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

Robert Egbert[1], Leandra Brettner[2], David Zong[2] and Eric Klavins[1]

1 - University of Washington, Electrical Engineering (Seattle, WA, USA);
2 - University of Washington, Bioengineering (Seattle, WA, USA)
Correspondence: Prof. Eric Klavins, Campus Box 352500, Seattle, WA 98195
Email: klavins@uw.edu

Abstract

Stochastic differentiation and programmed cell death are common developmental processes in microbes, driving diverse altruistic behaviors that promote cooperation. Utilizing cell death in developmental programs requires control over the rate of differentiation to balance cell proliferation against the utility of sacrifice. However, the regulatory networks that control these behaviors are often complex and have yet to be successfully harnessed as biotechnology. Here, we engineered a synthetic developmental gene network that couples stochastic differentiation with programmed cell death to implement a two-member division of labor. Progenitor cellobiose consumer cells were engineered to grow on cellobiose and differentiate at a controlled rate into self-destructive altruists that release an otherwise sequestered cellulase enzyme payload through autolysis to form a developmental *Escherichia coli* consortium that utilizes cellulose for growth. We used an experimentally parameterized model of task switching, payload delivery and cellulose conversion to nutrients to set key parameters to achieve overall population growth supported by cellulase release, liberating 14-23% of the available carbon. An inevitable consequence of engineering altruistic developmental behaviors is the emergence of cheaters that undermine cooperation. We observed cheater phenotypes for consumers and altruists, identified mutational hotspots and constructed a predictive model of circuit longeivity based on mutation rate estimates for each mode of evolutionary escape. This work introduces the altruistic developmental program as a new tool for synthetic biology, demonstrates the utility of population dynamics models to engineer complex phenotypes and provides a testbed for probing the evolutionary biology of self-destructive altruism.

Introduction

Compartmentalization of function across differentiated cell types was essential to the emergence of complexity in biological systems. Organogenesis in plants and animals[1], schizogamy in polychaete worms[2] and germ-soma differentiation in *Volvox* algae[3] are clear examples of these divisions of labor. Many microbial developmental programs utilize stochastic differentiation and programmed cell death as vital components of population fitness[4]. Selection for programmed cell death has been proposed to drive complex behaviors that delay commitment to costly cell fate decisions[5], enable adaptation to environmental fluctuations[6], eliminate competitor species[7], reinforce biofilm structure[8] and promote colonization of hostile environments[9]. These behaviors represent divisions of labor between subpopulations of progenitor cells that propagate the species and sacrificial cells that provide a public good, analogous to germline and somatic cells from multicellular organisms. Many aspects of the emergence of multi-cellular cooperation and the genetic circuits that control its complexity remain unclear. Limited understanding of the architectures and stimuli that control developmental gene networks constrains efforts to repurpose them for engineered cell behaviors.

Current engineering paradigms of DNA-encoded cellular logic and feedback control circuits fail to encompass the full suite of behaviors necessary to advance the fields of bioprocessing, bioremediation and cell-based therapeutics. Synthetic microbial consortia have been demonstrated to improve bioprocessing efficiency[10] or to explore other complex behaviors[11]. A major challenge to engineering microbial consortia is the control of population distributions for complex traits. While syntrophic interactions in defined communities may address some of these needs, they may not be sustainable in environments with fluctuating nutrient or microbial constituents. Further, efficient delivery of protein or small molecule payloads is constrained by the cell membrane, often requiring the expression of payload-specific pumps or secretion signals. Autolysis triggered by chemical[12,13] or autoinducer[14] signals allows release of protein payloads, but limits applications that may require continuous delivery. Synthetic developmental programs could address these challenges, enabling approaches to create and regenerate microbial communities seeded by individual cells that cooperatively carry out complex behaviors.

A synthetic altruistic developmental program

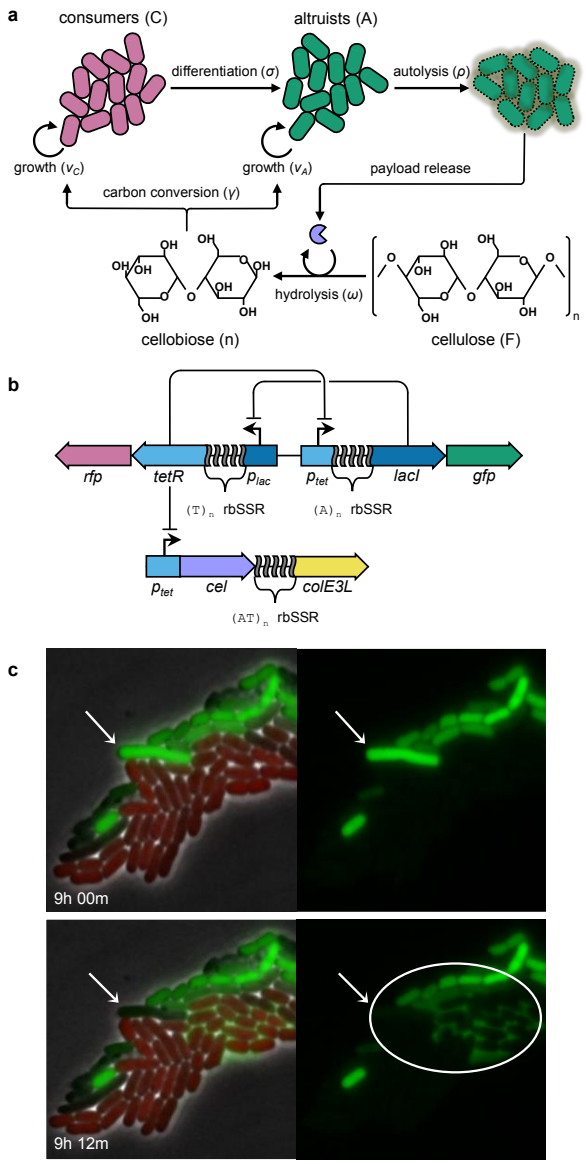

Figure 1: A synthetic developmental program for cooperative cellulose digestion. (a) Cellobiose consumers stochastically transition to self-destructive altruists. Altruists, in turn, produce and release cellulase payloads via autolysis to support the consumer population. (b) The genetic implementation of the SDAc developmental program includes a differentiation control plasmid (above) and a payload delivery plasmid (below). Cell states are mediated by a mutual inhibition toggle switch using transcriptional repressors LacI and TetR. TetR-dominant cells express RFP as consumers; LacI-dominant cells co-express cellulase and colE3 lysis (*colE3L*) proteins with GFP as altruists. Differentiation and lysis rates are fine-tuned with rbSSR sequences for *tetR* and *lacI* (differentiation) and *colE3L* (lysis). (c) Demonstration of differentiation and autolysis within a microcolony seeded by a single cell. A large altruist cell (white arrow) maintains its payload (upper panel) until it undergoes autolysis (lower panel), enabling GFP payload diffusion to surrounding cells (circle). See Supplementary Movie.

Here we present a synthetic developmental program that implements a germ-soma division of labor to cooperatively digest cellulose. The program links a synthetic differentiation controller with autolysis-mediated enzymatic payload delivery, balancing the rates of stochastic differentiation and programmed cell death to drive population growth overall. We constructed a genetic circuit to create cellobiose consumer cells that produce a sub-population of self-destructive altruists at a controlled rate to enable utilization of cellulose as a sole carbon source through extracellular release of cellulase payloads (Figure 1a). We refer to the system as SDAc, short for self-destructive altruism with a cellulase payload. We implemented SDAc in *Escherichia coli* by engineering a native operon to efficiently utilize cellobiose, introducing a genetic toggle switch tuned to function as a differentiation controller, and constructing a cellulase-lysis payload module to execute the altruist behavior (Figure 1b). We used dynamical systems analysis modeling to identify parameter values critical to achieving overall growth and demonstrated control over the circuit behavior by fine-tuning each parameter.

Using multiplexed mutagenesis and selection, we isolated a strain with a growth rate in cellobiose that is 63% its growth rate in glucose. Though *E. coli* does not natively digest cellobiose, we modified the *chb* operon in a recombinogenic MG1655 derivative[15] by replacing the native chitobiose-regulated promoter with a strong constitutive promoter[16]. We further improved growth on cellobiose by subjecting the constitutive *chb* variant to multiple cycles of single-strand DNA multiplexed recombineering targeting the *chb* genes and selected for cellobiose utilization in minimal cellobiose media (Supplementary Figure 1). We identified the variant with the highest growth rate, DL069, as a *chbR* deletion mutant.

To control differentiation rate, we constructed and sampled from a library of mutual inhibition toggle switch variants that exhibit regular stochastic state transitions. While genetic toggle switches are often designed to function as bistable memory devices[17], a quasi-steady state can be achieved by properly balancing expression levels of the repressor proteins[18]. Simple sequence repeats embedded in the ribosome binding site (rbSSR) allow predictable modulation of translation initiation rate to tune the balance between transcriptional repressors[19], which for this circuit are LacI and TetR.

We engineered altruist payload delivery by constructing a cellulase and lysis gene cassette. The operon was designed to maximize production of the cellulase payload with an efficient ribosome binding site and a poly-(AT) rbSSR to fine-tune lysis gene expression. Minimization of the fitness defects of self-destructive altruism dictates post-differentiation accumulation of the payload followed by autolysis. Colicin gene networks share this trait, using stochastic gene expression of colicin and lysis genes within subpopulations to deliver a toxin against ecological competitors[7,20]. We found that coupling the lysis gene from colicin E3 to the differentiation controller enabled stochastic state transitions and delayed lysis at the microcolony level, evidenced by post-differentiation accumulation of a GFP payload followed by autolysis (Figure 1c and Supplementary Movie).

Box 1. Systems of equations for modeling synthetic self-destructive altruism.

I.

$$\dot{C} = \frac{n}{k_C + n} v_C C - \frac{n}{k_\sigma + n} \sigma C$$

$$\dot{A} = \frac{n}{k_A + n} v_A A + \frac{n}{k_\sigma + n} \sigma C - \rho A$$

$$\dot{F} = -\omega \rho A F$$

$$\dot{n} = \omega \rho A F - \frac{n}{\gamma}\left(\frac{v_C}{k_C + n} C + \frac{v_A}{k_A + n} A\right)$$

II.

$$\dot{C} = (v_C - \sigma) C$$

$$\dot{A} = v_A A + \sigma C$$

III.

$$\dot{C} = (v_C - \sigma) C$$

$$\dot{A} = (v_A - \rho) A + \sigma C$$

IV.

$$\dot{F} = -\omega \rho A F$$

$$Y(n) = \gamma n + b$$

V.

$$\dot{C} = (v_C - \sigma - \chi_C) C$$

$$\dot{S} = v_C S + \chi_C C$$

$$\dot{A} = (v_A - \rho - \chi_A) A + \sigma C$$

$$\dot{P} = v_A P + \chi_A A$$

Module I. We constructed a population scale model composed of first order ordinary differential equations. The model contains four relevant species: consumers (C), altruists (A), cellulose (F, feedstock), and digestible nutrients (n) with corresponding units of colony forming units per mL for cells and grams per mL for molecules (1-4). Consumer and altruist cells grow in the presence of nutrients at rates v_C and v_A, respectively. Individual consumer cells differentiate to altruists at rate σ, and altruists lyse at rate ρ. Altruist payloads degrade feedstock to nutrients at rate ω, and nutrients yield biomass Y according to γ. Nutrient-dependent dynamics are controlled by half maximal rate constants for growth (k_C, k_A) and differentiation (k_σ). Autolysis is considered nutrient-independent. Additional model details are included in Supplementary Table 1 and Supplementary Note 1.

Experimental Parameter Measurements

The modularity of the synthetic SDAc developmental gene network allows experimental measurement of each parameter by systematic deconstruction of the full system. We constructed simplified sub-models to identify the key circuit parameters and measured the behaviors of defined sub-circuits to estimate the parameters. Experimental details are described in the Materials and Methods and modeling approaches to the parameter estimates are described in detail in the Supplementary Information.

Module II. Differentiation rate (σ)

A continuous growth model of consumer and pseudo-altruists was used to estimate differentiation rate for payload and lysis deficient SDAc strains. The temporal population fraction of consumers (RFP producers) and altruists (GFP producers) was measured by flow cytometry for strains pre-induced to the consumer state, washed and grown in cellobiose (Figure 2a). For each strain σ estimates were fit to this system of equations using growth rates measured independently. Unbounded growth in the dynamics represents growth for each passage over a finite duration of the periodic dilution.

Module III. Autolysis rate (ρ)

A continuous growth model of consumers and functional altruists was used to estimate autolysis rate for SDAc strains. Strains were initialized and measured as in Module I. For each strain ρ estimates were fit to equations this system of equations (Figure 2b) using growth rates and differentiation rates measured independently.

Module IV. Cellulose hydrolysis (ω)

A model of feedstock degradation and nutrient release was used to estimate cellulose hydrolysis rates for lysis deficient SDAc strains expressing cellulase payloads. Crude cell lysates were generated from cellulase producing strains. Nutrient release from PASC media inoculated with cell lysates was measured via growth of a cellulase and lysis deficient strain. For each strain ω estimates were fit to the feedstock equation F using nutrient estimates derived by applying the equation for Y to growth data.

Module V. Cheater dynamics

A continuous growth model of consumers, altruists, consumer cheaters that do not differentiate (S) and altruist cheaters that do not lyse (P) was used to estimate rates of escape from consumer (χ_C) and altruist (χ_A) states. Models that account for single or dual cheater subpopulations were used as fits to the population fraction data obtained for Module II (Figure 4d).

Model description for self-destructive altruism with enzymatic payload

Analysis of a population dynamics model of SDAc behavior suggested optimal parameter regimes for cellulose utilization and guided implementation of the developmental circuit. Though we observed the requisite behaviors of differentiation, payload accumulation and autolytic release at the microcolony level, it was not clear what expression levels of circuit components would enable cooperative growth on cellulose. To reason about the functional parameter space for the SDAc circuit we developed a population dynamics model using a system of ordinary differential equations that maps system parameters to experimentally tunable features of the genetic circuit. The model species are consumers, altruists, cellulose feedstock, and cellulose-derived nutrients. These species and the associated kinetic parameters are described in Box 1.

The core tunable parameters for SDAc cellulose utilization are the growth rate on cellobiose, differentiation rate, lysis rate and cellulase activity. Cells must utilize the hydrolysis products of the cellulase enzymes, including cellobiose. Minimal differentiation would limit growth via insufficient cellulase production, while excessive differentiation would incur unnecessary fitness defects for consumers or, at extreme rates, to population collapse. Low lysis rates would limit feedstock degradation through sequestration of intracellular cellulase and excessive rates would reduce payload burst size due to premature lysis. High cellulase activity improves growth titer by reducing the population fraction of altruists required to deconstruct the feedstock. Ultimately, SDAc performance is constrained by the tunability of the circuit components and large regions of parameter space predict no growth on cellulose (Supplementary Figure 2).

SDAc parameter estimates through modular system decomposition

The modularity of synthetic gene circuits allowed us to decompose the system model and its experimental components to estimate core parameters and predict cellulose utilization for the full circuit. We developed the modules described in Box 1 to measure growth rate on cellobiose, differentiation rate, lysis rate, and cellulase activity, each drawing from a small part library to sample a range of parameters.

We experimentally tuned differentiation rate over an order of magnitude with a collection of autolysis-deficient SDAc strains. Specifically, we used multiple poly-$(\mathtt{T})$ rbSSR variants controlling TetR expression to modulate the differentiation rate from TetR-dominant consumers to LacI-dominant altruists lacking genes for cellulase or autolysis (Figure 2a,b). We measured the population fraction of differentiated cells as a function of time using flow cytometry (Supplementary Figure 3) and fit a two-state, continuous growth model to the data for consumers transitioning to altruists at rate σ (Box1, Module II). We found that the repeat length was inversely proportional to differentiation rate, supporting previous results on a higher copy number plasmid[19] and resulting in σ estimates ranging from $2.7\times10^{-2} \pm 9.0\times10^{-3}\ h^{-1}$ for $(\mathtt{T})_{12}$ to $2.11\times10^{-1} \pm 1.5\times10^{-2}\ h^{-1}$ for $(\mathtt{T})_{18}$ (Figure 2c, Supplementary Table 4).

Lysis rates were modulated over a four-fold range using expression variants of the colicin E3 lysis gene. Using intermediate differentiator $(\mathtt{T})_{16}$, we tested a set of poly-$(\mathtt{AT})$ rbSSR variants to modulate lysis gene expression (Figure 2d,e). We used the same population fraction assay as for differentiation to fit a consumer growth and differentiation model that includes altruist lysis parameter ρ (Box1, Module III). We measured lysis rates from $7.3\times10^{-2} \pm 1.8\times10^{-2}\ h^{-1}$ for $(\mathtt{AT})_{11}$ to $1.9\times10^{-1} \pm 2.5\times10^{-2}\ h^{-1}$ for $(\mathtt{AT})_{8}$ (Figure 2f, Supplementary Table 5). As predicted by a model accounting for altruist cell death, the differentiated population fraction for each switch variant expressing the lysis gene is lower than for the equivalent autolysis-deficient strain (Supplementary Figure 5). We found, however, that the lysis rate did not correlate with rbSSR length (Supplementary Figure 6).

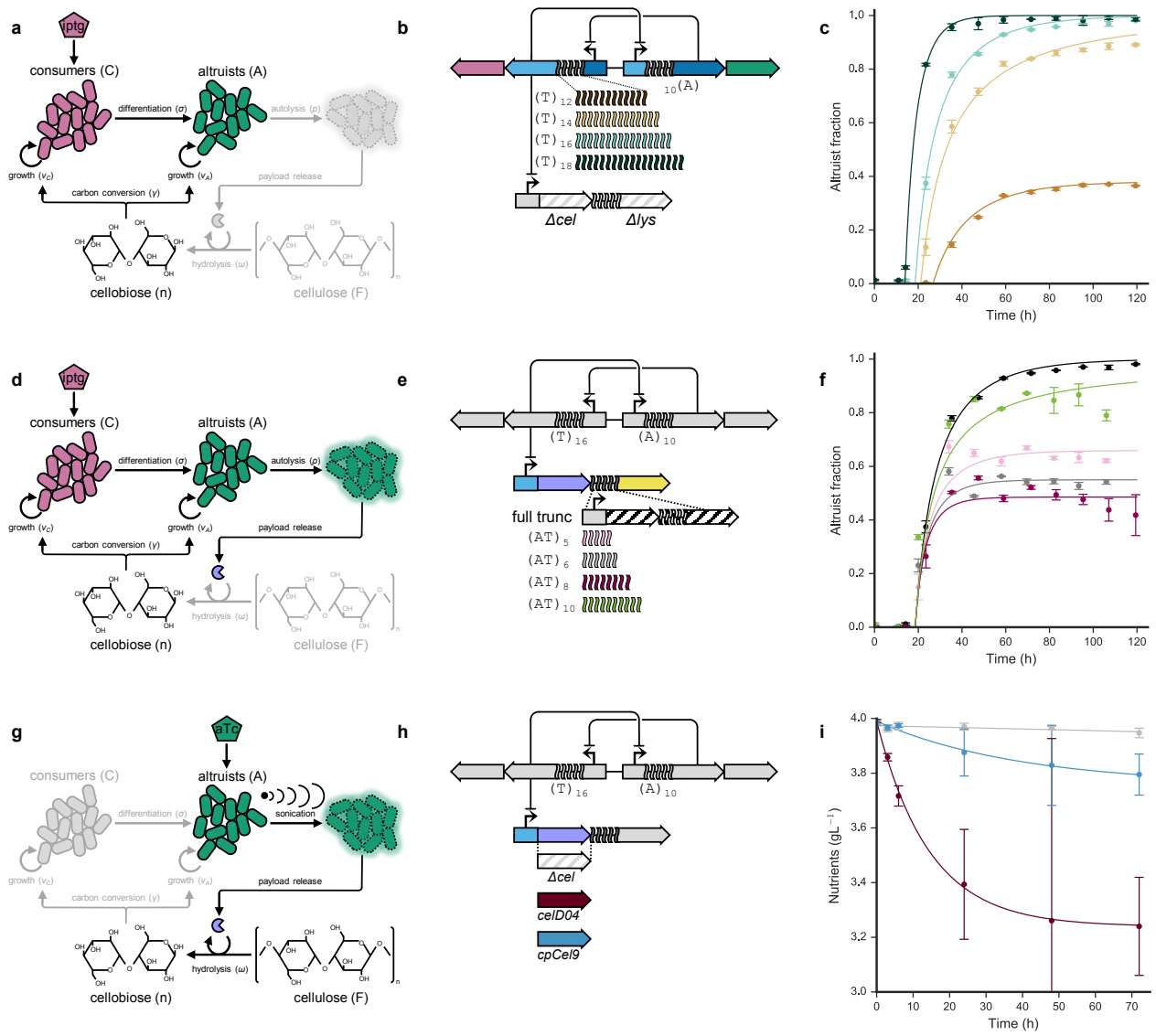

Figure 2. Parameter sweeps for differentiation, autolysis and cellulase activity. (a) For the differentiation assay, cultures were initialized to the consumer state using IPTG and grown continuously in cellobiose to differentiate into altruists deficient in cellulase and lysis. (b) Genetic variants tested for differentiation vary rbSSR repeat length to modulate TetR expression. Cellulase and lysis genes are removed from the payload plasmid. (c) Differentiation data and model fits to estimate $\boldsymbol{\sigma}$ (see Supplementary Note 4). (d) For the lysis assay, cultures were initialized as in (a) for consumers to differentiate into lysis-competent altruists. (e) Genetic variants tested for lysis used intermediate differentiator $(\mathtt{T})_{16}$, varying $(\mathtt{AT})$-rbSSR repeats driving lysis gene expression or using a control plasmid with no cellulase or lysis genes (ctl). (f) Lysis data and model fits to estimate $\boldsymbol{\rho}$ (see Supplementary Note 5). (g) For the cellulase activity assay, lysis-deficient strains were induced to the altruist state using aTc, growth to saturation and sonicated to generate crude cell extracts. (h) Genetic variants tests for cellulase activity by expressing different cellulases or maintaining a control plasmid (***Δcel***). (i) Cellulase activity data and model fits to estimate $\boldsymbol{\omega}$ (see Supplementary Note 7).

To estimate cellulase activity we quantified cellulose degradation from cell lysates of autolysis-deficient SDAc strains producing individual cellulases, observing hydrolysis rates over a three-fold range. We measured cellulose degradation and digestible nutrient release for three endoglucanases from two glycoside hydrolase (GH) families: CelD04[21] and BsCel5[22] from GH5; and CpCel9 from GH9[23]. We also measured the activity of multi-enzyme cocktails using each GH5 enzyme with CpCel9, combinations shown to have synergistic activities[24] (Figure 2f,g). We used Congo Red staining to observe cellulose degradation up to 23% (Supplementary Note 6) in M9 minimal phosphoric acid swollen cellulose (PASC) spiked with cell lysate and quantified cell growth on the resulting supernatant to estimate nutrient release of up to 14% of cellobiose equivalents (see Supplementary Figure 8, Supplementary Note 6). We used these cellulase activity measurements to fit a value for ω to the feedstock differential equation (Box 1, Module IV). Cellulase activity estimates range from 6.0×10^{-13} CFU^{-1} mL for CpCel9 to 1.8×10^{-12} CFU^{-1} mL for the BsCel5/CpCel9 cocktail (Figure 2i, Supplementary Table 7).

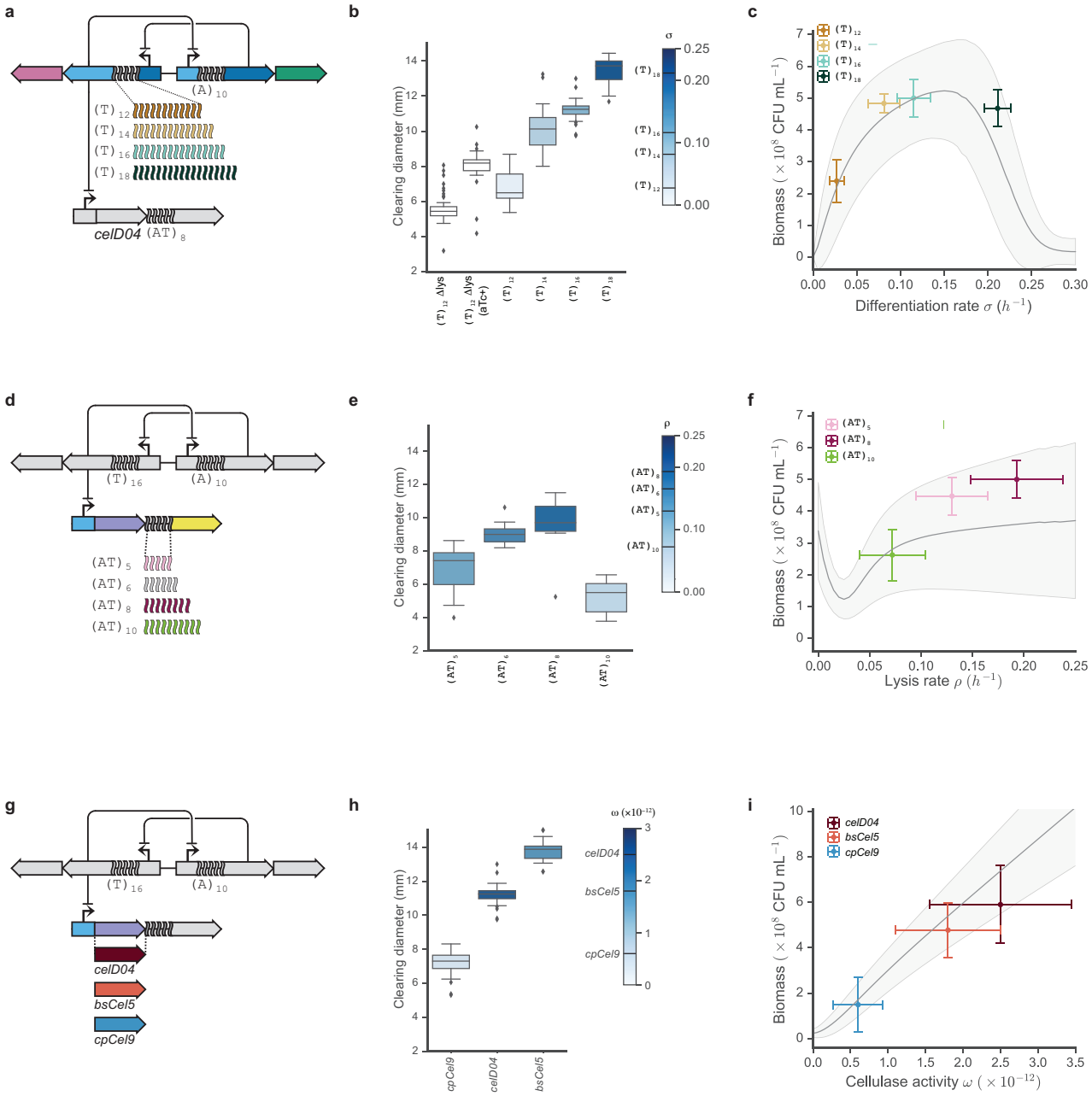

Figure 3. Characterization of cellulose hydrolysis and nutrient utilization reveals SDAc model prediction accuracy. (a) Differentiation rate variants expressing cellulase CelD04 with lysis rbSSR $(AT)_8$. (b) Clearing size distributions for individual colonies from differentiation variants in (a) (N = 31, 21, 17, 15, 6). (c) Growth titer in M9 minimal 0.4% PASC media after 72 hours for differentiation variants in (a). X-axis error bars represent standard deviation of the parameter estimate and y-axis error bars represent standard error for CFU counts from at least four replicates. One standard deviation of model uncertainty for cell growth is displayed as the shaded region. (d) Lysis rate variants for intermediate differentiator $(T)_{16}$ expressing cellulase CelD04. (e) Clearing size distributions for individual colonies from lysis variants in (d) (N = 20, 14, 11, 15). (f) Growth titer as in (c) for some lysis variants in (d). (g) Cellulase variants for intermediate differentiator $(T)_{16}$ and an optimal lysis rbSSR for each cellulase. (h) Clearing size distributions for individual colonies from cellulase variants in (g) (N = 20, 15, 16). (i) Growth titer as in (c) for cellulase variants in (g), except that error bars for the x-axis represent the interquartile range for each ω estimate. Whiskers shown for each box plot extend one interquartile range.

Full system performance & model predictability

To quantify the combined effects of SDAc differentiation and autolysis dynamics on feedstock degradation we measured cellulase activity on cellulose. Cellulose hydrolysis by individual colonies was measured by Congo Red clearing assays on agar plates supplemented with carboxymethylcellulose (Figure 3). We found that the clearing diameter for switch variants increased as a function of differentiation rate (Figure 3b). We observed no clearings for a control lacking cellulase. We also tested the effect of rbSSR lysis variants for cellulase CelD04 (Figure 3c,d) as well as for individual cellulases (Figure 3g,h) using intermediate

differentiator $(\mathrm{T})_{16}$. We found the GH5 cellulases generated larger clearings than CpCel9, consistent with the in vitro cellulase activity results.

Fine-tuning the differentiation, lysis and cellulase activity parameters is critical to realizing robust SDAc growth on cellulose as a sole carbon source. To determine fitness on cellulose and validate the full dynamics model (Box 1, module I), we measured viable cell counts in PASC for SDAc variants that span a range of values for each core parameter. We found optimal SDAc fitness at intermediate differentiation rates (Figure 3c), with efficient payload release (Figure 3f) and with high cellulase activity (Figure 3i), trends that are consistent with the naive model predictions from Supplementary Figure 2. Growth kinetics fits using parameter estimates from individual modules match observations for most variants (Supplementary figure 9), though the model predicted longer lags for differentiation variant $(\mathrm{T})_{18}$ and failed to capture the final cell density for the BsCel5-CpCel9 cellulase cocktail.

SDAc system cheaters

The full dynamics model of SDAc growth on cellulose predicts that high differentiation rates will lead to system collapse (Figure 3a), but does not account for mutational dynamics that could produce non-cooperative cheaters. Indeed, we found the SDAc system showed long-term functional instability for hyperdifferentiator switch variant $(\mathrm{T})_{18}$. The instability was manifested in the cellobiose differentiation and lysis assay as two distinct, temporally separated steady-state altruist fractions (Supplementary Figure 5), a result that was supported by distinct steady-state population growth rates coincident with each altruist fraction (see Figure 4e). We observed two additional colony phenotypes after extended growth of the same circuit variant in PASC media (Figure 4a), further suggesting functional instability at extreme differentiation rates.

Analysis of all potential dynamic models incorporating cheater states suggests that two cheater phenotypes independently deficient in differentiation and lysis are necessary to observe the observed dynamics. Four potential models are possible for the cheater dynamics, each with multiple potential initial conditions. The models account for differentiation-deficient cheaters, lysis-deficient cheaters, both cheater types, or no observable cheater phenotypes (Figure 4b). For initial conditions, cheaters may arise from distinct mutation events during growth or be pre-existent in the culture and rise in population fraction once the majority of cells differentiate and lyse. We derived analytical solutions for each of these models, finding that the only models that support the observed dynamics include both cheater states (Supplementary Note 11). These solutions do not distinguish between initial conditions that require *de novo* mutagenesis or pre-existing mutants subpopulations in the starting populations.

To examine SDAc dynamics with cheaters we developed a model that introduces new species for switch-deficient cheaters (S) and for lysis-deficient pseudo-altruist cheaters (P) and their respective escape rates, χ_C and χ_A (Box 1, module V). We fit the cheater model to the differentiation and lysis data in cellobiose for hyperdifferentiator $(\mathrm{T})_{18}$ (Figure 4c). The model fit estimates the emergence of switch-deficient cheaters at a rate of $5.8\times10^{-6} \pm 1.6\times10^{-6}\, h^{-1}$ and the emergence of altruist cheaters at a rate of $3.8\times10^{-5} \pm 1.7\times 10^{-5}\, h^{-1}$. Using the mutagenesis parameters fit from hyperdifferentiator $(\mathrm{T})_{18}$, intermediate differentiator $(\mathrm{T})_{16}$ is also predicted to accumulate cheaters within the measurement interval, which fits the trend of the cellobiose switching data. When applying these escape rates with differentiation and lysis dynamics to the combined population growth rate for differentiator variants, we found the model predicted the observed growth rate dynamics for variants with high differentiation rates (Figure 4d). Further, using the cheater rate estimates into a modified PASC growth model that incorporates cheater dynamics improves the growth fit for the differentiation variants, in particular $(\mathrm{T})_{18}$ (Supplementary Figure 10).

DNA sequencing of cheater isolates confirms genetic sources for both differentiation and lysis deficiencies (Figure 4e). We observed bright red colonies and bright green colonies – putative differentiation and lysis cheaters, respectively – in addition to the expected mixed-color colony on solid media after growth in PASC. Sequence analysis of the differentiation plasmid from red escape colonies isolated from six replicate cultures revealed mutations to two hypermutable loci with predictable effects (Supplementary Table 10): expansion or contraction of the tandem $(\mathrm{CTGG})_3$ mutational hotspot observed in four of six replicates should prevent altruist emergence through inactive, truncated Lac repressor[25]; deletions within the $(\mathrm{T})_{18}$ rbSSR controlling TetR expression (one of six replicates) should abolish differentiation by reducing σ; and a transposition event of IS2[26] (one of six replicates) internal to *lacI* should abolish state transition. Sequence analysis of the payload delivery plasmid revealed transposition of IS1[27] between the cellulase and lysis genes in one of six sequenced replicates, likely disrupting operon expression (Supplementary Table 11). The majority of the altruist cheater colonies we sequenced revealed no mutations in the payload delivery transcription unit, suggesting lysis evasion via mutations on the genome or elsewhere on the plasmids. Given that the lysis

gene is sourced from a colicin plasmid found in natural *E. coli* populations, it is possible the MG1655 genome encodes evolutionary paths to lysis immunity.

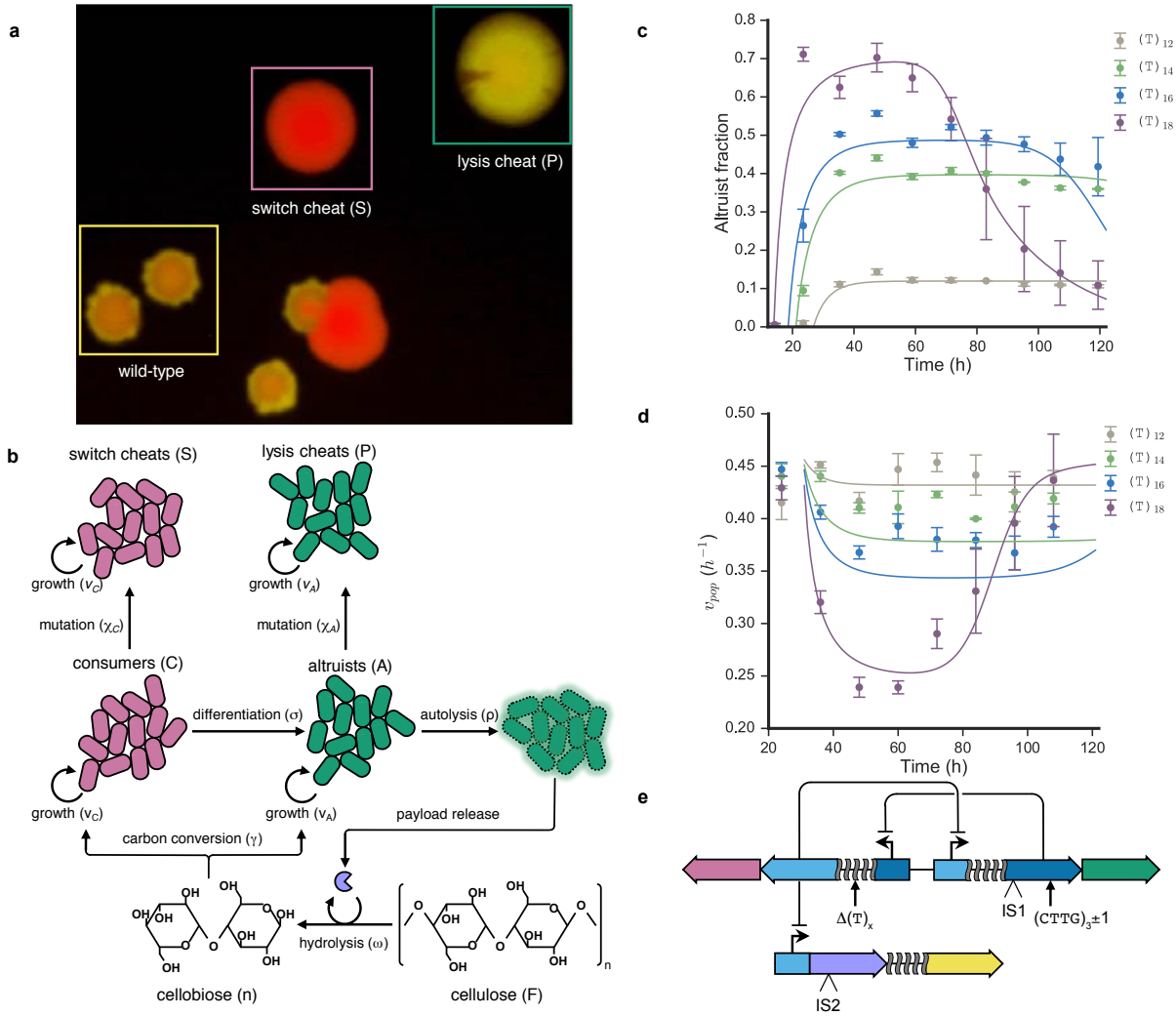

Figure 4. Mechanisms and rates of escape for SDAc cheaters. (a) Fluorescence image of wild-type and cheater colonies isolated from PASC cultures of hyperdifferentiator $(T)_{18}$ that no longer differentiate (S) or no longer lyse (P). (b) Representation of alternate SDAc model that incorporates mutational dynamics by including sink states for switch cheaters (S) and pseudo-altruist cheaters (P) with associated mutagenesis rates (see Box 1, module V). (c) Dual cheater model fit to altruist fraction measurements for differentiation (compare to Supplementary Figure 5). Escape rate estimates for hyperdifferentiator $(T)_{18}$ are used for all altruist fraction fits. (d) Continuous growth model fit for bulk population growth rate v_{pop} measurements of differentiation variants, incorporating dual cheater rates from hyperdifferentiator $(T)_{18}$. (e) Summary of observed mutations that produce switch or pseudo-altruist cheaters. Switch mutants included an expansion or contraction of a native simple sequence repeat element in the *lacI* coding sequence, repeat unit truncations of rbSSR $(T)_{18}$ that controls differentiation rate and insertion sequence disruption of LacI expression. Insertion sequence disruption of the CelD04 gene (green arrows) were observed for altruist cheaters. For additional details on observed mutations see Supplementary Tables 10 and 11.

Discussion

We have demonstrated a first-principles approach to construct simple developmental gene circuits for synthetic biology and implemented one two-member developmental system to cooperatively utilize the complex feedstock cellulose. In-depth system deconstruction and characterization enables model-guided optimization of growth on cellulose. At extreme differentiation rates, genetic instability drives the emergence of cheaters that fail to differentiate or fail to lyse.

Due to the observed functional instability for some variants, the SDAc program likely suffers from a tragedy of the commons[28]. In well-mixed cellulose media, emergent cheaters fully benefit from the public good provided by the altruists. Further, due to the costs of switching and lysis, the cheaters can out-compete cooperators and sweep the population. In the absence of altruists, cellulose degradation ceases, resulting in

population collapse. Previous work has shown that when the environment is organized spatially a communal benefit applies only to nearby, closely related cells who are likely fellow cooperators[29]. Indeed, research suggests the cellulosome evolved to localize the benefits of cellulase expression, as in sucrose utilization in yeast[30]. Cheaters are stranded with limited or no access to the shared resource. This phenomenon, attributed to kin selection, could preserve cooperative behavior for many more generations, potentially avoiding the functional instability currently observed. Future work could elucidate the role of structured environments in this synthetic system to reduce the impact of cheaters or to evolve more stable cooperator phenotypes.

While we only observed a high fraction of SDAc cheaters from hyperdifferentiation variant $(T)_{18}$, engineering developmental circuits for deployment in bioreactors or other complex environments would require long-term evolutionary stability to minimize cheaters and maintain engineered function. Interestingly, previous studies have shown that *lacI* tandem repeat mutations occur at a rate $> 10^{-6}$ events per generation[25] and transposon insertion elements jump at rates of 10^{-6}-10^{-5} insertions per generation[31]. These rates are consistent with our experimental estimates for mutagenesis, suggesting relatively simple modifications may considerably boost SDAc circuit longeivity. Analysis of the cheater model suggests that a reduction of cheater rates by 100-fold and 1000-fold for intermediate differentiator $(T)_{16}$ would increase circuit stability by 56% and 85%, respectively, boosting the functional period from 2.8 days to 5.1 days. Genetic strategies to boost evolutionary stability include recoding the repeat region of *lacI*, introducing stabilizing degeneracy into rbSSR sequences and porting the system to a low mutation rate strain deficient for insertion elements[32]. Further gains in system performance could be achieved by chromosomal integration of the SDAc network to prevent the fixation of mutant plasmids in the population[33] or plasmid loss. Finally, incorporating more efficient cellulase cocktails will reduce evolutionary pressure for cheating by decreasing the optimal altruist fraction.

The division of labor system outlined here is a template for the construction of other developmental programs to perform complex tasks in engineered microbial communities. This work can be extended in many ways. For SDAc, the developmental program could be triggered in response to nutrient depletion when the supply of simple sugars is depleted. Alternative protein and small molecule payloads from a general autolytic delivery system could be designed to mediate microbial interactions, aid in bioprocessing or bioremediation or as a cellular therapeutic. Further, stochastic strategies could be employed with or without self-destructive altruism to seed multicellular developmental programs for distributed metabolic engineering[34], evolutionary engineering[35] or to control distributions of multiple cell types in microbial communities[36,37]. Tunable developmental programs could also be applied to better understand the emergence and persistence of well-studied developmental programs, substituting complex regulation with tunable differentiation dynamics.

Acknowledgements

We thank Ben Kerr for helpful discussions regarding evolutionary dynamics of self-destructive altruism and for sharing a sample of the colicin E3 plasmid. We also thank Chris Takahashi for sharing strain CT009. This work was funded by National Science Foundation Molecular Programming Project grants 0832773 (R.G.E and E.K.) and 1317653 (L.M.B and E.K.), National Science Foundation Bio/computational Evolution in Action CONsortium (BEACON) grant 0939454 (L.M.B. and E.K.) and the University of Washington Mary Gates Research Scholarship (D.Z.). R.G.E. and E.K. conceived and designed the study. R.G.E., L.M.B. and D.Z. performed the experiments and analyzed the data. L.M.B., R.G.E. and E.K. performed the computational modeling and analyzed simulations. All authors wrote the manuscript.

References

1. Weissman, I. L. Stem cells: units of development, units of regeneration, and units in evolution. *Cell* **100,** 157–168 (2000).
2. Franke, H.-D. Reproduction of the Syllidae (Annelida: Polychaeta). *Hydrobiologia* **402,** 39–55 (1999).
3. Prochnik, S. E. *et al.* Genomic analysis of organismal complexity in the multicellular green alga Volvox carteri. *Science* **329,** 223–226 (2010).
4. Rice, K. C. & Bayles, K. W. Molecular control of bacterial death and lysis. *Microbiol Mol Biol Rev* **72,** 85–109, table of contents (2008).
5. González-Pastor, J. E., Hobbs, E. C. & Losick, R. Cannibalism by Sporulating Bacteria. *Science* **301,** 510–513 (2003).
6. Claverys, J.-P., Martin, B. & Håvarstein, L. S. Competence-induced fratricide in streptococci. *Mol Microbiol* **64,** 1423–1433 (2007).
7. Riley, M. A. & Gordon, D. M. The ecological role of bacteriocins in bacterial competition. *Trends Microbiol* **7,** 129–133 (1999).
8. Perry, J. A., Cvitkovitch, D. G. & Lévesque, C. M. Cell death in Streptococcus mutans biofilms: a link between CSP and extracellular DNA. *FEMS Microbiol Lett* **299,** 261–266 (2009).
9. Ackermann, M. *et al.* Self-destructive cooperation mediated by phenotypic noise. *Nature* **454,** 987–990 (2008).
10. Zhou, K., Qiao, K., Edgar, S. & Stephanopoulos, G. Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat Biotech* **33,** 377–383 (2015).
11. Chen, Y., Kim, J. K., Hirning, A. J., Josic, K. & Bennett, M. R. Emergent genetic oscillations in a synthetic microbial consortium. *Science* **349,** 986–989 (2015).
12. Tanji, Y., Asami, K., Xing, X.-H. & Unno, H. Controlled expression of lysis genes encoded in T4 phage for the gentle disruption of Escherichia coli cells. *Journal of Fermentation and Bioengineering* **85,** 74–78 (1998).
13. Huh, J. H., Kittleson, J. T., Arkin, A. P. & Anderson, J. C. Modular design of a synthetic payload delivery device. *ACS Synth Biol* **2,** 418–424 (2013).
14. Din, M. O. *et al.* Synchronized cycles of bacterial lysis for in vivo delivery. *Nature* **536,** 81–85 (2016).
15. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460,** 894–898 (2009).
16. Kachroo, A. H., Kancherla, A. K., Singh, N. S., Varshney, U. & Mahadevan, S. Mutations that alter the regulation of the chb operon of Escherichia coli allow utilization of cellobiose. *Mol Microbiol* **66,** 1382–1395 (2007).
17. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403,** 339–342 (2000).
18. Wu, M. *et al.* Engineering of regulated stochastic cell fate determination. *Proc Natl Acad Sci USA* **110,** 10610–10615 (2013).
19. Egbert, R. G. & Klavins, E. Fine-tuning gene networks using simple sequence repeats. *Proc Natl Acad Sci USA* **109,** 16817–16822 (2012).
20. Cascales, E. *et al.* Colicin Biology. *Microbiol Mol Biol Rev* **71,** 158–229 (2007).
21. 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).
22. Zhang, X.-Z., Sathitsuksanoh, N., Zhu, Z. & Percival Zhang, Y.-H. One-step production of lactate from cellulose as the sole carbon source without any other organic nutrient by recombinant cellulolytic Bacillus subtilis. *Metab Eng* **13,** 364–372 (2011).
23. Zhang, X.-Z., Sathitsuksanoh, N. & Zhang, Y.-H. P. Glycoside hydrolase family 9 processive endoglucanase from Clostridium phytofermentans: heterologous expression, characterization, and synergy with family 48 cellobiohydrolase. *Bioresour Technol* **101,** 5534–5538 (2010).
24. Liao, H., Zhang, X.-Z., Rollin, J. A. & Zhang, Y.-H. P. A minimal set of bacterial cellulases for consolidated bioprocessing of lignocellulose. *Biotechnol J* **6,** 1409–1418 (2011).
25. Farabaugh, P. J., Schmeissner, U., Hofer, M. & Miller, J. H. Genetic studies of the lac repressor. VII. On the molecular nature of spontaneous hotspots in the lacI gene of Escherichia coli. *J Mol Biol* **126,** 847–857 (1978).
26. Ghosal, D., Sommer, H. & Saedler, H. Nucleotide sequence of the transposable DNA-element IS2. *Nucleic Acids Res* **6,** 1111–1122 (1979).
27. Johnsrud, L. DNA sequence of the transposable element IS1. *Mol. Gen. Genet.* **169,** 213–218 (1979).
28. Hardin, G. The Tragedy of the Commons. *Science* **162,** 1243–1248 (1968).
29. Kerr, B., Neuhauser, C., Bohannan, B. J. M. & Dean, A. M. Local migration promotes competitive restraint in a host–pathogen 'tragedy of the commons'. *Nature* **442,** 75–78 (2006).
30. Gore, J., Youk, H. & van Oudenaarden, A. Snowdrift game dynamics and facultative cheating in yeast. *Nature* **459,** 253–256 (2009).
31. Sousa, A., Bourgard, C., Wahl, L. M. & Gordo, I. Rates of transposition in Escherichia coli. *Biology*

Letters **9,** 20130838–20130838 (2013).
32. Csörgo, B., Fehér, T., Tímár, E., Blattner, F. R. & Pósfai, G. Low-mutation-rate, reduced-genome Escherichia coli: an improved host for faithful maintenance of engineered genetic constructs. *Microb Cell Fact* **11,** 11 (2012).
33. Tyo, K. E. J., Ajikumar, P. K. & Stephanopoulos, G. Stabilized gene duplication enables long-term selection-free heterologous pathway expression. *Nat Biotech* **27,** 760–765 (2009).
34. Babson, D. M., Held, M. & Schmidt Dannert, C. *Designer Microbial Ecosystems – Toward Biosynthesis with Engineered Microbial Consortia*. *Natural Products* 23–38 (John Wiley & Sons, Inc., 2014). doi:10.1002/9781118794623.ch2
35. Esvelt, K. M., Carlson, J. C. & Liu, D. R. A system for the continuous directed evolution of biomolecules. *Nature* **472,** 499–503 (2011).
36. Brenner, K., You, L. & Arnold, F. H. Engineering microbial consortia: a new frontier in synthetic biology. *Trends Biotechnol* **26,** 483–489 (2008).
37. Bernstein, H. C. & Carlson, R. P. Microbial Consortia Engineering for Cellular Factories: in vitro to in silico systems. *Comput Struct Biotechnol J* **3,** e201210017 (2012).

Materials and methods

Media

Rich media was prepared as LB Miller broth (Cat #244620, Difco), supplemented with 15 g L^{-1} bacto agar (Cat# 214030, Difco) when making solid media and with bacto agar and 0.1% carboxymethylcellulose (C5678, Sigma) to quantify cellulase clearings. M9 minimal media was prepared with M9 salts (Cat# 248510, Difco), 1 mM $MgSO_4$, 100 μM $CaCl_2$, supplemented with 1 mg mL^{-1} biotin and a carbon source (glucose, cellobiose or PASC) at concentrations specified in the text. M9 minimal 0.4% cellobiose plates for isolating cellobiose utilizer strains were prepared as above supplemented with 15 g L^{-1} bacto agar. Ampicillin and kanamycin antibiotics were supplemented when required at 20 μg mL^{-1} and 20 μg mL^{-1}, respectively, unless indicated.

PASC preparation

Phosphoric acid swollen cellulose (PASC) was prepared following a reported protocol[1]. Briefly, a cellulose slurry was created from sterile water and cellulose powder (Sigmacell 20, Sigma), combined with ice-cold 85% phosphoric acid (Sigma) and incubated at 4°C for 2 hours. Cellulose was precipitated in sterile water and washed repeatedly to remove the acid and bring the slurry to neutral pH.

Strains and plasmid construction

All strains used for assays in this study were derived from a variant of EcNR1[2] with genomic deletions of the *fim* operon, *ampR* and *lacIZ* (CT009). Cellobiose utilizer DL069 was used as parent strain for all SDAc variants. Control strains for cellobiose utilization were derived from progenitor strain CT009. All strains are listed in Supplementary Table 2.

Plasmid construction was carried out using standard Gibson assembly protocols (Gibson et al. 2009). PCRs were performed using Phusion PCR Master Mix (NEB, M0531L) with a T-100 thermal cycler (Bio-Rad). Synthetic oligonucleotides for cloning were synthesized by Integrated DNA Technologies. Gibson assemblies were transformed into cloning strain DH5alpha. Transformations were performed by electroporation at 1250 V using an Eppendorf 2510 electroporator. Transformants were cultured on LB Miller agar plates with appropriate antibiotic or LB liquid media (Difco) at 37°C. Plasmids were extracted using the QIAprep Spin Miniprep Kit. Plasmids were sequenced by Genewiz (Seattle, WA). All plasmid assemblies were performed with vector pSC101 vector pGA4A5 or p15A vector pGA3K3 [3].

Cellulase genes were synthesized following standard polymerase cycling assembly (PCA) protocols[4] using synthetic oligonucleotides provided by OligoCo, Inc. PCA was performed with Phusion polymerase and cloned into plasmids for sequence verification using Gibson Assembly.

Genome engineering

Cellobiose utilizer strain RGE531 was engineered using multiple cycles of multiplex genome engineering following Wang *et al.*[2]. *E. coli* strain CT009 was grown in LB Miller broth to OD 0.4-0.6 at 30°C, then heat shocked in a water bath for 15 minutes at 42°C. The culture was chilled via ice slurry and centrifuged for 3 minutes at 5000 rpm in a centrifuge (Legend XR1, Thermo Scientific) pre-chilled to 4°C. The supernatant was decanted and cells were resuspended in 3 mL of ice-cold sterile water. The resuspension was centrifuged again for 3 minutes at 5000 rpm and the supernatant decanted. The cell pellet was resuspended in 800 μL of ice-cold sterile water and transferred to a 1.5 mL microcentrifuge tube. Cells were centrifuged at 4000 x g for 1 minute (Accuspin Micro 17R, Fisher Scientific) pre-chilled to 4°C. Supernatants were aspirated by pipette and cell pellets were resuspended in 40 mL ice-cold sterile water.

Competent cells were transformed with an oligonucleotide cocktail targeting ribosome binding site variation of *chbB/A/C/F/G* or deletion of *chbR* (synthesized by OligoCo, Inc., see Supplementary Table 10) by electroporation at 1800V using an Eppendorf electroporator 2510. Transformants were recovered in 10 mL of M9 0.4% cellobiose media in a shaker flask and grown for 48-72 hours at 30°C with shaking. This process was repeated for two additional cycles, using the saturated recovery culture from the previous as the inoculum for the next cycle of mutagenesis. Following each outgrowth in cellobiose, individual colonies were screened for rapid growth on M9 0.4% cellobiose agar plates, and several variants were isolated for *chb* operon sequencing and further use.

Microscopy

Fluorescence microscopy of SDAc microcolonies was carried out using equipment and procedures described previously[3]. *E. coli* strain 2.320 was transformed with a differentiation plasmid (*tetR*/*lacI* switch with $(T)_{10}/(A)_{12}$ repeats, p15A origin) and a payload delivery plasmid (colicin E3 lysis gene with $(AC)_{11}$ repeat, colE1 origin). Individual colonies were grown to saturation in minimal M9CA media (M8010, Teknova), 50 µg mL^{-1} kanamycin and 20 µg mL^{-1} ampicillin. Cultures were diluted 100:1, grown for two hours with shaking at 32°C and vortexed to break up cell aggregates. 1 µL of the culture was added to a glass-bottomed microscopy dish (GWSB-3512, WillCo Wells BV) and cells were immobilized on the dish surface using a nutrient agar slab prepared with M9CA media, as described above, and 1% bacto agar (Difco). Microscopy dishes were transferred to a Nikon Ti-S inverted microscope (Nikon Instruments) equipped with a Coolsnap-HQ camera (Roper Scientific) and an environmental chamber (In Vivo Scientific). Microcolony growth and fluorescent states were recorded every 15 min by capturing phase contrast, GFP and RFP images while the dishes were maintained at 32°C for 10 to 12 hours.

Measurements of growth, differentiation and lysis rates

Individual strains were initialized to the consumer state or altruist state (lysis-deficient control only) by inoculating individual colonies from a restreaked LB agar plate. The colonies were forced to the consumer state via IPTG induction (1 mM) or the altruist state via aTc induction (100 ng/mL) and grown at 37°C for 16 hours on LB agar media supplemented with ampicillin and kanamycin. Cultures were washed to remove inducer, inoculated 2000:1 into a well of a deep-well plate with 500 µL M9 minimal 0.4% cellobiose media supplemented with biotin and antibiotics as above, and grown at 37°C. To maintain exponential growth and low cell density, each culture was sampled every 12 hours and transferred to the cytometer (C6 with CSampler, Accuri) to measure culture density and the fluorescence distribution to determine the fraction of altruists. Growth rates between measurements were measured as a simple exponential fit using the initial cell density, the final cell density, and the number of hours between measurements. Each culture was diluted into fresh broth with a dilution factor defined by the culture density and the estimated growth rate. The dilution factors ranged from 5:1 to 50:1 every 12 hours. Growth rates and population fractions of consumers and altruists were calculated as described in Supplementary Notes 3 and 4, respectively.

Cell growth in PASC media

Cultures grown in PASC were initially inoculated into LB liquid culture with ampicillin and kanamycin from individual colonies restreaked on LB agar plates with the same antibiotics and IPTG to force the consumer state. Saturated cultures were transferred at 100:1 dilutions into M9 minimal PASC media and viable cell counts were periodically quantified by serial dilution and plating.

Measurements of carbon to biomass conversion efficiency

Carbon to biomass yield for cellobiose and PASC were measured using strains grown in M9 minimal media supplemented with increasing levels of each carbon source, up to 4 g L^{-1}. Cellobiose growth was quantified after 36 hours with cellulase and lysis deficient DL146 by serial dilution, plating and counting colonies. PASC growth was quantified after 72 hours with DL110 by the same method. All cultures were inoculated according to the PASC growth assay protocol and colony forming units were quantified for at least three replicates on LB plates supplemented with ampicillin and kanamycin. The estimate for cellobiose was fit only to the linear, carbon-limited regime observed at concentrations less than 1g L^{-1} (see Supplementary Figure XX and Supplementary Note 6)

Measurements of cellulose degradation

Cellulase production strains for individual cellulases and cellulase cocktails were initialized by restreaking on LB agar plates supplemented with kanamycin (50 µg/mL), ampicillin (20 µg/mL) and aTc (100 ng/mL). Single colonies were transferred to 3 mL LB supplemented with kanamycin and ampicillin and grown to saturation overnight. Turbid 1 mL cultures were sonicated on ice (Ultrasonic Cell Disruptor, SharperTek) to release intracellular cellulase. The sonicator settings were 15 minutes with 95% power, and 75% duty cycle of 20 seconds on and 10 seconds off. Viable cell counts were measured before and after sonication using serial dilution and CFU plating to measure cell lysis efficiency (ρ). Cell lysates were filtered (0.4 micron filter, Company) to generate sterile supernatant. 10 µL cellulase lysate was transferred to 1 mL of 0.4% M9 minimal PASC + biotin media and incubated at 37°C for 72 hours, periodically measuring cellulose degradation by Congo Red absorbance or nutrient release. Uniprot identifiers for cellulases are Q45430 (CelD04), P10475 (BsCel5) and A9KT90 (CpCel9).

Congo Red absorbance measurements were taken by sampling 100 uL of the PASC growth media with cellulase at 0, 6, 12, 24, 48 and 72 hours post-inoculation, adding 100uL 0.15% (w/v) aqueous Congo red. After staining for 30 minutes, samples were centrifuged at 5000 rpm (Legend XR1) and 100 µL of PASC

supernatant was transferred to a clear, flat-bottom microwell plate (Corning 2595). Absorbance readings were measured by plate reader (Biotek) at a wavelength of 480 nm.

Cellulolytic release of digestible nutrients was estimated by measuring DL146 growth on PASC supernatant after lysate incubation. 10 µL aliquots of sterile cell lysates, prepared as above, were inoculated in 1 mL M9 minimal 0.4% PASC media, prepared in parallel for each time measurement. At each time point, 1 mL PASC samples were centrifuged in triplicate to pellet the PASC and collect the supernatant. DL146 cultures were grown for 36 hours in the supernatant media and culture densities were measured by serial dilution and plating.

Congo Red clearing assay

Individual colonies were generated on LB agar containing cellulose and the plates were stained with Congo Red dye to observe clearings. Strains were incubated for 14 hours on LB Miller agar supplemented with kanamycin, ampicillin and 0.1% carboxymethylcellulose. Colonies were removed from plates by washing with 70% ethanol. Plates were stained with 0.1% Congo Red for 20 minutes and destained with 1 M NaCl in two rinse steps, destaining for 5 minutes and 30 minutes, respectively. Clearing diameter for each colony was quantified from bright-field images of the stained plates by measuring clearing diameter using ImageJ software. Clearing diameter was converted from pixels to millimeters by normalizing to known plate diameter.

Sequencing SDAc switching or lysis cheaters

Bright, monochromatic red or green fluorescent colonies isolated from six replicate cultures of $(T)_{18}/(A)_{10}$ grown in M9 minimal 0.4% PASC media for 72 hours were identified via blue light transillumination. Each colony was replated separately on an LB agar plate and an LB agar plate supplemented with kanamycin and ampicillin to test for plasmid loss. Sequencing was performed on individual colonies by Genewiz, Inc. Differentiation plasmids were sequenced at the dual promoter region (BSSregSeq_2809) and the *lacI* C-terminal sequence (lacIrptSeq_2810). Payload plasmids were sequenced across the payload transcription unit (oligos Vf2 & Vr). Sequencing oligos are listed in Supplementary Table 13.

Materials and Methods References

1. Zhang, Y. H. P., Cui, J., Lynd, L. R. & Kuang, L. R. A transition from cellulose swelling to cellulose dissolution by o-phosphoric acid: evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules* **7,** 644–648 (2006).
2. Wang, H. H. *et al.* Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460,** 894–898 (2009).
3. Egbert, R. G. & Klavins, E. Fine-tuning gene networks using simple sequence repeats. *Proc Natl Acad Sci USA* **109,** 16817–16822 (2012).
4. Hoover, D. M. & Lubkowski, J. DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res* **30,** e43 (2002).