SUPPORTING INFORMATION

Trade-offs in engineering sugar utilization pathways for titratable control

Taliman Afroz, Konstantinos Biliouris, Kelsey E. Boykin, Yiannis Kaznessis, and Chase L. Beisel

TABLE OF CONTENTS

DESCRIPTION OF SIMPLE MODEL

We developed a simple deterministic model comprising six ordinary differential equations to capture the dynamical behavior of utilization pathways that catabolize the inducing sugar. The model accounts for the regulator (R_t), the extracellular (S_0) and intracellular (S) sugar, two transporters (T_H, T_L), and one enzyme (E). Binding of S to R_t generates an active regulator (R) that upregulates the expression of the transporters and the enzyme following the Hill equation. The transporters import sugar into the cell while the enzyme irreversibly degrades the sugar. The model also allows for the breakdown of S_0 through sugar catabolism. The volume ratio of cells and medium converts between the amounts of intracellular and extracellular sugar. The following equations describe the amount of R (eq. 1), S_0 (eq. 2) and S (eq. 3), T_H and T_L, (eq. 4, 5), and E (eq. 6). The equations of the model follow mass-action and Michaelis-Menten kinetics, similar to previously published studies describing sugar utilization pathway.[1,2]

$$R = \underbrace{R_t \frac{S}{k_0 + S}}_{\text{active regulator}} \tag{1}$$

$$\frac{dS_0}{dt} = \underbrace{-\alpha_3 \frac{S}{k_3 + S} E}_{\text{catabolism of sugar}} \underbrace{\frac{V_C}{V_M}}_{\text{volume ratio}} \tag{2}$$

$$\frac{dS}{dt} = \underbrace{\alpha_1 \frac{S_0}{k_1 + S_0} T_L}_{\text{active transport of sugar}} + \underbrace{\alpha_2 \frac{S_0}{k_2 + S_0} T_H}_{\text{active transport of sugar}} - \underbrace{\alpha_3 \frac{S}{k_3 + S} E}_{\text{catabolism of sugar}} - \underbrace{d\,S}_{\text{dilution of sugar}} \tag{3}$$

$$\frac{dT_L}{dt} = \underbrace{b_{T_L}}_{\substack{\text{basal expression} \\ \text{of low-affinity transporter}}} + \underbrace{\alpha_4 \frac{R^n}{k_4^n + R^n}}_{\text{induction of low-affinity transporter by sugar}} - \underbrace{d\,T_L}_{\text{dilution of low-affinity transporter}} \tag{4}$$

$$\frac{dT_H}{dt} = \underbrace{b_{T_H}}_{\substack{\text{basal expression} \\ \text{of high-affinity transporter}}} + \underbrace{\alpha_5 \frac{R^n}{k_5^n + R^n}}_{\text{induction of high-affinity transporter by sugar}} - \underbrace{d\,T_H}_{\text{dilution of high-affinity transporter}} \tag{5}$$

$$\frac{dE}{dt} = \underbrace{b_E}_{\substack{\text{basal expression} \\ \text{of enzyme}}} + \underbrace{\alpha_6 \frac{R^n}{k_6^n + R^n}}_{\text{induction of enzyme by sugar}} - \underbrace{d\,E}_{\text{dilution of enzyme}} \tag{6}$$

The model does not explicitly account for DNA looping[1,2] or auto-repression of the regulator[3]; instead, the model assumes that the total amount of regulator is constant and DNA looping and regulator auto-repression were indirectly incorporated by manipulating the Hill coefficient in the

expression terms. The model also neglects differences in growth rate due to the small (~20%) difference in growth rate in the presence or absence of L-arabinose (data not shown). Each parameter in the model is described below.

Symbol	Description (Units)
R	Concentration of active regulator (M)
R_t	Concentration of total regulator (M)
S	Intracellular concentration of sugar (M)
S_0	Extracellular concentration of sugar (M)
t	Time (s)
V_C	Volume of the cell (M)
V_M	Volume of the medium (M)
T_L	Low-affinity transporter concentration (M)
T_H	High-affinity transporter concentration (M)
E	Enzyme concentration (M)
k_0	Half maximum regulator concentration (M)
k_1	Half-maximum concentration of sugar (M)
k_2	Half-maximum concentration of sugar (M)
k_3	Half-maximum concentration of sugar (M)
k_4	Half-maximum concentration of regulator (M)
k_5	Half-maximum concentration of regulator (M)
k_6	Half-maximum concentration of regulator (M)
b_{TL}	Basal expression of first transporter (M/s)
b_{TH}	Basal expression of second transporter (M/s)
b_E	Basal expression of enzyme (M/s)
α_1	Maximum rate of sugar import from low-affinity transporter (1/s)
α_2	Maximum rate of sugar import from high-affinity transporter (1/s)
α_3	Maximum rate of sugar catabolism (M/s)
α_4	Maximum rate of low-affinity transporter production (M/s)
α_5	Maximum rate of high-affinity transporter production (M/s)
α_6	Maximum rate of enzyme production (M/s)
d	Dilution rate due to cell division (1/s)
n	Hill coefficient

To reduce the total number of parameters, we first non-dimensionalized the variables as shown on the next page.

Dimensionless variable	Formula
R'	R/R_t
S'	S/k_0
S_0'	S_0/k_1
T_L'	$(d/\alpha_4)\cdot T_1$
T_H'	$(d/\alpha_5)\cdot T_2$
E'	$(d/\alpha_6)\cdot E$
τ	$t\cdot d$

Non-dimensionalizing the variables resulted in the following equations:

$$R' = \frac{S'}{1+S'} \tag{1b}$$

$$\frac{dS_0'}{d\tau} = -\frac{\alpha_3\alpha_6}{k_0 d^2}\frac{S'}{\frac{k_3}{k_0}+S'}E'\frac{k_0}{k_1}\cdot\frac{V_C}{V_M} \tag{2b}$$

$$\frac{dS'}{d\tau} = \frac{\alpha_1\alpha_4}{k_0 d^2}\frac{S_0'}{1+S_0'}T_1' + \frac{\alpha_2\alpha_5}{k_0 d^2}\frac{S_0'}{\frac{k_2}{k_1}+S_0'}T_2' - \frac{\alpha_3\alpha_6}{k_0 d^2}\frac{S'}{\frac{k_3}{k_0}+S'}E' - S' \tag{3b}$$

$$\frac{dT_L'}{d\tau} = \frac{b_L}{\alpha_4} + \frac{(R')^n}{\left(\frac{k_4}{R_t}\right)^n + (R')^n} - T_L' \tag{4b}$$

$$\frac{dT_H'}{d\tau} = \frac{b_H}{\alpha_5} + \frac{(R')^n}{\left(\frac{k_5}{R_t}\right)^n + (R')^n} - T_H' \tag{5b}$$

$$\frac{dE'}{d\tau} = \frac{b_E}{\alpha_6} + \frac{(R')^n}{\left(\frac{k_6}{R_t}\right)^n + (R')^n} - E' \tag{6b}$$

We next assigned single constants to the lumped parameters to generate the following set of equations:

$$R' = \frac{S'}{1 + S'} \tag{1c}$$

$$\frac{dS_0'}{d\tau} = -\alpha_E \frac{S'}{K_E + S'} E'\upsilon \tag{2c}$$

$$\frac{dS'}{d\tau} = \alpha_L \frac{S_0'}{1 + S_0'} T_1' + \alpha_H \frac{S_0'}{K_H + S_0'} T_2' - \alpha_E \frac{S'}{K_E + S'} E' - S' \tag{3c}$$

$$\frac{dT_L'}{d\tau} = b_L' + \frac{(R')^n}{K_{RL}^n + (R')^n} - T_L' \tag{4c}$$

$$\frac{dT_H'}{d\tau} = b_H' + \frac{(R')^n}{K_{RH}^n + (R')^n} - T_H' \tag{5c}$$

$$\frac{dE'}{d\tau} = b_E' + \frac{(R')^n}{K_{RE}^n + (R')^n} - E' \tag{6c}$$

The final constants along with the equivalent lumped parameters are shown in the table on the next page. We selected values for each constant to capture the general behavior of the L-arabinose utilization pathway: a high-affinity/low-capacity transporter expressed at lower sugar concentrations and a low-affinity/high-capacity transporter expressed at higher sugar concentrations ($\alpha_L > \alpha_H$, $K_L > K_H$, $K_{RL} > K_{RH}$), low basal expression (b_L', b_H', $b_E << 1$), proteins and sugar that are diluted through cell division, and sufficient protein expression and activity (α_E, K_E) to yield a bistable response.

Formula	Dimensionless parameter	Value
$\frac{\alpha_1\alpha_4}{k_0 d^2}$	α_L	300
$\frac{k_2}{k_0}$	K_L	0.2
$\frac{\alpha_2\alpha_5}{k_0 d^2}$	α_H	100
$\frac{k_3}{k_0}$	K_E	0.01
$\frac{\alpha_3\alpha_6}{k_0 d^2}$	α_E	0.6
ν	$\frac{k_0}{k_1}\cdot\frac{V_C}{V_M}$	0 – 0.01
$\frac{k_4}{R_t}$	K_{RL}	0.2
$\frac{k_5}{R_t}$	K_{RH}	0.005
$\frac{k_6}{R_t}$	K_{RE}	0.02
$\frac{b_{T_L}}{\alpha_4}$	b_L'	0.0002
$\frac{b_{T_H}}{\alpha_5}$	b_H'	0.0002
$\frac{b_E}{\alpha_6}$	b_E'	0.0002
n	n	2

These values were used in all MATLAB simulations unless indicated otherwise. The simulation results in Figure 2 neglected sugar breakdown, which involved removing the differential equation for extracellular sugar (eq. 2c). To capture constitutive expression of either transporter, we replaced the corresponding expression term $R'^2/(K_2 + R'^2)$ with the indicated value. To delete either transporter or the enzyme, we set the maximal activity term equal to zero (e.g. $\alpha_L = 0$). As part of the simulations that neglected sugar breakdown, we integrated the equations using Euler's

method and calculated the steady-state curves using the arc-length continuation method. When accounting for sugar breakdown, we integrated the equations to $\tau = 10$, which was sufficient to reach steady-state in the absence of breakdown. Altering the final value of τ did not change the qualitative impact of sugar breakdown.

SUPPORTING FIGURES

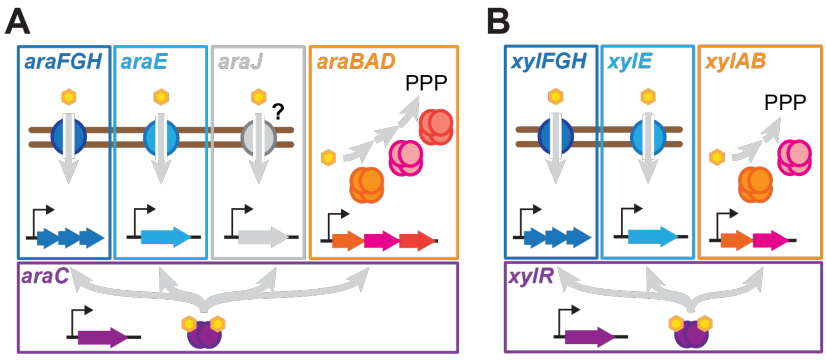

Figure S1. The L-arabinose and D-xylose utilization pathways in *E. coli*. (A) The pathway is composed of one regulator (*araC*), one high-affinity transporter (*araFGH*), one low-affinity transporter (*araE*), one putative transporter (*araJ*), and three enzymes (*araBAD*) that shunt L-arabinose into the pentose phosphate pathway (PPP). The AraC protein activates transcription at each promoter when bound to L-arabinose. Although the AraJ protein resembles a transporter, previous deletion of *araJ* had no measurable impact on the behavior of the pathway.[4] (B) The D-xylose utilization pathway is structured similarly to the L-arabinose utilization pathway, with a high-affinity transporter (*xylFGH*), a low-affinity transporter (*xylE*), two enzymes (*xylAB*) that shunt D-xylose into the pentose phosphate pathway, and a single regulator (*xylR*) responsible for activating the operons in the presence of D-xylose.

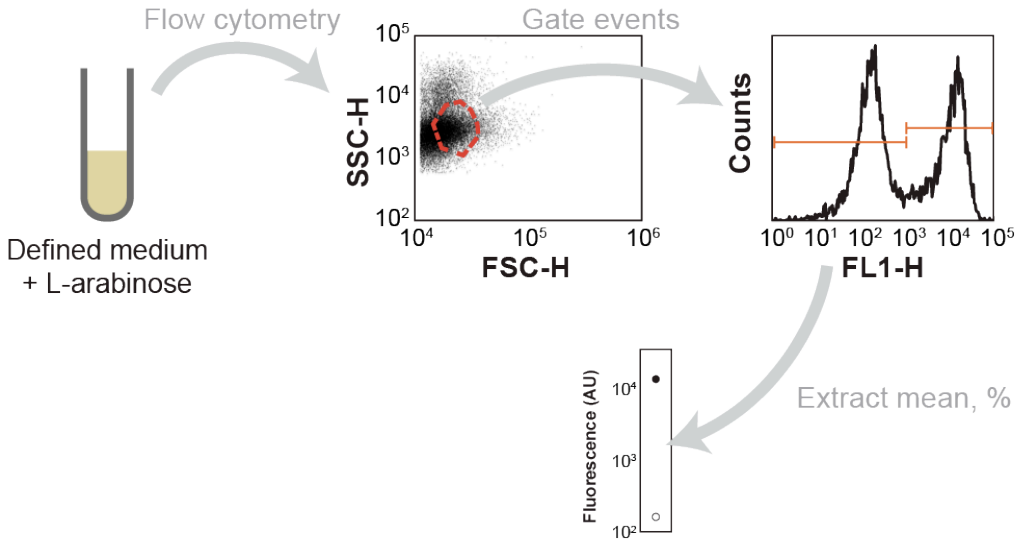

Figure S2. Overview of flow cytometry analysis. Cultures were diluted 1:100 in 1X PBS and run on a BD Accuri C6 personal flow cytometer. Cells were gated based on forward scatter (FSC-H) and side scatter (SSC-H). The gate (red) was selected to capture cells and no "noisy" events based on initial runs with DRAQ5 dye. Once gated, the mean fluorescence was extracted for unimodal distributions, while the mean fluorescence and relative number of events for each subpopulation were extracted for bimodal distributions.

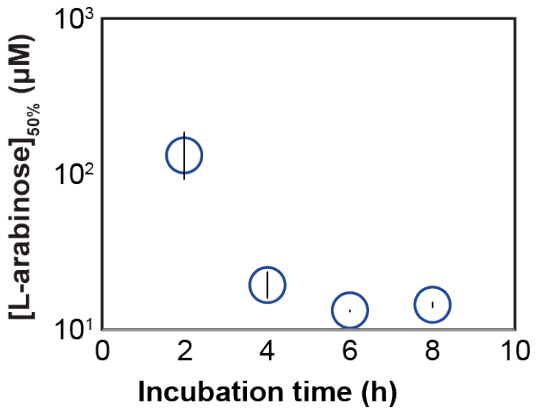

Figure S3. Impact of incubation time on the response to L-arabinose. Cultures of the WT strain (CB26) were grown in varying L-arabinose concentrations for the indicated amount of time to mid-log phase (ABS_{600} ~0.4) and then subjected to flow cytometry analysis. The concentration to achieve 50% induced cells was estimated by extracting the percentage of the population that was induced and fitting these values to a Hill equation. Note that this approach was selected over calculating the EC_{50} value because the response curve did not saturate at the maximal L-arabinose concentration that we used (10 mM). The reported concentration values represent the geometric mean and S.E.M. from at least three independent experiments. We selected an incubation time of 6 hours for all other experiments because increasing the incubation time had no obvious effect on the response.

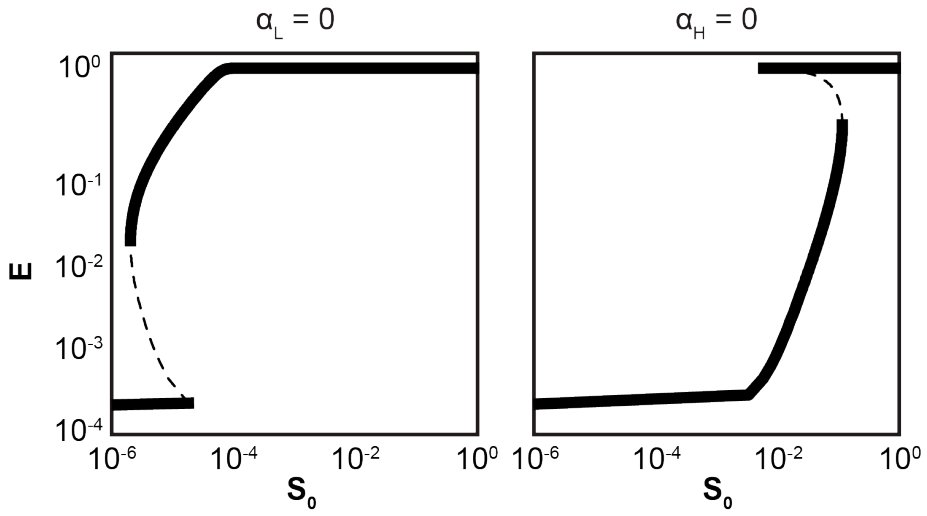

Figure S4. Predicted response when only one transporter is present. Model predictions reflect the base model in Figure 2 when the maximal activity of either transporter (α_H, α_L) is set equal to zero. Both transporters are sufficient for full induction of the pathway, where much lower concentrations of the extracellular sugar are needed for the high-affinity transporter than for the low-affinity transporter. Note that all variables were non-dimensionalized as part of the model derivation.

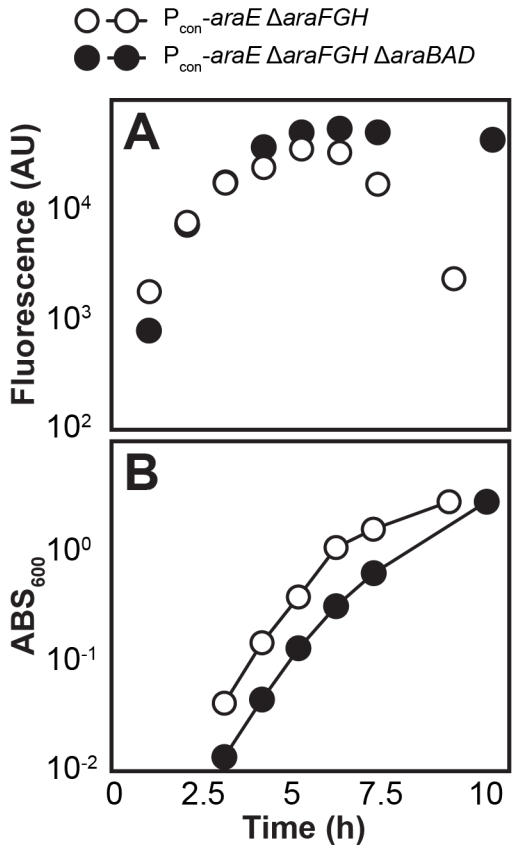

Figure S5. Response dynamics in the presence or absence of L-arabinose catabolism. The P_{con}-*araE* Δ*araFGH* and P_{con}-*araE* Δ*araFGH* Δ*araBAD* strains harboring the pUA66-ParaB plasmid were grown in M9 minimal medium with 292 μM or 8 μM of L-arabinose, respectively. These L-arabinose concentrations reflect the approximate EC_{50} value for each modified pathway. The ABS_{600} (A) and single-cell fluorescence (B) were measured at the indicated times. The values represent the mean and S.E.M. from experiments conducted with three independent colonies. The error bars are smaller than each symbol at all time points. The ABS_{600} was too low to accurately measure prior to 3 hours.

SUPPORTING TABLES

Table S1. Measured growth rates for each strain. Measurements were performed in M9 glycerol with 0.2% casamino acids and either no L-arabinose (- ara) or 10 mM L-arabinose (+ ara) during exponential growth.

Alteration	Growth rate[a]	
	- ara	**+ ara**
None (WT)	38.6 ± 1.2	34.2 ± 0.7
Δ*araBAD*	38.9 ± 0.2	48.4 ± 0.4
P_{con}-*araFGH*	38.7 ± 0.8	37.8 ± 1.6
P_{con}-*araFGH* Δ*araBAD*	35.7 ± 0.8	44.6 ± 0.6
P_{con}-*araFGH* Δ*araE*	32.2 ± 1.2	33 ± 1.9
P_{con}-*araFGH* Δ*araE* Δ*araBAD*	36.9 ± 1.4	37.4 ± 0.5
P_{con}-*araE*	41.5 ± 2.5	40.3 ± 2.3
P_{con}-*araE* Δ*araBAD*	37.9 ± 1.2	42.0 ± 1.3
P_{con}-*araE* Δ*araFGH*	34.5 ± 0.5	31 ± 0.6
P_{con}-*araE* Δ*araFGH* Δ*araBAD*	37.7 ± 0.5	41.6 ± 0.8

[a] Growth rates reflect the mean and S.E. from the measurement of three independent colonies.

Table S2. Strains used in this study

Name	Genotype	Stock #
WT	*Escherichia coli* K-12 substrain MG1655	CB168
Δ*araBAD*	MG1655 Δ*araBAD::cat*	CB225
NM400	MG1655 mini-λ-*cat*	CB25
NM500	MG1655 mini-λ-*tetA*	CB26
P_{con}-*araE*	MG1655 P_{araE}::[*cat* P_{con}]	CB325
P_{con}-*araE* Δ*araBAD*	MG1655 P_{araE}::[*cat* P_{con}] Δ*araBAD*	CB326
P_{con}-*araE* Δ*araFGH*	MG1655 P_{araE}::[*cat* P_{con}] Δ*araFGH::bla*	CB327
P_{con}-*araE* Δ*araFGH* Δ*araBAD*	MG1655 P_{araE}::[*cat* P_{con}] Δ*araBAD* Δ*araFGH::bla*	CB328
P_{con}-*araFGH*	MG1655 P_{araF}::[*cat* P_{con}]	CB329
P_{con}-*araFGH* Δ*araBAD*	MG1655 P_{araF}::[*cat* P_{con}] Δ*araBAD*	CB330
P_{con}-*araFGH* Δ*araE*	MG1655 P_{araF}::[*cat* P_{con}] Δ*araE::bla*	CB331
P_{con}-*araFGH* Δ*araE* Δ*araBAD*	MG1655 P_{araF}::[*cat* P_{con}] Δ*araBAD* Δ*araE::bla*	CB332
P_{con}-*xylFGH*	MG1655 P_{xylF}::[*cat* P_{con}]	CB342
P_{con}-*xylFGH* Δ*xylAB*	MG1655 [Δ*xylAB* P_{xylF}]::[*cat* P_{con}]	CB343

Table S3. Plasmids used in this study.

Name	Description	Source	Stock #
pUA66	Low-copy plasmid (pSC101 ori) with MCS upstream of *gfp*	OpenBiosystems	pCB198
pUA66ParaB	pUA66 with the *araB* promoter	OpenBiosystems	pCB208
pUA66PxylA	pUA66 with the *xylA* promoter	OpenBiosystems	pCB289
pKD3	Plasmid encoding *cat* cassette flanked by FRT sites and a *bla* cassette.	Ref. 5	pCB333
pKD46	L-arabinose-inducible expression of λ-red genes on a plasmid with a heat-sensitive origin-of-replication	Ref. 5	pCB13
pCP20	Heat-inducible expression of FRT on a plasmid with a heat-sensitive origin-of-replication	Ref. 6	pCB44

Table S4. Oligonucleotides used in this study. Bases shown in red are part of the synthetic promoter (J23110) are shown red. Bolded bases are the homology regions for recombination into MG1655.

Name	Sequence
del-araBAD.fwd	**TCCATACCCGTTTTTTTGGATGGAGTGAAACG**CATGGTCCATATGAATATCCTCCTTAG
del-araBAD.rev	**GTTTGCTGCATATCCGGTAACTGCGGCGCTAACTGACGGCAG**GTAGGCTGGAGCTGCTT
ParaE.fwd	**TCTGCTGTAAAATTAGGTGGTTAATAATAATCTCAATAATTCA**GTAGGCTGGAGCTGCTT
ParaF.fwd	**CCATTTCAAAAACTCAACGCCAGGTAATGCGGCCTATTGACTG**GTAGGCTGGAGCTGCTT
J23110.rev	**GCTAGCATTGTACCTAGGACTGAGCTAGCCGTAAA**CATATGAATATCCTCCTTAG
ParaE-J23110.rev	**GATAGTAACCATTTTTTCCTGCCAGCAGAGAGTAAGAC**GCTAGCATTGTACCTAGGACTG
ParaF-J23110.rev	**AAAGCATTACCTTTTAACTAAAAGATAAGTGACTGTGT**GCTAGCATTGTACCTAGGACTG
del-araE.fwd	**TGTTCACGTATTTTTTCACTATGTCTTACTCTCTGCTGGC**TGATGCCTGGCAGTTCCCTA
del-araE.rev	**GACGCCGATATTTCTCAACTTCTCGCCTGCCATCAGTT**TTACCAATGCTTAATCAGTGAG
del-araFGH.fwd	**TCATTCGTTTTTGCCCTACACAAAACGACACTAAAGCTGG**TGATGCCTGGCAGTTCCCTA
del-araFGH.rev	**GACAGTGCGTTTCGCTTTTTGCTTGTAACGGTCGAAGA**TTACCAATGCTTAATCAGTGAG
PxylA.fwd	**TATCACGAGGCCCTTTCGTCTTCAC**GGTGTAGGGCCTTCTGTAGT
PxylA.rev	**TGTATATCTCCTTCTTAAATCTAGA**TTACATACGCTTACCCAACACCA
PxylF.fwd	**GCAATAAGTACAATTGCGCAACAAAAGTAAGATCTCGGTCATAG**TAGGCTGGAGCTGCTT
PxylF-J23110.rev	**GGACAGTTTTAATAAGTAACAATCACCGCGATAAACGT**GCTAGCATTGTACCTAGGACTG
del-xylAB.rev	**GGCAACAATTCAATGAGCGATTTCTCTGGATTCGCCGCGATC**GTAGGCTGGAGCTGCTT

SUPPORTING REFERENCES

(1) Kuhlman, T.; Zhang, Z.; Saier, M. H., Jr; Hwa, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6043–6048.

(2) Huo, L.; Martin, K. J.; Schleif, R. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5444–5448.

(3) Madar, D.; Dekel, E.; Bren, A.; Alon, U. *BMC Syst Biol* **2011**, *5*, 111.

(4) Reeder, T.; Schleif, R. *J. Bacteriol.* **1991**, *173*, 7765–7771.

(5) Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6640–6645.

(6) Cherepanov, P. P.; Wackernagel, W. *Gene* **1995**, *158*, 9–14.