Elementary Gene Circuits: Design Principles and Context-Dependent Expression

Michael A. Savageau
The University of Michigan
April 2001

The problems faced by pre- and post-genomic genetics are ... much the same -- they all involve bridging the chasm between genotype and phenotype.

-- Sydney Brenner, *Science* **287**:2173-4 (2000)

Human Genome Sequence

- Fewer genes than expected relative to number of proteins
- Importance of intervening circuitry and regulation
- Evolution and robustness of regulatory mechanisms

Outline

- Background
  - Gene Circuitry
  - Molecular Foundation
  - Systems Goal
  - Two Ways of Approaching Function
- Examples of Design Principles
- Integration and Evolution

Background

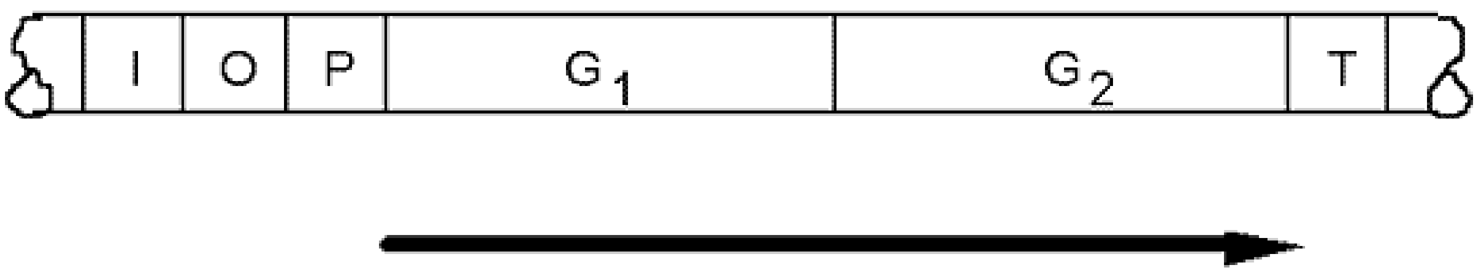
Function of Gene Circuitry

- Superficial answer
  - Genotype determined by the information encoded in the DNA sequence
  - Phenotype by the context-dependent expression of the genome
  - Circuitry interprets context and orchestrates expression
- Deeper answer
  - Hierarchy of mechanisms
  - Diversity of design issues
  - Accident and rule

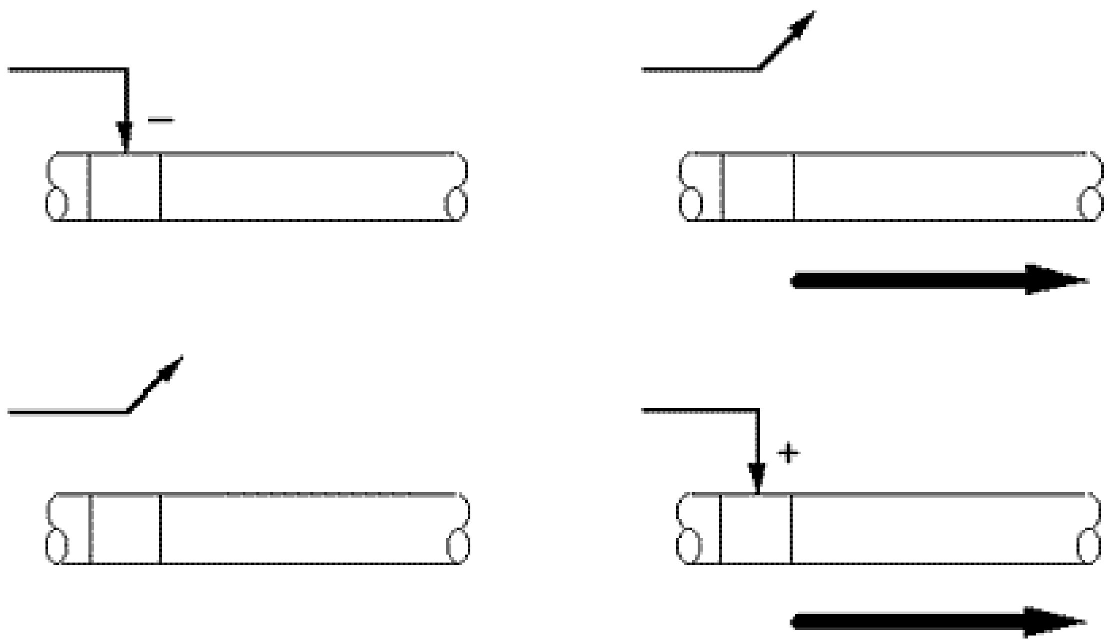
DNA Sequence

```
G G T C A T A G C T G T T T C C T G T G T G A A A T T G T T A T C C G C T C A C
                 720                 730                 740                 750

A A T T C C A C A C A A C A T A C G A G C C G G A A G C A T A A A G T G T A A A
                 760                 770                 780                 790

G C C T G G G G T G C C T A A T G A G T G A G C T A A C T C A C A T T A A T T.G
                 800                 810                 820                 830
```

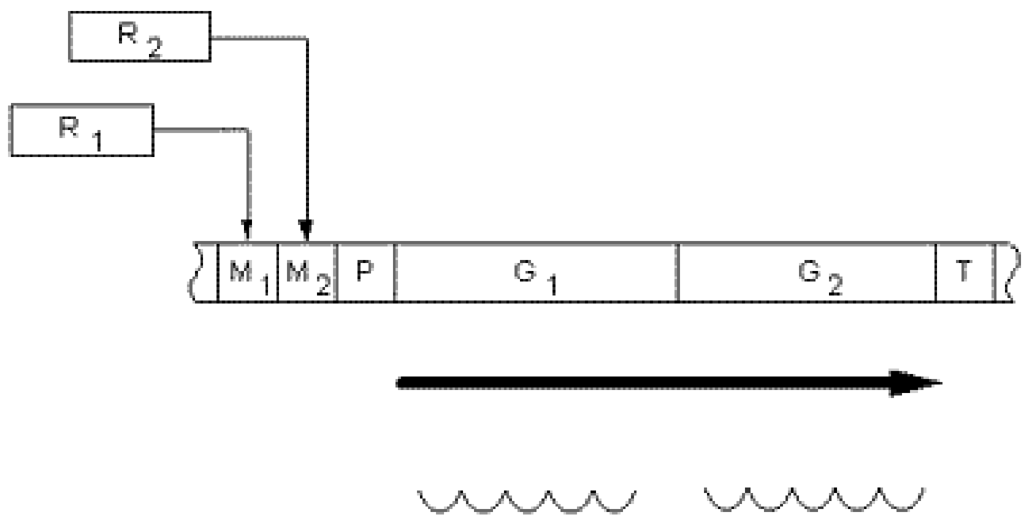
Unit of Transcription



Mode of Control

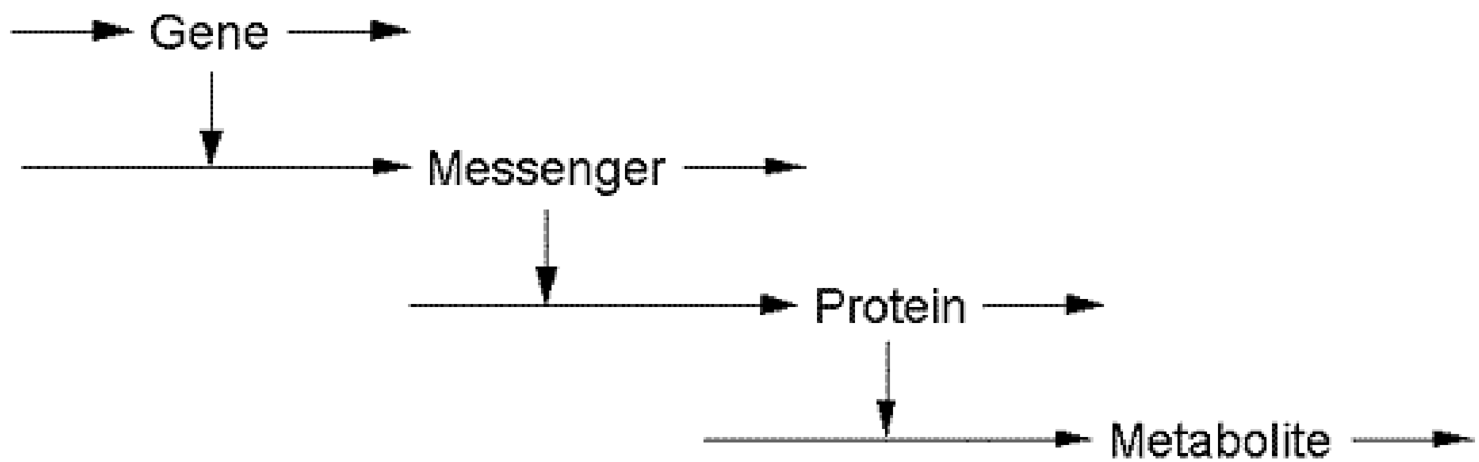


Logic

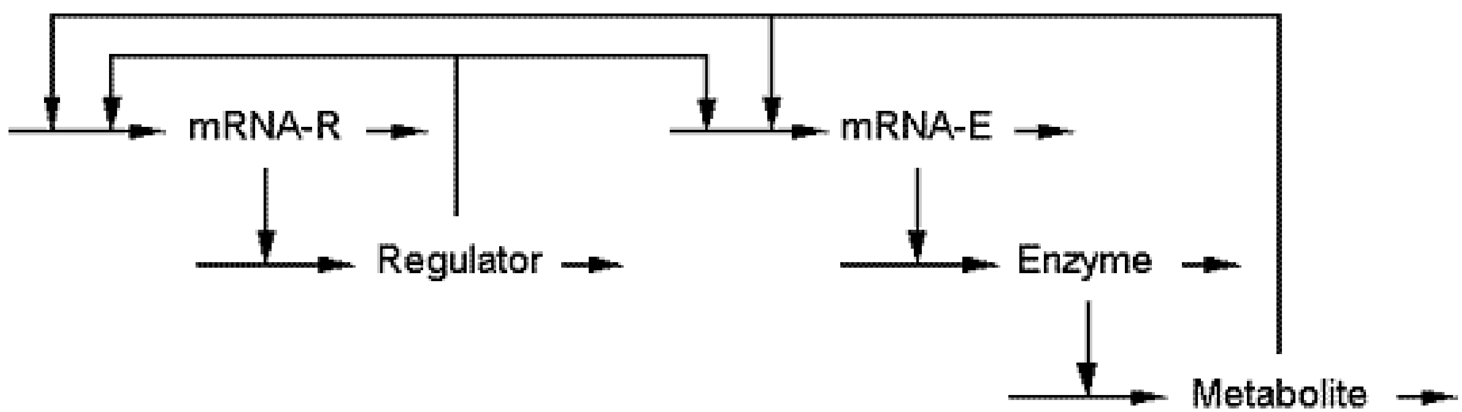


Mode		Presence		Expression	
R_1	R_2	R_1	R_2	AND	OR
+	-	Yes	Yes	OFF	ON
		Yes	No	ON	ON
		No	Yes	OFF	OFF
		No	No	OFF	ON

Expression Cascade



Connectivity



Molecular Foundation

- Molecular revolution
- Completing the catalog of components
- Lack of systemic understanding
  - No predicting the response to a novel change in the genome
  - No predicting the response to a novel environment

Systems Goal

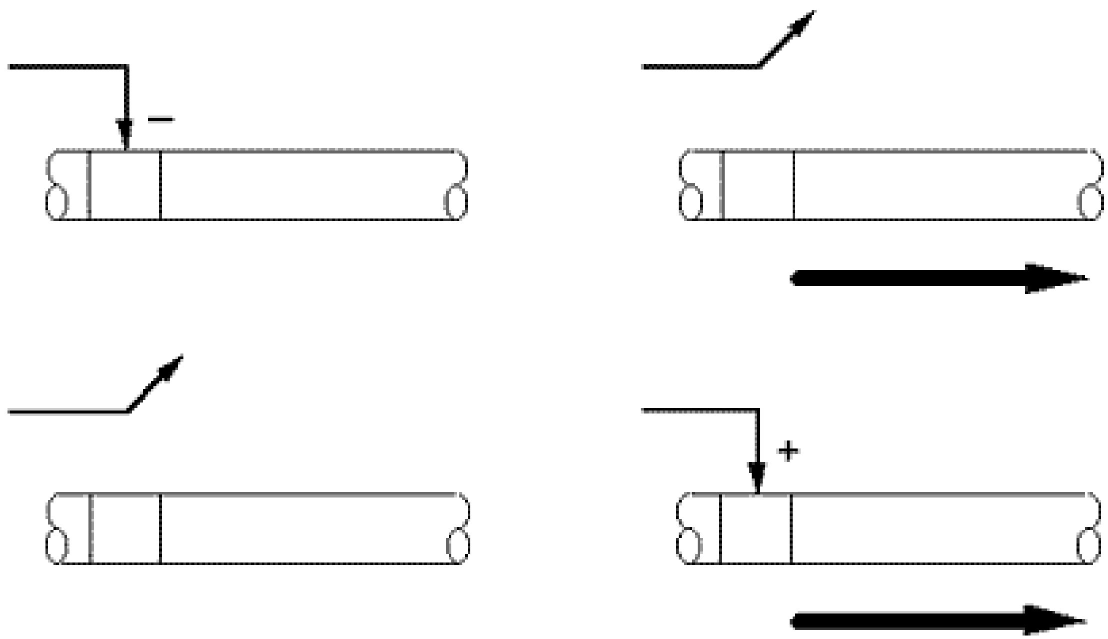
- To elucidate and to quantify
- Complex web of interactions
- Linking numerous hierarchical levels of organization
- From DNA sequence to integrated behavior
- Intact organism in its environment

Two Ways of Approaching Function

- Molecular function vs. systemic function
- Functions exhibited by a class of systems
  - Nearly universal designs
  - Alternative designs
- Functions exhibited by a particular system

Examples of Design Principles

Molecular Mode of Gene Control



Molecular Mode of Gene Control

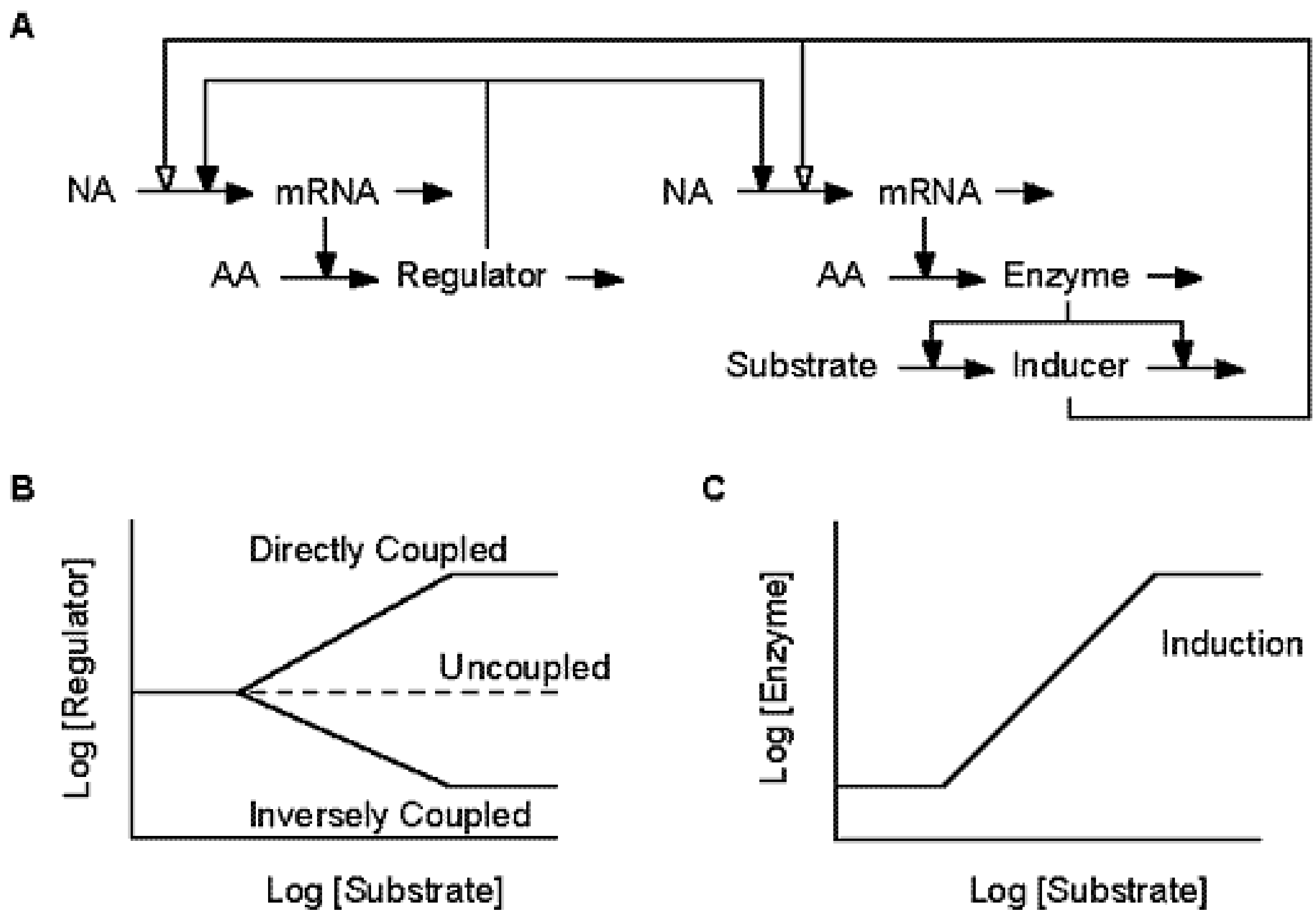
- Predictions
  - A **positive** mode of control is predicted when there is a *high* demand for expression of a gene
  - A **negative** mode of control is predicted when there is a *low* demand for expression of a gene

Molecular Mode of Gene Control

- Experimental evidence
  - Single demand functions >100
  - Logical coupling of functions ~ 20
  - Differentiated cell-specific functions ~ 6

Savageau (1989)

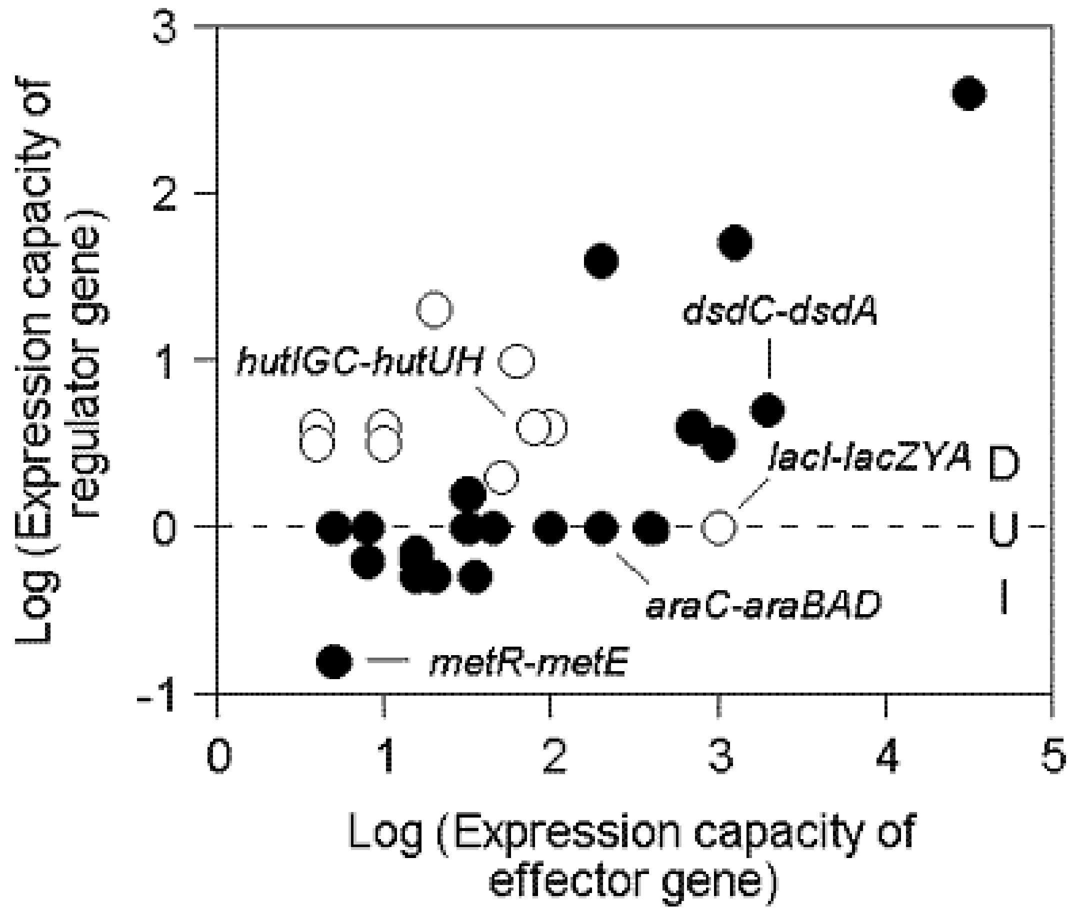
Coupling of Gene Expression in Elementary Circuits



Coupling of Gene Expression in Elementary Circuits

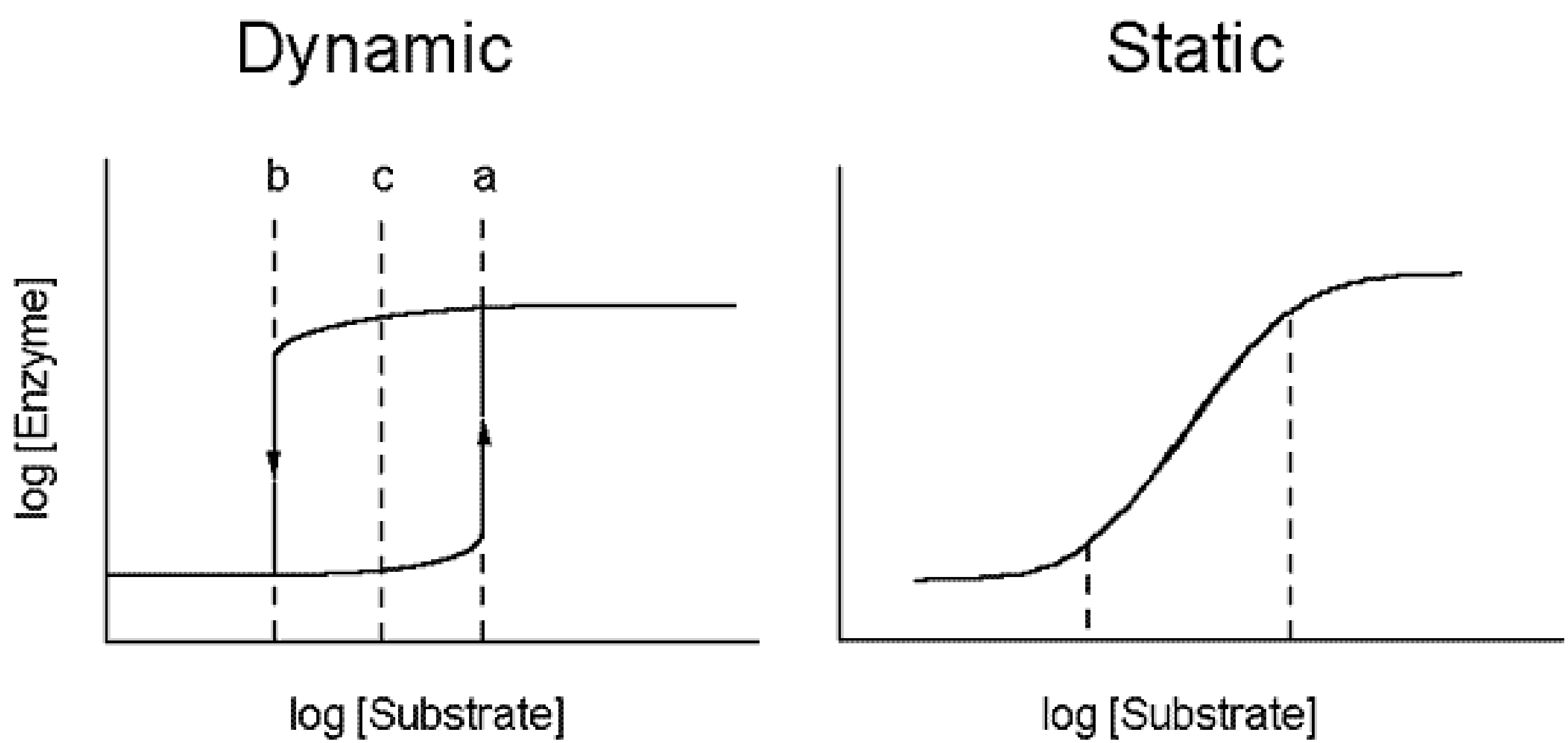
Mode	Capacity	Predicted coupling
Positive	Low	Inversely coupled
Positive	Intermediate	Uncoupled
Positive	High	Directly coupled
Negative	Low	Directly coupled
Negative	Intermediate	Uncoupled
Negative	High	Inversely coupled

Experimental Evidence for Coupling of Gene Expression in Elementary Circuits



Hlavacek & Savageau, *J.Mol.Biol.* **255**:121-139 (1996)

Two Types of Switches

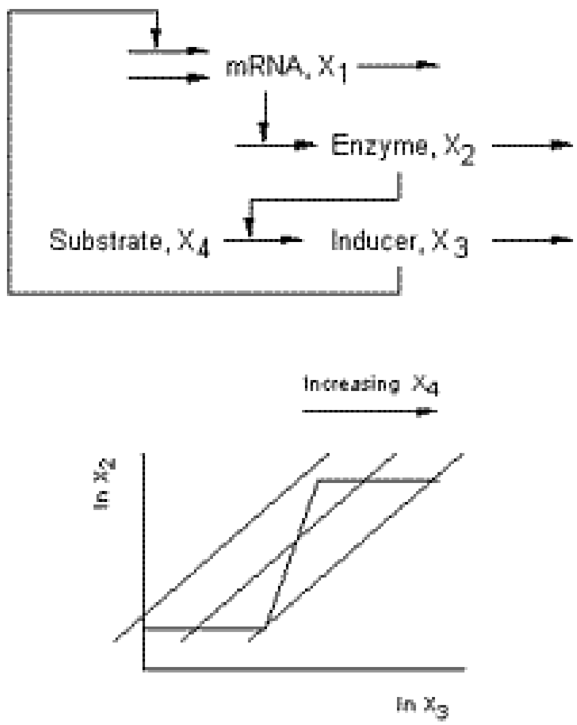


Two Types of Switches

Predictions

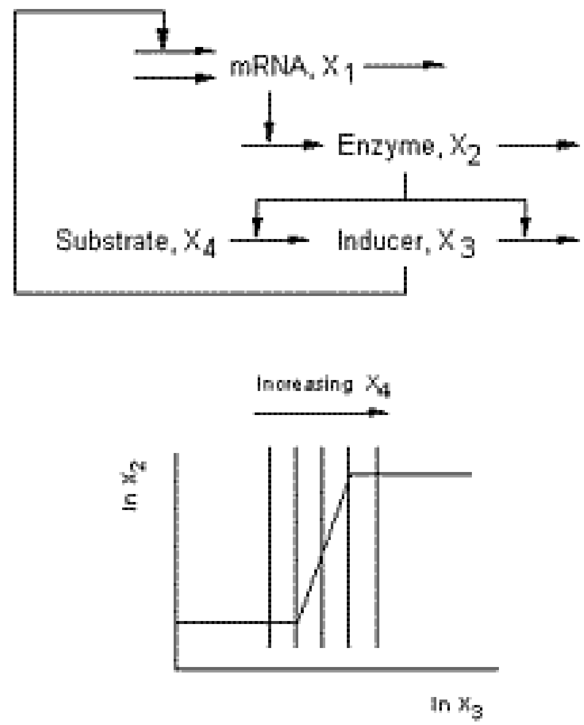
Gratuitous inducer

$$(g_{32} - h_{32}) > h_{33} / g_{13}$$



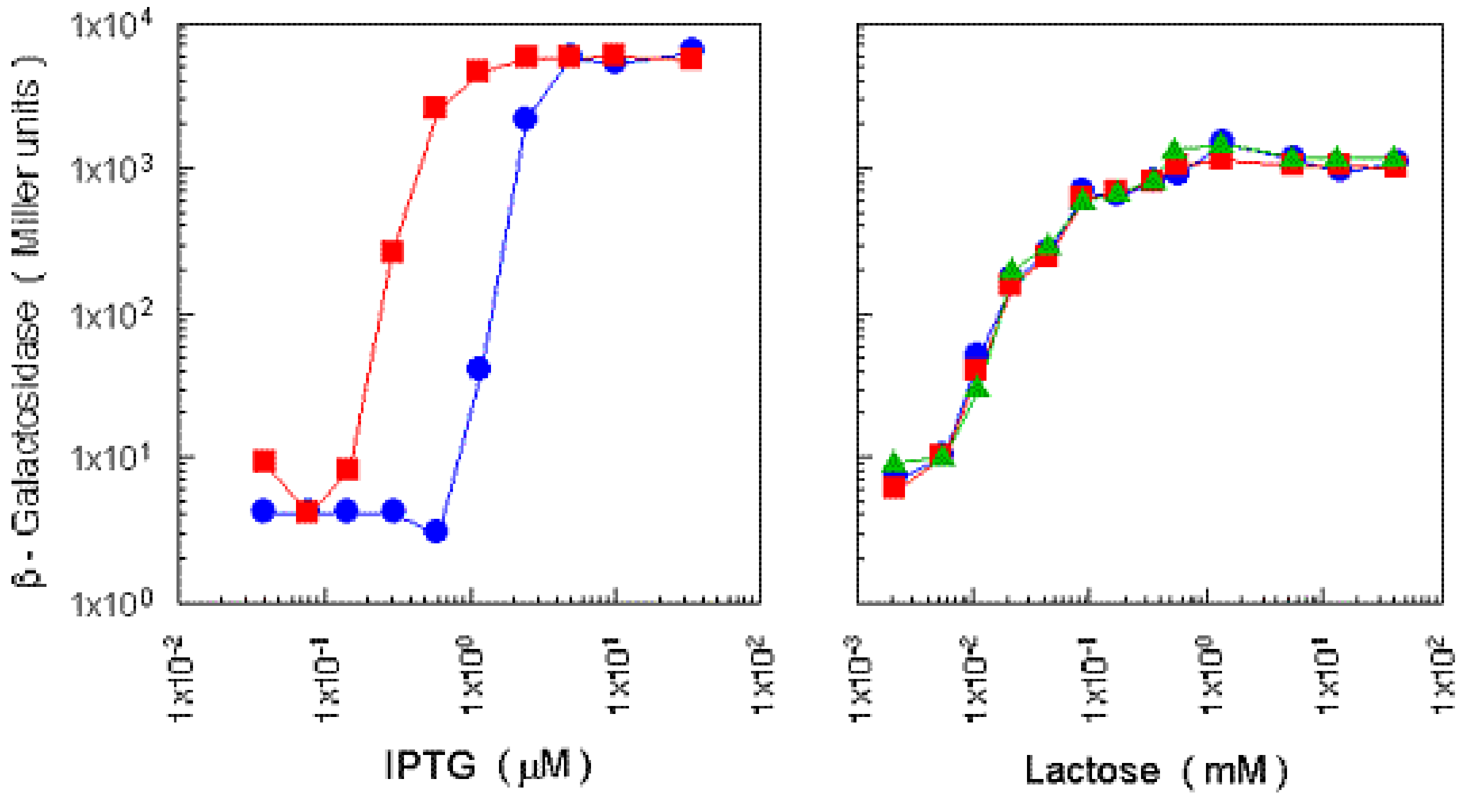
Natural inducer

$$(g_{32} - h_{32}) < h_{33} / g_{13}$$



Two Types of Switches

Experimental evidence



What is Common to These Successful Explanations of Design?

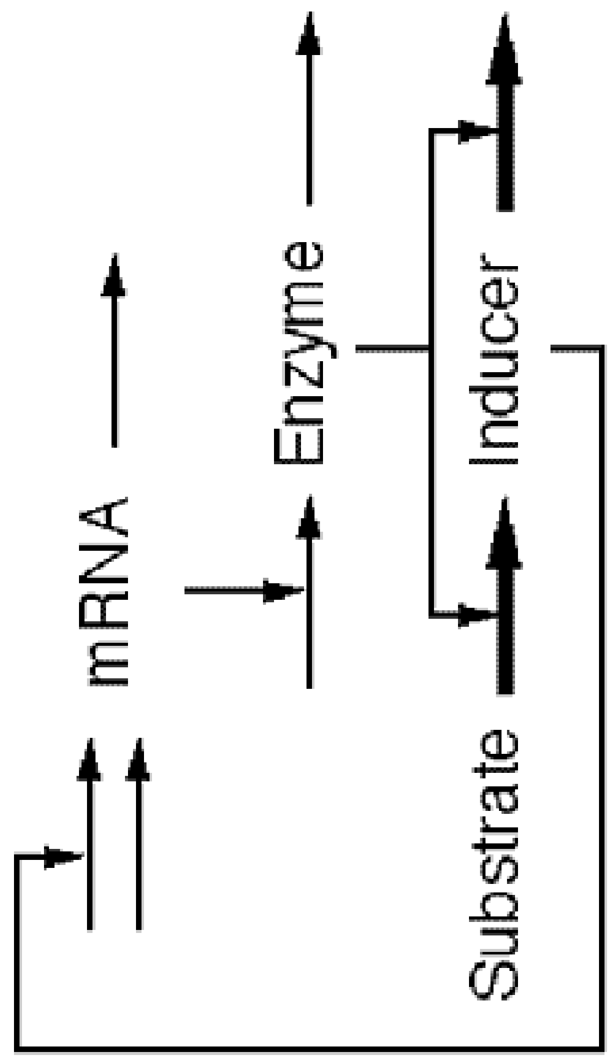
- A limited number of possible variations on a theme
- Simple equations whose structure was amenable to qualitative analysis (and to exhaustive numerical analysis when necessary)

Integration and Evolution

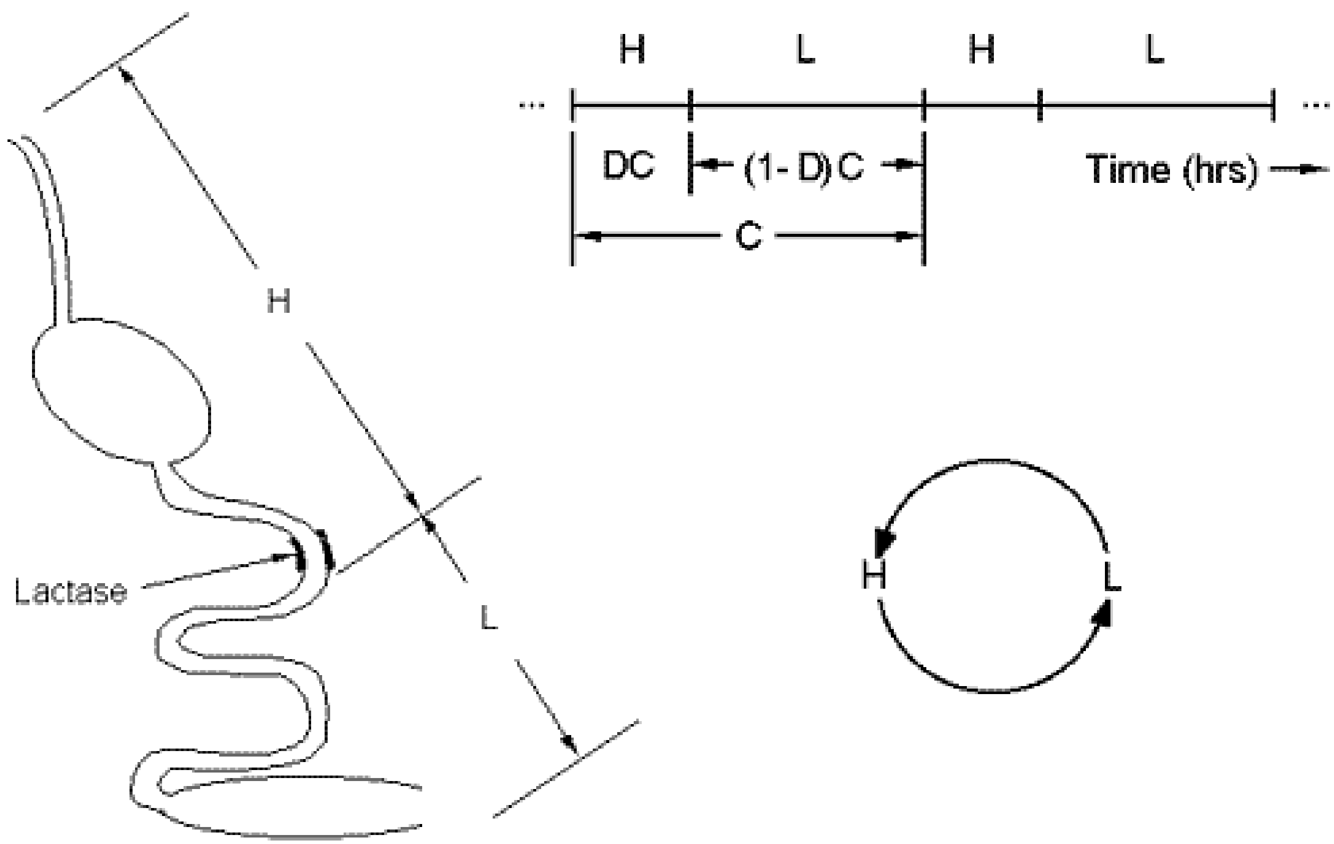
Lac operon of *E. coli*

- *Lac* circuitry
- Life cycle and demand for expression
- Mutation
- Population dynamics
- Mathematical analysis
- Quantify rules of demand theory
- Predictions relating genotype and phenotype

Lac Circuitry

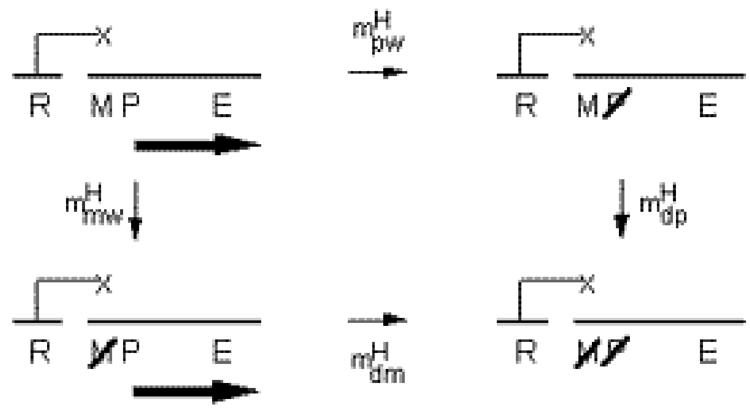


Life Cycle of *Escherichia coli*

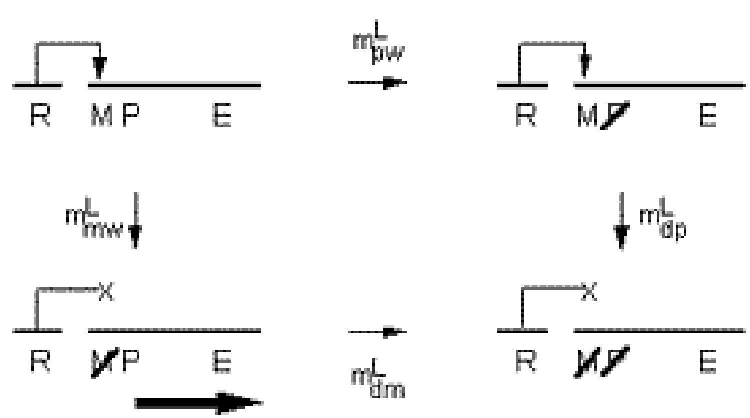


Mutations Rates

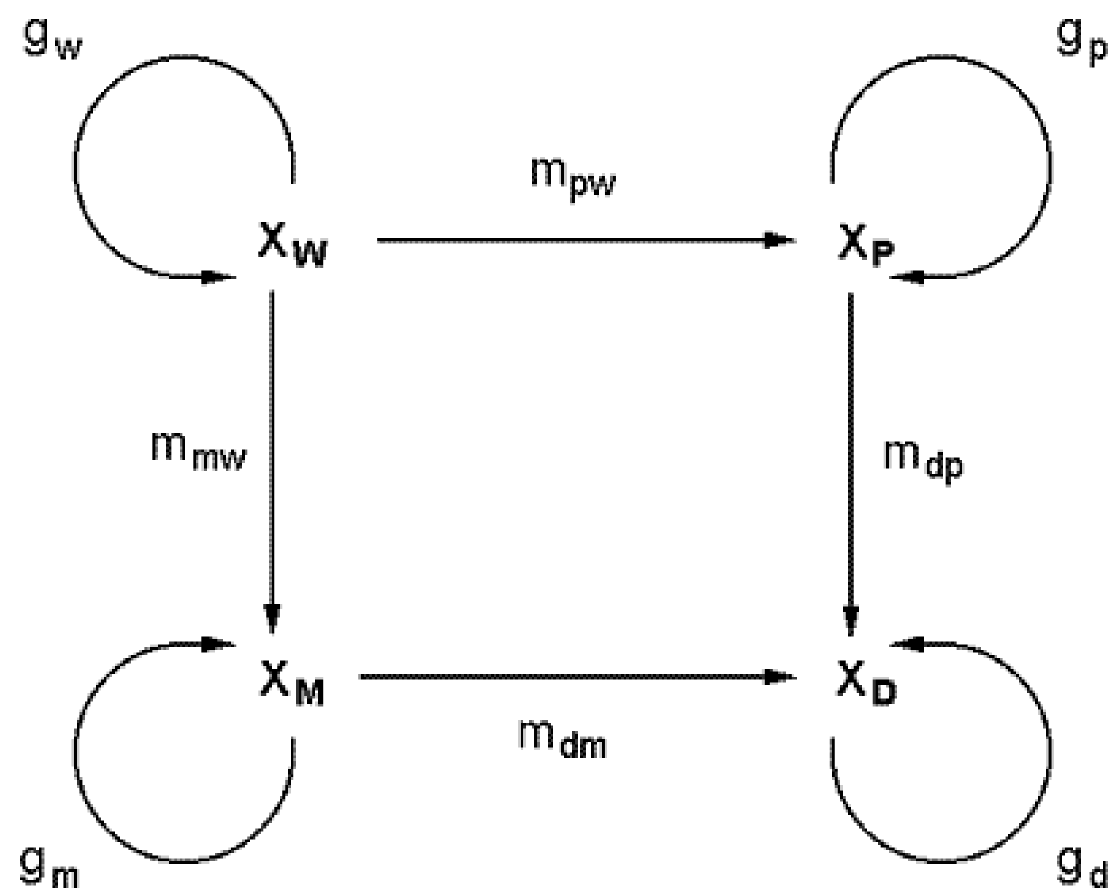
Negative Regulation in Environment H



Negative Regulation in Environment L



Population Dynamics



Equations and Solution

$$dX_w/dt = \alpha_{ww} X_w \quad (1)$$

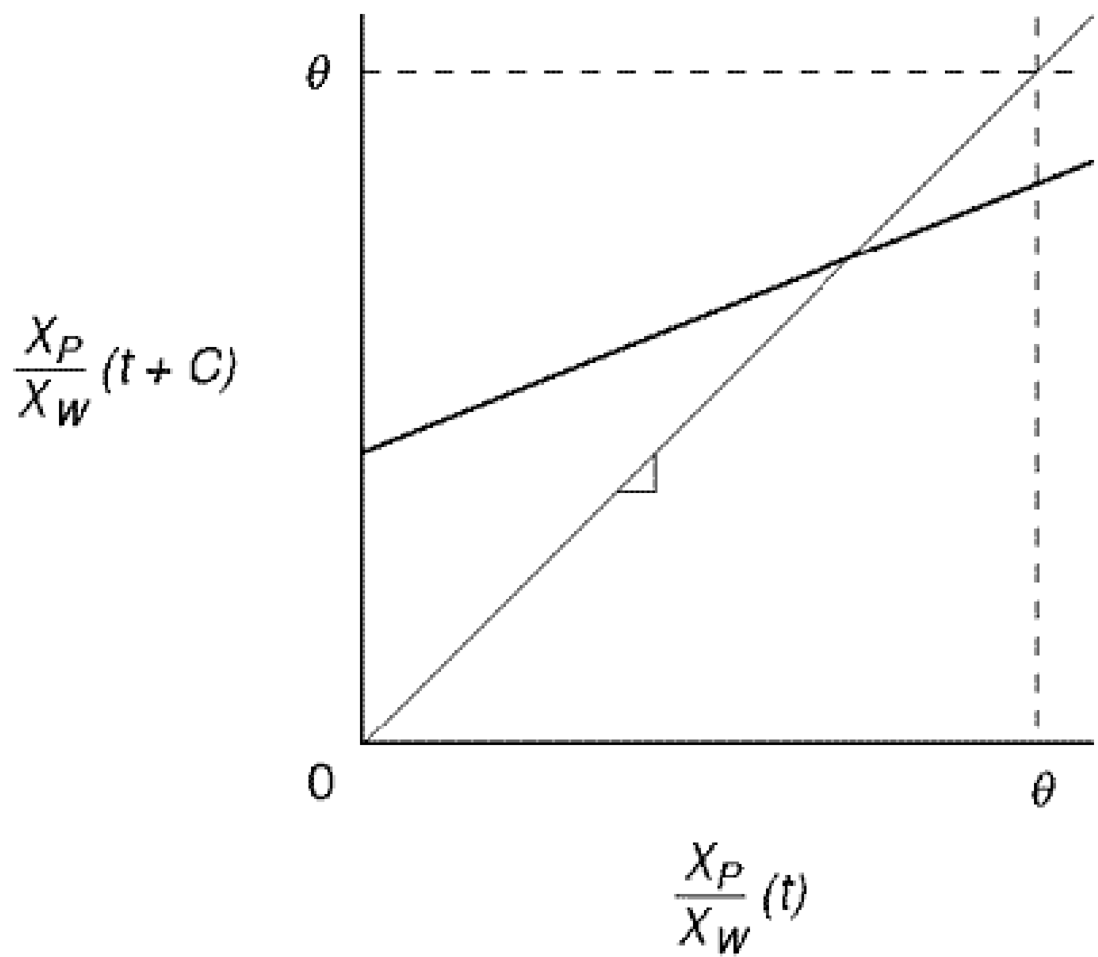
$$dX_p/dt = \alpha_{pw} X_w + \alpha_{pp} X_p \quad (2)$$

$$dX_m/dt = \alpha_{mw} X_w + \alpha_{mm} X_m \quad (3)$$

$$dX_d/dt = \alpha_{dm} X_m + \alpha_{dp} X_p + \alpha_{dd} X_d \quad (4)$$

$$X_p(t+C)/X_w(t+C) = \{[\alpha_{pw}^L/(\alpha_{ww}^L - \alpha_{pp}^L)]\{1 - \exp[(\alpha_{pp}^L - \alpha_{ww}^L)(1-D)C]\} + [\alpha_{pw}^H/(\alpha_{ww}^H - \alpha_{pp}^H)]\{1 - \exp[(\alpha_{pp}^H - \alpha_{ww}^H)DC]\} * \exp[(\alpha_{pp}^L - \alpha_{ww}^L)(1-D)C]\} + \{\exp[(\alpha_{pp}^H - \alpha_{ww}^H)DC] \exp[(\alpha_{pp}^L - \alpha_{ww}^L)(1-D)C]\} * X_p(t)/X_w(t)$$

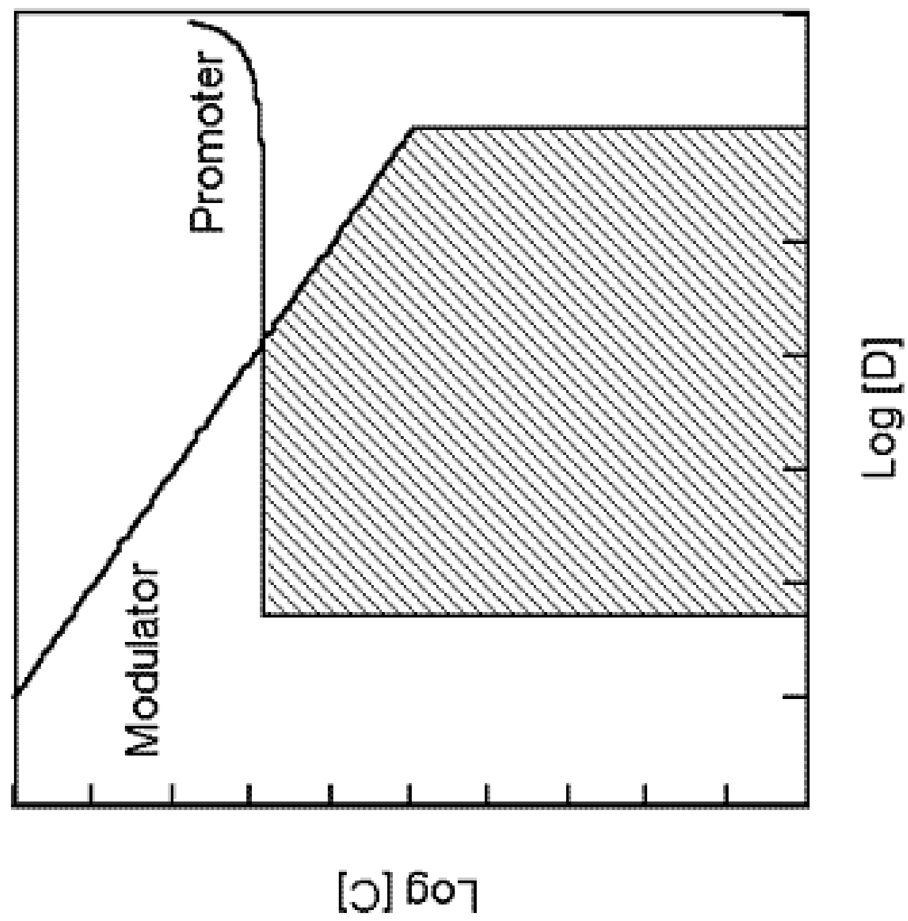
Graphical Solution



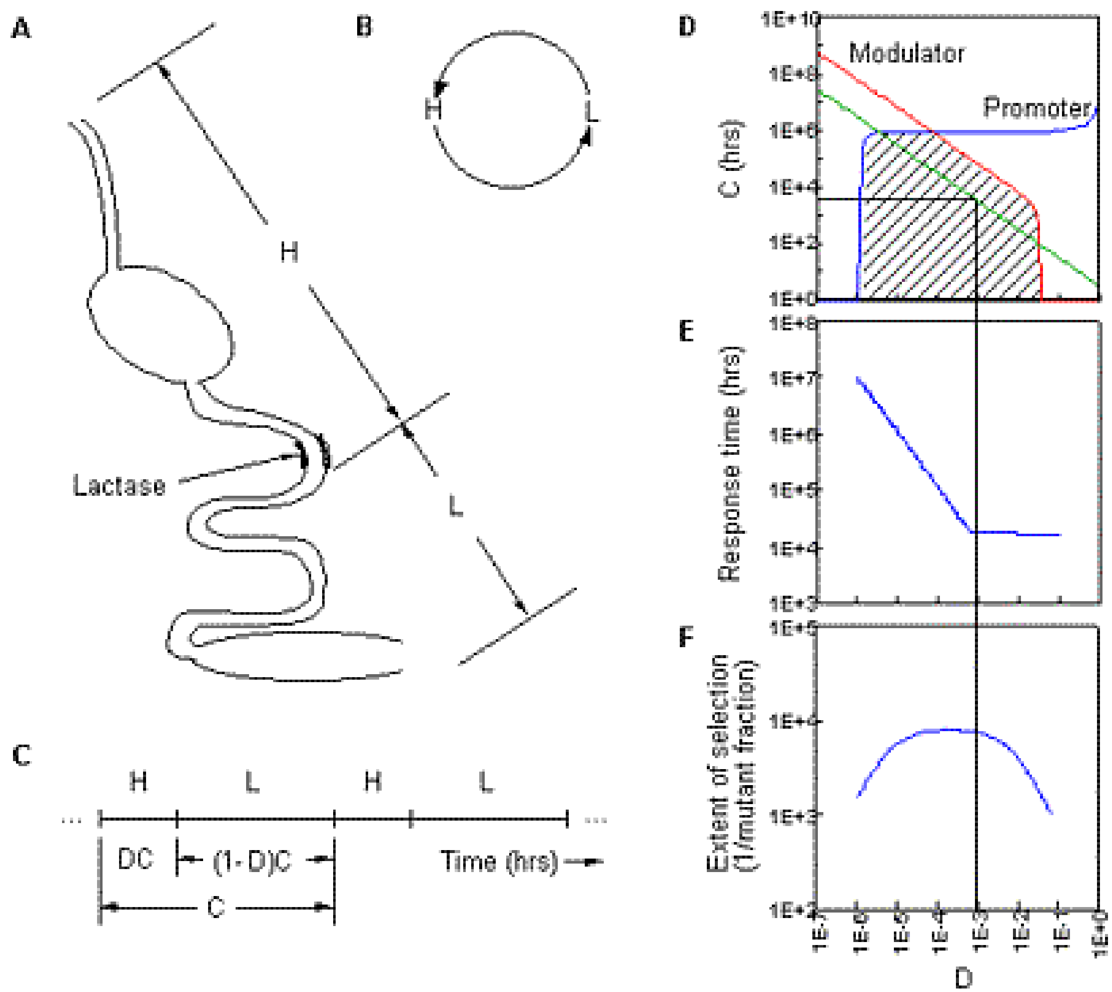
Steady State

$$X_p/X_w = \{[\alpha_{pw}{}^{L}/(\alpha_{ww}{}^{L}-\alpha_{pp}{}^{L})]\{1- \exp[(\alpha_{pp}{}^{L}- \alpha_{ww}{}^{L})(1-D)C]\} +$$
$$[\alpha_{pw}{}^{H}/(\alpha_{ww}{}^{H}-\alpha_{pp}{}^{H})]\{1- \exp[(\alpha_{pp}{}^{H}- \alpha_{ww}{}^{H})DC]\} *$$
$$\exp[(\alpha_{pp}{}^{L}- \alpha_{ww}{}^{L})(1-D)C]\} /$$
$$\{1-\exp[(\alpha_{pp}{}^{H}- \alpha_{ww}{}^{H})DC + (\alpha_{pp}{}^{L}- \alpha_{ww}{}^{L})(1-D)C]\}$$

Region of Realizability



Rate and Extent of Selection



Predictions

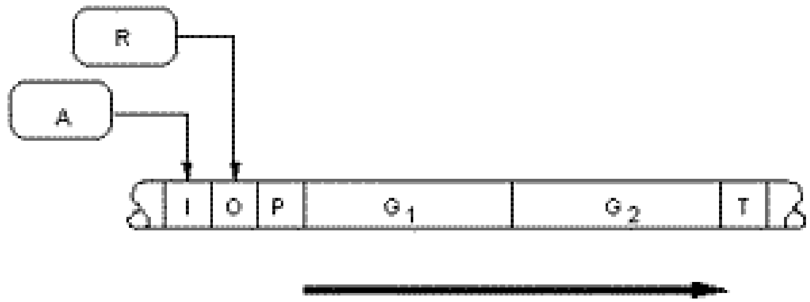
- Cycling without colonization $\approx$ 26 hours
- Colonization without cycling $\approx$ 66 years
- Rate of re-colonization $\approx$ 4 months

Savageau, *Genetics* **149**:1677-1691 (1998)

Regulation Involving Combinatorial Processes

- Logic
- Phasing
- Evolvability

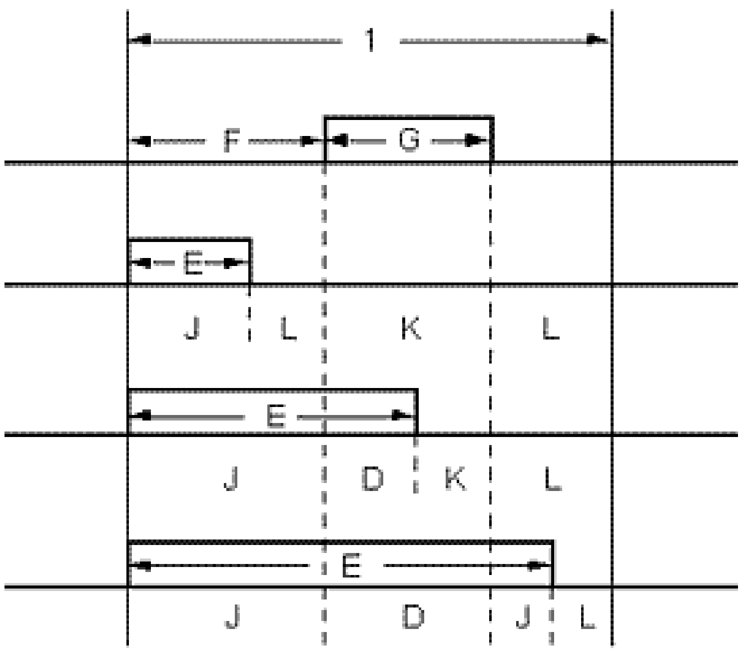
Logic of *Lac* Control



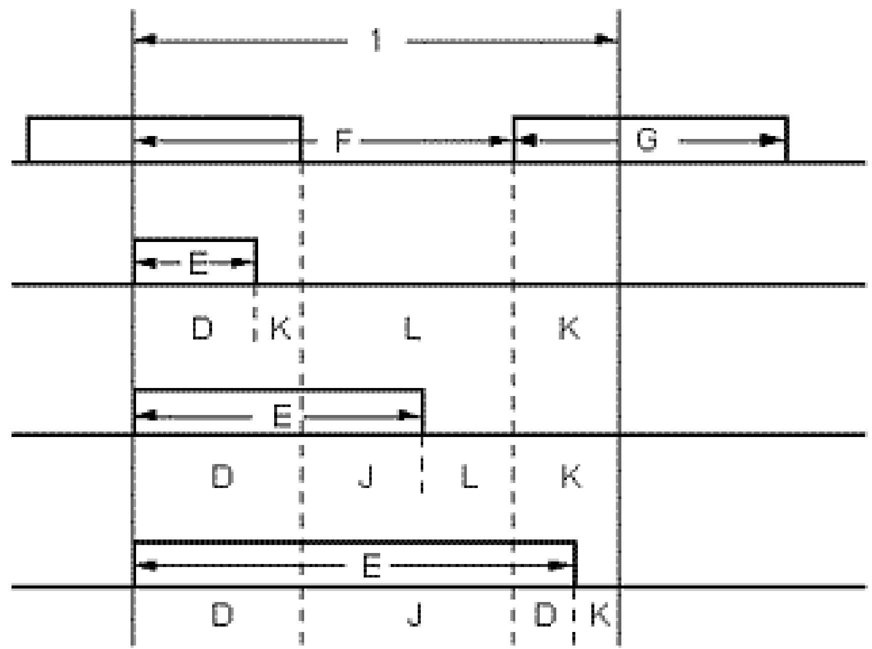
Four Possible States

Regulator		Expression	State
Activator	Repressor		
+	+	-	J
+	-	+	D
-	+	-	L
-	-	-	K

Six Possible Life Cycles

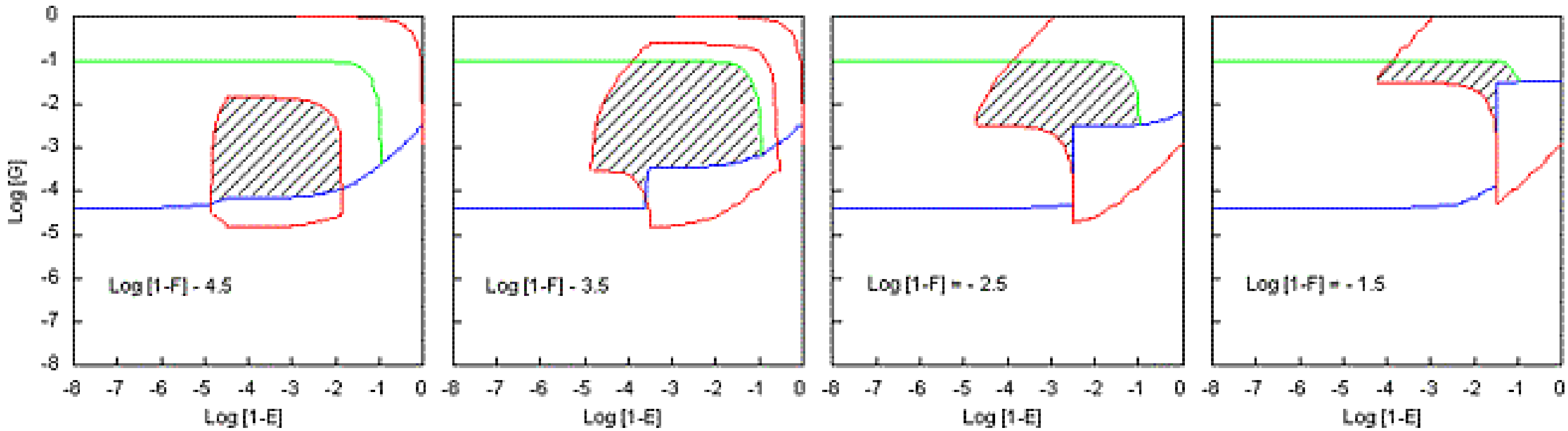


F + G < 1

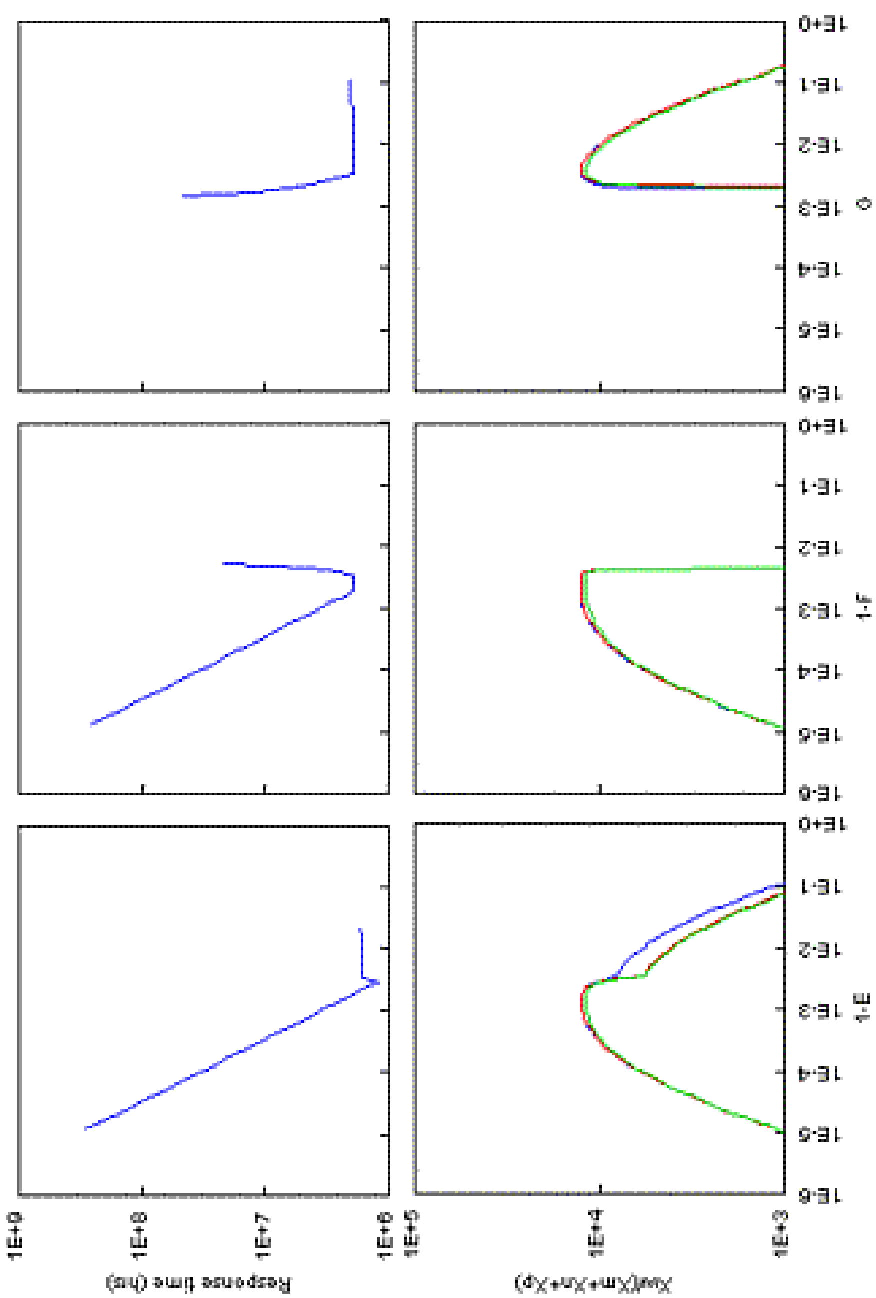


F + G > 1

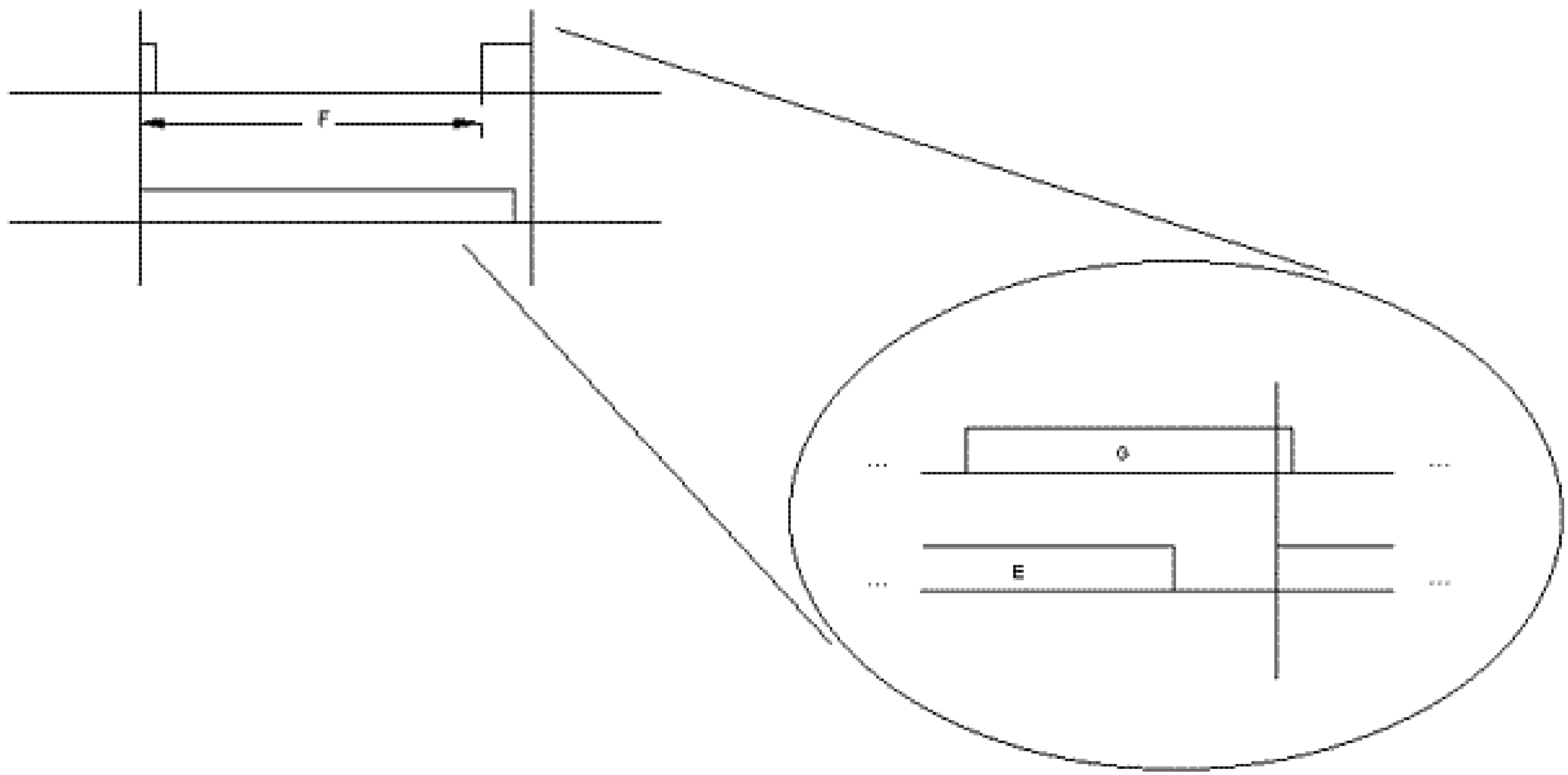
Region of Realizability



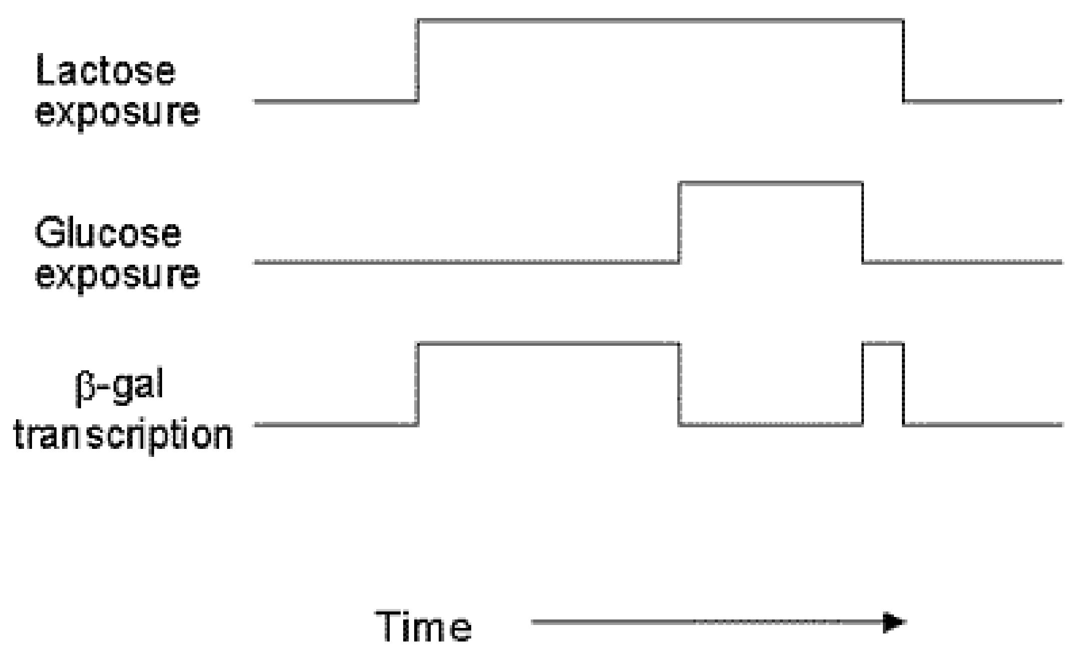
Rate & Extent of Selection



Predicted Phasing of Expression



Interpretation



Goal Restated

- To elucidate and to quantify
- Complex web of interactions
- Linking numerous hierarchical levels of organization
- From DNA sequence to integrated behavior
- Intact organism in its environment

"There is one important piece of information that is almost totally missing: the sequence information that specifies when and where and for how long a gene is turned on or off."

-- Sydney Brenner, *Science* 287:2173-4 (2000)

Summary

Sequence
Molecular mechanisms
Circuitry
Physiology
Ecology
Selection

Mathematical modeling → Integration
Computation

Methodological Issues

- Canonical nonlinear representations
- Mathematically controlled comparisons
- Quantitative criteria for functional effectiveness

Systemic Manifestations of the Power-Law Formalism

- Generalized-Mass-Action Representation

$$\frac{dX_i}{dt} = \sum_{k=1}^{r} \alpha_{ik} \prod_{j=1}^{n} X_j^{g_{ijk}} - \sum_{k=1}^{r} \beta_{ik} \prod_{j=1}^{n} X_j^{h_{ijk}}$$

- Synergistic-System Representation

$$\frac{dX_i}{dt} = \alpha_i \prod_{j=1}^{n} X_j^{g_{ij}} - \beta_i \prod_{j=1}^{n} X_j^{h_{ij}}$$

Mathematically Controlled Comparisons

- Two designs are restricted to having differences in a single specific process
- One design is chosen as the reference
- Internal equivalence
- External equivalence
- The systems can be rigorously compared by mathematical and computer analysis

Quantitative Criteria for Functional Effectiveness

- Decisiveness
- Efficiency
- Selectivity
- Robustness
- Stability
- Responsiveness

Power-Law Formalism is Canonical from Three Different Perspectives

- Fundamental
- Recast
- Local

Hierarchy of Fundamental Representations

- Power law - fractal kinetics $$\frac{dX_i}{dt} = \sum_{k=1}^{r} \alpha_{ik} \prod_{j=1}^{n} X_j^{g_{ijk}}$$
- Mass action - chemical kinetics $$\frac{dX_i}{dt} = \sum_{k=1}^{r} \alpha_{ik} \prod_{j=1}^{n} X_j^{l_{ijk}}$$
- Volterra - ecosystem dynamics $$\frac{dX_i}{dt} = X_i \left(b_i + \sum_{j=1}^{n} a_{ij} X_j \right)$$
- Linear- compartmental analysis $$\frac{dX_i}{dt} = b_i + \sum_{j=1}^{n} a_{ij} X_j$$

Recast Representation: Example 1

$$dx/dt = 0.343 - (y + 17.15)e^{-x} \qquad x(0) = 3.85$$

$$dy/dt = e^{-x} - (50 + z) \qquad y(0) = 7.16$$

$$dz/dt = 1.82 + (y - 9.75)z \qquad z(0) = 7.98$$

- $x_1 = e^x$, $x_2 = y + 17.15$ and $x_3 = z + 50$

$$dx_1/dt = 0.343x_1 - x_2 \qquad x_1(0) = 46.87$$

$$dx_2/dt = x_1 - x_3 \qquad x_2(0) = 24.31$$

$$dx_3/dt = 1346.82 + x_2x_3 - 50x_2 - 26.9x_3 \qquad x_3(0) = 57.98$$

Recast Representation: Example 2

$$dx / dt = 0.343 - (y + 17.15)e^{-x} \qquad x(0) = 3.85$$

$$dy / dt = e^{-x} - (50 + z) \qquad y(0) = 7.16$$

$$dz / dt = 1.82 + (y - 9.75)z \qquad z(0) = 7.98$$

- $X_1 = e^x$, $X_2 = y + 17.15$ and $X_3 X_4 = z + 50$

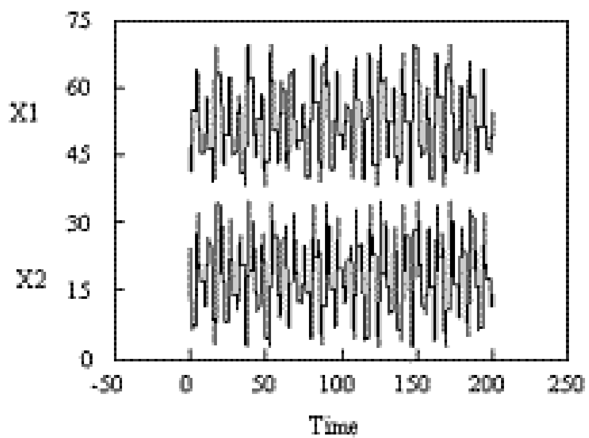
$$dx_1 / dt = 0.343 x_1 - x_2 \qquad x_1(0) = 46.87$$

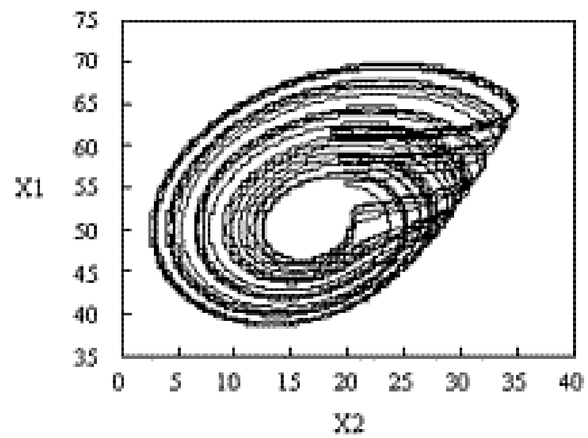
$$dx_2 / dt = x_1 - x_3 x_4 \qquad x_2(0) = 24.31$$

$$dx_3 / dt = 1346.82 x_4^{-1} - 50 x_2 x_4^{-1} \qquad x_3(0) = 57.98$$

$$dx_4 / dt = x_2 x_4 - 26.9 x_4 \qquad x_4(0) = 1$$

Global Accuracy of Recast Representations





Local Representation

$$\ln v = f(\ln x)$$

$$\ln v \approx \ln v_0 + \left.\frac{d(\ln v)}{d(\ln x)}\right|_0 (\ln x - \ln x_0)$$

$$\ln v \approx \ln v_0 + \sum_{j=1}^{n} \left.\frac{d(\ln v)}{d(\ln x_j)}\right|_0 (\ln x_j - \ln x_{j0})$$

$$v(x) = \alpha \prod_{j=1}^{n} x_j^{g_j}$$

Some References

- Savageau, M.A., and E.O. Voit, Recasting nonlinear differential equations as S-systems: a canonical nonlinear form. *Math. Biosci.* **87**, 83-115 (1987).
- Savageau, M.A., Are there rules governing patterns of gene regulation?, *in* "Theoretical Biology -- Epigenetic and Evolutionary Order," B.C. Goodwin and P.T. Saunders, Eds., Edinburgh University Press, Edinburgh, pp 42-66 (1989).
- Irvine, D.H., and M.A. Savageau, Efficient solution of nonlinear ordinary differential equations expressed in S-system canonical form, *SIAM J. Numerical Anal.* **27**, 704-735 (1990).
- Savageau, M.A., Power-law formalism: A canonical nonlinear approach to modeling and analysis, *in* "World Congress of Nonlinear Analysts 92," Vol. 4, V. Lakshmikantham, Ed., Walter de Gruyter Publishers, Berlin, pp. 3323-3334 (1996).
- Hlavacek, W.S., and M.A. Savageau, Rules for coupled expression of regulator and effector genes in inducible circuits, *J. Mol. Biol.* **255**, 121-139 (1996).
- Savageau, M.A., Development of fractal kinetic theory for enzyme-catalyzed reactions and implications for the design of biochemical pathways, *BioSystems* **47**, 9-36 (1998).
- Savageau, M.A., Demand theory of gene regulation: Quantitative development of the theory, *Genetics* **149**, 1665-1676 (1998).
- Savageau, M.A., Demand theory of gene regulation: Quantitative application to the lactose and maltose operons of *Escherichia coli*, *Genetics* **149**, 1677-1691 (1998).
- Savageau, M.A., Design principles for elementary gene circuits, *Chaos* **11**, 142-159 (2001).