Quantitative Microbiology

PHYS 176/276
Instructor: Terry Hwa
Winter 2023

course website:
https://matisse.ucsd.edu/courses/w23-quant-microb/

What is quantitative biology?

⇒ quantitative biology ≠ biology + numbers/equations
   ≠ application of quant tools to bio
⇒ use numbers to gain predictive understanding of living systems

Why quantitative biology?

• because biology is quantitative
• needed to formulate and test falsifiable predictions
• demanded by synthetic biology

Role of theory

• formulate expectation and predictions (via quantitative model)
• guide the design of new experiments and technology
• power: the generality of (falsifiable) ideas, not necessarily math
  [e.g., Cupernicus, Darwin, Einstein]
• “cost” : the simplifying assumptions, not necessarily forced by math,
  but required in order to reveal principles

⇒ This course: quantitative (molecular) microbiology
Life of a bacterium:

**TABLE 1. Typical elemental composition of biological specimen**

<table>
<thead>
<tr>
<th>Element</th>
<th>Tissuea</th>
<th>Bacteriaa</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.50</td>
<td>0.47</td>
</tr>
<tr>
<td>N</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>H</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>O</td>
<td>0.25</td>
<td>0.23</td>
</tr>
<tr>
<td>P + S + others</td>
<td>0.02</td>
<td>0.10b</td>
</tr>
</tbody>
</table>

[Heldal et al, 1985]

- molar composition: $\text{CH}_{1.5}\text{O}_{0.35}\text{N}_{0.24} (+\text{S, P, Mg, Fe, ...})$
- algae (photosynthesis):
  $\text{CO}_2 + \text{H}_2\text{O} + \text{N}_2 + \text{photons} \rightarrow \text{biomass} + \text{O}_2$
- *E. coli* (minimal medium):
  $\text{glucose} + \text{NH}_3 \rightarrow \text{biomass} + \text{CO}_2$

Learning from the growth curve

[Monod, Ann Rev Microb. 1949]

- $\text{OD}_{600} = \text{biomass content}$
  
  
  $\text{[1 OD-ml} = 0.5\text{mg CDW} \sim 10^9\text{cells]}$
- saturation $\text{OD} \rightarrow \text{yield}$
- (lag: transition from pre-shift phase)

Can we predict GR & yield?

- environmental factors: nutrient types & conc, temperature, pH, osmolarity, drugs, ...
- genetic factors: enzymes & regulation
growth of E. coli

Can we predict GR & yield?

environmental factors:
- nutrient types & conc
- temperature, pH, osmolarity, drugs, …

genetic factors:
enzymes & regulation

What does it take to replicate a cell?

condition-dependent

biomass

DNA

RNA

protein

other

water

protein = defined sequence of 20 amino acids

protein synthesis: ribosomes

What does it take to replicate a cell?

condition-dependent

biomass

DNA

RNA

protein

other

water
protein = defined sequence of 20 amino acids

protein synthesis: ribosomes

>85% of all RNA
up to 1/3 of all proteins
protein = defined sequence of 20 amino acids

protein synthesis: ribosomes

- amino acids & ATP from metabolic reactions

代谢
- 代谢：
  - 固定和分解营养物质
    - 获得能量
    - 生成碳前体
    - 固定N、S、P、金属
- 糖酵解 (2x pyruvate):
  - 生成4x ATP
- 有氧呼吸 (6x CO₂) + 生成32x ATP
  - (需要O₂)

- 无氧发酵 (2x acetate + 2x CO₂):
  - 生成8x ATP

- 但许多生物即使有氧也会用发酵 (Crabtree效应)，为什么？
metabolism
- sequester & breakdown nutrients
  - derive energy
  - generate carbon precursors
  - sequester N, S, P, metals
- biosynthesis of building blocks
  - amino acid
  - nucleic acid
  - lipids
  - co-enzymes
- degradation/recycling (e.g., mRNA)
- typical biochemical reaction:
  \[ S + C\cdot b \leftrightarrow S\cdot b + C \]
  
  S: substrate
  b: component (e.g., CH$_3$, NH$_2$, e$^-$)
  C: co-enzyme
  (needed for difficult reactions)

  ➔ most reactions catalyzed by enzymes (proteins)
  ➔ flux of the products and “by-products” need to be balanced

metabolic control via coordinated regulation of enzyme abundance/activity

protein = defined sequence of 20 amino acids

protein synthesis: ribosomes

tRNA charging

free AA & uncharged tRNAs
protein = defined sequence of 20 amino acids

protein synthesis: ribosomes

\[ J_i = k_i [RS_i] \frac{[a_i]}{K_{a,i} + [a_i]} \frac{[tR_i]}{K_{tR,i} + [tR_i]} \]

~100 variables
~500 parameters

free AA & uncharged tRNAs

~100 variables
~500 parameters

parameter explosion

~500 enzymes
~500 metabolites
~5000 parameters

~10,000 parameters

regulation: when and how much proteins to make
dependence on temp, pH, osmolarity, …
How to deal with exploding no. of parameters?

Newtonian Mechanics \[ \frac{d^2 \vec{r}_i}{dt^2} = \vec{f}_{ij}(\vec{r}_i - \vec{r}_j) \]

Thermodynamics

\[ PV = nRT \]

Need moles of parameters:
\[ \vec{r}_i(t=0) = ... \]
\[ \vec{v}_i(t=0) = ... \]

Statistical Mechanics
\[ \rho(\vec{q}_i, \vec{p}_i; t) \xrightarrow{t \to \infty} \rho_{eq}(\mathcal{H}(\vec{q}_i, \vec{p}_i)) \]
This course: explore cellular strategies of dimension reduction using bacterial mechanisms
Molecular aspect: Central Dogma + regulation

- tRNA
- rRNA
- tRNA
- mRNA
- transcribed

- RNA synthesis
- transcribed

- transcription
- elongation and termination

- how does it all work together? what can it be predictively manipulated? what parameters are needed?

- tsx initiation control by transcription factors (TF)
- tsl initiation control by sRNA and RNA-binding proteins
- tsx termination control by sRNA and anti-terminators
- control of mRNA and protein degradation
- control of enzyme activity by metabolites

- coupled to environmental signals; coord growth program

- sugar, NH₃, O₂

molecular biology of transcription (RNA synthesis)

- transcriptional initiation

- elongation and termination
transcriptional initiation control

- modulation of RNAP-promoter affinity via activators and repressors

⇒ net result: rate of transcriptional initiation dependent on cellular conc of activators and repressors controlled by metabolites and signaling molecules

Gene regulatory networks

⇒ A grand challenge of Systems Biology
  - map out the complete wiring diagram of the cell
  - predictive computational model of the cell
Circuit diagram as system-level descriptor?

<table>
<thead>
<tr>
<th></th>
<th>electronic circuits</th>
<th>genetic circuits</th>
</tr>
</thead>
<tbody>
<tr>
<td>components</td>
<td>simple &amp; well-characterized; many (~10^9); fast (10^-9 sec)</td>
<td>heterogeneous, most rates unknown; few (~1000); slow (&gt; 10 min)</td>
</tr>
<tr>
<td>connectivity</td>
<td>physical interconnect between well-insulated components (1~2 inputs per node)</td>
<td>multiply-connected (1~10 inputs per node); regulation at all stages</td>
</tr>
<tr>
<td>network complexity</td>
<td>iterated cascades from complex network wiring</td>
<td>combinatorial signal integration from complex molecular control</td>
</tr>
</tbody>
</table>

Gene regulatory networks

- A grand challenge of Systems Biology
  - map out the complete wiring diagram of the cell
  - predictive computational model of the cell
**scope of this course**

- focus on simple systems (bacterial gene regulation)
- role of theory, modeling, and computation
- coarse-grained description at multiple scales (telescoped description)
- **quantitative connections between molecular mechanisms and physiological (functional or behavioral) characteristics**
- power of functional and mechanistic constraints

**Course content**

- molecular interactions: ligand-protein, protein-DNA, and protein-protein
- transcriptional control: activation, repression, and combinatorial
- modeling genetic circuits: bi-stability, oscillation, and stochasticity
- post-transcriptional control and functional enhancement
- from molecular interaction to cell physiology