2 classes

- model-driven design, Jim Collins

incremental model construction and calibration

- from "a bottom-up approach to gene regulation"

Nature (2006) vol. 439 (7078) pp. 856-860

the idea of the paper: model-driven design

incremental modeling with careful calibration at each stage

- both a deterministic (uses a thermo model as we have studied earlier)

- and stochastic model (next class)

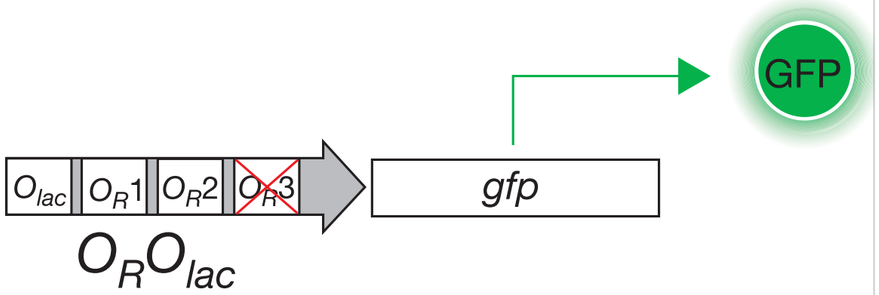
model of (random exponential with 20 min doubling time) growth and division (with binomial for plasmid allocation to daughter cells)

incude the dynamics of plasmid duplication (linked to plasmid's origin of replication ...)

(prok) transcriptional logic - Terminology

- [ ] plasmid copy numbers - origin of replication (type)
- [ ] dose-response curve
- [ ] Gamma distribution
- [ ] CV = coefficient of variation = $Var(X)^{1/2}/E(X)$
- [ ] noise sources: growth and division, synthesis machinery

a deliberate simplification



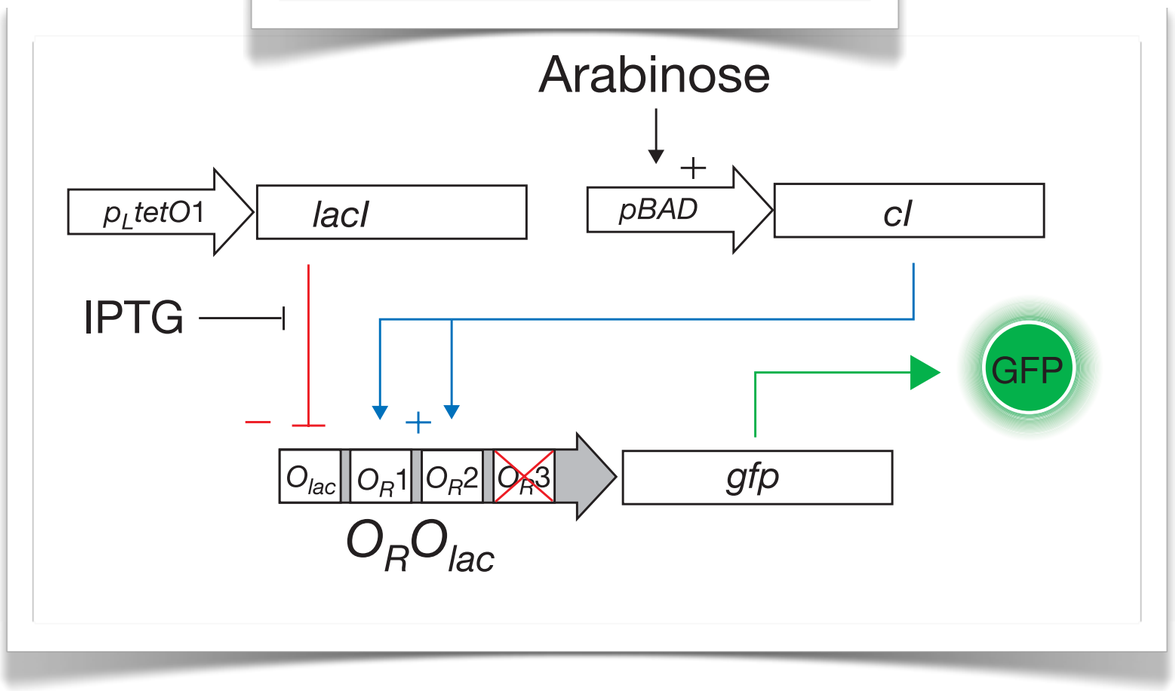

```
TGTTGTGTGGAATTGTGAGCGGATAACAATTCACACACCCGGCCGGGGCAACCATTATCACCGCCAGAGGTAAAATAGTCAACACGCACGGTGTTAGATATTTATCCCTTGTGGTGATAGATT
ACAACACACCTTAACACTCGCCTATTGTTAAGTGTGTGGGCCGGCCCCGTTGGTAATAGTGGCGGTCTCCATTTTATCAGTTGTGCGTGCCACAATCTATAAATAGGGAACACCACTATCTAA
```

Olac OR1 OR2 -35 OR3 -10

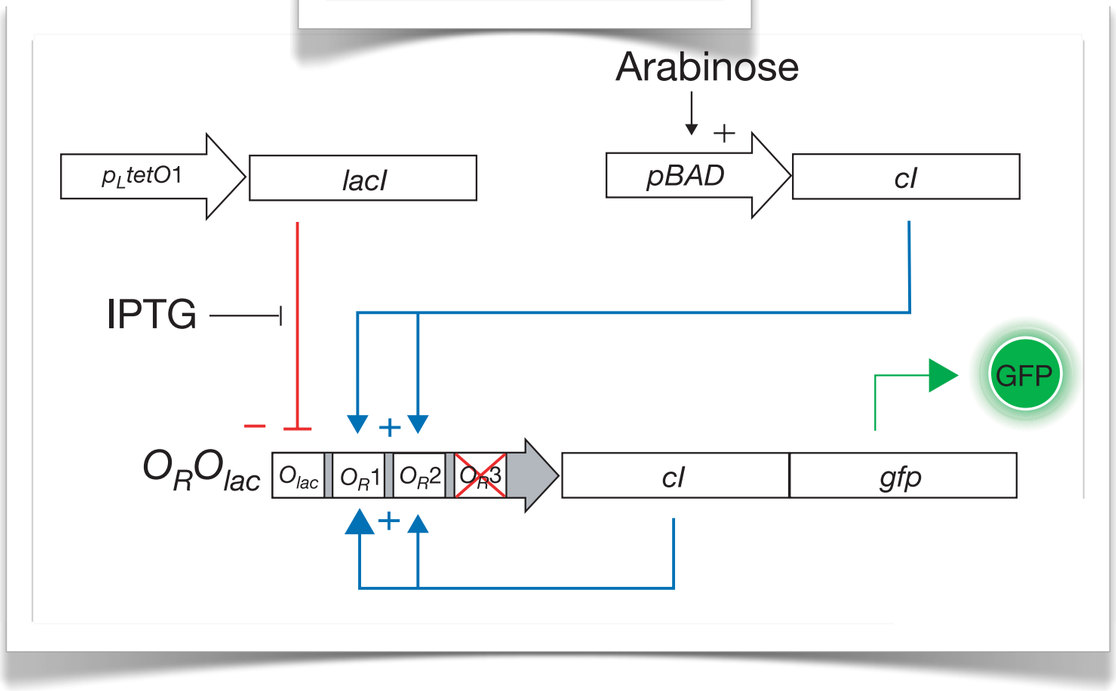
Supplementary Figure 12: The promoter sequence for the O_RO_{lac} promoter. The lac operator site (OLac) and the three λ sites (O_R1, O_R2 and the mutated O_R3) are indicated in the grey boxes. The blue boxes indicate the -10 and -35 RNA polymerase binding sites.

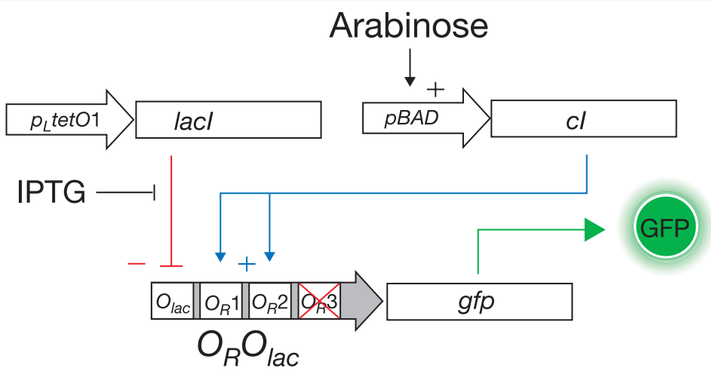
4 + 1 scenarios

baseline, act, rep, act+rep



positive feedback





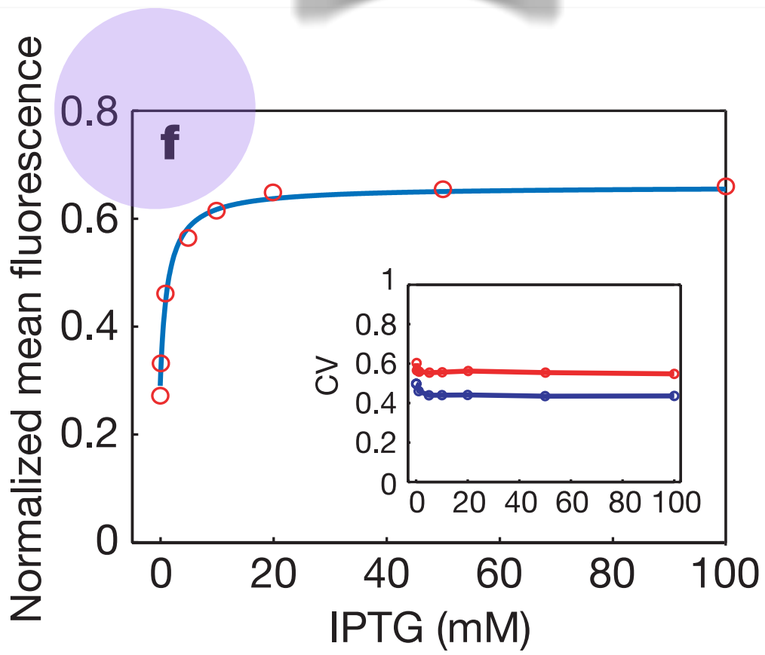
4 data sets

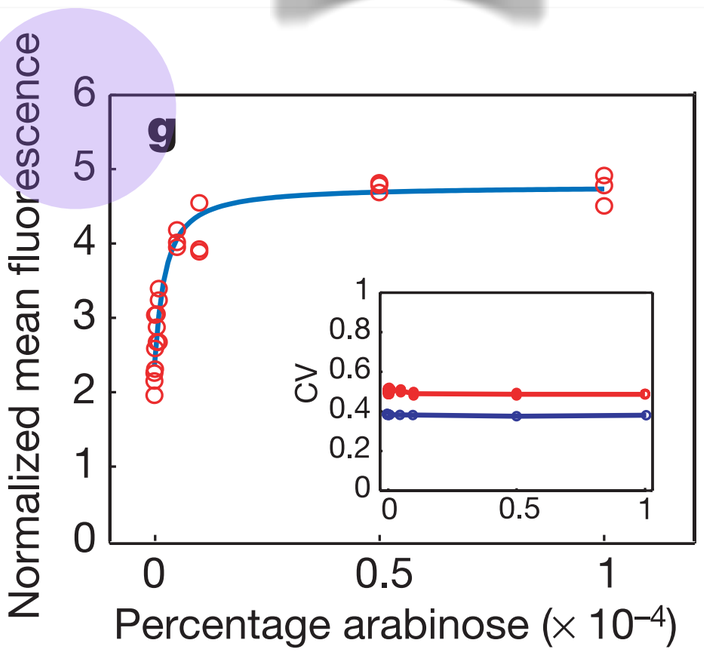
dose-response curves/ normalised to the baseline case

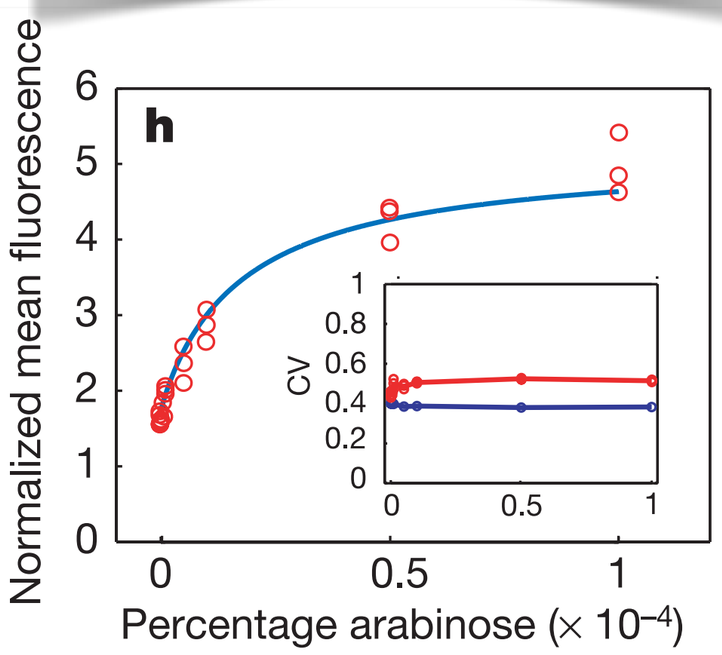
rep only

act only

act and rep - IPTG=10nM







model vs. data - strategy:

we want to compute the equilibrium transcription rate, relative to baseline rate (with no TFs at all)

supposing fluorescence is proportional to GFP numbers, and GFP numbers are proportional to said rate; the rate ratio should be the fluo ration: <--- mapping model output to data

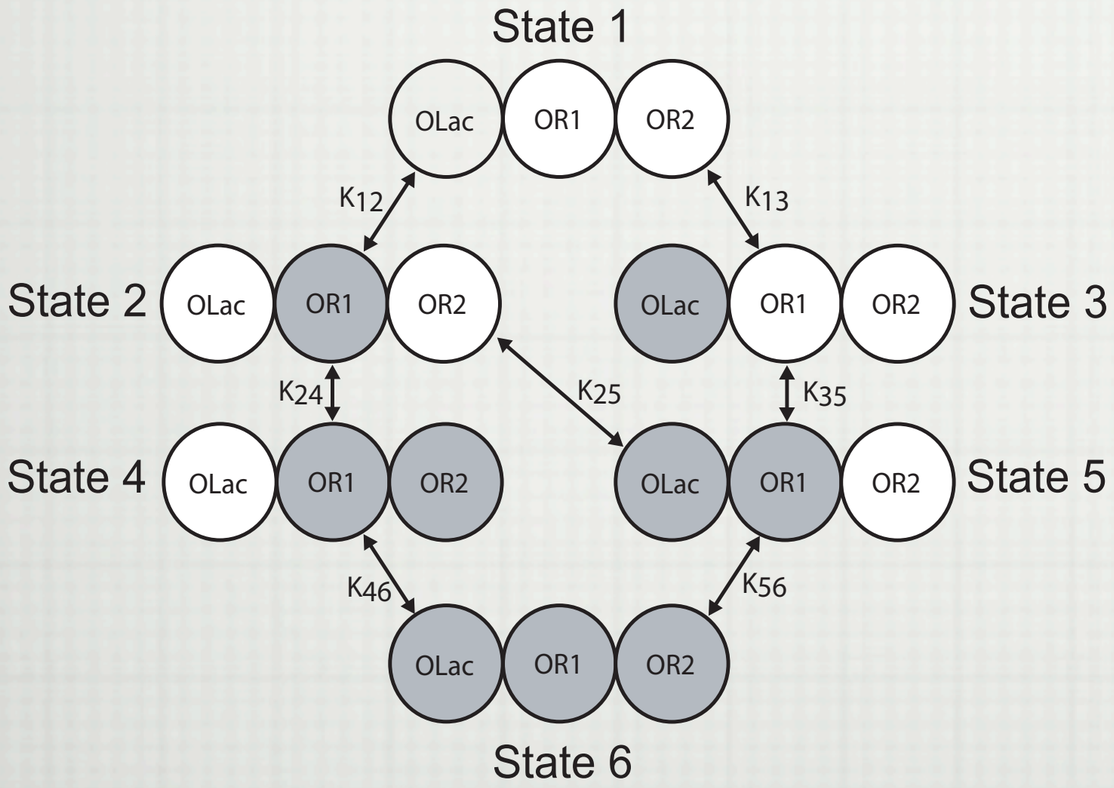
we need to:

1- compute the eq. probability of occupancy of the promoter (as a function of the TFs' concentration) in each 3 cases, rep, act and both TFs

2- compute the TF's concentration in each 3 cases as a function of the inputs (IPTG, arabinose) <--- mapping data inputs to model

occupancy states of the promoter, equilibrium probability

O_R "sequential", 6 states only



we suppose (active) TFs A, B are in constant nb n(A), n(B),...

$$p(i)/p(j=i+A) = n(i)/n(i+A)\ \exp(-\beta(E(i)-E(j))) = 1/n(A)\exp(-\ldots) = K_{i,A}/[A] =: K_{ij}$$

this gives p(i) as a function of [A], [B], the active TFs, hence the promoter activity: <gamma(i);p(i)>

equilibrium for active TFs: the repressor-only case

repressor-only:

LacI is a 4-mer, comes in 2 kinds:

- with IPTG, TI (weaker binding to O_{lac}, increases the dissoc rate),
- and without T (stronger)

[I]: short for IPTG concentration

$[I] \sim [I]_{tot}$

f fraction of 4-mer w/o IPTG $= [T]/[T]_{tot}$

$K_{d,TI}[TI] = [T][I]$ (1 new param)

$f = K_{d,TI}/(K_{d,TI} + [I])$

k_{31} modified off-rate for TI:O_{lac}

$= k'_{31}(f + (1-f)*alpha)$

$= k'_{31}(K_{d,TI} + alpha\ [I])/(K_{d,TI} + [I])$

NB: alpha>1, since IPTG inactivates the repressor (1 new param)

$$p(3)/p(1) = [T]/K_{d,TP} * (K_{d,TI} + [I])/(K_{d,TI} + alpha*[I])$$

equilibrium for active TFs: the activator-only case

activator-only: cI_2 a dimer under the indirect influence of arabinose [A]

assume

$[cI2] = [cI2]_0 + s\,[A]$ (2 new params)

$$p(2)/p(1) = [cI2]/K_{1,cI}$$
$$= ([cI2]_0 + s\,[A])/K_{1,cI}$$

equilibrium for active TFs: the repressor and activator case

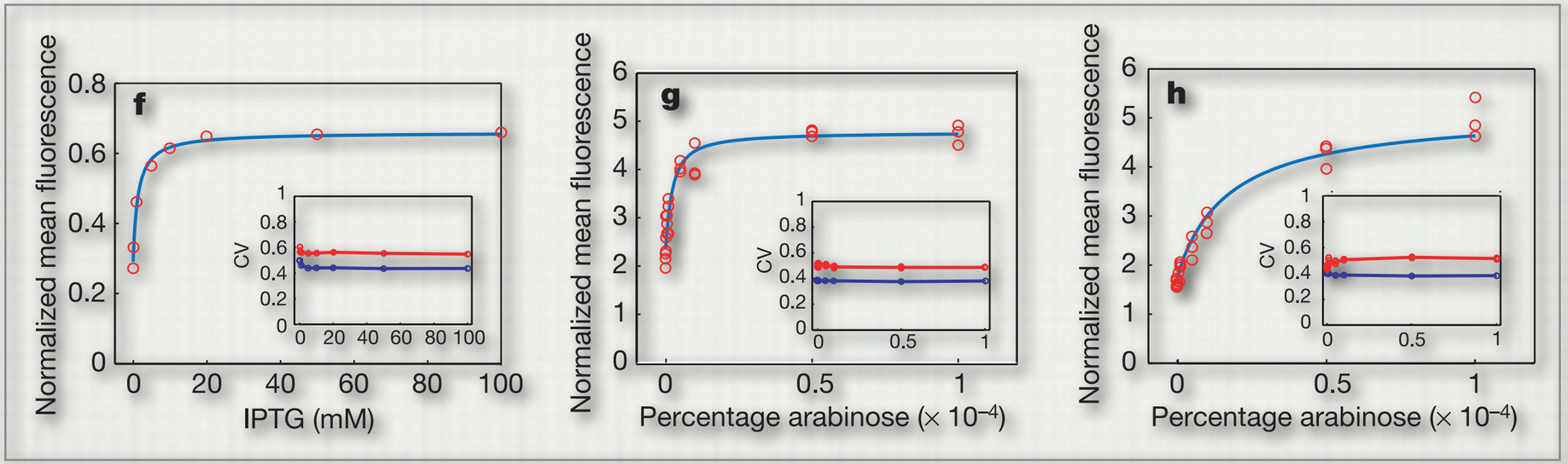
- combine the two preceding cases

no new parameter is needed

$p(i)/p(i+A) =$ ... complicated expression

but we know how to write it

- fit parameters to 3 data sets



parameters 1: promoter equilibrium

Parameter	Description	Value	mean	std dev
K_{13}^{eq}	Equilibrium constant for LacI/IPTG with no CI	0.93 nM^{-1}	1.24	4.21×10^{-01}
K_{12}^{eq}	Equilibrium constant for first CI site with no LacI bound	0.006 nM^{-1}	5.97×10^{-03}	1.43×10^{-04}
K_{24}^{eq}	Equilibrium constant for the second CI site with no LacI bound	0.00138 nM^{-1}	1.41×10^{-02}	4.59×10^{-03}
K_{35}^{eq}	Equilibrium constant for the first CI site with LacI bound	0.0117 nM^{-1}	1.12×10^{-01}	6.80×10^{-03}
K_{46}^{eq}	Equilibrium constant for LacI bound with 2 CI bound	0.00444 nM^{-1}	6.27×10^{-01}	2.13×10^{-01}

parameters 2: promoter activities

g_2	Relative production rate for promoter state S_2	1	1.00	5.64×10^{-03}
g_3	Relative production rate for promoter state S_3	0.292	2.92×10^{-01}	8.67×10^{-04}
g_4	Relative production rate for promoter state S_4	4.78	4.79	1.03×10^{-02}
g_5	Relative production rate for promoter state S_5	1.31	1.30	3.84×10^{-03}
g_6	Relative production rate for promoter state S_6	3.48	3.48	1.82×10^{-02}

parameters 3: IPTG vs lacI

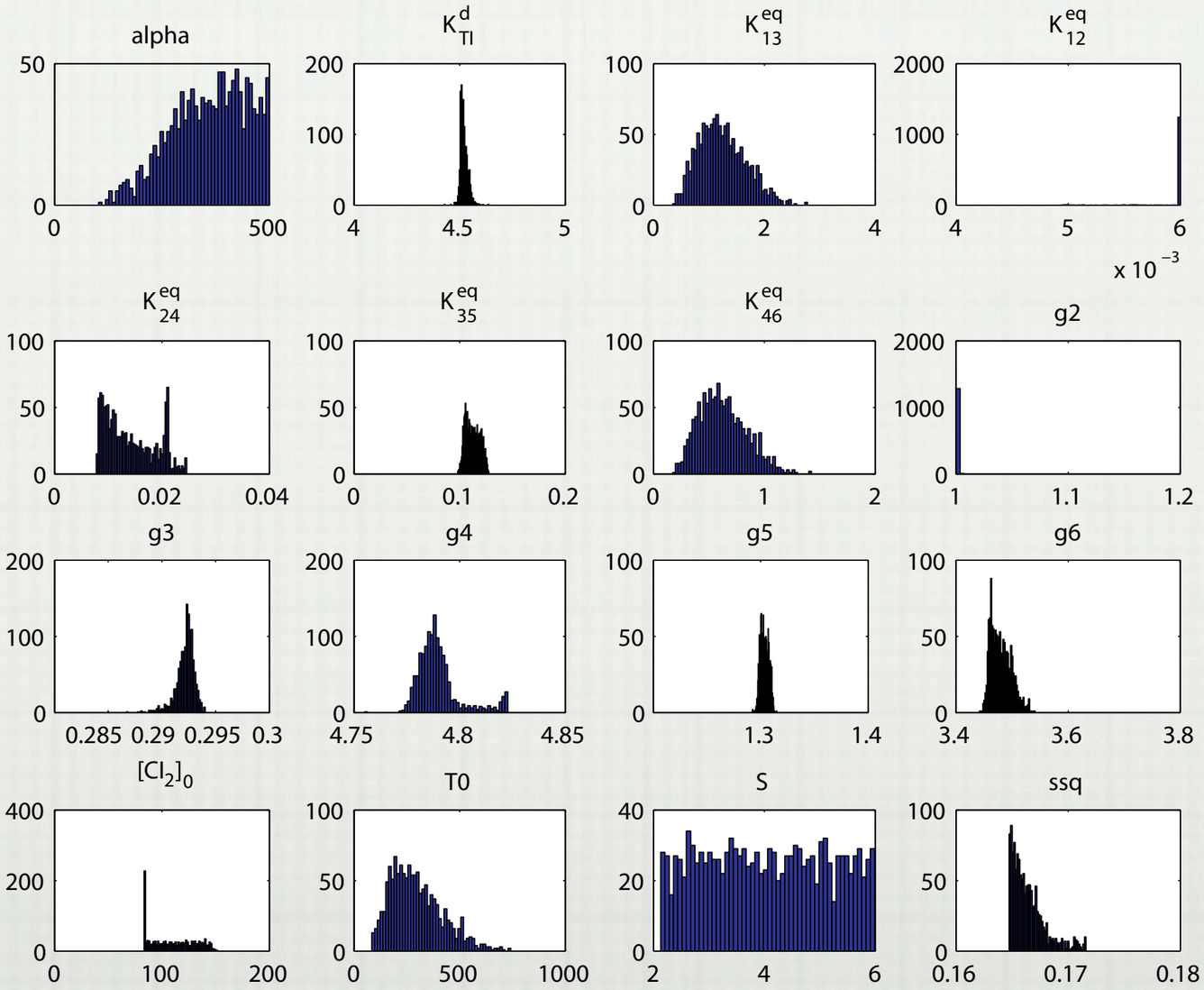
α	Ratio of LacI off rates with and without IPTG	330	3.59×10^{02}	8.82×10^{01}
K_{TI}^{d}	Equilibrium constant for LacI/IPTG	4.52 nM	4.52	1.08×10^{-02}

T0	LacI tetramer concentration	$325nM$	2.96×10^{02}	1.22×10^{02}

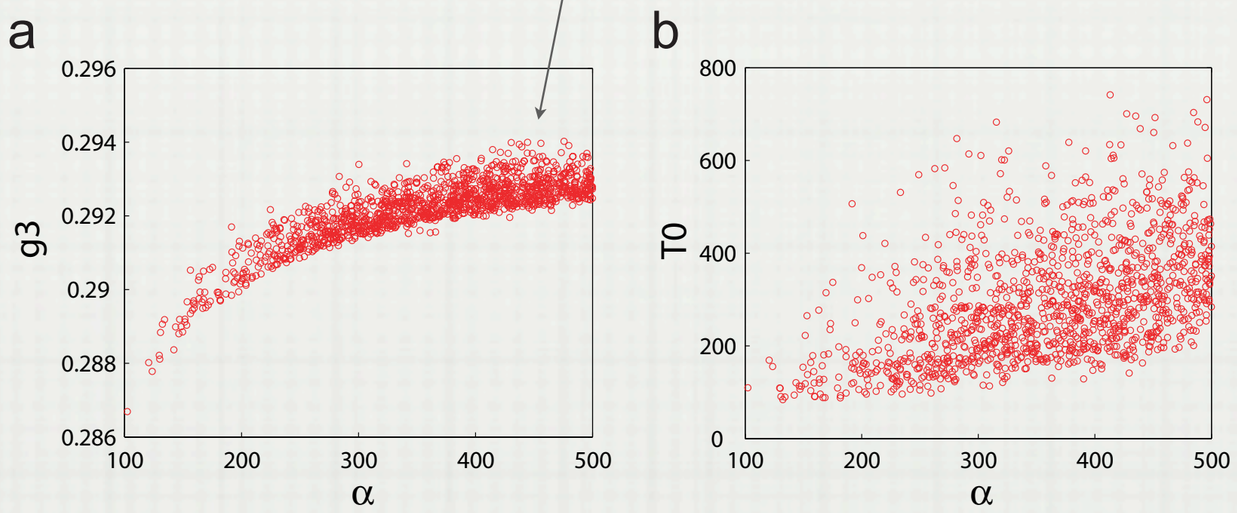
parameters 4: arabinose vs cI

$[CI_2]_0$	CI dimer concentration with no arabinose	$105nM$	1.10×10^{02}	2.08×10^{01}
S	Coefficient related to CI induction	3.85e7	4.05×10^{07}	1.11×10^{07}

parametrizations - best fit



g3 becomes constant in large alphas



Supplementary Figure 3: Scatter plots of the parameter α with the parameters g_3 and $T0$.

- model-driven design 2, the stochastic case

1- inputs -> steady state of TF concentrations, cI2, TI and T

2- TF concentrations -> transition rates (Q matrix) of the promoter CTMC

3- CTMC state -> transcription rate for mRNA -> translation GFP

4- V(t) random growth volume with exponential law of which mean V(t) = V(0) exp(-ln 2 t) (doubling time 1, so time unit = cell cycle; + binomial for plasmid allocation to daughter cells)

5- high copy plasmid Gamma(alpha,beta): mean = alpha*beta = 50, var = alpha*beta^2 fitted (50 comes from plasmid's origin of replication ...)

sources of noise

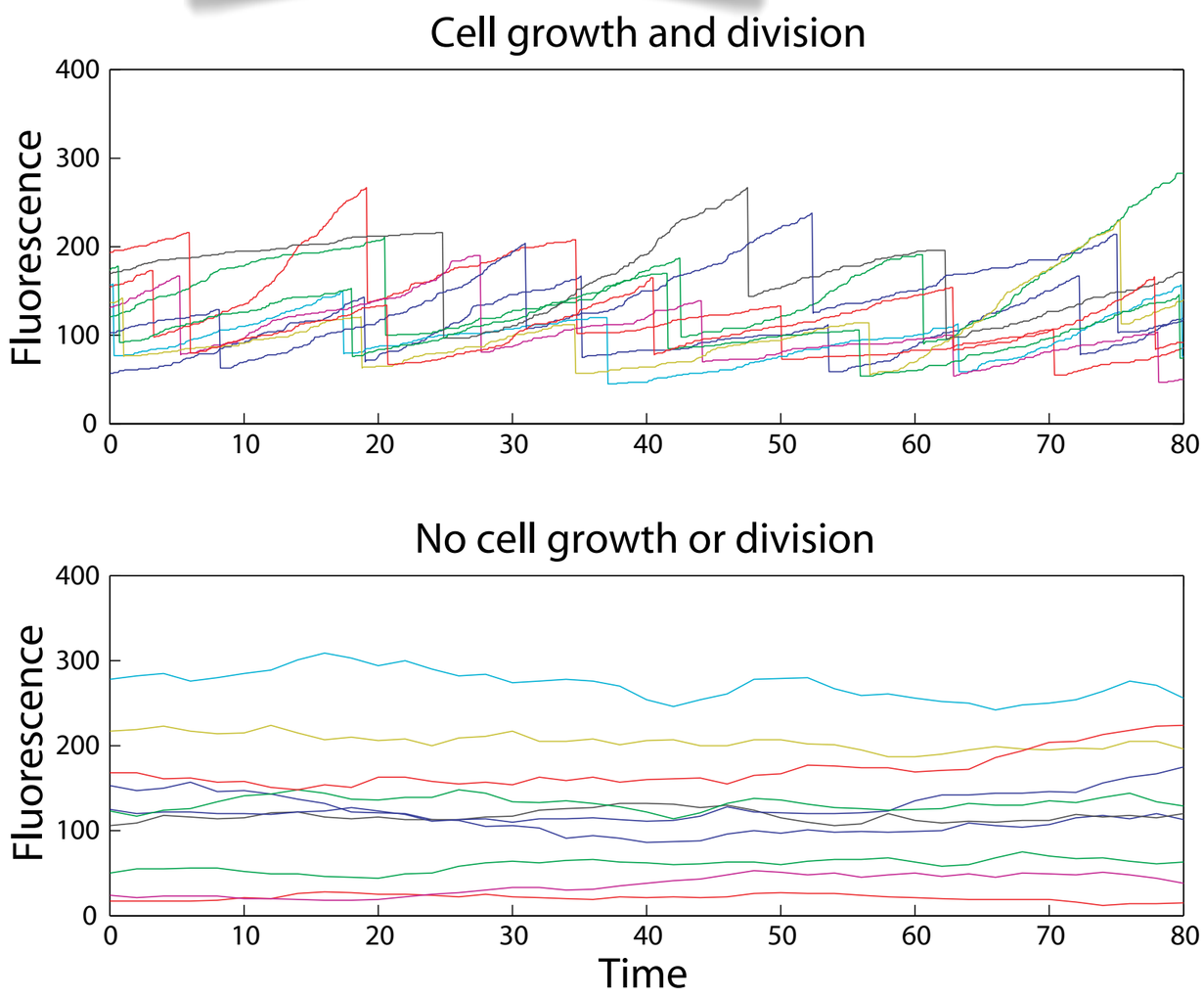
1- transitions of the promoter (negligible)

2- transcription/translation

3- V(t) random growth

4- binomial allocation of mRNAs, plasmids, GFPs

growth & division



Supplementary Figure 8: Model simulations of the repressor-activator system with a high copy plasmid and 5×10^{-6} % arabinose. **a**, Fluorescence level of cells with growth and division. **b**, Fluorescence level of cells without growth or division.

parameters 2: additional stochastic parameters

promoter Q matrix

Parameter	Description	Value
k_{21}	CI dissociation rate	100
k_{42}	CI dissociation rate	100
k_{53}	CI dissociation rate	10
k_{65}	CI dissociation rate	100
k_{31}	LacI dissociation rate with IPTG	10
k_{52}	LacI dissociation rate with IPTG	1
k_{64}	LacI dissociation rate with IPTG	100
γ	Constitutive mRNA production rate	0.5
δ_m	mRNA degradation rate	3.5
γ_g	Rate constant for GFP synthesis	10
δ_g	Rate constant for GFP degradation	0
k_v	Rate constant for volume growth	0.693147
α, β	gamma distribution parameters	4, 12.5

transcription,

translation

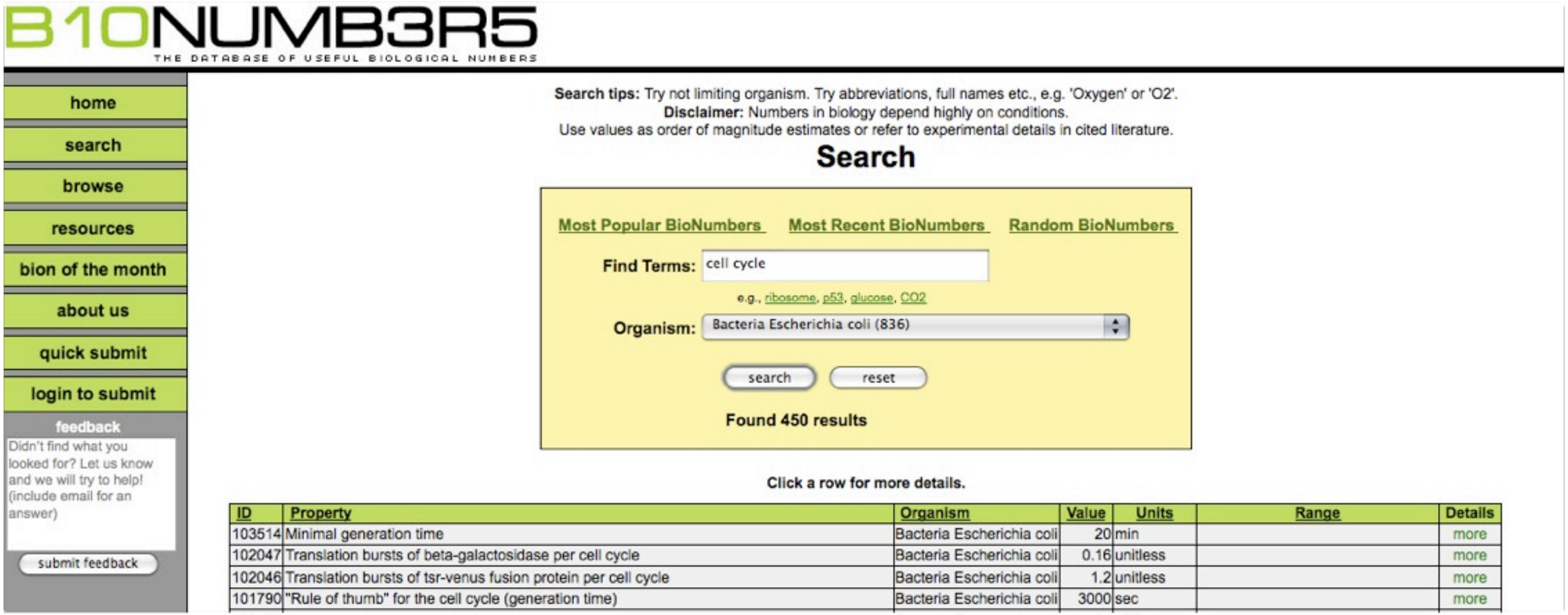
time unit = cell cycle = 20 minutes

volume/plasmid evolution

parameters 3: additional parameters for stochastic model of FB model

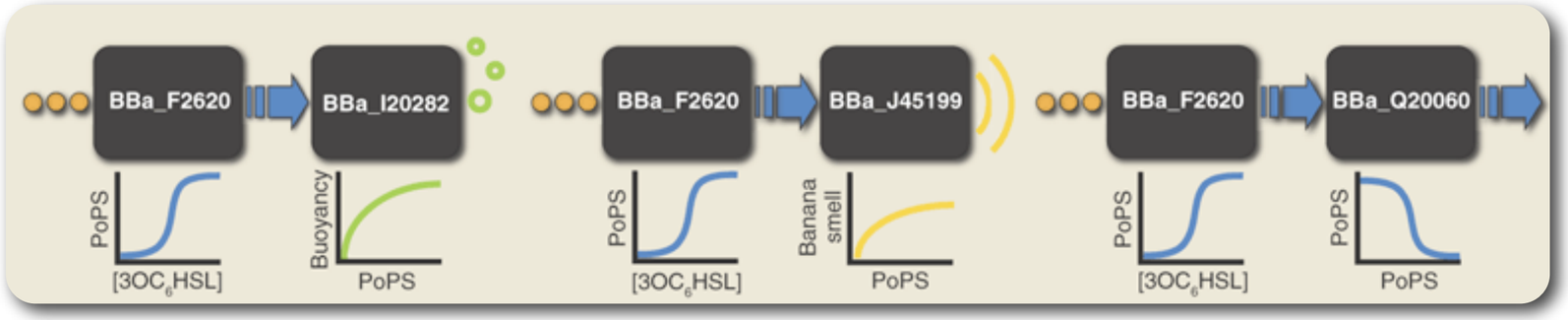
Parameter	Description	Value
γ'	Constitutive CI mRNA synthesis rate from OROLac	3
pCI	CI mRNA synthesis rate from $pBAD$	$0.006 * (CI_0 + s * ARA) * kb/kf$
dCI	CI mRNA degradation rate	3.5
$pLacI$	LacI mRNA synthesis rate	3.4
$dLacI$	LacI mRNA degradation rate	3.5
γ_{CI}	CI synthesis rate	13.5
δ_{CI}	CI degradation rate	0
γ_{LacI}	LacI synthesis rate	17.5
δ_{LacI}	LacI degradation rate	0
kf_1	CI association rate	1
kb_1	CI dissociation rate	100
kf_2	LacI association rate	1
kb_2	LacI dissociation rate	100
kf_3	LacI2 association rate	1
kb_3	LacI2 dissociation rate	10

digression: use bionumbers! (bionumbers.hms.harvard.edu)



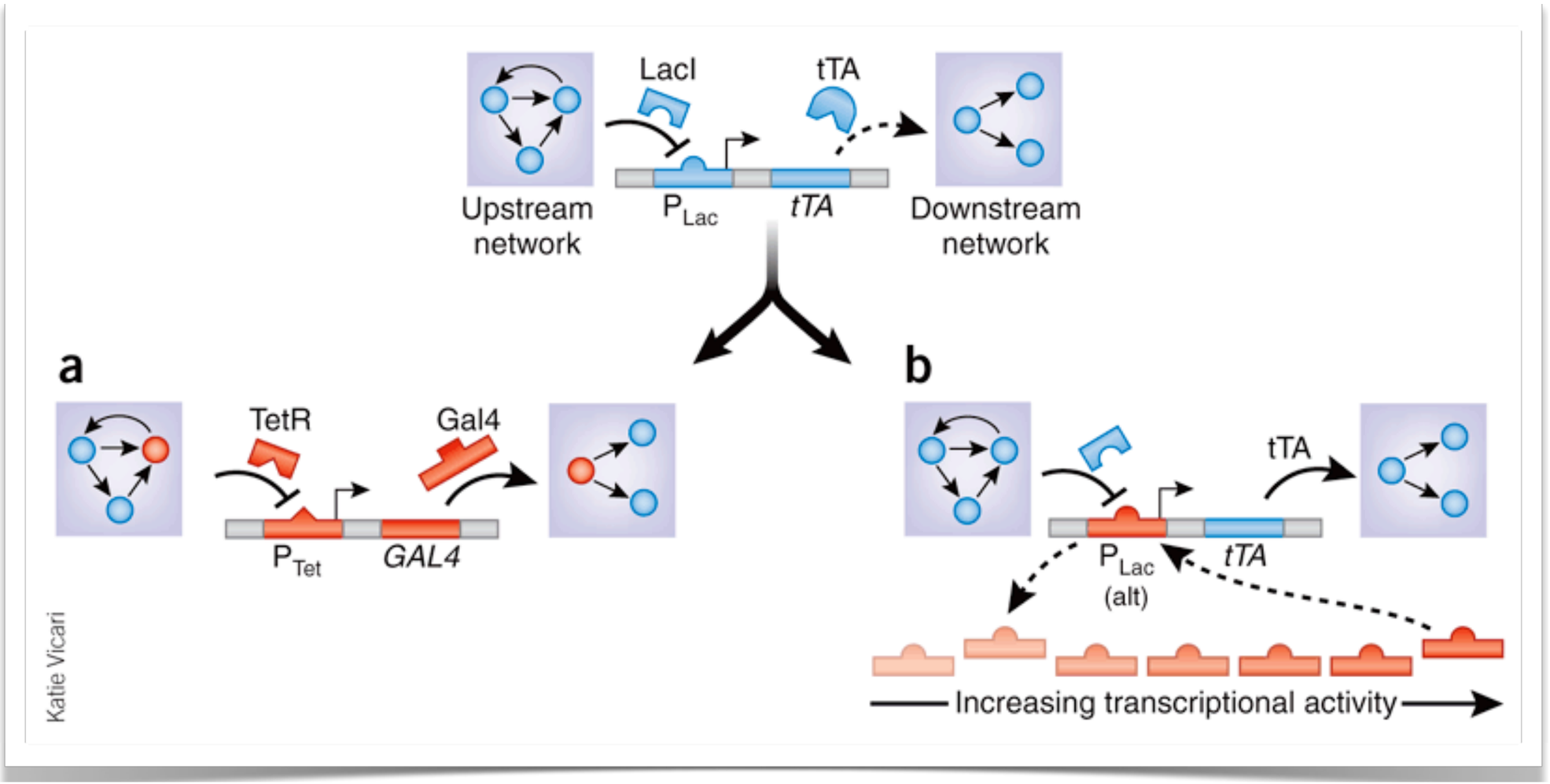
ID	Property	Organism	Value	Units	Range	Details
103514	Minimal generation time	Bacteria Escherichia coli	20	min		more
102047	Translation bursts of beta-galactosidase per cell cycle	Bacteria Escherichia coli	0.16	unitless		more
102046	Translation bursts of tsr-venus fusion protein per cell cycle	Bacteria Escherichia coli	1.2	unitless		more
101790	"Rule of thumb" for the cell cycle (generation time)	Bacteria Escherichia coli	3000	sec		more

physical vs functional composition



pops

impedance matching - how it is useful to have many versions of a promoter



- The upstream network must be reconfigured to produce TetR instead of LacI
- the downstream network must receive a new input Gal4p (Fig. 1a) because TetR and tTA interfere