21st Century Biology:
Informatics in the Post-Genome Era

Robert J. Robbins
Fred Hutchinson Cancer Research Center
1124 Columbia Street, LV-101
Seattle, Washington 98104

rrobbins@fhcrc.org
(206) 667 4778
• Biotechnology will be the “magic” technology of the 21st Century.
• Biotechnology will be the “magic” technology of the 21st Century.
• Information Technology (IT) has a special relationship with biology.
Topics

- Biotechnology will be the “magic” technology of the 21st Century.
- Information Technology (IT) has a special relationship with biology.
- Moore’s Law constantly transforms IT (and everything else).
Topics

• Biotechnology will be the “magic” technology of the 21st Century.
• Information Technology (IT) has a special relationship with biology.
• Moore’s Law constantly transforms IT (and everything else).
• Current funding mechanisms for bio-information infrastructure are hopelessly inadequate to meet future needs and must be radically reformed.
Introduction

Magical Technology
To a person from 1897, much current technology would seem like magic.
To a person from 1897, much current technology would seem like magic.

What technology of 2097 would seem magical to a person from 1997?
To a person from 1897, much current technology would seem like magic.

What technology of 2097 would seem magical to a person from 1997?

**Candidate:** Biotechnology so advanced that the distinction between living and non-living is blurred.
IT-Biology Synergism
IT is Special

Information Technology:

- affects the performance and the management of tasks
Information Technology:

- affects the performance and the management of tasks
- allows the manipulation of huge amounts of highly complex data
IT is Special

Information Technology:

• *affects the performance and the management of tasks*

• *allows the manipulation of huge amounts of highly complex data*

• *is incredibly plastic*
  (programming and poetry are both exercises in pure thought)
IT is Special

Information Technology:

• affects the performance and the management of tasks
• allows the manipulation of huge amounts of highly complex data
• is incredibly plastic (programming and poetry are both exercises in pure thought)
• improves exponentially (Moore’s Law)
Biology is Special

Life is Characterized by:

- individuality
Biology is Special

Life is Characterized by:

- *individuality*
- *historicity*
Life is Characterized by:

- individuality
- historicity
- contingency
Biology is Special

Life is Characterized by:

- *individuality*
- *historicity*
- *contingency*
- *high (digital) information content*
Biology is Special

Life is Characterized by:

• *individuality*

• *historicity*

• *contingency*

• *high (digital) information content*

No law of large numbers...
Biology is Special

Life is Characterized by:

- *individuality*
- *historicity*
- *contingency*
- *high (digital) information content*

No law of large numbers, since every living thing is genuinely unique.
• Physics needs calculus, the method for manipulating information about statistically large numbers of vanishingly small, independent, equivalent things.
IT-Biology Synergism

- Physics needs calculus, the method for manipulating information about statistically large numbers of vanishingly small, independent, equivalent things.

- Biology needs information technology, the method for manipulating information about large numbers of dependent, historically contingent, individual things.
For it is in relation to the statistical point of view that the structure of the vital parts of living organisms differs so entirely from that of any piece of matter that we physicists and chemists have ever handled in our laboratories or mentally at our writing desks.

Erwin Schrödinger. 1944. *What is Life.*
Genetics as Code

[The] chromosomes ... contain in some kind of code-script the entire pattern of the individual's future development and of its functioning in the mature state. ... [By] code-script we mean that the all-penetrating mind, once conceived by Laplace, to which every causal connection lay immediately open, could tell from their structure whether [an egg carrying them] would develop, under suitable conditions, into a black cock or into a speckled hen, into a fly or a maize plant, a rhodo-dendron, a beetle, a mouse, or a woman.

Erwin Schrödinger. 1944. *What is Life.*
We now know that Schrödinger’s mysterious human “code-script” consists of 3.3 billion base pairs of DNA.
We now know that Schrödinger’s mysterious human “code-script” consists of 3.3 billion base pairs of DNA.

Typed in 10-pitch font, one human sequence would stretch for more than 5,000 miles. Digitally formatted, it could be stored on one CD-ROM. Biologically encoded, it fits easily within a single cell.
Bio-digital Information

DNA is a highly efficient digital storage device:

- There is more mass-storage capacity in the DNA of a side of beef than in all the hard drives of all the world’s computers.
Bio-digital Information

DNA is a highly efficient digital storage device:

- There is more mass-storage capacity in the DNA of a side of beef than in all the hard drives of all the world’s computers.

- Storing all of the (redundant) information in all of the world’s DNA on computer hard disks would require that the entire surface of the Earth be covered to a depth of three miles in Conner 1.0 gB drives.
Genomics: An Example
Computers are not just tools for cataloging existing knowledge. They are instruments that change the way we can see the biological world. Computers allow us to see genomes, just as radio telescopes let us see quasars and microscopes let us see cells.
Human Genome Project - Goals

- construction of a high-resolution genetic map of the human genome;

Human Genome Project - Goals

- construction of a high-resolution genetic map of the human genome;
- production of a variety of physical maps of all human chromosomes and of the DNA of selected model organisms;

Human Genome Project - Goals

- construction of a high-resolution genetic map of the human genome;
- production of a variety of physical maps of all human chromosomes and of the DNA of selected model organisms;
- determination of the complete sequence of human DNA and of the DNA of selected model organisms;

Human Genome Project - Goals

- construction of a high-resolution genetic map of the human genome;
- production of a variety of physical maps of all human chromosomes and of the DNA of selected model organisms;
- determination of the complete sequence of human DNA and of the DNA of selected model organisms;
- development of capabilities for collecting, storing, distributing, and analyzing the data produced;

Human Genome Project - Goals

- construction of a high-resolution genetic map of the human genome;
- production of a variety of physical maps of all human chromosomes and of the DNA of selected model organisms;
- determination of the complete sequence of human DNA and of the DNA of selected model organisms;
- development of capabilities for collecting, storing, distributing, and analyzing the data produced;
- creation of appropriate technologies necessary to achieve these objectives.

Growth in GenBank is exponential. More data were added in the last 10 weeks than were added in the first 10 years of the project.
Infrastructure and the HGP

Progress towards all of the [Genome Project] goals will require the establishment of well-funded centralized facilities, including a stock center for the cloned DNA fragments generated in the mapping and sequencing effort and a data center for the computer-based collection and distribution of large amounts of DNA sequence information.

[The] database developer should provide, in some real sense, an intellectual focus for the interpretation of genomic data.

NIH-DOE Ad Hoc Committee on Genome Databases
Paradigm Shift in Biology

The new paradigm, now emerging, is that all the ‘genes’ will be known (in the sense of being resident in databases available electronically), and that the starting point of a biological investigation will be theoretical. An individual scientist will begin with a theoretical conjecture, only then turning to experiment to follow or test that hypothesis.

To use [the] flood of knowledge, which will pour across the computer networks of the world, biologists not only must become computer literate, but also change their approach to the problem of understanding life.

Human Resources Issues

- Reduction in need for non-IT staff
Human Resources Issues

- Reduction in need for non-IT staff
- Increase in need for IT staff, especially “information engineers”
Human Resources Issues

- Reduction in need for non-IT staff
- Increase in need for IT staff, especially “information engineers”

In modern biology, a general trend is to convert expert work into staff work and finally into computation. New expertise is required to design, carry out, and interpret continuing work.
Human Resources Issues

Elbert Branscomb: “You must recognize that some day you may need as many computer scientists as biologists in your labs.”
Human Resources Issues

Elbert Branscomb: “You must recognize that some day you may need as many computer scientists as biologists in your labs.”

Craig Venter: “At TIGR, we already have twice as many computer scientists on our staff.”

Exchange at DOE workshop on high-throughput sequencing.
21st Century Biology

The Science
The fundamental dogma of molecular biology is that genes act to create phenotypes through a flow of information from DNA to RNA to proteins, to interactions among proteins, and ultimately to phenotypes.

Collections of individual phenotypes, of course, constitute a population.
Although a few databases already exist to distribute molecular information,
Although a few databases already exist to distribute molecular information, the post-genomic era will need many more to collect, manage, and publish the coming flood of new findings.
21st Century Biology

The Literature
** Locus Detail View **

Symbol: HBB
Name: hemoglobin, beta
MIM Num: 141900
Location: 11p15.5
Created: 01 Jan 86 00:00

** Polymorphism Table **

<table>
<thead>
<tr>
<th>Probe</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta-globin cDNA</td>
<td>Rsal</td>
</tr>
<tr>
<td>beta-globin cDNA,JW10+</td>
<td>Avall</td>
</tr>
<tr>
<td>Pstbeta,JW102,BD23,pB+</td>
<td>BamHI</td>
</tr>
<tr>
<td>pRK29,Unknown</td>
<td>HindII</td>
</tr>
<tr>
<td>beta-IVS2 probe</td>
<td>Hphi</td>
</tr>
<tr>
<td>IVS-2 normal</td>
<td>Hphi</td>
</tr>
<tr>
<td>Unknown</td>
<td>AvrII</td>
</tr>
<tr>
<td>beta-IVS2 probe</td>
<td>Asul</td>
</tr>
</tbody>
</table>
** Locus Detail View **

<table>
<thead>
<tr>
<th>Symbol:</th>
<th>HBB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name:</td>
<td>hemoglobin, beta</td>
</tr>
<tr>
<td>MIM Num:</td>
<td>141900</td>
</tr>
<tr>
<td>Location:</td>
<td>11p15.5</td>
</tr>
<tr>
<td>Created:</td>
<td>01 Jan 86 00:00</td>
</tr>
</tbody>
</table>

** Polymorphism Table **

<table>
<thead>
<tr>
<th>Probe</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta-globin cDNA</td>
<td>Rsal</td>
</tr>
<tr>
<td>beta-globin cDNA,JW10+</td>
<td>AvaII</td>
</tr>
<tr>
<td>Pstbeta,JW102,BD23,pB+</td>
<td>BamHI</td>
</tr>
<tr>
<td>pRK29,Unknown</td>
<td>HindII</td>
</tr>
<tr>
<td>beta-IVS2 probe</td>
<td>Hphl</td>
</tr>
<tr>
<td>IVS-2 normal</td>
<td>Hphl</td>
</tr>
<tr>
<td>Unknown</td>
<td>AvrII</td>
</tr>
<tr>
<td>beta-IVS2 probe</td>
<td>Asul</td>
</tr>
</tbody>
</table>

The alpha and beta loci determine the two types of polypeptide chains in adult hemoglobin, Hb A. By autoradiography using heavy-labeled hemoglobin-specific messenger RNA, Price et al. (1972) found labeling of a chromosome 2 and a group B chromosome. They concluded, incorrectly as it turned out, that the beta-gamma-delta linkage group was on a group B chromosome since the zone of labeling was longer on that chromosome than on chromosome 2 (which by this reasoning would be identified as group A).
**O M I M -- Beta Hemoglobin**

**Locus Detail View**

---

Symbol: HBB  
Name: hemoglobin, beta  
MIM Num: 141900  
Location: 11p15.5  
Created: 01 Jan 86 00:00

---

**Polymorphism Table**

<table>
<thead>
<tr>
<th>Probe</th>
<th>Enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta-globin cDNA</td>
<td>Rsal</td>
</tr>
<tr>
<td>beta-globin cDNA,JW10+</td>
<td>Avall</td>
</tr>
<tr>
<td>Pstbeta,JW102,BD23,pB+</td>
<td>BamHI</td>
</tr>
<tr>
<td>pRK29,Unknown</td>
<td>HindII</td>
</tr>
<tr>
<td>beta-IVS2 probe</td>
<td>HphI</td>
</tr>
<tr>
<td>IVS-2 normal</td>
<td>HphI</td>
</tr>
<tr>
<td>Unknown</td>
<td>AvrII</td>
</tr>
<tr>
<td>beta-IVS2 probe</td>
<td>Asul</td>
</tr>
</tbody>
</table>

---

OMIM -- Beta Hemoglobin

GENETICINFORMATION

**Locus**

*141900 HEMOGLOBIN--BETA LOCUS  
[HEMOGLOBIN, BETA; SICKLE CELL ANEMIA, INCLUDED; BETA-THALASSEMIAS, INCLUDED; HEINZ BODY ANEMIAS, BETA-GLOBIN TYPE, ...]

The alpha and beta loci determine the types of polypeptide chains in adult hemoglobin, Hb A. By using heavy-labeled human hemoglobin-specific messenger RNA, Price et al. (1972) found labeling of a chromosome 2 and a group B chromosome. They concluded, incorrectly as it turned out, that the beta-gamma-delta linkage group was on a group B chromosome than on chromosome 2 (which by this reasoning would be the alpha-gamma-delta linkage group).
Beta Hemoglobin

**DEFINITION**

[HBHU] Hemoglobin beta chain
- Human, chimpanzee, pygmy chimpanzee, and gorilla

**SUMMARY**

<table>
<thead>
<tr>
<th>Type</th>
<th>Protein</th>
<th>Molecular-weight</th>
<th>Length</th>
<th>Checksum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>15867</td>
<td>146</td>
<td>1242</td>
</tr>
</tbody>
</table>

**SEQUENCE**

V H L T P E E K S A V T A L W G
K K V L G A F S D G L A H L D N
L K G T F A T L S E L H C D K LH V D P E N F R L L G N V L V C
V L A H H F G

**GenBank -- Beta Hemoglobin**

**DEFINITION**

[HUMHBB] Human beta globin region

**LOCUS**

HUMHBB

**ACCESSION NO.**

J00179 J00093 J00094 J00096 J00158 J00159 J00160 J00161

**KEYWORDS**

Alu repetitive element; HPFH; KpnI repetitive sequence; RNA polymerase III; allelic variation; alternate cap site;

**SEQUENCE**

gaattctaatctccctctcactactgtctagt
atccctcaaggagtggtggctcatgtcttgagctcaagagtttgatataaaaaaaaattagcca
ggcaaatgggaggatcccttgagcgcactcca
gcct

**OMIM -- Beta Hemoglobin**

**Title**

*141900 HEMOGLOBIN--BETA LOCUS
[HBB; SICKLE CELL ANEMIA, INCLUDED; BETA-THALASSEMIAS, INCLUDED; HEINZ BODY ANEMIAS, BETA-GLOBIN TYPE, ...]

The alpha and beta loci determine the types of polypeptide chains in adult hemoglobin, Hb A. By autoradiography using heavy-labeled messenger RNA, Price et al. (1972) found labeling of a chromosome 2 and a group B chromosome. They concluded, incorrectly as it turned out, that the beta-gamma-delta linkage group was on a group B chromosome, since the zone of labeling was longer on that chromosome than on chromosome 2 (which by this reasoning...
The ESP site is dedicated to the electronic publishing of scientific and other scholarly materials. Of particular interest are the history of science, genetics, computational biology, and genome research.
The Classical Genetics: Foundations series provides ready access to typeset-quality, electronic editions of important publications that can otherwise be very difficult to find.
“Hardy” (of Hardy-Weinberg) is a name well known to most students of biology.
But how many have read, or even seen, **all** of Hardy’s biological writings?

This is it: A single, one-page letter to the editor of *Science*.
Electronic Scholarly Publishing

http://www.blocks.fhcrc.org/~kinesin

Electronic publishing is especially appropriate for some kinds of dynamic review papers.
Kinesin

Kinesin is a mechanochemical protein capable of utilizing chemical energy from ATP hydrolysis to generate mechanical force. In the presence of ATP, kinesin can bind to and move on microtubules (see Motility). The ability to translocate along the microtubule lattice has led to the classification of kinesin as a microtubule motor protein. Kinesin is unrelated in sequence to the other known class of microtubule motor proteins, the dyneins, and is thought to perform functions in the cell distinct from the dyneins.

The mechanism by which molecular motor proteins convert energy from ATP hydrolysis into mechanical force is not known. A problem related to their mechanism of function is the molecular basis of polarity of translocation along the microtubule: some kinesin motors move toward microtubule minus ends, instead of plus ends like 'conventional' kinesin. The coupling of the ATPase cycle to force generation and the determinants of motor polarity are actively being investigated using biochemical, biophysical, and molecular approaches.

Other areas of investigation include the regulation of kinesin function in the cell and identification of proteins that enable kinesin to interact with intracellular vesicles and organelles.

Recent Reviews

Return to the Kinesin Home Page.
and it is easy to provide large amounts of in-depth supporting and related data.
The C-terminal Motor Subfamily

The C-terminal motor subfamily in the most recent analysis is no longer supported by a high bootstrap value. This is an indication of divergence within the group due to newly discovered members of the subfamily. The tree shown is taken from a tree built in a search of 73 kinesin motor domains. The kinesin proteins in this group have in common a C-terminal motor domain, and 4 members of the group (DmNcd, ScKAR3, CgCHO2, AtKCBP) have now been demonstrated to be minus-end directed motors.

<table>
<thead>
<tr>
<th>Species/protein</th>
<th>Molecular mass (kDa)</th>
<th>Motor polarity &amp; velocity</th>
<th>Subcellular localization</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. thaliana KatA</td>
<td>89</td>
<td>ND</td>
<td>ND</td>
<td>Calmodulin binding</td>
</tr>
<tr>
<td>A. thaliana KatB</td>
<td>82</td>
<td>ND</td>
<td>ND</td>
<td>Calmodulin binding</td>
</tr>
<tr>
<td>A. thaliana KatC</td>
<td>84</td>
<td>ND</td>
<td>ND</td>
<td>Calmodulin binding</td>
</tr>
<tr>
<td>A. thaliana KCBP</td>
<td>143</td>
<td>Minus, 8-10</td>
<td>ND</td>
<td>Calmodulin binding</td>
</tr>
</tbody>
</table>
Electronic Scholarly Publishing

Entire monographs can be made instantly available to readers world-wide.

http://www.esp.org/books/darwin/beagle
Today’s computer technology was nearly unimaginable just ten years ago. The technology of ten years from now will also bring many surprises.

How is it that IT can maintain such an amazing rate of sustained change?

And what, if any, are the implications of that rate of change for biology?
Moore’s Law

Transforms InfoTech
(and everything else)
Moore’s Law: The Statement

Every eighteen months, the number of transistors that can be placed on a chip doubles.

Gordon Moore, co-founder of Intel...
Moore’s Law: *The Effect*

![Graph showing Moore's Law](image)

- **Performance (constant cost)**
- **Time**

Legend: The line indicates a linear increase in performance over time, with time on the x-axis and performance on the y-axis.
Moore’s Law: The Effect

The graph shows the relationship between performance and cost over time. The x-axis represents time, while the y-axis for performance shows constant cost, and the y-axis for cost shows constant performance. The graph illustrates how performance increases as cost decreases over time.
Moore’s Law: The Effect

Three Phases of Novel IT Applications

- It’s Impossible
Moore’s Law: *The Effect*

Three Phases of Novel IT Applications

- It’s Impossible
- It’s Impractical
Moore’s Law: *The Effect*

Three Phases of Novel IT Applications

- It’s Impossible
- It’s Impractical
- It’s Overdue
Moore’s Law: The Effect

![Graph showing the relationship between performance, cost, and time according to Moore's Law. The graph illustrates a linear increase in performance and cost over time.]
Moore’s Law: The Effect

![Diagram showing the relationship between performance, cost, and time according to Moore's Law.]

- Performance: constant cost
- Cost: constant performance
- Time: increasing
Moore’s Law: The Effect

![Graph showing the relationship between Performance and Cost over Time.](image-url)
Moore’s Law: *The Effect*

![Graph showing the relationship between time, performance, and cost according to Moore's Law.](image-url)
Moore’s Law: The Effect

![Graph illustrating Moore's Law with axes labeled: Performance (constant cost) on the y-axis, Cost (constant performance) on the x-axis, and Time on a horizontal axis. The graph shows a linear increase in Performance with Time and a linear decrease in Cost with Time. Points labeled P and C on the graph represent specific performance and cost levels over time.]
Moore’s Law: The Effect

![Diagram showing the relationship between performance, cost, and time according to Moore's Law. The graph illustrates how performance increases with time while cost decreases, maintaining a constant performance level. Key points labeled D, A, and C indicate specific performance and cost perspectives at different time intervals.]
Moore’s Law: The Effect

<table>
<thead>
<tr>
<th>Time</th>
<th>Performance (constant cost)</th>
<th>Cost (constant performance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000</td>
<td>10,000</td>
<td>100,000</td>
</tr>
<tr>
<td>100</td>
<td>1,000</td>
<td>10,000</td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>1,000</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td>0.1</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>
Moore’s Law: *The Effect*

- **Performance (constant cost)**
- **Cost (constant performance)**
- **Time**

Relevance for biology?
Cost (constant performance)
Cost (constant performance)
Cost (constant performance)
Cost (constant performance)
Funding for Information Infrastructure
The Problem

- IT moves at “Internet Speed” and responds rapidly to market forces.
The Problem

• IT moves at “Internet Speed” and responds rapidly to market forces.

• IT will play a central role in 21st Century biology.
The Problem

• IT moves at “Internet Speed” and responds rapidly to market forces.
• IT will play a central role in 21st Century biology.
• Current levels of support for public bio-information infrastructure are too low.
The Problem

- IT moves at “Internet Speed” and responds rapidly to market forces.
- IT will play a central role in 21st Century biology.
- Current levels of support for public bio-information infrastructure are too low.
- Reallocation of federal funding is difficult, and subject to political pressures.
The Problem

- IT moves at “Internet Speed” and responds rapidly to market forces.
- IT will play a central role in 21st Century biology.
- Current levels of support for public bio-information infrastructure are too low.
- Reallocation of federal funding is difficult, and subject to political pressures.
- Federal-funding decision processes are ponderously slow and inefficient.
Federal Funding of Bio-Databases

The challenges:
Federal Funding of Bio-Databases

The challenges:

• providing adequate funding levels
Federal Funding of Bio-Databases

The challenges:

• providing adequate funding levels
• making timely, efficient decisions
Federal Funding of Bio-Databases

Appropriate funding level:
Federal Funding of Bio-Databases

Appropriate funding level:

• approx. 10% of research funding
Federal Funding of Bio-Databases

Appropriate funding level:

- approx. 10% of research funding
- *i.e.*, 1 - 2 billion dollars per year
Federal Funding of Bio-Databases

Appropriate funding level:

- approx. 10% of research funding
- *i.e.*, 1 - 2 billion dollars per year

Source of estimate:
- Experience of IT-transformed industries.
- Current support for IT-rich biological research.
In a simple market economy, vendors try to anticipate the needs of buyers and offer products and services to meet those needs.
Market Forces

In a simple market economy, vendors try to anticipate the needs of buyers and offer products and services to meet those needs.

Real users decide whether or not to buy a product or service, depending upon whether or not it meets a real need at a reasonable price.
In a simple market economy, vendors try to anticipate the needs of buyers and offer products and services to meet those needs.

Real users decide whether or not to buy a product or service, depending upon whether or not it meets a real need at a reasonable price.

**Business 101 Insight:**
Successful vendors target a niche and excel at meeting the needs of that niche.
Market Forces

Funding to initiate the development of products and services come from investors, not from buyers.
Funding to initiate the development of products and services come from investors, not from buyers. Investors decide whether or not to provide start-up funding based upon the estimated ability of the vendor to create products and services that will meet real needs at competitive prices.
If biological databases were driven by market forces, individual users would choose what services they need and individual database providers would choose what services to make available.
If biological databases were driven by market forces, individual users would choose what services they need and individual database providers would choose what services to make available.

Investors would provide start-up money on the likelihood of successful products and services being developed.
If biological databases were driven by market forces, individual users would choose what services they need and individual database providers would choose what services to make available.

Investors would provide start-up money on the likelihood of successful products and services being developed.

Ultimate success would depend on meeting the needs of real users. Decisions could be made rapidly, in response to changing needs and emerging opportunities.
Instead, funding decisions for biological databases can follow a ponderously slow course, with almost no opportunity for input from real users.

Those most knowledgeable about a particular database are often excluded from participating in the review process because of a possible “conflict of interest” status with the database provider.
Federal Funding of Bio-Databases

Possible solution - create market forces:
Federal Funding of Bio-Databases

Possible solution - create market forces:

• stop supporting the supply side of biodatabases through slow, inefficient processes.
Federal Funding of Bio-Databases

Possible solution - create market forces:

- stop supporting the supply side of biodatabases through slow, inefficient processes.
- start supporting the demand side through fast, efficient processes.
Federal Funding of Bio-Databases

Possible solution - create market forces:

• stop supporting the supply side of biodatabases through slow, inefficient processes.
• start supporting the demand side through fast, efficient processes.
• provide guaranteed supplementary funding, redeemable only for access to bio-databases.
Federal Funding of Bio-Databases

Possible solution - create market forces:

• stop supporting the supply side of biodatabases through slow, inefficient processes.
• start supporting the demand side through fast, efficient processes.
• provide guaranteed supplementary funding, redeemable only for access to bio-databases.
• data stamps
Federal Funding of Bio-Databases

Possible solution - create market forces:

- stop supporting the supply side of biodatabases through slow, inefficient processes.
- start supporting the demand side through fast, efficient processes.
- provide guaranteed supplementary funding, redeemable only for access to bio-databases.
- data stamps, AKA food (for-thought) stamps ?!
Food (for thought) Stamps

Funding Agencies could:

- provide a 10% supplement to every research grant in the form of “stamps” redeemable only at database providers.
Food (for thought) Stamps

Funding Agencies could:

• provide a 10% supplement to every research grant in the form of “stamps” redeemable only at database providers.

• allow the “stamps” to be transferable among scientists, so that a market for them could emerge.
Funding Agencies could:

• provide a 10% supplement to every research grant in the form of “stamps” redeemable only at database providers.

• allow the “stamps” to be transferable among scientists, so that a market for them could emerge.

• provide funding only after the stamps have been redeemed at a database provider.
Food (for thought) Stamps

Problems:

- how to estimate the amount of FFT stamps that would actually be redeemed (and thus the required budget set-aside).
Food (for thought) Stamps

Problems:

• how to estimate the amount of FFT stamps that would actually be redeemed (and thus the required budget set-aside).

• how to identify “approved” database providers.
Food (for thought) Stamps

Problems:

• how to estimate the amount of FFT stamps that would actually be redeemed (and thus the required budget set-aside).

• how to identify “approved” database providers.

• how to initiate the FFT system.
Food (for thought) Stamps

Problems:

• how to estimate the amount of FFT stamps that would actually be redeemed (and thus the required budget set-aside).
• how to identify “approved” database providers.
• how to initiate the FFT system.
• etc etc
Alternatives (if no solution emerges):

- increasingly inefficient research activities (abject failure will occur when it becomes easier to redo research than to obtain the results of prior work).
Alternatives (if no solution emerges):

- increasingly inefficient research activities (abject failure will occur when it becomes easier to redo research than to obtain the results of prior work).
- loss of access to bio-databases for public-sector research.
Alternatives (if no solution emerges):

- increasingly inefficient research activities (abject failure will occur when it becomes easier to redo research than to obtain the results of prior work).
- loss of access to bio-databases for public-sector research.
- movement of majority of “important” biological research into the private sector.
Food (for thought) Stamps

Alternatives (if no solution emerges):

• increasingly inefficient research activities (abject failure will occur when it becomes easier to redo research than to obtain the results of prior work).
• loss of access to bio-databases for public-sector research.
• movement of majority of “important” biological research into the private sector.
• loss of American pre-eminence (if other countries solve the problems first).
Food (for thought) Stamps

Bottom Line:

• This might not be the answer, but
Food (for thought) Stamps

Bottom Line:

• This might not be the answer, but
• it’s FOOD FOR THOUGHT
Slides:

http://www.esp.org/rjr/beckman.pdf