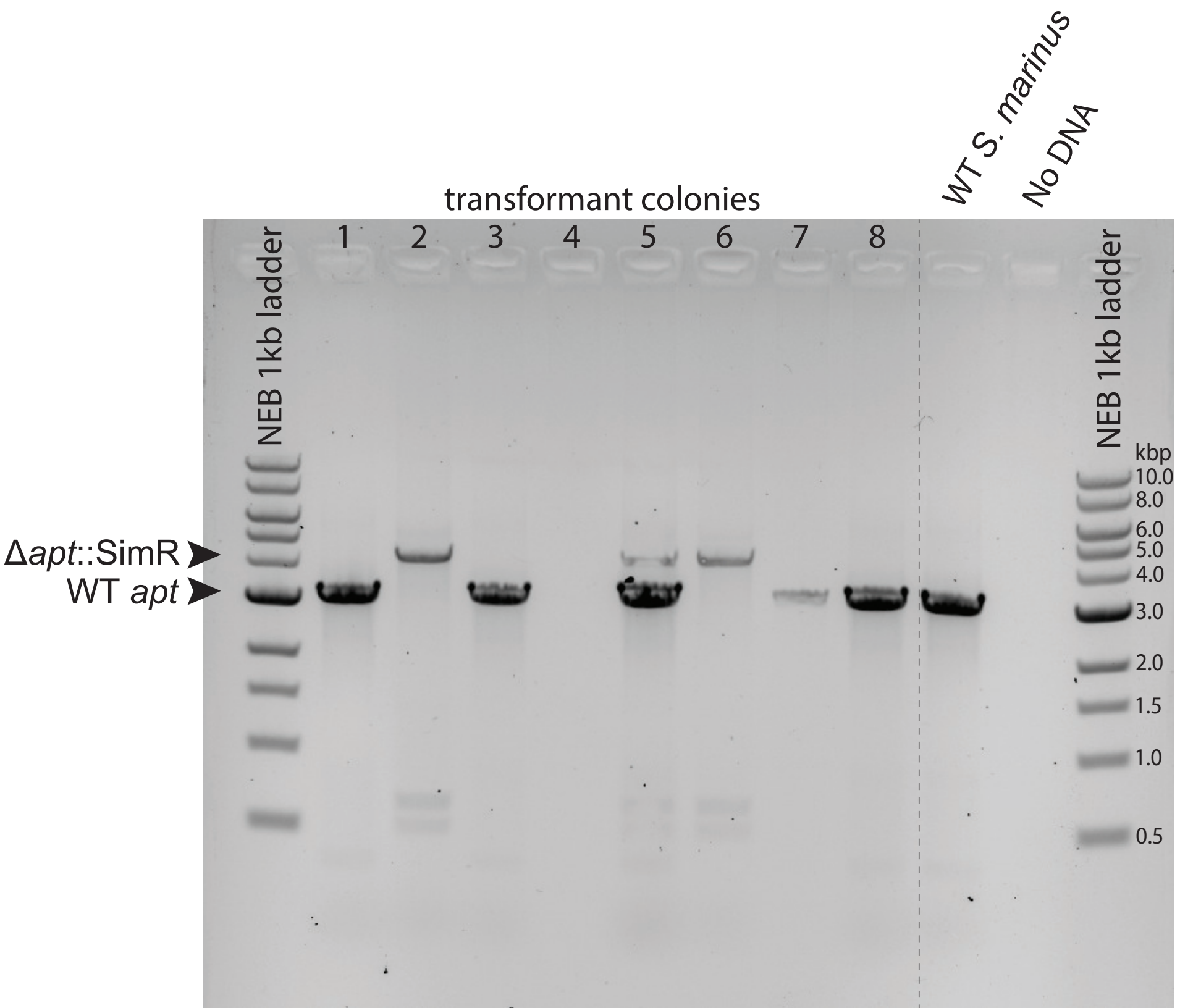

Homology arm 1
Selectable marker
oriT1
Homology arm 2
oriT2
Origin of replication

B)


Donor cell
Recipient cell
Transferred DNA
pT33-3
poriT
Homology arm 1
Selectable marker
oriT1/2
Homology arm 2

C)


Donor cell
Recipient cell
Transferred DNA
pT33-3
poriT
oriT2/1
Origin of replication


transformant colonies
1
2
3
4
5
6
7
8
WT S. marinus
No DNA
NEB 1kb ladder
NEB 1kb ladder
kbp
10.0
8.0
6.0
5.0
4.0
3.0
2.0
1.5
1.0
0.5
Δapt::SimR
WT apt