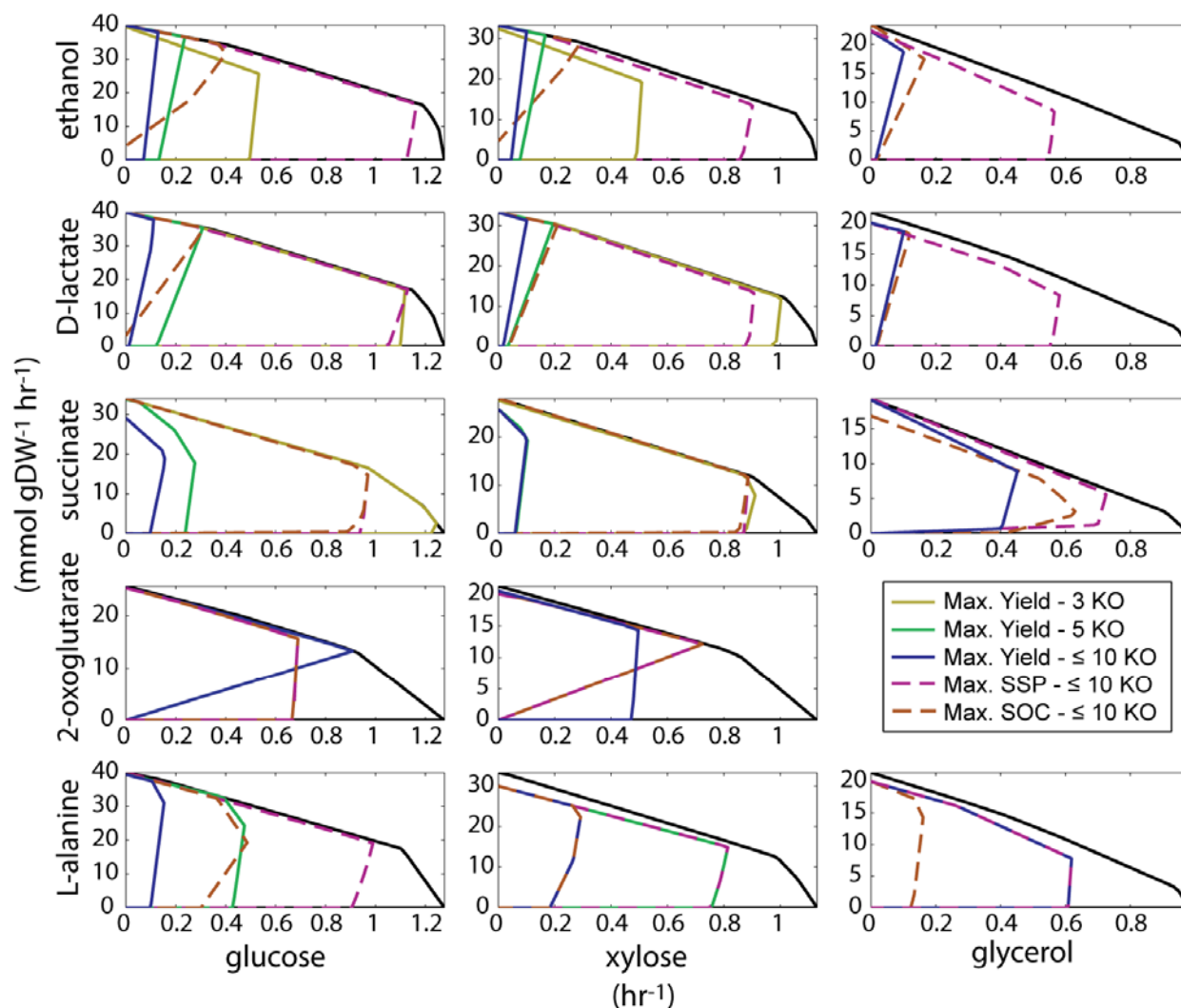**Supplementary Text**

List of Supplementary Data Files

File 1: All of the designs identified in the OptKnock analysis.

Supplementary Figure 1: The strain designs generated for five different targets from glucose, xylose, and glycerol aerobically

Supplementary Figure 1: A set of graphs that give the production envelopes for different substrate / target pairs that were calculated during the analysis under anaerobic conditions. The different target production rates (mmol gDW^{-1} hr^{-1}) are shown on the y-axis and the growth rate (hr^{-1}) is given on the x-axis. Shown on each plot (if a solution exists) are the maximum yields for 3 knockouts (yellow, solid line), 5 knockouts (green, solid line), up to 10 knockouts (with a 99.99% deletion penalty, blue, solid line), the maximum substrate-specific productivity (SSP, pink, dashed line), and the maximum strength of growth coupling (SOC, orange, dashed line) design. For example, there are no valid solutions for 2-oxoglutarate production on glycerol given the minimum growth rate of 0.1 hr^{-1}. Some of the solutions are identical for multiple objectives.

Supplementary Figure 2: Theoretical maximum production achievable for different substrate / target pairs for under aerobic conditions
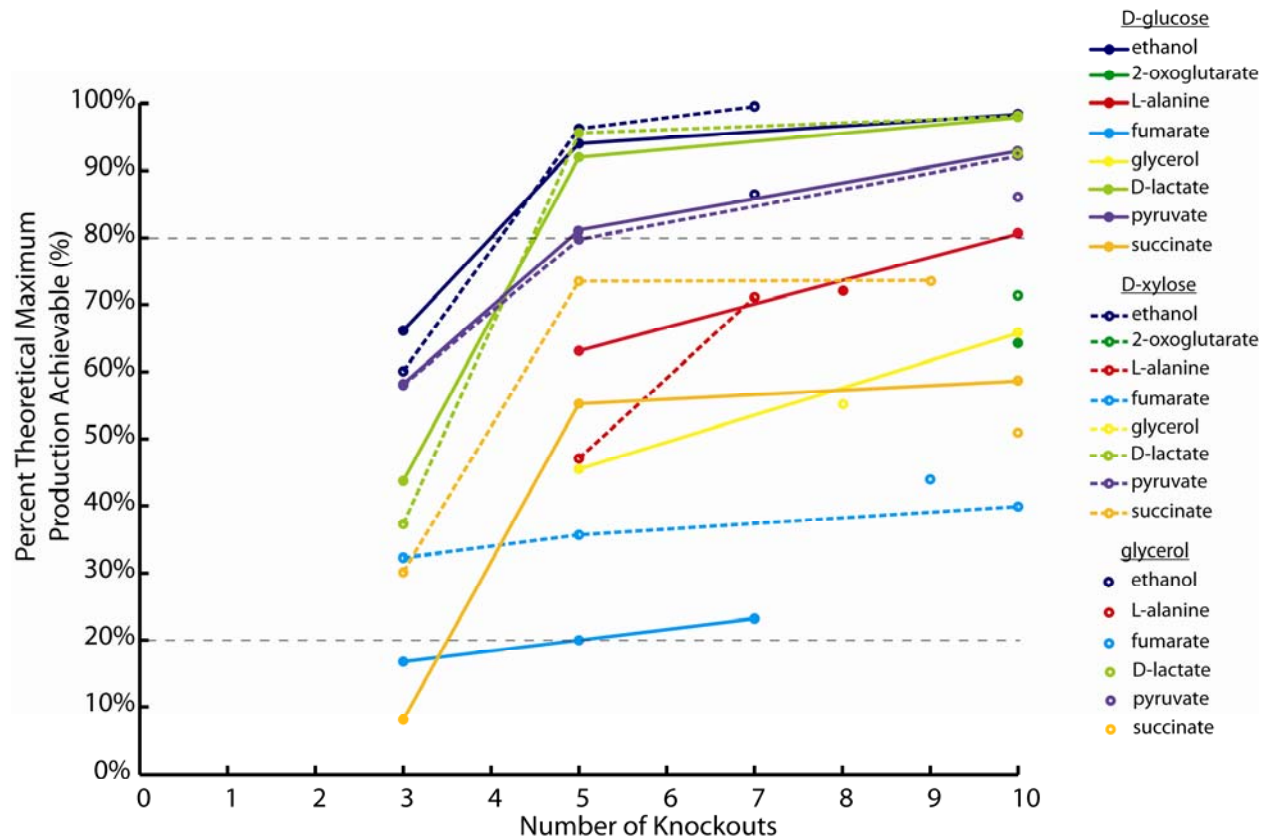
Supplementary Table 1: Changes from iAF1260 to make iAF1260b

Reaction added (abbreviation)	Reason for addition
DHORTfum	Added to no longer make FRD essential
MALt3pp	Means of excretion
ALAt2rpp	Means of excretion
GLYt2rpp	Means of excretion
CITt3pp	Means of excretion
ASPt2rpp	Means of excretion

Supplemental Table 2: Theoretical Yields - Molar Yields

product	no. of carbons	Glucose Anaerobic molar yield	Xylose Anaerobic molar yield	Glycerol Anaerobic molar yield	Glucose Aerobic molar yield	Xylose Aerobic molar yield	Glycerol Aerobic molar yield
Ethanol	2	193%	159%	98%	193%	159%	108%
D-Lactate	3	193%	159%	13%	193%	159%	101%
Glycerol	3	73%	45%		148%	121%	
L-Alanine	3	193%	127%	13%	193%	157%	99%
L-Serine	3	81%	50%	5%	198%	163%	102%
Pyruvate	3	147%	104%	7%	208%	171%	106%
Fumarate	4	86%	52%	4%	174%	142%	94%
L-Malate	4	86%	52%	4%	174%	142%	94%
Succinate	4	144%	105%	10%	161%	131%	88%
2-Oxoglutarate	5	50%	34%	2%	122%	100%	64%
L-Glutamate	5	55%	37%	2%	113%	93%	61%

Supplementary Table 3: Substrate conditions

Carbon Substrate(s)	Aerobicity	Wild Type Growth Rate	Target Reactions After Reduction of Scope
Glucose	Anaerobic	0.4589	142
Xylose	Anaerobic	0.3194	141
Glycerol	Anaerobic	0.1192	140
Glucose	Aerobic	1.2759	170
Xylose	Aerobic	1.1305	165
Glycerol	Aerobic	0.9828	166
Glucose / Acetate	Anaerobic	0.4589	142
Glucose / Acetate	Aerobic	1.2759	170
Glucose / AAs	Anaerobic	1.0313	162
Glucose / AAs	Aerobic	2.0996	189

Maximum uptake rates for primary carbon sources and acetate were set to 20 mmol/gDW^{-1} hr^{-1}. Maximum amino acid (AA) uptake rates were set to 5 mmol/gDW^{-1} hr^{-1}. In aerobic simulations, maximum oxygen uptake rate was set to 20 mmol/gDW^{-1} hr^{-1}.

Methods

Description of OptGene procedure

OptGene was implemented as previously described [ref], with the following modifications, i) the genotype of the population was changed from genes to reactions. This was done to reduce the number of total targets needed as the number of gene associated with the set of target reactions was larger than the number of reactions (effectively decreasing search time) and to make optGene input have the same format as optKnock, ii.) the mutation function was modified as described below, iii.) some parameters were modified as described below, this was performed in a conservative fashion to increase the chance of finding an optimal solution, and iv) solutions from OptKnock which had a positive production were used as an initial population for OptGene simulation, in order to increase the efficiency of the algorithm by inputting reasonably good initial guesses for further refinement. OptGene was implemented in the Matlab software framework using the Genetic Algorithms and Directed Search toolbox.

Specific Change: *Mutation Function*

It was noted that having an independent mutation function where at each generation, each reaction is toggled independently with probability 1/nrxns will tend to increase the number of mutations until half the reactions are knocked out. As this is undesirable, a small modification was added where with 50% knockouts would be randomly removed until the child genome has fewer KO's than the parent. This heuristic makes the probability of increasing the number of KOs roughly equal to the probability of removing one.

Specific Change: Hashing of genotypes

A speed improvement made was the hashing of previously explored genotypes. All genotypes for which a fitness has been computed are put into a hash table (the key being a concatenation of the genotype). When a new genotype needs to be evaluated, the hash table is queried first to see if it has been previously computed. This results in a significant speedup in performance without affecting the results of the algorithm. The hashtable is periodically purged to prevent memory overflow. Note that this improvement has no effect on the outcome of the algorithm, only the speed.

Supplementary Table 4: Parameters and values

Name	Value	Value in [ref]	Description	Comments
Simulation Parameters				
Population size	4x125	125	Multiple populations were used. The total number of individuals was increased by a factor of 4	Increasing the number of individuals increases the chance of finding the optimum.
MigrationDirection	Forward	n/a	Migration occurs between sub-populations. Individuals only move forward from one population to the next	
MigrationFraction	.1	n/a	The probability that any individual migrates to the next population at a migration event.	1/10th of the population migrates.
MigrationInterval	100	n/a	How often migration intervals occur. In this case	100 is sufficiently high that for the most part the sub-populations are independent
MaxKO	10	varries	Maximum number of KO's allowed.	
Mutation rate	1/nrxns *	1/nrxns	The mutation rate indicates the probability that a child will have a different genome as its parent. The value 1/nrxns indicates that on average there will be one mutation per offspring (although there could be zero or more)	This was modified slightly to not accumulate mutations (see below). We looked at the distribution of mutations in the population and with this change there are more individuals with fewer mutations.
Cross over fraction	.8	1	This value represents the fraction of offspring which are produced by crossing over between two parents (as opposed to mutations from a single parent)	20% of individuals are now generated by a simple mutation process.
Cross over mutation rate	.2*mutation rate	1/nrxns	After crossing over, there is an additional probability of a mutation.	

Cross over strategy	uniform	uniform or 1 point or 2 point	Crossovers are performed on a gene by gene basis. The offspring of two parents has an equal probability of receiving any given gene from either parent.	
Selection strategy	roulette	roulette	Selection for the next generation is based on an individual's fitness. Greater fitness gives a higher probability.	
Elite count	2	n/a	Number of top performing genotypes which automatically to pass to the next generation.	This is to keep track of the global best scores.
Terminating Conditions				
MaxGenerations	10000	5000	The maximum number of generations performed.	
StallGenLimit	10000	n/a	The maximum number of generations between finding an improved score.	
Time Limit	2 days	n/a	The maximum time limit	This limit is not reached.
StallTimeLimit	1 day	n/a	The maximum time between finding a score improvement	This limit is not reached.