Electronic Supplementary Information (ESI)

Low carbon fuels and commodity chemicals from waste gases – Systematic approach to understand energy metabolism in a model acetogen

Esteban Marcellin[a†], James B. Behrendorff[b†], Shilpa Nagaraju[b], Sashini DeTissera[b], Simon Segovia[b], Robin W. Palfreyman[a], James Daniell[b,c], Cuauhtemoc Licona-Cassani[a], Lake-ee Quek[a], Robert Speight[a,d], Mark P. Hodson[a], Sean D. Simpson[b], Wayne P. Mitchell[b], Michael Köpke[b*] and Lars K. Nielsen[a]

[a] Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Brisbane, QLD 4072, Australia.

[b] LanzaTech Inc, 8045 Lamon Ave, Suite 400, Skokie, IL 60077, USA.

[c] School of Biological Sciences, University of Auckland, Auckland, New Zealand.

[†]Equal contributors

*Corresponding author Dr. Michael Köpke, LanzaTech Inc, 8045 Lamon Ave, Suite 400, Skokie, IL 60077, USA. Phone: +1 (847) 324-2498, Email: Michael.Koepke@lanzatech.com

SUPPLEMENTARY INFORMATION:

Supplementary Materials and Methods

Strain and Media

Clostridium autoethanogenum strain DSM 10061 was obtained from DSMZ (German Collection of Microorganisms and Cell Cultures).

Growth medium PETC-MES contained (per L) 1 g NH_4Cl, 0.1 g KCl, 0.2 g KH_2PO_4, 0.2 g $MgSO_4 \cdot 7\ H_2O$, 0.02 g $CaCl_2$, 1.5 g yeast extract, 0.5 ml of 2 g/L resazurin, 20 g 2-(*N*-morpholino) ethanesulfonic acid (MES), 0.05 g Fe $(SO_4)_2(NH_4)_2.6H_2O$, 0.05 g nitriolotriacetic acid and 5 g of fructose (only for heterotrophic growth), 10 mL trace element solution (TSE) and 10 mL of Wolfe's vitamin solution. The TSE solution composition (per L) was: 2 g nitrilotriacetic acid, 1 g $MnSO_4.H_2O$, 0.8 g Fe $(SO_4)_2(NH_4)_2.6H_2O$, 0.2 g $CoCl_2.6H_2O$, 0.2 mg $ZnSO_4.7H_2O$, 0.02 g $CuCl_2.2H_2O$, 0.02 g $NaMoO_4.2H_2O$, 0.02 g Na_2SeO_3, 0.02 g $NiCl_2.6H_2O$ and 0.02 g $Na_2WO_4.2H_2O$. The vitamin solution composition (per L) was: 2 mg biotin, 2 mg folic acid, 10 mg pyridoxine hydrochloride, 5 mg thiamine HCl, 5 mg riboflavin, 5 mg nicotinic acid, 5 mg calcium pantothenate, 0.1 mg vitamin B12, 5 mg *p*-aminobenzoic acid and 5 mg thioctic acid. The medium was prepared anaerobically [1] and the pH of the medium was adjusted to 5.7. Unless otherwise specified, all chemicals were purchased from Sigma (Sigma Aldrich, United States). Prior to inoculation, 5 mL of reducing solutions 1 and 2 were added into the medium. Reducing solution 1 contained 228 mM L-cysteine-HCl and solution 2, contained 228 mM L-cysteine-HCl and 167 mM of Na_2S. For heterotrophic growth conditions the carbon source was 5 g fructose/L. For autotrophic conditions basic oxygen furnace (BOF) gas with a composition of carbon monoxide (35 %), carbon dioxide (10 %), hydrogen (2 %) and the remainder as nitrogen was used at 25 psi. Gas was collected at Glenbrook steel mill, New Zealand.

Culture condition and sampling for omics experiments

Samples were extracted from three different fermentations in 1 L Schott bottles sealed with butyl rubber stoppers (Glasgerätebau Ochs) with a working volumes of 250 mL, shaking at 100 rpm (32 mm shake orbit). Temperature remained constant at 37 °C throughout the fermentation. Growth was monitored by following the optical density in a spectrophotometer at a wavelength of 600nm and production of standard metabolites acetate, ethanol, 2,3-butanediol and lactate followed by HPLC measurements (see below). Cells were grown up to mid log phase (OD_{600nm} = 0.5-0.6) and then sampled for the Omics analysis. Around 50 mL culture volume was used

for proteomics and transcriptomics each and 5-10 mL for metabolomics. Sampling was performed in parallel and is described below.

Measurements of standard metabolites

HPLC measurements with refractive index detection using the following method: culture samples (400 µL) were mixed with 100 µL of 5-sulfosalicylic acid aqueous solution (1 % w/v). Samples were centrifuged at 12,000 *g* for 3 min and the supernatant was removed for analysis. Metabolites were separated by isocratic elution on a cation exchange column for organic acids (AllTech IOA-2000 organic acid column, 150 x 6.5 mm, 8 µm) (Grace, USA), column temperature 60 °C, flow rate 0.7 mL/min, mobile phase: 0.005 N sulphuric acid, detector temperature 35 °C.

Proteomics

Proteins were extracted from cell pellets sampled in triplicate. 50 mL of cells were lysed using a bead beater (Mini Beadbeater 24, Biospec) with glass beads (0.1 mm diameter) for 5 minutes. 200 mg of proteins were digested overnight with trypsin (Promega) and analysed via Nano-LC MS/MS. The first separation was conducted offline on an Agilent Off-gel 3100 fractionator as described in the manufacturer's protocol. A 24 cm GE Healthcare Immobiline DryStrip pH 3-10 was used. Fractionation was carried out using the default program for peptides and 24 cm strips OG24PE01. The 24 resulting fractions were combined into 12 fractions, desalted using a Sep-Pak tC_{18} 1 cc Vac Cartridge (Waters), concentrated to remove ACN from elution buffer and resuspended in 0.1% FA prior to injection. The mass spectrometer, 5600 (ABSciex), was equipped with a nano-spray ESI sources operated in positive ion mode coupled to a Nano-LC (Shimadzu Prominence). Peptides were separated using a flow rate of 30 µL/min on a Vydac Everest C18 column (300 A, 5 µm, 150 mm x 150 µm) at a flow rate of 1 µL/min and a gradient of 10-60% mobile phase B over 90 min. Analyst® Software (version 1.5.2, AB Sciex) was used for peak picking with a method searched for masses of 300 to 1800 Da. Information Dependent Acquisition (IDA) selected for +2 to +4 charges which exceeded 150 counts using Enhanced Resolution scans. The two most abundant ions in each of these scans (or with unknown charge) were subjected to MS/MS. An Enhanced Product Ion scan was used to collate fragment ions and present the product ion spectrum for subsequent database analysis. Protein Pilot Software v 4.5 (Applied Biosystems) and the Paragon Algorithm were used for peptide identification using a fasta-formatted file with all protein sequences reported for the *C. autoethanogenum*. The theoretical ions and peaks were

matched using the tolerance used by the Paragon Algorithm search, based on information about the mass accuracy of the instrument chosen in the Paragon Method dialog box. Search parameters included iodoacetamide as cysteine modification, trypsin as the enzyme for protein digestion and 'Thorough ID' search effort using a detected protein threshold of 95% allowing for false discovery rate analysis (FDR). Only proteins with a confidence score of 95% or better (estimated global FDR 5% or lower) were accepted. For a protein to be identified, at least two 95% confident independent peptide identification were required.

Quantitation was performed using iTRAQ [2]. After trypsin digestion 50 µg of protein were labeled at room temperature for one hour, combined and partially evaporated using a vacuum centrifuge to remove excess ethanol from the iTRAQ reagents. Samples were subsequently cleaned using an ICAT cation exchange cartridge (AB Sciex) as described in the iTRAQ protocol. Finally, samples were re-concentrated to remove residual ACN and desalted with a Sep-Pak tC_{18} 1 cc Vac Cartridge (Waters). Samples were finally evaporated for the last time to eliminate residual ACN and re-suspended in 0.1% formic acid. All statistical analysis was done using Protein Pilot 4.5.

Transcriptomics

Total RNA was extracted from OD_{600nm} = 25 of culture (around 50 mL of a culture with OD_{600nm} = 0.5), centrifuged at 4000 *g* for 10 min at room temperature and resuspended in 1 mL of TRIzol reagent. Cells in TRIzol reagent were transferred to 2 mL screw cap tubes containing 0.8 g of RNAse-free glass beads. Cells were disrupted using a bead beater with 3 cycles of 1 min at 4800 rpm. After removing beads by centrifugation, 100 µL of 1-Bromo-3-Chloropropane (BCP) was added followed by room temperature incubation with occasional vortex mixing. The mixture was centrifuged for 15 minutes at 4 °C at 13000 rpm and the aqueous phase was transferred to a fresh tube. The RNA was washed twice in 70 % ice cold ethanol and resuspended in RNase free water. DNA was removed using DNase TURBO (Life Sciences) followed by column purifications using RNAeasy (Qiagen). RNA quality was evaluated using BioAnalyzer (Agilent) and Nanodrop 1000 (Thermo Scientific) prior analysis. Ribosomal RNA was removed using magnetic beads from the ribo-zero meta-bacterial kit (Epicentre). mRNA libraries for sequencing were prepared using TruSeq RNA sample preparation kits from Illumina. All sequencing was performed on the Illumina Hiseq-2000.

Reads were aligned to the genome using Bowtie 2 [3], with two mismatches allowed per read alignment. Before mapping to the genome the 100 base pair reads were trimmed to 75 base pair to avoid reading errors. Transcript abundance was estimated using FPKM using upper-quantile normalization, Cuffdif [4] was used to estimate differentially expressed transcripts. A $\log_2$ fold change of +2 or -2 and 0.05 false discovery rate was used to estimate significant changes.

Metabolomics.

Intracellular metabolites were extracted from 5 and 10 mL of culture after removing media by centrifugation at 18,700 *g* for 2 minutes at -20 °C inside and outside the anaerobic chamber for comparison. Cells were rapidly washed in 10 mL of PBS buffer, centrifuged again under the same conditions and resuspended in lysis buffer: ice-cold 50 % acetonitrile. Cellular debris was removed by centrifugation and metabolites were freeze dried and resuspended in 1 mL of MilliQ water. Intracellular metabolites were measured in a Dionex Ultimate 3000 coupled to a QTRAP 4000 (ABSciex) mass spectrometer, operated in negative ion mode. Chromatographic separation was achieved on a Gemini-NX C18 column operated at 55 °C. Mobile phase A was 7.5 mM tributylamine aqueous solution (adjusted to pH=4.95 with acetic acid) and B was acetonitrile. Eluent was infused directly into the mass spectrometer. All samples were injected in triplicate in 10 μL volume. The samples were run with sample- and analyte-relevant calibration standards and pooled QC to control for reproducibility of data acquisition and to ensure data integrity. Bio-triplicated samples were performed for the intracellular metabolites measured (Fig. SI 4 and Supplementary Table SI9). AZT was used as internal standard and used for normalization (Bar Chart in supplementary Table 9). All samples were injected twice.

Genome scale model assembly.

The genome-scale metabolic reconstruction was adapted from the method applied to the GEM of *Mus musculus* [5], *Arabidopsis* (AraGEM) [6], and *S. erythraea* [7]. The core of the genome-scale model was reconstructed using the SEED model annotation pipeline [8]. The reconstruction retained all reaction attributes from SEED model, including unique reactions, compound IDs and the reversibility of reactions. The model was manually gap filled in Excel (Microsoft Corporation) for ease of annotation and commenting. From this gene-centric database, a 2D reaction-centric SBML (System Biology Markup Language, http://www.sbml.org) representation was generated using an in-house Java (Oracle Corporation) application.

Constraint-based reconstruction and analysis was performed using the COBRA toolbox (http://opencobra.sourceforge.net/) [9]; a set of scripts for constraint-based modelling run within the MATLAB environment. The XML model can be downloaded from the supplementary material.

Mutant construction and analysis

ClosTron constructs were designed for disruption of GAPDH genes CAETHG_3424 and CAETHG_1760, Nfn complex gene (CAETHG_1580) and Phosphoenolpyruvate carbxykinase gene (CAETHG_2721) using the web-based design tool (http://clostron.com/clostron2.php) and the intron targeting plasmids were synthesized by DNA 2.0 (USA). Plasmids were transferred into *C. autoethanogenum* by bacterial conjugation from *E. coli* strain HB101 carrying plasmid R702 (CA434) according to previously-published methods [10,11]. Briefly, 5 mL of *E. coli* CA434 bearing the ClosTron plasmid was grown to $OD_{600\ nm}$ = 0.5, washed with phosphate-buffered saline and mixed with 1 mL of *C. autoethanogenum* grown to $OD_{600\ nm}$ = 0.5. The cell mixture was spotted onto PETC-MES agar plates (15 g agar/L) containing 5 g fructose/L. Conjugations were incubated at 37 °C for 24 h under 25 psi RMG in custom built, pressure rated and gas-tight jars. Cells were subsequently transferred onto PETC-MES agar plates containing 5 g fructose/L, 5 mg clarithromycin/L for selection of transformants bearing the ClosTron plasmid, and 10 mg trimethoprim/L for counter-selection of *E. coli*. Single colonies were re-streaked and then screened using colony PCR. A list of primers and intron targeting sequences are available in (Supplementary Table S1 and Supplementary Table S2). Where the desired genetic disruptions were confirmed by colony PCR [12] and Sanger sequencing across the targeted locus (Macrogen, USA), colonies were grown in PETC-MES medium with fructose and clarithromycin under 25 psi of RMG to a density of $OD_{600\ nm}$ = 2 and were cryopreserved in 15 % (v/v) glycerol.

Growth rates and a selection of extracellular metabolites were compared between the wild-type *C. autoethanogenum* and *C. autoethanogenum* mutant strains under three different growth conditions: PETC-MES with fructose, PETC-MES with fructose and BOF gas, and PETC-MES with BOF gas only. Strains were inoculated from glycerol stocks (n = 3 biological replicates) into 2 mL PETC-MES containing fructose in 12-well plates (Nunc, Denmark), and grown under 25 psi of BOF gas in gas-tight jars at 37 °C with shaking. Growing cultures were subcultured (1 in 25 dilution) into 6 mL PETC-MES in 6-well plates (Nunc, Denmark) and grown under the three different conditions: PETC-MES with 10 g fructose/L, PETC-MES with

10 g fructose/L and 25 psi BOF gas, and PETC-MES with 25 psi BOF gas only. *C. autoethanogenum* wild-type and mutant strains were subcultured twice under each of the three media conditions before inoculation into 40 mL medium in pressure resistant 1 L Schott bottles (Schott, USA). Where BOF gas was used in Schott bottles the pressure was limited to 20 psi. Actively growing cultures were inoculated to an initial density of $OD_{600\ nm}$ = 0.025 and incubated at 37 °C with shaking. Fructose, acetate, ethanol, lactate, and 2,3-BDO concentrations were measured via HPLC as described above.

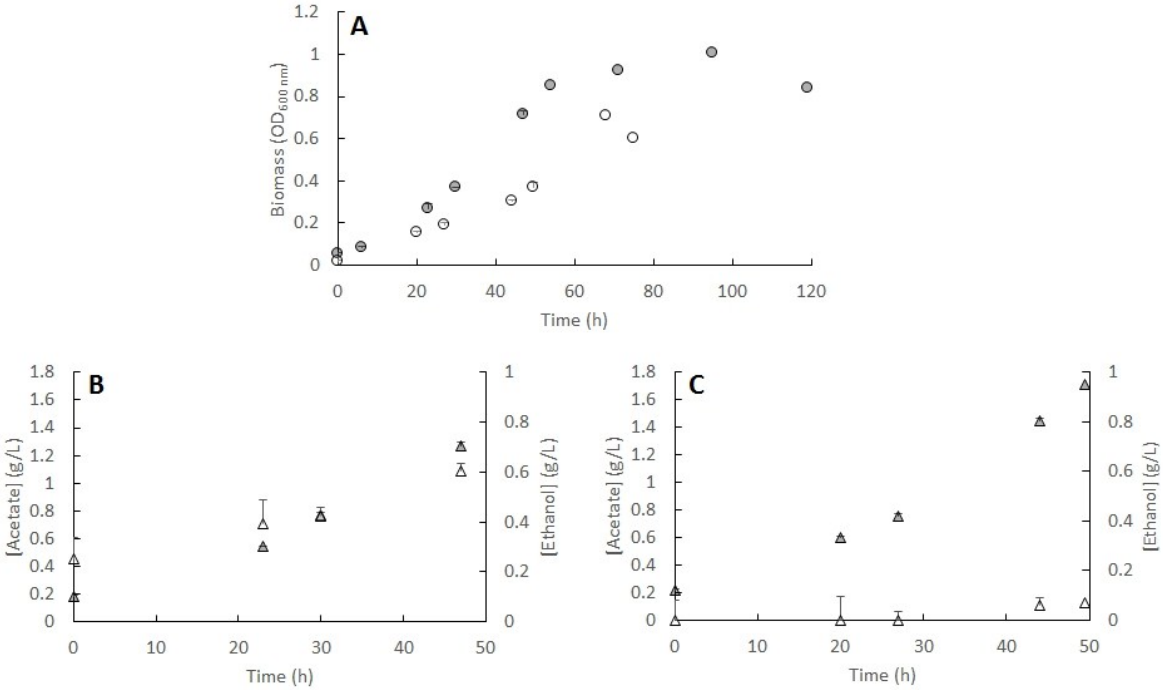


Figure SI 1. (A) Growth profile (measured by OD_{600nm}) for *C. autoethanogenum* grown on fructose (grey) and BOF gas (white). (B) Fructose fermentation products and (C) BOF gas fermentation products. Acetate (grey, primary axis), ethanol (white, secondary axis).

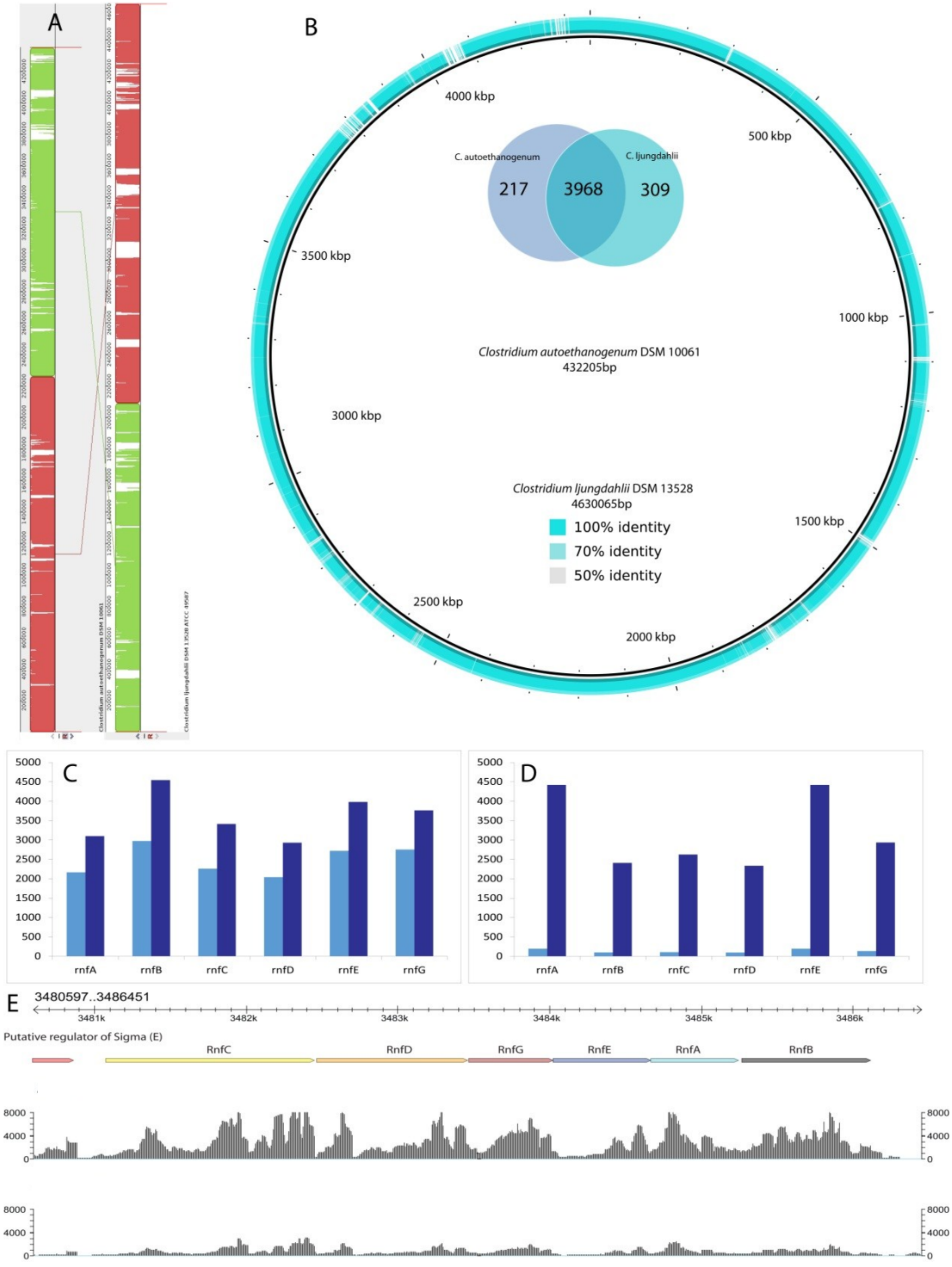
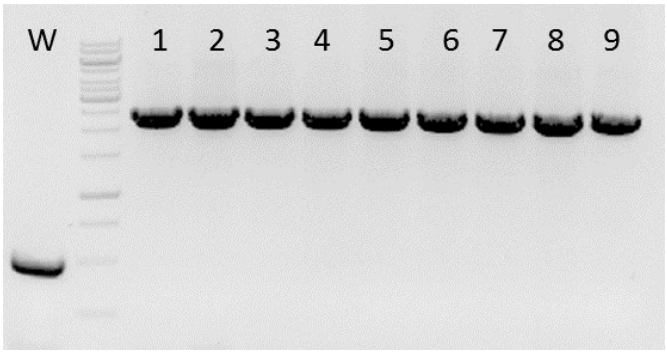


Figure SI 2. Comparative analysis between *Clostridium autoethanogenum* and *Clostridium ljungdahlii*[13] highlight the importance of the Rnf complex in the energy metabolism of acetogens. While no differences are observed at the genomic level for the Rnf complex, at the transcriptional level, significant changes were observed in the Rnf complex levels of expression. A) Alignment of the two genomes using MAUVE (ref) and B) BRIG (ref). RNA-seq data (FPKM) of C) *C. autoethanogenum* and D) *C. ljungdahlii* under autotrophic (light blue) and heterotrophic (blue) conditions of the genes of the Rnf complex[13]. E) Mapped RNA-seq reads from autotrophic (top) and heterotrophic (bottom) conditions to Rnf complex of *C. autoethanogenum*.

A)



B)

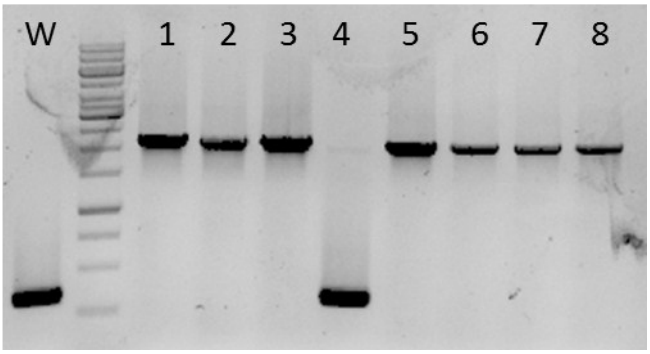


Figure SI 3. Colony PCR to screen and confirm ClosTron mutants. A) Nfn gene disruption in *C. autoethanogenum*. Nine colonies (1-9) were screened for Nfn gene disruption using primers CAETHG_1580_F and CAETHG_1580_R. Amplification of PCR product of 2260bp from all 9 clones indicates the disruption of Nfn gene. Wildtype *C. autoethanogenum* (W) was used as a control to amplify unmodified Nfn gene fragment of 460bp. B) PEP carboxykinase (PCK) gene disruption in *C. autoethanogenum*. Eight colonies (1-9) were screened for PCK gene disruption using primers CAETHG_2721_F and CAETHG_2721_R that would PCR amplify a fragment of 361bp from wildtype *C. autoethanogenum*. Amplification of PCR product of 2161bp from all four colonies indicates the disruption of PEP carboxykinase gene. Gene Ruler 1kb DNA ladder (Thermo) was used.

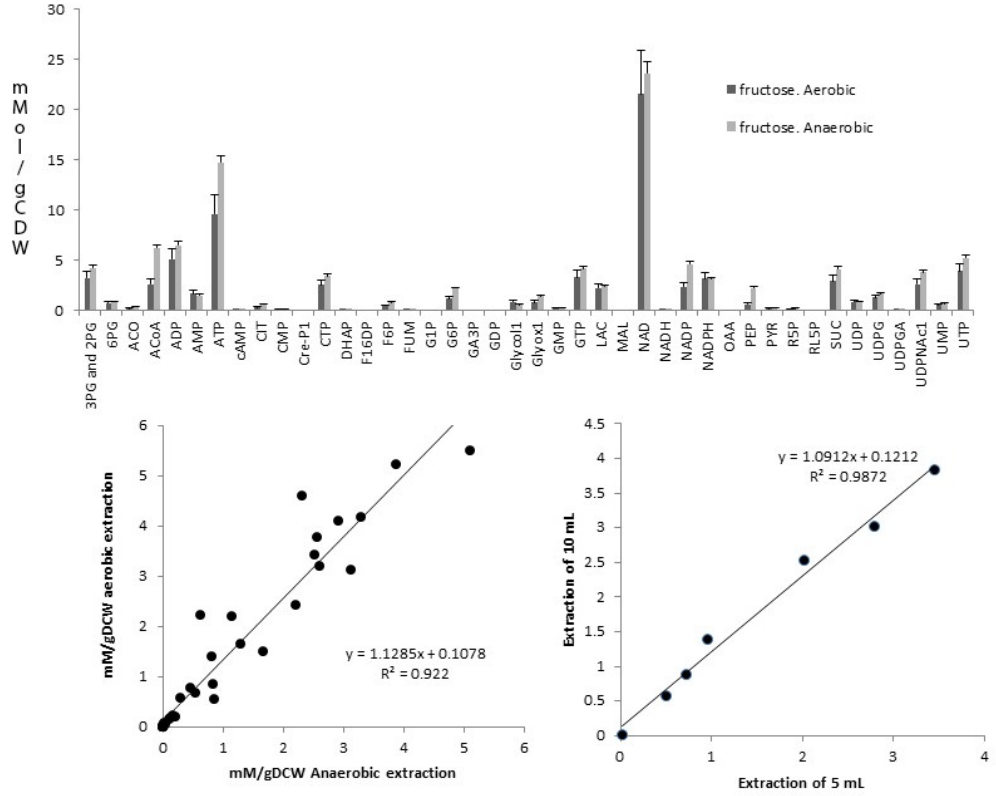


Figure SI 4. To evaluate different cell separation methods, cells extracted in anaerobic conditions were compared to cells extracted outside the anaerobic chamber for 55 intracellular metabolites from three replicate fermentations (A and B). In addition, extraction of 5 mL of culture was compared to extraction of 10 mL of culture (C). After normalizing for dry cell weight, no differences were observed for both volumes extracted (p=0.8). Direct comparison of the two extraction protocols also showed no significant differences between the techniques (p=0.56). Abbreviations: 3PG and 2PG, 3- and 2-phospho-D-glyceric acid; 6PG, 6-phosphogluconic acid; ACO, aconitic acid; ACoA, acetyl coenzyme A; ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; cAMP, Cyclic adenosine monophosphate; CMP, cytidine monophosphate; Cre-P1, creatine phosphate; CTP, cytidine triphosphate; DHAP, dihydroxyacetone phosphate; F16DP, fructose 1,6,-bisphosphate; F6P, fructose 6-phosphate; FUM, fumaric acid; G1P, glucose 1-phosphate; G6P, glucose 6-phosphate; GA3P, glyceraldehyde 3-phosphate; GDP, guanosine diphosphate; Glycol1, glycolic acid; Glyox1, glyoxylic acid; GMP, guanosine monophosphate; GTP,guanosine triphosphate; LAC, lactic acid; MAL, malic acid; NAD, nicotinamide adenine dinucleotide (oxidised); NADH, nicotinamide adenine dinucleotide (reduced); NADP, nicotinamide adenine dinucleotide phosphate (oxidised); NADPH, nicotinamide adenine dinucleotide phosphate (reduced), OAA, oxaloacetic acid; PEP, phosphoenolpyruvic acid; PYR, pyruvic acid; R5P, D-ribose 5-phosphate; SUC, succinic acid; UDP, uridine diphosphate; UDPG, uridine diphosphate glucose; UDPGA, uridine

diphosphate glucuronic acid; UDPNAc1, uridine diphosphate N-acetylglucosamine; UMP, uridine monophosphate; UTP, uridine triphosphate.

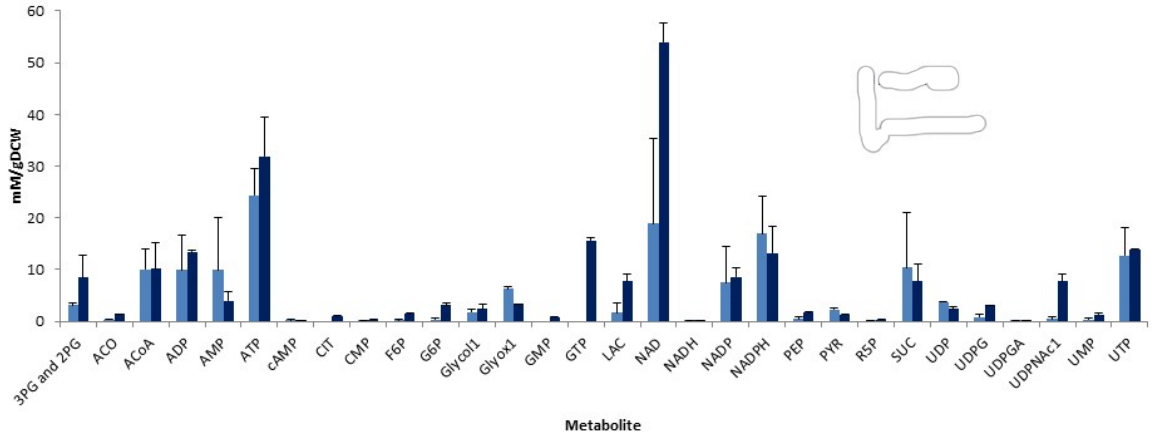


Figure SI 5. Rapid centrifugation was used to compare autotrophic (light blue) and heterotrophic growth (dark Blue). Error bars display standard deviation from biological triplicates. Minimal differences for metabolites levels were observed for energy metabolites and redox metabolites when cells were grown under heterotrophic and autotrophic conditions. The obvious exception was for glycolytic intermediates which were higher when cells grew in the presence of fructose. 3PG and 2PG, 3- and 2-phospho-D-glyceric acid; 6PG, 6-phosphogluconic acid; ACO, aconitic acid; ACoA, acetyl coenzyme A; ADP, adenosine diphosphate; AMP, adenosine monophosphate; ATP, adenosine triphosphate; cAMP, Cyclic adenosine monophosphate; CMP, cytidine monophosphate; Cre-P1, creatine phosphate; CTP, cytidine triphosphate; DHAP, dihydroxyacetone phosphate; F16DP, fructose 1,6,-bisphosphate; F6P, fructose 6-phosphate; FUM, fumaric acid; G1P, glucose 1-phosphate; G6P, glucose 6-phosphate; GA3P, glyceraldehyde 3-phosphate; GDP, guanosine diphosphate; Glycol1, glycolic acid; Glyox1, glyoxylic acid; GMP, guanosine monophosphate; GTP,guanosine triphosphate; LAC, lactic acid; MAL, malic acid; NAD, nicotinamide adenine dinucleotide (oxidised); NADH, nicotinamide adenine dinucleotide (reduced); NADP, nicotinamide adenine dinucleotide phosphate (oxidised); NADPH, nicotinamide adenine dinucleotide phosphate (reduced), OAA, oxaloacetic acid; PEP, phosphoenolpyruvic acid; PYR, pyruvic acid; R5P, D-ribose 5-phosphate; SUC, succinic acid; UDP, uridine diphosphate; UDPG, uridine diphosphate glucose; UDPGA, uridine

diphosphate glucuronic acid; UDPNAc1, uridine diphosphate N-acetylglucosamine; UMP, uridine monophosphate; UTP, uridine triphosphate.

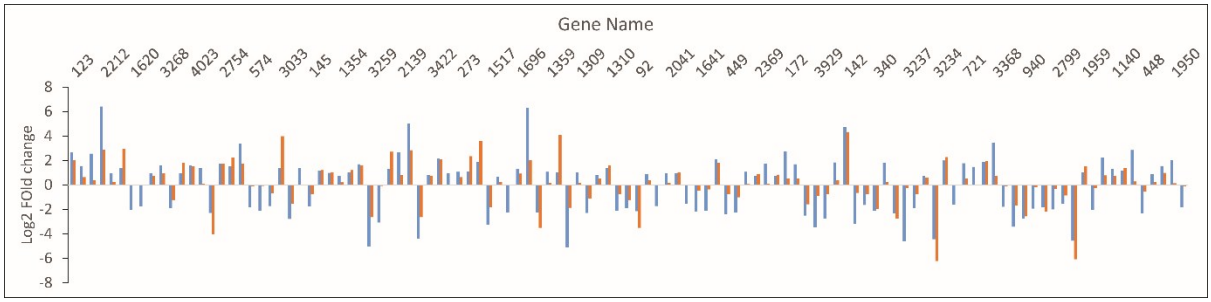


Figure SI 6. Transcriptional and translational comparison of significantly up-regulated enzymes from the iTRAQ experiment. The data shows iTRAQ (blue)and RNA-seq (orange)log 2 fold change for the genes found in proteomics and transcriptomics. Y-axis shows CAETHG gene numbers according to the NCBI annotation.

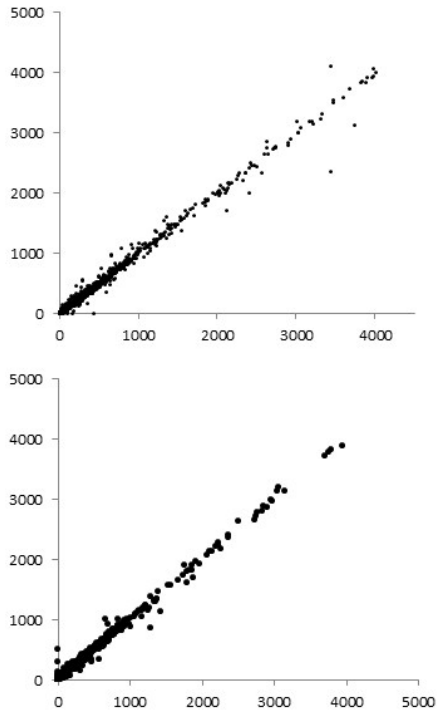


Figure SI 7. RNA-seq data for biological replicates grown on gas and fructose.

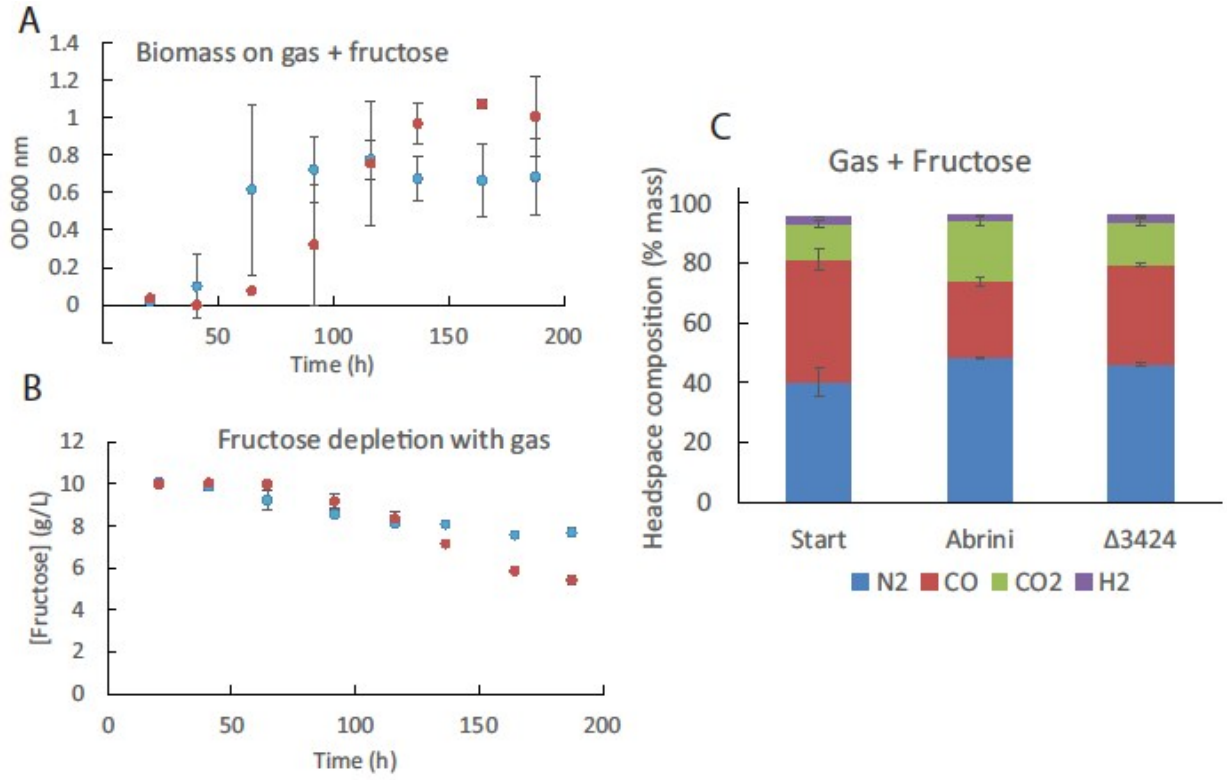


Figure SI 8. Growth curve for CAETHG_3424 mutants on Fructose and RMG for the WT Abrini (blue) and the knock out CAETHG_3424 strain (red) (A), scattered plot for fructose consumption (B) stacked column plot for the overhead gas composition before and after the fermentation (C).

Supplementary Table SI 1. Information for ClosTron experiments for GAPDH genes CAETHG_3424 and CAETHG_1760

Nucleotide sequence
ATGTCCATAAAAGTTGGAATTAATGGTTTTGGAAGAATAGGAAGGTTGGTTTTGAGAATT GCTCAAGAAAGACTGAAGGACAATATTGAGATTGTAGCAATTAACGCAAGAGCAGATAAC GAAACTCTGGCACATCTTCTAAAATACGACTCCTGCTACGGAAGATTTAATGGAAATGTA AAGGCAGCAGAAGATTCAATTATTGTAAATGGTTCTAAAATAAAAGTGTTTAGAGAAAAT GATCCTGAAGATTTACCTTGGGGTAAATTGGATGTAGATATAGTTATAGAATCCACGGGA AAGTTTAAAGATAGAGAAAGCTTATACAAGCATATAAAAGCTGGAGCAAAAAAAGTTATA ATTACAGCTCCTGCAAAGGACGAAGATATAACTATAGTTATGGGAGTAAATCAAGAAAAA TTTGATGTTGAATCTCATAATATTATCTCAAATGCATCCTGTACAACTAATTGTTTGGCA CCTTTTGCTAAAGTATTAGATAAGAATTTTGGAATTATAAGAGGCCTTATGACTACAGTT CACTCTTATACAAATGACCAACAATTGCTTGATAAAACTCACAAGGATTTGAGAAGAGCT AGAGCTGCAGGAGAATCTATAATTCCGACTACTACTGGAGCAGCTAAGGCAGTGGCAAAA GTATTGCCTAATTTAGAGGGAAAATTAAGCGGTTTTGCACTTAGAATACCTACACCTACA GTTTCTATTACAGACTTGGTCTGTGAACTTTCAAAATCTGTTACGGTAGAAGAAGTAAAT GAAGCTTTTAAAAGAGCTTCTGAGGGATCTATGAAAGGAATACTGGGATATTCTGAAGAA CCACTGGTATCTATAGATTATAAGGGAGACGATAGGTCTTCTATAGTAGATGGTCTTTCA ACTATGGTATTAAAGGATAATATGATTAAAGTAGTATCCTGGTATGATAATGAATGGGCT TATTCCTGTAGAACTGTAGATTTGGCTAATTATATTGCAGGTGAGATGGAGAAGGCTAGT GAAAAACCTATTGATATGGTTGCAGAAGGAGTTTAA
ClosTron insertion targeted between nucleotides 135 and 136
Primers for insertion screening: CAETHG_135_F: TGCTCAAGAAAGACTGAAGGACA CAETHG_135_R: GCTGCTCCAGTAGTAGTCGG

Nucleotide sequence
ATGGTTAAGGTTGGAATTAATGGTTTTGGAAGAATAGGAAGGAATGTATTTAAAGCATTA GTAAAAAGATATGCTAATGAATTACAGGTTGTAGGTATCAATGATTTAACTGATGCAAAA ACATTAGCTCACTTATTAAAATATGATTCTTTATATGGCAGATTTGATGGAACTGTAGAA GCAAAGGAAAATTCTATAGTTGTAAATGGAAATGAAATAAAAATATTTGCAGAAAGAGAT CCTAAAAATATAGATTGGAATTCATTAGGTGCTGAAATAGTTATAGAATCCACTGGACTT TTCACAGATGCTACAAAAGCCAATGCACATTTGGGAGGATCAGTTAAAAAAGTTCTTATA TCTGCTCCAGCTAAAAATGAAGATGCAACAATAGTTATGGGAGTTAATGAAGAAACTTAT GATGCTTCTAAACATAATATAATTTCAAATGCATCTTGTACTACAAACTGTCTTGCACCA TTTGCTAAAATATTAGATAGAGAATTTAAGATAGTAAAGGGACTTATGACTACAGTACAT TCATATACAGGAGACCAAAGACTTTTAGATGCTCCTCACAGAGATATGAGAAGAGCTAGA GCAGCTTGTGAATCTATGATTCCAACTACTACAGGAGCAGCAAAGGCAGTTGCACTTGTA CTTCCTCAATTAAAGGGAAAATTAAATGGATTTTCTCTTAGAGTACCTACACCTACAGTA TCATGTACAGACTTAGTTGCAGAACTTTCAAGAGATGTAACAGTAGAAGAAGTAAATGAA GCATTCAAAAAAGCTGCTGAAAGTGATATGAAGGGAATATTAGGATACAGTGATGAACCT TTGGTATCTGTAGATTATAGAGGAGATGAAAGATCTTCTATAGTGGATGGACTTTCAACT ATGGCTATAGGCGGAAACATGGTTAAAGTTGTTGCTTGGTATGATAACGAATTTGGATAT TCAAATAGACTAGCAGATTTAACTAAATATGTTGCCGACAGACTATAA
ClosTron insertion targeted between nucleotides 399 and 400
Primers for insertion screening: CAETHG_1760_399F: TGCTAATGAATTACAGGTTGTAGGT CAETHG_1760_399R: CCATGTTTCCGCCTATAGCC
ClosTron insertion targeted between nucleotides 273 and 274
Primers for insertion screening: CAETHG_1760_273F: TGGTTAAGGTTGGAATTAATGGTTT CAETHG_1760_273R: TGCCTTTGCTGCTCCTGTAG

Supplementary Table SI 2. Information for ClosTron experiment for Nfn complex gene CAETHG_1580 and Pyruvate carboxykinase gene CAETHG_2721

Nucleotide sequence
TTATTTTTTCAGATATTCATCTATAGCCTTAGCAGCTTTTTTACCTGCACCCATCGCAAG TATTACAGTAGCAGCACCTGTTACTGCATCTCCACCTGCATATATTCCTTCTCTTGTAGT AAGTCCAGTCTCTTCTTCTGCTATTAAGCATTTACGCTTATTCATTTCAAGTCCTTTTGT TGTAGTTGATATAAGTGGATTTGGTGAAGTACCAAGGGCCATAATTACAGTATCTACATC TAAATCAAATTCTGAGCCCTTTATTGCAACTGGTTTTCTTCTTCCAGATGCATCTGGTTC TCCAAGTTCCATCTTAATACAGCGTATTCCATTAACCCAACCTTTTTCATCTCCTAATAT TTCTACAGGATTAGTTAAAAGATTGAACTTTATTCCTTCCTCTTTTGCATGGTGTACTTC CTCTGCTCTTGCTGGAAGTTCTGCTTCGGATCTTCTGTATACTATGTATACTTCTGCTCC AAGTCTTAGAGCTGTCCTTGCAGAGTCCATAGCTACGTTTCCGCCTCCTACTACAGCTAC TTTCTTACCAGCTTTTATAGGAGTTGCATAATCATCCCTATATGCTTTCATTAAATTACT TCTTGTTAAGAATTCATTTGCAGAGAATACTCCATTTAAGTTTTCTCCAGGTATTCCCAT AAACCTTGGTAGTCCTGCTCCTGAACCTATAAATACAGCATCAAATTTTTCTTTTTCTAC TAGTTCATCTATAGTAACAGTTCTTCCTATTATTACATCTGTCTCTATTTTTACTCCTAA TTTCTTTACATTTTCAACTTCATGTTTTACTACAGTATCCTTTGGAAGTCTGAACTCTGG AATACCATATACAAGTACTCCTCCTGCTTCATGAAGTGCTTCAAATATAGTAACGTCATA TCCAAGCTTTGCTAAATCTCCTGCACAAGTAAGTCCTGAAGGACCACTTCCTATAACAGC TACTTTTTTGCCGTTTTTAGGTAAAGTCTTAGATAAATCGATATTATTCTTTCTTGACCA ATCTGCTACAAATCTTTCCAGCTTACCTATAGCAACAGCATCACCTTTTTTGCCAAGTAC ACATTTTCCTTCACACTGAGTTTCCTGAGGACATACTCTTCCACATACAGCAGGAAGTGC ACTTGATTCTGCTATTATTTTAGCAGCTTCTTCAAAGTTTCTATTTTTTACCTGTTCAAC AAACTTAGGTATAGTTATTGTAACAGGGCACTGAGTAACACACATAGGCTTTTTACAATT CAAGCATCTTGAAGCTTCTTTTACAGCTTCTTCTTCAGTGTATCCCAAGCAAACTTCATC AAAATTAGTAGCTCTAACTTTAGGATCCTGTTCCTTTATAGGTACTCTTTTCATCCTAGC AGCTTTATCTTCATTGTGACACTCACAGCCTTTTCTATCAAAAGTGTCTCCTTCTTCTTC TTTCAGAAGTTTCTTTCCTTCTTCAGTTTTATACATAGCCTGTCTTCTCATAGCTTCGTC AAAATTTACAAGATGACCATCAAATTCAGGACCATCTACACAGGCAAATTTTAATTCTCC ACCTACAGTAACCCTACAAGCTCCGCACATTCCAGTTCCATCTACCATTATAGTATTTAA ACTTACTATTGTTTTAATTCCGTATTGTTTCGTAACTTGAGTTATAAACTTCATCATTAT CATAGGCCCTATGGCAACTACACAATCGTACTCTTTATCTTTCTTGTCAATAAGCTCTTT TAAAAGATCCGTAACAAATCCTTTATATCCATAACTTCCATCATCCGTTGCTATATATAA ATTCCCACAAATTGGTTTTAATTCATCTTCAAATAATAGATAATCTTTTGATTTACATCC AACTATAACATCTGCCTCTACCCCATGCTGATGAAGCCACTTTACCTGTGGATAAACTGG TGCAGTACCTACACCACCTGCTACAAACATTATATTTTTCTTCTTAAGTTCTTTAAGATC CATCTCTACAAGTTCTGAGCAATGTCCAAGTGGTCCAACAAAATCTTCAAAATATTCTCC CACTTCATATTTAGCCATTTTTTTAGTTGAAGCTCCTAATGTCTGAAATACAATAGTAAC AGTTCCTTTTCCTGCATCATAATCACAGATAGTAAGAGGTATTCTTTCCCCTTTATCATC CATTTTGACTATGATAAATTGTCCTGGTAAACAGGACTTGGCAACTCTTGGTGCCTCAAT ATCCAT
ClosTron insertion targeted between nucleotides 1515 and 1516

Primers for insertion screening:

CAETHG_1580_F: AGGGCCTATGATAATGATGAAGTTT

CAETHG_1580_R: CTGAGTAACACACATAGGCTTTTTAC

Nucleotide sequence

ATGAATATTGACTTATCATATTTAAATATTAATAAATATAGAAATATGTATAAAAATTTA
TCACCATCAGAATTAACGGAATTTTCAATTAAAAGAGGAGAAGGATTTTTATCAAATAAG
GGAGCTCTTATGATTAATACTGGAAAGTACACAGGAAGATCTCCTAAAGATAGATTTATA
GTTAATCAAGAAAGCATTAGGAACAAAATAAACTGGGGAAATGTAAATCTTTCTATAGAA
GAAGATATTTTTAATAAAATGTATGATAAGATTTTAAATTATATAAGTGATAAAGATATT
TTTGTGTTTGATGGATTTGTTGGAGCTTTAAAAAAATATACCCTTCCTATAAGAGTAATA
TGCGAAAGGGCATCCCAGGCGTTGTTTGCAAATCAATTGTTTAGAAGACCAACGGAGGAG
GATTTAAAGTGTTTTACTCCTGAATTTAATATTATATCGGCACCTGGATTTAAGGCTAAG
GGCAAAGAAGACGGTTTAAATTCAGATGCCTTTATTTTAGTAAATTTCGATAAAAAAATT
ATATTAATAGGTGGAACCAGTTACTCGGGAGAAATAAAAAAATCAGTATTTTCAGTAATG
AACTTCTTGCTTCCACAAAAAGGAGTCATGCCTATGCACTGTTCTGCTAATATAGGACAA
GATAATAAAACTTGCTTATTTTTTGGATTGTCAGGAACAGGAAAAACTACTTTATCAGCA
GATGGTGAAAGAAGACTGATTGGTGATGACGAACATGGATGGTCTAATGAAGGTGTATTT
AATTTTGAGGGTGGATGTTATGCTAAAACTATAAGACTTGATAAGGAAAAGGAAAGTCAG
ATATACAATGCCATAAAATTTGGAACTGTAGTTGAAAATGTAGTGGCAGATGGGAATAGA
GTACCTGATTATAATGATGCTAGATATACTGAAAATACAAGGGCAGCATATCCTATAAAT
TATATAGATAATATAGAAGAAAGTGGTGTAGGAGGAAATCCAGAGACTATAATATTTTTA
ACCGCAGATGCTTTTGGTGTAATGCCACCTATATCAAGGCTTTCTAAAGAAGCAGCAATG
TATCACTTTATGTCTGGATATACTAGCAAGATAGCTGGAACTGAAAGAGGAATAATTGAA
CCTCAAGCTACTTTTTCCTCTTGCTTTGGTGAACCGTTTATGTTAATGAATCCTGCTGTC
TATGCAAAATTGTTAGGCGAAAGAATAGACAAGTATAACACTCAGGTATATTTAGTGAAT
ACTGGATGGCTATCTGGAGGATATGGAAATGGAGATAGAATAAAACTTTCCTATACAAGG
GCTATGATTAGAGAAGCTTTGAAAGGGAAGTTCAAGGATGTTGAATTTGTGGAACATCCT
GTATTTAAAGTAATGATGCCTAAAAGATGTCCAGGTGTACCTGATGAAATATTAAATCCT
AGAAATATATGGGAAGATAAAGAAGCATATGATGAGACAGCGAGAAAGCTGGCGCTGAAG
TTTAGTAAAAACTTTGAGAAGTTTAAAGATGTTTCCGAAGATATAGCAAAAGCTGGACCT
GAAGCTTAA

ClosTron insertion targeted between nucleotides 47 and 48

Primers for insertion screening:

CAETHG_2721_F:ATCACCATCAGAATTAACGGAATTT

CAETHG_2721_R:CTCCTCCGTTGGTCTTCTAAACAAT

Supplementary Table SI 3. RNA-sequenicng data. mRNA libraries for sequencing were prepared using TruSeq RNA sample preparation kits from Illumina. Transcript abundance was estimated using FPKM using upper-quantile normalization, Cuffdif (4) was used to estimate differentially expressed proteins.

Supplementary Table SI 4. Proteins detected. Proteins were extracted from cell pellets sampled in triplicate. Cells were lysed using glass beads for 5 minutes at 4800 rpm. The first separation was conducted offline on an Agilent Off-gel 3100 fractionator. The 24 resulting fractions were combined into 12 fractions, desalted using a Sep-Pak tC_{18} 1 cc Vac Cartridge (Waters), concentrated to remove ACN from elution buffer and resuspended in 0.1% FA prior to injection. The mass spectrometer, 5600 (ABSciex), was equipped with a nano-spray ESI sources operated in positive ion mode coupled to a Nano-LC (Shimadzu Prominence).

Supplementary Table SI 5. Protein quantitation was done using iTRAQ as described in (2). 84 proteins were significantly up or down expressed ($p<0.05$).

Supplementary Table SI 6. Most relevant fluxes. Condition 1, growth on 10 mMol of fructose and max biomass. Condition 2, same simulation as condition 1 and stopping acetate. Condition 3, same simulation as condition 2 and stopping ethanol. Condition 4, same simulation as condition 1 but stop aldehyde ferredoxin oxidoreductase, allow nadh version. Condition 5, same simulation as condition 1 but stop both aldehyde dehydrogenases versions. Condition 6, same simulation as condition 3 and stopping lactate. Condition 7, growth on 10 mMol of fructose and max biomass, with a reversible gapdh (NADH). Condition 8, growth on 10 mMol of fructose and max biomass, with a reversible gapdh (NADPH). Condition 9, growth on gas (CO,CO_2 and H_2) and max biomass. Condition 10, growth on gas (CO and CO2) and max biomass. Condition 11, growth on gas (CO_2 and H_2) and max biomass. Condition 12, growth on gas (CO,CO_2 and H_2) and max biomass but no ethanol. Condition 13, growth on gas (CO,CO_2 and H_2) and max biomass but no acetate or ethanol. Condition 14, same simulation as condition 9 but stop aldehyde ferredoxin oxidoreductase. Condition 15, same simulation as condition 9 but stop aldehyde ferredoxin oxidoreductase and stop aldehyde dehydrogenase. Condition 16, growth on gas (CO,CO_2 and H_2) GAPDH (NADH) toward gluconeogenesis. Condition 17, growth on gas (CO,CO_2 and H_2) GAPDH (NADPH) toward gluconeogenesis. Condition 18, growth on gas (CO,CO_2 and H_2) and set BDO to 2.

Supplementary Table SI 7. All fluxes. Condition 1, growth on 10 mMol of fructose and max biomass. Condition 2, same simulation as condition 1 and stopping acetate. Condition 3, same simulation as condition 2 and stopping ethanol. Condition 4, same simulation as condition 1 but stop aldehyde ferredoxin oxidoreductase, allow nadh version. Condition 5, same simulation as condition 1 but stop both aldehyde dehydrogenases versions. Condition 6, same simulation as condition 3 and stopping lactate. Condition 7, growth on 10 mMol of fructose and max biomass, with a reversible gapdh (NADH). Condition 8, growth on 10 mMol of fructose and max biomass, with a reversible gapdh (NADPH). Condition 9, growth on gas (CO,CO_2 and H_2) and max biomass. Condition 10, growth on gas (CO and CO2) and max biomass. Condition 11, growth on gas (CO_2 and H_2) and max biomass. Condition 12, growth on gas (CO,CO_2 and H_2) and max biomass but no ethanol. Condition 13, growth on gas (CO,CO_2 and H_2) and max biomass but no acetate or ethanol. Condition 14, same simulation as condition 9 but stop aldehyde ferredoxin oxidoreductase. Condition 15, same simulation as condition 9 but stop aldehyde ferredoxin oxidoreductase and stop aldehyde dehydrogenase. Condition 16, growth on gas (CO,CO_2 and H_2) GAPDH (NADH) toward gluconeogenesis. Condition 17, growth on gas (CO,CO_2 and H_2) GAPDH (NADPH) toward gluconeogenesis. Condition 18, growth on gas (CO,CO_2 and H_2) and set BDO to 2.

Supplementary Table SI 8. RNA-seq data used in Figure 1 with gene names and FPKMs.

SUPPLEMENTARY REFERENCES FOR MATERIALS AND METHODS

1 W. E. Balch, G. E. Fox, L. J. Magrum, C. R. Woese and R. S. Wolfe, *Microbiol. Rev.*, 1979, **43**, 260–96.
2 C. A. Orellana, E. Marcellin, B. L. Schulz, A. S. Nouwens, P. P. Gray and L. K. Nielsen, *J. Proteome Res.*, 2015, **14**, 609–18.
3 B. Langmead and S. L. Salzberg, *Nat. Methods*, 2012, **9**, 357–9.
4 C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn and L. Pachter, *Nat. Protoc.*, 2012, **7**, 562–78.
5 L.-E. Quek and L. K. Nielsen, *Genome Inform.*, 2008, **21**, 89–100.
6 C. G. de Oliveira Dal'Molin, L.-E. Quek, R. W. Palfreyman, S. M. Brumbley and L. K. Nielsen, *Plant Physiol.*, 2010, **152**, 579–89.
7 C. Licona-Cassani, E. Marcellin, L.-E. Quek, S. Jacob and L. K. Nielsen, *Antonie Van Leeuwenhoek*, 2012, **102**, 493–502.
8 C. S. Henry, M. DeJongh, A. A. Best, P. M. Frybarger, B. Linsay and R. L. Stevens, *Nat. Biotechnol.*, 2010, **28**, 977–82.
9 J. Schellenberger, R. Que, R. M. T. Fleming, I. Thiele, J. D. Orth, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke and B. Ø. Palsson, *Nat. Protoc.*, 2011, **6**, 1290–307.
10 D. Purdy, T. a T. O'Keeffe, M. Elmore, M. Herbert, A. McLeod, M. Bokori-Brown, A. Ostrowski and N. P. Minton, *Mol. ...*, 2002, **46**, 439–52.
11 D. R. Williams, D. I. Young and M. Young, *J. Gen. Microbiol.*, 1990, **136**, 819–26.
12 L. Plourde-Owobi, D. Seguin, M. A. Baudin, C. Moste and B. Rokbi, *Appl. Environ. Microbiol.*, 2005, **71**, 5604–5606.
13 H. Nagarajan, M. Sahin, J. Nogales, H. Latif, D. R. Lovley, A. Ebrahim and K. Zengler, *Microb. Cell Fact.*, 2013, **12**, 118.