ENGINEERING AND COMPUTATIONAL ISSUES IN INDUSTRIAL BIOLOGY

Implications for Commercializing Biology in the 21st Century

Babatunde Ogunnaike
DuPont CS&E

1. INTRODUCTION

- AN ENDURING PARADIGM
  - ➢First, Science, *then* its commercialization;
  - ➢Technology–the enabling “middleman”;
  - ➢Inexorable driving force: meeting societal needs

Evolutionary Trends?

- 19th Century
  - ➢ Science: PHYSICS
    - ◆ Origin in Newton’s 18th Century;
    - ◆ Spilling over into 20th Century;
  - ➢ Commercialization
    - ◆ from Edisonian light bulb to locomotive engine, to refrigeration, to skyscrapers …
  - ➢ Enabling Technology: (Mech.) Engineering

Evolutionary Trends?

- 20th Century
  - ➢Science: CHEMISTRY
    - ◆ with significant influence from Physics;
  - ➢Commercialization
    - ◆ modern chemical industry, still viable, but in need of re-energization…
    - ◆(Late dominance of "Parallel Universe" of "Electronics" arising from *integrating Physics*)
  - ➢Enabling Technology: (Chem.) Engineering

Evolutionary Trends?

- 21st Century
  - ➢Science: BIOLOGY
    - ◆ with significant influence from Physics *and* Chemistry;
  - ➢Commercialization
    - ◆Ongoing “gold rush”…
  - ➢Enabling Technologies: Still evolving
    - ◆Tantalizing observation: *any fundamental connection between “Information” and “Cells”?*
- Any clues from the past?

Issues to Consider

- “Proximity” of the Science to the relevant Engineering.
  - ➢How far does the “enabler” have to travel?
- Tools of the Engineering trade
  - ➢What is needed? Is it available?

2. COMMERCIALIZATION

- What is going on?
  - ➢Biology Research Drivers
  - ➢Biology R & D Activities
- Engineering Involvement

Biology Research Drivers

NIH funding ($15.6 billion in 1999, 82% extramural)

Pharmaceutical Industry(est. $26 billion U.S. R&D 2000)
-part of healthcare - largest market segment
-huge unmet needs - cancer, aging, heart disease
-most profitable industry
-highest investors in R&D as percent of sales
-baby boomers will pay anything to stay young forever

Agriculture ($X billion R&D)
-world population now 5E9 will be even higher in 2030

Materials ($Y billion R&D)
-sustainable alternatives to oil
-higher value products

Activities of Biology R&D

Laboratory work

hypothesis testing (drug target validation)
data generation (to discover new genes)

Hypothesis generation *(deep thinking by scientists/domain experts)*

"..we propose that stabilizing selection has maintained phenotypic constancy for *eve* expression but has allowed mutational turnover…"

Data organization

assembling genomes (human has 3E9 bases, ~140,000 genes)

Data analysis

developing and testing algorithms for recognizing genes and their control elements in DNA sequence

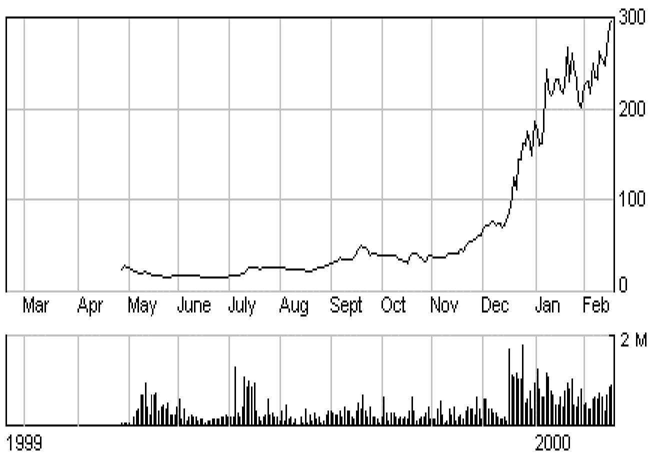
Technique development

new ways to turn genes off and on experimentally

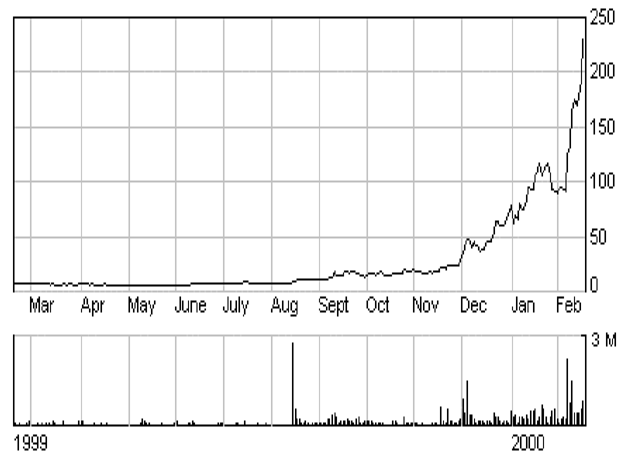
Highest Value Activities Currently

High throughput data gathering, organization, analysis
For example 2 companies with IPO's last year:

Celera Genomics ($8.2E9 mkt. cap)
faster sequencing,
first to finish human genome



Curagen Corporation ($3.6E9 mkt. cap)
first to map protein-protein interactions
for an organism (yeast)



Engineering-like activities

Industrialization of Biology:

Celera -- old sequencing machines, new machines
Curagen -- industrialized gene expression measurement with gels, now industrializing 2-hybrid assay.

Both cases:

Machines/robotics to aid data generation
Computer software to organize, search, analyze, present data

Involvement of Engineers

Not intense domain knowledge of expert;
Understand generic molecular biology;
Know promising new techniques;
Be first to scale-up techniques;
Have expertise in computer technology --
web tools, databases, interfaces, algorithm design
(industry average for bioinformatician is $94K/yr)

Future Targets for Engineers in Biology

- Invent important techniques and methods
- Models of biological systems -- make biology predictive
  - ➢Entire cells, Diseases (cf. Entelos, Inc)
- Invent in “Molecular Biology” too?
  - ➢ See Yeast 2-hybrid (*Nature*:**340**,1989)!

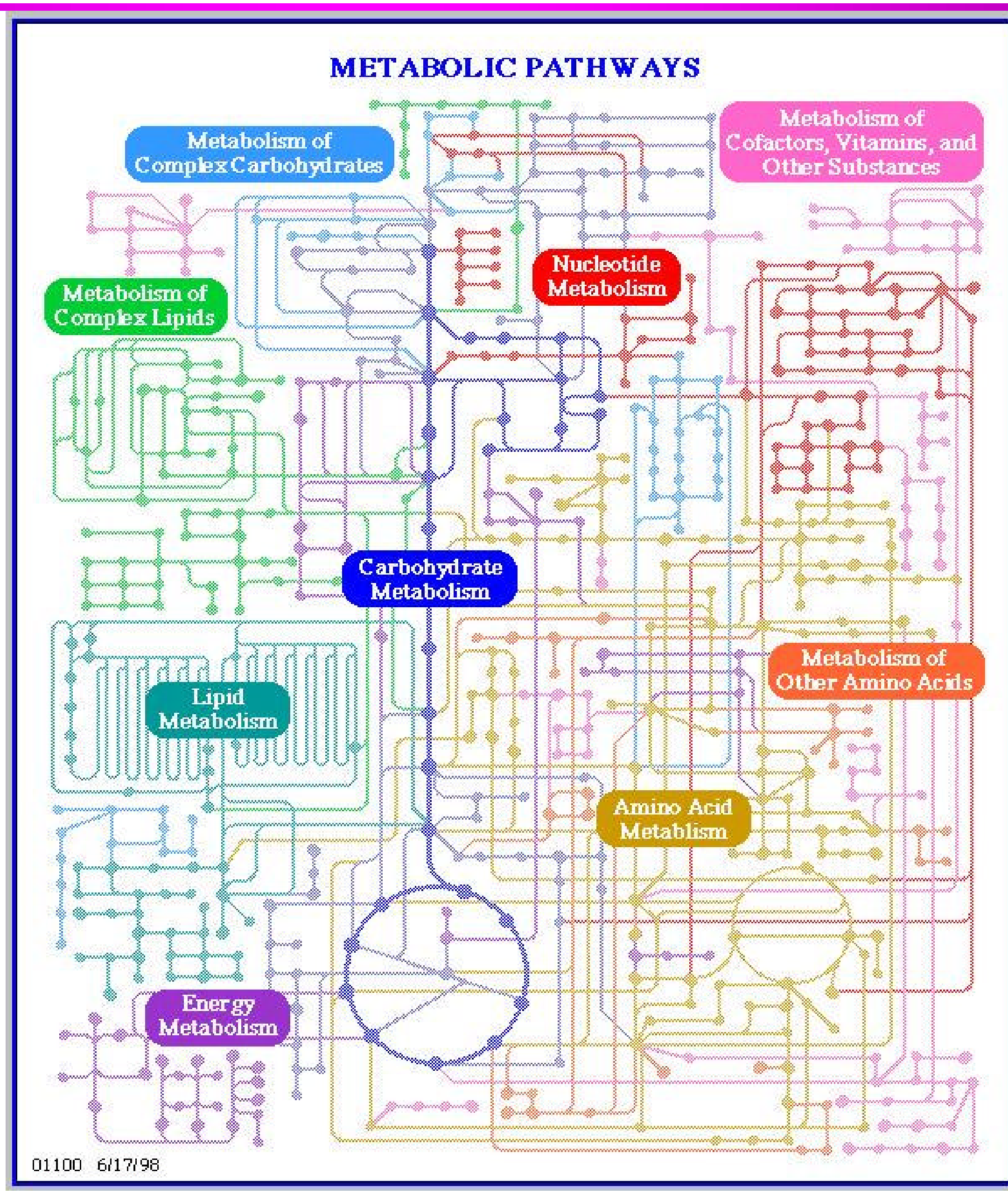
Models of Biological Systems

Just metabolism:

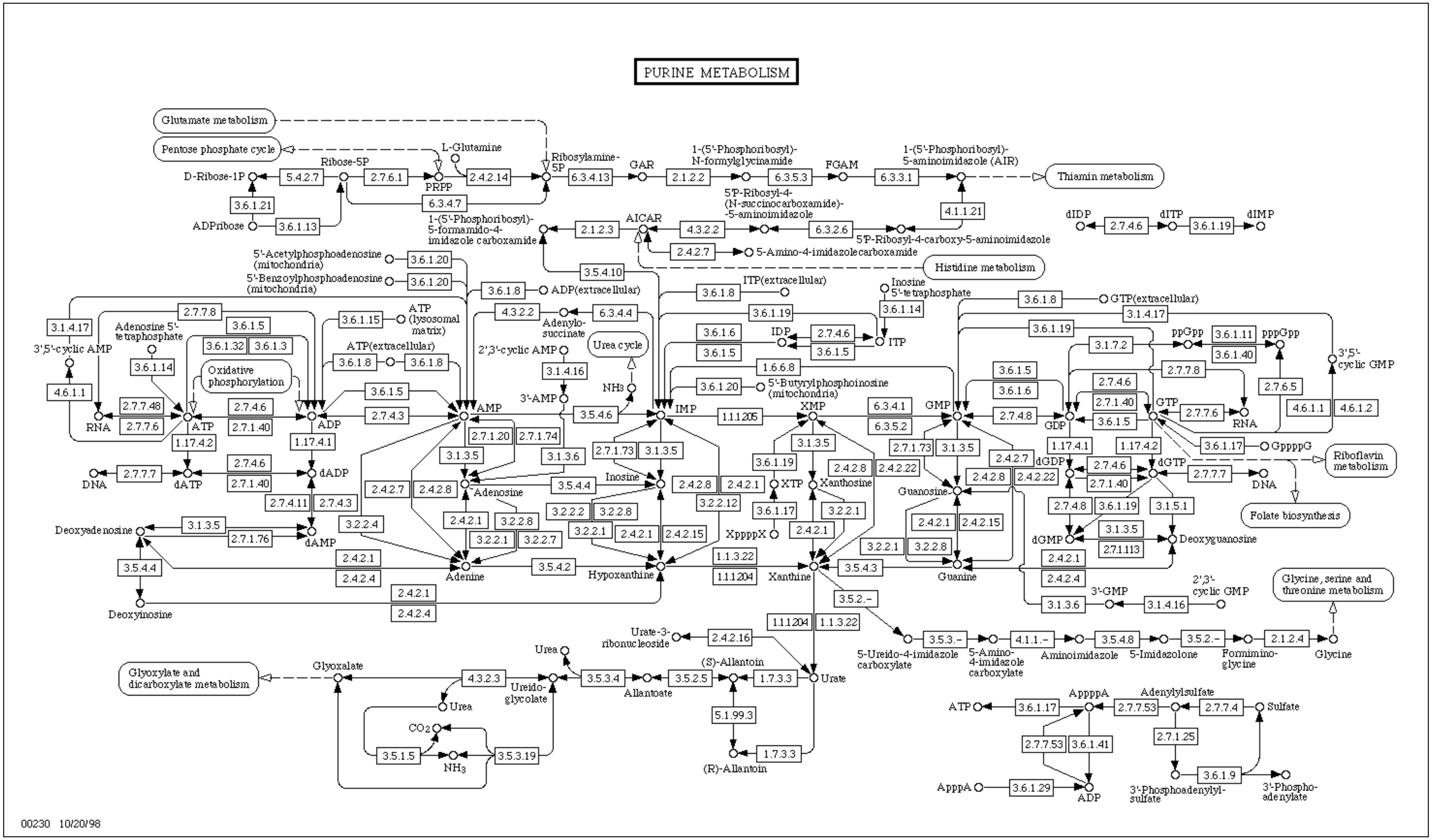
thousands of enzymes

thousands of reactants and products

reaction rates depend on concentrations of non-participants



One of ~80 such maps for metabolism:



Can we formulate and solve $dx/dt = f(x,t) + \mathrm{flux}(t)$?

Not Yet

Unknown and poorly defined parts:

Human genes, a poorly characterized system:
Human has ~100,000 genes; We understand ~10,000.
(We just don't know all the pieces yet, But we have to start thinking about how we will take on these complex systems.)

Basic Metabolism, a well studied system:
Know the genes, but can only poorly measure their concentrations and in-vivo kinetics.

Know the reactants and products, but cannot measure their in-vivo concentrations.

3. SUMMARY & CONCLUSIONS

- Commercializing biology in the 21st century:
  - ➢requires integration of sciences, and engineering;
  - ➢Can we learn something from the past?
  - ➢analogy to “gold rush” (e.g. providing infrastructure for the “forty-niners”);
- Role of Engineers
  - ➢Wide variety of roles; mostly computational

ACKNOWLEDGEMENTS

- Bob LaROSSA (DuPont BCS & E):
- Bill FOSTER (DuPont Pharma)