Supplementary Information

Self-destructive altruism in a synthetic developmental program enables complex feedstock utilization

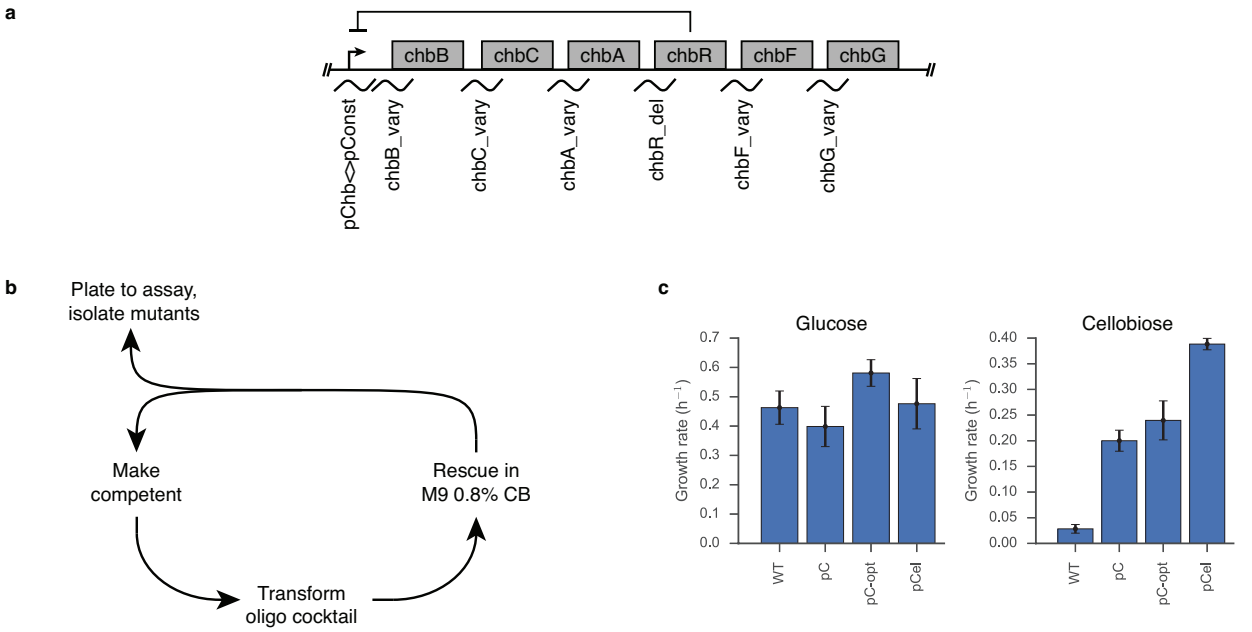
Robert G. Egbert[1], Leandra M. Brettner[2], David Zong[2], Eric Klavins[1]

1. Electrical Engineering Department, University of Washington, Seattle, WA, USA.
2. Bioengineering Department, University of Washington, Seattle, WA, USA.

Table of Contents:

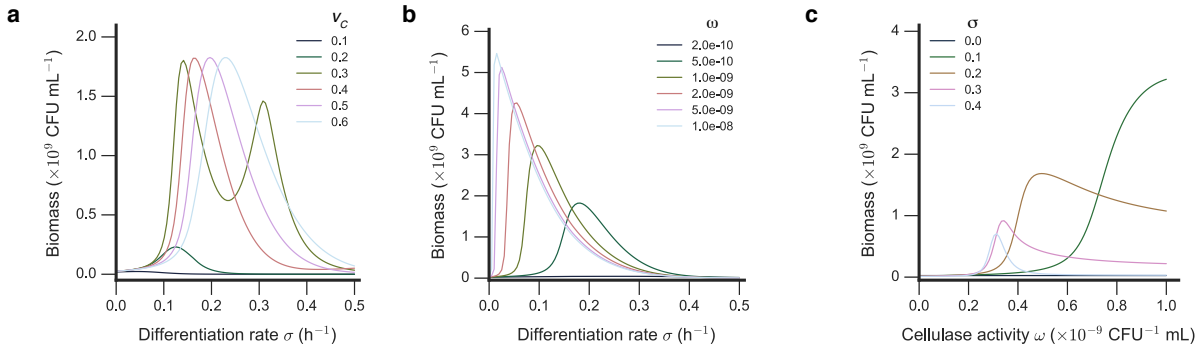
Supplementary Figure 1. Engineering cellobiose utilization for SDAc strains **3**
Supplementary Figure 2. Parameter sweeps for cellulose utilization model **4**
Supplementary Figure 3. Data collection for growth, differentiation and lysis rates **5**
Supplementary Figure 4. Mapping differentiation rate to differentiation delay **6**
Supplementary Figure 5. Altruist fraction fits combining growth, differentiation and lysis parameters **7**
Supplementary Figure 6. Mapping *colE3L* (AT) rbSSR length to lysis rate **8**
Supplementary Figure 7. Mapping cellulose degradation clearings to differentiation and lysis rates **9**
Supplementary Figure 8. Carbon to biomass yield estimates **10**
Supplementary Figure 9. Cellulose growth dynamics predictions for SDAc strains **11**
Supplementary Figure 10. Model fits to altruist fraction and overall growth for SDAc switch variants **12**
Supplementary Figure 11. Model fit of initial conditions to differentiation rate **13**
Supplementary Movie **14**
Supplementary Table 1. SDAc model parameters **15**
Supplementary Table 2. List of strains used **16**
Supplementary Table 3. Cellobiose growth rates for SDAc strains **17**
Supplementary Table 4. Differentiation rates for lysis-deficient SDAc switch variants **17**
Supplementary Table 5. Lysis rates for SDAc switch and lysis variants **17**
Supplementary Table 7. Cellulase activity measurements for cellulase production strains **18**
Supplementary Table 8. Half-maximal rate constants for growth and differentiation **18**
Supplementary Table 9. Cheater rates for expanded SDAc model **18**
Supplementary Table 10. Genotypes for differentiation cheaters **19**
Supplementary Table 11. Genotypes for lysis cheaters **20**
Supplementary Table 12. Oligonucleotides used to modify *chb* operon and for cheater sequencing **21**
Supplementary Note 1. Model assumptions for SDAc population dynamics **22**
Supplementary Note 2. Parameter bootstrapping for SDAc models **22**
Supplementary Note 3. Estimating growth rates from cytometry data **22**
Supplementary Note 4. Estimating differentiation rates from cytometry data **22**
Supplementary Note 5. Estimating lysis rates from cytometry data **23**
Supplementary Note 6. Estimating biomass yield for cellobiose and PASC **23**
Supplementary Note 7. Estimating cellulase activity rates **24**
Supplementary Note 8. Estimating half-maximal rate constants for growth and differentiation **25**
Supplementary Note 9. Initial conditions for nutrients and population size in full dynamics model **25**
Supplementary Note 10. Estimating SDAc cheater rates **25**
Supplementary Note 11. Modeling cheater dynamics **26**
Supplementary References **29**

Supplementary Figure 1. Engineering cellobiose utilization for SDAc strains



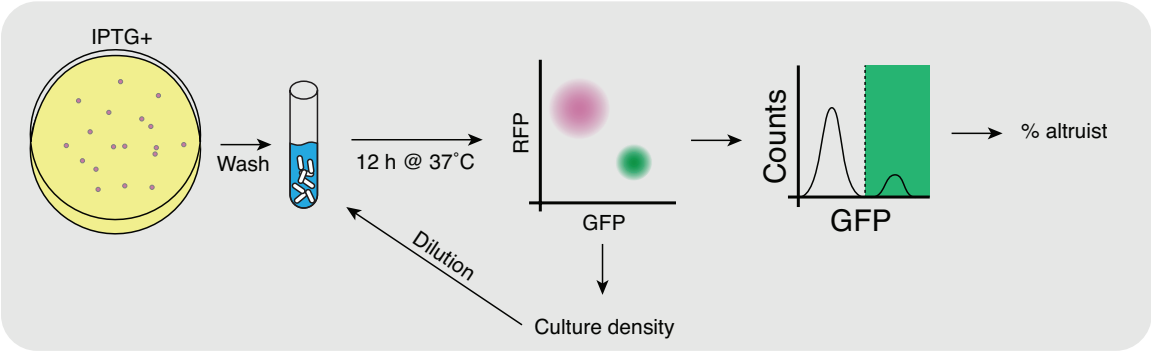
(a) Chitobiose (*chb*) operon with targeted sites for promoter replacement (pChb), overexpression (*chbA*, *chbB*, *chbC*, *chbF*, *chbG*) or deletion (*chbR*). See Supplementary Table 12 for oligonucleotide sequences. **(b)** Workflow to select for variants that grow faster on cellobiose. **(c)** Growth rates for a selection of variants: wild-type (WT, EcNR1 Δ*fim* Δ*lacIZ*), the native *chb* promoter replaced with a constitutive promoter (pC, RGE531), the fastest growing recombinant mutant (pC-opt, DL069) and a control strain (pCel, DL180) that expresses a membrane-bound beta-glucosidase from *Cellvibrio japonicas*[1]. Growth rates were measured at 32°C in a Biotek Synergy plate reader (n=3) using 200 µL cultures grown in 96-well microwell plates.

Supplementary Figure 2. Parameter sweeps for cellulose utilization model



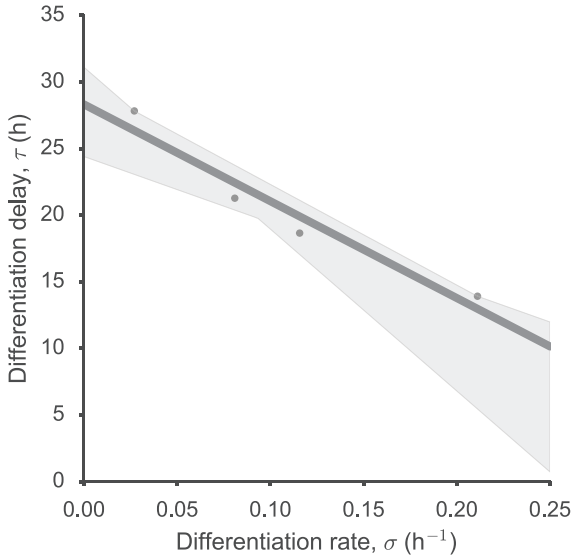
Initial parameter sweeps of core SDAc model parameters for predicted biomass yield to explore circuit behavior and guide system implementation. When not varied in a plot, nominal values for parameters are $v_C = v_A = 0.45, k = k_\sigma = 10^{-4}, \sigma = 0.116, \rho = 0.164, \gamma = 1.5{\times}10^{12}$ and $\omega = 5{\times}10^{-10}$. Initial cell populations were set to 10^7 from which the altruist fraction was set to $\frac{\sigma}{\rho}$ or 1 when $\sigma > \rho$ (see Supplementary Note 11). Initial nutrients were set to 10^{-5}. **(a)** Expected biomass yield as a function of differentiation rate for individual growth rates. **(b)** Expected biomass yield as a function of differentiation rate for individual cellulase activity levels. **(c)** Expected biomass yield as a function of cellulase activity for individual differentiation rates.

Supplementary Figure 3. Data collection for growth, differentiation and lysis rates



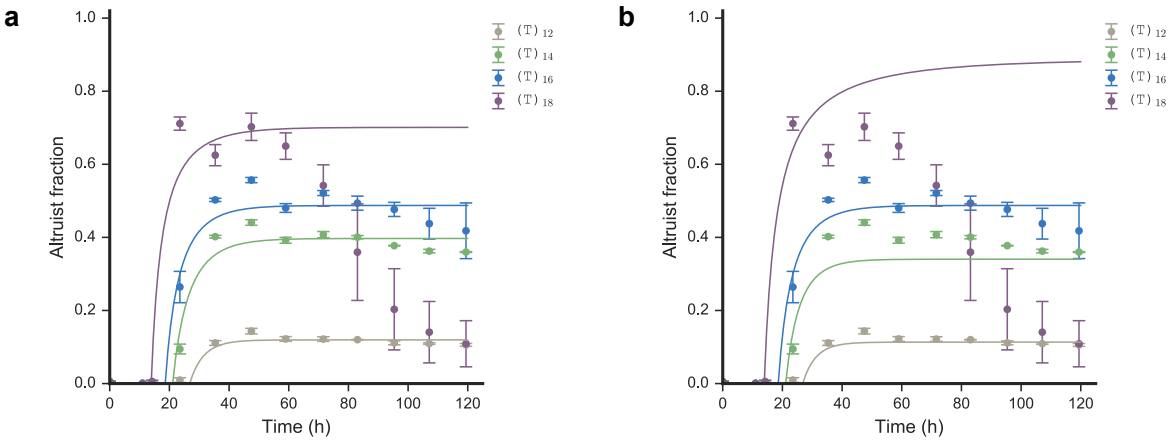
Experimental workflow to estimate growth, differentiation and lysis rates using low-density continuous culture and flow cytometry. Cultures were initialized in the consumer state by growing individual colonies on LB agar supplemented with IPTG. Colonies were washed and diluted into M9 minimal cellobiose media and periodically sampled in the flow cytometer to calculate the population fraction of differentiated cells and the dilution factor for the next growth passage.

Supplementary Figure 4. Mapping differentiation rate to differentiation delay



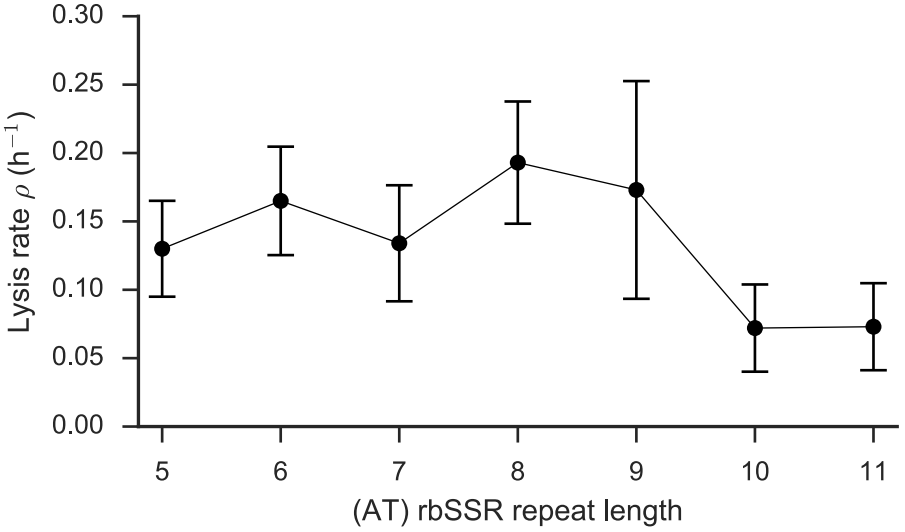
Relationship between the differentiation rate and the delay in the initiation of switching following the removal of IPTG inducer for strains DL046, DL112, DL110 and DL108. Linear fit line is $y = 28.3 - 72.7x\ h$. The shaded area represents the 95% confidence interval for the fit.

Supplementary Figure 5. Altruist fraction fits combining growth, differentiation and lysis parameters



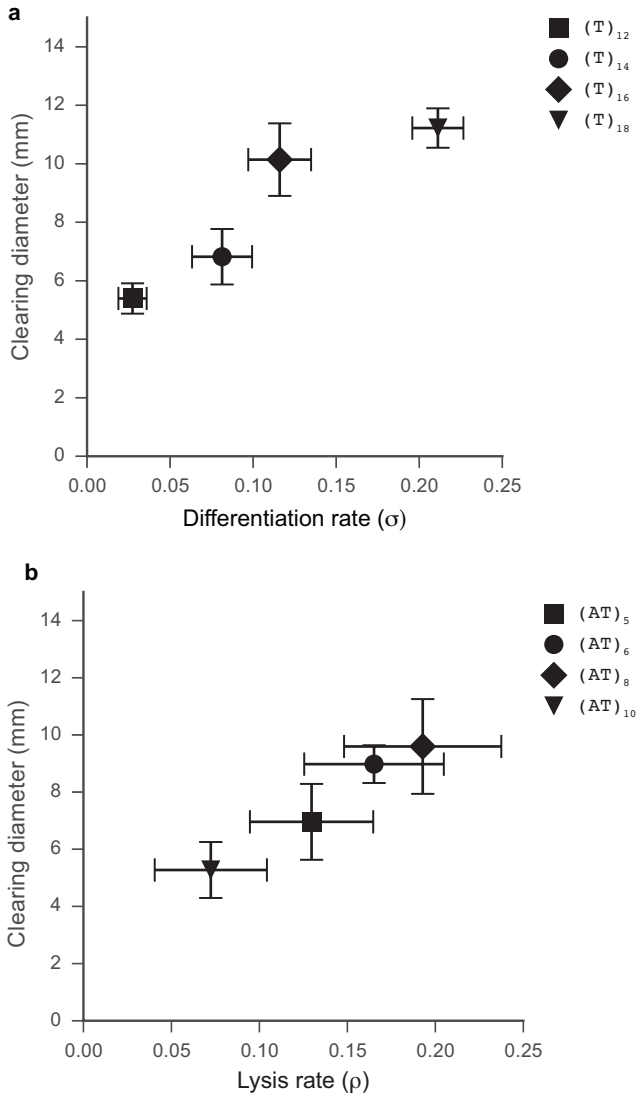
(a) Altruist fractions as a function of time for multiple differentiation variants. Altruist fractions were calculated from flow cytometry measurements of the continuous culture in cellobiose. Curves are derived from model predictions using growth rates of DL147, differentiation rates as estimated for the lysis-deficient switch variants and lysis rate fits for each strain. The lysis rate fit for $(T)_{18}$ was truncated at 65 h. Each strain uses rbSSR $(AT)_8$ for lysis gene expression. Data for labels $(T)_{12}$, $(T)_{14}$, $(T)_{16}$ and $(T)_{18}$ correspond to measurements for strains DL046, DL112, DL110 and DL108, respectively (n=3). **(b)** Fits as in **(a)**, but with a single lysis rate of $0.192\ h^{-1}$ measured from DL110.

Supplementary Figure 6. Mapping *colE3L* `(AT)` rbSSR length to lysis rate



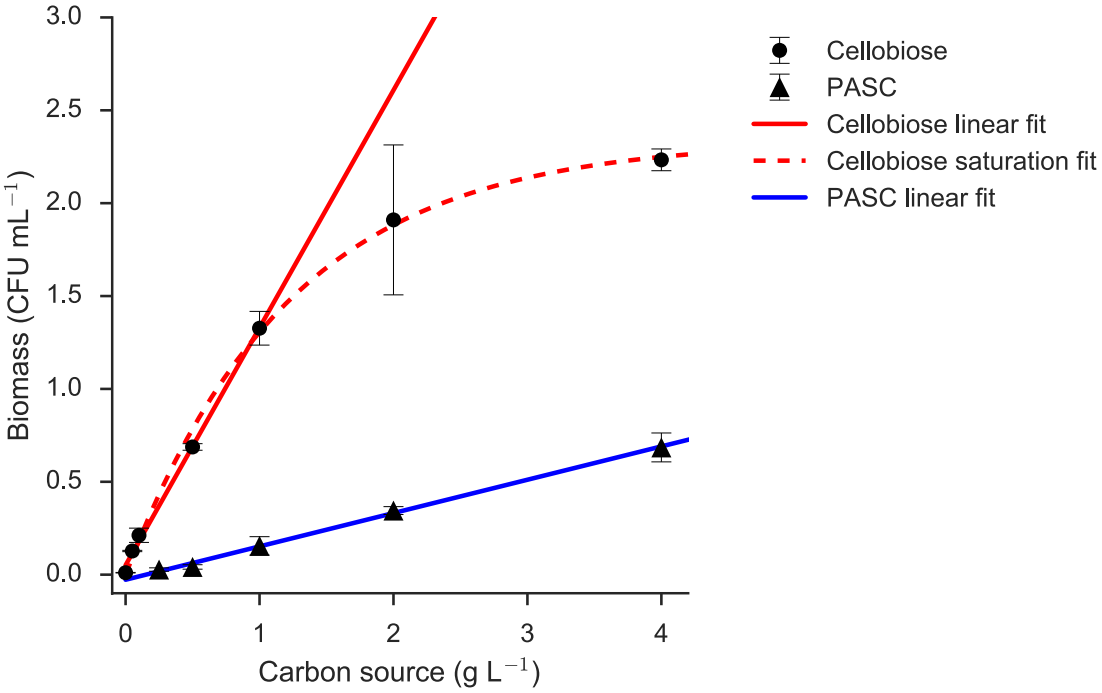
Lysis rate fits as a function of (AT) repeat length for multiple strains with switch $(T)_{16}$. Error bars represent parameter estimate uncertainty, inclusive of measurement error and parameter bootstrapping from growth rates and differentiation rates.

Supplementary Figure 7. Mapping cellulose degradation clearings to differentiation and lysis rates



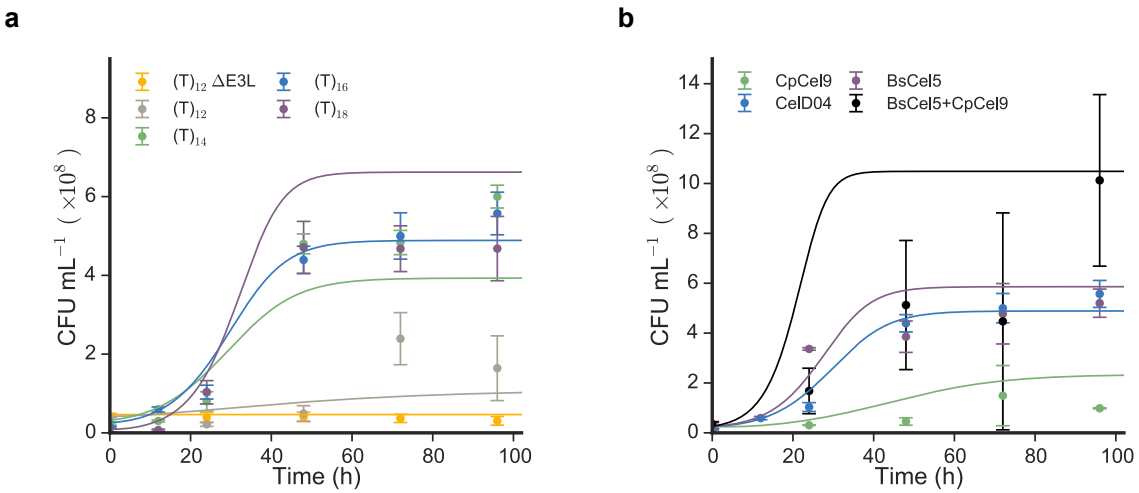
(a) Plot of cellulose degradation clearings as a function of differentiation rate for SDAc variants of the poly-(T) *tetR* rbSSR. **(b)** Plot of cellulose degradation clearings as a function of lysis rate for SDAc variants of the *colE3L* lysis gene rbSSR. For each panel the x-axis error bars represent uncertainty of experimental measurements and parameter bootstrapping and the y-axis error bars represent standard error from at least six replicate clearings of individual colonies (see Figure 3).

Supplementary Figure 8. Carbon to biomass yield estimates



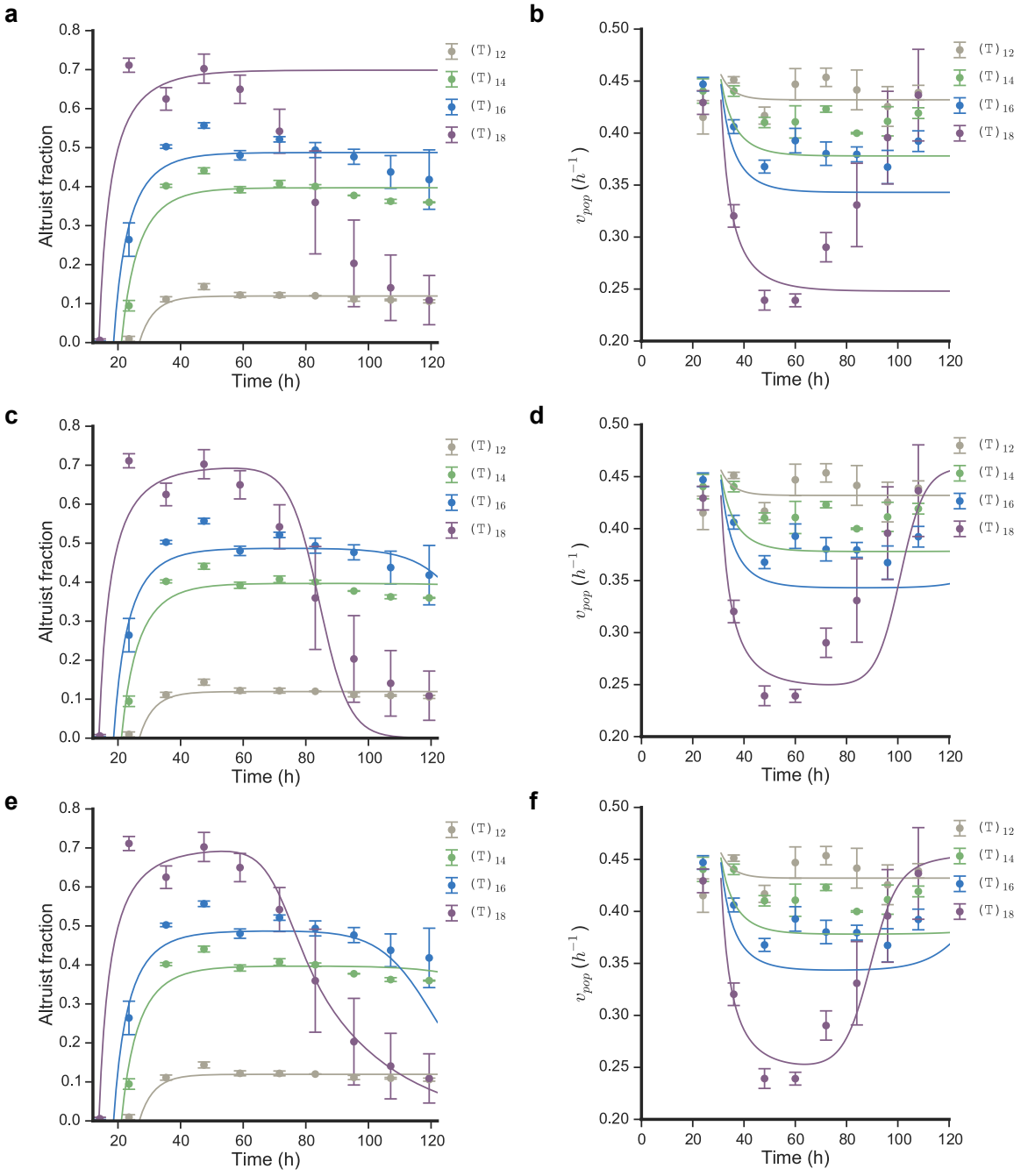
Carbon utilization measurements and fits for DL146 (cellobiose) and DL110 (PASC).

Supplementary Figure 9. Cellulose growth dynamics predictions for SDAc strains



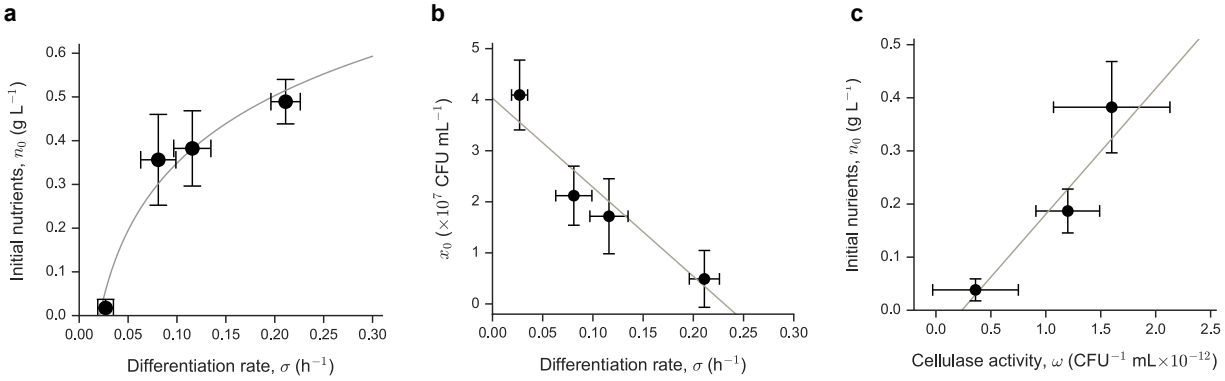
(a) Growth dynamics and model predictions for multiple differentiation variants and a lysis control. **(b)** Growth dynamics and model predictions for multiple cellulase variants using intermediate differentiator $(T)_{16}$.

Supplementary Figure 10. Model fits to altruist fraction and overall growth for SDAc switch variants



Model fits to altruist fraction dynamics from continuous growth in cellobiose for **(a)** no cheaters, **(c)** differentiation escapees and **(e)** both differentiation and lysis escapees. Model fits to growth rate dynamics for **(b)** no cheaters, **(d)** differentiation escapees and **(f)** both differentiation and lysis escapees.

Supplementary Figure 11. Model fit of initial conditions to differentiation rate



(a) Logarithmic fit of differentiation rate to initial nutrients generated in PASC media, measured in growth of cellulase and lysis deficient DL146 using supernatant from toggle switch variants. **(b)** Linear fit of differentiation rate to initial viable population size at initiation of PASC growth assay, measured by CFU counts. **(c)** Linear fit of cellulase activity to initial nutrients generated in PASC media, measured as in **(a)**.

Supplementary Movie

Movie S1 shows the growth and differentiation dynamics of individual microcolonies derived from individual cells of the (T)10/(A)12 rbSSR-BSS switch variant with an inactive cellulase gene and colicin E3 lysis gene expression controlled by an $(AC)_{11}$ rbSSR, transformed in strain 2.320. The movie demonstrates the stochastic differentiation and delayed lysis dynamics underlying SDAc behavior. The movie scale bar is 5 μm.

Supplementary Table 1. SDAc model parameters

Parameter	Description	Units
v_C, v_A	Consumer, altruist growth rates	h^{-1}
k_C, k_A	Half maximal growth constants	g
σ	Differentiation rate	h^{-1}
k_σ	Half maximal differentiation constant	g
ρ	Lysis rate	h^{-1}
ω	Cellulase activity	CFU^{-1} mL
γ	Nutrient to biomass conversion efficiency	CFU g^{-1}

Supplementary Table 2. List of strains used

Strain ID	Switch variant*	Payload	colE3 lysis repeat	Parent strain	Modifications
CT009	N/A	N/A	N/A	EcNR1	Δ*fimAICDFGH* Δ*ampR* Δ*lacIZ*
RGE531	N/A	N/A	N/A	CT009	pChb<>pConst,
DL046	$(T)_{12}/(A)_{10}$	CelD04	$(AT)_8$	DL069	
DL069	N/A	N/A	N/A	CT009	pChb<>pConst, Δ*chbR*, Δ(λ-Red)
DL108	$(T)_{18}/(A)_{10}$	CelD04	$(AT)_8$	DL069	
DL110	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_8$	DL069	
DL112	$(T)_{14}/(A)_{10}$	CelD04	$(AT)_8$	DL069	
DL146	$(T)_{12}/(A)_{10}$	None	No lysis gene	DL069	
DL147	$(T)_{12}/(A)_{10}$	CelD04	No lysis gene	DL069	
DL180	N/A	CelD04	No lysis gene	CT009	Δ(λ-Red), pCellulose[1] (ampR), mRFP1 (kanR)
DL182	$(T)_{16}/(A)_{10}$	None	No lysis gene	DL069	
DL183	$(T)_{14}/(A)_{10}$	None	No lysis gene	DL069	
DL188	$(T)_{18}/(A)_{10}$	None	No lysis gene	DL069	
DL202	$(T)_{16}/(A)_{10}$	BsCel5	$(AT)_7$	DL069	
DL261	$(T)_{16}/(A)_{10}$	CpCel9	$(AT)_7$	DL069	
DL268	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_5$	DL069	
DL269	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_6$	DL069	
DL270	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_7$	DL069	
DL271	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_9$	DL069	
DL272	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_{10}$	DL069	
DL273	$(T)_{16}/(A)_{10}$	CelD04	$(AT)_{11}$	DL069	
DL292	$(T)_{12}/(A)_{10}$	CpCel9	No lysis gene	DL069	
DL294	$(T)_{12}/(A)_{10}$	BsCel5, CpCel9	No lysis gene	DL069	
DL296	$(T)_{12}/(A)_{10}$	CelD04, CpCel9	No lysis gene	DL069	
DL307	$(T)_{12}/(A)_{10}$	BsCel5	No lysis gene	DL069	
DL322	$(T)_{16}/(A)_{10}$	BsCel5, CpCel9	$(AT)_9$	DL069	
DL329	$(T)_{12}/(A)_{12}$	CelD04	colE3	DL069	
DL327	$(T)_{16}/(A)_{10}$	CelD04, CpCel9	$(AT)_9$	DL069	

* $(T)_x/(A)_y$ indicates *x* poly-T rbSSR repeats driving TetR expression and *y* poly-A repeats driving LacI expression for the differentiation plasmid.

Supplementary Table 3. Cellobiose growth rates for SDAc strains

Strain ID	Switch variant	Cellulase	Lysis	Induction state*	Growth rate (hr^{-1})	StDev
DL146	$(T)_{12}/(A)_{10}$	None	None	Consumer	0.463	0.019
DL146	$(T)_{12}/(A)_{10}$	None	None	Altruist	0.391	0.019
DL147	$(T)_{12}/(A)_{10}$	CelD04	None	Consumer	0.459	0.007
DL147	$(T)_{12}/(A)_{10}$	CelD04	None	Altruist	0.413	0.023
DL046	$(T)_{12}/(A)_{10}$	CelD04	colE3	Consumer	0.438	0.018
DL112	$(T)_{14}/(A)_{10}$	CelD04	colE3	Consumer	0.420	0.017
DL183	$(T)_{14}/(A)_{10}$	None	None	Consumer	0.430	0.012
DL110	$(T)_{16}/(A)_{10}$	CelD04	colE3	Consumer	0.389	0.018
DL182	$(T)_{16}/(A)_{10}$	None	None	Consumer	0.423	0.019
DL108	$(T)_{18}/(A)_{10}$	CelD04	colE3	Consumer	0.266	0.041
DL188	$(T)_{18}/(A)_{10}$	None	None	Consumer	0.395	0.003

* Induction to consumer state implies IPTG induction (1 mM). Induction to altruist state implies aTc induction (100 ng/mL).

Supplementary Table 4. Differentiation rates for lysis-deficient SDAc switch variants

Strain ID	Switch variant	σ (h^{-1})	StDev	τ (h)	StDev
DL146	$(T)_{12}/(A)_{10}$	0.027	0.009	26.8	4.56
DL183	$(T)_{14}/(A)_{10}$	0.081	0.018	21.1	0.97
DL182	$(T)_{16}/(A)_{10}$	0.116	0.019	18.6	0.78
DL188	$(T)_{18}/(A)_{10}$	0.211	0.015	13.9	0.03

Supplementary Table 5. Lysis rates for SDAc switch and lysis variants

Strain ID	Switch variant	Lysis variant	σ (h^{-1})	ρ (h^{-1})	StDev
DL046	$(T)_{12}/(A)_{10}$	$(AT)_8$	0.027	0.180	0.080
DL112	$(T)_{14}/(A)_{10}$	$(AT)_8$	0.081	0.159	0.052
DL108	$(T)_{18}/(A)_{10}$	$(AT)_8$	0.211	0.256	0.036
DL268	$(T)_{16}/(A)_{10}$	$(AT)_5$	0.116	0.130	0.035
DL269	$(T)_{16}/(A)_{10}$	$(AT)_6$	0.116	0.165	0.040
DL270	$(T)_{16}/(A)_{10}$	$(AT)_7$	0.116	0.134	0.042
DL110	$(T)_{16}/(A)_{10}$	$(AT)_8$	0.116	0.192	0.048
DL271	$(T)_{16}/(A)_{10}$	$(AT)_9$	0.116	0.173	0.080
DL272	$(T)_{16}/(A)_{10}$	$(AT)_{10}$	0.116	0.072	0.032
DL273	$(T)_{16}/(A)_{10}$	$(AT)_{11}$	0.116	0.073	0.032

Supplementary Table 6. Carbon to biomass yields for cellobiose and PASC

Media	Linear range	γ (CFU g^{-1})	StDev	R^2
Cellobiose	0 – 1 g L^{-1}	1.279e+12	3.344e+10	0.9919
PASC	0 – 4 g L^{-1}	1.795e+11	8.475e+09	0.9825

Supplementary Table 7. Cellulase activity measurements for cellulase production strains

Strain ID	Cellulase	ω (hydrolysis)	StDev	ω (biomass)	StDev
DL292	CpCel9	4.6e-12	2.4e-13	3.6e-13	3.9e-13
DL147	CelD04	5.7e-12	6.8e-13	1.6e-12	5.3e-13
DL294	CelD04 + CpCel9	6.8e-12	9.5e-13	1.8e-12	5.6e-13
DL307	BsCel5	6.1e-12	9.8e-13	1.2e-12	2.9e-13
DL294	BsCel5 + CpCel9	7.4e-12	4.3e-13	1.9e-12	3.0e-13

Note: units for ω are CFU^{-1} mL

Supplementary Table 8. Half-maximal rate constants for growth and differentiation

Parameter	Fit	StDev
k_C, k_A	0.00112	0.00045
k_σ	0.09999	0.00019

Supplementary Table 9. Cheater rates for expanded SDAc model

Cheater type(s)	χ_C (h^{-1})	Std Error	χ_A (h^{-1})	Std Error
Differentiation	2.473e-07	5.857e-08	n/a	n/a
Differentiation and lysis	1.775e-06	6.279e-07	1.908e-05	8.207e-06

Supplementary Table 10. Genotypes for differentiation cheaters

Plate Replicate	Isolation condition	Promoter mutation	*lacI* mutation
2	LB	None	None
2	LB kan/amp	None	$(CTGG)_2$
2	LB kan/amp	None	None
3	LB kan/amp	None	IS2 insertion at base 15 of *lacI*
3	LB kan/amp	None	IS2 insertion at base 15 of *lacI*
3	LB kan/amp	None	IS2 insertion at base 34 of *lacI*
4	LB	None	$(CTTG)_4$
4	LB	None	148 bp *lacI* deletion (A351)
4	LB kan/amp	None	None
4	LB kan/amp	None	$(CTTG)_4$
5	LB kan/amp	None	$(CTTG)_4$
5	LB kan/amp	None	$(CTGG)_2$
5	LB kan/amp	None	$(CTTG)_4$
5	LB kan/amp	None	20 bp *lacI* deletion (A303)
6	LB kan/amp	$\Delta(T)_5$	None
6	LB kan/amp	None	$(CTGG)_2$
6	LB kan/amp	$\Delta(T)_5$	None
6	LB kan/amp	$\Delta(T)_5$	None

Sequencing results of differentiation plasmid mutations from monochromatic RFP-positive $(T)_{18}/(A)_{10}$ escape colonies (six replicate plates). The first replicate plate had no mutations at the sequenced loci for the differentiation plasmid.

Supplementary Table 11. Genotypes for lysis cheaters

Plate Replicate	Isolation condition	Payload plasmid mutations
1	LB	None
1	LB Kan/Amp	None
2	LB	None
2	LB	None
2	LB	None
2	LB Kan/Amp	None
2	LB Kan/Amp	None
2	LB Kan/Amp	None
2	LB Kan/Amp	None
3	LB	None
4	LB	IS1 insertion at A514 in celD04
4	LB	IS1 insertion at A514 in celD04
4	LB Kan/Amp	IS1 insertion at A514 in celD04
4	LB Kan/Amp	IS1 insertion at A514 in celD04
5	LB	None
6	LB	None
6	LB	None
6	LB Kan/Amp	None

Sequencing results of payload plasmid mutations from monochromatic GFP-positive $(T)_{18}/(A)_{10}$ escape colonies (six replicate plates).

Supplementary Table 12. Oligonucleotides used to modify *chb* operon and for cheater sequencing

Oligo name	Sequence
pChb<>pConst	CTTCCATGCTCTGGGTAACTTGCGAAACCAACATGATGAATCTATTATAAAAAAAAAAAAAAAAA AAGTCAATACTCTATCGAACTCAGGCCAAAAAAAACCGGCGCAATGGCCGGTTTC
chbR_del	ggacataaataatccagcaacaggacagatatgtgaattgtcaggtataacgacttactgcatcg actccttatgccttcagttttcatgaagctcaattaattcagtaatcagttcac
chbA_A15	agctcttcagcttccgtttgcgtatcgggaatgttatcgagatccatcatTTTTTTTTTTTTTTT CCTCCTCttttcttaccggcacgattacccgtaccggcatcgattaaaatttcag
chbC_T15	actgcaaaagggaggagtacctttcaagcgatgcaataacattactcatAAAAAAAAAAAAAAA CCTCCTaaaaaccgcaatttaaatattgcggtattgatttatgaaataactcttt
chbB_A15	gaggtagacatgcccgcagaacaaaacagataaatgtgtttcttttccatTTTTTTTTTTTTTTT ACCTCCTGATatcgacgattatctgtcagccagacactccgcaagccttaacctg
chbF_A5	gtatagctgctcccgccaccaatagtgacgacttttaatttctggctcatTTTTTACCTCCTagt acagaatactgatatctggcatatctgccccccggacataaata
chbF_A15	gtatagctgctcccgccaccaatagtgacgacttttaatttctggctcatTTTTTTTTTTTTTTT ACCTCCTagtacagaatactgatatctggcatatctgccccccggacataaata
chbG_T15	cctttgcttaagccaaaatcatcggcattaacaatcagtaagcgttccatAAAAAAAAAAAAAAA CCTCCttaatgtgctttttaagctctgcgatgcagtcggcaaagtttggcagcc
BSSregSeq_2809	GGCTGCTCTACACCTAGCTTCTGG
lacIrptSeq_2810	GGTTTGTTGAAAACCGGACATGG
Vf2	TGCCACCTGACGTCTAAGAA
Vr	ATTACCGCCTTTGAGTGAGC

Supplementary Note 1. Model assumptions for SDAc population dynamics

The model for synthetic self-destructive altruism dynamics relies on multiple assumptions to minimize the number of parameters. First, intracellular enzyme accumulation is considered independent of the lysis rate, resulting in a fixed payload size per altruist. Second, cellulase dynamics were removed from the system of equations based on an assumption that cellulase enzymatic activity occurs much faster than its degradation, allowing parameters to be consolidated. Cellulase activity ω is normalized to the altruist population through the product ρA. Third, while we assume nutrient-dependent rates for growth and switching, altruist commitment to lysis is considered to be nutrient independent.

Supplementary Note 2. Parameter bootstrapping for SDAc models

Bootstrapping based on Monte Carlo simulations was used to generate many parameter estimates for SDAc models. Uncertainty in the parameter estimates depend both on experimental measurement error and independently measured experimental parameters. Parameter estimates should reflect both the fit to the experimental data and uncertainty propagated from parameters used to constrain the fit. Thousands of parameter fits were generated to propagate the uncertainty measures, randomly sampling from a normal distribution of the existing parameters for each fit[2]. We calculated the mean and standard deviation of the ensemble of fits to generate the parameter estimate.

Supplementary Note 3. Estimating growth rates from cytometry data

As described in the methods section, the growth rates for consumer cells and altruist cells were estimated by applying an exponential fit to cell densities measured via cytometry at 12 hour intervals in periodically diluted samples. The growth rate for each strain over three passage intervals (36, 48, 60 h) was measured in triplicate to obtain mean and standard error measurements for each strain (Supplementary Table 2).

For each strain, growth rate $v = \alpha/\ln(2)$ was estimated from time *t1* to time *t2* using equations

$$P_{t2} = P_{t1}\left(\frac{V_{tx}}{V_{tx}+V_w}\right)e^{\alpha(t2-t1)}, \tag{S1}$$

$$\alpha = \ln\left(\frac{P_{t2}}{P_{t1}}\left(\frac{V_{tx}+V_w}{V_{tx}}\right)\right)(t_2-t_1)^{-1}, \tag{S2}$$

for which P_{t1} and P_{t2} are the cell densities measured in gated events per second at times *t1* and *t2*, respectively, V_{tx} is the transfer volume in µL between passages and V_w is the media volume in µL.

Supplementary Note 4. Estimating differentiation rates from cytometry data

As described in Box 1 and the methods section, the differentiation rate σ for individual strains was estimated by applying parameter bootstrapping for growth rate estimates to fit Equations (5-6) to the fraction of lysis-deficient altruists. Cytometry samples were processed using custom Matlab scripts as previously reported[3]. Briefly, samples were filtered using a fixed gate for forward scatter values between 16,000 and 100,000 and side scatter values between 10 and 7,000. The altruist subpopulation was determined by using a fixed threshold in the GFP channel (488 nm excitation with 533 nm bandpass emission filter) of 600, which captures less that 0.1% of the population for a non-fluorescent strain.

An analytical solution to consumer and lysis-deficient altruist populations that accounts for a temporal delay observed in differentiation experiments was generated from Equations (5-6) of Box 1:

$$C(t) = e^{(v_C-\sigma)(t-\tau)}, \quad \text{(S3)}$$

$$A(t) = \frac{\sigma\left(e^{v_A(t-\tau)} - e^{(v_C-\sigma)(t-\tau)}\right)}{v_A - v_C + \sigma}. \quad \text{(S4)}$$

Differentiation rate σ and corresponding delay τ were estimated from measurements of $\frac{A(t)}{A(t)+C(t)}$ fit to Equations (S3, S4) for the cellulase and lysis deficient SDAc strains shown in Figure 2c (see Supplementary Table 3 for values). Fits were generated using the Levenberg-Marquardt nonlinear least squares algorithm (nlsLM) in R. Differentiation delay was observed to be inversely proportional to the differentiation rate (see Supplementary Figure 4). Parameter bootstrapping was performed to create 10,000 fits for each of three biological replicates, randomly sampling a normal distribution (n=10,000) of v_C and v_A from control strain DL146. The mean and standard deviation of the fit ensemble were calculated as the mean and standard deviation differentiation and delay measures for each strain.

Supplementary Note 5. Estimating lysis rates from cytometry data

As described in Box 1 and the methods section, the lysis rate ρ for individual strains was estimated by applying parameter bootstrapping for growth rate, differentiation and delay estimates to fit Equation Module I to the fraction of GFP-expressing altruists. Cytometry samples were processed as described in Supplementary Note 4.

An analytical solution to consumer and altruist populations that includes a lysis term and accounts for delayed differentiation as in Supplemental Equations S3 and S4 was generated from Equation Module II of Box 1:

$$C(t) = e^{(v_C-\sigma)(t-\tau)}, \quad \text{(S3)}$$

$$A(t) = \frac{\sigma\left(e^{(v_A-\rho)(t-\tau)} - e^{(v_C-\sigma)(t-\tau)}\right)}{v_A - v_C - \rho + \sigma}. \quad \text{(S5)}$$

Lysis rate ρ was estimated from measurements of $\frac{A(t)}{A(t)+C(t)}$ fit to Equations (S3, S5) for the SDAc strains shown in Figure 2f (see Supplementary Table 4 for values). Fits and parameter bootstrapping were performed as described for differentiation rate estimates in Supplementary Note 4. Random growth rate estimates for bootstrapping were sampled from cellulase production variant DL147 and differentiation rates estimates were sampled from rates of the appropriate switch variant.

Supplementary Note 6. Estimating biomass yield for cellobiose and PASC

The Monod model of bacterial growth assumes a linear relationship between the saturated density of cells and the limiting nutrient and states that the slope of the line is the yield coefficient for the nutrient[4]. Here, we depict that relationship as

$$Y(n) = \gamma n + b, \quad \text{(S6)}$$

with Y as the biomass yield, γ as the yield coefficient, n as the nutrient level and b as an inoculation-dependent y. DL146 cultures grown in cellobiose appeared to be carbon limited only at concentrations below $1\ g\ L^{-1}$ (Supplementary Figure 7a). When all cellobiose concentrations are included in the fit, carbon utilization matches an exponential saturation model,

$$Y(n) = S(1 - e^{\alpha n}), \tag{S7}$$

with *S* as the maximum biomass yield afforded by the limiting nutrient. No saturation was observed for PASC (Supplementary Figure 7b), suggesting the media remains carbon limited up to $4\ g\ L^{-1}$. Comparison of the slopes for each carbon source ($\gamma_{PASC}/\gamma_{cellobiose}$) suggests approximately 14% of available carbon in the cellulose is released in a digestible form for cellulase CelD04. Estimates for γ are summarized in Supplementary Table 6.

Supplementary Note 7. Estimating cellulase activity rates

Culture densities before and after sonication were used to estimate altruist concentration *A* and lysis efficiency ρ from the feedstock equation of Module I in Box 1. As mentioned in the Methods section, two approaches were used to estimate cellulose hydrolysis: Congo Red staining of residual cellulose and biomass conversion from released nutrients. Congo Red absorbance measurements consistently suggested higher cellulose degradation rates than the nutrient release assay, suggesting some hydrolysis products, including oligomers larger than cellobiose, might not be digestible by SDAc strains.

A spectrophotometric calibration curve was established to estimate residual cellulose by measuring absorbance at 480 nm for PASC supernatant stained with 0.15% Congo Red. The calibration curve was generated for PASC concentrations from $2.68\ g\ L^{-1}$ to $4\ g\ L^{-1}$. Absorbance readings were converted to cellulose concentrations using a linear fit to the calibration curve ($R^2 = 0.8869$).

$$[Cellulose] = -(0.595 \pm 0.0125) \times ABS_{480} + (4.7 \pm 0.03) \tag{S8}$$

Cellulose concentrations were estimated from released nutrients by measuring viable cell counts of cultures grown in PASC supernatant and converting cell counts to nutrients using the γ parameter and Supplementary Equation S6.

$$F = F_0 - n_i = F_0 - \frac{CFU(n_i) - b}{\gamma} = 0.004 - \frac{CFU(n_i) - (5.1\text{e}07 \pm 1.68\text{e}07)}{(1.28\text{e}12 \pm 3.34\text{e}10)} \tag{S9}$$

Parameter bootstrapping was used to estimate feedstock degradation for measurements using each method by generating mean and standard deviation statistics for 10,000 fits, each randomly sampling from a normal distribution of the error for the appropriate calibration curve.

An analytical solution to feedstock levels was generated from feedstock equation of Module I in Box 1 to estimate cellulase activity rates from Congo Red absorbance or nutrient release measurements,

$$F = Pe^{-\alpha\rho At} + f, \tag{S10}$$

with parameters P, α, and f. By the rules of derivation, the cellulase activity rate ω is the product of P and α. Fits were generated using initial altruist concentrations (A) and lysis efficiency (ρ) measured for each strain (Supplementary Table 7).

Supplementary Note 8. Estimating half-maximal rate constants for growth and differentiation

Nutrient-specific half-maximal rate constants for consumer and altruist growth – (k_C and k_A, respectively) and differentiation (k_σ) were fit to the growth curve for strain DL110. 10,000 fits were generated using parameter distributions with the mean and standard deviation for σ, ρ and ω associated with switch variant $(\mathrm{T})_{16}/(\mathrm{A})_{10}$, lysis rbSSR variant $(\mathrm{AT})_8$ and CelD04 payload variant from Supplementary Tables 4, 5 and 7, respectively. Parameter bootstrapping distributions for growth rates (v_C, v_A) were generated from DL147 (Supplementary Table 3):

$$\begin{aligned} v_C &= 0.459 \pm 0.004 \\ v_A &= 0.413 \pm 0.012 \\ \sigma &= 0.116 \pm 0.009 \\ \rho &= 0.191 \pm 0.025 \\ \gamma &= 1.279 * 10^{12} \pm 3.344 * 10^{10} \\ \omega &= 1.6 * 10^{-12} \pm 5.3 * 10^{-13} \end{aligned}$$

The half maximal constants for C and A are assumed to be the same and were fit as one parameter. The resulting parameter fits are shown in Supplementary Table 8. These half-maximal rate constants were adopted for use as necessary in models for SDAc strains.

Supplementary Note 9. Initial conditions for nutrients and population size in full dynamics model

SDAc system dynamics, both in the model and the physical implementation, were sensitive to initial conditions. For example, PASC growth media was inoculated with SDAc cultures directly from a saturated culture in rich media. Thus, the initial concentration of viable cells was dependent on differentiation rate (cultures with high differentiation rates likely generate more altruists, reducing the concentration of viable cells) and the inoculum included cellulase payload to begin converting PASC to nutrients. To account for these details, initial nutrients (with a concomitant reduction in initial feedstock) and initial population size were estimated as a function of the differentiation rate σ. Nutrients levels were estimated by fitting a logarithmic curve to the growth yield of DL146 in PASC media with filtered supernatant from SDAc strains DL046, DL112, DL110 and DL108, applying the biomass conversion equation from Module IV of Box 1 to convert CFUs to nutrients. Parameter bootstrapping from uncertainty in the γ parameter (Supplementary Equation S6) was applied to the fit (10,000 samples), which produced an estimate for initial nutrients as a function of differentiation rate (see Supplementary Figure 11): $n_0 = (1.952 * 10^{-4} \pm 2.952 * 10^{-5}) \log(\sigma) + (8.529 * 10^{-4} \pm 9.463 * 10^{-5})$.

A linear curve fit was used to map differentiation rate to initial population size. Viable cell count mean and standard deviation measurements from the same SDAc strains above were used to map differentiation rate σ to initial cell concentration x_0: $x_0 = -(1.7{\times}10^8 \pm 4.14{\times}10^7)\sigma + (3.66{\times}10^7 \pm 5.43{\times}10^6)$. Initial population estimates were distributed among consumers and altruists by applying the steady state solutions to the equations of Module II from Box 1 for DL147 growth rates and strain-specific differentiation rates.

Supplementary Note 10. Estimating SDAc cheater rates

Cooperation escape rates for consumers and for altruists were estimated by applying parameter bootstrapping for growth, differentiation and lysis rates to fit the equations of Module V from Box 1 to the fluorescence distribution and overall growth rate data from SDAc hyperdifferentiator strain DL108. The model was fit to the fluorescence distribution data (collected as described in Supplementary Note 4) separately with different assumptions: no cheaters ($\chi_C = \chi_A = 0$, simplifies to Equations 5-6), switch-deficient cheaters only ($\chi_A = 0$) and both switch and lysis-deficient cheaters. Cheater rate estimates for the latter two assumptions are shown in Supplementary Table 9. The fits were generated using the

Levenberg-Marquardt nonlinear least squares algorithm (nlsLM) in R. The standard error represents residuals error reported by the algorithm.

Supplementary Note 11. Modeling cheater dynamics

Consumers and altruists only

Steady-state analysis of the mutagenesis-naïve model of SDAc dynamics cannot account for the observed altruist fraction dynamics. The initial model for SDAc dynamics included cell states for only consumers (C) and altruists (A). Examination of the altruist fraction dynamics with differentiation and lysis for hyperdifferentiator $\texttt{(T)}_{18}/\texttt{(A)}_{10}$ grown in cellobiose (Supplementary Figure 5) clearly shows that the initial models do not track the measured data. While the population appears to reach an initial steady-state fraction around 40 hours, the altruist fraction subsequently decreases to settle at a new, non-zero steady state at 120 hours.

For this model, the fraction of cells expressing RFP (f_{red}) and GFP (f_{green}) represent the fraction of consumers and altruists, respectively. We can derive the rate of change of each relative fraction from the ODEs of Module II in Box 1 of the main text by applying the quotient rule:

$$f_{red} = f_C = \frac{C}{C+A}, \qquad f_{green} = f_A = \frac{A}{C+A}.$$

$$\dot{f_C} = \frac{\dot{C}(C+A) - C\left(\dot{C}+\dot{A}\right)}{(C+A)^2} = \frac{vC - \sigma C}{C+A} - \frac{f_C(vC - \sigma C + vA - \rho A + \sigma C)}{C+A} = f_C(\rho f_A - \sigma)$$

$$f_C^*: f_C(\rho f_A - \sigma) = 0$$

$$f_C^* = 0$$

$$f_A = (1 - f_C), \qquad f_C(\rho(1 - f_C) - \sigma) = 0$$

$$f_C^* = 1 - \frac{\sigma}{\rho}$$

$$\dot{f_A} = \frac{\dot{A}(C+A) - A\left(\dot{C}+\dot{A}\right)}{(C+A)^2} = \frac{vA - \rho A + \sigma C}{C+A} - \frac{f_A(vC - \sigma C + vA - \rho A + \sigma C)}{C+A} = \sigma f_C - \rho f_A(1 - f_A)$$

$$f_A^*(f_C^* = 0): 0 - \rho f_A(1 - f_A) = 0$$

$$f_A^*(f_C^* = 0) = 0, 1$$

$$f_A^*\left(f_C^* = 1 - \frac{\sigma}{\rho}\right): f_C = 1 - f_A, \qquad 1 - f_A = 1 - \frac{\sigma}{\rho}$$

$$f_A^*\left(f_C^* = 1 - \frac{\sigma}{\rho}\right) = \frac{\sigma}{\rho}$$

$$(f_C^*, f_A^*) = \begin{cases} (0,0) & \text{stable} \\ \left(1 - \frac{\sigma}{\rho}, \frac{\sigma}{\rho}\right) & \text{saddle for } \sigma < \rho, \text{unstable for } \sigma > \rho \\ (0,1) & \text{unstable for } \sigma < \rho, \text{saddle for } \sigma > \rho \end{cases}$$

Steady state analysis of this model constrains the consumer and altruist fractions to simple, saturating dynamics from the initial condition $f_A = 0$, as used in experiments. For this model, the fractions stabilize at

$\left(1-\frac{\sigma}{\rho},\frac{\sigma}{\rho}\right)$ if σ is less than ρ, or $(0,1)$ if σ is greater than or equal to ρ. The steady state with both population fractions at zero is trivial.

Consumers, altruists, and switch cheaters

Let us now consider a system in which a subpopulation of consumer cells no longer differentiates. We will refer to this cell state as switch cheaters (S). The resultant three cell states can be described by the following system of equations, with an introduced consumer defection rate χ:

$$\dot{C} = (v-\sigma-\chi)C$$
$$\dot{A} = (v-\rho)A+\sigma C$$
$$\dot{S} = vS+\chi C$$

For this model, the red cell fraction includes consumers and switch cheaters:

$$f_{red} = f_C + f_S = \frac{C+S}{C+A+S}, \qquad f_{green} = f_A = \frac{A}{C+A+S}.$$

Deriving the fractional dynamics, as above, results in the following system of equations:

$$\dot{f_C} = f_C(\rho f_A - \sigma - \chi)$$
$$\dot{f_A} = \sigma f_C - \rho f_A(1-f_A)$$
$$\dot{f_S} = \chi f_C + \rho f_A f_S$$

The altruist and switch cheater fractions can be rewritten to remove the variable for the consumer fraction by assuming the three fractions sum to 1:

$$\dot{f_A} = \sigma\big(1-(f_A+f_S)\big) - \rho f_A(1-f_A)$$
$$\dot{f_S} = \chi\big(1-(f_A+f_S)\big) + \rho f_A f_S$$

$$f_C^*: f_C(\rho f_A - \sigma - \chi) = 0$$
$$f_C^* = 0, 1-f_S-\frac{\sigma+\chi}{\rho}$$

$$f_S^*: \chi(1-f_A) + f_S(\rho f_A - X) = 0$$
$$f_S^* = \frac{\chi(f_A-1)}{\rho f_A - \chi}$$

$$f_A^*: \sigma\left(1-f_A-\left(\frac{\chi(f_A-1)}{\rho f_A-\chi}\right)\right) - \rho f_A(1-f_A) = 0$$
$$f_A^* = 0, \frac{\sigma+\chi}{\rho}, 1$$

The only equilibrium point for this system is $(f_C^*, f_A^*, f_S^*) = (0,0,1)$. In this model, the green cell fraction dynamics peak at $\frac{\sigma+\chi}{\rho}$ or 1 and then decrease to zero. The red cell fraction, however, stabilizes at 1 because of the switch cheater population, which acts as a sink state.

Cheaters, altruists, switch cheaters, lysis cheaters

In the data used to fit lysis rate, f_{green} appears to stabilize at a non-zero value suggesting the presence of an additional cheater subpopulation that differentiate, but cease to lyse; we will refer to these cells as

pseudo-altruists (P). The resultant four cell states can be described by the following system of equations, with consumer defection rate χ_C and altruist defection rate χ_A:

$$\dot{C} = (v - \sigma - \chi_C)C$$
$$\dot{A} = (v - \rho - \chi_A)A + \sigma C$$
$$\dot{S} = vS + \chi_C C$$
$$\dot{P} = vP + \chi_A A$$

For this model, the green cell fraction now includes the lysis cheaters:

$$f_{red} = f_C + f_S = \frac{C + S}{C + S + A + P}, \qquad f_{green} = f_A + f_P = \frac{A + P}{C + S + A + P}.$$

Deriving the fractional dynamics, as above, results in the following relationships:

$$\dot{f_C} = f_C(\rho f_A - \sigma - \chi_C)$$
$$\dot{f_A} = \sigma f_C + f_A(\rho f_A - \rho - \chi_A)$$
$$\dot{f_S} = \chi_C f_C + \rho f_A f_S$$
$$\dot{f_P} = f_A(\chi_A + \rho f_P)$$

If we substitute the consumer fraction in the altruist fraction dynamics as $f_C = 1 - (f_A + f_S + f_P)$, we see that the altruist fraction has a clear steady state at zero if the combined switch and lysis cheater fractions equal one:

$$\dot{f_A} = \sigma\big(1 - (f_A + f_S + f_P)\big) + f_A(\rho f_A - \rho - \chi_A)$$

$$\dot{f_A} = \rho f_A^2 - (\sigma + \rho + \chi_A)f_A + \sigma\big(1 - (f_S + f_P)\big)$$

$$f_A^*: \rho f_A^2 - (\sigma + \rho + \chi_A)f_A + \sigma\big(1 - (f_S + f_P)\big) = 0,$$

$$f_A(\rho f_A - \sigma + \rho + \chi_A) = -\sigma\big(1 - (f_S + f_P)\big)$$

$$f_A^*(f_S + f_P = 1) = 0, \frac{\sigma - \rho - \chi_A}{\rho}$$

Additionally, the switch and lysis cheater fractions only have clear non-negative steady states when the consumer and altruist fractions are zero. Unless the lysis cheater rate is zero, this system will settle on a final non-zero ratio of switch and lysis cheaters. Thus, the dynamics of the green cell fraction follow similar trajectories to the model with only switch cheaters, but with a non-zero steady state.

Supplementary References

1. Bokinsky, G. *et al.* Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered Escherichia coli. *Proc Natl Acad Sci USA* **108,** 19949–19954 (2011).
2. Efron, B. & Tibshirani, R. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical science* (1986). doi:10.2307/2245500
3. Egbert, R. G. & Klavins, E. Fine-tuning gene networks using simple sequence repeats. *Proc Natl Acad Sci USA* **109,** 16817–16822 (2012).
4. Monod, J. The growth of bacterial cultures. *Annual Reviews in Microbiology* (1949).