Electronic Supplementary Material (ESI) for Integrative Biology.

Supplementary Information 1

Investigating *Moorella thermoacetica* Metabolism with a Genome-Scale Constraint-Based Metabolic Model

M. Ahsanul Islam[a], Karsten Zengler[b], Elizabeth A. Edwards[c], Radhakrishnan Mahadevan[c,*], and

Gregory Stephanopoulos[a,*]

[a]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[b]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA

[c]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, Ontario M5S 3E5, CANADA

*Corresponding authors: gregstep@mit.edu and krishna.mahadevan@utoronto.ca

Table of Contents

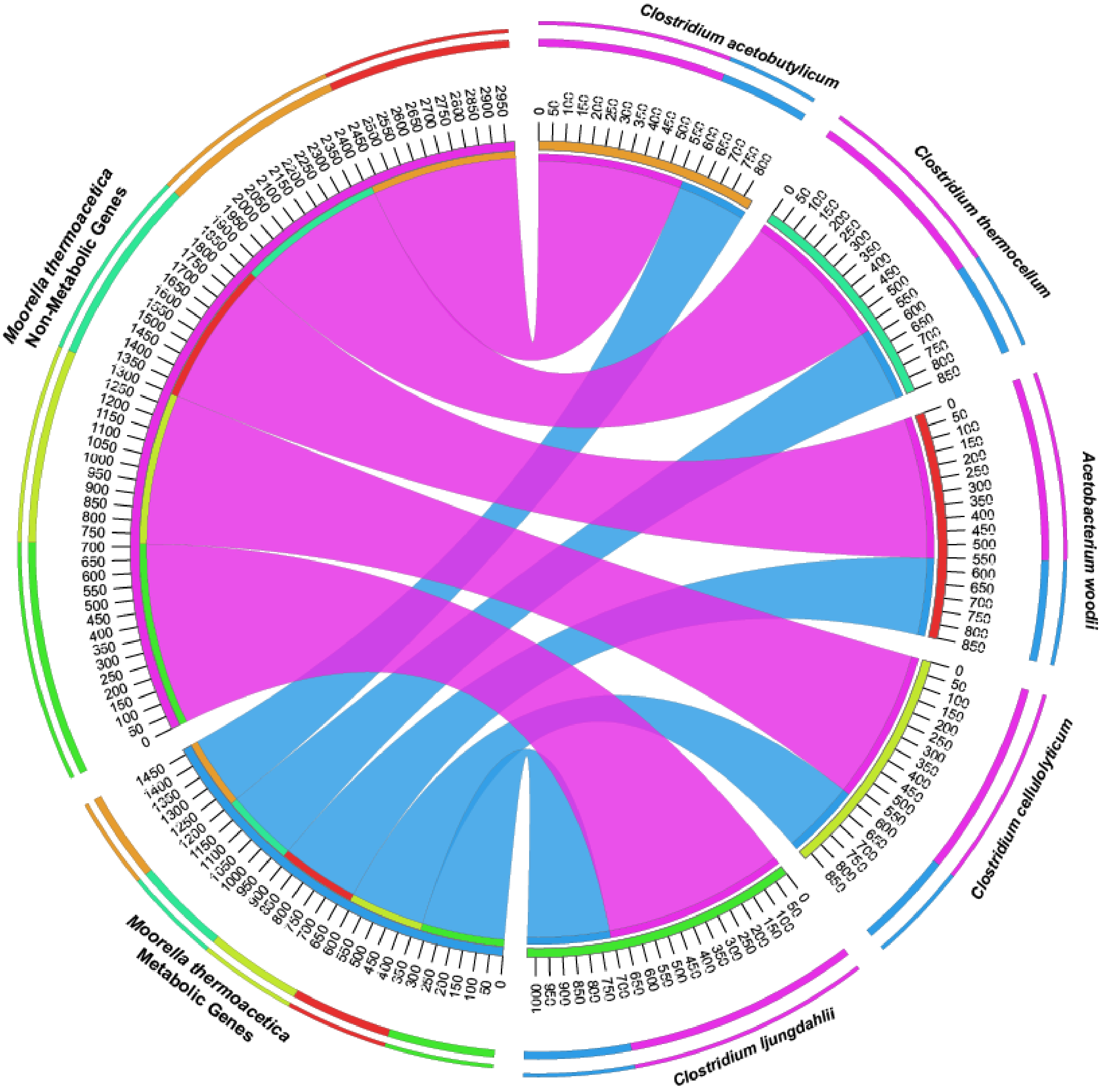

Figure S1. Comparison of the *M. thermoacetica* genome with other model acetogen and model clostridial genomes. The orthologous genes of *M. thermoacetica* metabolic and non-metabolic genes in the 5 other compared genomes are illustrated in this circos figure. The colors are indicating the metabolic and non-metabolic genes, while the ribbons are showing the total number of metabolic and non-metabolic orthologs identified in a genome.

Table S1. General features of *Moorella thermoacetica* metabolic reconstruction

Genes	
Total number of Protein-Coding Genes	2523
Genes Included in the Model	558 (22%)
Genes Excluded from the Model	1965 (78%)
Intra-System Reactions	
Total Number of Model Reactions	705
Gene-Associated Model Reactions	680
Non-Gene Associated Model Reactions	25
Exchange Reactions	
Total Number of Exchange Reactions	60
Input-Output Reactions	59
Biomass Demand Reaction	1
Metabolites	
Total Number of Metabolites	698
Number of Extracellular Metabolites	67
Number of Biomass Metabolites	71

Table S2. Overall cellular composition of a *Moorella thermoacetica* cell[1]

Metabolite Class	% w/w composition
Protein	52.85
DNA	2.60

Metabolite Class	% w/w composition
RNA	6.55
Lipid	7.60
Lipoteichoic acid	3.04
Cell wall components	22.42
Ion and metabolite	4.94
Sum	100.00

Table S3. Protein composition of a *Moorella thermoacetica* cell[2]

Amino Acid	Count (all ORFs) †	% Prevalence	MW (g/mol)	%P*MW	% (by weight)	mmol/gDCW*
Alanine (A)	79490.00	10.37	89.05	923.89	7.23	0.4288
Arginine (R)	52645.00	6.87	175.11	1203.21	9.41	0.2840
Asparagine (N)	23361.00	3.05	132.05	402.63	3.15	0.1260
Aspartic acid (D)	34673.00	4.53	132.04	597.55	4.67	0.1870
Cysteine (C)	8138.00	1.06	121.02	128.54	1.01	0.0439
Glutamate (E)	51805.00	6.76	146.05	987.52	7.72	0.2795
Glutamine (Q)	27224.00	3.55	146.07	519.02	4.06	0.1469
Glycine (G)	65648.00	8.57	75.03	642.88	5.03	0.3541
Histidine (H)	13400.00	1.75	155.07	271.21	2.12	0.0723
Isoleucine (I)	46347.00	6.05	131.09	792.99	6.20	0.2500
Leucine (L)	86082.00	11.24	131.09	1472.84	11.52	0.4644
Lysine (K)	33094.00	4.32	147.11	635.43	4.97	0.1785
Methionine (M)	17155.00	2.24	149.05	333.73	2.61	0.0925
Phenylalanine (F)	26267.00	3.43	165.08	565.95	4.43	0.1417
Proline (P)	38383.00	5.01	115.06	576.42	4.51	0.2071

Amino Acid	Count (all ORFs)[†]	% Prevalence	MW (g/mol)	%P*MW	% (by weight)	mmol/gDCW[*]
Serine (S)	33503.00	4.37	105.04	459.32	3.59	0.1807
Threonine (T)	38340.00	5.00	119.06	595.79	4.66	0.2068
Tryptophan (W)	8681.00	1.13	204.09	231.24	1.81	0.0468
Tyrosine (Y)	23295.00	3.04	181.07	550.53	4.31	0.1257
Valine (V)	58640.00	7.65	117.08	896.09	7.01	0.3163
Sum	**766171**	**100**	**2736.31**	**12786.79**	**100**	

[*] gDCW = gram dry cell weight

Table S4. DNA composition of a *Moorella thermoacetica* cell[2]

DNA	% Prevalence	MW (g/mol)	%P*MW	% (by weight)	mmol/gDCW[*]
dATP	22.105	487.00	10765.15	22.30	0.0119
dCTP	27.895	462.99	12915.09	26.75	0.0150
dGTP	27.895	503.00	14031.07	29.06	0.0150
dTTP	22.105	477.99	10565.95	21.89	0.0119
Sum	**100**		**48277.27**	**100**	

[*] gDCW = gram dry cell weight

Table S5. RNA composition of a *Moorella thermoacetica* cell[2]

RNA	Count (all ORFs)	% Prevalence	MW (g/mol)	%P*MW	% (by weight)	mmol/gDCW[*]
ATP	582554.0	22.2	503.00	11146.68	22.49	0.0293
CTP	732124.0	27.9	478.98	13339.86	26.91	0.0368

RNA	Count (all ORFs)	% Prevalence	MW (g/mol)	%P*MW	% (by weight)	mmol/gDCW[*]
GTP	734515.0	27.9	518.99	14501.25	29.25	0.0369
UTP	579591.0	22.0	479.97	10582.29	21.35	0.0291
Sum	**2628784**	**100**		**49570.07**	**100**	

[*] gDCW = gram dry cell weight

Table S6. Fatty acid composition of a *Moorella thermoacetica* cell[3]

Fatty Acid Components	% (by weight)	Normalized % w/w	MW (g/mol)	Averaged MW (g/mol)	% molar composition (mmol/gDCW[*])
C14:0 normal	1.4	1.4156	227.4	3.2190	0.0127
C15:0 branched (iso)	39.7	40.1416	241.4	96.9017	0.3812
C15:0 normal	2.4	2.4267	241.4	5.8580	0.0230
C16:0 normal	21	21.2335	255.4	54.2305	0.2133
C16:1	1.9	1.9211	253.4	4.8681	0.0191
C17:0 branched (anteiso)	28	28.3114	269.4	76.2709	0.3000
C18:0 normal	3.3	3.3367	283.5	9.4595	0.0372
C18:1	1.2	1.2133	281.5	3.4155	0.0134
Sum	**98.9**	**100**	**2053.4**	**254.2235**	**1**

[*] gDCW = gram dry cell weight

Table S7. Lipid composition of a *Moorella thermoacetica* cell[1]

Components (id)	**Content % (w/w) ***	**MW (g/mol)**	**%P*MW**	**% (by weight)**	**mmol/gDCW**[*]
Lipoteichoic acid (n=24), linked, glucose substituted	19	8445.3175	1604.6103	22.6795	0.0007
Lipoteichoic acid (n=24), linked, N-acetyl-D-glucosamine	19	9430.5895	1791.8120	25.3254	0.0006
Lipoteichoic acid (n=24), linked, D-alanine substituted	40	6692.1895	2676.8758	37.8348	0.0018
Lipoteichoic acid (n=24), linked, unsubstituted	22	4553.9335	1001.8654	14.1603	0.0015
Sum	**100.0**		**7075.2**	**100.0**	

[*] gDCW = gram dry cell weight

Table S8. Cell wall composition of a *Moorella thermoacetica* cell[1]

Components (id)	**Content % (w/w) ***	**MW (g/mol)**	**%P*MW**	**% (mole)**	**mmol/gDCW**[*]
Peptidoglycan subunit (peptido_MT)	45	991	445.9500	6.2786	0.1018
Glycerol teichoic acid (n=45), unlinked, unsubstituted (gtca1_45_MT)	11.9	7373.6	877.4584	12.3539	0.0036
Glycerol teichoic acid (n=45), unlinked, D-ala substituted (gtca2_45_MT)	11.9	11382.8	1354.5532	19.0711	0.0023
Glycerol teichoic acid (n=45), unlinked, glucose substituted (gtca3_45_MT)	11.9	14688	1747.8720	24.6087	0.0018

Components (id)	**Content % (w/w) ***	**MW (g/mol)**	**%P*MW**	**% (mole)**	**mmol/gDCW***
Minor teichoic acid (acetylgalactosamine glucose phosphate, n=30) (tcam_MT)	19.3	13869.6	2676.8328	37.6877	0.0031
Sum	**100.0**		**7102.7**	**100.0**	

*gDCW = gram dry cell weight

Table S9. Ions and metabolites of a *Moorella thermoacetica* cell[1]

Components (id)	**Content % (w/w) ***	**MW (g/mol)**	**%P*MW**	**% (by weight)**	**mmol/gDCW***
K	86	39.1	33.626	5.3396	1.0865
Mg	7.7	24.3	1.8711	0.2971	0.1565
Fe(+3)	0.6	55.9	0.3354	0.0533	0.0053
Ca	0.4	40.1	0.1604	0.0255	0.0049
Phosphate	4.3	96	4.128	0.6555	0.0221
Diphospate	0.5	174.9	0.8745	0.1389	0.0014
Menaquinol 7	1	651	6.51	1.0337	0.0008
10-Formyltetrahydrofolate	1	471.4	4.714	0.7485	0.0010
NAD	61.9	662.4	410.0256	65.1089	0.0462
AMP	9.3	345.2	32.1036	5.0978	0.0133
ATP	8.7	503.2	43.7784	6.9517	0.0085

Components (id)	Content % (w/w) *	MW (g/mol)	%P*MW	% (by weight)	mmol/gDCW[*]
ADP	6.3	424.2	26.7246	4.2437	0.0073
CMP	1.9	321.2	6.1028	0.9691	0.0029
NADP	4	740.4	29.616	4.7028	0.0027
CTP	1.5	479.1	7.1865	1.1412	0.0015
GMP	1.1	361.2	3.9732	0.6309	0.0015
GTP	1.3	519.1	6.7483	1.0716	0.0012
CDP	0.6	400.2	2.4012	0.3813	0.0007
NADPH	0.9	741.4	6.6726	1.0596	0.0006
GDP	0.5	440.2	2.201	0.3495	0.0006
Sum	**100.0**		**629.8**	**100.0**	

[*] gDCW = gram dry cell weight

Moorella thermoacetica biomass equation

55 h2o[c] + 55 atp[c] + 0.2795 glu-L[c] + 0.0006 nadph[c] + 0.1469 gln-L[c] + 0.0133 amp[c] + 0.0291 utp[c] + 0.0925 met-L[c] + 0.1807 ser-L[c] + 0.0006 gdp[c] + 0.0369 gtp[c] + 0.0015 gmp[c] + 0.187 asp-L[c] + 0.4288 ala-L[c] + 0.1785 lys-L[c] + 0.0029 cmp[c] + 0.0007 cdp[c] + 0.284 arg-L[c] + 0.0368 ctp[c] + 0.126 asn-L[c] + 0.0468 trp-L[c] + 0.1417 phe-L[c] + 0.1257 tyr-L[c] + 0.0439 cys-L[c] + 0.3541 gly[c] + 0.001 10fthf[c] + 0.4644 leu-L[c] + 0.0199 datp[c] + 0.0723 his-L[c] + 0.3163 val-L[c] + 0.2071 pro-L[c] + 0.2068 thr-L[c] + 0.015 dgtp[c] + 0.0199 dttp[c] + 0.25 ile-L[c] + 0.015 dctp[c] + 0.0034 fe3[c] + 0.7063 k[c] + 0.1017 mg2[c] + 0.0008 mqn8[c] + 0.001 pydx5p[c] + 0.001 ribflv[c] + 0.001 mlthf[c] + 0.001 5mthf[c] + 0.001 succoa[c] + 0.000279 accoa[c] + 0.000167 coa[c] + 0.001 adocbl[c] + 0.0462 nad[c] + 0.0027 nadp[c] + 0.0032 ca2[c] + 0.011 d12dg_MT[c] + 0.0084 m12dg_MT[c] + 0.0066 t12dg_MT[c] + 0.0177 pgly_MT[c] + 0.0005 cdlp_MT[c] + 0.0022 lysylpgly_MT[c] + 0.0562 psetha_MT[c] + 0.0007 lipo1_24_MT[c] + 0.0006 lipo2_24_MT[c] + 0.0018 lipo3_24_MT[c] + 0.0015 lipo4_24_MT[c] + 0.0036 gtca1_45_MT[c] + 0.0023 gtca2_45_MT[c] + 0.0018 gtca3_45_MT[c] + 0.0031 tcam_MT[c] + 0.1018 peptido_MT[c] --> 0.0009 ppi[c] + 55 pi[c] + 55 h[c] + 55 adp[c]

Calculation of NGAM and GAM parameters of *i*AI558

Non-growth associated (NGAM) and growth associated maintenance (GAM) parameters for *M. thermoacetica* were estimated using published data[4, 5] and equation[6-8], as well as model simulations using *i*AI558.

The non-growth associated ATP maintenance is given by

$$m = \frac{b}{Y_G}$$

where, b = specific maintenance rate or decay rate (day^{-1})

Y_G = True growth yield or yield without maintenance (g DCW/electron equivalence, eeq)

Assuming Y_G = Y= Observed growth yield for *M. thermoacetica*,

$$m = \frac{b}{Y}$$

The reported growth yield of *M. thermoacetica* on glucose[5] is 40 gDCW/ mol glucose or 40/24 = 1.67 gDCW/eeq glucose because 24 electrons[9] are released from the oxidation of 1 mole of glucose and the typical decay rate for an anerobe[9] is 0.02 day^{-1}. Also, it is assumed that 4 protons (H^+) are translocated for 1 mole of ATP production by the ATP synthase reaction in the model. Using these data in the above equation, we obtained the NGAM for *M. thermoacetica*,

$$m = \frac{0.02 \times 1 \times 1000}{1.67 \times 24 \times 4}$$

$$= 0.12 \text{ mmol ATP/g DCW.h}$$

References

1. Y. K. Oh, B. Palsson, S. Park, C. Schilling and R. Mahadevan, *J. Biol. Chem.*, 2007, 282, 28791-28799.
2. E. Pierce, G. Xie, R. Barabote, E. Saunders, C. Han, J. Detter, P. Richardson, T. Brettin, A. Das, L. Ljungdahl and S. Ragsdale, *Environ. Microbiol.*, 2008, 10, 2550-2573.
3. K. Yamamoto, R. Murakami and Y. Takamura, *FEMS (Fed Eur Microbiol Soc) Microbiol Lett*, 1998, 161, 351-358.
4. S. L. Daniel, T. Hsu, S. I. Dean and H. L. Drake, *J. Bacteriol.*, 1990, 172, 4464-4471.
5. H. Huang, S. Wang, J. Moll and R. K. Thauer, *J. Bacteriol.*, 2012, 194, 3689-3699.
6. J. B. Russell and G. M. Cook, *Microbiol Rev*, 1995, 59, 48-62.
7. S. J. Pirt, *Proc R Soc Biol Sci Ser B*, 1965, 163, 224-231.
8. S. J. Pirt, *Arch. Microbiol.*, 1982, 133, 300-302.
9. B. E. Rittmann and P. McCarty, *Environmental biotechnology: principles and applications*, McGraw-Hill, New York, 2001.