AEM Accepts, published online ahead of print on 22 July 2011
Appl. Environ. Microbiol. doi:10.1128/AEM.05057-11
Copyright © 2011, American Society for Microbiology and/or the Listed Authors/Institutions. All Rights Reserved.

AEM Accepts published online ahead of print

1

2 # Defining Components of the Chromosomal Origin of Replication
3 # of the Hyperthermophilic Archaeon, *Pyrococcus furiosus,* Needed
4 # for Construction of a Stable Replicating Shuttle Vector

5

6 Joel Farkas[1,3], Daehwan Chung[1,3], Megan DeBarry[1,3],
7 Michael WW Adams[2,3] and Janet Westpheling[1,3] *

8

9 Departments of Genetics[1] and Biochemistry and Molecular Biology[2]
10 University of Georgia, Athens, Georgia
11 and
12 The BioEnergy Science Center[3], Department of Energy,
13 Oak Ridge National Laboratory, Oak Ridge, Tennessee

14

15

16

17

18 *Corresponding author: Janet Westpheling
19 Department of Genetics, University of Georgia, Athens, Georgia 30602
20 Office phone: 706-542-1436 FAX: 706-542-3910
21 Email: janwest@uga.edu
22 Running title: Replicating Shuttle Vectors for *Pyrococcus furiosus*

AEM Accepts published online ahead of print

23 **ABSTRACT**

24 We report the construction of a series of replicating shuttle vectors that consist of a low copy
25 number cloning vector for *Escherichia coli* and functional components of the origin of
26 replication (*oriC*) of the chromosome of the hyperthermophilic archaeon, *Pyrococcus*
27 *furiosus*. In the process of identifying the minimum replication origin sequence required for
28 autonomous plasmid replication in *P. furiosus,* we discovered that several features of the
29 origin predicted by bioinformatic analysis and *in vitro* binding studies were not essential for
30 stable autonomous plasmid replication. A minimum region required to promote plasmid DNA
31 replication was identified and plasmids based on this sequence readily transformed *P.*
32 *furiosus*. The plasmids replicated autonomously and existed in single copy. In contrast to
33 shuttle vectors based on a plasmid from the closely related hyperthermophile *Pyrococcus*
34 *abyssi* for use in *P. furiosus*, plasmids based on the *P. furiosus* chromosomal origin were
35 structurally unchanged after transformation and were stable without selection for more than
36 one hundred generations.

37 **INTRODUCTION**

38 *Pyrococcus* species are hyperthermophilic marine archaea that grow anaerobically
39 at temperatures near and above 100 °C (19). Their interesting biology, evolutionary history
40 and potential commercial utility make them an important group to study (9, 47). We recently
41 reported an efficient method for DNA transformation for *P. furiosus* (27), which provides the
42 basis for development of methods for genetic manipulation. A markerless deletion of the
43 OMP decarboxylase (*pyrF*) locus was constructed in the *P. furiosus* genome generating a
44 mutant that is a uracil auxotroph and resistant to 5-fluroorotic acid (5FOA).
45 Complementation by the wild type *pyrF* allele restored uracil prototrophy and 5FOA
46 sensitivity to this strain. We took advantage of this marker to construct replicating shuttle

AEM Accepts published online ahead of print

47 vectors for use in *P. furiosus*.

48 Replicating vectors facilitate a variety of genetic manipulations and vectors capable
49 of shuttling between any host of interest and *Escherichia coli*, perhaps the most convenient
50 biological host for DNA manipulation, are especially useful. Naturally occurring plasmids
51 have been identified in archaeal species (7, 10-12, 16, 22, 51), and some of them have been
52 used to develop shuttle vectors in the haloarchaea (25) and methanogens (49). A shuttle
53 vector between *E. coli* and *Thermococcus kodakaraensis*, a close relative of *P. furiosus* (4,
54 36), was reported that combined a commercial *E. coli* cloning vector with a naturally
55 occurring plasmid, pTN1, from *Thermococcus nautilis* (44, 46). This plasmid was used
56 successfully to express a hemagglutinin (HA) epitope-tagged version of RpoL, a subunit of *T.*
57 *kodakaraensis* RNA polymerase, illustrating the utility of such vectors (44). The pGT5
58 plasmid, a naturally occurring plasmid in *Pyrococcus abyssi* (17), was used to create a shuttle
59 vector capable of replicating in both *P. abyssi* and *E. coli* (29) and most recently, a shuttle
60 vector based on pGT5 was used to express a gene encoding His6-tagged subunit D of RNA
61 polymerase in *P. furiosus* (50). Our attempts to construct a stable shuttle vector for *P. furiosus*
62 based on pGT5 were unsuccessful.

63 *Pyrococcus* species are predicted to have a single origin of replication with both
64 bacterial and eukaryotic features (28, 37). Other archaea (5, 30, 42) have multiple replication
65 origins and eukaryotes have hundreds or even thousands of replication origins (41).
66 Eukaryotic replication origins are poorly defined in terms of sequence as well as the number
67 and nature of auxiliary proteins that facilitate their function and there is increasing evidence
68 that origin maintenance in many eukaryotes is controlled by epigenetic factors whose
69 function is also poorly understood (3, 45). *Saccharomyces cerevisiae* is a notable exception
70 in that it has well-defined origins called autonomously replicating sequence (ARS) elements,

AEM Accepts published online ahead of print

71 generally 100-200 base pairs in length containing multiple *cis*-acting sequence elements (38),
72 and binding sites for the origin recognition complex (6). This complex is composed of the
73 Origin Recognition Complex (Orc) proteins 1-6 (6) and a DNA binding protein, Cdc6, first
74 identified in yeast as a gene whose product is involved in cell cycle control (15, 23). Most
75 bacteria have a single origin sequence flanked by polymerases and *dnaA* (35), which encodes
76 a protein that binds DNA within the origin (20) and functions to recruit DnaB and DnaC to
77 form the replication initiation complex. While DnaA is only distantly related to Cdc6, it
78 provides the same function in nucleating the protein complex. Replication origin regions in
79 both bacteria and archaea are AT-rich (5, 41).

80 The origin of replication in several archaeal genera was first predicted by
81 cumulative GC skew, an early bioinformatic method used to find prokaryotic replication
82 origins (21). In *P. abyssi* the location of the chromosomal replication origin was predicted to
83 be within an ~80 kb region that contained a large intergenic space and genes for several
84 putative replication proteins (28). This intergenic space, neighboring genes, and features are
85 conserved in *P. furiosus* (Figure 1). While the organization of putative protein binding sites
86 and replicative proteins around the *P. furiosus oriC* is bacterial in nature, the proteins that
87 likely bind the sequences are eukaryotic. In the location where *dnaA* often resides in bacteria,
88 there is a single gene homologous to the eukaryotic replication proteins *cdc6/orc1* (40).
89 Studies of the function of the origin sequences of *Halobacterium* NRC-1 *in vivo,* showed that
90 a 750 bp region containing *oriC* requires the adjacent *orc7* gene in *cis* for autonomous
91 plasmid replication (8). In *Sulfolobus solfataricus*, there are three origins of replication (30,
92 42), and for each a *cdc6* gene is adjacent but is not required in *cis* for the origin to function in
93 plasmid replication (14). Evidence for the location of the *P. furiosus* origin and for the role of
94 putative replication proteins is based entirely on bioinformatic and *in vitro* analysis. Cells

AEM Accepts published online ahead of print

95 emerging from replication arrest were shown to incorporate radiolabeled nucleotides in the
96 DNA sequence at the predicted origin (37). Chromatin immunoprecipitation of the
97 Cdc6/Orc1 protein showed significant enrichment of the predicted *oriC* DNA region (32) and
98 DNA binding studies showed that the *Sulfolobus* Cdc6/Orc1 protein binds this region of *P.*
99 *abyssi* as well as *Sulfolobus* DNA *in vitro* (33, 42). The sequence of the sites of protein
100 binding (referred to as origin recognition boxes - ORBs) suggested that inverted repeat
101 elements flanking the predicted origin, conserved among archaea, were involved in
102 replication complex formation and binding. Binding of the Cdc6/Orc1 protein to the
103 intergenic space resulted in unwinding of the DNA supporting the notion that this was the
104 location of the origin (34). While the intergenic space is AT-rich, the DNA sequence is not
105 conserved across *Pyrococcus* species. Further evidence for the position of the origin comes
106 from the fact that DNA at the origin contains a transient replication bubble produced by
107 bidirectional DNA polymerization and the structure may be isolated by displaying digested
108 total genomic DNA, from actively dividing cells, on 2-dimensional gels. Sequences
109 containing a replication bubble were located to a 1 kb fragment that included the intergenic
110 space, possibly overlapping the *cdc6* gene (37). Bioinformatic analysis identified several 13
111 bp mini-ORB repeats in and around the intergenic space suggested to be involved in origin
112 recognition and potential Cdc6/Orc1 binding (26, 33).

113 Here we present an *in vivo* analysis of sequences at the origin of DNA replication in the
114 *P. furiosus* chromosome and their function in autonomous plasmid replication. DNA frag-
115 ments containing the putative chromosomal origin as well as predicted protein binding sites
116 and the *cdc6/orc1* gene were cloned into an *E. coli* plasmid vector and tested for the ability to
117 promote autonomous DNA replication in *P. furiosus*. Cloning of the origin sequence without
118 the *cdc6/orc1* gene did not affect its ability to function suggesting that this protein, while

119 likely required for DNA replication, is not required in *cis*. The location of the origin as pre-
120 dicted from *in vitro* analysis was confirmed. Only two of the predicted ORB sequences, how-
121 ever, are required for autonomous plasmid replication. We used this origin sequence in com-
122 bination with the pSC101 origin from *Salmonella panama* (13) for replication in *E. coli* to
123 construct a replicating shuttle vector for *P. furiosus* that is stable in single copy without selec-
124 tion for more than one hundred generations and is structurally unchanged after transforma-
125 tion into *P. furiosus* and back-transformation into *E. coli*.
126

127 **METHODS & MATERIALS**

128 **Strains, media and growth conditions.** *E. coli* strain DH5α was used for plasmid DNA
129 constructions and preparations. Standard techniques for *E. coli* were performed as described
130 (43). Apramycin was used for selection at 50 mg/ml. *P. furiosus* DSM 3638 (19) wild type
131 and *P. furiosus* COM1 Δ*pyrF* strains (27) were grown anaerobically in a defined medium
132 with cellobiose as carbon source (27) at 90 °C for 16-20 hours in 100 mL serum bottles
133 containing 50 mL of liquid medium or on medium solidified with phytagel (1% w/v) for 60
134 hours. The *P. furiosus* COM1 Δ*pyrF* strain was used as host for all DNA transformation
135 experiments. *P. abyssi* strain GE5 (16, 18) was grown in a liquid base salts medium (1)
136 containing 0.5% (w/v) casein hydrolysate and 0.2% (w/v) elemental sulfur for 40-48 hours at
137 90 °C under anaerobic conditions. Total genomic DNA was isolated as described previously
138 (27) except that DNA was precipitated with isopropanol and resuspended with 50 μL TE
139 Buffer (10mM Tris, 1mM EDTA) containing RNaseA (100 ng/ml).
140 **Construction of vectors and transformation of *P. furiosus*.** To construct pJFW027 and
141 pJFW018, PCR products containing the indicated regions of the chromosome (Figure 1) were
142 ligated into a linear DNA fragment containing the entire pJFW017 plasmid (Figure 2) also

AEM Accepts published online ahead of print

AEM Accepts published online ahead of print

143 generated by PCR using primers JF266 and JF267. To generate plasmids pJFW031 to
144 pJFW044, primers with restriction sites added to the 5'ends were used to allow digestion and
145 subsequent directional cloning of origin containing fragments into pJFW017. The 5' end of
146 each fragment contained a *Bam*HI site and the 3' end contained a *Cla*I site. PCR
147 amplification of pJFW017 used primers JF266.2 and JF267.2 with the same restriction sites.
148 Primers used in these constructions are listed in Table 1 and DNA sequences of the primers
149 are shown in Supplemental Material (Table S1). *E. coli* strain DH5α was transformed by
150 electroporation in a 2 mm gap cuvette, at 2.5 V. Plasmid DNA was isolated from liquid
151 cultures using QIAprep Spin Miniprep columns (Qiagen Inc.). For DNA transformations, the
152 *P. furiosus* COM1 Δ*pyrF* strain was grown for 16-20 hours in the defined liquid medium
153 containing 20 μM uracil. Plasmid DNA (100 to 200 ng) was added to 100 μL of culture and
154 plated onto the defined medium without uracil. Prototrophic colonies were inoculated into
155 liquid medium for DNA isolation. The presence of plasmid sequences in *P. furiosus* was
156 confirmed by PCR amplification of the *aac* gene, present only on the plasmid, from *P.*
157 *furiosus* total genomic DNA using primers JF263 and JF264 (Table1).
158 **Assessment of plasmid maintenance, stability and copy number.** To assess plasmid
159 maintenance, *P. furiosus* transformants were serially subcultured every 24 hours for 10 days
160 in selective and nonselective liquid media. After each passage, the culture was diluted 100-
161 fold with base salts and 30 μL of diluted culture was plated onto selective medium to
162 determine the number of prototrophic colonies, i.e. those maintaining the plasmid. The cell
163 density of the liquid culture was determined by direct cell counting using a Petroff-Hausser
164 counting chamber. To assess the structural stability of the plasmid, total genomic DNA
165 isolated from five independent *P. furiosus* transformants containing pJFW027 was used to
166 back-transform *E. coli* for plasmid isolation and restriction digestion analysis. To determine

AEM Accepts published online ahead of print

167 plasmid copy number, total genomic DNA was isolated from *P. furiosus* plasmid
168 transformants and digested twice with 10 U of HpaI for 120 minutes at 37 °C. The
169 restriction fragments were separated by electrophoresis in a 1.0% (w/v) agarose gel, and
170 transferred to nylon membranes (Roche, Manheim Germany). Primers GL021 and GL023
171 (27) were used to amplify the glutamate dehydrogenase (*gdh*) promoter from wild type *P.*
172 *furiosus* total genomic DNA to generate a DIG-labeled probe by random priming with DIG
173 High Prime DNA Labeling and Detection Starter Kit I (Roche, Manheim Germany). The
174 membrane was incubated at 42 °C and washed at 65 °C. Band intensities were determined
175 using a Storm 840TM PhospoImager (GE Healthcare) equipped with ImageQuant v.5.4.
176 software (Molecular Dynamics).
177
178 **RESULTS AND DISCUSSION**
179 **The *P. furiosus* chromosomal replication origin functions for stable autonomous**
180 **plasmid replication.** Attempts to construct a stable replicating shuttle vector based on the
181 pGT5 plasmid from *P. abyssi* for use in *P. furiosus* were unsuccessful. Plasmids based on
182 pGT5 exist in high copy number in *P. abyssi* (29, 50) but show a significantly reduced copy
183 number in *P. furiosus* (50) and cannot be used for selection of transformants in the closely
184 related *Thermococcus kodakarensis* (44). In an attempt to construct a shuttle vector based on
185 pGT5, the entire plasmid was cloned into pJFW017 that contained a pSC101 origin for
186 replication in *E. coli*, an apramycin resistance gene for selection in *E. coli* and a wild type
187 copy of the *P. furiosus pyrF* gene for selection of transformants in the *P. furiosus* COM
188 Δ*pyrF* strain. A fragment containing the entire pGT5 plasmid sequence was produced by
189 PCR amplification with primers JF254 and JF270, linearizing the plasmid at a site within
190 pGT5 previously shown not to interfere with replication functions (17, 31), to produce

AEM Accepts published online ahead of print

191 pJFW019 (Supplemental figure S1). This plasmid readily transformed *P. furiosus* but was
192 rapidly lost without selection (Table 2) and showed internal deletions after transformation
193 into *P. furiosus* and subsequent back-transformation into *E. coli* (Supplemental figure S2).
194 Other attempts to use pGT5 for the construction of shuttle vectors in *T. kodakarensis* were
195 similarly unsuccessful (44).

196 To test whether the predicted *P. furiosus* chromosomal origin of replication could
197 promote stable autonomous plasmid replication, a fragment of the chromosome containing
198 the predicted origin sequence and the gene encoding Cdc6/Orc1 (Figure 1) was cloned into
199 an *E. coli* plasmid, pJFW017 (Figure 2) to make pJFW027. We used transformation
200 efficiency as an assay for plasmid replication (8). As shown in Table 1, transformants of
201 pJFW027 were observed at a frequency of 5.8×10^5 transformants per µg of plasmid DNA.
202 No transformants were observed in the absence of added plasmid DNA and while some
203 transformants were obtained in experiments with pJFW017, which does not contain origin
204 sequence, (8.2×10^2 transformants per µg of plasmid DNA), this is most likely due to
205 integration by homologous recombination between the *gdh* promoter region (283 bp) driving
206 transcription of the *pyrF* gene on the plasmid and the *gdh* locus in the chromosome. In fact,
207 we have observed integration of non-replicating plasmid DNA by homologous recombination
208 at the same frequency (27). The transformation frequency of pJFW027 was a thousand-fold
209 greater than that of pJFW017 indicating that the plasmid was replicating autonomously. PCR
210 amplification of the apramycin resistance gene, contained only on the plasmids, was used to
211 confirm the presence of plasmid DNA in the transformants. A 950 bp product containing this
212 sequence was obtained from transformant total genomic DNA but not from the wild type or
213 the *P. furiosus* COM1 Δ*pyrF* strain (Supplemental figure S3).

214 Attempts to isolate a significant quantity of pJFW027 plasmid DNA from *P.*

215 *furiosus* were unsuccessful. This is perhaps not surprising since plasmids based on the
216 chromosomal origin would be expected to be low or one copy per chromosome. In lieu of
217 direct plasmid isolation, we chose to rescue the plasmid by back-transformation to *E. coli*.
218 That transformants contained a replicating copy of the plasmid was shown using total
219 genomic DNA isolated from 5 independent plasmid transformants to back-transform *E. coli*
220 strain DH5α selecting apramycin resistance. Back-transformants were obtained at
221 frequencies of 10^4 transformants per μg of DNA, an underestimate of plasmid transformation
222 since this frequency is based on the number of transformants per microgram of total genomic
223 DNA, and only covalently closed circular plasmid DNA is capable of transforming *E. coli*
224 strain DH5α at this frequency (24). Plasmid DNA isolated from these back-transformants
225 was indistinguishable from the pJFW027 plasmid DNA used to transform *P. furiosus* using
226 restriction digestion analysis, indicating that there were no gross rearrangements during
227 transformation and replication in *P. furiosus* or subsequent back-transformation to *E. coli*.
228 When the P_{gdh} fragment, specific to the pJFW027 plasmid, was used as probe in Southern
229 hybridization of total genomic DNA from the *P. furiosus* transformants with DNA digested
230 with either EcoRV (data not shown) or HpaI (Figure 3), which have a single cleavage site
231 within the plasmid, a single band was detected showing that the plasmid DNA was not
232 integrated into the chromosome and existed as an autonomously replicating molecule.
233 To examine plasmid maintenance, transformants of pJFW027 and pJFW018 were
234 serially subcultured in liquid medium with or without selection (i.e. in the absence or
235 presence of uracil) for more than one hundred generations and then plated on minimal
236 medium without uracil to assay plasmid maintenance. No loss of plasmids with *oriC* was
237 detected even without selection (Table 2). In addition, the restriction pattern of plasmid DNA
238 isolated from *E. coli* after transformation into *P. furiosus* and subsequent transformation back

AEM Accepts published online ahead of print

239 into *E. coli* remained unchanged indicating that no rearrangements of the plasmid DNA
240 occurred (Figure 4, one of 10 transformants tested is shown).
241
242 **The *cdc6/orc1* open reading frame is not required in *cis* for replication origin function.**
243 In bacteria, *oriC* is often, but not always adjacent to *dnaA*. In *E. coli*, *oriC* is between *gidA*
244 and *mioC* (which encodes another replication protein), approximately 43 kb from *dnaA*, and
245 can function for autonomous replication of plasmids without *cis*-acting replicating
246 components (39, 48). In the chromosome of *Halobacterium* NRC-1, the *oriC* requires the
247 adjacent *orc7* gene in *cis* for autonomous plasmid replication (8). In *Sulfolobus solfataricus*,
248 there are three origins of replication and for each a *cdc6* gene is adjacent but is not required
249 in *cis* for the origin to function in autonomous plasmid replication (14). In the sequenced
250 *Pyrococcus* species, *P. furiosus*, *P. abyssi*, and *P. horikoshii* as well as the closely related *T.*
251 *kodakarensis* there is a single *oriC* adjacent to a *cdc6/orc1* homologue but nothing is known
252 about the requirement of this protein for *oriC* function. To test whether *cdc6/orc1* was
253 required in *cis* for autonomous plasmid replication in *P. furiosus*, a fragment containing only
254 *oriC* was cloned into the parent pJFW017 plasmid to generate pJFW18 (Figure 1). As shown
255 in Table 1, plasmids containing the fragment with only the *oriC* sequence without the
256 *cdc6/orc1* gene transformed *P. furiosus* as efficiently and were maintained as stably as the
257 pJFW27 plasmid containing the *cdc6/orc1* gene suggesting that the *cdc6/orc1* gene is not
258 required in *cis* for stable autonomous plasmid replication.
259
260 **Only two of the predicted ORB sequences and part of the predicted chromosomal**
261 **origin sequence are required for plasmid replication.** The origin region was predicted to
262 contain several ORB and mini-ORB sequences (33), suggested to be binding sites for the

AEM Accepts published online ahead of print

AEM Accepts published online ahead of print

263 replication initiation protein Cdc6/Orc1, which is presumed to facilitate nucleation of the
264 replication complex. Our analysis using the BLASTN 2.2.24+ algorithm (2) identified three
265 ORB repeats and several mini-ORB repeats by self-alignment of the sequence of the genomic
266 region containing the *oriC* and neighboring genes. These results are similar, but not identical
267 to the analysis of Matsunaga et al (33), in that we found a clustering of mini-ORB repeats in
268 and around the *oriC*, but the exact number and position of these mini-ORB repeats was
269 different. In addition, we identified two conserved palindromic sequences (Table 3)
270 conserved in all sequenced *Pyrococcus* species. One of them contains compensating changes
271 within the sequence that retain the perfect palindromic structure, suggesting that these are not
272 random sequences within this highly repetitive region of DNA and may potentially be
273 binding sites for other replication proteins, or have a structural role in replication. These
274 palindromes are not present in the *oriC* region of the closely related *Thermococcus* species,
275 however, suggesting that if they have a function it may be specific to *Pyrococcus.* To test
276 whether these sequences were required for autonomous plasmid replication, plasmids
277 containing various portions of the region around the predicted origin were constructed and
278 tested for the ability to replicate. The smallest insert able to promote autonomous plasmid
279 replication was the 653 bp fragment cloned into pJFW033. As shown in Table 1, only two of
280 the three ORB sequences were required for plasmid replication, and only a part of the
281 sequence predicted to contain the origin. The predicted unwinding site, for example, is
282 apparently not required for autonomous plasmid replication.
283
284 **Replicating shuttle vectors based on the chromosomal origin exist in single copy.** To
285 determine the approximate copy number of the *oriC*-based plasmids, a PCR product
286 generated from the P_{gdh} promoter was used as probe in Southern hybridization experiments

287 with total genomic DNA from *P. furiosus* wild type cells and pJFW027 transformants. Since

288 P_{gdh} is present in one copy on both pJFW027 and in the *P. furiosus* chromosome, a

289 densitometry analysis of the amount of DIG-labeled probe hybridized to each allowed an

290 estimation of the number of plasmid copies per chromosome (Figure 3). The relative

291 intensities of plasmid to chromosomal copy of P_{gdh} for the EcoRV and HpaI digests ranged

292 from 1.4 to 1.8 for 10 transformants tested indicating that the *oriC*-based plasmids exist in

293 single copy.

294

295 **CONCLUSIONS**

296 The functional analysis of the replication origin of the *P. furiosus* chromosome reported here

297 showed that only two of the three ORB sequences, those flanking an AT-rich sequence most

298 conserved in arrangement and sequence among the Archaea (42), and no more than three of

299 the mini-ORB sites are required for autonomous plasmid replication. In particular, the DNA

300 unwinding site, predicted by P1 endonuclease assays (34), is not required for autonomous

301 plasmid replication nor are any of the predicted ORB or mini-ORB sequences within the

302 DNA polymerase small subunit open reading frame. We emphasize that we have not ruled

303 out the possibility that these sequences are important for chromosomal replication and that

304 they may serve to promote additional Cdc6/Orc1 binding for chromosomal replication

305 initiation. The open reading frame encoding the Cdc6/Orc1 protein present adjacent to the

306 predicted origin sequence is not required in *cis* for autonomous plasmid replication. Vectors

307 based on the *P. furiosus oriC* were stably maintained for more than 100 generations without

308 selection and show no evidence of rearrangement after replication and transformation

309 between *E. coli* and *P. furiosus*. The smallest *oriC* fragment identified in this study capable of

310 conferring autonomous replication was 653 bp in length and vectors based on the origin exist

AEM Accepts published online ahead of print

311 in single copy per chromosome in the cell. Two conserved short palindromes were identified
312 within the origin region that are conserved among *Pyrococcus* species but not in the closely
313 related species *Thermococcus kodakaraensis*, suggesting that if they have a function it may
314 be specific to *Pyrococcus* species. We anticipate that these vectors will have utility for
315 homologous and heterologous gene expression, as well as providing a tool for the study of
316 natural competence, and *in vivo* studies of replication and recombination in *P. furiosus*.
317

318 **ACKNOWLEDGEMENTS**

319 We are especially grateful to Gina Lipscomb and Karen Stirrett for discussions and advice
320 throughout the course of this work and to Sidney Kushner for guidance in the design of
321 experiments testing DNA replication. We also thank Jennifer Huddleston for critical review
322 of the manuscript. This work was supported by a grant to MWWA and JW from the
323 BioEnergy Science Center (DE-PS02-06ER64304) administered by Oak Ridge National
324 Laboratory and by the Office of Biological and Environmental Research (FG02-08ER64690)
325 in the DOE Office of Science.
326

327 **REFERENCES**

328 1. **Adams, M. W. W., J. F. Holden, A. L. Menon, G. J. Schut, A. M. Grunden, C. Hou, A. M.
329 Hutchins, F. E. J. Jr., C. Kim, K. Ma, G. Pan, R. Roy, R. Sapra, S. V. Story, and M. F. J.
330 M. Verhagen.** 2001. Key role for sulfur in peptide metabolism and in regulation of three
331 hydrogenases in the hyperthermophilic archaeon *Pyrococcus furiosus*. Journal of Bacteriology
332 **183:**716-724.

333 2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J.
334 Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search

AEM Accepts published online ahead of print

AEM Accepts published online ahead of print

335 programs. Nucleic Acids Res **25:**3389-402.

336 3. **Antequera, F.** 2004. Genomic specification and epigenetic regulation of eukaryotic DNA
337 replication origins. EMBO J **23:**4365-70.

338 4. **Atomi, H., T. Fukui, T. Kanai, M. Morikawa, and T. Imanaka.** 2004. Description of
339 *Thermococcus kodakaraensis* sp. nov., a well studied hyperthermophilic archaeon previously
340 reported as *Pyrococcus* sp. KOD1. Archaea **1:**263-267.

341 5. **Barry, E. R., and S. D. Bell.** 2006. DNA Replication in the Archaea. Microbiology and
342 Molecular Biology Reviews **70:**876-887.

343 6. **Bell, S. P., and B. Stillman.** 1992. ATP-dependent recognition of eukaryotic origins of DNA
344 replication by a multiprotein complex. Nature **357:**128-34.

345 7. **Benbouzid-Rollet, N., P. Lopez-Garcia, L. Watrin, G. Erauso, D. Prieur, and P. Forterre.**
346 1997. Isolation of new plasmids from hyperthermophilic Archaea of the order *Thermococcales*.
347 Research in Microbiology **148:**767-775.

348 8. **Berquist, B. R., and S. DasSarma.** 2003. An archaeal chromosomal autonomously replicating
349 sequence element from an extreme halophile, *Halobacterium* sp. Strain NRC-1. Journal of
350 Bacteriology **185:**5959-5966.

351 9. **Blumer-Schuette, S. E., I. Kataeva, J. Westpheling, M. W. W. Adams, and R. M. Kelly.**
352 2008. Extremely thermophilic microorganisms for biomass conversion: status and prospects.
353 Current Opinion in Biotechnology **19:**210-217.

354 10. **Bokranz, M., A. Klein, and L. Meile.** 1990. Complete nucleotide sequence of plasmid
355 pME2001 of Methanobacterium thermoautotrophicum (Marburg). Nucleic Acids Res **18:**363.

356 11. **Charbonnier, F., G. Erauso, T. Barbeyron, D. Prieur, and P. Forterre.** 1992. Evidence that a
357 plasmid from a hyperthermophilic archaebacterium is relaxed at physiological temperatures.
358 Journal of Bacteriology **174:**6103-6108.

AEM Accepts published online ahead of print

359 12. **Charlebois, R. L., W. L. Lam, S. W. Cline, and W. F. Doolittle.** 1987. Characterization of
360 pHV2 from *Halobacterium volcanii* and its use in demonstrating transformation of an
361 archaebacterium. Proceedings of the National Academy of Science **84:**8530-8534.
362 13. **Cohen, S. N., A. C. Chang, H. W. Boyer, and R. B. Helling.** 1973. Construction of
363 biologically functional bacterial plasmids in vitro. Proc Natl Acad Sci U S A **70:**3240-4.
364 14. **Contursi, P., F. M. Pisani, A. Grigoriev, R. Cannio, S. Bartolucci, and M. Rossi.** 2004.
365 Identification and autonomous replication capability of a chromosomal replication origin from
366 the archaeon *Sulfolobus solfataricus*. Extremophiles **8:**385-391.
367 15. **Culotti, J., and L. H. Hartwell.** 1971. Genetic control of the cell division cycle in yeast. 3.
368 Seven genes controlling nuclear division. Exp Cell Res **67:**389-401.
369 16. **Erauso, G., F. Charbonnier, T. Barbeyron, P. Forterre, and D. Prieur.** 1992. Preliminary
370 characterization of a hyperthermophilic archaebacterium with a plasmid, isolated from a north
371 Fiji basin hydrothermal vent. Les Comptes Rendus de l'Académie des sciences **314:**387-393.
372 17. **Erauso, G., S. Marsin, N. Benbouzid-Rollet, M.-F. Baucher, T. Barbeyron, Y. Zivanovic, D.**
373 **Prieur, and P. Forterre.** 1996. Sequence of plasmid pGT5 from the archaeon *Pyrococcus*
374 *abyssi*: evidence for rolling-circle replication in a hyperthermophile. Journal of Bacteriology
375 **178:**3232-3237.
376 18. **Erauso, G., A.-L. Reysenbach, A. Godfroy, J.-R. Meunier, B. Crump, F. Partensky, J. A.**
377 **Baross, V. Marteinsson, G. Barbier, N. R. Pace, and D. Prieur.** 1993. *Pyrococcus abyssi* sp.
378 nov., a new hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent. Archives
379 of Microbiology **160:**338-349.
380 19. **Fiala, G., and K. O. Stetter.** 1986. *Pyrococcus furiosus* sp. nov. represents a novel genus of
381 marine heterotrophic archaebacteria growing optimally at 100°C. Archives of Microbiology
382 **145:**56-61.

AEM Accepts published online ahead of print

383 20. **Fuller, R. S., B. E. Funnell, and A. Kornberg.** 1984. The DnaA protein complex with the *E.*
384 *coli* chromosomal replication origin (*oriC*) and other DNA sites. Cell **38:**889-900.
385 21. **Grigoriev, A.** 1998. Analyzing genomes with cumulative skew diagrams. Nucleic Acids
386 Research **26:**2286-2290.
387 22. **Hall, M. J., and N. R. Hackett.** 1989. DNA sequence of a small plasmid from *Halobacterium*
388 strain GN101. Nucleic Acids Res **17:**10501.
389 23. **Hartwell, L. H., R. K. Mortimer, J. Culotti, and M. Culotti.** 1973. Genetic control of the cell
390 division cycle in yeast: V. genetic analysis of *cdc* mutants. Genetics **74:**267-86.
391 24. **Hoekstra, W. P., J. E. Bergmans, and E. M. Zuidweg.** 1980. Role of *rec*BC nuclease in
392 *Escherichia coli* transformation. Journal of Bacteriology **143:**1031-1032.
393 25. **Lam, W. L., and W. F. Doolittle.** 1989. Shuttle vectors for the archaebacterium *Halobacterium*
394 *volcanii*. Proceedings of the National Academy of Science **86:**5478-5482.
395 26. **Li, Z., T. J. Santangelo, L. Cubonova, J. N. Reeve, and Z. Kelman.** 2010. Affinity
396 purification of an archaeal DNA replication protein network. MBio **1**.
397 27. **Lipscomb, G. L., K. Stirrett, G. J. Schut, F. Yang, J. Francis E. Jenney, R. A. Scott, M. W.**
398 **W. Adams, and J. Westpheling.** 2011. Natural competence in the hyperthermophilic archaeon
399 *Pyrococcus furiosus* facilitates genetic manipulation: construction of multiple markerless
400 deletions of genes encoding the two cytoplasmic hydrogenases. Applied and Environmental
401 Microbiology **77:**2232-2238.
402 28. **Lopez, P., H. Philippe, H. Myllykallio, and P. Forterre.** 1999. Identification of putative
403 chromosomal origins of replication in Archaea. Molecular Microbiology **32:**883-886.
404 29. **Lucas, S., L. Toffin, Y. Zivanovic, D. Charlier, H. Moussard, P. Forterre, D. Prieur, and G.**
405 **Erauso.** 2002. Construction of a shuttle vector for and spheroplast transformation of the
406 hyperthermophilic archaeon *Pyrococcus abyssi*. Applied and Environmental Microbiology

AEM Accepts published online ahead of print

407 **68:**5528-5536.

408 30. **Lundgren, M., A. Andersson, L. Chen, P. Nilsson, and R. Bernander.** 2004. Three
409 replication origins in *Sulfolobus* species: synchronous initiation of chromosome replication and
410 asynchronous termination. Proc Natl Acad Sci U S A **101:**7046-51.

411 31. **Marsin, S., and P. Forterre.** 1998. A rolling circle replication initiator protein with a
412 nucleotidyl-transferase activity encoded by the plasmid pGT5 from the hyperthermophilic
413 archaeon *Pyrococcus abyssi*. Molecular Microbiology **27:**1183-1192.

414 32. **Matsunaga, F., P. Forterre, Y. Ishino, and H. Myllykallio.** 2001. *In vivo* interactions of
415 archaeal Cdc6/Orc1 and minichromosome maintenance proteins with the replication origin.
416 Proceedings of the National Academy of Science **98:**11152–11157.

417 33. **Matsunaga, F., A. Glatigny, M.-H. Mucchielli-Giorgi, N. Agier, H. Delacroix, L. Marisa, P.**
418 **Durosay, Y. Ishino, L. Aggerbeck, and P. Forterre.** 2007. Genomewide and biochemical
419 analyses of DNA-binding activity of Cdc6/Orc1 and Mcm proteins in *Pyrococcus* sp. Nucleic
420 Acids Research **35:**3214-3222.

421 34. **Matsunaga, F., K. Takemura, M. Akita, A. Adachi, T. Yamagami, and Y. Ishino.** 2010.
422 Localized melting of duplex DNA by Cdc6/Orc1 at the DNA replication origin in the
423 hyperthermophilic archaeon *Pyrococcus furiosus*. Extremophiles **14:**21-31.

424 35. **Messer, W.** 2002. The bacterial replication initiator DnaA. DnaA and *oriC*, the bacterial mode
425 to initiate DNA replication. FEMS Microbiol Rev **26:**355-74.

426 36. **Morikawa, M., Y. Izawa, N. Rashid, T. Hoaki, and T. Imanaka.** 1994. Purification and
427 characterization of a thermostable thiol protease from a newly isolated hyperthermophilic
428 *Pyrococcus* sp. Applied and Environmental Microbiology **60:**4559-4566.

429 37. **Myllykallio, H., P. Lopez, P. López-García, R. Heilig, W. Saurin, Y. Zivanovic, H. Philippe,**
430 **and P. Forterre.** 2000. Bacterial mode of replication with eukaryotic-like machinery in a

431 hyperthermophilic archaeon. Science **288:**2212-2215.

432 38. **Newlon, C. S., and J. F. Theis.** 2002. DNA replication joins the revolution: whole-genome
433 views of DNA replication in budding yeast. Bioessays **24:**300-4.

434 39. **Oka, A., K. Sugimoto, M. Takanami, and Y. Hirota.** 1980. Replication origin of the
435 Escherichia coli K-12 chromosome: the size and structure of the minimum DNA segment
436 carrying the information for autonomous replication. Molecular & General Genetics **178:**9-20.

437 40. **Robb, F. T., D. L. Maeder, J. R. Brown, J. DiRuggiero, M. D. Stump, R. K. Yeh, R. B.**
438 **Weiss, and D. M. Dunn.** 2001. Genomic sequence of hyperthermophile, *Pyrococcus furiosus*:
439 Implications for physiology and enzymology. Methods in Enzymology **330:**134-157.

440 41. **Robinson, N. P., and S. D. Bell.** 2005. Origins of DNA replication in the three domains of life.
441 The FEBS Journal **272:**3757-3766.

442 42. **Robinson, N. P., I. Dionne, M. Lundgren, V. L. Marsh, R. Bernander, and S. D. Bell.** 2004.
443 Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus*
444 *solfataricus*. Cell **116:**25-38.

445 43. **Sambrook, J., and D. W. Russell.** 2006. The condensed protocols from Molecular cloning : a
446 laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

447 44. **Santangelo, T. J., L. Cubonova, and J. N. Reeve.** 2008. Shuttle vector expression in
448 *Thermococcus kodakaraensis*: contributions of *cis* elements to protein synthesis in a
449 hyperthermophilic archaeon. Applied and Environmental Microbiology **74:**3099-3104.

450 45. **Schwob, E.** 2004. Flexibility and governance in eukaryotic DNA replication. Curr Opin
451 Microbiol **7:**680-90.

452 46. **Soler, N., A. Justome, S. Quevillon-Cheruel, F. Lorieux, E. L. Cam, E. Marguet, and P.**
453 **Forterre.** 2007. The rolling-circle plasmid pTN1 from the hyperthermophilic archaeon
454 *Thermococcus nautilus* Molecular Microbiology **66:**357-370.

AEM Accepts published online ahead of print

AEM Accepts published online ahead of print

455 47. **Stetter, K. O.** 1996. Hyperthermophilic procaryotes. FEMS Microbiology Reviews **18:**149-
456 158.
457 48. **Sugimoto, K., A. Oka, H. Sugisaki, M. Takanami, A. Nishimura, Y. Yasuda, and Y. Hirota.**
458 1978. Nucleotide sequence of *Escherichia coli* K-12 replication origin. Proceedings of the
459 National Academy of Science **76:**575-579.
460 49. **Tumbula, D. L., T. L. Bowen, and W. B. Whitman.** 1997. Characterization of pURB500 from
461 the archaeon *Methanococcus maripaludis* and construction of a shuttle vector. Journal of
462 Bacteriology **179:**2976-2986.
463 50. **Waege, I., G. Schmid, S. Thumann, M. Thomm, and W. Hausner.** 2010. A shuttle vector
464 based transformation system for *Pyrococcus furiosus*. Applied and Environmental Microbiology
465 **76:**3308-3313.
466 51. **Zillig, W., A. Kletzin, C. Schleper, I. Holz, D. Janekovic, J. Hain, M. Lanzendorfer, and J.**
467 **K. Kristjansson.** 1994. Screening for Sulfolobales, their Plasmids and their viruses in Icelandic
468 solfataras. Systematic and Applied Microbiology **16:**609-628.
469

470 FIGURE LEGENDS

471 Figure 1. The region of the *P. furiosus* chromosome predicted to contain the origin of replication, *oriC*.
472 (A) with an expanded view of the intergenic space (B). ORBs are indicated in red, mini-ORBS in pink,
473 and the putative unwinding site in green. The chromosomal regions that were cloned into pJFW017 to
474 produce various plasmids are indicated by black lines below the diagrams. Conserved palindromes are
475 marked by inverted black arrows. Inserts that resulted in plasmids capable of autonomous replication in
476 *P. furiosus* are indicated by thick black lines.
477

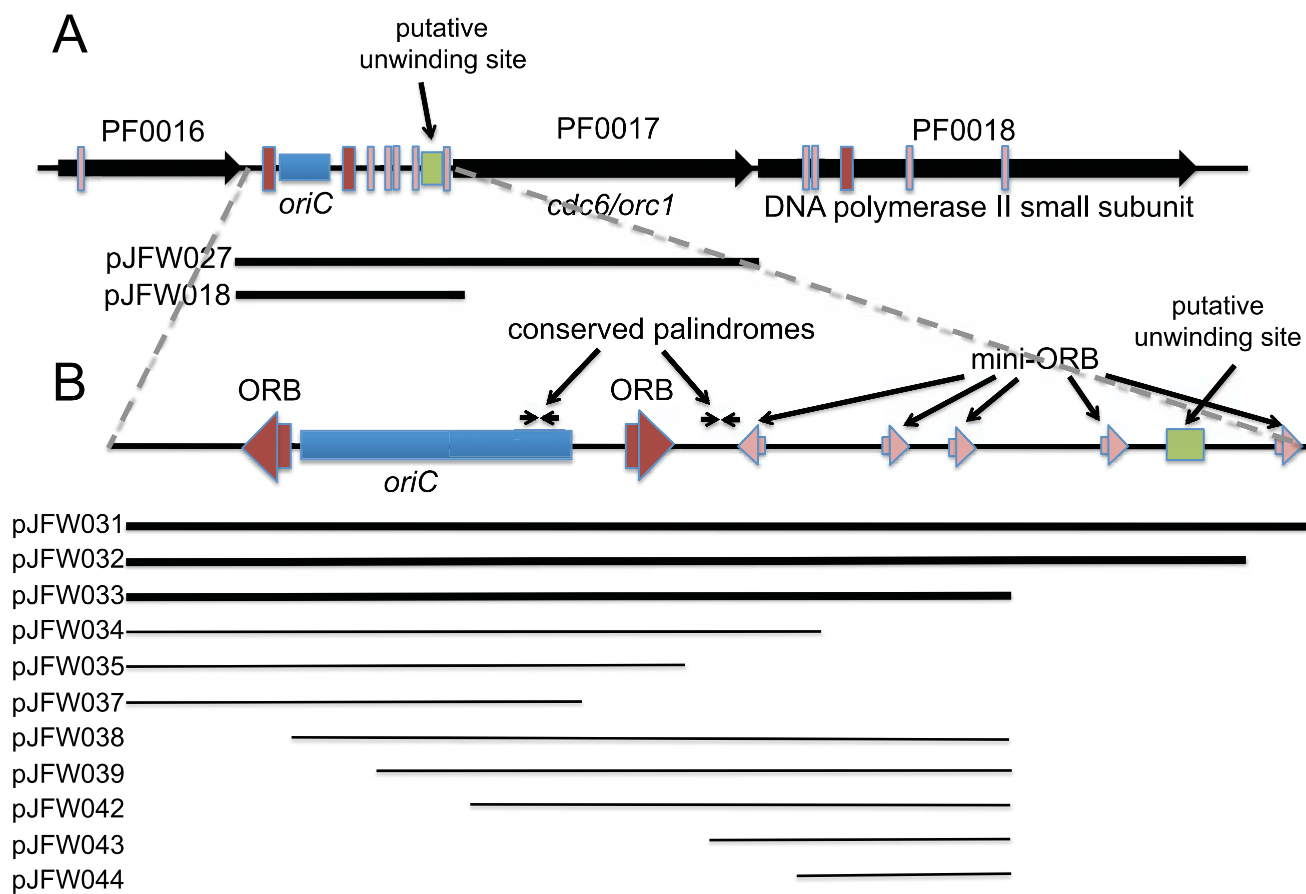
478 Figure 2. Construction of pJFW027. A linear DNA fragment containing the entire sequence of

AEM Accepts published online ahead of print

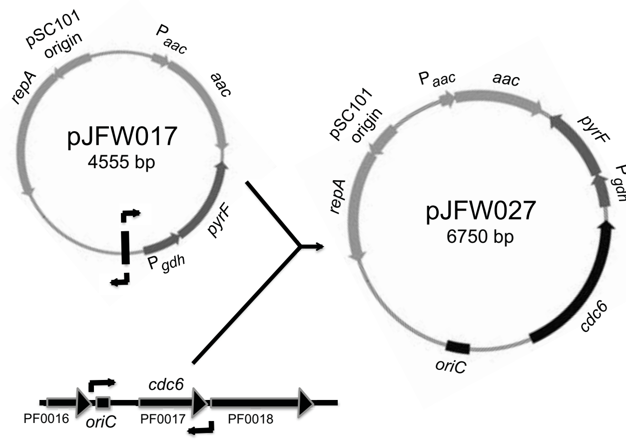
479 pJFW017 was generated by PCR amplification using primers JF266 and JF267 and ligated to the origin
480 fragment indicated in Figure 1 also generated by PCR amplification using primers JF268 and JF282.
481 Plasmids containing the various origin fragments described in Figure 1 were cloned into pJFW017 for
482 testing.
483 Figure 3. Determination of copy number for pJFW027 in *P. furiosus*. A. Diagram of the chromosomal
484 region including the *gdh* ORF. HpaI sites are indicated, as well as the location of primers used to
485 generate the *gdh* hybridization probe. B. Southern blot of pJFW027 transformants. Lanes 1-10, DNA
486 isolated from transformants and digested with HpaI. Lanes 11 and 12, DNA from *P. furiosus* wild type
487 and COM1 Δ*pyrF*, respectively. Lane 13, pJFW027 plasmid DNA purified from *E coli*.
488

489 Figure 4. Restriction analysis of plasmid DNA before and after transformation of *P. furiosus* and back-
490 transformation to *E. coli*. 1kb DNA ladder (Lanes M). pJFW018 plasmid DNA isolated from *E. coli*
491 DH5α (Lane 1), and digested with AccI (Lane 2), AvaI (Lane 3), and HindIII (Lane 4). Plasmid DNA
492 isolated from *E. coli* DH5α back-transformed from *P. furiosus* transformants (Lane 5), and digested
493 with AccI (Lane 6), AvaI (Lane 7), and HindIII (Lane 8).
494 Table 1. Footnote. Genomic locations are based on the numbering convention of the *Pyrococcus*
495 *furiosus* (NC_003413.1) genome sequence deposited in GenBank. The detection threshold was 4.7 x
496 10^2 transformants per μg of plasmid DNA.
497 Table 2. Footnote. Transformants containing each plasmid were serially passaged in liquid media with
498 or without uracil. Following each passage, diluted culture was plated onto selective media to determine
499 the number of prototrophic cells remaining.
500 Table 3. Footnote. Base differences are indicated by lowercase letters. Genomic locations are based on
501 the numbering of the genome sequences deposited in GenBank.

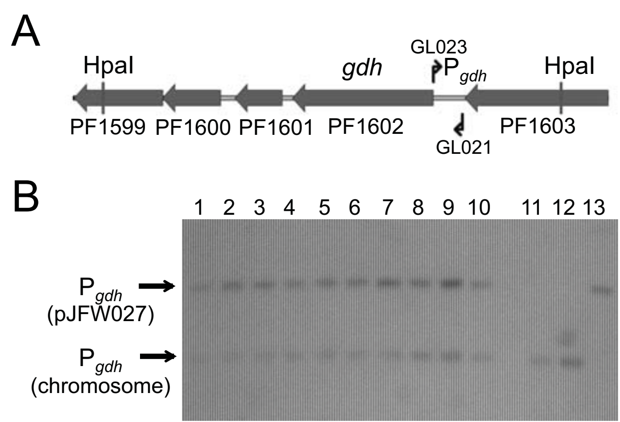
AEM Accepts published online ahead of print



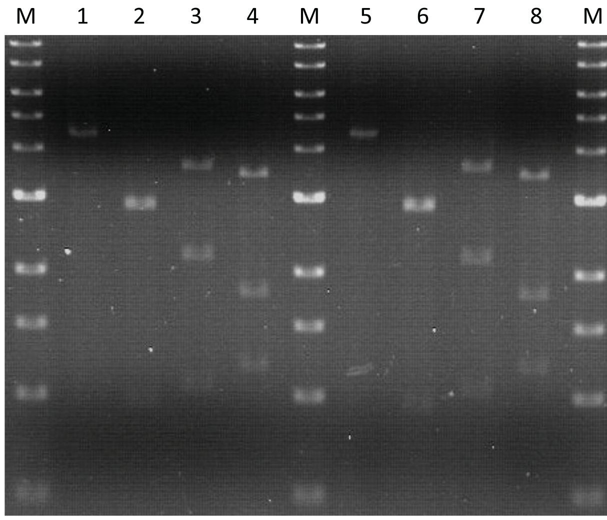
AEM Accepts published online ahead of print



AEM Accepts published online ahead of print



AEM Accepts published online ahead of print



AEM Accepts published online ahead of print

Table 1. Plasmid transformation efficiencies

Plasmid name	*oriC* insert	5' primer	3' primer	Transformation Efficiency
pJFW017	none			$8.2x10^{2}$
pJFW018	15382 to 16226	JF268	JF269	$9.1x10^{5}$
pJFW027	15382 to 17576	JF268	JF282	$5.8x10^{5}$
pJFW031	15382 to 16228	JF268.2	JF269.2	$6.6x10^{5}$
pJFW032	15382 to 16187	JF306.2	JF269.2	$8.0x10^{4}$
pJFW033	15382 to 16034	JF305.2	JF269.2	$6.5x10^{5}$
pJFW034	15382 to 15890	JF304.2	JF269.2	$5.7x10^{2}$
pJFW035	15382 to 15786	JF303.2	JF269.2	$1.4x10^{3}$
pJFW037	15382 to 15705	JF301.1	JF269.2	$1.4x10^{2}$
pJFW038	15492 to 16034	JF305.2	JF339	$4.7x10^{1}$
pJFW039	15561 to 16034	JF305.2	JF345	$< 4.7x10^{1}$
pJFW042	15746 to 16034	JF305.2	JF348	$< 4.7x10^{1}$
pJFW043	15813 to 16034	JF305.2	JF349	$< 4.7x10^{1}$

AEM Accepts published online ahead of print

Table 2. Maintenance of plasmids in *P. furiosus*.

	pJFW018		pJFW027		pJFW019	
Passage	+ura	-ura	+ura	-ura	+ura	-ura
1	187	256	190	194	73	112
2	132	217	97	203	54	268
3	112	232	132	143	11	138
4	146	117	154	165	6	83
5	87	276	100	113	0	77
6	138	197	151	201	0	104
7	144	263	111	122	0	91
8	118	242	87	131	0	94
9	93	213	114	218	0	97
10	107	183	112	169	0	87

AEM Accepts published online ahead of print

Table 3. Conserved Palindromic Sequences within the *Pyrococcus oriC* region.

Species	Sequence	Genomic location
Palindrome #1		
P. furiosus (NC_003413.1)	ATATTTAAATAT	15,641-15,674
P. abyssi (NC_000868.1)	TATTTAAATA	123,223-123,232
P. horikoshii (NC_000961.1)	TATTTAAATA	111,307-111,316
Palindrome #2		
P. furiosus (NC_003413.1)	ATTAgaTTAAtcTAAT	15,809-15,824
P. abyssi (NC_000868.1)	ATTAagTTAAccTAAT	123,072-123,087
P. horikoshii (NC_000961.1)	ATTAagTTAActTAAT	111,155-111,170