Appendix

Plasmid construction.

Generation of vectors using yeast gap repair. Please see Table 2 for the list of plasmids used in this study. All vectors created in this study, except pMU102, pMU245, pMU482, pMU612, and pMU1162, were made using yeast gap repair cloning, as described by Shanks *et al* (1). In brief, a vector capable of replication in yeast was linearized by restriction digestion. Through primer design, PCR amplicons were engineered to contain ~40-bp at the 5' and 3' ends that are homologous to DNA on either side of the restriction site used to linearize the vector. When multiple, adjacent amplicons were cloned into a single vector, primer design was such that the 3' end of the upstream fragment and the 5' end of the downstream fragment contained ~40-bp of homology. The digested vector and PCR amplicons were transformed into *S. cerevisiae* InvSc1 (*ura3-52/ura3-52* uracil auxotroph) using a modified "lazy bones" transformation protocol (1). All yeast vectors used in the study have a copy of the *URA3* gene capable of complementing the *ura3-52* mutation. The vectors also contain the *CEN6* and *ARSH4* elements but do not contain telomeres which would be required to maintain the plasmid in a linear form. Thus, when the transformed yeast were plated on selective, uracil "drop-out" medium (SD-ura), only those that have generated a circular, plasmid-based replicon via homologous recombination of the linearized plasmid and PCR amplison(s) can grow. Yeast transformants growing on SD-ura medium were pooled and plasmid DNA was recovered using the "smash and grab" method (1) and transformed into *E. coli* NEB 5-alpha electrocompetent cells (New England Biolabs, Ipswich) following manufacturer protocols.

Creation of pMU102 and pMU245 – E. coli-C. thermocellum shuttle vectors. The vector pMU102 is derived from pMU104. The vector pMU104 is an *E. coli-C. thermocellum* shuttle vector that contains ~350 bp repeat sequence. To minimize the size of the vector and to prevent complications of unwanted recombination in subsequent versions that would be used as *S. cerevisiae-E. coli-C. thermocellum* shuttle vectors the repeat region was deleted by digesting pMU104 with FokI and EcoRI, followed by Klenow treatment and re-ligation to produce pMU102. To create pMU245: Primers X00967 and X00968, with engineered BglII and EcoRI sites, were used with pMU102 as template to generate an amplicon that consisted of the pMU102 backbone devoid of the *cat* gene. Primers X00969 and X00970, also with engineered BglII and EcoRI sites, were used to amplify the *amp* gene from pUC19. The two amplicons were digested with the enzymes mentioned and ligated together to create a version of pMU102 in which the *cat* gene was replaced with the *amp* gene. This vector still provides selection in *E.coli* but avails the *cat* open reading frame (orf) for use in subsequent cloning strategies into this vector.

Creation of pMU113 –generation of a cassette (gapDHp-cat) used for positive selection in C. thermocellum. The *aacC1* gene of pMU110 was replaced with the *cat* gene from pM102. The vector pMU110 was linearized with BglII and the *cat* gene from pMU102 was amplified with primers X00105 and X00106. The digested vector and amplicon were used in a gap repair cloning to generate pMU111. The native *gapDH*p from *C. thermocellum* was amplified from genomic DNA using primers X00109 and X00110 and

fused to the *cat* gene in pMU111 that was digested with BsrGI via gap repair cloning, resulting in pMU113.

Creation of pMU357 and pMU749 – S. cerevisiae-E. coli-C. thermocellum shuttle vectors

The vector pMU357 contains the native *C. thermocellum gapDH*p and thus can be used as an expression vector. To create pMU357, the *URA3, CEN6* and *ARSH4* elements and the native *C. thermocellum gapDH*p were amplified from pMU113 using primers X00973 and X00974 and inserted into pMU102 that was digested with BspHI via gap repair cloning, resulting in pMU357. The vector pMU749 is a *S. cerevisiae-E. coli-C. thermocellum* shuttle vectors that contains the *amp* gene as the only selectable marker for bacteria, thereby allowing the *cat* orf to be included in subsequent cloning strategies into this vector. To create pMU749, the *URA3, CEN6* and *ARSH4* elements were amplified from pMU110 using primers X03852 and X03853 and inserted into pMU245 that was digested with SacI via gap repair cloning, resulting in pMU749.

Creation of pMU482 – a replicating vector used to delete the C. thermocellum pyrF gene.

To generate the Δ*pyrF* fragment, ~1 kb of DNA that flank the 5' and 3' region of *pyrF* were cloned into pMU357 using gap repair cloning. Primers X01840 and X01841 were used to amplify the 5'-flanking DNA, and primers X01842 and X01843 were used to amplify the 3'-flanking DNA. The 5' and 3' flanking fragments were cloned into pMU357 that was linearized with BamHI using gap repair cloning resulting in pMU440. The Δ*pyrF* fragment from pMU440 was amplified with primers X01893 and X01905 which

have engineered XbaI and KpnI sites, respectively and cloned into pMU102 digested with those same enzymes, resulting in pMU482.

Creation of pMU612 – a vector for complementing the C. thermocellum ΔpyrF strain. The *C. thermocellum cbp* promoter and orf were cloned upstream of the *E. coli* rrnB terminator using gap repair cloning. Primers X03061 and X03062 were used to amplify the *C. thermocellum cbp* promoter and orf from genomic DNA and X03063 and X03064 were used to amplify the rrnB terminator from a synthesized DNA fragment (Codon Devices). The two amplicons were fused via gap repair cloning with a gel purified, 3.9-kb fragment of pMU110 that resulted from BamHI and XhoI digestion, resulting in pMU582. The *cbp* orf of pMU582 was then replaced with the *pyrF* orf of *C. thermocellum* as follows: The vector pMU582 was linearized with SmaI and the *pyrF* orf was amplified with primers X03079 and X03080. The SmaI digested vector and *pyrF* amplicon were used in a gap repair cloning to generate the *pyrF* complementing cassette (*cbp*p-*pyrF*-rrnBT1-T2), resulting in pMU597. The *pyrF* complementing cassette was released from pMU597 by digestion with XbaI and recovered by gel purification and cloned into pMU102 digested with XbaI. The orientation of insertion was confirmed by DNA sequencing and restriction digestion. The resulting plasmid is pMU612.

Creation of pMU769 – a replicating vector used to delete the C. thermocellum pyrF gene and replace with a gapDHp-cat cassette. To generate the DNA cassette used to create a the *C. thermocellum* Δ*pyrF::gapDH*p-*cat* strain, ~1 kb of DNA that flank the 5' and 3' region of *pyrF* were cloned into pMU749 such that they flank the *gapDH*p-*cat* cassette

via gap repair cloning. Primers X03886 and X03898 were used to amplify the 3'-flanking DNA, and primers X03897 and X03889 were used to amplify the 5'-flanking DNA using pMU440 as template. Primers X03896 and X03899 were used to amplifiy the *gapDH*p-*cat* cassette from pMU113. The 5' and 3' flanking fragments, and the *gapDH*p-*cat* cassette were cloned into pMU749 that was linearized with BamHI using gap repair cloning resulting in pMU769.

Creation of pMU1162 - a replicating vector used to delete the C. thermocellum pta gene and replace it with a gapDHp-cat cassette. To generate the DNA cassette used to create a the *C. thermocellum* Δ*pta::gapDH*p-*cat* strain, ~1 kb of DNA that flank the 5' and 3' region of *pta* were cloned into pMU749 such that they flank the *gapDH*p-*cat* cassette via gap repair cloning. Primers X05109 and X05110 were used to amplify the 5'-flanking DNA, and primers X05113 and X05114 were used to amplify the 3'-flanking DNA using genomic DNA as template. Primers X05111 and X05112 were used to amplifiy the *gapDH*p-*cat* cassette from pMU113. The 5' and 3' flanking fragments, and the *gapDH*p-*cat* cassette were cloned into pMU749 that was linearized with BamHI using gap repair cloning resulting in pMU1016. To add the *pyrF* complementing cassette to pMU1016, the *cbp*p-*pyrF*-rrnBT1-T2 was released from pMU612 by digestion with XbaI and recovered by gel purification and cloned into pMU1016 digested with XbaI. The orientation of insertion was confirmed by DNA sequencing and restriction digestion. The resulting plasmid is pMU1162.

References:

1. Shanks RMQ, Caiazza NC, Hinsa SM, Toutain CM, & O'Toole GA (2006) *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from Gram-negative bacteria. *Appl Environ Microbiol* 72(7):5027-5036.

Supporting information figure legends:

Fig. A1: DNA constructs that directly resulted in the creation of strains used in this study.

Replication origins – ColE1 (*E. coli*), *CEN6/ARSH4*(*S. cerevisiae*), dso (*C. thermocellum*).

Orf's – *repB*- replication initation protein that function in *C. thermocellum*, *cat* – chloramphenicol acetyltransferase, *amp* – ampicillin resistance gene, *pyrF* – *C. thermocellum* orotidine 5-phosphate decarboxylase, *URA3* - *S. cerevisiae* orotidine 5-phosphate decarboxylase,

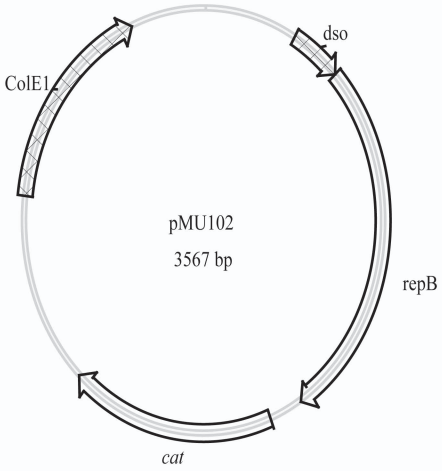
Promoters – *gapDH*p- from *C. thermocellum* glyceraldehyde 3-phosphate dehydrogenase, *cbp*p - from *C. thermocellum* cellobiose phosphorylase

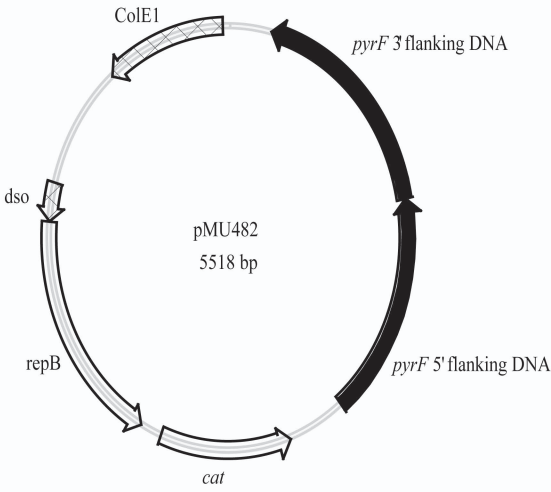
Terminators – T1, T2

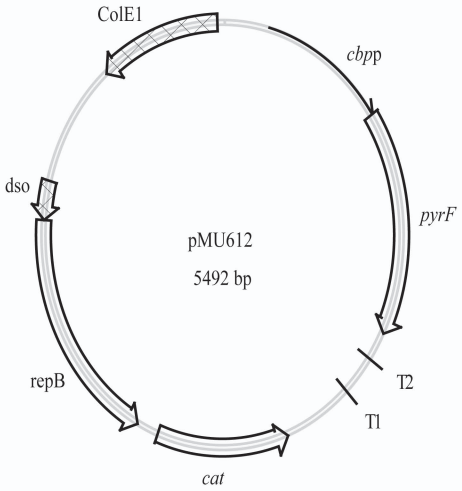
Flanking DNA used to target integration – *pyrF* 5' and 3', *pta* 5' and *pta* 3'

Fig. A2: Square pulse Diagram. A series of 45 pulses, each with a duration of 30us, a period of 300us and an amplitude of 1.9kV, were applied to a standard 1mm electroporation cuvette with a custom-built pulse generator (see material and methods for details).

Fig. A1:


ColE1
dso
pMU102
3567 bp
repB
cat


ColE1
pyrF 3' flanking DNA
dso
pMU482
5518 bp
repB
pyrF 5' flanking DNA
cat


ColE1
cbpp
dso
pyrF
pMU612
5492 bp
repB
T2
T1
cat

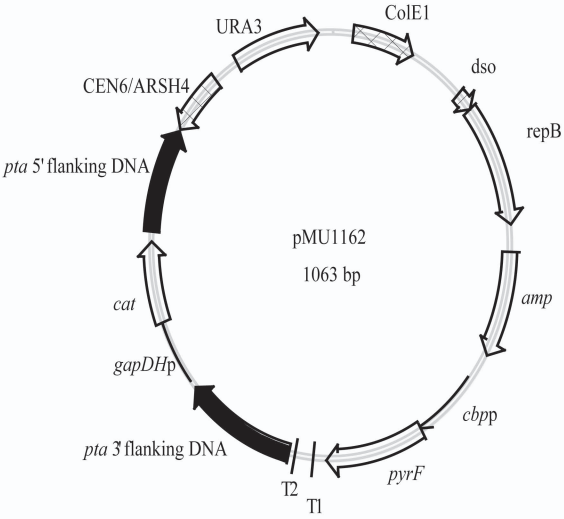

URA3
ColE1
dso
CEN6/ARSH4
repB
pta 5' flanking DNA
pMU1162
1063 bp
amp
cat
gapDHp
cbpp
pta 3' flanking DNA
pyrF
T2
T1

Fig. A2