Adversity, Diversity and Design

Signal Processing in Cellular Regulatory Networks: Evolution, Design and Control and Physics

Prediction, Control and Design
Talk Overview

1. Review of central dogma and evolution
2. Review of basic cellular dynamics
3. The central questions in cellular network design
4. The game of life and optimal design
5. An early “pulling” application
Talk Overview

1. Review of central dogma and evolution
2. Review of basic cellular dynamics
3. The central questions in cellular network design
4. The game of life and optimal design
5. An early “pulling” application
"Nothing in biology makes sense except in the light of evolution."

The Advent of Molecular Biology

Genome

Macromolecules

Metabolites
All life uses the same genetic code
Through RNA
Feedback & Feedforward

Biochemistry
Myxococcus xanthus

• Even cells as “simple” as bacteria are highly social, differentiating, sensing/actuation systems

Images courtesy of Dale Kaiser
Immune cells

• They perform amazing engineering feats under the control of complex cellular networks
Evolution as a tool for understanding

- Evolution acts on phenotype
- Phenotype is created by genotype
- Genotype is thus the substrate for evolution

- By comparing and contrasting “homologous” systems in different organisms perhaps design principles can be learned
Cells as “Bags O’ Parts”

- **Primordial Lego® Kit**
  - Assembly using simple interconnect protocol (explicit but underdetermined)

- **First Cellular Life**
  - Self-assembly using genetic code protocol (implicit and mostly determined)
Cartoon of Evolution

Primordial Lego® Kit

First Cellular Life

Point mutation

Insertion mutation

gene loss

gene duplication

Insertion and point mutation
Primordial Lego® Kit

Cartoon of Evolution

How does this new part work in a diverged context?

Is the duplicated and diverged gene “adaptive”?

Did the dark blue box HGT from f.1.2.2 or mutate from f2.2.1?
Are there designs that make it more probable that mutation will not destroy critical old function and will possibly create good new function?
Designs for Evolvability

Building blocks shouldn’t be too atomistic:

Limits the space of function
Designs for Evolvability

Parts shouldn’t be too specialized or too complex

Or else you limit potential functionality and are more fragile to mutation
Designs for Evolvability

Building blocks should be designed for compatibility and utility:

- Incompatible
- Hard to put together, odd shapes for utility

Or else you limit coherence of the machines in the cell, prevent reuse of parts evolutionarily, and make it harder to build complex objects from smaller ones.
1. Review of central dogma and evolution
2. Review of basic cellular dynamics
3. The central questions in cellular network design
4. The game of life and optimal design
5. An early “pulling” application
Basic Kinetics

$[A]_{ss}$

E.g. Focal Adhesion Kinase Alternative Splice
\[
\frac{dA}{dt} = \text{Dephosphorylation} - (\text{Phosphorylation}_B + \text{Phosphorylation}_{A-p})
\]

**Phosphorylation** 
\[
\text{Phosphorylation}_B = k_{\text{cat-f}} \ast [B - p] \ast \frac{[A]}{K_{Af} + [A]}
\]

**Dephosphorylation** 
\[
\text{Dephosphorylation} = V_{\text{max-r}} \ast \frac{[A - p]}{K_{Ar} + [A - p]}
\]

**Phosphorylation** 
\[
\text{Phosphorylation}_{A-p} = k_{\text{cat-fA}} \ast [A - p] \ast \frac{[A]}{K_{Af2} + [A]}
\]
Bistability

A simple model of the positive feedback

$k_c$ – catalytic constant for the trans-autophosphorylation.
Diversification via Stochastics
(aka random fluctuations or noise)

**Deterministic process (large numbers)**

![Graph showing population size over time](image)

Population Size $N$

Time

Large numbers = (largely) predictable behavior

**Stochastic process (small numbers)**

![Graph showing populations in a stochastic environment](image)

B. Populations in a Stochastic Environment

Time, $t$

Abundance, $N(t)$

Small numbers = random molecular (and phenotypic) fluctuation
Molecular Reaction Noise

- One gene
- Growing cell, 45 minutes division time
- Average ~60 seconds between transcripts
- Average 10 proteins/transcript:

Monte Carlo simulation data

Feedback → Stochastic Fate


You don’t even need feedback!

\[
\frac{1}{2} = p_1 = p_0 = 0.3
\]

\[
E + 0.005 \quad 0.01 \quad 0.05 \quad 0.1 \quad 0.5 \quad 1
\]

\[
E_0 \quad 0.3 \quad 0.5 \quad E_{\frac{1}{2}} \quad E_1 \quad 1.5 \quad 2
\]

Deterministic Stationary State Solution

Langevin Solutions

More rigorous: Master Equation

\[ \frac{dP(\vec{X}_i, t)}{dt} = \sum_{i' \neq i} W_{i \rightarrow i'} P(\vec{X}_{i'}, t) - W_{i \rightarrow i} P(\vec{X}_i, t) \]
Molecular Reaction Noise

Prediction of noise in bacterial gene expression:

\[ \frac{dP(\tilde{X}_i, t)}{dt} = \sum_{i \neq i} W_{i \rightarrow i} P(\tilde{X}_i, t) - W_{i \leftarrow i} P(\tilde{X}_i, t) \]


Prediction and demonstration of noise in expression from the HIV promoter.

## Discriminating among 16 Models

<table>
<thead>
<tr>
<th>MODELS TESTED &amp; RESULTING TRAJECTORIES</th>
<th>Acetylation Neglected</th>
<th>Acetylation Included</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_{\text{BASAL}} &gt; 10^{-4}$</td>
<td>$k_{\text{BASAL}} &lt; 10^{-4}$</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{BASAL}} &gt; 10^{-8}$</td>
<td>$k_{\text{BASAL}} &lt; 10^{-8}$</td>
</tr>
<tr>
<td><strong>No Tat Cooperativity</strong> (1 molecule of Tat required)</td>
<td>Tat$_{e} = 0$</td>
<td>Bright only</td>
</tr>
<tr>
<td></td>
<td>Tat$_{e} \geq 1$</td>
<td>OFF only</td>
</tr>
<tr>
<td><strong>Tat Cooperativity</strong> (≥ 2 molecules of Tat required)</td>
<td>Tat$_{e} = 0$</td>
<td>Bright Only</td>
</tr>
<tr>
<td></td>
<td>Tat$_{e} \geq 1$</td>
<td>Bright -Off Uniform</td>
</tr>
</tbody>
</table>

**TIME (each plot is from 0 → 3 weeks)**
Spatial PDE Analysis

Elowitz

Colella/Arkin

Onsum/Arkin

Colella/Arkin

Baumeister
Bacterial Cytoskeleton

Biological dynamics

• Naively: biological systems are high dimensional, spatially and mechanically operating, stochastic dynamical systems. (Eek!)

• Abstractions are possible but the formalisms are early.

• There is fundamental theory to be done in architecture, protocol layers and communication (John Doyle and Munther Dahleh, Ron Pinter and others)

• Models are complex but there is some progress on model reduction (Mustafa Khammash)
Talk Overview

1. Review of central dogma and evolution
2. Review of basic cellular dynamics
3. The central questions in cellular network design
4. The game of life and optimal design
5. An early “pulling” application
Questions

• How do we use evolution to infer the principles of design and evolvability of cellular networks?

• How do we link evolutionary analyses to dynamical analyses of cellular systems?

• What is being optimized? Can we infer the “internal model” of the environment in cellular control networks?

• Can we use deduced and inferred principles for design of new organisms and functions that we can prove are effective and fail-safe?
Talk Overview

1. Review of central dogma and evolution
2. Review of basic cellular dynamics
3. The central questions in cellular network design
4. The game of life and optimal design
5. An early “pulling” application
The set up: Niches are Dynamic

• Characteristic times may be spent in each environment.

• Environments themselves are variable.
The game of life
Logic of *B. subtilis* stress response

- Network organization has a functional logic.
- There are different levels of abstraction to be found.
B. subtilis starvation responses

Staging of modules during transition to S.P.  Diversification of Individual Response
Clustered phylogenetic profile shows blocks of conserved genes

1. methyl-processing receptors and chemotaxis genes in motile bacteria
2. methyl-processing receptors and chemotaxis genes in motile Archaea
3. flagellar genes in motile bacteria
4. type III secretion system (virulence) in non-motile pathogenic bacteria
5. motility genes in spore-forming bacteria
6. late-stage sporulation genes in spore-forming bacteria
7. spore coat and germination response genes in spore-forming bacteria that are not competent
8. late-stage sporulation genes in spore-forming bacteria that are also competent
9. DNA uptake genes in Gram positive bacteria
10. DNA uptake genes in Gram negative bacteria
Clusters are functionally coherent

<table>
<thead>
<tr>
<th>Receptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal Transduction (che)</td>
</tr>
<tr>
<td>Hook and Flagellar Body</td>
</tr>
<tr>
<td>Flagellar export/Type III secretion</td>
</tr>
<tr>
<td>Flagellar length and motor control</td>
</tr>
<tr>
<td>Hypothetical receptors</td>
</tr>
<tr>
<td>Cross-Regulation with Sporulation/Cell Cycle</td>
</tr>
</tbody>
</table>
Different modules for different lives

<table>
<thead>
<tr>
<th>Archeal Extremophiles</th>
<th>Animal pathogens</th>
<th>Plant pathogens</th>
<th>Endopathogens</th>
<th>Sporulators</th>
<th>Endopathogens</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Different Modules/Different Conservation

Sensors

Cross-Talk

Regulators

Actuators
Differences in robustness

E. Coli

B. subtilis

Chris Rao/John Kirby
Sporulation Initiation
Bistability

[ 

\[ A_R (\text{protein/mRNA-s}) \]

\[ \text{[Sin]} (\text{nM}) \]

\[ k_3 \]

\[ \text{region affecting } k_1 \]

\[ K_1 \]

\[ \text{Comparison of five strains of Bacillus anthracis} \]

In anthracis:
Mutations mostly affect \( K_1 \) and \( k_1 \)
Threshold of the switch is most affected.

Across all sporulators
Very variable.

Voigt, Wolf, APA, Genetics (2005)
Evolutionary Modules in *B. subtilis*

- Simple clustering of phylogenetic profiles demonstrates modularity at different levels of resolution.

- Homologous modules can have very different mechanisms and properties even when general behavior is similar.

- Even fully “conserved” modules may be functionally flexible.

- Comparative genomics may provide a clue as to what features of a circuit are under selection.
Adaptability via Diversification

• Standard Paradigm: Cell senses environment, implements a particular response.

• Heterogeneity arises largely due to uncontrolled/local effects and mutation.

• However, non-genetic individuality can be an evolutionarily stable strategy.
Examples of Stochastic Choice

• Controlled random phase variation
  – Type-1 and Pap pili phase variation
  – Antigenic surface variation
    • Flagellar phase variation in Salmonella
    • Surface lipid phase variation in Neisseria
  – Fate choice in HIV-1 proviral latency?
  – Persistent switching in E. coli antibiotic resistance.

• What are the different flavors of phase variation?
• When exactly are such strategies fit?


Example: two environments, two moves

e.g. x=pili; y=no pili
E1=in host; E2=out
Two environments, two moves, perfect sensor

IF E1: selects for x, against y
E2: selects against x, for y

e.g. x=pili; y=no pili
E1=in host; E2=out
With no sensor, the options are...

1. ALL cells in state x
2. ALL cells in state y
3. Statically mixed population (some x, some y)
4. Phase variation of individual cells between x and y
With no sensor, the options are...

1. ALL cells in state x
2. ALL cells in state y
3. Statically mixed population (some x, some y)
4. Phase variation of individual cells between x and y
With no sensor, the options are...

1. ALL cells in state \( x \)
2. ALL cells in state \( y \)
3. Statically mixed population (some \( x \), some \( y \))
4. Phase variation of individual cells between \( x \) and \( y \)
With no sensor, the options are...

1. ALL cells in state x
2. ALL cells in state y
3. Statically mixed population (some x, some y)
4. Phase variation of individual cells between x and y
1. ALL cells in state x
2. ALL cells in state y
3. Statically mixed population (some x, some y)
4. Phase variation of individual cells between x and y
Phase variation for survival

This is a Devil’s compromise: Phase-variation behaviors is not optimal in any one environment but necessary for survival with noisy sensors in a fluctuating environment.
Evolutionary Game Model


Parametric Search For Fit Strategies

Low dimensional cuts through the Game Parameter Space

Pure
LPF-Pure
“symmetric lifestyle”

Pure
All x
“asymmetric lifestyle”
Learning Environment from Cell State

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Sensor profile</th>
<th>Environmental profile</th>
</tr>
</thead>
</table>
| No sensors            | O=Low prob. observable transitions over DC or extinction set. | • Devil’s Compromise (DC) lifecycle: time varying environment with different environmental states selecting for different cell states.  
                          | D=Long delays relative to env. transition times.                                                       | • Optimal switching rates a function of lifecycle asymmetries and environmental autocorrelation.  
                          |                                                      | • Time variation required (spatial variation insufficient).                                              |
| Sensor Based Mixed    | O=High prob. observable transitions; A=Poor accuracy  | Temporally or spatially varying environment with each environmental state selecting for a single cell state.                                           |
| Sensor Based Pure     | O=High prob. observable transitions; A=High accuracy; or moderate accuracy and low noise N.          |                                                                                                                                                        |
| Sensor Based Pure; LPF| O=High prob. observable transitions; A=Moderate accuracy. N=High additive noise.                       |                                                                                                                                                        |
Summary 1

• We are attempting to link evolutionary modularity to dynamic modularity.

• Molecular evolution provides a clue as to what parts of a network/module are under selection.

• Evolutionary Game Theory tells us about what strategies are likely to survive under different conditions.
Talk Overview

1. Review of central dogma and evolution
2. Review of basic cellular dynamics
3. The central questions in cellular network design
4. The game of life and optimal design
5. An early “pulling” application
Design of Therapeutic Bacteria

J. Christopher Anderson, Christopher Voigt, Adam Arkin

UC-Berkeley
UC-San Francisco
Many bacteria localize to tumors

- Intravenous injection of *Salmonella typhimurium*

- Localization occurs for many bacteria
  - *Vibrio cholerae*
  - *Clostridium*
  - *Bifidobacterium*
  - *Listeria monocytogenes*
  - *Bordetella pertussis*
  - *Escherichia coli DH5α*

- Phase I clinical trials in humans
- High doses have toxic side effects

*(Yu et al., 2004; Dang et al., 2001; Low et al., 1999; Yu et al., 2004)*

*(Tjivajev et al., 2001)*
Really smart drugs

“Smart” Liposome

- Recognition (Antibody)
- Liposome (programmed release)
- Drug

(Park et al., J. Controlled Release, 2001)

Bacteria

- Environmental Sensors
- Binding Specificity
- Protein Delivery
- Logic & Communication
- Protein & Chemical Synthesis
- Motility
The Abstraction Hierarchy

The System

Processes

Modules (Devices) 1
Modules (Devices) 2
Modules (Devices) ..n

E. coli strain JCA0934893

Parts
What's a Process?

Sensing → Signal Processing → Actuation

- Oxygen
- Glucose
- Lactate
- pH
- Touch sensor
- Specific adhesion
- Inverters
- AND Gates
- OR Gates
- Filters
- Switches
- Invasion
- O-Antigen
- K-Capsule
- Iron Acquisition
Engineering for Distinct Environments

Bloodstream  Tumor  Cell Surface  Intracellular
Engineering for Distinct Environments

Bloodstream
- High Glucose
- High $O_2$
- Low lactate

Tumor
- Low Glucose
- Low $O_2$
- Low lactate
- High Density

Cell Surface
- $\beta$1-Integrin
- TF Antigen
- Her2
- Sialyl Lewis X

Intracellular
- Vacuole
- Metabolites

Sensing
Engineering for Distinct Environments

Bloodstream: Shielding, No Growth
Tumor: Activation, Deshielding, No Growth
Cell Surface: Adhesion, Secretion, Invasion, No Growth
Intracellular: Growth, Lysis, Killing

Actuation
Environment 1: The Bloodstream
Single-Cell Targets

Leukemia
HIV-infected Cells
Autoimmune Disease
Design: Interfaces, Logic, Transfer Functions

Blood stream
Iron restriction for growth

Cell Surface
Intracellular

Tumor

Cell density

Regulation of O Antigen

Relative CUG

Pret Induction

O - \frac{W}{a + W}

Invasion

Cholesterol

Phospholipids

Lactate

Glucose

Amino Acids

O2 (mmHg)

pH

Vascular

Interstitial

Cellular

Iron restriction for growth

Blood

Stream
Metabolic safety catch

DAP Necessary for cell-wall biosynthesis and lysine production

No DAP in mammalian metabolism

Tight auxotroph; requires 40 ug/mL in LB

Infection of HeLa cells with MC892 (DdapD)/pAC-TetInv
Challenges

• Specificity of function
  – Selecting for sensors
  – Evolving optimal thresholds in the switches
  – Combining multiple signals at the right thresholds

• Preservation of function
  – Predicting load of circuit components on cell growth/fitness
  – Robust Design
  – Harmless failure modes

• Engineering for safety
  – Multiple sensitivities/dependencies
  – Genetic incompatibility?
Maximizing Efficacy: Time to takeover of dysfunctional mutants
Summary cont.

• The TKB provides a testbed for our ability to predict, control and design behaviors.

• It tests our ability to understand
  - Specific sensing of the microenvironment
  - Engineering for both robustness and sensitivity
  - Engineering against evolution
  - Engineering for game with the host immune system.
  - How to build a real synthetic biology application
Overall

- There is a critical need for new analytical methods for inferring, analyzing and designing cellular networks.

- Both pure understanding and (safe) applications are suffering from lack of a theoretical framework.

- But evolutionary and dynamic analyses are suggesting that there really are principles of cellular network design if only to improve evolvability.

- Data quality and quantity is roadblock but a smaller one that the lack of development of a theoretical foundation that can optimize experimental designs and derive understanding from integrated datasets.
Acknowledgements

Evolutionary Game Theory
- Denise Wolf
- Vijay Vazirani

B. Subtilis
- Denise Wolf
- Lisa Fontaine-Bodin
- Jay Keasling

Stochastic Theory
- Michael Samoilov

GTL/VIMSS
- http://vimss.lbl.gov

HIV Expression
- Leor Weinberger
- David Schaffer
- Jared Toettcher
- John Burnett

Tumor Killing Bacterium
- J. Chris Anderson
- Christopher Voigt

Funding
- DOE
- HHMI
- NIH
- DARPA