Physiology, genomics, and pathway engineering of an ethanol-tolerant strain of *Clostridium phytofermentans*

Supplementary information

Page 2: Fig S1 Glucose consumption by *C. phytofermentans* WT and ET strains.

Page 3: Fig S2 *C. phytofermentans* WT and ET strain morphologies.

Page 4: Table S1 Generation times (hours) and maximum cell densities (OD_{600}) of *C. phytofermentans* wild-type (WT) and ethanol-adapted (ET) strains grown in GS2 medium supplemented with different ethanol concentrations.

Page 5-6: Supplementary Text Analysis of the genome sequence of the *C.phytofermentans* ethanol tolerant (ET) strain

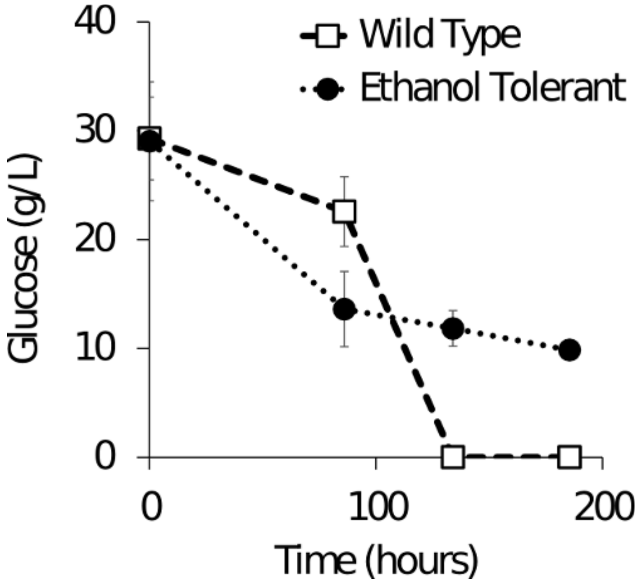

Fig S1 *C. phytofermentans* ET strain shows reduced glucose consumption relative to wild-type when growing in GS2 medium initially containing 30 g/l glucose. Data points are averages of 4 independent cultures and error bars are 1 standard deviation.

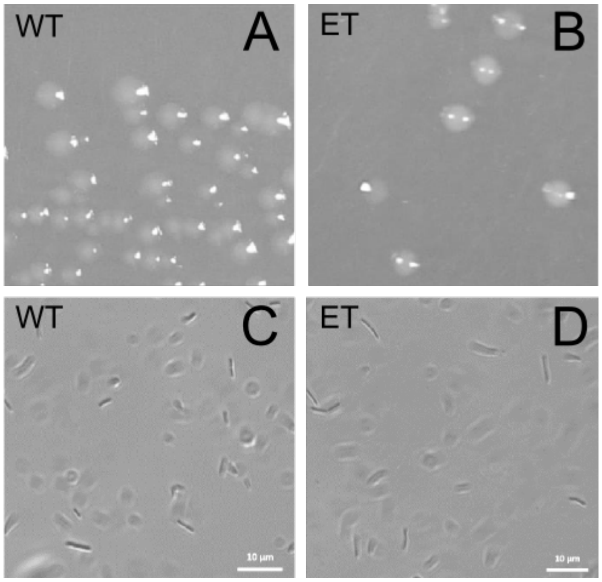

Fig S2 *C. phytofermentans* WT and ET strains have similar morphologies for **A-B** colonies growing on GS2 agar plates and **C-D** cells visualized by bright field microscopy using a Zeiss AxioScope-A1 with a Zeiss EC Plan-173 NEOFLUAR 100x/1.3 oil immersion objective. Scale bar is 10 micrometers.

Table S1 Generation times (hours) and maximum cell densities (OD_{600}) of *C. phytofermentans* wild-type (WT) and ethanol-adapted (ET) strains grown in microtiter plates at 30ºC in GS2 medium supplemented with 0, 2, 4, 6, 7% (v/v) ethanol.

Strain	Ethanol added (v/v)	Generation time (h)	Maximum cell density (OD600) Mean	Maximum cell density (OD600) Standard Deviation
WT	0	2.47	0.87	0.02
WT	2	2.51	0.82	0.06
WT	4	ND	0.19	0.05
WT	6	ND	0.08	0.05
WT	7	ND	0.07	0.04
ET	0	4.15	0.62	0.04
ET	2	4.22	0.66	0.05
ET	4	6.42	0.46	0.04
ET	6	5.76	0.30	0.06
ET	7	ND	0.15	0.02

Supplementary text: Analysis of the genome sequence of the *C.phytofermentans* ethanol tolerant (ET) strain

FASTQ-formatted DNA sequencing files for the *C. phytofermentans* ET genome were submitted to the European Nucleotide Archive under Primary Accession PRJEB7255.

DNA sequencing reads were quality-filtered using Picard tools (https://github.com/broadinstitute/picard). Specifically, reads in fastq format were merged into bam files and reverted with RevertSam to restore original qualities from the OQ field to the QUAL field and to remove duplicate read flags and all alignment information. 'MarkIlluminaAdapters' was run to clear any existing adapter-trimming tags for PAIRED_END adapter sequences. The reads were aligned to the NCBI reference genome NC_010001.1 (http://www.ncbi.nlm.nih.gov/nuccore/NC_010001.1) with bwaMem version 0.7.7 using the -M flag to mark shorter split hits as secondary. After alignment, duplicate records and optical duplicates were flagged using MarkDuplicates.

DNA variants in ET genome were identified using the Broad Institute GATK (https://www.broadinstitute.org/gatk/). SNP and indel calling was performed using the GATK version 3.1 UnifiedGenotyper with the following parameters:

```
java -jar GenomeAnalysisTK-3.1-144-g00f68a3.jar -R [fasta file of Clostridium phytofermentans
reference genome] -I [bam file of Clostridium phytofermentans ET genome sequence] -T
UnifiedGenotyper -ploidy 1 -stand_call_conf 30 -stand_emit_conf 10
```

Raw output produced by the above GATK analysis:

```
#CHROM  POS     ID     REF     ALT     QUAL     FILTER  INFO     FORMAT  Cphy_Ethanol_Tolerant

Cphy    144943  .       C       T       8841    .
AC=1;AF=1.00;AN=1;BaseQRankSum=-
0.584;DP=247;Dels=0.00;FS=2.508;HaplotypeScore=11.9565;MLEAC=1;MLEAF=1.00;MQ=60.00;
MQ0=0;MQRankSum=-0.277;QD=28.17;ReadPosRankSum=1.745 GT:AD:DP:GQ:MLPSAC:MLPSAF:PL
1:12,234:247:99:1:1.00:8871,0

Cphy    263809  .       T       A       5569    .
AC=1;AF=1.00;AN=1;BaseQRankSum=-
0.337;DP=249;Dels=0.00;FS=1.221;HaplotypeScore=12.5372;MLEAC=1;MLEAF=1.00;MQ=60.00;
MQ0=0;MQRankSum=0.187;QD=22.37;ReadPosRankSum=-0.005 GT:AD:DP:GQ:MLPSAC:MLPSAF:PL
1:55,192:249:99:1:1.00:5599,0

Cphy    289172  .       C       T       9246    .
AC=1;AF=1.00;AN=1;BaseQRankSum=-
0.107;DP=250;Dels=0.00;FS=2.586;HaplotypeScore=7.1707;MLEAC=1;MLEAF=1.00;MQ=60.00;M
Q0=0;MQRankSum=-0.060;QD=32.72;ReadPosRankSum=1.224  GT:AD:DP:GQ:MLPSAC:MLPSAF:PL
1:11,239:250:99:1:1.00:9276,0

Cphy    333974  .       G       A       7470    .
AC=1;AF=1.00;AN=1;BaseQRankSum=0.656;DP=209;Dels=0.00;FS=0.000;HaplotypeScore=7.196
4;MLEAC=1;MLEAF=1.00;MQ=60.00;MQ0=0;MQRankSum=-1.343;QD=27.32;ReadPosRankSum=-0.272
GT:AD:DP:GQ:MLPSAC:MLPSAF:PL    1:13,196:209:99:1:1.00:7500,0

Cphy    697063  .       C       A       8661    .
AC=1;AF=1.00;AN=1;BaseQRankSum=-
1.428;DP=248;Dels=0.00;FS=4.733;HaplotypeScore=12.7853;MLEAC=1;MLEAF=1.00;MQ=60.00;
MQ0=0;MQRankSum=-0.991;QD=34.92;ReadPosRankSum=1.131 GT:AD:DP:GQ:MLPSAC:MLPSAF:PL
1:18,230:248:99:1:1.00:8691,0
```

```
Cphy    1644823 .       G       T       52      .
AC=1;AF=1.00;AN=1;DP=18;Dels=0.00;FS=0.000;HaplotypeScore=0.0000;MLEAC=1;MLEAF=1.00
;MQ=20.00;MQ0=16;QD=2.89     GT:AD:DP:GQ:MLPSAC:MLPSAF:PL 1:16,2:18:82:1:1.00:82,0

Cphy    3712793 .       T       C       9011    .
AC=1;AF=1.00;AN=1;DP=224;Dels=0.00;FS=0.000;HaplotypeScore=13.5533;MLEAC=1;MLEAF=1.
00;MQ=60.00;MQ0=0;QD=31.94 GT:AD:DP:GQ:MLPSAC:MLPSAF:PL
1:0,223:224:99:1:1.00:9041,0

Cphy    3717529 .       C       T       8214    .
AC=1;AF=1.00;AN=1;BaseQRankSum=0.774;DP=249;Dels=0.00;FS=0.000;HaplotypeScore=10.20
34;MLEAC=1;MLEAF=1.00;MQ=60.00;MQ0=0;MQRankSum=0.596;QD=32.99;ReadPosRankSum=0.780
GT:AD:DP:GQ:MLPSAC:MLPSAF:PL    1:22,226:249:99:1:1.00:8244,0

Cphy    3960216 .       G       GTTGCATGAGAAAACCTGCAT   15937.97        .
AC=1;AF=1.00;AN=1;BaseQRankSum=1.517;DP=235;FS=0.547;MLEAC=1;MLEAF=1.00;MQ=60.00;MQ
0=0;MQRankSum=-
0.769;QD=10.17;RPA=1,2;RU=TTGCATGAGAAAACCTGCAT;ReadPosRankSum=5.925;STR
GT:AD:DP:GQ:MLPSAC:MLPSAF:PL    1:76,112:235:99:1:1.00:15977,0

Cphy    4533687 .       C       A       8824    .
AC=1;AF=1.00;AN=1;BaseQRankSum=1.103;DP=249;Dels=0.00;FS=0.000;HaplotypeScore=5.461
4;MLEAC=1;MLEAF=1.00;MQ=60.00;MQ0=0;MQRankSum=1.764;QD=29.71;ReadPosRankSum=-0.791
GT:AD:DP:GQ:MLPSAC:MLPSAF:PL    1:16,233:249:99:1:1.00:8854,0

Cphy    4631034 .       TA      T       11686.97        .
AC=1;AF=1.00;AN=1;BaseQRankSum=0.155;DP=249;FS=1.645;MLEAC=1;MLEAF=1.00;MQ=60.00;MQ
0=0;MQRankSum=0.870;QD=26.93;RPA=7,6;RU=A;ReadPosRankSum=-0.184;STR
GT:AD:DP:GQ:MLPSAC:MLPSAF:PL    1:6,238:248:99:1:1.00:11726,0

Cphy    4823500 .       G       A       8929    .
AC=1;AF=1.00;AN=1;BaseQRankSum=-
2.003;DP=246;Dels=0.00;FS=3.092;HaplotypeScore=5.1933;MLEAC=1;MLEAF=1.00;MQ=60.00;M
Q0=0;MQRankSum=0.412;QD=31.80;ReadPosRankSum=-2.821  GT:AD:DP:GQ:MLPSAC:MLPSAF:PL
1:8,238:246:99:1:1.00:8959,0
```