



**HAL**  
open science

## 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

### ► To cite this version:

Ryan J Catchpole, Valérie Barbe, Ghislaine Magdelenat, Evelyne Marguet, Michael Terns, et al.. A self-transmissible plasmid from a hyperthermophile that facilitates genetic modification of diverse Archaea. *Nature Microbiology*, 2023, 8 (7), pp.1339 - 1347. 10.1038/s41564-023-01387-x . hal-04274679

**HAL Id: hal-04274679**

**<https://hal.science/hal-04274679>**

Submitted on 6 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

15

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

18 Ryan J. Catchpole<sup>1,2,3</sup>; Valérie Barbe<sup>4</sup>; Ghislaine Magdelenat<sup>4</sup>, Evelyne Marguet<sup>2</sup>; Michael  
19 Terns<sup>3</sup>; Jacques Oberto<sup>2</sup>; Patrick Forterre<sup>1,2</sup>; Violette Da Cunha<sup>1,2,4</sup>.

20 <sup>1</sup>Institut Pasteur, Unité de Biologie Moléculaire du Gène chez les Extrêmophiles (BMGE), Département de Microbiologie F-75015 Paris,  
21 France

22 <sup>2</sup>Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Univ. Paris-Saclay, 91198, Gif-sur-Yvette Cedex,  
23 France

24 <sup>3</sup>Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602, USA

25 <sup>4</sup>Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Univ Evry, Université Paris-Saclay, 91057 Evry, France

26

27 Corresponding authors

28 Ryan J. Catchpole: [ryan.catchpole@uga.edu](mailto:ryan.catchpole@uga.edu)

29 Violette Da Cunha: [vdacunha@genoscope.cns.fr](mailto:vdacunha@genoscope.cns.fr)

30

31 **ABSTRACT**

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

46

47 **MAIN**

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

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

72

## 73 RESULTS

### 74 Identification and characterization of pT33-3

75 We sequenced the complete genome of the archaeal strain *Thermococcus sp. 33-3*, a  
76 hyperthermophilic euryarchaeon isolated from the East Pacific Ridge (depth 2,330m)<sup>11</sup> and  
77 assembled a complete closed circular genome of 2,013,327 bp (ENA accession  
78 GCA\_946300405). The 16S rRNA of *T. sp. 33-3* is identical to that of the published species  
79 *T. nautili*<sup>12</sup> and *T. sp. 26-2*<sup>13</sup>, suggesting they are strains of the same species; however, other  
80 common phylogenetic markers are non-identical (*rpoB* at 97% nucleotide identity, *pheS* at  
81 92%), and whole genome alignments reveal significant large-scale rearrangements between  
82 these isolates (Extended Data Figure 1), as observed in other Thermococcales species<sup>6</sup>.  
83 Phylogenetic reconstruction using RpoB places *T. sp. 33-3* in a clade with *T. nautili*,  
84 *T. henrieti*<sup>14</sup> and the as yet unnamed *T. sp. 21S9*<sup>15</sup> (Figure 1B). In addition to the circular  
85 chromosome, we identified an extra-chromosomal circular plasmid of 103,230 bp (hereafter  
86 referred to as pT33-3) (Figure 1A), representing the largest known plasmid in the  
87 Thermococcales order. Illumina read depth for the plasmid was approximately equal to that  
88 of the chromosome, suggesting pT33-3 exists at single copy number (1.47±0.33). Marker  
89 frequency analysis did not reveal the presence of an active replication origin (unsurprising in  
90 Thermococcales<sup>16</sup>), though indicators of replication initiation and termination regions (GC  
91 skew and keto excess<sup>17</sup>) give clear inflection points at 3.6 kbp and 19.0 kbp (Extended Data  
92 Figure 2).

93

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

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

110

### 111 **Determination of pT33-3 transmissibility**

112 If pT33-3 is mobile between Thermococcales species/strains (or has been in the past), we  
113 predict a record of these encounters to be found in the CRISPR arrays of such cells<sup>22</sup>. Indeed,  
114 of the 70 available Thermococcales genome assemblies, 12 strains encode a total of 45  
115 CRISPR spacers potentially targeting pT33-3 at >85% identity (Figure 1A). Additionally, the  
116 45 spacers target different sites on pT33-3 (Figure 1B), indicating multiple spacer acquisition  
117 events and thus multiple infections by pT33-3, rather than a single ancestral infection<sup>23</sup>.

118

119 The potential T4SS-like machinery of pT33-3, combined with its apparent historical  
120 promiscuity prompted us to investigate its ability to transfer under laboratory conditions.  
121 Although genetic tools have been developed for a few select Thermococcales species<sup>24</sup>,  
122 *T. sp. 33-3* is not naturally competent for DNA uptake, so we instead developed the plasmid  
123 in a *T. kodakarensis* host. However, the presence of a potential anti-pT33-3 CRISPR spacer  
124 in *T. kodakarensis* (CRISPR1 spacer 12 with 87% match to *p0132*) required generation of a  
125 strain devoid of CRISPR-Cas activity. To achieve this, we disrupted the entire 20 kb cas gene  
126 locus and the upstream CRISPR array ( $\Delta TK0445-TK0464, \Delta CR1$ ) in *T. kodakarensis*<sup>25</sup>. The  
127 resulting strain indeed lacked CRISPR-Cas activity against incoming plasmid DNA  
128 (Extended Data Figure 3) and was thus used as a host for pT33-3 in subsequent experiments.

129

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

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

153

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

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

168

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

179

### 180 **Mobilization of shuttle vectors by pT33-3**

181 A further defining feature of conjugative elements is the use of an *oriT* (origin of transfer)  
182 sequence as the initiation point for transfer<sup>31</sup>. In bacterial CPs, *oriTs* are often encoded  
183 adjacent to the transfer genes<sup>32</sup>. We identified two candidate *oriT* sequences (hereafter named  
184 *oriT1* and *oriT2*, Figure 1) of ~2 kbp due to their proximity to potential transfer genes, being  
185 relatively long intergenic spaces, encoding significant potential for stem-loop formation, and  
186 being at inflection points of GC-skew (points 3.6 kb and 19.0 kb (Extended Data Figure 2)).  
187 We cloned each region into a small pTN1-based replicative vector encoding the mevinolin  
188 resistance marker (*poriT1/2*) (Extended Data Figure 6). We then tested the ability of  
189 *T. kodakarensis* donors with and without pT33-3 to mobilise these plasmids to other  
190 Thermococcales. Strikingly, *oriT2* facilitated transfer to *T. nautili* and *T. gammatolerans* at a  
191 higher frequency than pT33-3 itself (Figure 3, Extended Data Figure 6). The elevated transfer  
192 frequency may be due to the smaller size of mobilised plasmids, and the higher copy number  
193 relative to pT33-3. This transfer was reliant upon the presence of pT33-3, suggesting  
194 mobilization of shuttle vectors occurs through the pT33-3 encoded conjugation apparatus. We  
195 were able to recover intact plasmid DNA from cultures of *T. litoralis*, *T. gammatolerans*, *T.*  
196 *pacificus*, and *T. sp. 5-4* (Figure 3). These species are not amenable to genetic transformation  
197 with purified DNA, demonstrating the utility of this conjugation/mobilization-based  
198 approach. We further narrowed this *oriT*-encoding region to a 300 bp sequence which  
199 facilitated mobilization (Extended Data Figure 6). Deletion of this 300bp *oriT* region from

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

202

### 203 **Evolutionary origin of pT33-3**

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

221

### 222 **pT33-3 for genomic manipulation of diverse Archaea**

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



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

236

237 Additionally, we were able to extend this plasmid delivery/genetic manipulation approach  
238 beyond Euryarchaeota to the Crenarchaeal species *Staphylothermus marinus*<sup>35</sup>. Using  
239 *T. kodakarensis* as a mobilization donor, we replaced the putative adenosine  
240 phosphoribosyltransferase (*apt*) gene (*hpt1/Smar\_0750/ABN69851.1*) with a mevinolin-  
241 resistance marker from *Sulfolobus islandicus*<sup>36</sup>. Transformants were selected on a  
242 combination of 6-methylpurine and mevinolin, and colonies screened by PCR. In this case,  
243 spontaneous resistance was more frequent than true allelic exchange, but knockouts were  
244 detectable by PCR (Figure 4, Extended Data Figure 9). To our knowledge, this is the first  
245 example of inter-phylum conjugation in Archaea, and the first example of genetic  
246 modification of *Staphylothermus marinus*.

247

#### 248 **Conjugation at high temperatures**

249 Finally, we were intrigued by the ability of pT33-3 to transfer to *Pyrococcus abyssi*  
250 (maximum growth temperature 102°C<sup>37</sup>) in our experiments, as well as the presence of pT33-  
251 3 CRISPR spacers in various *Pyrococcus* species, suggesting an impressive ability to transfer  
252 at very high temperature. We thus sought to test the maximum and minimum temperatures at  
253 which we can detect pT33-3 transfer between species. Transconjugants of both *P. abyssi* and  
254 *T. litoralis* were isolated following co-culturing at 100°C but not at 105°C; however, this is  
255 outside the growth temperature range of the *T. kodakarensis* donor, thus host survival may be  
256 the limiting factor rather than thermostability of the pT33-3 transfer machinery. It is difficult  
257 to say whether conjugation occurs during the 100°C incubation period, or at some  
258 intermediate temperature while cultures warm to 100°C. Regardless, isolation of  
259 transconjugants at 100°C sets new upper bounds on the function of known conjugative  
260 plasmids. At lower temperatures, transfer was successful at 55°C, but not detected at 50°C.  
261 Again, this is outside the growth temperatures of both donor and recipient cells. Such high  
262 thermostability makes pT33-3 an ideal candidate for studying the structure of archaeal  
263 conjugative apparatus.

264

#### 265 **DISCUSSION**

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

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

272

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

288

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

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

304

305

306

307

## 308 **Methods**

### 309 **Culturing**

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

320

321

### 322 **DNA sequencing and analyses**

323 Total DNA was isolated from cultures by phenol-chloroform extraction as described  
324 previously<sup>11</sup>. Sequencing of the *Thermococcus sp.* 33-3 strain (ENA accession  
325 GCA\_946300405) was carried out at Genoscope (Centre National de Séquençage, France),  
326 using Illumina MiSeq and Oxford Nanopore MinION. Reads were assembled with NaS  
327 (<https://github.com/institut-de-genomique/NaS>) and corrected with Canu<sup>46</sup>.  
328 Mapping of Illumina reads to assembled genomes (as an estimate of copy number) was  
329 performed and visualised using Geneious (BioMatters). 16S, RpoB and PheS sequences were  
330 each aligned with MAFFT<sup>47</sup>. Whole genomes were aligned with Gepard v2.1<sup>48</sup>.  
331 Functional annotations of pT33-3 were predicted using each ORF as a query sequence for  
332 BLASTp<sup>49</sup> against the nr database (cutoff E=0.01); HHblits (HHSuite v3)<sup>50</sup> against the  
333 UniRef30\_2020\_06 database (n=2 iterations, cutoff E=0.01).  
334 To analyse CRISPR spacer matches, CRISPR arrays were predicted in all Thermococcales  
335 genomes available through NCBI (November 2021) using CRT<sup>51</sup> in Geneious. Spacers were  
336 extracted and used as query sequences for BLASTn search against pT33-3 (cutoff parameters  
337 of 85% identity over the entire spacer length).

338

### 339 **Strain generation**

340 To generate a *T. kodakarensis* strain devoid of CRISPR activity, we deleted the region of the  
341 genome comprising *TK0445-TK0464* and adjacent the CRISPR1 array<sup>25</sup> from strain TS559  
342 using techniques described previously<sup>52</sup>. Activity was tested by cloning spacers from each of  
343 the 3 *T. kodakarensis* CRISPR arrays<sup>25</sup> in to pLC70<sup>53</sup>, and used to transform CRISPR-  
344 knockout strains (and parental strains) of *T. kodakarensis* as described previously<sup>40</sup>.  
345 To generate a *T. kodakarensis* strain harboring pT33-3, we generated a pTP2-derived  
346 plasmid<sup>54</sup> encoding TK0664 (conferring 6-methylpurine sensitivity), and a recombination  
347 cassette encoding 1kb homology arms identical to pT33-3 flanking a *pdaD* agmatine marker  
348 gene. A triparental mating was carried out by mixing cell pellets from 50mL of early  
349 stationary phase *T. sp. 33-3*, *T. kodakarensis* (*trp*-) harboring the recombination substrate  
350 plasmid, and the *T. kodakarensis* (*trp*+) CRISPR-null strain. The mixed suspension was  
351 spotted on solid TRM media (supplemented with agmatine) and incubated anaerobically (N<sub>2</sub>  
352 atmosphere) at 85°C overnight. The following day, plates were chilled to 4°C before  
353 suspending the plated cell mass in 200 µL 0.8x ASW, and spreading on solid ASW-CH  
354 media containing 6-methylpurine (media devoid of tryptophan to select for recipient cells  
355 over recombination donors; containing 6-methylpurine to select against the full  
356 recombination substrate plasmid and against *T. sp. 33-3*; devoid of agmatine to select for the  
357 presence of the agmatine marker cassette). Colonies were screened by PCR for the presence  
358 of pT33-3 with the agmatine marker cassette, grown in liquid culture, and total DNA  
359 sequenced as described above.

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

363

#### 364 **Mating assays**

365 25 mL cultures of donor and recipient cells were grown to early stationary phase in  
366 appropriate media (see above). Cultures were chilled on ice before pelleting cell mass at  
367 3,500xg for 20 minutes at 4°C. Media was decanted and pellets resuspended in 100 µL 0.8x  
368 ASW. Appropriate cell suspensions were mixed and spotted on the centre of a solid TRM  
369 plate containing agmatine. Where donors and recipients were separated by a membrane filter,  
370 donor suspension was first spotted directly on the media then overlaid with a 0.2 µm  
371 nitrocellulose filter, atop which the recipient suspension was spotted. Plates were incubated at  
372 85°C (except when testing temperature limits) for 16 hours to allow mating, before chilling to  
373 4°C. Cell mass was scraped from the plate and suspended in 200 µL 0.8x ASW. Serial

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

378

### 379 **Mobilization assays**

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

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

394

### 395 **Phylogeny**

396 For RpoB phylogenetic reconstruction, sequences aligned with MAFFT<sup>47</sup>, alignment gaps  
397 removed with BMGE<sup>55</sup>, evolutionary model selected with ModelFinder<sup>56</sup>, and phylogenetic  
398 trees generated with IQ-TREE v1.6.12<sup>57</sup>. Trees were visualised with iTol<sup>58</sup>.

399 For VirB4 phylogenetic reconstruction, sequences of VirB4 proteins from MPF<sub>FATA</sub>, and  
400 MPF<sub>FA</sub> used previously<sup>10,33</sup> were retrieved from Genbank. Additional archaeal sequences  
401 were retrieved by BLASTp search. Sequences of the MPF<sub>F</sub> class were used as an outgroup.  
402 Phylogenetic trees were generated as for RpoB as above. The maximum likelihood tree was  
403 constructed using IQ-TREE with the MFP option for model selection. Branch supports were  
404 either computed through SH-like aLRT<sup>59</sup> (10,000 replicates) and UFBoot<sup>60</sup> (10,000  
405 replicates).

406

407

408

409

410 **Data Availability**

411 The genome sequence for *Thermococcus sp.* 33-3 has been deposited with ENA number  
412 GCA\_946300405. Sequences of all plasmids, genomes, and primers used in this study are  
413 available on Dryad with doi 10.5061/dryad.2jm63xst9.

414

415

416 **Acknowledgements**

417 We thank Nicolas Soler for advice while writing this manuscript. This project has received  
418 funding from the European Research Council (ERC) under the European Union's Horizon  
419 2020 Research and Innovation Program (grant agreement no.772725, project HiPhore to PF),  
420 under the European Union's Seventh Framework Program (grant agreement no. 340440,  
421 project EVOMOBIL to PF) and by the National Institutes of Health (R35GM118160 to MT).

422

423 **Author Contributions**

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

428

429

430 **Competing interests**

431 The authors declare no competing interests.

432

433

434 **References**

435

436 1. Lederberg, J., Cavalli, L. L. & Lederberg, E. M. Sex Compatibility in Escherichia Coli.

437 *Genetics* **37**, 720–730 (1952).

438 2. Grohmann Elisabeth, Muth Günther & Espinosa Manuel. Conjugative Plasmid Transfer

439 in Gram-Positive Bacteria. *Microbiol. Mol. Biol. Rev.* **67**, 277–301 (2003).

440 3. De La Cruz, F., Frost, L. S., Meyer, R. J. & Zechner, E. L. Conjugative DNA metabolism

441 in Gram-negative bacteria. *FEMS Microbiol. Rev.* **34**, 18–40 (2010).

442 4. Ramirez, M. S., Traglia, G. M., Lin, D. L., Tran, T. & Tolmasky, M. E. Plasmid-

443 Mediated Antibiotic Resistance and Virulence in Gram-negatives: the Klebsiella

444 pneumoniae Paradigm. *Microbiol Spectr* **2**, 1–15 (2014).

445 5. Montero, I., Herrero-Fresno, A., Rodicio, R. & Rodicio, M. R. Efficient mobilization of a

446 resistance derivative of pSLT, the virulence plasmid specific of Salmonella enterica

447 serovar Typhimurium, by an IncI1 plasmid. *Plasmid* **70**, 104–109 (2013).

448 6. Cossu, M. *et al.* Flipping chromosomes in deep-sea archaea. *PLoS Genet.* **13**, e1006847

449 (2017).

450 7. Tamminen, M., Virta, M., Fani, R. & Fondi, M. Large-scale analysis of plasmid

451 relationships through gene-sharing networks. *Mol. Biol. Evol.* **29**, 1225–1240 (2012).

452 8. Basta, T., Smyth, J., Forterre, P., Prangishvili, D. & Peng, X. Novel archaeal plasmid

453 pAH1 and its interactions with the lipothrixvirus AFV1. *Mol. Microbiol.* **71**, 23–34

454 (2009).

455 9. Prangishvili, D. *et al.* Conjugation in Archaea: Frequent Occurrence of Conjugative

456 Plasmids in Sulfolobus. *Plasmid* **40**, 190–202 (1998).

457 10. Guglielmini, J., de la Cruz, F. & Rocha, E. P. C. Evolution of conjugation and type IV

458 secretion systems. *Mol. Biol. Evol.* **30**, 315–331 (2013).

459 11. Lepage, E. *et al.* Molecular diversity of new Thermococcales isolates from a single area



- 460 of hydrothermal deep-sea vents as revealed by randomly amplified polymorphic DNA  
461 fingerprinting and 16S rRNA gene sequence analysis. *Appl. Environ. Microbiol.* **70**,  
462 1277–1286 (2004).
- 463 12. Gorlas, A. *et al.* Thermococcus nautili sp. nov., a hyperthermophilic archaeon isolated  
464 from a hydrothermal deep-sea vent. *Int. J. Syst. Evol. Microbiol.* **64**, 1802–1810 (2014).
- 465 13. Soler, N. *et al.* Two novel families of plasmids from hyperthermophilic archaea encoding  
466 new families of replication proteins. *Nucleic Acids Res.* **38**, 5088–5104 (2010).
- 467 14. Alain, K. *et al.* Thermococcus henrietii sp. nov., a novel extreme thermophilic and  
468 piezophilic sulfur-reducing archaeon isolated from a deep-sea hydrothermal chimney. *Int.*  
469 *J. Syst. Evol. Microbiol.* **71**, (2021).
- 470 15. Strang, L. C. Genomic Insights and Ecological Adaptations of Deep-Subsurface and Near  
471 Subsurface Thermococcus Isolates. (Western Washington University, 2020).
- 472 16. Gehring, A. M. *et al.* Genome Replication in Thermococcus kodakarensis Independent of  
473 Cdc6 and an Origin of Replication. *Front. Microbiol.* **8**, 2084 (2017).
- 474 17. Freeman, J. M., Plasterer, T. N., Smith, T. F. & Mohr, S. C. Patterns of Genome  
475 Organization in Bacteria. *Science* (1998) doi:10.1126/science.279.5358.1827a.
- 476 18. Díaz-Orejas, R., Espinosa, M. & Yeo, C. C. The Importance of the Expendable: Toxin-  
477 Antitoxin Genes in Plasmids and Chromosomes. *Front. Microbiol.* **8**, 1479 (2017).
- 478 19. Guynet, C. & de la Cruz, F. Plasmid segregation without partition. *Mob. Genet. Elements*  
479 **1**, 236–241 (2011).
- 480 20. Krupovic, M., Gribaldo, S., Bamford, D. H. & Forterre, P. The evolutionary history of  
481 archaeal MCM helicases: a case study of vertical evolution combined with hitchhiking of  
482 mobile genetic elements. *Mol. Biol. Evol.* **27**, 2716–2732 (2010).
- 483 21. Grohmann, E., Christie, P. J., Waksman, G. & Backert, S. Type IV secretion in Gram-  
484 negative and Gram-positive bacteria. *Mol. Microbiol.* **107**, 455–471 (2018).

- 485 22. Terns, M. P. & Terns, R. M. CRISPR-based adaptive immune systems. *Curr. Opin.*  
486 *Microbiol.* **14**, 321–327 (2011).
- 487 23. Pinilla-Redondo, R. *et al.* Type IV CRISPR-Cas systems are highly diverse and involved  
488 in competition between plasmids. *Nucleic Acids Res.* (2019) doi:10.1093/nar/gkz1197.
- 489 24. Atomi, H., Imanaka, T. & Fukui, T. Overview of the genetic tools in the Archaea. *Front.*  
490 *Microbiol.* **3**, 337 (2012).
- 491 25. Elmore, J. R. *et al.* Programmable plasmid interference by the CRISPR-Cas system in  
492 *Thermococcus kodakarensis*. *RNA Biol.* **10**, 828–840 (2013).
- 493 26. Matsumi, R., Manabe, K., Fukui, T., Atomi, H. & Imanaka, T. Disruption of a sugar  
494 transporter gene cluster in a hyperthermophilic archaeon using a host-marker system  
495 based on antibiotic resistance. *J. Bacteriol.* **189**, 2683–2691 (2007).
- 496 27. Naor, A. & Gophna, U. Cell fusion and hybrids in Archaea: prospects for genome  
497 shuffling and accelerated strain development for biotechnology. *Bioengineered* **4**, 126–  
498 129 (2013).
- 499 28. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Targeted gene disruption by homologous  
500 recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J.*  
501 *Bacteriol.* **185**, 210–220 (2003).
- 502 29. Forterre, P., Da Cunha, V. & Catchpole, R. Plasmid vesicles mimicking virions. *Nat*  
503 *Microbiol* **2**, 1340–1341 (2017).
- 504 30. Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M. J. & Cavicchioli, R. A plasmid  
505 from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and  
506 infect plasmid-free cells. *Nat Microbiol* **2**, 1446–1455 (2017).
- 507 31. Willetts, N. & Wilkins, B. Processing of plasmid DNA during bacterial conjugation.  
508 *Microbiol. Rev.* **48**, 24–41 (1984).
- 509 32. Zatyka, M. & Thomas, C. M. Control of genes for conjugative transfer of plasmids and

- 510 other mobile elements. *FEMS Microbiol. Rev.* **21**, 291–319 (1998).
- 511 33. Guglielmini, J. *et al.* Key components of the eight classes of type IV secretion systems  
512 involved in bacterial conjugation or protein secretion. *Nucleic Acids Res.* **42**, 5715–5727  
513 (2014).
- 514 34. Draper, O., César, C. E., Machón, C., de la Cruz, F. & Llosa, M. Site-specific  
515 recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc.*  
516 *Natl. Acad. Sci. U. S. A.* **102**, 16385–16390 (2005).
- 517 35. Fiala, G., Stetter, K. O., Jannasch, H. W., Langworthy, T. A. & Madon, J.  
518 *Staphylothermus marinus* sp. nov. Represents a Novel Genus of Extremely Thermophilic  
519 Submarine Heterotrophic Archaeobacteria Growing up to 98 °C. *Syst. Appl. Microbiol.* **8**,  
520 106–113 (1986).
- 521 36. Zhang, C., Cooper, T. E., Krause, D. J. & Whitaker, R. J. Augmenting the genetic  
522 toolbox for *Sulfolobus islandicus* with a stringent positive selectable marker for agmatine  
523 prototrophy. *Appl. Environ. Microbiol.* **79**, 5539–5549 (2013).
- 524 37. Erauso, G. *et al.* *Pyrococcus abyssi* sp. nov., a new hyperthermophilic archaeon isolated  
525 from a deep-sea hydrothermal vent. *Arch. Microbiol.* **160**, 338–349 (1993).
- 526 38. Wagner, A. *et al.* Mechanisms of gene flow in archaea. *Nat. Rev. Microbiol.* **15**, 492–501  
527 (2017).
- 528 39. Albers, S.-V. & Meyer, B. H. The archaeal cell envelope. *Nat. Rev. Microbiol.* **9**, 414–  
529 426 (2011).
- 530 40. Sato, T., Fukui, T., Atomi, H. & Imanaka, T. Improved and versatile transformation  
531 system allowing multiple genetic manipulations of the hyperthermophilic archaeon  
532 *Thermococcus kodakaraensis*. *Appl. Environ. Microbiol.* **71**, 3889–3899 (2005).
- 533 41. Lipscomb, G. L. *et al.* Natural competence in the hyperthermophilic archaeon  
534 *Pyrococcus furiosus* facilitates genetic manipulation: construction of markerless deletions

535 of genes encoding the two cytoplasmic hydrogenases. *Appl. Environ. Microbiol.* **77**,  
536 2232–2238 (2011).

537 42. Kim, M.-S. *et al.* CO-dependent H<sub>2</sub> production by genetically engineered *Thermococcus*  
538 *onnurineus* NA1. *Appl. Environ. Microbiol.* **79**, 2048–2053 (2013).

539 43. Li, X. *et al.* Genetic tools for the piezophilic hyperthermophilic archaeon *Pyrococcus*  
540 *yayanosii*. *Extremophiles* **19**, 59–67 (2015).

541 44. Zeng, X. *et al.* *Pyrococcus* CH1, an obligate piezophilic hyperthermophile: extending the  
542 upper pressure-temperature limits for life. *ISME J.* **3**, 873–876 (2009).

543 45. Akpan, I., Bankole, M. O. & Adesemowo, A. M. A rapid plate culture method for  
544 screening of amylase producing micro-organisms. *Biotechnol. Tech.* **13**, 411–413 (1999).

545 46. Koren, S. *et al.* Canu: scalable and accurate long-read assembly via adaptive k-mer  
546 weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).

547 47. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:  
548 improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

549 48. Krumsiek, J., Arnold, R. & Rattei, T. Gepard: a rapid and sensitive tool for creating  
550 dotplots on genome scale. *Bioinformatics* **23**, 1026–1028 (2007).

551 49. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local  
552 alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

553 50. Steinegger, M. *et al.* HH-suite3 for fast remote homology detection and deep protein  
554 annotation. *bioRxiv* 560029 (2019) doi:10.1101/560029.

555 51. Bland, C. *et al.* CRISPR recognition tool (CRT): a tool for automatic detection of  
556 clustered regularly interspaced palindromic repeats. *BMC Bioinformatics* **8**, 209 (2007).

557 52. Santangelo, T. J., Cubonová, L. & Reeve, J. N. *Thermococcus kodakarensis* genetics:  
558 TK1827-encoded beta-glycosidase, new positive-selection protocol, and targeted and  
559 repetitive deletion technology. *Appl. Environ. Microbiol.* **76**, 1044–1052 (2010).

- 560 53. Santangelo, T. J., Cubonová, L. & Reeve, J. N. Shuttle vector expression in  
561 *Thermococcus kodakaraensis*: contributions of cis elements to protein synthesis in a  
562 hyperthermophilic archaeon. *Appl. Environ. Microbiol.* **74**, 3099–3104 (2008).
- 563 54. Catchpole, R., Gorlas, A., Oberto, J. & Forterre, P. A series of new *E. coli*-*Thermococcus*  
564 shuttle vectors compatible with previously existing vectors. *Extremophiles* **22**, 591–598  
565 (2018).
- 566 55. Criscuolo, A. & Gribaldo, S. BMGE (Block Mapping and Gathering with Entropy): a  
567 new software for selection of phylogenetic informative regions from multiple sequence  
568 alignments. *BMC Evol. Biol.* **10**, 210 (2010).
- 569 56. Kalyaanamoorthy, S., Minh, B. Q. & Wong, T. K. F. ModelFinder: fast model selection  
570 for accurate phylogenetic estimates. *Nature* (2017).
- 571 57. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: A Fast and  
572 Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol.*  
573 *Biol. Evol.* **32**, 268–274 (2014).
- 574 58. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and  
575 annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242-5 (2016).
- 576 59. Guindon, S. *et al.* New Algorithms and Methods to Estimate Maximum-Likelihood  
577 Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).
- 578 60. Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q. & Vinh, L. S. UFBoot2:  
579 Improving the Ultrafast Bootstrap Approximation. *Mol. Biol. Evol.* **35**, 518–522 (2018).
- 580

581 **FIGURE LEGENDS**

582

583

**Figure 1**

584

**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.

588

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).

591

592

**Figure 2**

593

**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.

595

596

597

598

599

600

601

602

603

**Figure 3**

604

**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* (*pori2*) is mobilized to non-competent donor species in the presence of pT33-3. C) Plasmid DNA isolated from transconjugants has identical *HindIII* 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.

610

611

**Figure 4**

612

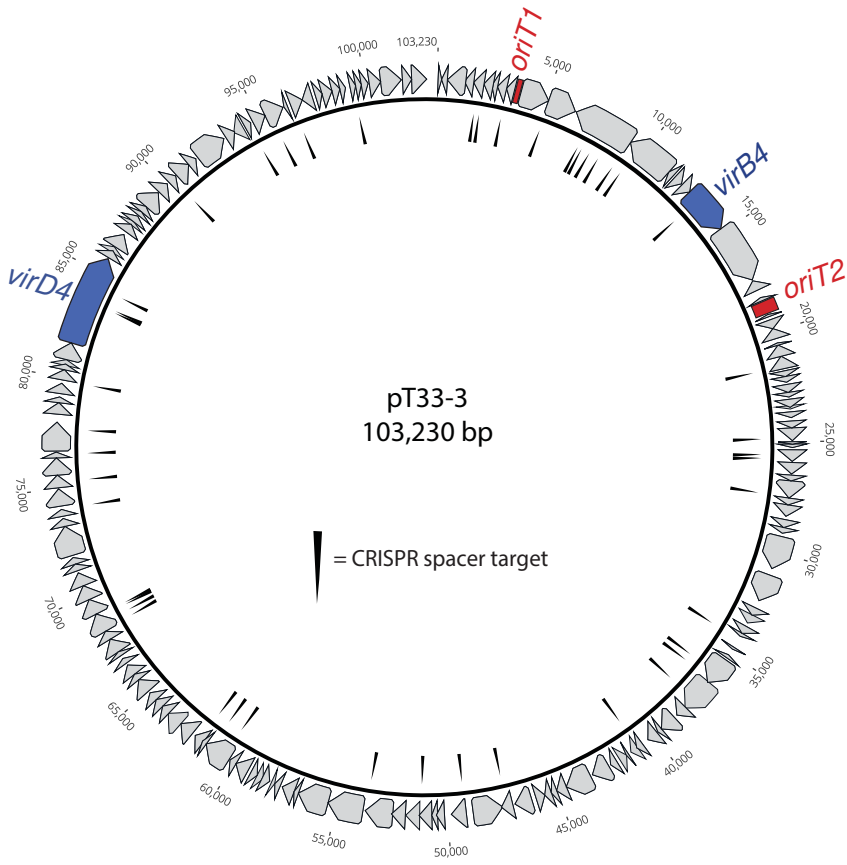
**pT33-3 mediated genome modification.** A) pT33-3 *oriT* allows mobilisation of non-replicative DNA to recipient organisms. Here, homology arms allow allelic exchange with the chromosome, facilitating genetic modification of recipients. Phenotypes selected in this work are indicated. B) Deletion of *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  $\Delta pyrF$  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.

622

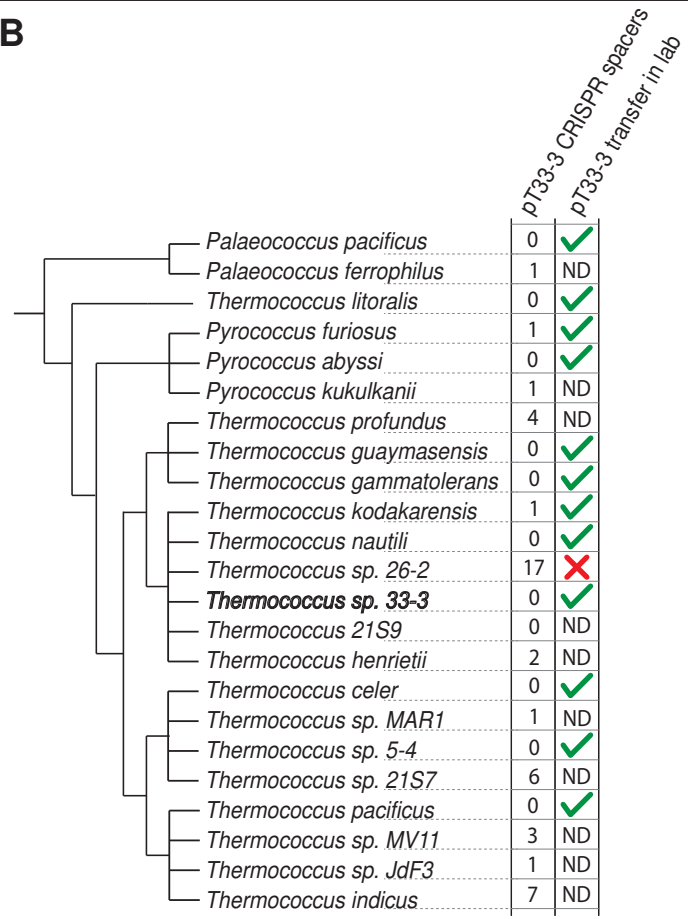
623

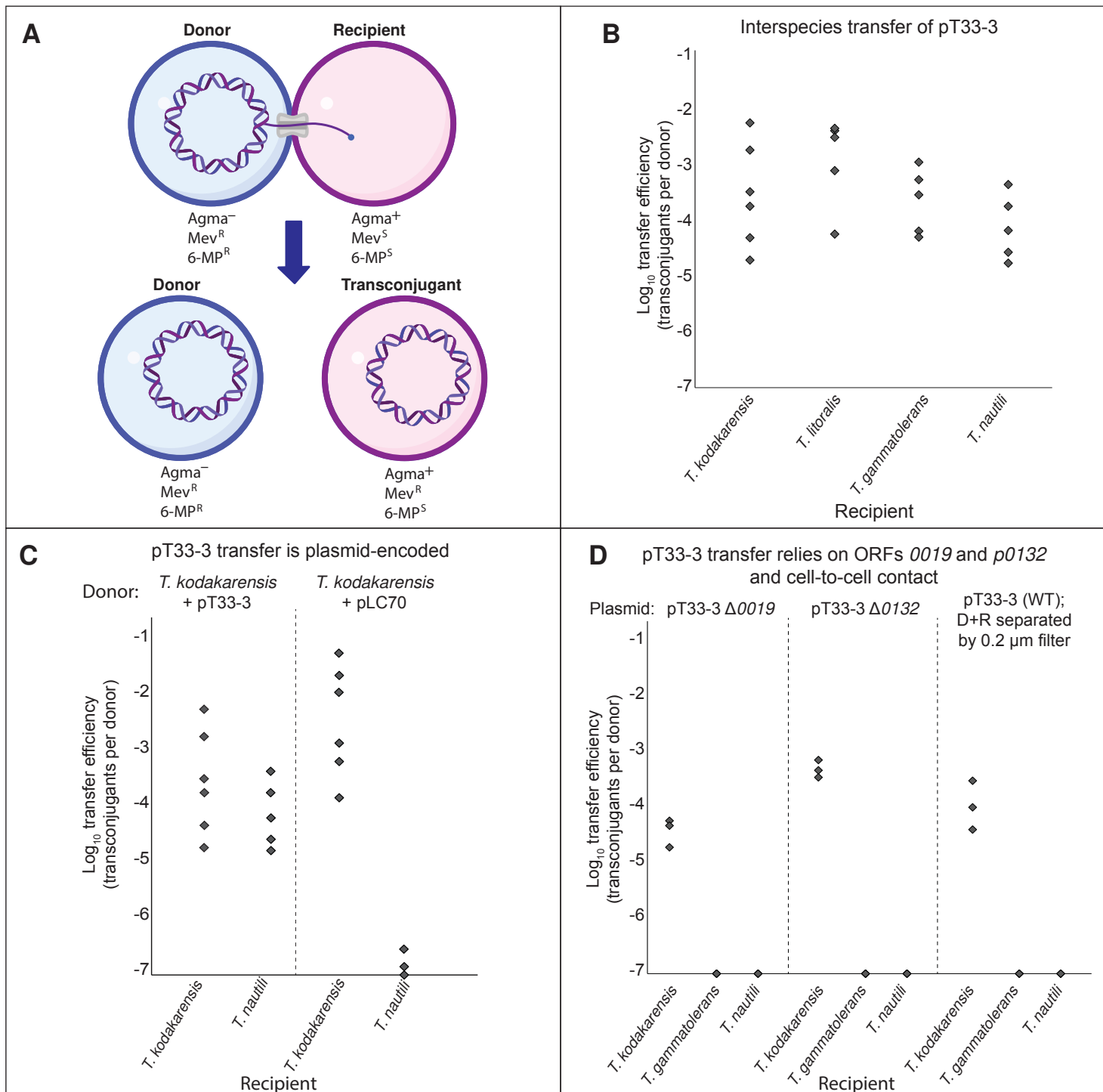
624

**A**

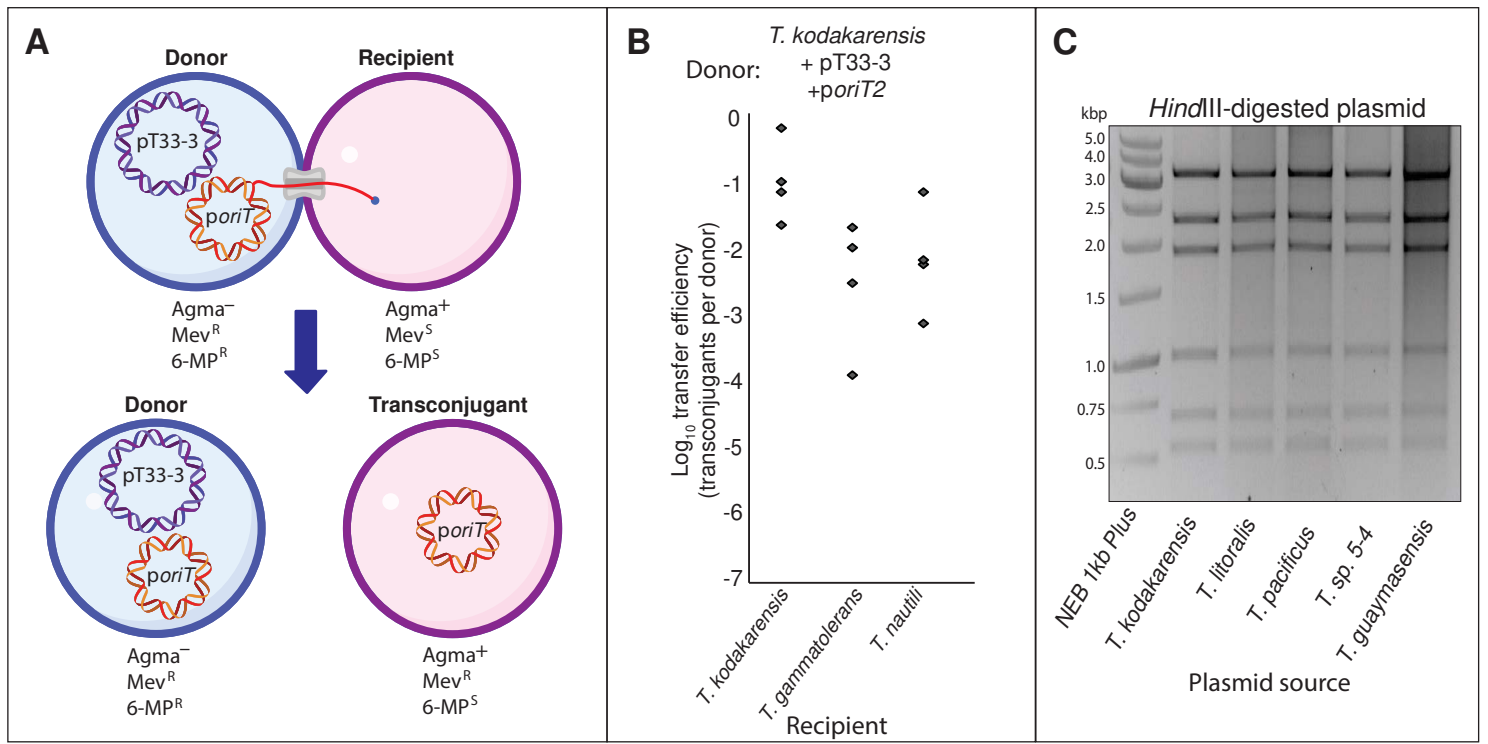


**B**









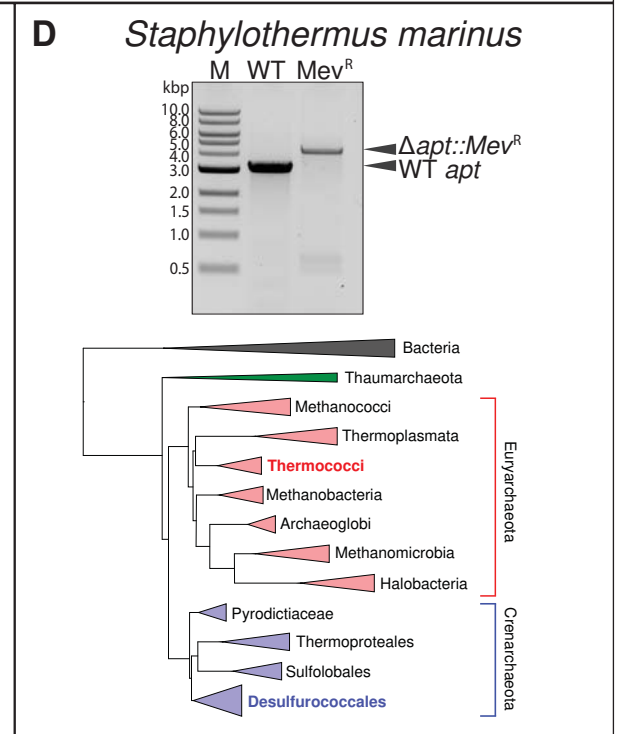
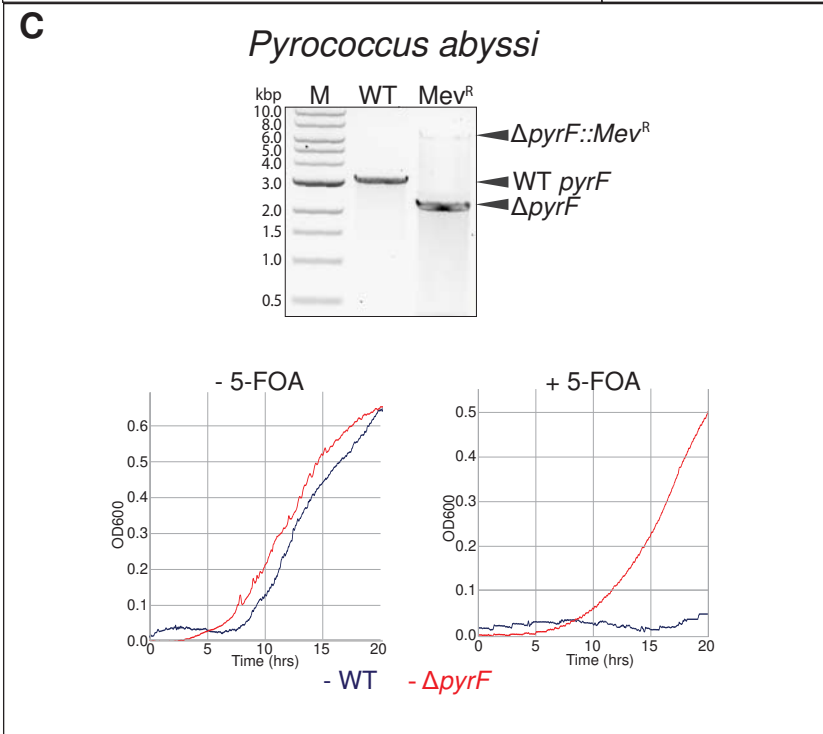
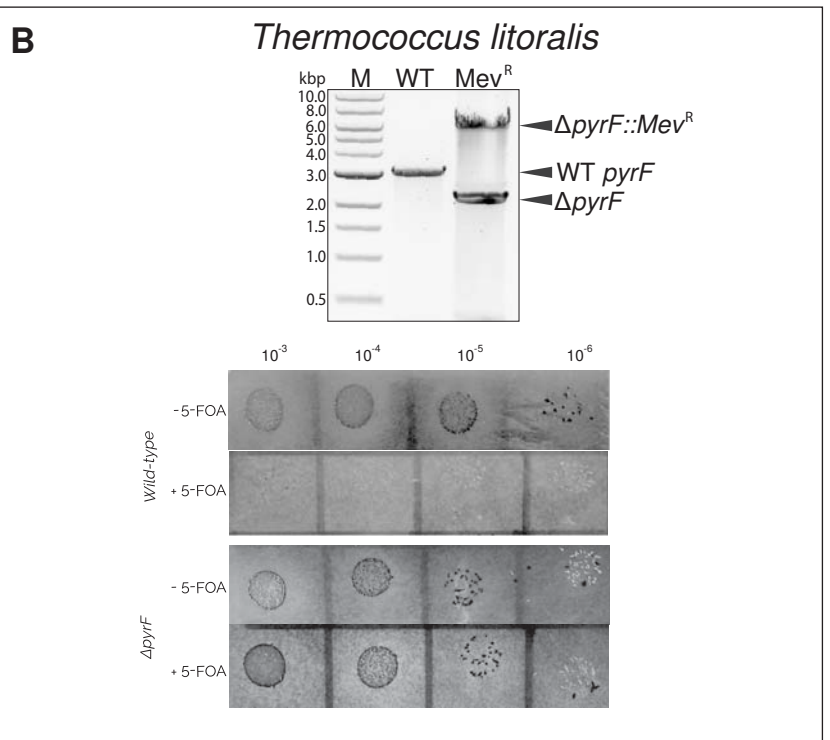
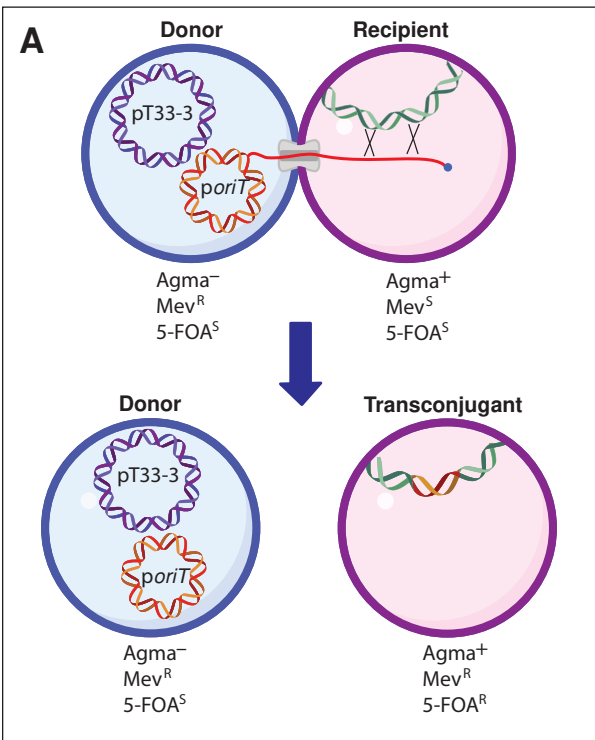


Figure #	Figure title One sentence only	Filename This should be the name the file is saved as when it is uploaded to our system. Please include the file extension. i.e.: <i>Smith_ED_Fig1.jpg</i>	Figure Legend If you are citing a reference for the first time in these legends, please include all new references 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 <i>T. sp.</i> 33-3 to other species.	Ex_data_fig_1.eps	Genome alignment between <i>Thermococcus. sp.</i> 33-3 and closely related strains or species, <i>T. nautili</i> (A); <i>T. henrieti</i> EXT12c (B); and <i>T. sp.</i> 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 <i>T. kodakarensis</i>	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) <i>T. kodakarensis</i> 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 $\sim 100\text{cfu}/\mu\text{g}$ 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, <i>p0019</i> and <i>p0132</i> . 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 <i>T. kodakarensis</i> strains does not require conjugation	Ex_data_fig_5.eps	A) <i>T. kodakarensis</i> incubated with purified plasmid DNA (presented as transformants per fg DNA for comparable scale) - <i>T. kodakarensis</i> is naturally competent for DNA uptake. B) Transfer of a non-conjugative plasmid from a <i>T. kodakarensis</i> donor to a plasmid-free recipient – plasmids transfer between <i>T. kodakarensis</i> strains occurs

			<p>by simple co-culturing. C) Transfer of a chromosomal prototrophic marker from a <i>T. kodakarensis</i> donor to an auxotrophic recipient - chromosomal markers transfer between <i>T. kodakarensis</i> strains. D) Transfer of a chromosomal prototrophic marker from a <i>T. gammatolerans</i> donor to an auxotrophic <i>T. kodakarensis</i> recipient - chromosomal markers are unable to transfer between <i>T. gammatolerans</i> and <i>T. kodakarensis</i>, suggesting allelic exchange is mediated by homologous recombination. E) Transfer of a chromosomal prototrophic marker from a <i>T. kodakarensis</i> 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.</p> <p>G-I) In contrast to <i>T. kodakarensis</i>, <i>T. nautili</i> is unable to receive a shuttle vector by co-culturing, whereas pT33-3 readily transfers (see main text and Figure 2C).</p>
Extended Data Fig. 6	Identification of <i>oriT</i> -encoding region of pT33-3.	Ex_data_fig_6.eps	<p>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 <i>T. kodakarensis</i> donors. A) <i>oriT1</i> candidate region. While plasmids encoding this region readily transferred between <i>T. kodakarensis</i> strains, minimal transfer was observed to <i>T. nautili</i> and <i>T. gammatolerans</i> recipients. B) <i>oriT2</i> candidate region. Plasmids encoding this region readily transferred from <i>T. kodakarensis</i> to <i>T. kodakarensis</i>, <i>T. nautili</i>, and <i>T. gammatolerans</i>, indicating this region encodes the <i>oriT</i> of pT33.3. C) <i>oriT2</i> region was split into four overlapping fragments. Plasmids encoding both <i>oriT2.1</i> and <i>oriT2.2</i> regions transferred to the non-competent recipients, whereas <i>oriT2.3</i> and <i>oriT2.4</i> did not, indicating the <i>oriT</i> is encoded by the overlap region between <i>oriT2.1</i> and <i>oriT2.2</i>. D) The 300bp overlap between <i>oriT2.1</i> and <i>oriT2.2</i> (named <i>oriT300</i>) also confers mobilization ability to plasmids, indicating that this region encodes the <i>oriT</i> of pT33.3.</p>

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 <sub>FATA</sub> group. B) Re-creation of MPF <sub>FATA</sub> and MPF <sub>FA</sub> phylogeny including VirB4 sequences from integrated elements in other archaeal genomes, and the VirB4 homologue from pT33-3. Phylogeny rooted with MPF <sub>F</sub> 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 <i>oriT</i> sequences and homologous recombination + selectable marker cassette for allelic exchange. B) Transfer of pMob initiating at <i>oriT1</i> and terminating at <i>oriT2</i> 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 <i>oriT2</i> and terminating at <i>oriT1</i> 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 <i>S. marinus</i> colonies following conjugation mediated mobilisation of a recombination substrate	Ex_data_fig_9.eps	Eight colonies were screened using primers binding to the <i>S. marinus</i> chromosome, outside the homology arms used in allelic exchange. Clean allelic exchange ( $\Delta apt::SimR$ ) 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.
------	----------	---	--

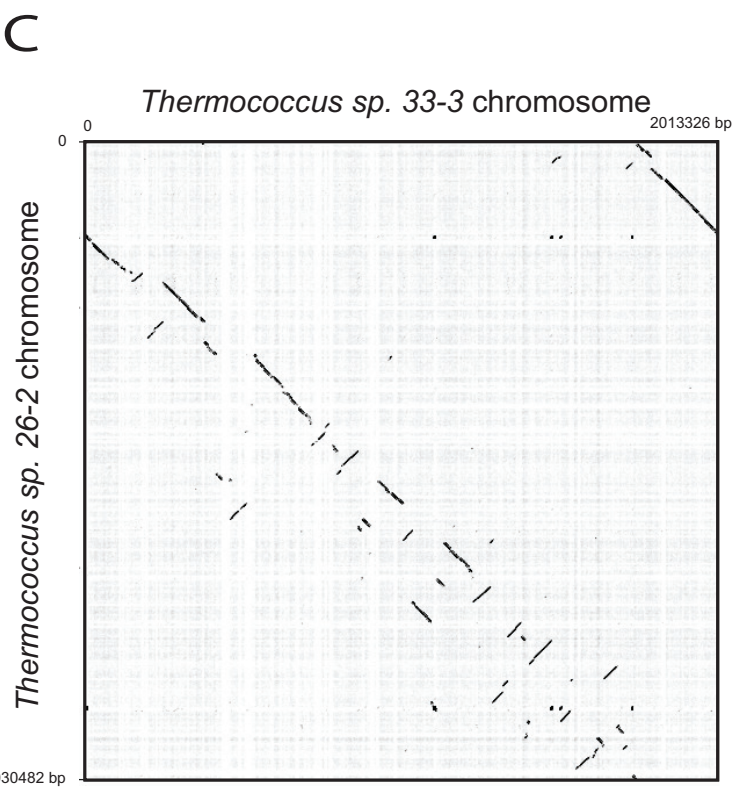
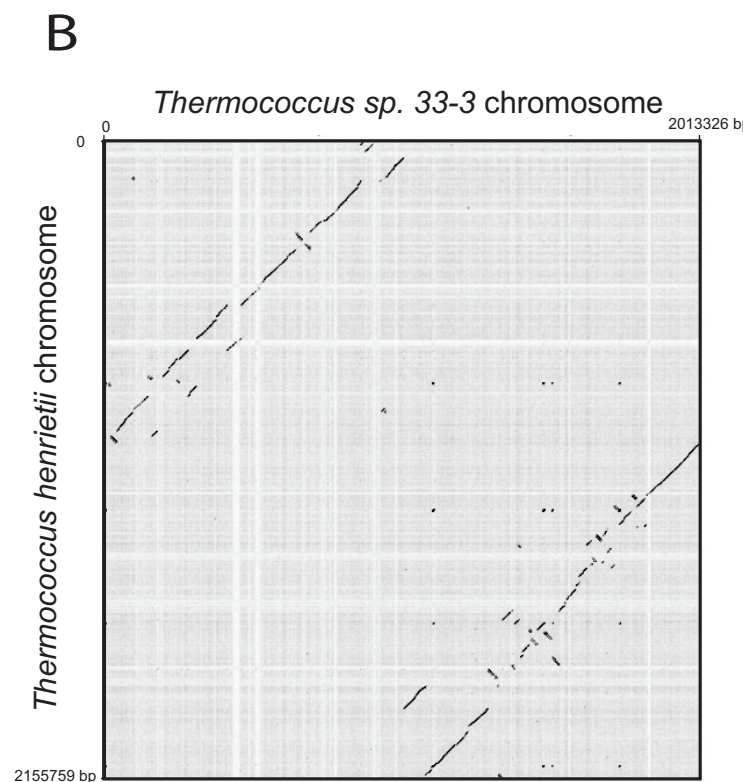
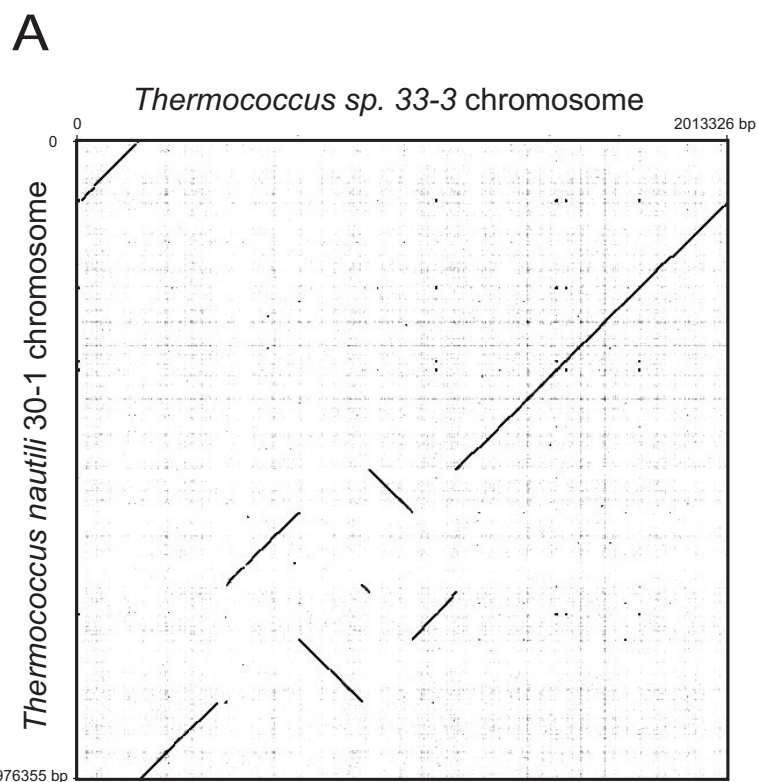
		and should include the file extension. The extension must be .pdf	
Supplementary Information	No		
Reporting Summary	Yes	Catchpole Flat RS.pdf	
Peer Review Information	No	OFFICE USE ONLY	

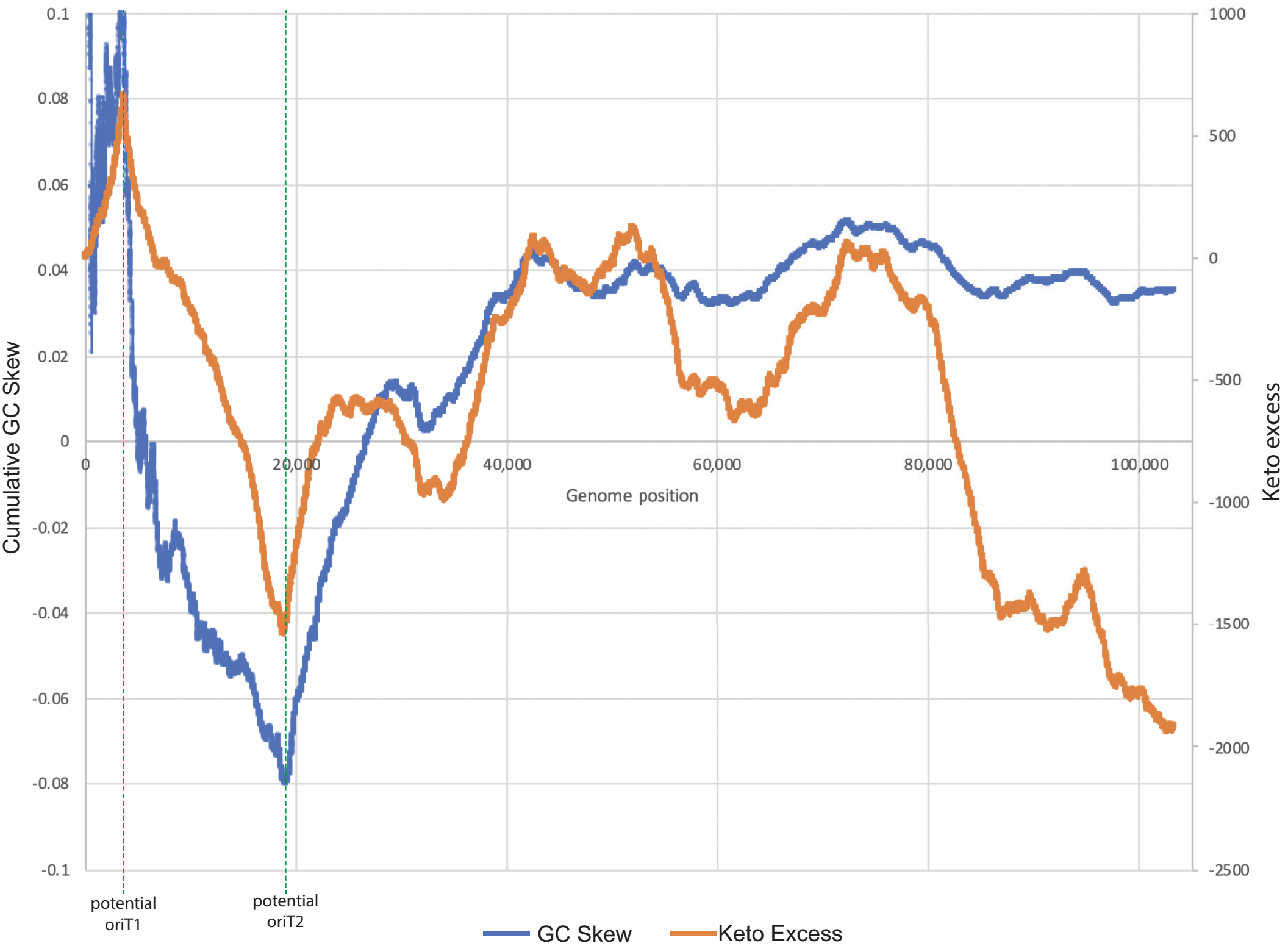
12

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.: <i>Smith_Supplementary_Video_1.mov</i>	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.

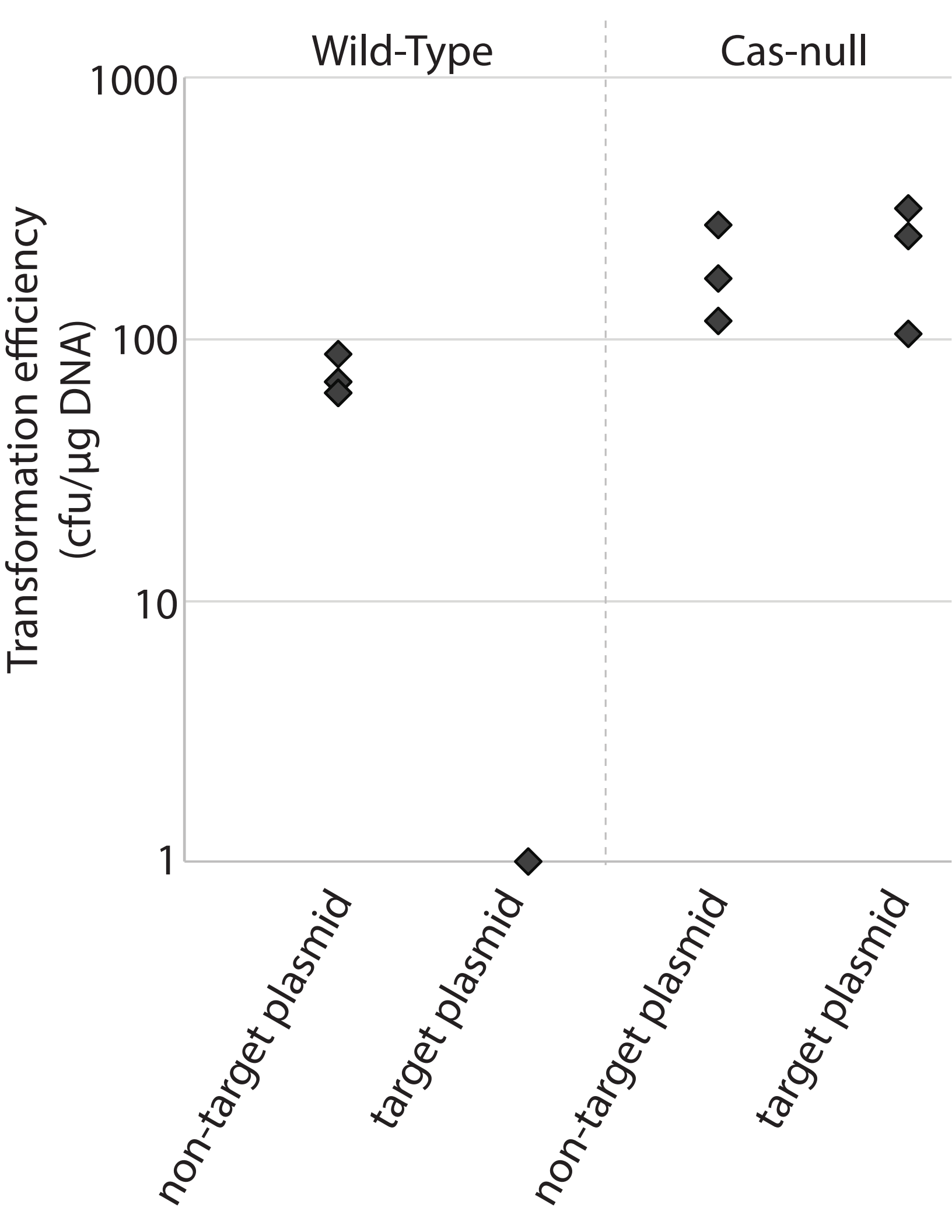
13

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.
------------------------	--	--

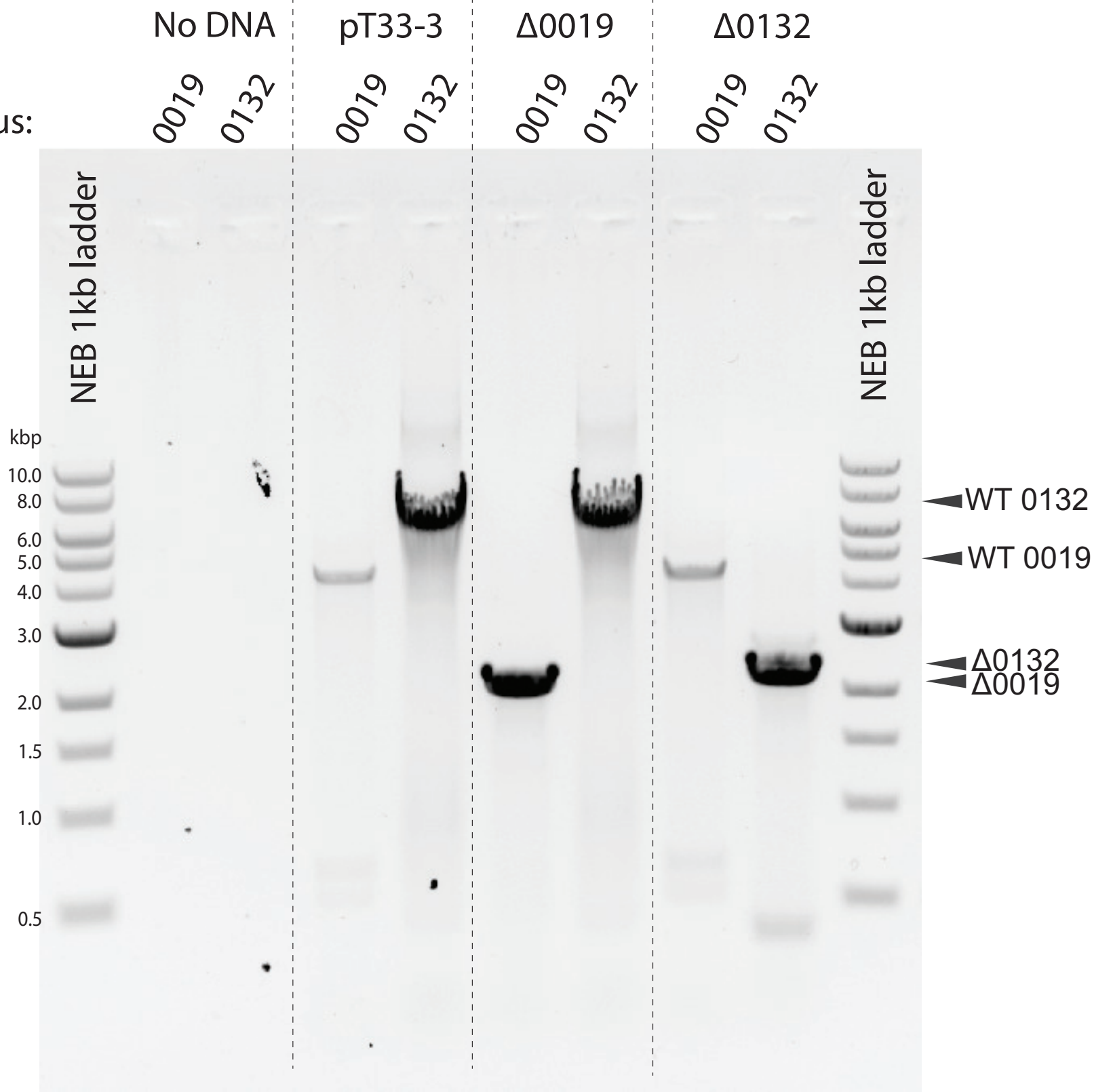


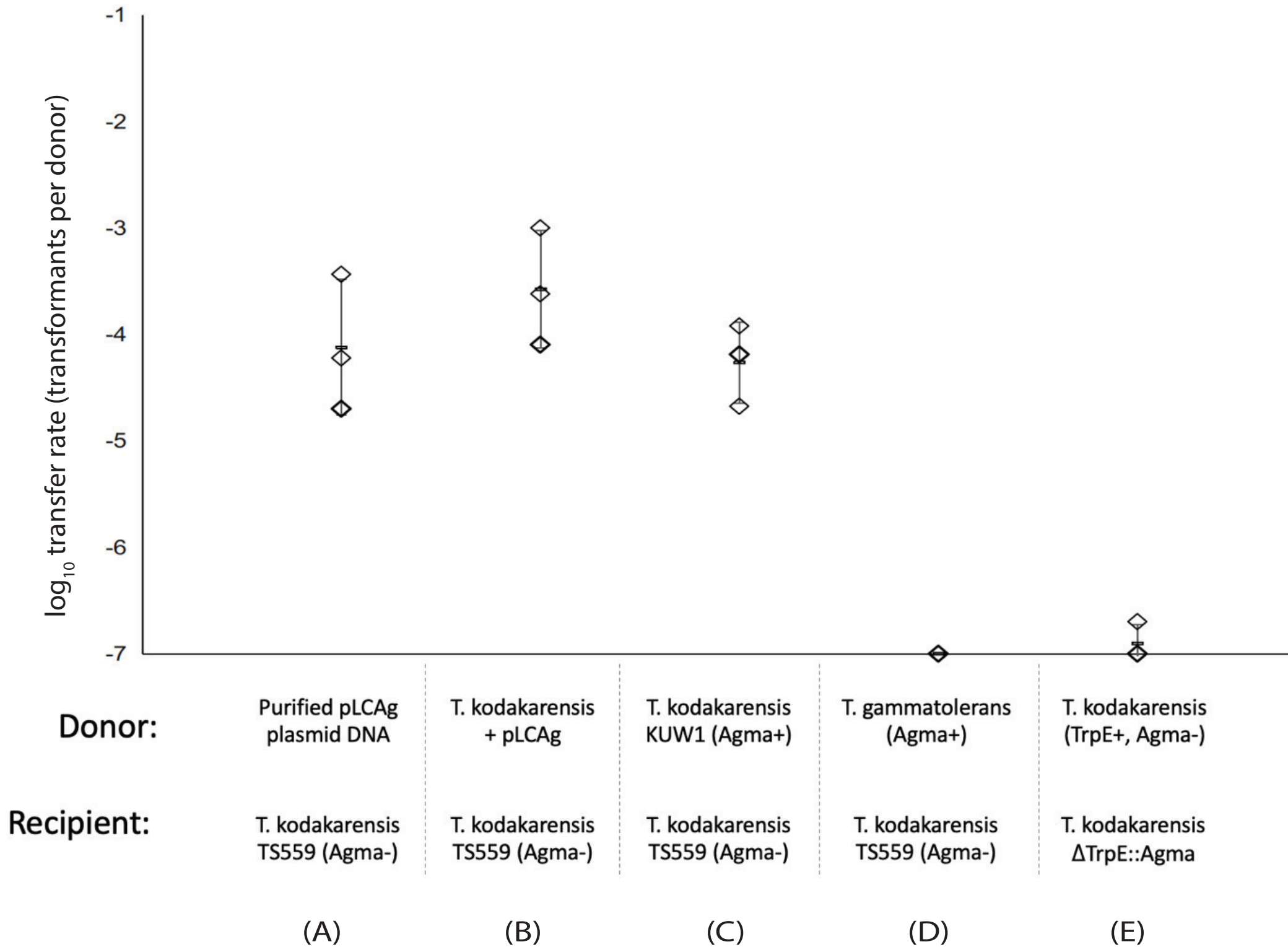


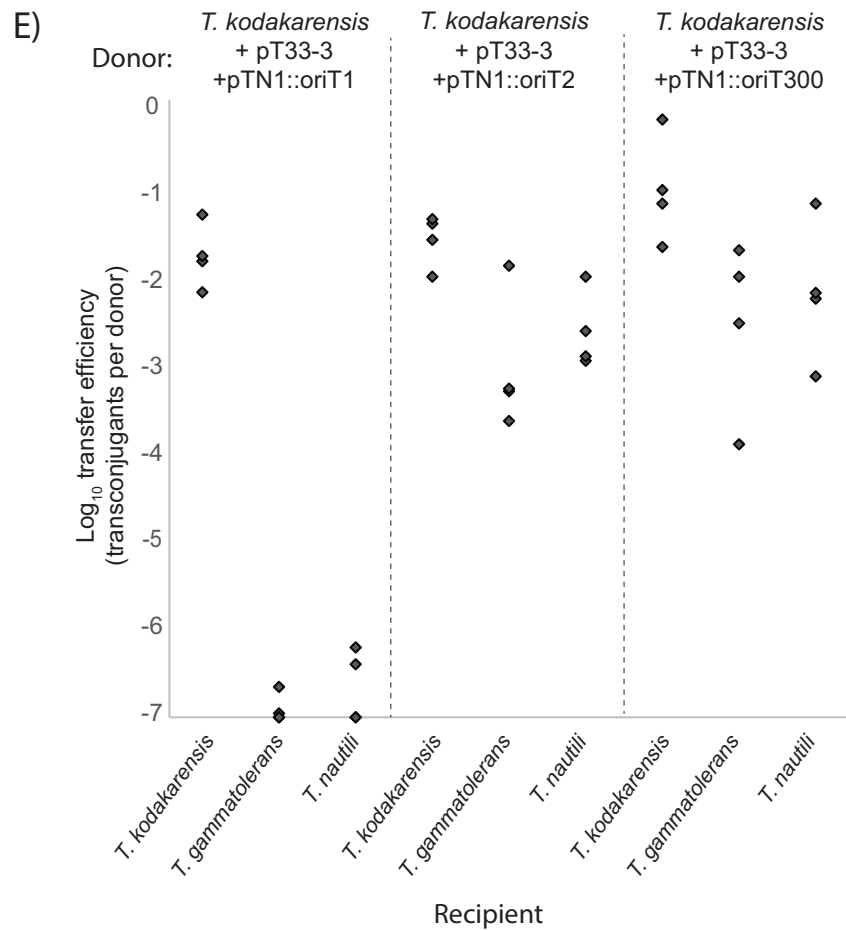
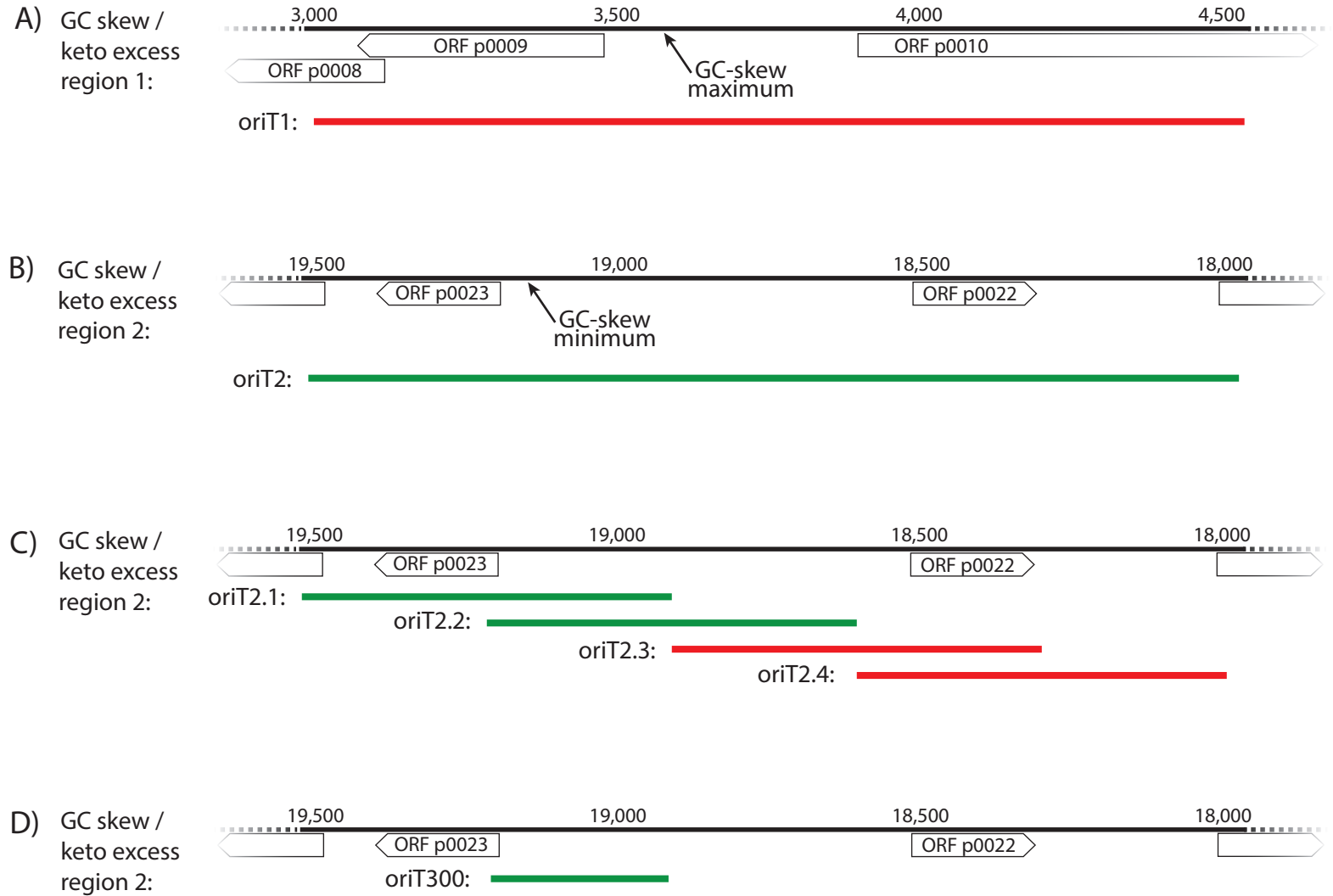


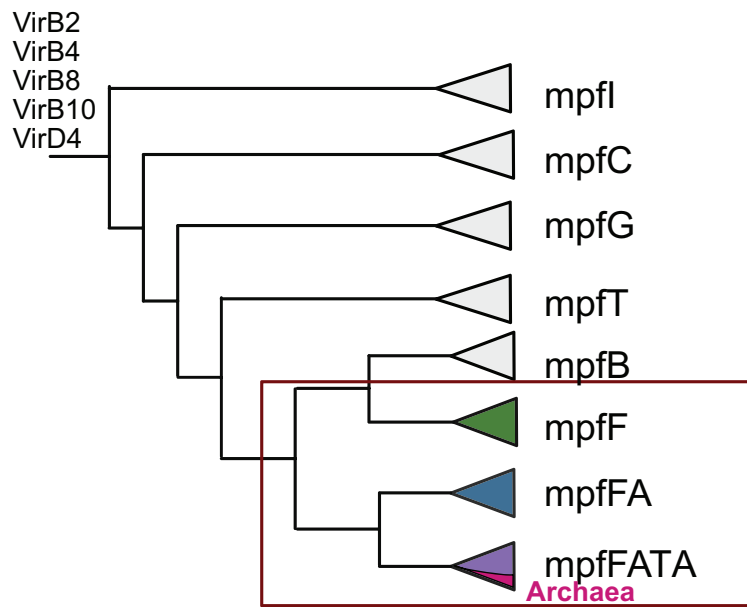
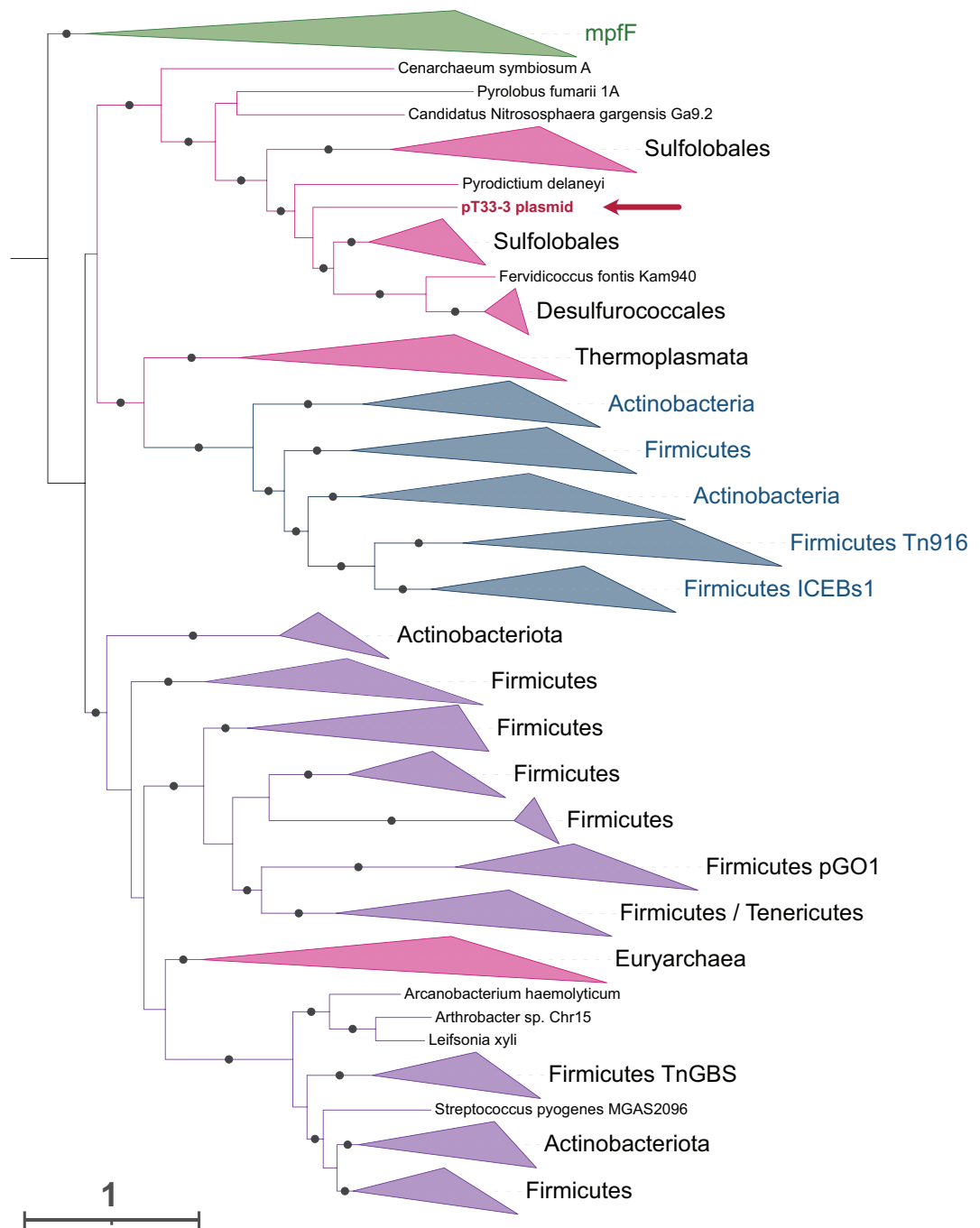


PCR locus:

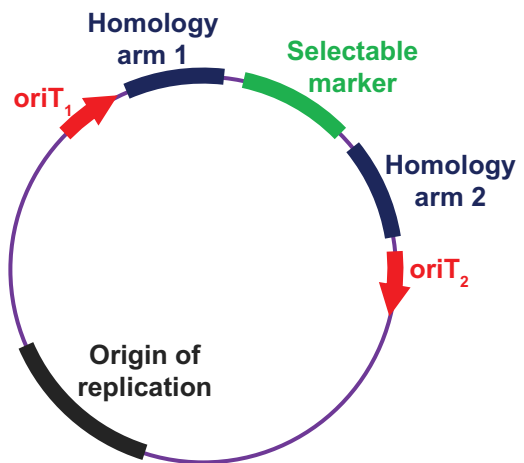




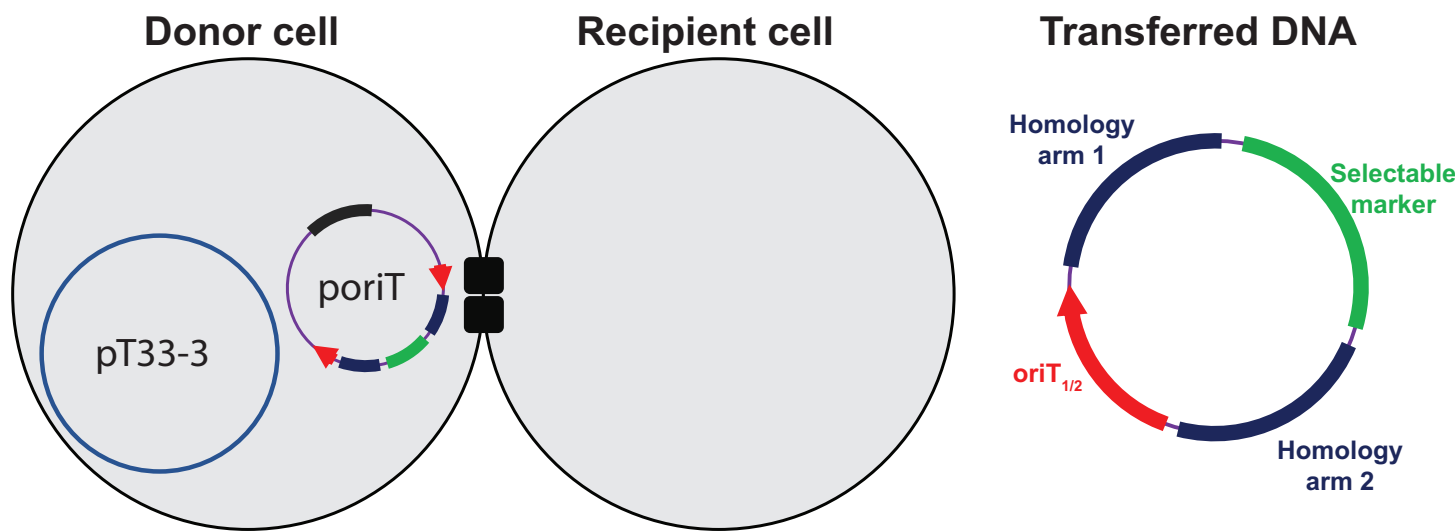


**A****B**

A)



B)



C)

