Supplementary Material for

Role of free DNA ends and protospacer adjacent motifs for CRISPR DNA uptake in *Pyrococcus furious*

Masami Shiimori, Sandra C. Garrett, Dwain P. Chambers, Claiborne V. C. Glover Ⅲ, Brenton R. Graveley, Michael P. Terns

Supplementary Figure S1.

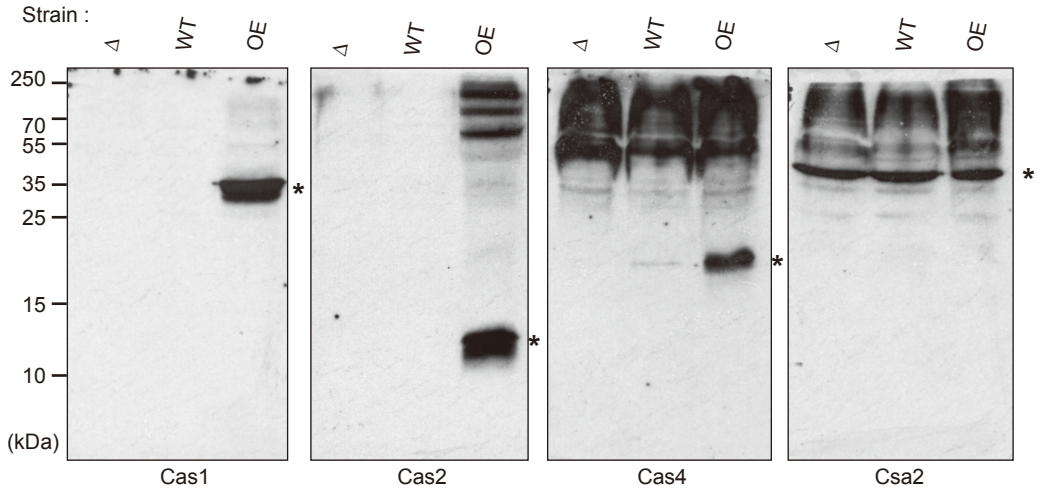

Figure S1. Cas1, Cas2 and Cas4 protein expression.

Western blot of *P. furiosus* Δ, WT and OE strains with Cas1, Cas2, Cas4 and Csa2 antibodies. Csa2 is used as a loading control. The indicated antibodies detect a major band (indicated with asterisk) of the expected molecular weight for each antigen. The predicted molecular weight of each protein is: Cas1 (37.5 kDa), Cas2 (10.0 kDa), Cas4 (20.2 kDa), and Csa2 (37.4 kDa).

Supplementary Figure S2

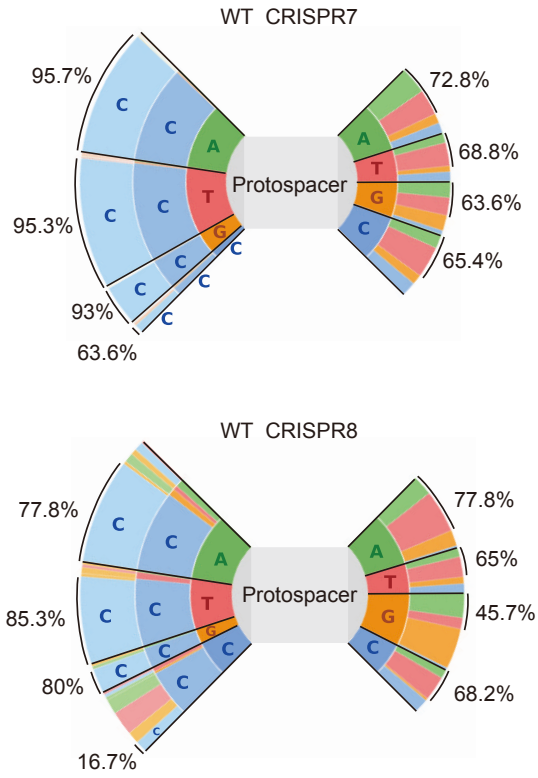

Figure S2. Analysis of PAM for the CRISPR8 locus. Newly-acquired spacers in the CRISPR7 and CRISPR8 array were used to search the genome and plasmids in order to identify the corresponding protospacers, and upstream and downstream sequences were extracted and used to generate consensus motifs on both strands of DNA. The plot shows the frequencies of three base pairs of flanking sequences upstream of the protospacer and two base pairs downstream of the protospacer for CRISPR7 and CRISPR8 in the WT strain. Percentages on left side indicate proportion of spacers with CC at the second and third position and percentages on right side indicate proportion of spacers with W (A/T) at the second position.

Supplementary Figure S3

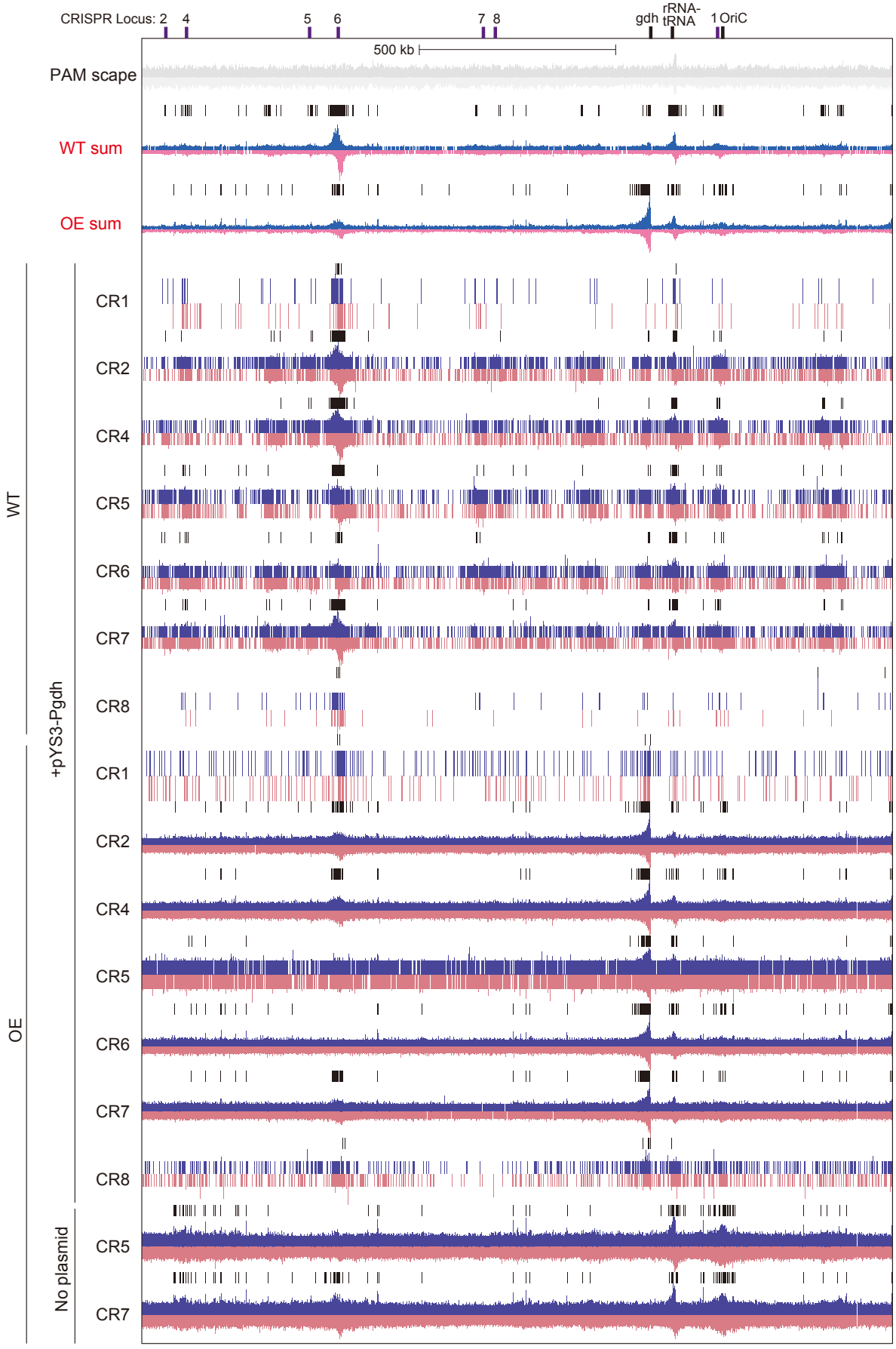

Figure S3. Distribution patterns of aligned protospacers are similar for the seven CRISPR loci. New spacers were analyzed for all seven CRISPR loci, in both WT and OE strains. Protospacers were deduced by aligning these new spacers to the genome. Protospacers on the plus and minus strand are indicated in blue and pink respectively. The 5'-CCN-3'/5'-NGG-3' PAMscape across the chromosome is shown in gray (top). Black bars show regions where protospacers are significantly enriched over the background, as detected by the peak-calling software from the HOMER package. At the top, the "sum" tracks for WT and OE show the combined pool of spacers from all seven CRISPR loci. All identical spacers (with their identical alignment coordinates) are collapsed to a single representative.

Supplementary Figure S4

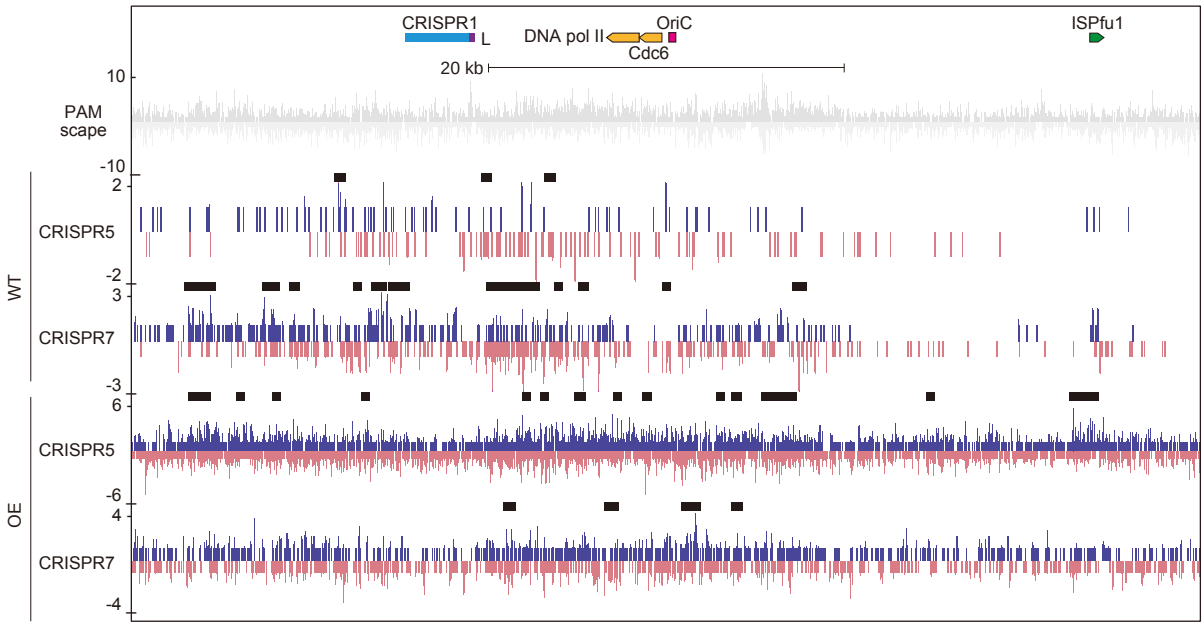

Figure S4. Protospacers are enriched around the DNA replication origin.

Protospacers are significantly enriched around the OriC DNA replication origin. Protospacers on the plus and minus strand are indicated in blue and pink, respectively. Protospacers were deduced from aligning new spacers acquired into the CRISPR5 and 7 loci of WT and OE strains. New spacers acquired into other CRISPR arrays showed similar trends. Black bars show areas were protospacers are significantly enriched, as detected by peaking finding software in the HOMER package.

Supplementary Figure S5

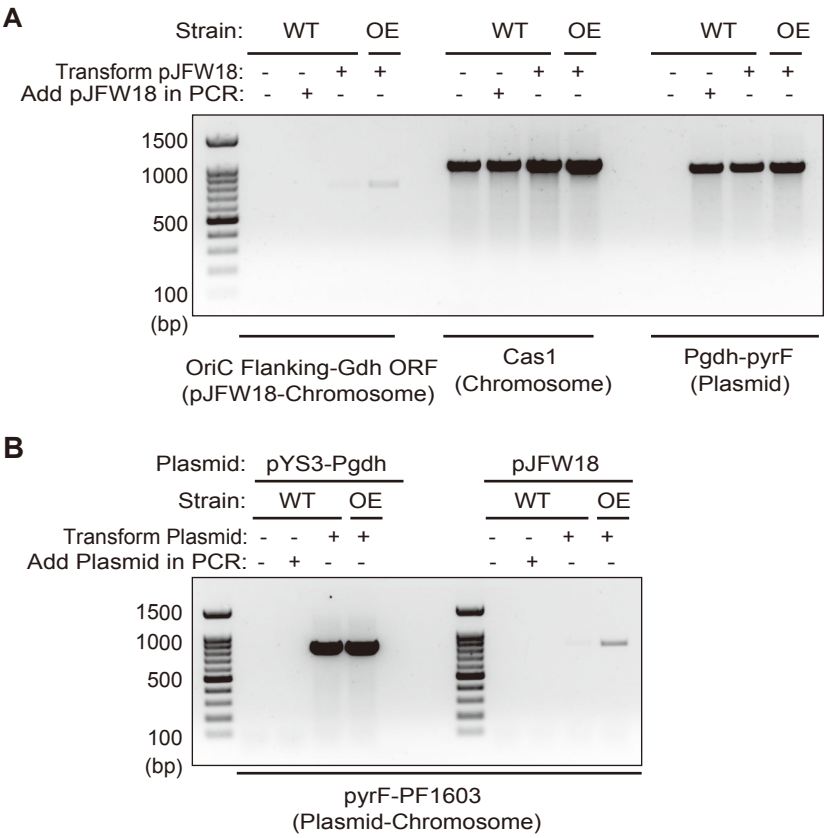

Figure S5. Recombination between the plasmid and the chromosome.

(A and B) PCR results confirm that homologous recombination between the plasmid and chromosome occurs. (A) Homologous recombination between pJFW18 plasmid and chromosome. This homologous recombination occurs preferentially when Cas proteins are overexpressed. (B) Homologous recombination between plasmids and chromosome. This homologous recombination occurs preferentially with rolling circle type replication plasmids.

Supplementary Figure S6

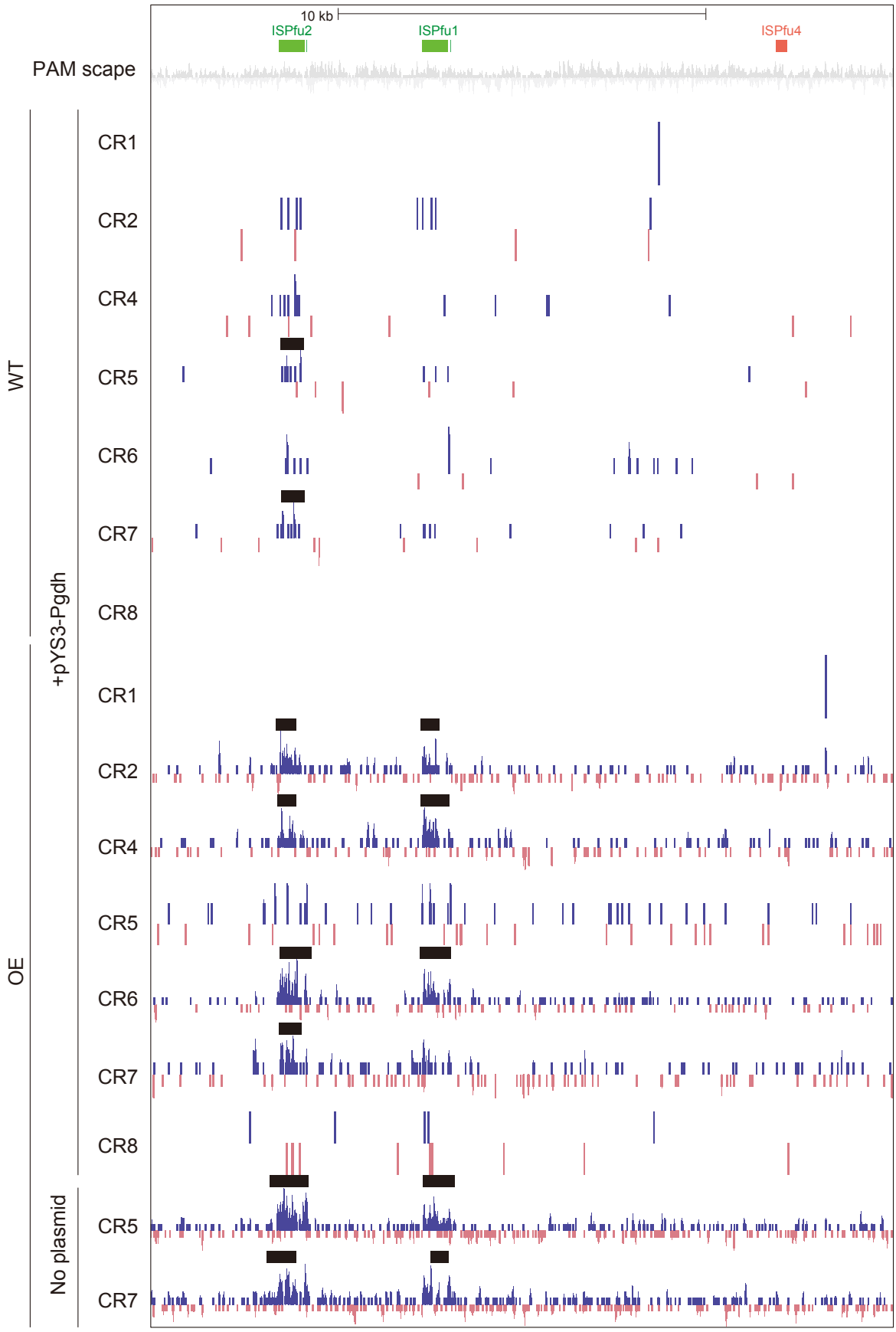

Figure S6. Protospacers are enriched for transposons. Protospacers are significantly enriched around active transposons but not around predicted inactive transposons. Black bars show areas were protospacers are significantly enriched, as detected by peak finding software in the HOMER package. Transposon peaks were observed for CR2, CR4, CR5, CR6, and CR7 loci, but were less apparent in the low efficiency loci CR1 and CR8. These peaks were not dependent on the presence of exogenous (plasmid) DNA.

Supplementary Figure S7

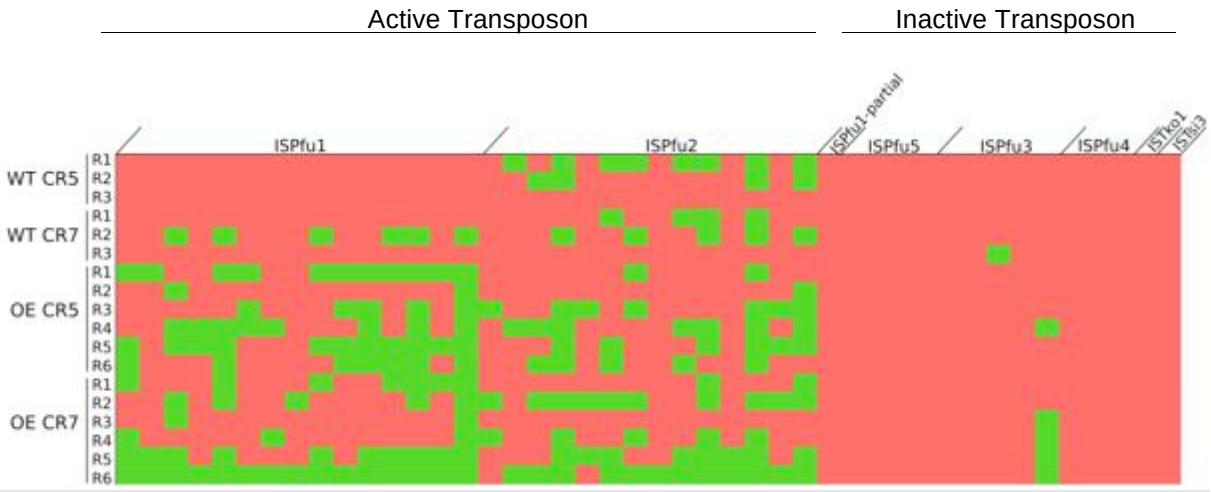

Figure S7. Protospacers hotspots at transposons.

Protospacers are significantly enriched at active transposons. Green spots show transposons for which protospacers are significantly enriched, as detected by peak finding software in the HOMER package. Red spots show transposons for which protospacers are not enriched. Protospacers were deduced by aligning new spacers acquired into the CRISPR5 and 7 loci of WT and OE strains.

Supplementary Figure S8

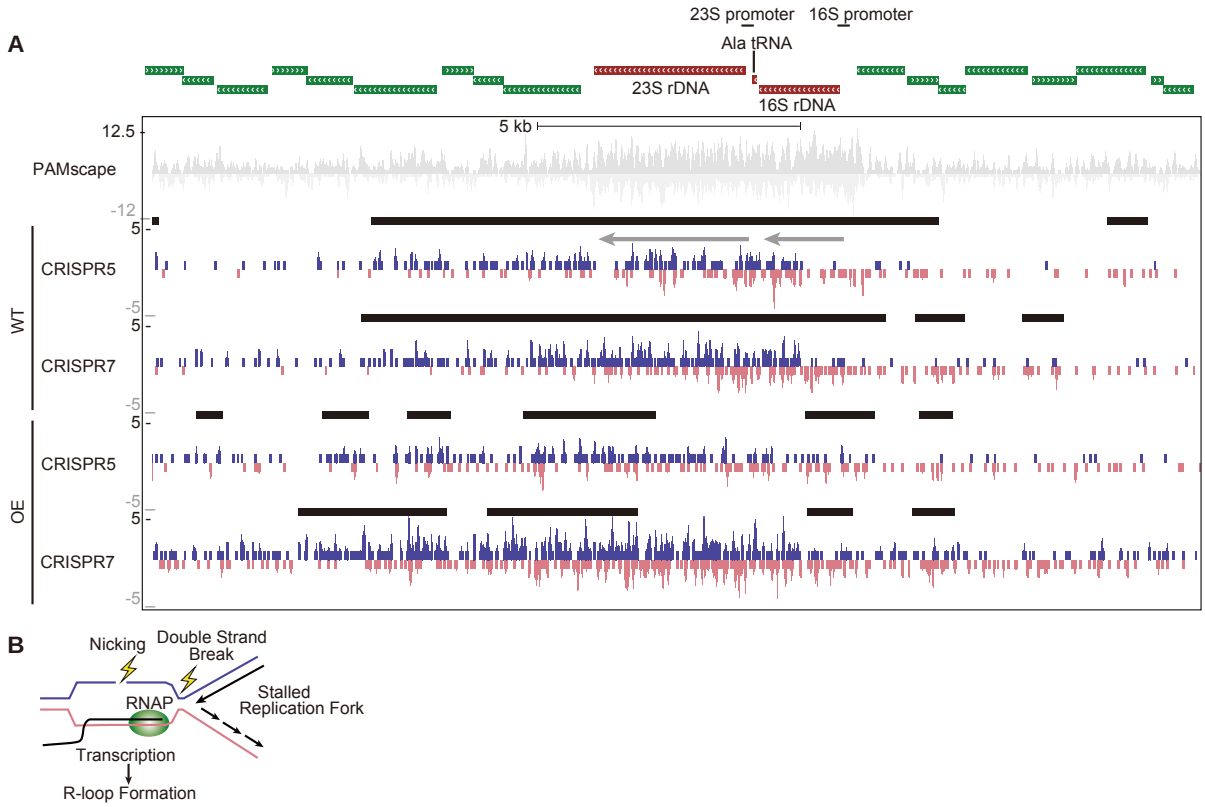

Figure S8. **Protospacers are enriched around rDNA gene cluster.**

(A) Protospacers are significantly enriched around the rRNA gene cluster. Protospacers on the plus and minus strand are indicated in blue and pink respectively. Protospacers were deduced from aligning new spacers acquired into the CRISPR5 and 7 loci of WT and OE strains. New spacers acquired into other CRISPR arrays showed similar trends. Black bars show areas were protospacers are significantly enriched, as detected by peaking finding software in the HOMER package. (B) Cartoon model for formation of DNA nicks in the rRNA gene cluster.

Supplemental Table S1. Oligonucleotide used in this study

PCR oligonucleotide

Name	Sequence (5'-3')
CRISPR1 fow	CCGCTCGAGGGATGCCCCGTAAAGGTTAT
CRISPR1 rev	CCATCGATGATCAGATTGCTTAAGACAAGAAATGAAT
CRISPR2 fow	CCGCTCGAGCTTGGGTAAGGTTGATGTCC
CRISPR2 rev	CCATCGATGTCAGAAGAAGAAGTGCGATGAAC
CRISPR4 fow	CCGCTCGAGAAGGTGGATGTCCCGTAAGG
CRISPR4 rev	CCATCGATCCCTTTGCATTGAGCAACATTC
CRISPR5 fow	CCGCTCGAGTAAGTTGGATGCCCCGTAAG
CRISPR5 rev	CCATCGATGCATGTCCAGAAACGAAGTCCTG
CRISPR6 fow	CCGCTCGAGGGTAAGTTGGATGCCCCGTAAG
CRISPR6 rev	CCATCGATGTTCCACTAAGGACATTTGTACGTCA
CRISPR7 fow	CCGCTCGAGTTCTGGGGTAAGGTTGATGC
CRISPR7 rev	CCATCGATCCGCAAGGACCGCATACTAC
CRISPR8 fow	CCGCTCGAGAGGGTTTGGTAAGGGGTATTTGTAA
CRISPR8 rev	CCATCGATGAGCAGCTATTTGCTTCGGC

PCR oligonucleotide for Illumina Miseq

Name	Sequence (5'-3')
CRISPR1 fow HTS	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGAGTGATATAGTTACTCCGTAGGAGTATTGG
CRISPR1 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGATCAGATTGCTTAAGACAAGAAATGAAT
CRISPR2 fow HTS	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCGAGTGATATAGTTACTCTGTAGGAGTATTGG
CRISPR2 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGTCAGAAGAAGAAGTGCGATGAAC
CRISPR4 fow HTS	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTAATTACCCCAATGGGGAATTGG
CRISPR4 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCCTTTGCATTGAGCAACATTC
CRISPR5_6 fow HTS	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAAGTGATATAAT

	TACTCCATAGGAGTATTGG
CRISPR5 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCCAGAAACGAAGTCCTGGC
CRISPR6 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGACATTTGTACGTCAAATTCTTCAC
CRISPR7 fow HTS	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGATGTAGTTACTCCGTAGGAGGATTGG
CRISPR7 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCGCAAGGACCGCATACTAC
CRISPR8 fow HTS	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCAAGTGATATAATTACTCCATAGGAGTATTGG
CRISPR8 rev HTS	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGAGCAGCTATTTGCTTCGGC

Oligonucleotide for generating overexpression and deletion strain

Name	**Sequence (5’-3’)**
Cas O/E_Cas2_fow	ACCCAAGGAGGTGTTGTCATATGTACATTGTAGTTGTCTACG
Cas O/E_Cas2_rev	AAAAAAAATCATTAAATTATTTCCTCTATTGGATTCTTTTC
Cas O/E_Cas4_fow	ATGGAAGAGTACGCCTTAGAG
Cas O/E_Cas4_rev	AAAAAAAATCACTACCCCCAACAAAGT
Cas O/E_Cas1_fow	ATGAGAAAAAAGTCTTTAACAATTTTTTCTG
Cas O/E_Cas1_rev	GGAGAAGAGAGGGGGGATCCTCAAAACCACGCGACAAGGG
Cas O/E_Pslp_fow	ATAGAGGAAATAATTTAATGATTTTTTTTCTGGCAGAATAGAAATAGATATTATC
Cas O/E_ Pslp _rev	CTCTAAGGCGTACTCTTCCATTTTTCTCCACCTCCCAATAA
Cas O/E_ Ppsp _fow	ACTTTGTTGGGGGTAGTGATTTTTTTTGTCAAATGCTCATCATTTAGTT
Cas O/E_ Ppsp _rev	CAGAAAAAATTGTTAAAGACTTTTTTCTCATATCGATTCACCTCCGCCTTA
P-gdh-PyrF_fow	GATTGAAAATGGAGTGAGCTGAG
P-gdh-PyrF_rev	TTATCTTGAGCTCCATTCTTTCACC
PF1120_CR7_1	CCAGTGCCAAGCTTGCATGCACCGGTACGCGTTTAAACACTCAAAGAGAGCGGAGCAA
PF1120_CR7_2	CTCAGCTCACTCCATTTTCAATCTAGAGAATATCCCCCTCACGG

PF1120_CR7_5	GGTGAAAGAATGGAGCTCAAGATAATGAAGCAATTCCAGGAGCTT
PF1120_CR7_6	GCATTACTAGTTTCGTTGACGTCTAGAGAATATCCCCCTCACGG
PF1120_CR7_7	GACGTCAACGAAACTAGTAATGCGGCCGCAGATATCGTGAAACAC CCAAAAATTTTTCGA
PF1120_CR7_8	TATGACATGATTACGAATTCTACGTATCCGGATTAATTAATTTAAAT TTATTGGAACTTGCACTCACG

Oligonucleotide for generating plasmids

Name	**Sequence (5’-3’)**
Pslp_fow	CGGGATCCACTCTTTCGGAAATAGCAAGTATAGTC
Pslp_rev	CGTCCAACGCTAGTACAATCATTTTTCTCCACCTCCCAATAATCTG
PyrF_fow	ATGATTGTACTAGCGTTGGACG
PyrF_rev	AAGATATCGTCGATTGGCTGAGCTCATG

Supplementary Table S2. Plasmids used in this study

Name	Relevant Characteristic	Source or reference
pYS3	AmpR general cloning vector with E. coli pUC19 ori and SimR for selection and Rep75 gene for replication in P. furious	(1)
pYS3-Pgdh	pYS3-derivative, Pgdh-pyrf repraced with SimR	(2)
pJFW17	AprR general cloning vector with *E.coli* OriT, and *Pfu* Pgdh-pyrF cassette	(3)
pJFW18	pJFW17-derivative, *P. furious* OriC for replication in P. furiosus	(3)
pYS3-Pslp	pYS3-derivative, Pslp-pyrf repraced with SimR	This study
pHSG298	The kanamycin resistant pUC vectors	Takara
pJE47	pJFW18 derivative, Tk-csg promoter-chiA terminator expression cassette	(4)
pMS32	pHSG298-derivative, containing Cas1, Cas2, Cas4 overexpression cassette	This study

Supplementary Table S3. Strains used in this study

Name	Relevant Characteristic	Source or reference
JFW02 (5)	Δ*pyrF* Δ*trpAB*	(6)
TPF43	JFW02 with Cas1, Cas2, Cas4 overexpression	This study
TPF55	JFW02 ΔCas1ΔCas2ΔCas4	This study

Materials and Methods

Western blot

P. furiosus strains (WT, OE, ∆) were grown overnight. Cells in 1 ml liquid culture were pelleted and lysed with 20 µl SDS loading buffer. Cells were incubated at for 98 ºC 10 min and then on ice for 2 min. Whole lysate was separated on a 15% SDS-PAGE gel and blotted onto a nitrocellulose membrane (Bio-Rad). The blots were incubated with 1:12,500 - 1:25,000 dilution of polyclonal IgY antibodies and 1:25,000 dilution of HRP-conjugated anti-IgY secondary antibody (Gallus Immunotech). The protein bands on the blot were detected using an enhanced chemiluminescent substrate for HRP (horse radish peroxidese) activity (ECL Prime, GE helthcare).

Reference

1. Waege, I., Schmid, G., Thumann, S., Thomm, M. and Hausner, W. (2010) Shuttle vector-based transformation system for Pyrococcus furiosus. *Appl Environ Microbiol*, **76**, 3308-3313.
2. Swarts, D.C., Hegge, J.W., Hinojo, I., Shiimori, M., Ellis, M.A., Dumrongkulraksa, J., Terns, R.M., Terns, M.P. and van der Oost, J. (2015) Argonaute of the archaeon Pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA. *Nucleic acids research*, **43**, 5120-5129.
3. Farkas, J., Chung, D., DeBarry, M., Adams, M.W. and Westpheling, J. (2011) Defining components of the chromosomal origin of replication of the hyperthermophilic archaeon Pyrococcus furiosus needed for construction of a stable replicating shuttle vector. *Appl Environ Microbiol*, **77**, 6343-6349.
4. Elmore, J., Deighan, T., Westpheling, J., Terns, R.M. and Terns, M.P. (2015) DNA targeting by the type I-G and type I-A CRISPR-Cas systems of Pyrococcus furiosus. *Nucleic acids research*, **43**, 10353-10363.
5. Antal, C.E., Hudson, A.M., Kang, E., Zanca, C., Wirth, C., Stephenson, N.L., Trotter, E.W., Gallegos, L.L., Miller, C.J., Furnari, F.B. *et al.* (2015) Cancer-associated protein kinase C mutations reveal kinase's role as tumor suppressor. *Cell*, **160**, 489-502.
6. Farkas, J., Stirrett, K., Lipscomb, G.L., Nixon, W., Scott, R.A., Adams, M.W. and Westpheling, J. (2012) Recombinogenic properties of Pyrococcus furiosus strain COM1 enable rapid selection of targeted mutants. *Appl Environ Microbiol*, **78**, 4669-4676.