Half of Our Biosphere Is Missing

Half of Our Biosphere Is No Longer Missing

No Longer

Half of Our Biosphere Is Missing

Implications for Understanding Biodiversity, Bioinformatics, and Life Itself

No Longer

Half of Our Biosphere Is Missing

Implications for Understanding Biodiversity, Bioinformatics, and Life Itself

Robert J. Robbins

rrobbins@gmail.com

(http://www.rj-robbins.com/slides/RJR-GBIC-2012.pdf)

Abstract:

In the last few years, improving genomic and metagenomic tools have been revealing details about the previously invisible microbial world. Although these new findings are yielding very important insights about global biodiversity, some may be difficult to accommodate in current biological models. Half of the world's biomass and by far the majority of its biodiversity are, for the first time, becoming available for study. Every month, startling results appear in the literature. Not only do free-living microbes represent a distinctly alien way of life, commensal microbes are proving to have profound effects on their host organisms, affecting things ranging from mate choice in Drosophila, to pain tolerance in mice, to niche partitioning in ants. The emerging pervasive influence of commensal microbes suggests that to fully understand the biology of any organism we must take into account its interactions with its associated microbiota. It seems, we are all lichens now. The resulting conceptual adjustments will offer great opportunities for expanding our understanding of the biosphere, but will offer real challenges to our current view of biodiversity and will greatly complicate the informatics tools needed to document biodiversity. Not only will biodiversity informatics projects need to deal with a explosion in the amount of biodiversity-relevant data, they may well need to accommodate data that are of a conceptually different form. As any informatics professional knows, making changes to an underlying data model is always difficult and fraught with risk. Making changes to conceptual base classes is the hardest of all. Welcome to the world of 21st Century biodiversity.

GBIF and Biodiversity:

- What is “biodiversity” and how should biodiversity information be managed?
- Sequencing is getter better and faster at an incredible rate. What is the relevance to biodiversity studies?
- Science is a “light’s better” endeavor. When the light changes, the science changes.
- The light IS changing: Biological dark matter is becoming visible.
- Reality is not negotiable I: Examples from genetics & genomics.
- Reality is not negotiable II: The future of biodiversity.
- Welcome to the world of 21st Century biodiversity.

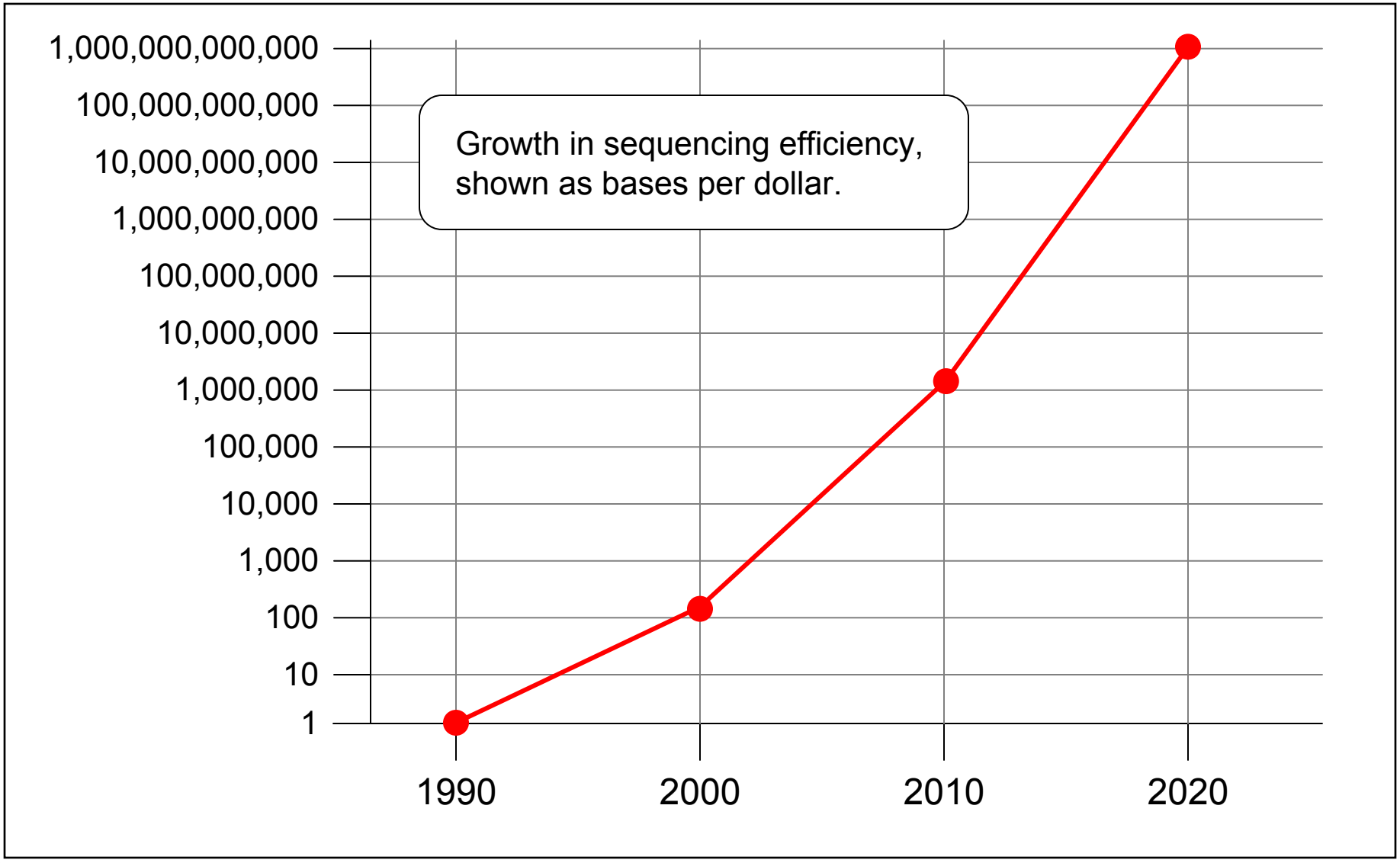
What is "biodiversity" and how should biodiversity information be managed?

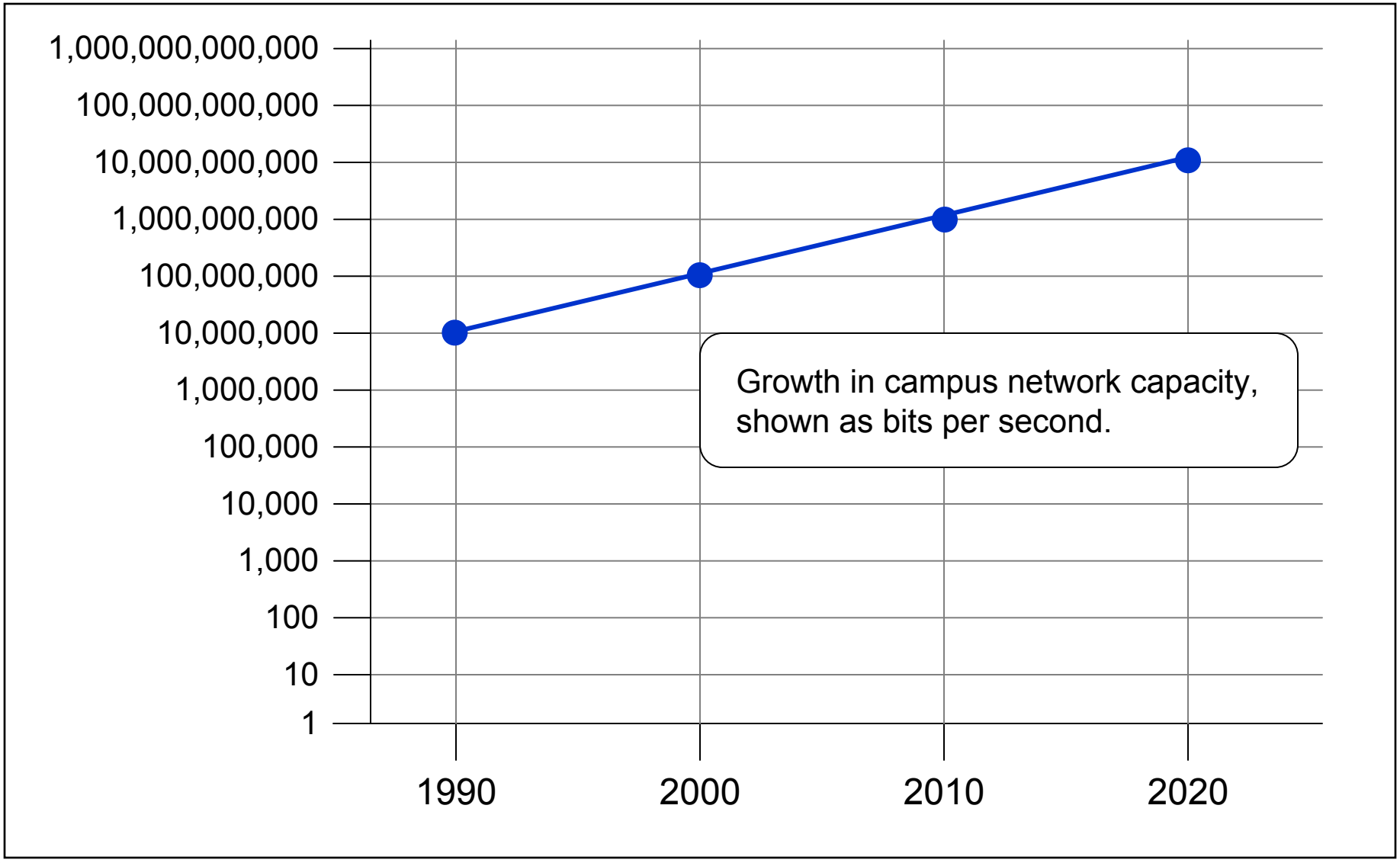
What is "biodiversity" and how should biodiversity information be managed?

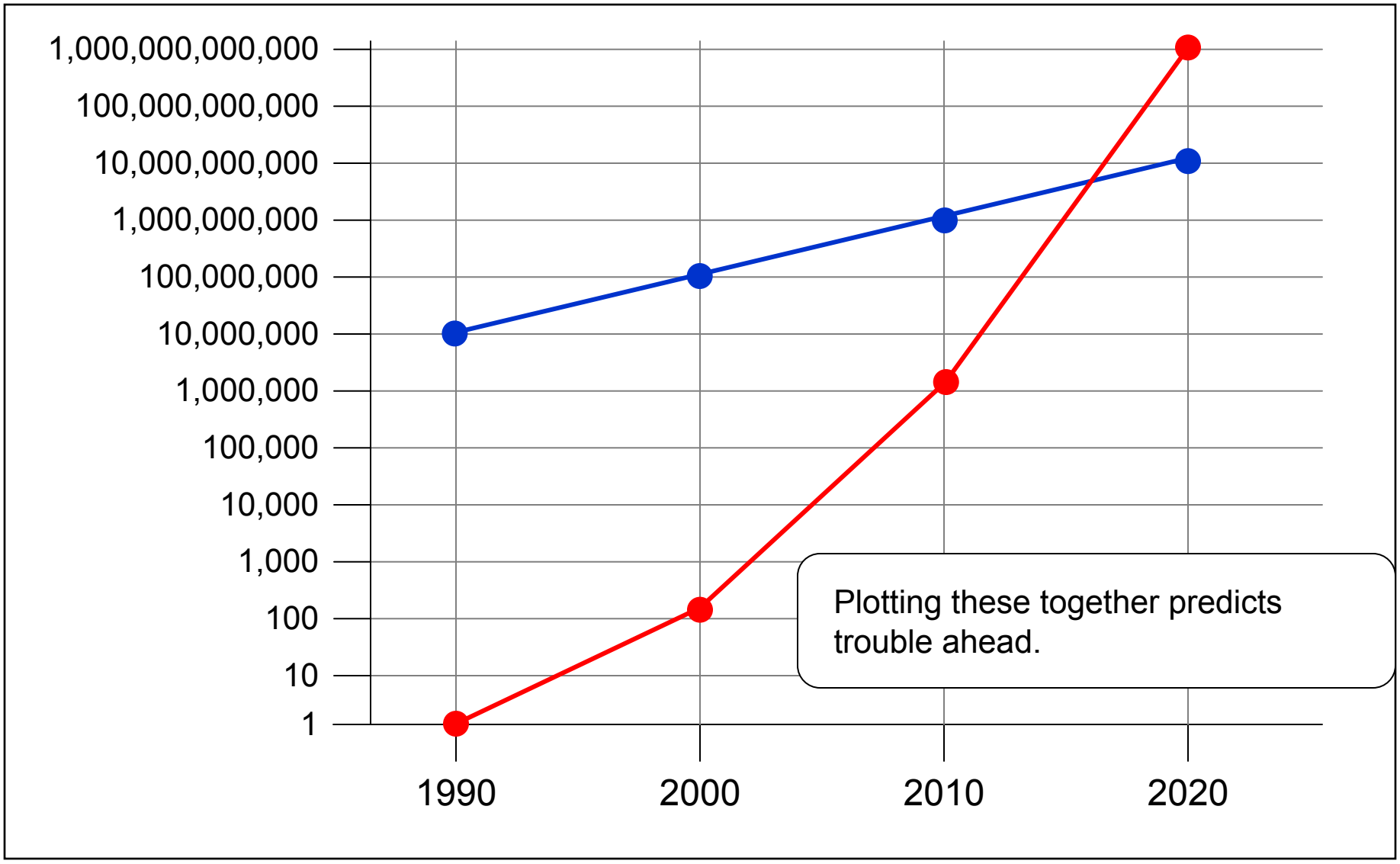
Addressing this question is, essentially, the purpose of this meeting.

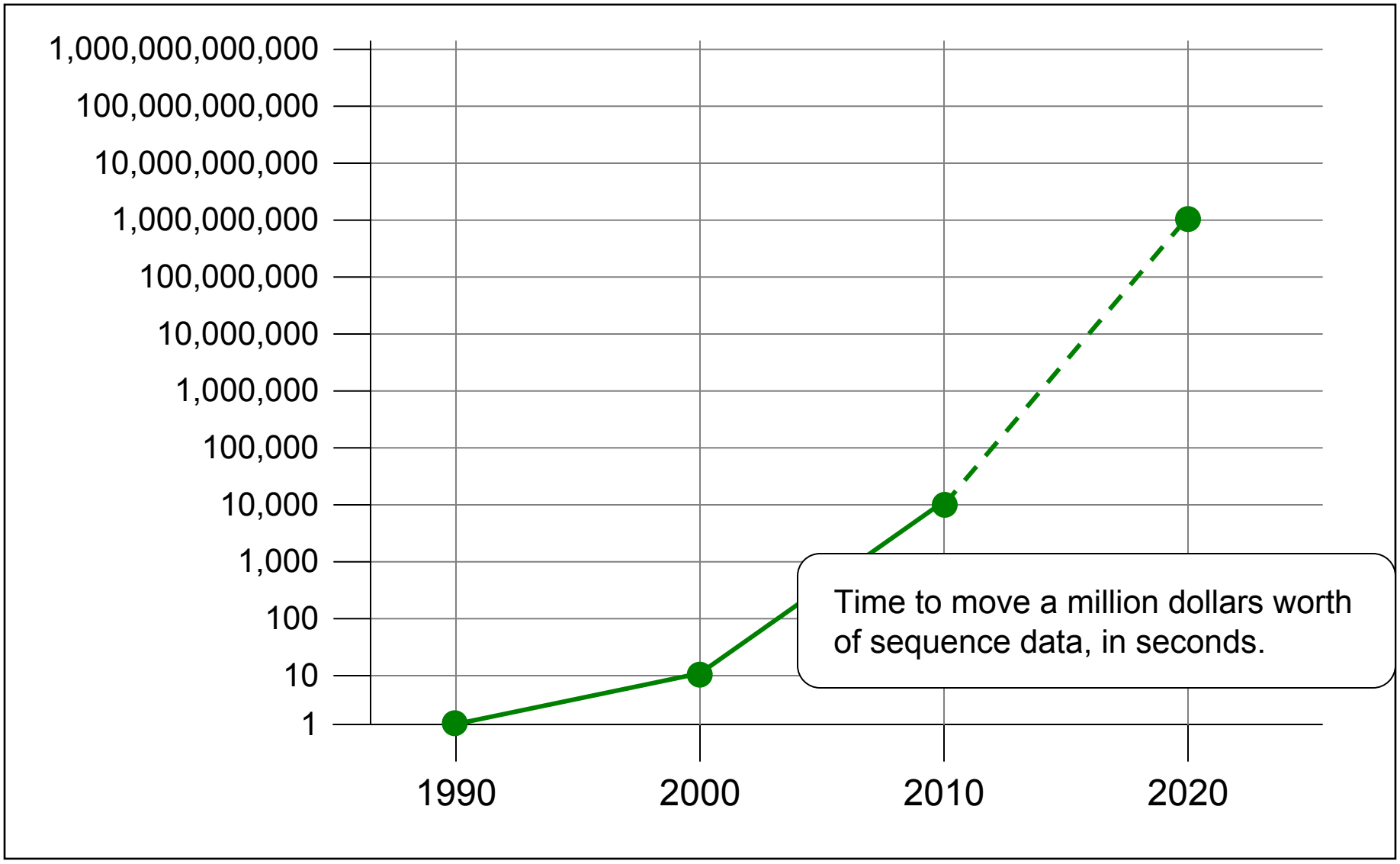
Sequencing Improvement

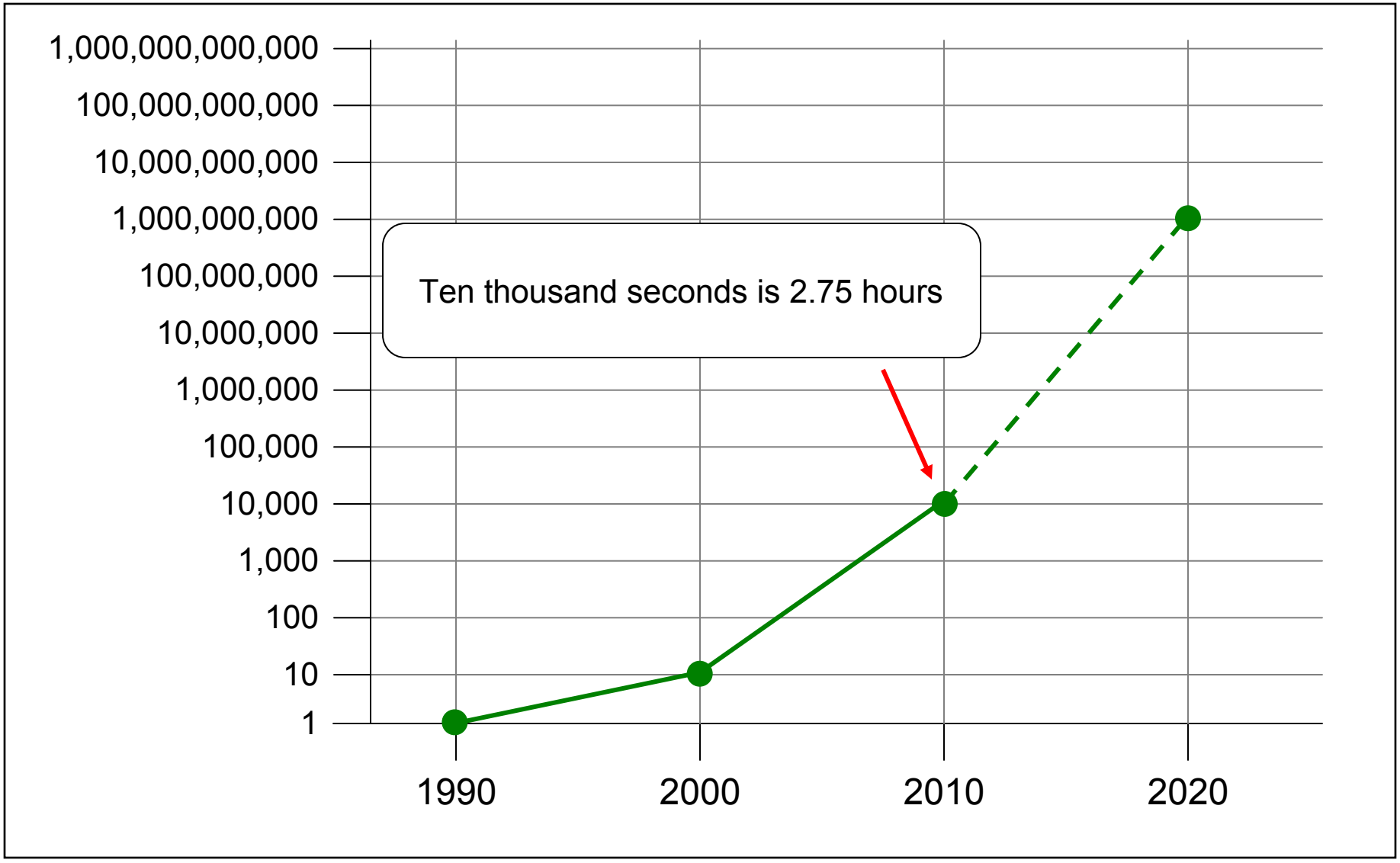
Is Astounding !!

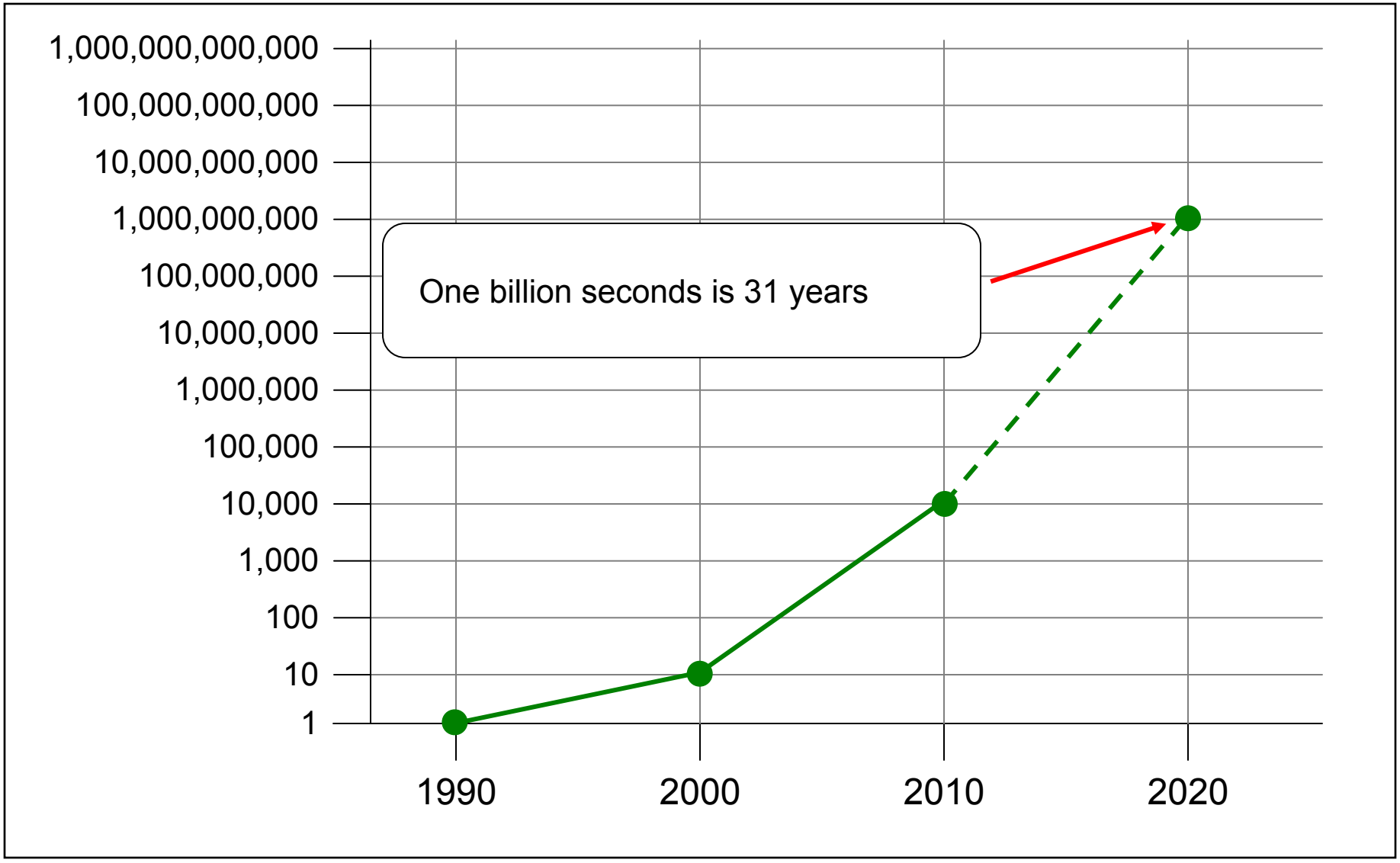

1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
Growth in sequencing efficiency, shown as bases per dollar.

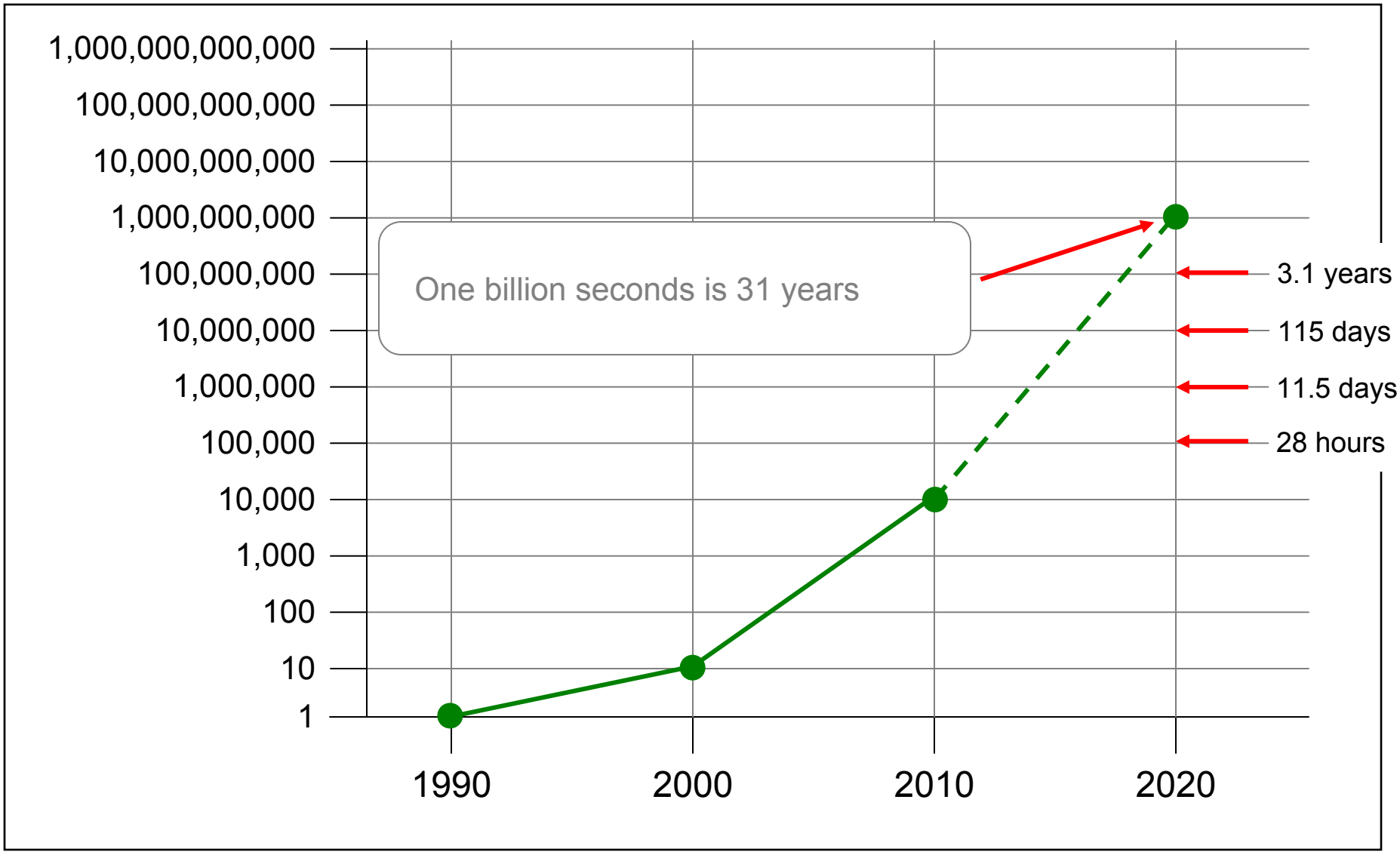

1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
Growth in campus network capacity, shown as bits per second.


1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
Plotting these together predicts trouble ahead.


1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
Time to move a million dollars worth of sequence data, in seconds.


1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
Ten thousand seconds is 2.75 hours


1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
One billion seconds is 31 years


1,000,000,000,000
100,000,000,000
10,000,000,000
1,000,000,000
100,000,000
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
1990
2000
2010
2020
One billion seconds is 31 years
3.1 years
115 days
11.5 days
28 hours

Bottom Line:

- There will be a lot more sequence data in the future than there is now.

Bottom Line:

- There will be a lot more sequence data in the future than there is now.
- Incorporating generic sequence data into biodiversity informatics will be technically and logistically challenging.

Bottom Line:

- There will be a lot more sequence data in the future than there is now.
- Incorporating generic sequence data into biodiversity informatics will be technically and logistically challenging.
- Incorporating biodiversity-specific sequence data into biodiversity informatics may also be conceptually and logically challenging.

Sequencing and Biodiversity

I. The Basics

Why (meta)genomics and biodiversity?

- Biodiversity is less a field of biology than a perspective (that of variance) into biology.
- Diversity is a sine qua non of biology; no diversity, no evolution.
- Genetics / genomics are equally central to biology – genetics is the study of the hereditary machinery, the basis of heritable variation, the raw material for evolution, the ultimate source of biodiversity.

Why (meta)genomics and biodiversity?

- Biodiversity is less a field of biology than a perspective

The connection between genomics/ metagenomics and biodiversity seems obvious and profound.

evolution, the ultimate source of biodiversity.

The Basics:

- As sequencing gets cheaper, its practical applicability to biodiversity will increase.

The Basics:

- As sequencing gets cheaper, its practical applicability to biodiversity will increase.
- Barcode-type data are useful in a diversity-diagnostic sense.

The Basics:

- As sequencing gets cheaper, its practical applicability to biodiversity will increase.
- Barcode-type data are useful in a diversity-diagnostic sense.
- Metagenomic tools allow a broad diversity assessment in a single test.

The Basics:

- As sequencing gets cheaper, its practical applicability to biodiversity will increase.
- Barcode-type data are useful in a diversity-diagnostic sense.
- Metagenomic tools allow a broad diversity assessment in a single test.
- Traditional biodiversity thinking will need to be extended to include sequence data; e.g., will we need a type sequence?

The Basics:

- Geo-referenced genetic data can provide evidence of patterns of origin and distribution of new genes.

The Basics:

- Geo-referenced genetic data can provide evidence of patterns of origin and distribution of new genes.
- Should species-abundance maps be extended to include geno-clines?

The Basics:

- Geo-referenced genetic data can provide evidence of patterns of origin and distribution of new genes.
- Should species-abundance maps be extended to include geno-clines?
- and so on…

Sequencing and Biodiversity

II. New Insights

New Insights:

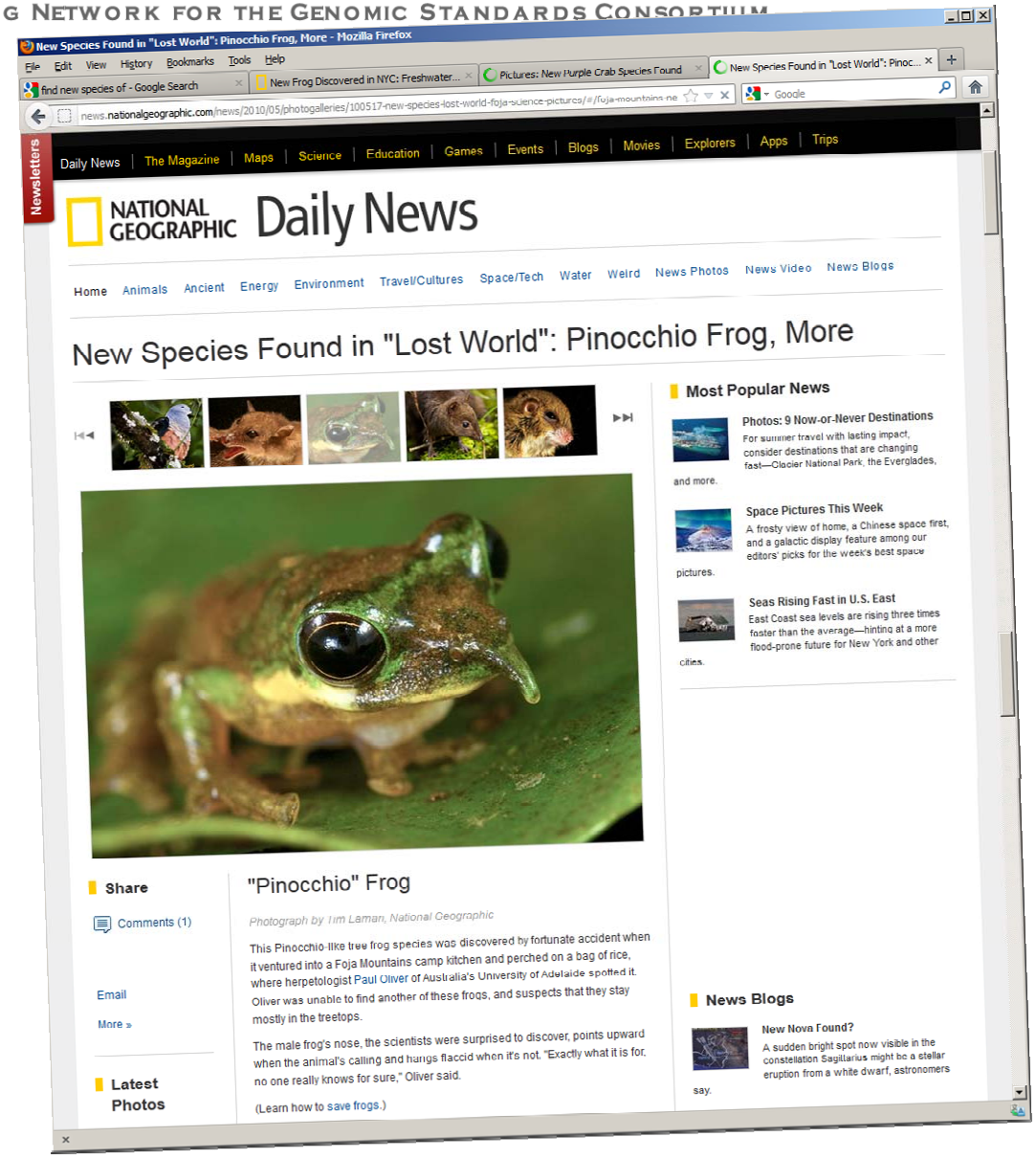
- Sequence analysis was responsible for the most important biodiversity discovery of the last hundred years.

New Insights:

- Sequence analysis was responsible for the most important biodiversity discovery of the last hundred years.
- Newly emerging sequence analysis tools will allow us to study vast swaths of biodiversity that have previously been invisible.

New Insights:

- Sequence analysis was responsible for the most important biodiversity discovery of the last hundred years.
- Newly emerging sequence analysis tools will allow us to study vast swaths of biodiversity that have previously been invisible.
- Findings from genomic and metagenomic studies of biodiversity may force some MAJOR reassessments of basic biological concepts.


New Species Found in "Lost World": Pinocchio Frog, More - Mozilla Firefox
Daily News | The Magazine | Maps | Science | Education | Games | Events | Blogs | Movies | Explorers | Apps | Trips
NATIONAL GEOGRAPHIC Daily News
Home Animals Ancient Energy Environment Travel/Cultures Space/Tech Water Weird News Photos News Video News Blogs
New Species Found in "Lost World": Pinocchio Frog, More
Most Popular News
Photos: 9 Now-or-Never Destinations
Space Pictures This Week
Seas Rising Fast in U.S. East
Share
Comments (1)
Email
More »
Latest Photos
"Pinocchio" Frog
Photograph by Tim Laman, National Geographic
This Pinocchio-like tree frog species was discovered by fortunate accident when it ventured into a Foja Mountains camp kitchen and perched on a bag of rice, where herpetologist Paul Oliver of Australia's University of Adelaide spotted it. Oliver was unable to find another of these frogs, and suspects that they stay mostly in the treetops.
The male frog's nose, the scientists were surprised to discover, points upward when the animal's calling and hangs flaccid when it's not. "Exactly what it is for, no one really knows for sure," Oliver said.
(Learn how to save frogs.)
News Blogs
New Nova Found?

Major Biodiversity Discovery !

Major Biodiversity Discovery !

Not really.

It's just another frog …

Major Biodiversity Discovery !

Proc. Natl. Acad. Sci. USA
Vol. 74, No. 11, pp. 5088–5090, November 1977
Evolution

Phylogenetic structure of the prokaryotic domain: The primary kingdoms

(archaebacteria/eubacteria/urkaryote/16S ribosomal RNA/molecular phylogeny)

CARL R. WOESE AND GEORGE E. FOX*

Department of Genetics and Development, University of Illinois, Urbana, Illinois 61801

Communicated by T. M. Sonneborn, August 18, 1977

ABSTRACT A phylogenetic analysis based upon ribosomal RNA sequence characterization reveals that living systems represent one of three aboriginal lines of descent: (*i*) the eubacteria, comprising all typical bacteria; (*ii*) the archaebacteria, containing methanogenic bacteria; and (*iii*) the urkaryotes, now represented in the cytoplasmic component of eukaryotic cells.

The biologist has customarily structured his world in terms of certain basic dichotomies. Classically, what was not plant was animal. The discovery that bacteria, which initially had been considered plants, resembled both plants and animals less than plants and animals resembled one another led to a reformulation of the issue in terms of a yet more basic dichotomy, that of eukaryote versus prokaryote. The striking differences between

to construct phylogenetic classifications between domains: Prokaryotic kingdoms are not comparable to eukaryotic ones. This should be recognized by an appropriate terminology. The highest phylogenetic unit in the prokaryotic domain we think should be called an "urkingdom"—or perhaps "primary kingdom." This would recognize the qualitative distinction between prokaryotic and eukaryotic kingdoms and emphasize that the former have primary evolutionary status.

The passage from one domain to a higher one then becomes a central problem. Initially one would like to know whether this is a frequent or a rare (unique) evolutionary event. It is traditionally assumed—without evidence—that the eukaryotic domain has arisen but once; all extant eukaryotes stem from a common [illegible] itself eukaryotic (2). A [illegible] holds

Major Biodiversity Discovery !

Proc. Natl. Acad. Sci. USA
Vol. 74, No. 11, pp. 5088–5090, November 1977
Evolution

Phylogenetic structure of the prokaryotic domain: The primary kingdoms

(archaebacteria/eubacteria/urkaryote/16S ribosomal RNA/molecular phylogeny)

CARL R. WOESE AND GEORGE E. FOX*

Department of Genetics and Development, University of Illinois, Urbana, Illinois 61801

Communicated by T. M. Sonneborn, August 18, 1977

ABSTRACT A phylogenetic analysis based upon ribosomal RNA sequence characterizatio... represent one of three aborigi... bacteria, comprising all typical ... containing methanogenic bacte... represented in the cytoplasm... cells.

Now this **was** a very big deal…

The biologist has customarily structured his world in terms of certain basic dichotomies. Classically, what was not plant was animal. The discovery that bacteria, which initially had been considered plants, resembled both plants and animals less than plants and animals resembled one another led to a reformulation of the issue in terms of a yet more basic dichotomy, that of eukaryote versus prokaryote. The striking differences between

to construct phylogenetic classifications between domains: Prokaryotic kingdoms are not comparable to eukaryotic ones. ...logy. The ... we think ... "primary ...distinction ...emphasize ... that the former have primary evolutionary status.

The passage from one domain to a higher one then becomes a central problem. Initially one would like to know whether this is a frequent or a rare (unique) evolutionary event. It is traditionally assumed—without evidence—that the eukaryotic domain has arisen but once; all extant eukaryotes stem from a common ... itself eukaryotic (2). A ... holds

LOGIC:

One could measure the similarity between two text documents by chopping them up into, say, ten-word phrases, and then asking what percentage of ten-word phrases were present in common between two documents.

If **every** phrase occurred in both documents, the score would be 1.0; if **no** phrases occurred in both documents, the score would be 0.0.

Table 1. Association coefficients (S_{AB}) between representative members of the three primary kingdoms

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. *Saccharomyces cerevisiae*, 18S	—	0.29	0.33	0.05	0.06	0.08	0.09	0.11	0.08	0.11	0.11	0.08	0.08
2. *Lemna minor*, 18S	0.29	—	0.36	0.10	0.05	0.06	0.10	0.09	0.11	0.10	0.10	0.13	0.07
3. L cell, 18S	0.33	0.36	—	0.06	0.06	0.07	0.07	0.09	0.06	0.10	0.10	0.09	0.07
4. *Escherichia coli*	0.05	0.10	0.06	—	0.24	0.25	0.28	0.26	0.21	0.11	0.12	0.07	0.12
5. *Chlorobium vibrioforme*	0.06	0.05	0.06	0.24	—	0.22	0.22	0.20	0.19	0.06	0.07	0.06	0.09
6. *Bacillus firmus*	0.08	0.06	0.07	0.25	0.22	—	0.34	0.26	0.20	0.11	0.13	0.06	0.12
7. *Corynebacterium diphtheriae*	0.09	0.10	0.07	0.28	0.22	0.34	—	0.23	0.21	0.12	0.12	0.09	0.10
8. *Aphanocapsa* 6714	0.11	0.09	0.09	0.26	0.20	0.26	0.23	—	0.31	0.11	0.11	0.10	0.10
9. Chloroplast (*Lemna*)	0.08	0.11	0.06	0.21	0.19	0.20	0.21	0.31	—	0.14	0.12	0.10	0.12
10. *Methanobacterium thermoautotrophicum*	0.11	0.10	0.10	0.11	0.06	0.11	0.12	0.11	0.14	—	0.51	0.25	0.30
11. *M. ruminantium* strain M-1	0.11	0.10	0.10	0.12	0.07	0.13	0.12	0.11	0.12	0.51	—	0.25	0.24
12. *Methanobacterium* sp., Cariaco isolate JR-1	0.08	0.13	0.09	0.07	0.06	0.06	0.09	0.10	0.10	0.25	0.25	—	0.32
13. *Methanosarcina barkeri*	0.08	0.07	0.07	0.12	0.09	0.12	0.10	0.10	0.12	0.30	0.24	0.32	—

The 16S (18S) ribosomal RNA from the organisms (organelles) listed were digested with T1 RNase and the resulting digests were subjected to two-dimensional electrophoretic separation to produce an oligonucleotide fingerprint. The individual oligonucleotides on each fingerprint were then sequenced by established procedures (13, 14) to produce an oligonucleotide catalog characteristic of the given organism (3, 4, 13–17, 22, 23; unpublished data). Comparisons of all possible pairs of such catalogs defines a set of association coefficients (S_{AB}) given by: $S_{AB} = 2N_{AB}/(N_A + N_B)$, in which N_A, N_B, and N_{AB} are the total numbers of nucleotides in sequences of hexamers or larger in the catalog for organism A, in that for organism B, and in the interreaction of the two catalogs, respectively (13, 23).

RESULTS

Range: 0.29-0.36

Range: 0.05-0.13

Table 1. Association coefficients (S_{AB}) between representative members of the three primary kingdoms

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. *Saccharomyces cerevisiae*, 18S	—	0.29	0.33	0.05	0.06	0.08	0.09	0.11	0.08	0.11	0.11	0.08	0.08
2. *Lemna minor*, 18S	0.29	—	0.36	0.10	0.05	0.06	0.10	0.09	0.11	0.10	0.10	0.13	0.07
3. L cell, 18S	0.33	0.36	—	0.06	0.06	0.07	0.07	0.09	0.06	0.10	0.10	0.09	0.07
4. *Escherichia coli*	0.05	0.10	0.06	—	0.24	0.25	0.28	0.26	0.21	0.11	0.12	0.07	0.12
5. *Chlorobium vibrioforme*	0.06	0.05	0.06	0.24	—	0.22	0.22	0.20	0.19	0.06	0.07	0.06	0.09
6. *Bacillus firmus*	0.08	0.06	0.07	0.25	0.22	—	0.34	0.26	0.20	0.11	0.13	0.06	0.12
7. *Corynebacterium diphtheriae*	0.09	0.10	0.07	0.28	0.22	0.34	—	0.23	0.21	0.12	0.12	0.09	0.10
8. *Aphanocapsa* 6714	0.11	0.09	0.09	0.26	0.20	0.26	0.23	—	0.31	0.11	0.11	0.10	0.10
9. Chloroplast (*Lemna*)	0.08	0.11	0.06	0.21	0.19	0.20	0.21	0.31	—	0.14	0.12	0.10	0.12
10. *Methanobacterium thermoautotrophicum*	0.11	0.10	0.10	0.11	0.06	0.11	0.12	0.11	0.14	—	0.51	0.25	0.30
11. *M. ruminantium* strain M-1	0.11	0.10	0.10	0.12	0.07	0.13	0.12	0.11	0.12	0.51	—	0.25	0.24
12. *Methanobacterium* sp., Cariaco isolate JR-1	0.08	0.13	0.09	0.07	0.06	0.06	0.09	0.10	0.10	0.25	0.25	—	0.32
13. *Methanosarcina barkeri*	0.08	0.07	0.07	0.12	0.09	0.12	0.10	0.10	0.12	0.30	0.24	0.32	—

The 16S (18S) ribosomal RNA from the organisms (organelles) listed were digested with T1 RNase and the resulting digests were subjected to two-dimensional electrophoretic separation to produce an oligonucleotide fingerprint. The individual oligonucleotides on each fingerprint were then sequenced by established procedures (13, 14) to produce an oligonucleotide catalog characteristic of the given organism (3, 4, 13–17, 22, 23; unpublished data). Comparisons of all possible pairs of such catalogs defines a set of association coefficients (S_{AB}) given by: $S_{AB} = 2N_{AB}/(N_A + N_B)$, in which N_A, N_B, and N_{AB} are the total numbers of nucleotides in sequences of hexamers or larger in the catalog for organism A, in that for organism B, and in the interreaction of the two catalogs, respectively (13, 23).

RESULTS

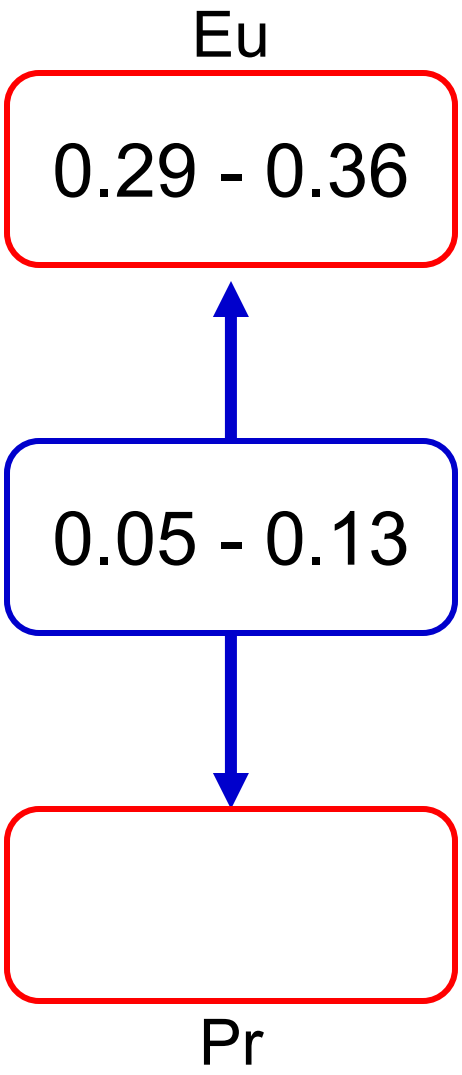
Eu

The eukaryotes were all similar to each other,

0.29 - 0.36

Range for all eu-eu comparisons

The eukaryotes were all similar to each other,



Range for all eu-eu comparisons

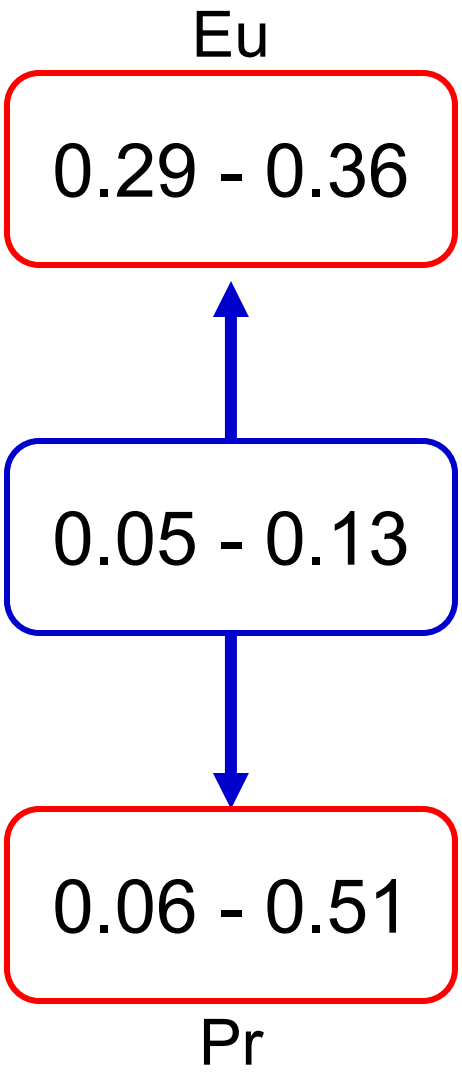
and they were different from the prokaryotes,

Range for all eu-pr comparisons

The eukaryotes were all similar to each other,

and they were different from the prokaryotes,

but the prokaryotes were wildly inconsistent with each other.



Range for all eu-eu comparisons

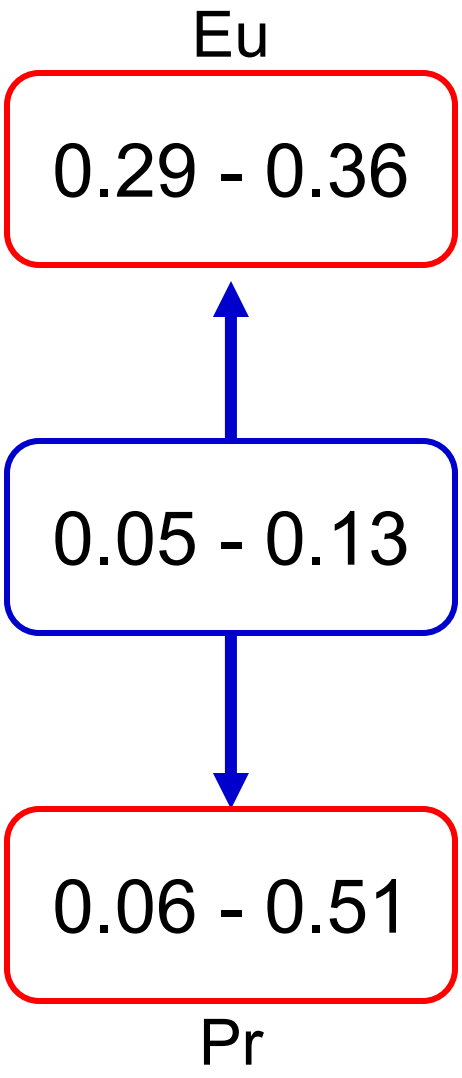
Range for all eu-pr comparisons

Range for all pr-pr comparisons

The eukaryotes were all similar to each other,

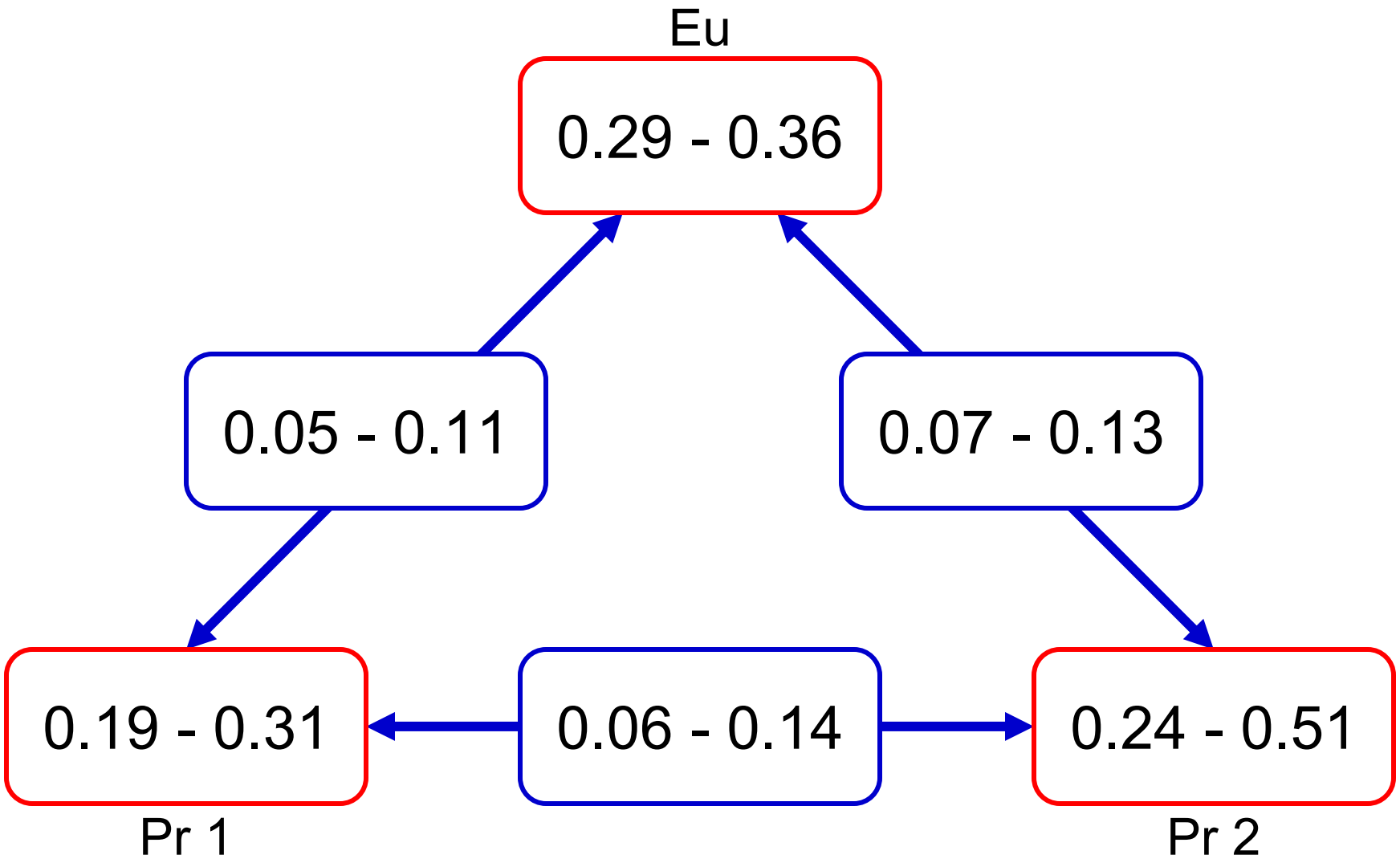
and they were different from the prokaryotes,

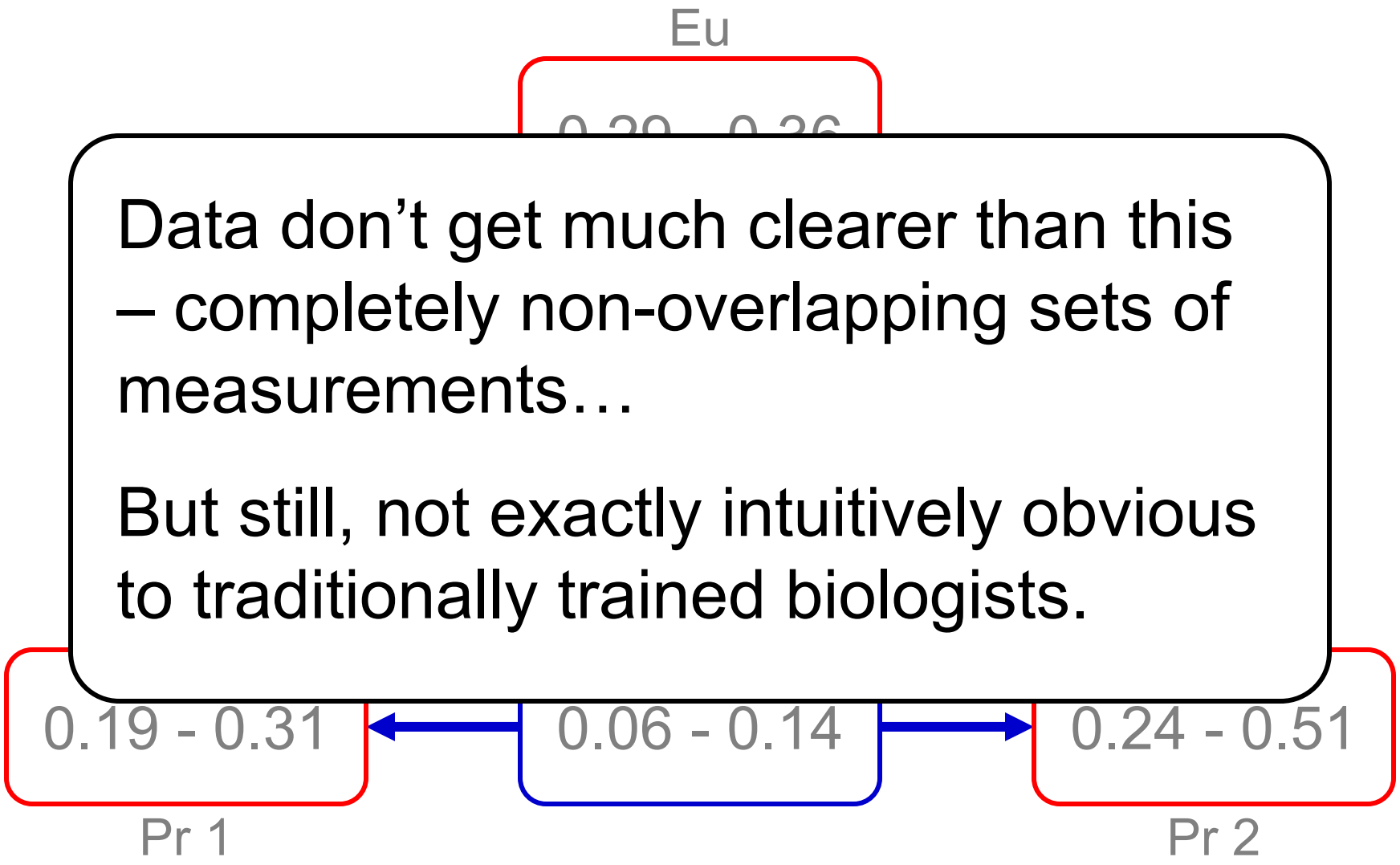
but the prokaryotes were wildly inconsistent with each other.



What to make of this?

Perhaps there was more than one kind of prokaryote…


Eu
0.29 - 0.36
0.05 - 0.11
0.07 - 0.13
0.19 - 0.31
0.06 - 0.14
0.24 - 0.51
Pr 1
Pr 2



Data don't get much clearer than this – completely non-overlapping sets of measurements…

But still, not exactly intuitively obvious to traditionally trained biologists.

Aside:

Most textbooks will tell you that, in 1610, Galileo Galilei became the first person to observe Saturn's rings.

But what did he really see?

This?

or…

The generation of important new insights while handicapped with limited technology, indirect measurement, and fuzzy data is the **mark** of scientific greatness.

What we can see is affected by (determined by?):

What we can see is affected by (determined by?):

Where we look.

What we can see is affected by (determined by?):

Where we look.

The “illumination” available to us.

actual lighting, instruments, analytical methods, other tools, …

What we can see is affected by (determined by?):

- Where we look.
- The "illumination" available to us.
  - actual lighting, instruments, analytical methods, other tools, …
- What we expect to see
  - our theories, past experiences, biases, prejudices, …

What we make of what we see is affected by:

Our ability to appreciate the details.

Our ability to see the big picture.

The context (vision) of our approach .

Our creativity.

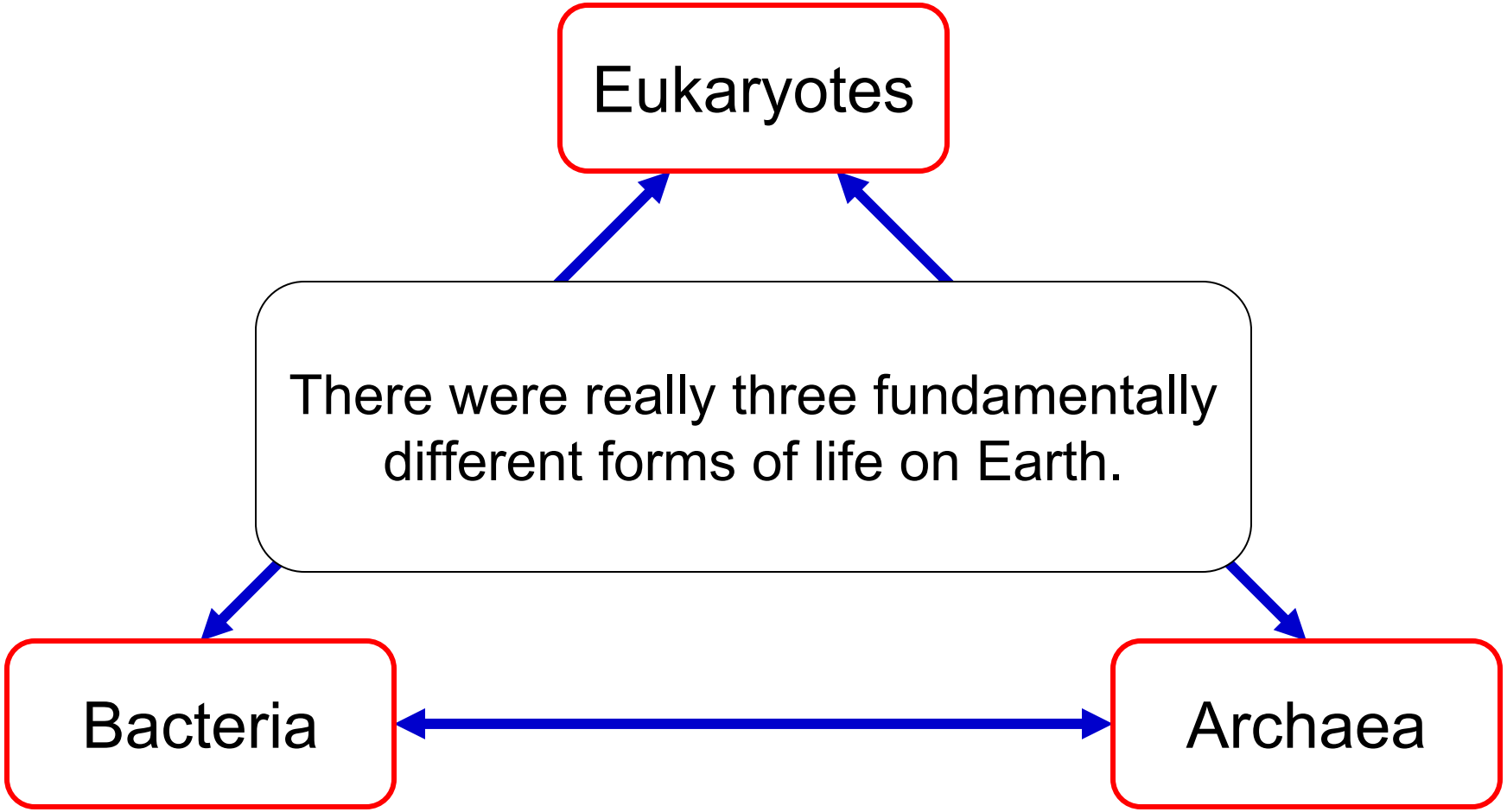
. . . .

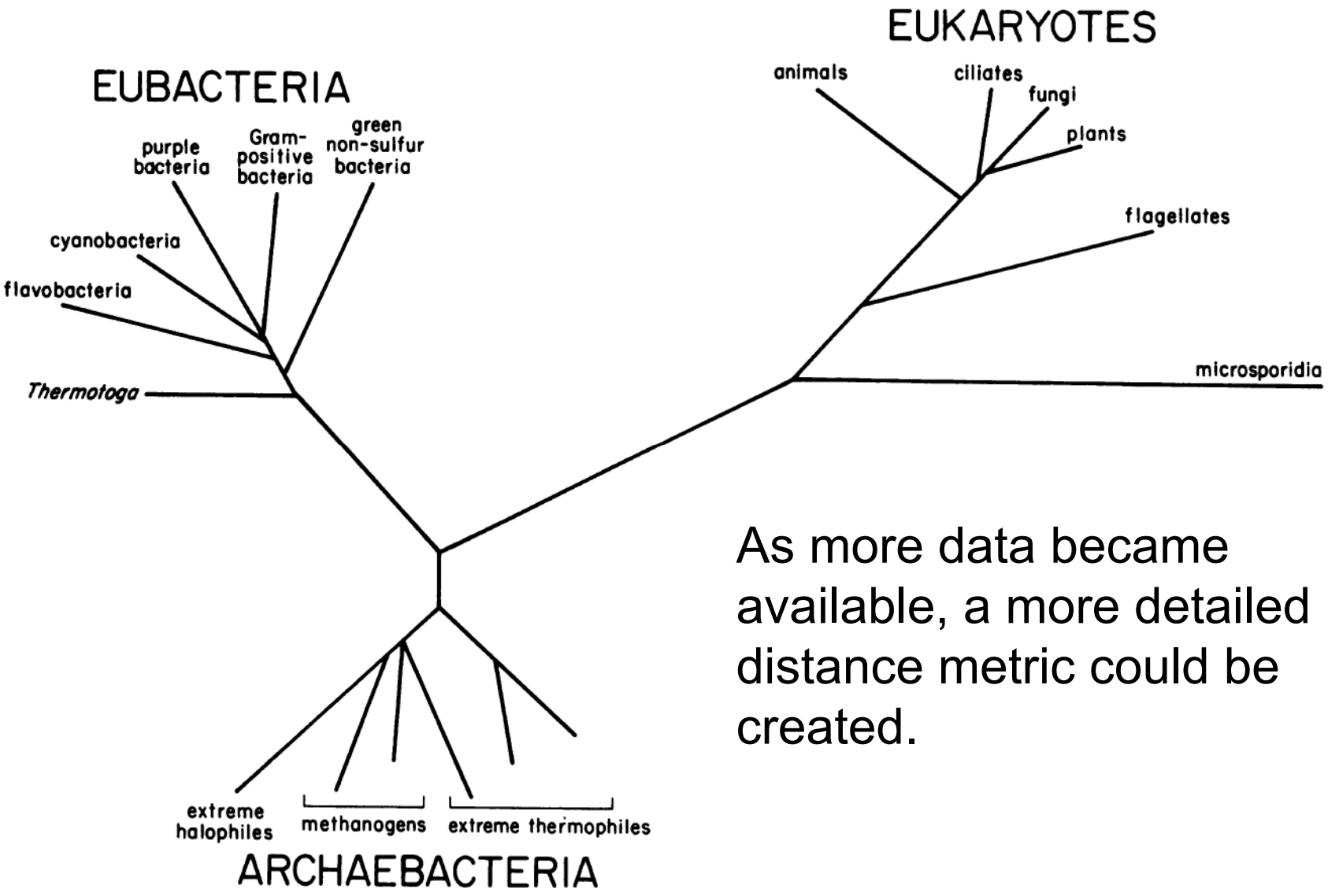
Table 1. Association coefficients (S_{AB}) between representative members of the three primary kingdoms

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. *Saccharomyces cerevisiae*, 18S	—	0.29	0.33	0.05	0.06	0.08	0.09	0.11	0.08	0.11	0.11	0.08	0.08
2. *Lemna minor*, 18S	0.29	—	0.36	0.10	0.05	0.06	0.10	0.09	0.11	0.10	0.10	0.13	0.07
3. L cell, 18S	0.33	0.36	—	0.06	0.06	0.07	0.07	0.09	0.06	0.10	0.10	0.09	0.07
4. *Escherichia coli*	0.05	0.10	0.06	—	0.24	0.25	0.28	0.26	0.21	0.11	0.12	0.07	0.12
5. *Chlorobium vibrioforme*	0.06	0.05	0.06	0.24	—	0.22	0.22	0.20	0.19	0.06	0.07	0.06	0.09
6. *Bacillus firmus*	0.08	0.06	0.07	0.25	0.22	—	0.34	0.26	0.20	0.11	0.13	0.06	0.12
7. *Corynebacterium diphtheriae*	0.09	0.10	0.07	0.28	0.22	0.34	—	0.23	0.21	0.12	0.12	0.09	0.10
8. *Aphanocapsa* 6714	0.11	0.09	0.09	0.26	0.20	0.26	0.23	—	0.31	0.11	0.11	0.10	0.10

With his early tools, the best Woese could see were tables of laboriously created association coefficients…,

but the implications were huge.

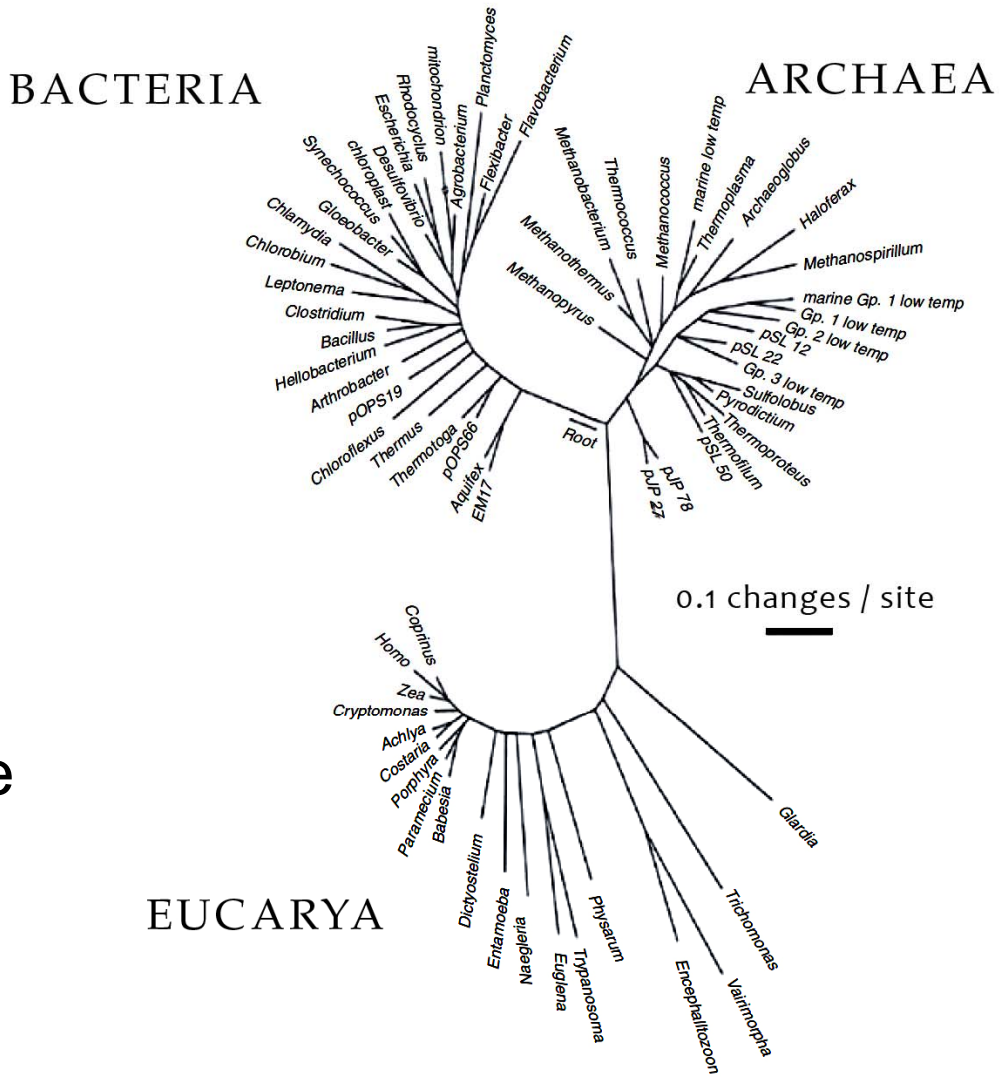

Eukaryotes
There were really three fundamentally different forms of life on Earth.
Bacteria
Archaea



As more data became available, a more detailed distance metric could be created.

Now it is much easier to generate full sequences, and from those full sequences to compute ever more detailed relationship trees.

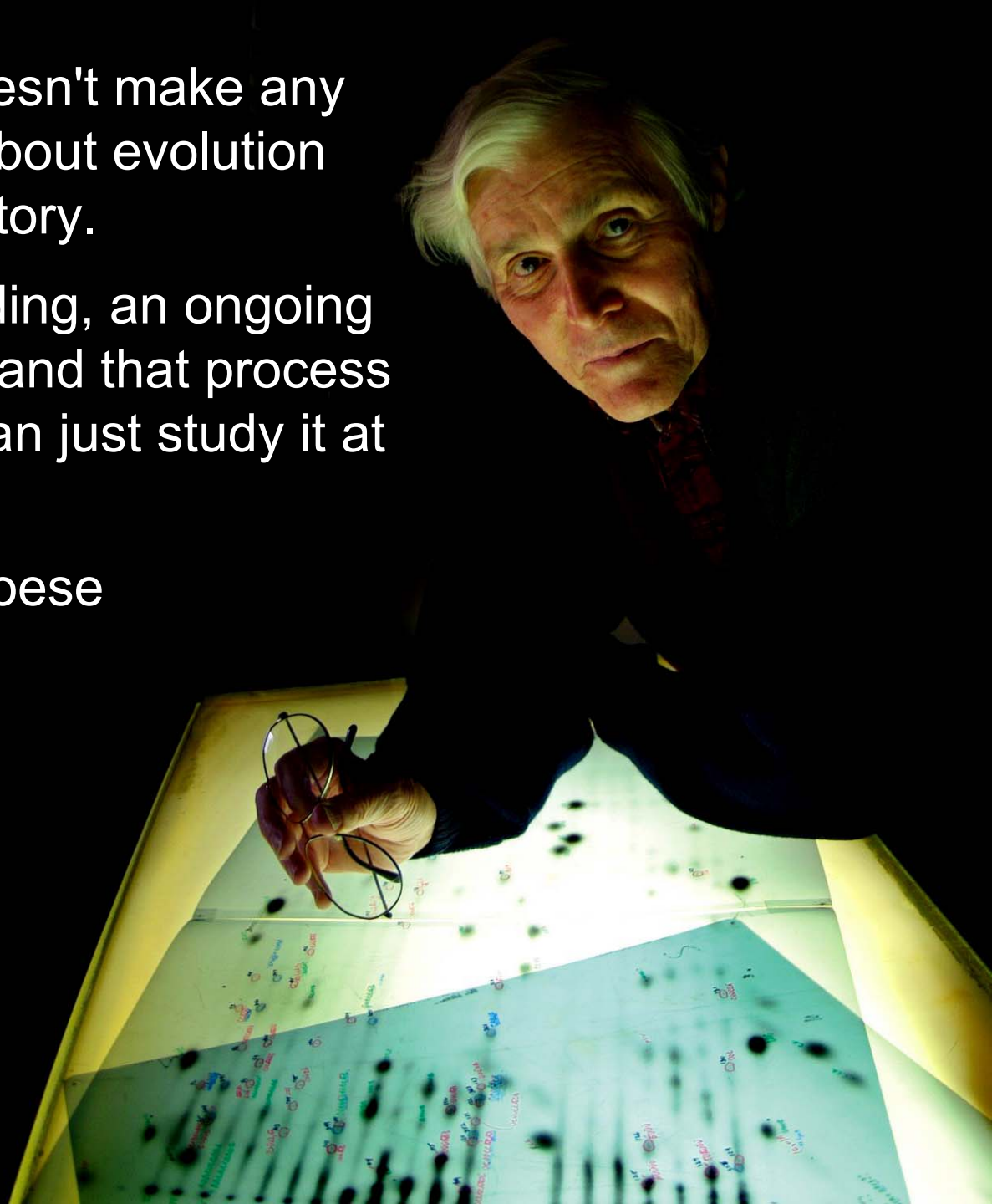
Although reality hasn't changed, our ability to see it, and to understand it, certainly has.



The study of life just doesn't make any sense unless you talk about evolution throughout its entire history.

Life is a historical unfolding, an ongoing process, and to understand that process you have to do more than just study it at any given point in time.

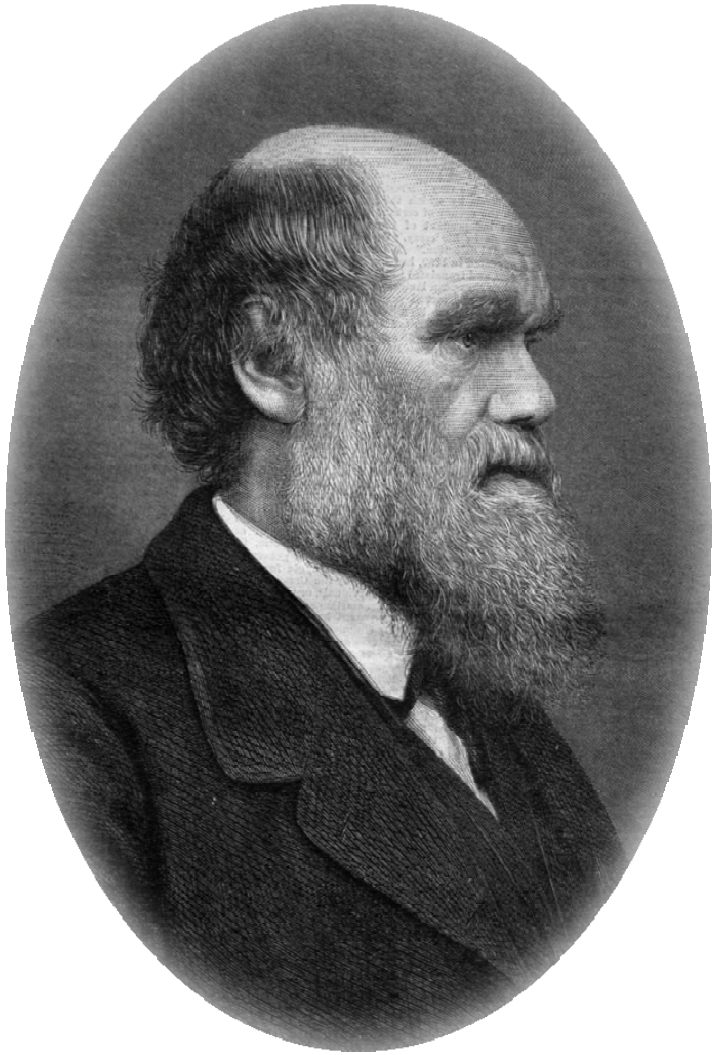
- Carl Woese



The study of life just doesn't make any sense unless you talk about evolution throughout its entire history.

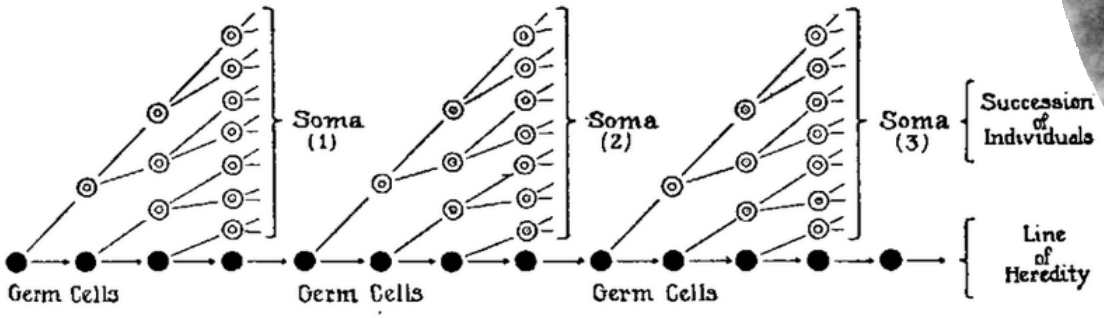
Life is a historical unfolding, an ongoing process, and to understand that process you have to do more than just study it at any given point in time.

If the goal of biodiversity studies is to understand all life in the biosphere in its full historical, biogeographical, and functional context, then Woese is right. We must strive to understand the complete evolutionary context in which life arose and diversified, not merely catalog recent changes.



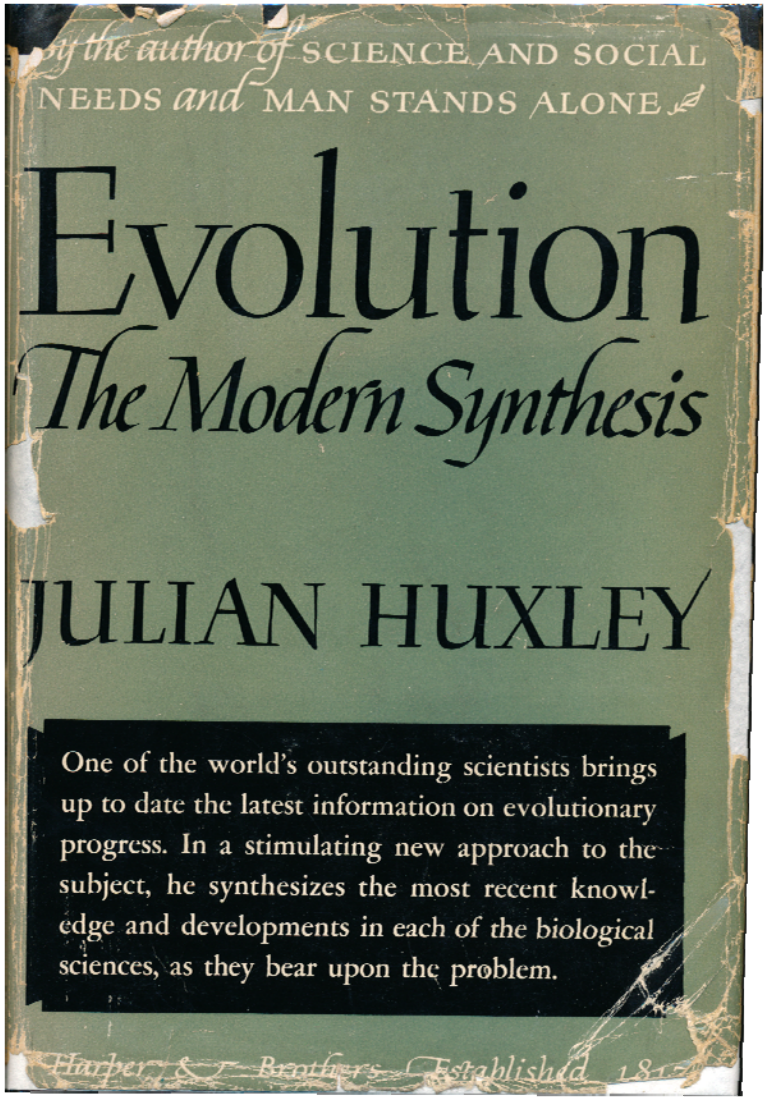
Darwin knew that his model required some hereditary mechanism that could supply the variation upon which selection could work, but which would also be resistant to dilution through "blending." He never developed a working model of his own, and some of his provisional ideas flirted with Lamarckism.

August Weismann's work on the germ-plasm theory assumed the hereditary stuff was in the cell nucleus and showed how this ruled out Lamarckian-style inheritance.



George Romanes thought Weismann's germ-plasm work ruling out the inheritance of acquired characteristics was an important extension to Darwin's own thinking, and so coined the phrase neo-Darwinism to describe this improved evolutionary model.

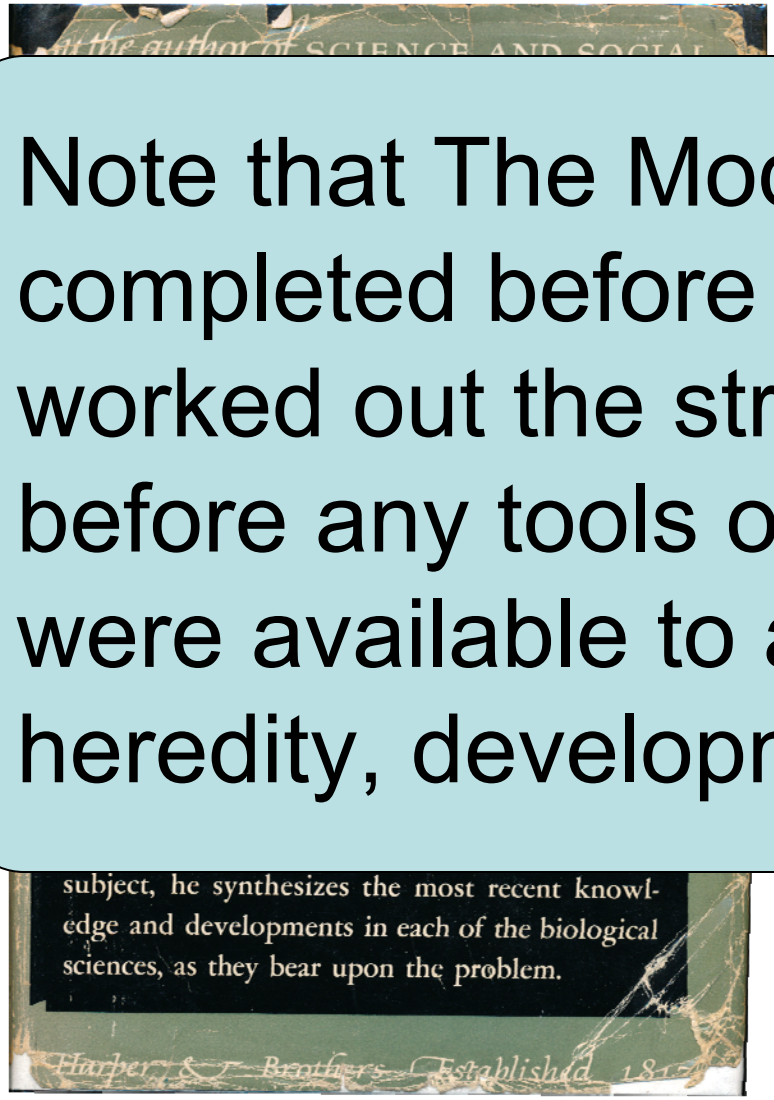
Mendel's work could have provided the hereditary model, but his work was unknown to Darwin (and unknown to most of science) until 1900, when the rediscovery of his work triggered an explosion of new research, establishing the field of classical genetics.



The combination of neo-Darwinism and Mendelism produced

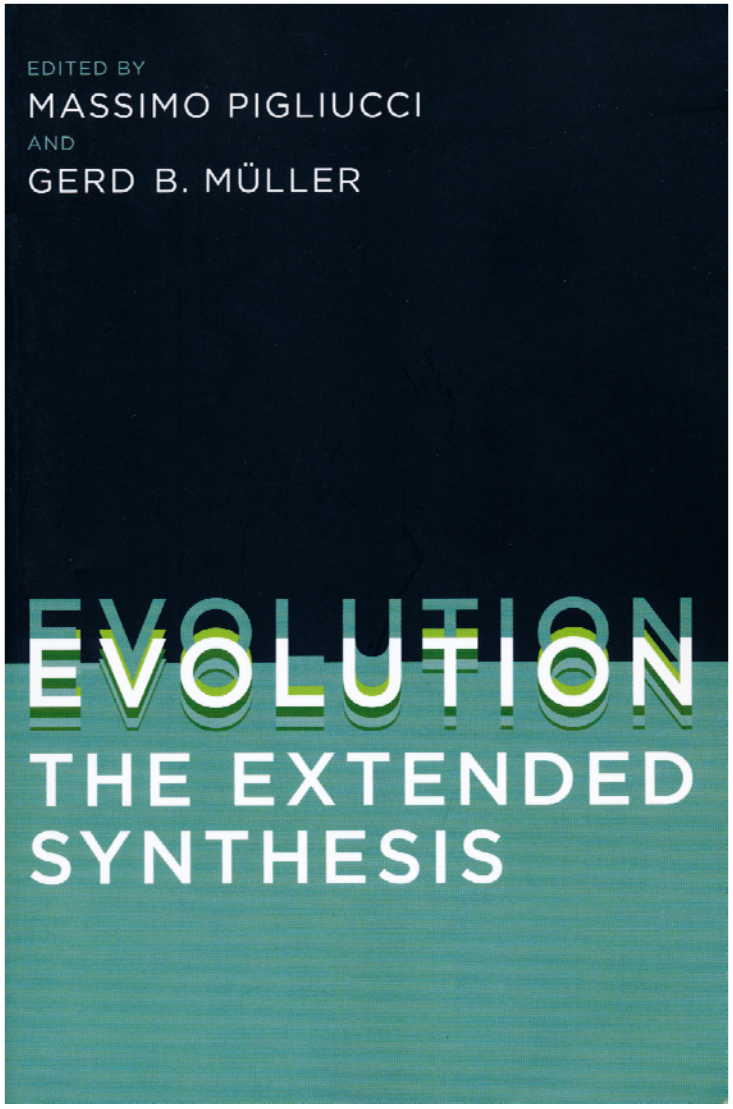
The Modern Synthesis

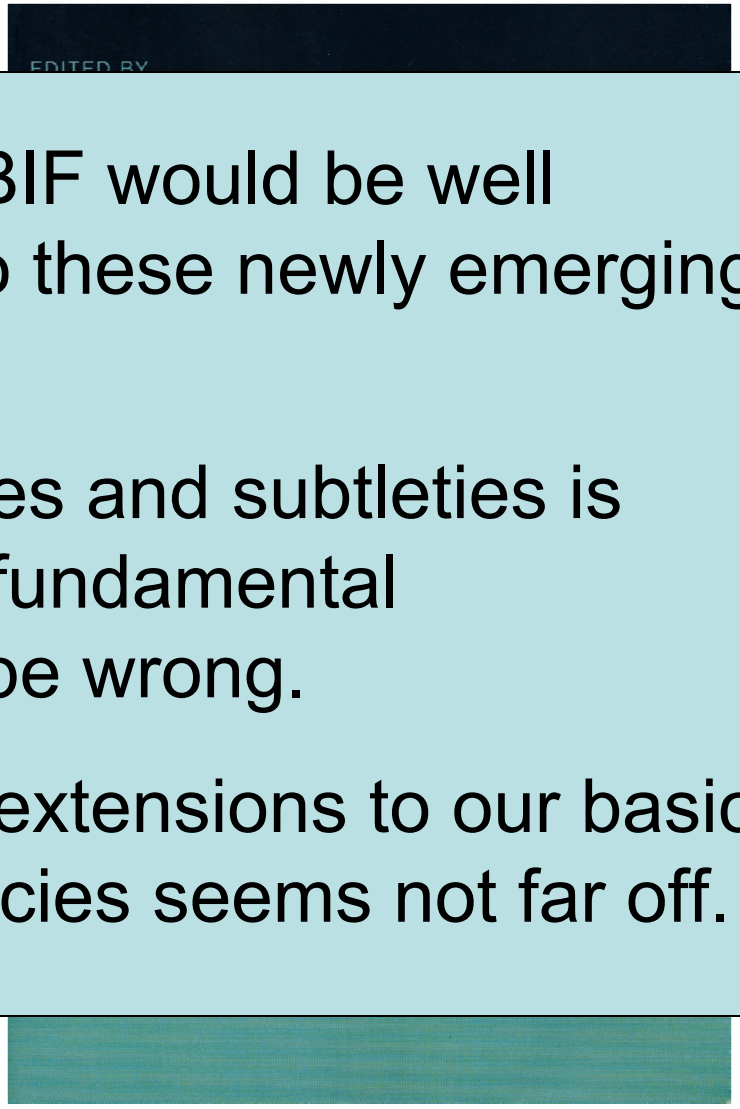
which has provided the intellectual foundation of most evolutionary thought from 1940 to the present.



Note that The Modern Synthesis was completed before Watson and Crick worked out the structure of DNA and before any tools of molecular biology were available to address problems of heredity, development, or evolution.

Today, some researchers are attempting to integrate newer findings from genomics and other fields to yield an improved and extended synthesis, suitable for 21st-century biology.

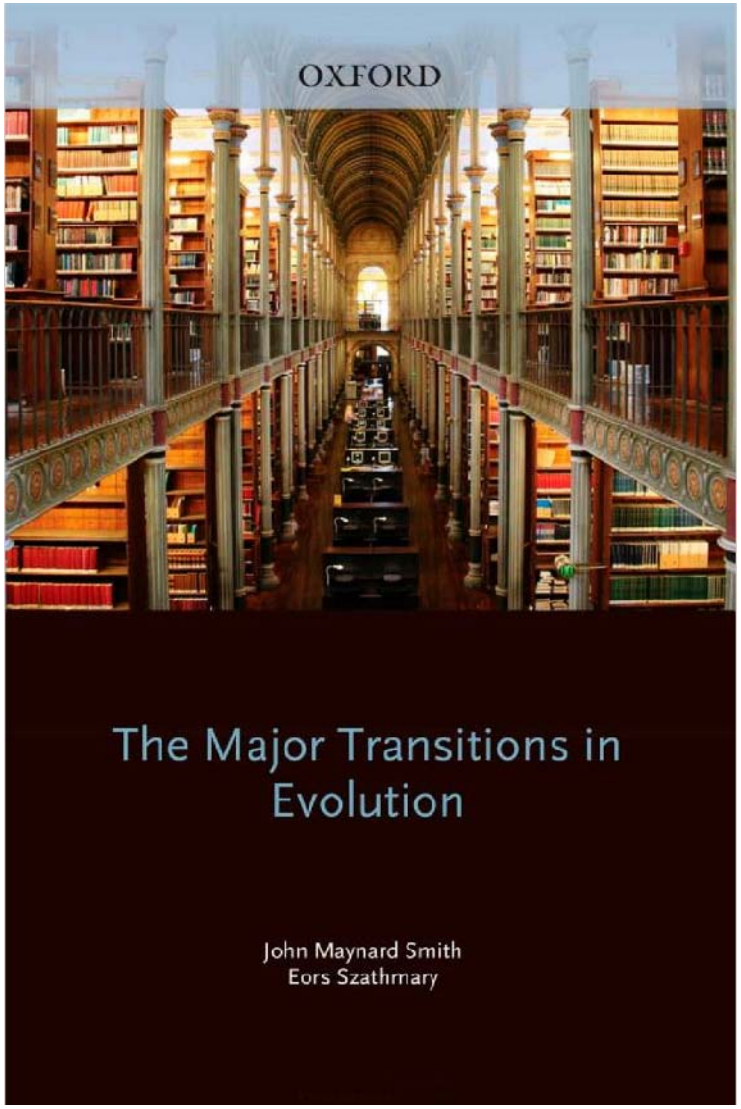




In planning for the future, GBIF would be well advised to attend carefully to these newly emerging evolutionary concepts.

Evidence for new complexities and subtleties is growing, while some earlier fundamental assumptions are proving to be wrong.

The possibility of significant extensions to our basic notions of organism and species seems not far off.



In the history of life on Earth, several major transitions have occurred.

These transitions were significant enough to change the nature of the evolutionary process itself, making it impossible to apply assumptions and analyses from one side of a transition to the other.

Replicating molecules → Populations of molecules

Independent replicators → Chromosomes

RNA → DNA

Prokaryotes → Eukaryotes

Asexual clones → Sexual populations

Unicellularity → Multicellularity

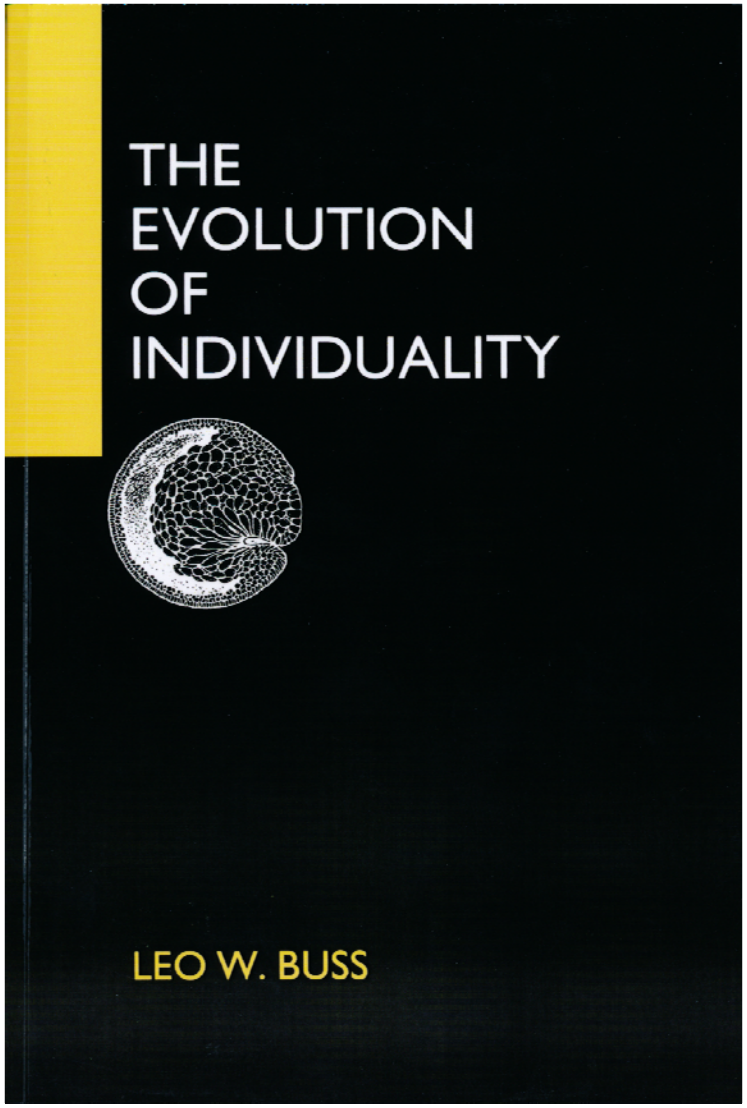
Solitary individuals → Colonies

Primate societies → Human societies (language)

For more than 80% of the time life has been evolving on Earth, multicellular "individuals" did not exist.

Even now, they occur in only a handful of top-level taxa.

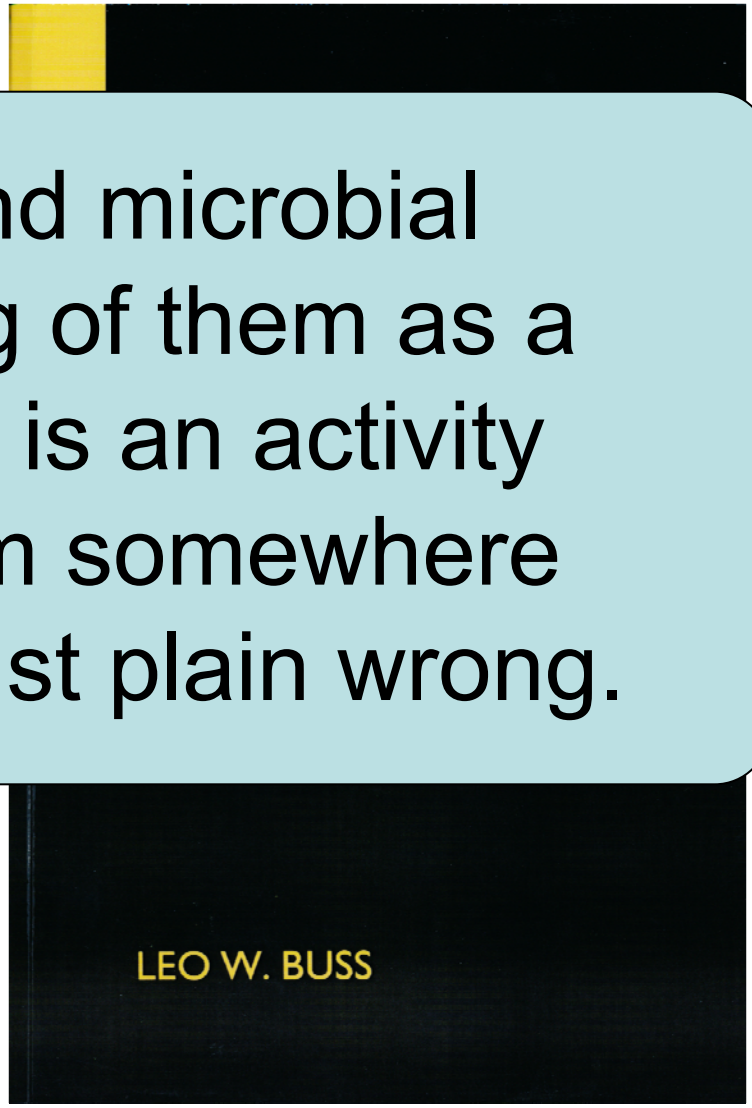
Thus, making the "individual" the centerpiece for under-standing evolution and for classifying life on Earth seems problematic.



For more than 80% of the

Attempting to understand microbial communities by thinking of them as a bunch of little bitty mice is an activity that falls on a continuum somewhere between fruitless and just plain wrong.

standing and classifying life on Earth seems problematic.



Science = Light's Better

Old Joke:

A drunk is crawling around a lamp post on his hands and knees.

A cop comes along …

Cop: What are you doing?

Drunk: Looking for my car keys.

Cop: Are you sure you dropped them here?

Drunk: No, I dropped them in the alley.

Cop: So why are you looking here?

Drunk: Because the light's better.

Old Joke:

Science is a light's better endeavor in that research effort is not directed at areas where the work is technically infeasible. Research is directed where real, interpretable results may be obtained.

We do, in fact, conduct research where the light's better.

But, when the light changes, so does science.

With better illumination, we look in new areas.

We find new things…

Drunk: Because the light's better.

Old Joke:

Science is a light's better endeavor in that research effort is not directed at areas where the work is technically infeasible. Research is directed where real, interpretable results may be obtained.

We do, in fact, conduct research where the light's better.

But, when the light changes, so does science.

With better illumination, we look in new areas.

We find new things…

Drunk: Bec

The light IS changing…

Before the Light Changed:

Biochemistry

If the genes are conceived as chemical substances, only one class of compounds need be given to which they can be reckoned as belonging, and that is the proteins in the wider sense, on account of the inexhaustible possibilities for variation which they offer. ... Such being the case, the most likely role for the nucleic acids seems to be that of the structure-determining supporting substance.

T. Caspersson. 1936. Über den chemischen Aufbau der Strukturen des Zellkernes. *Acta Med. Skand.*, 73, Suppl. 8, 1-151.

Before the Light Changed:

Classical Genetics

The genes are arranged [on chromosomes] in a manner similar to beads strung on a loose string.

Sturtevant, A.H., and Beadle, G.W., 1939, *An Introduction to Genetics*. W. B. Saunders Company, Philadelphia, p. 94.

Before the Light Changed:

Biodiversity

? ? ?

Before the Light Changed:

Biodiversity

Candidates:

The centrality of the individual organism – the specimen – to biodiversity thinking.

The centrality of the single-rooted tree of life as a device for representing our understanding of how life arose and diversified.

Biological Dark Matter Heaves into View

Biological Dark Matter Heaves into View

and what a view !

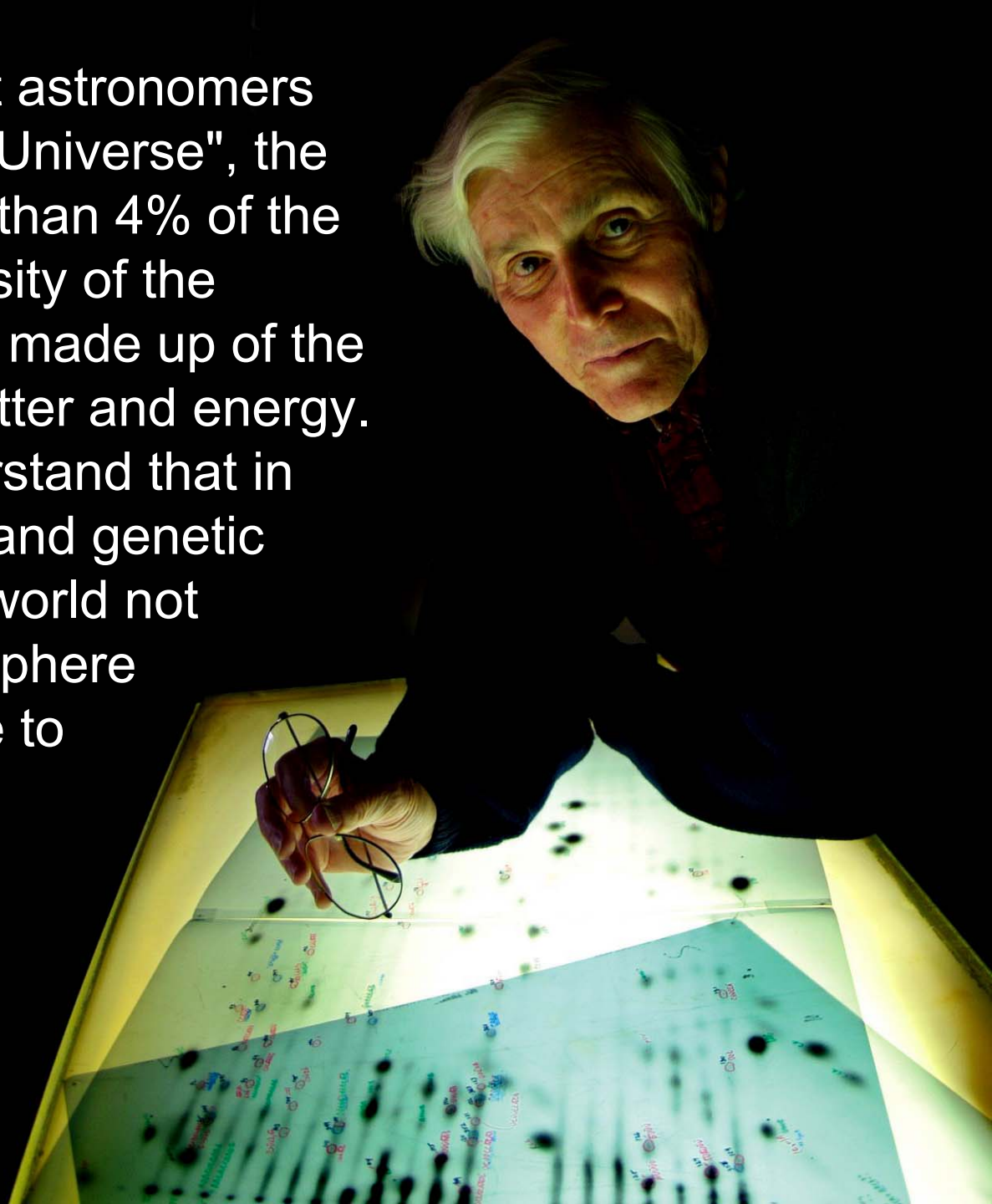
- Metagenomics tools are showing that, compared to macro-scale organisms, the diversity of microbial communities is staggering.

  Intra-species bacterial genetic diversity is greater than that among the great apes; intra-genus bacterial diversity is greater than that among all the mammalia.

- Metagenomics tools are showing that a full understanding of macro-scale organisms will depend on an understanding of their interactions with their associated microbiomes.

  Understanding how different ants optimize nutrient acquisition, and thus how they function in their niches, depends on an understanding of their associated gut microbiomes.

We now know that what astronomers used to think of as "the Universe", the visible universe, is less than 4% of the total matter/energy density of the universe, the rest being made up of the still mysterious dark matter and energy. Similarly, we now understand that in terms of both numbers and genetic diversity, the microbial world not only dominates the biosphere but is almost impossible to sample properly.



We now know that what astronomers used to think of as "the Universe", the visible universe, is less than 4% of the total matter/energy density of the universe, the rest being made up of the still mysterious dark matter and energy. Similarly, we now understand that in terms of both numbers and genetic

Understanding this biological dark matter must be a top goal for 21st century biology. In the 20th century we found that classical physics was only an approximation of reality – an incredibly useful approximation, but an approximation nonetheless. So, too, with classical biology.

Reality is NOT Negotiable, I

Genomics Example

Evolving Definition of a Gene:

Classical Definition: fundamental unit of heredity, mutation, and recombination (beads on a string).

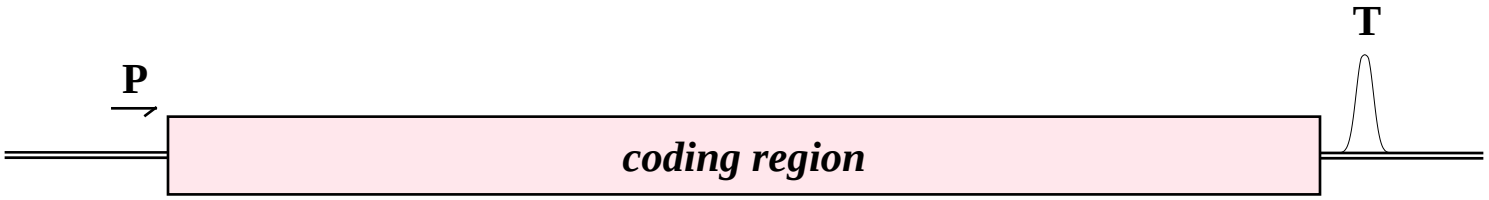
Physiological Definition: fundamental unit of function (one gene, one enzyme).

Cistronic Definition: fundamental unit of expression (cis-trans test).

Sequence Definition: the smallest segment of the gene-string consistently associated with the occurrence of a specific genetic effect.

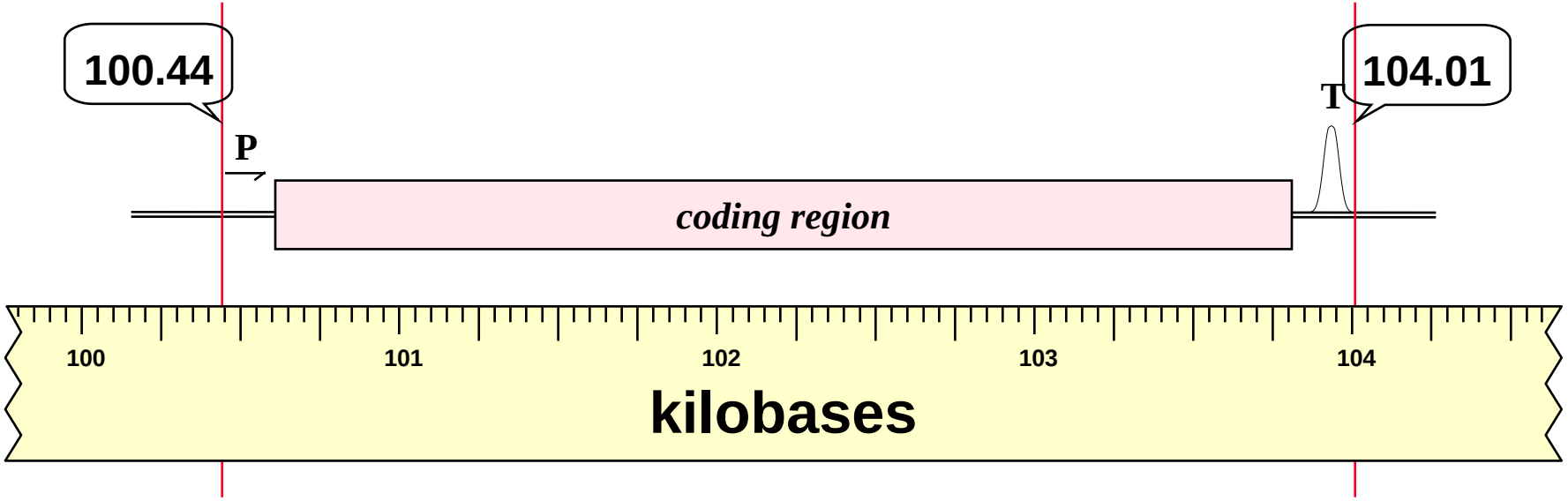
Current Definition: ???

Gene as Sequence (simplistic view)



A gene is a transcribed region of DNA, flanked by upstream start regulatory sequences and downstream stop regulatory sequences.

Gene as Sequence (simplistic view)

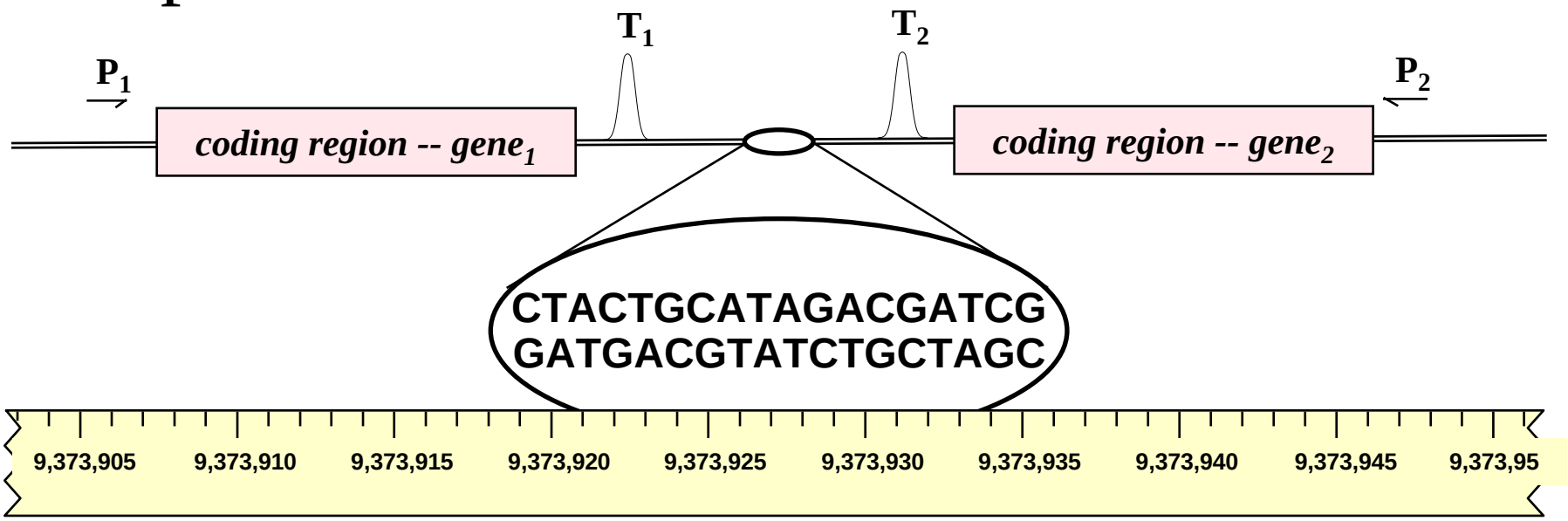


The location of a gene can be designated by specifying the base-pair location of its beginning and end.

Simplistic View of Genome

DNA may be transcribed in either direction. Therefore, fully specifying a gene's position requires noting its orientation as well as its start and stop positions.

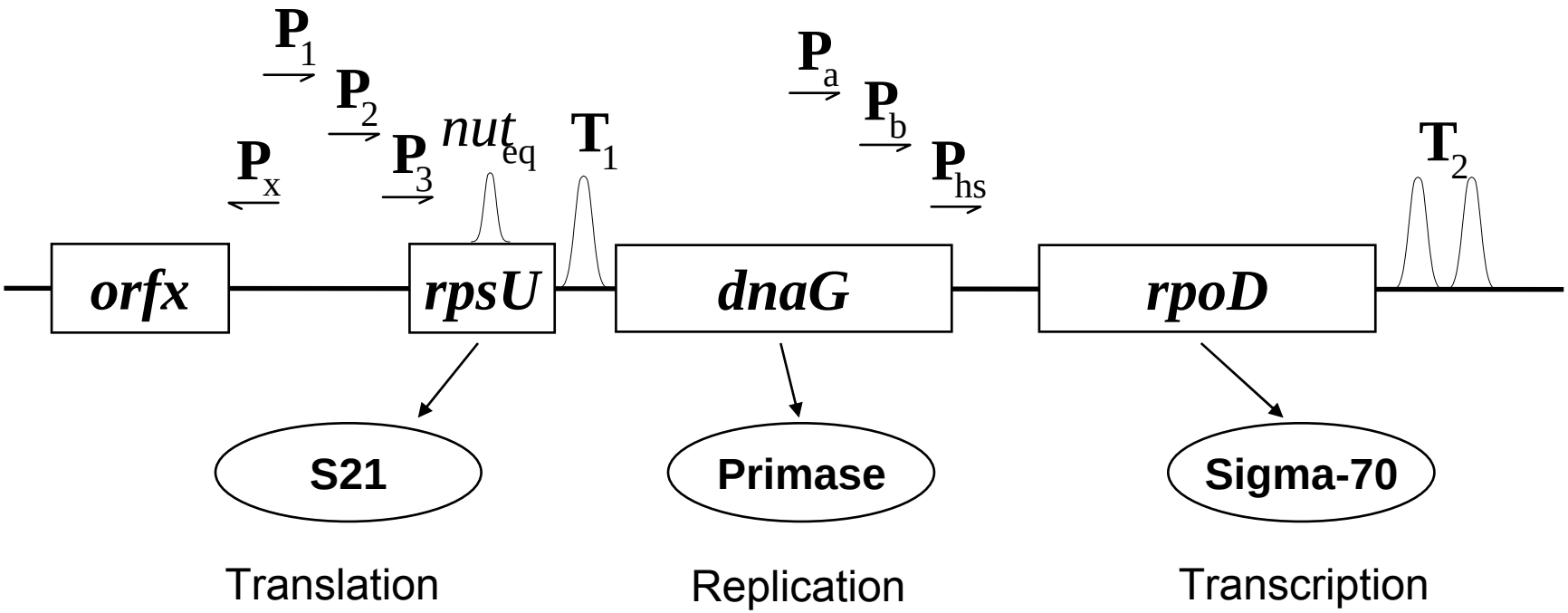
Simplistic View of Genome



A naive view holds that a genome can be represented as a continuous linear string of nucleotides, with landmarks identified by the chromosome number followed by the offset number of the nucleotide at the beginning and end of the region of interest. This simplistic approach ignores the fact that human chromosomes may vary in length by tens of millions of nucleotides.

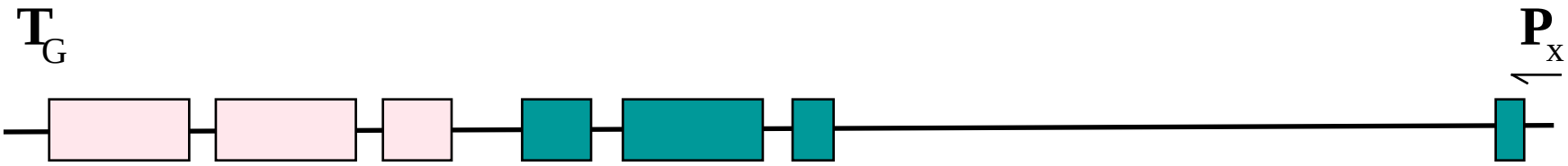
Complicated Sequences

Escherichia coli: the MMS Operon

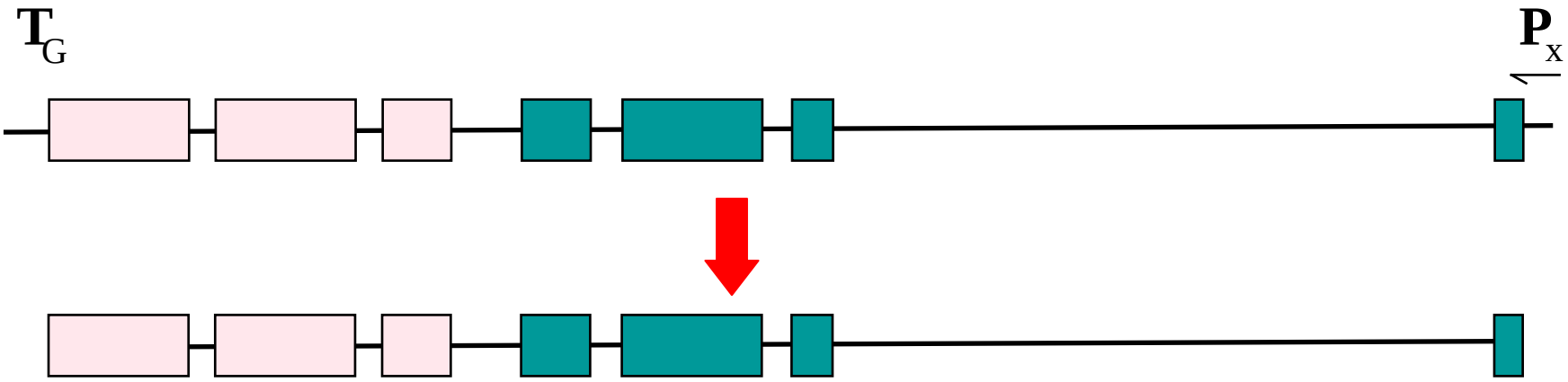


Lupski, J.R., Godson, G.N., 1989, DNA→DNA, and DNA→ RNA→Protein: Orchestration by a single complex operon, *BioEssays*, 10:152-157.

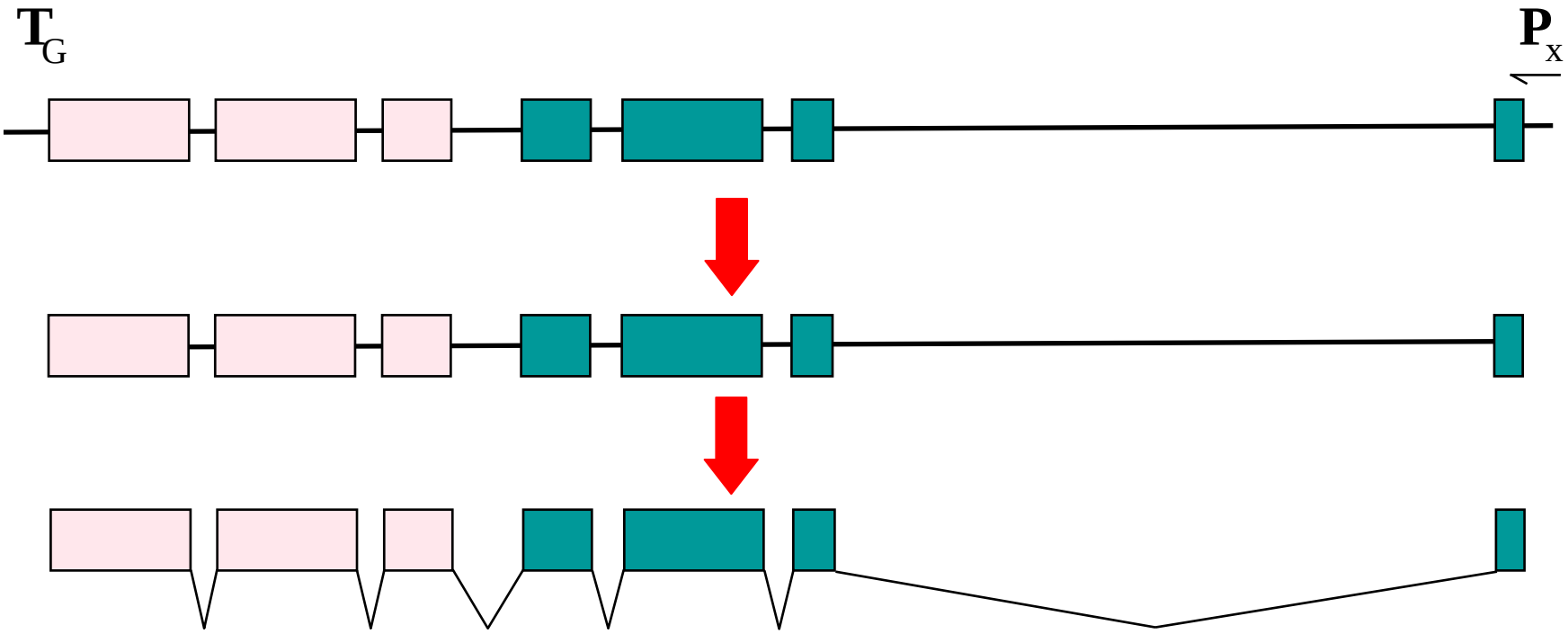
Introns: Gart



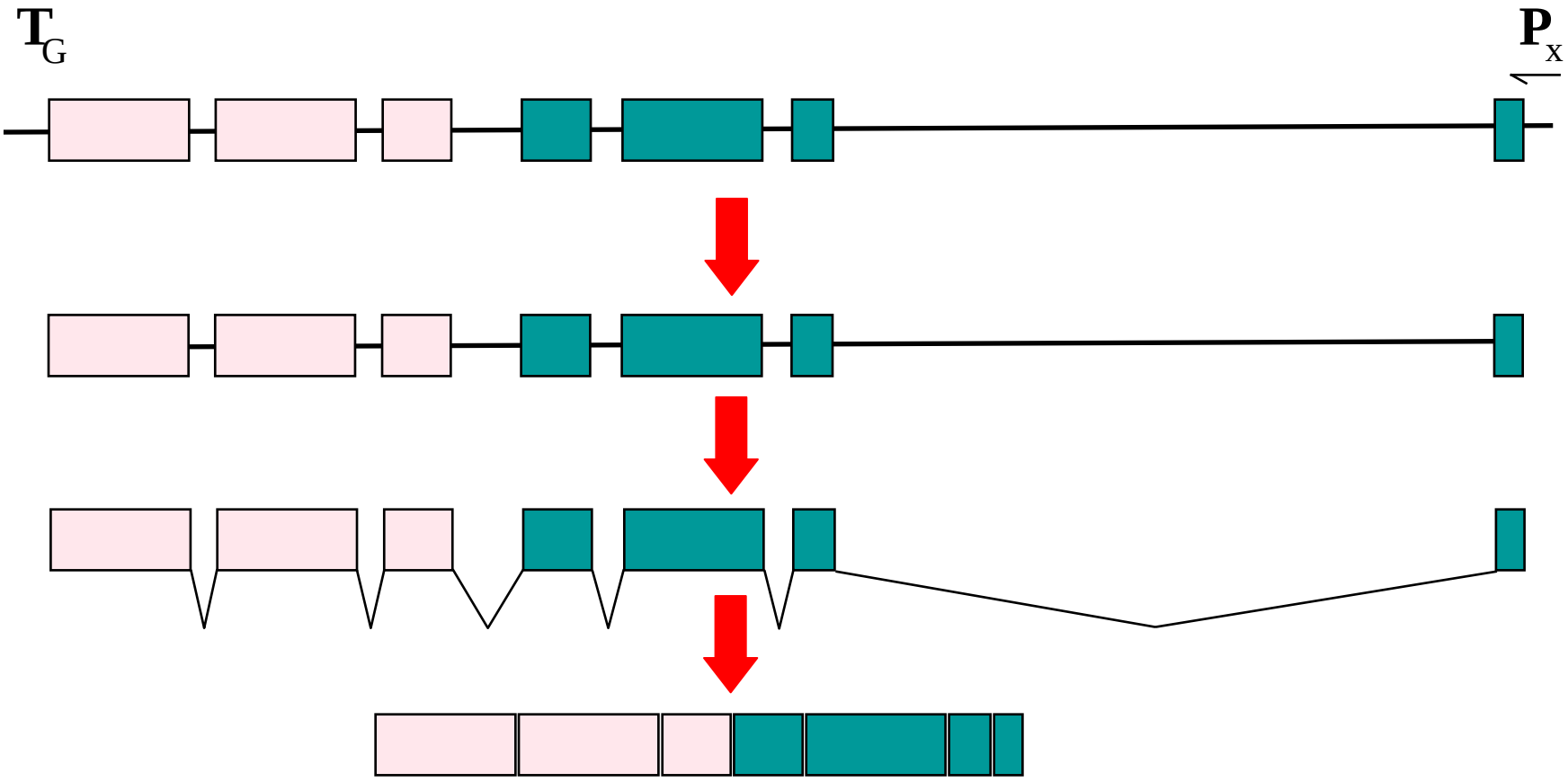
Introns: Gart



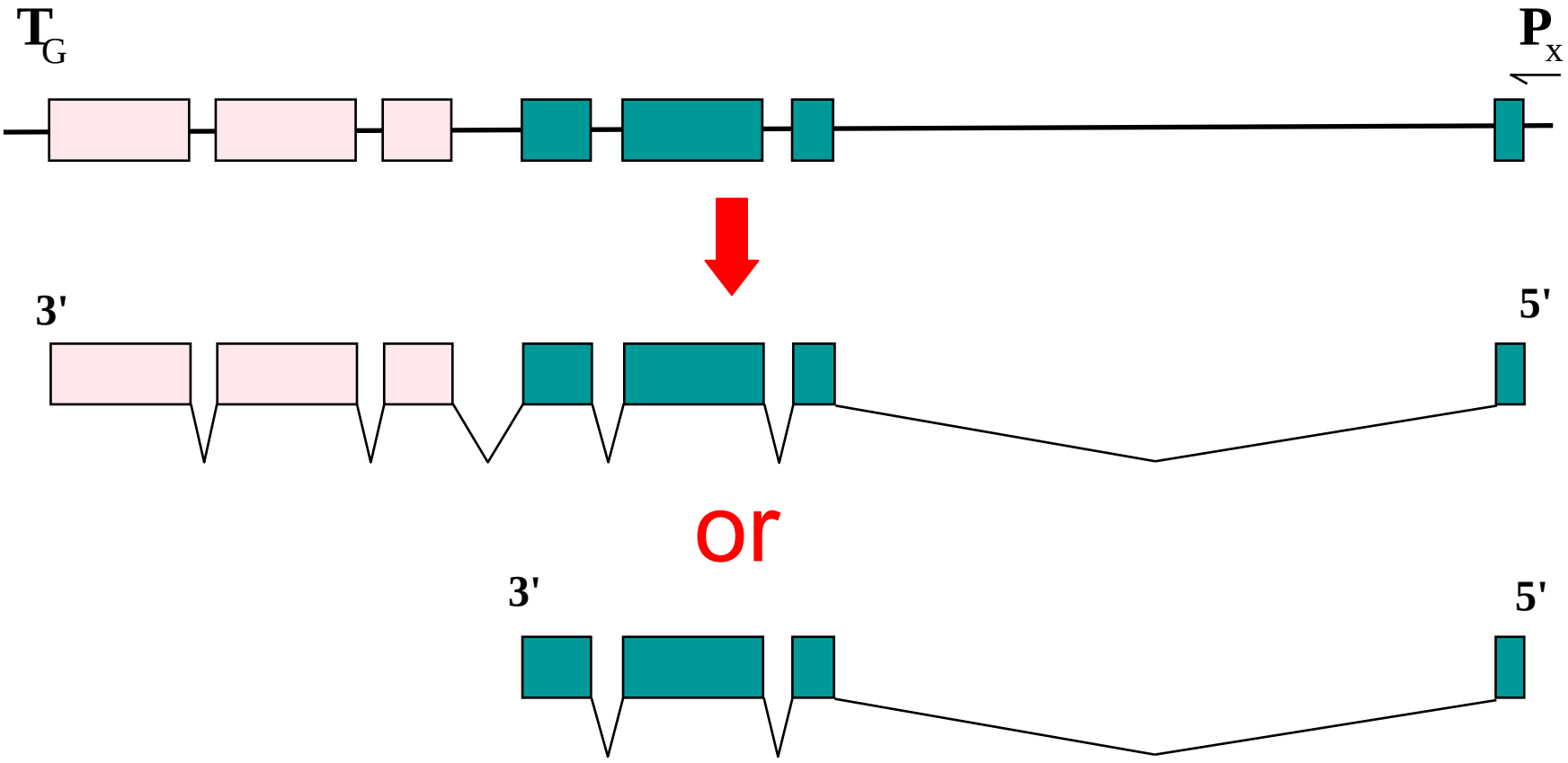
Introns: Gart



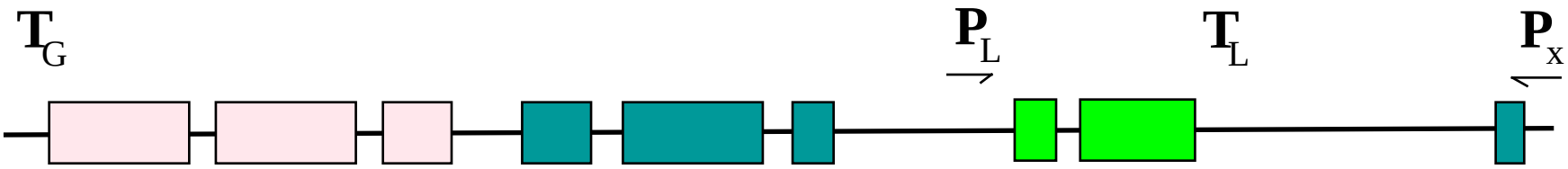
Introns: Gart



Alternative Splicing: Gart

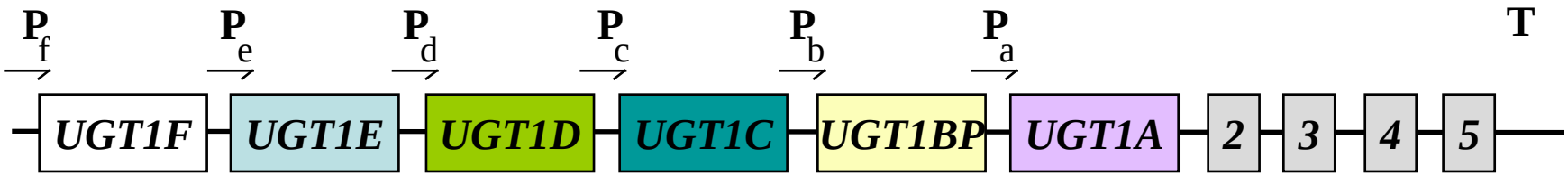


Nested Genes: Gart/Lcp



Henikoff, S., Keene, M.A., Fechtel, K., and Fristrom, J.W., 1986, Gene within a gene: Nested Drosophila genes encode unrelated proteins on opposite strands, *Cell* 44:33.

Nested Gene Families: UGT1



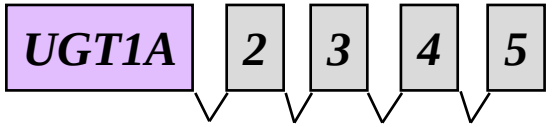
The UGT1 locus yields multiple transcripts through alternative promotion. Each promoter produces a transcript that is spliced so that the exon immediately adjacent to the promoter is joined with the four terminal exons shared by all of the transcripts.

Ritter, J.K., Chen, F., et al., 1992, A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini, *J. Biol. Chem.* 267:3257.

Nested Gene Families: UGT1

phenol UDP-glucuronosyltransferase:

bilirubin UDP-glucuronosyltransferases:

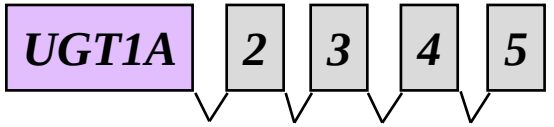


Nested Gene Families: UGT1

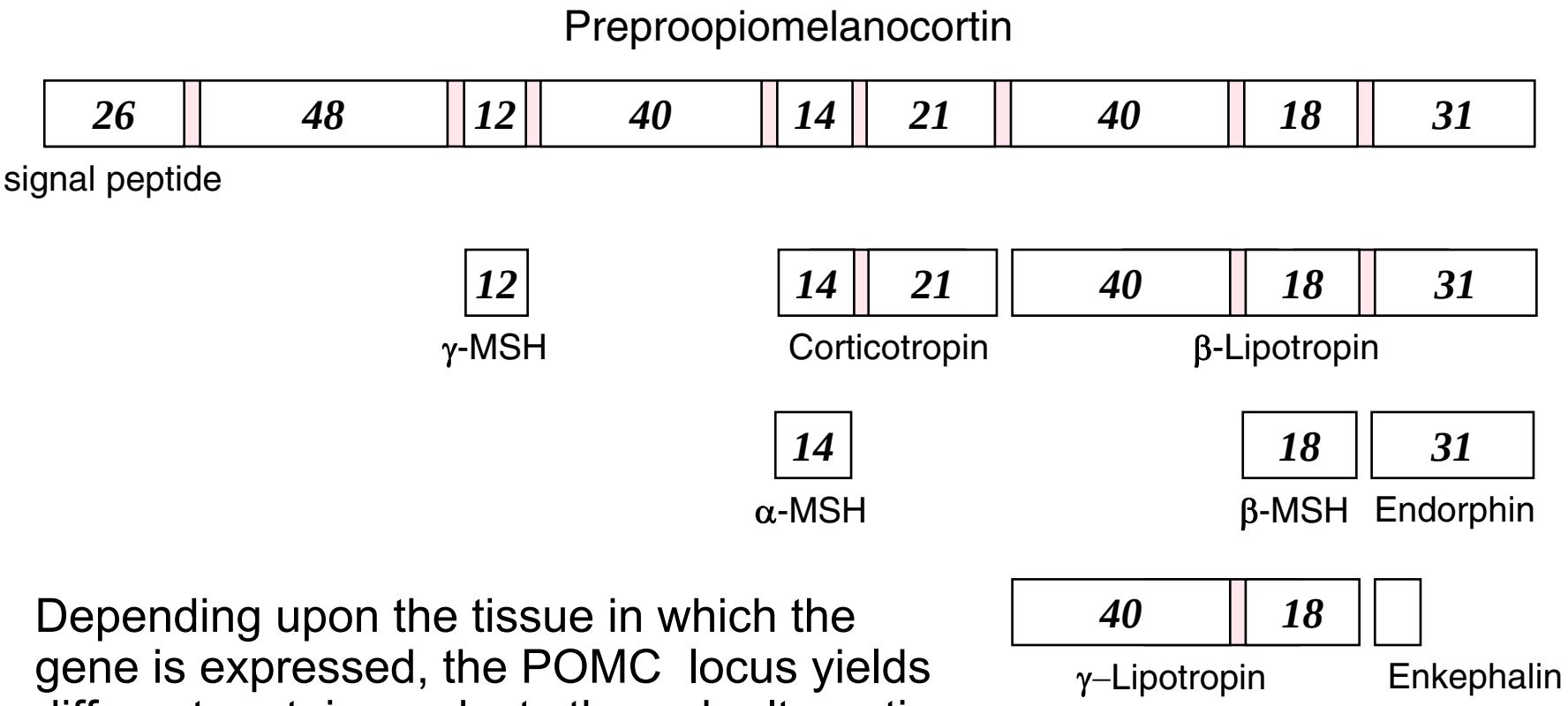
phenol UDP-glucuronosyltransferase:

bilirubin UDP-glucuronosyltransferases:

I challenge anyone to try to produce a definition of cistron that makes sense here.

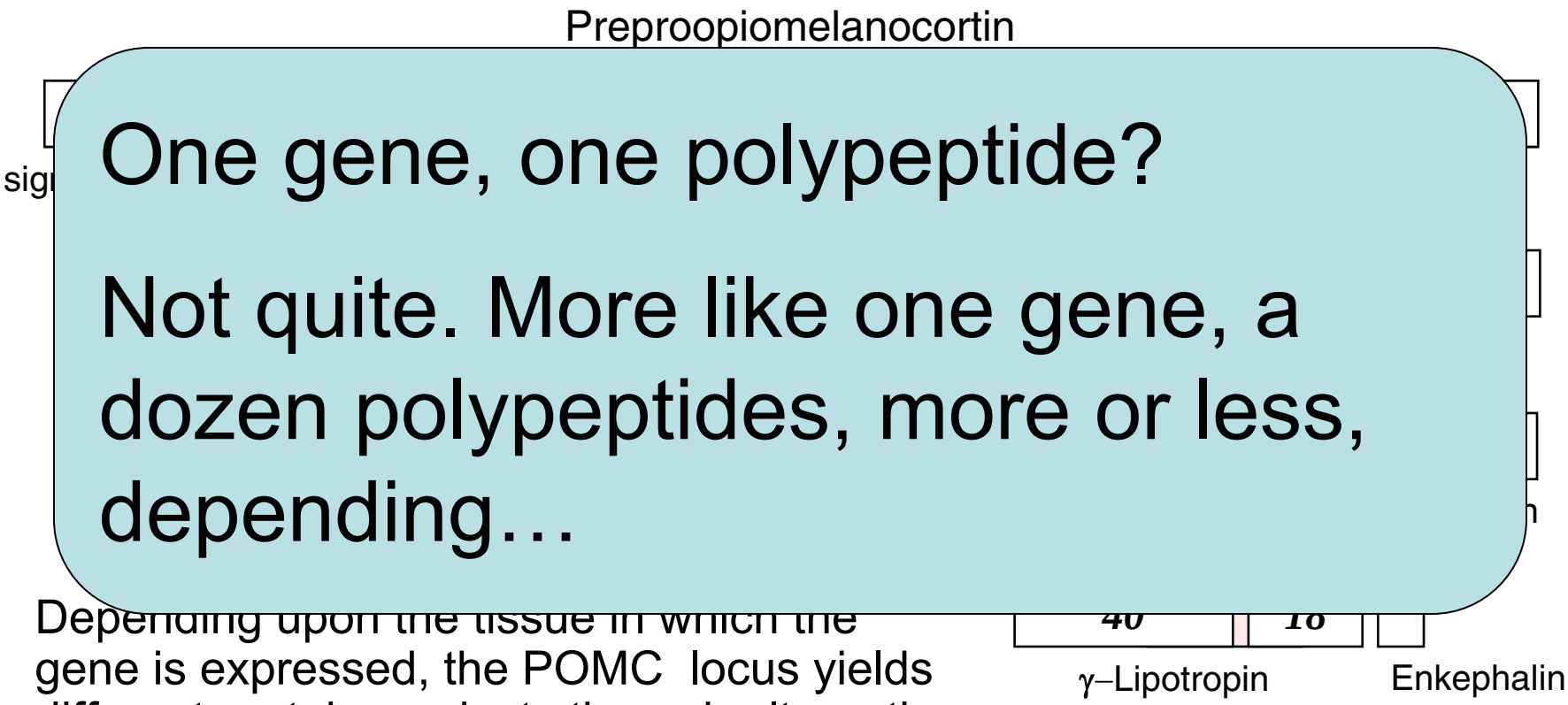


Multiple Gene Products: POMC



Depending upon the tissue in which the gene is expressed, the POMC locus yields different protein products through alternative processing of the resulting polypeptide.

Multiple Gene Products: POMC



Depending upon the tissue in which the gene is expressed, the POMC locus yields different protein products through alternative processing of the resulting polypeptide.

Before Molecular Biology:

Genes are the fundamental units of mutation, recombination, and heredity; they are arranged on the chromosomes likes beads on a string.

Before Molecular Biology:

Genes are the fundamental units of mutation, recombination, and heredity; they are arranged on the chromosomes likes beads on a string.

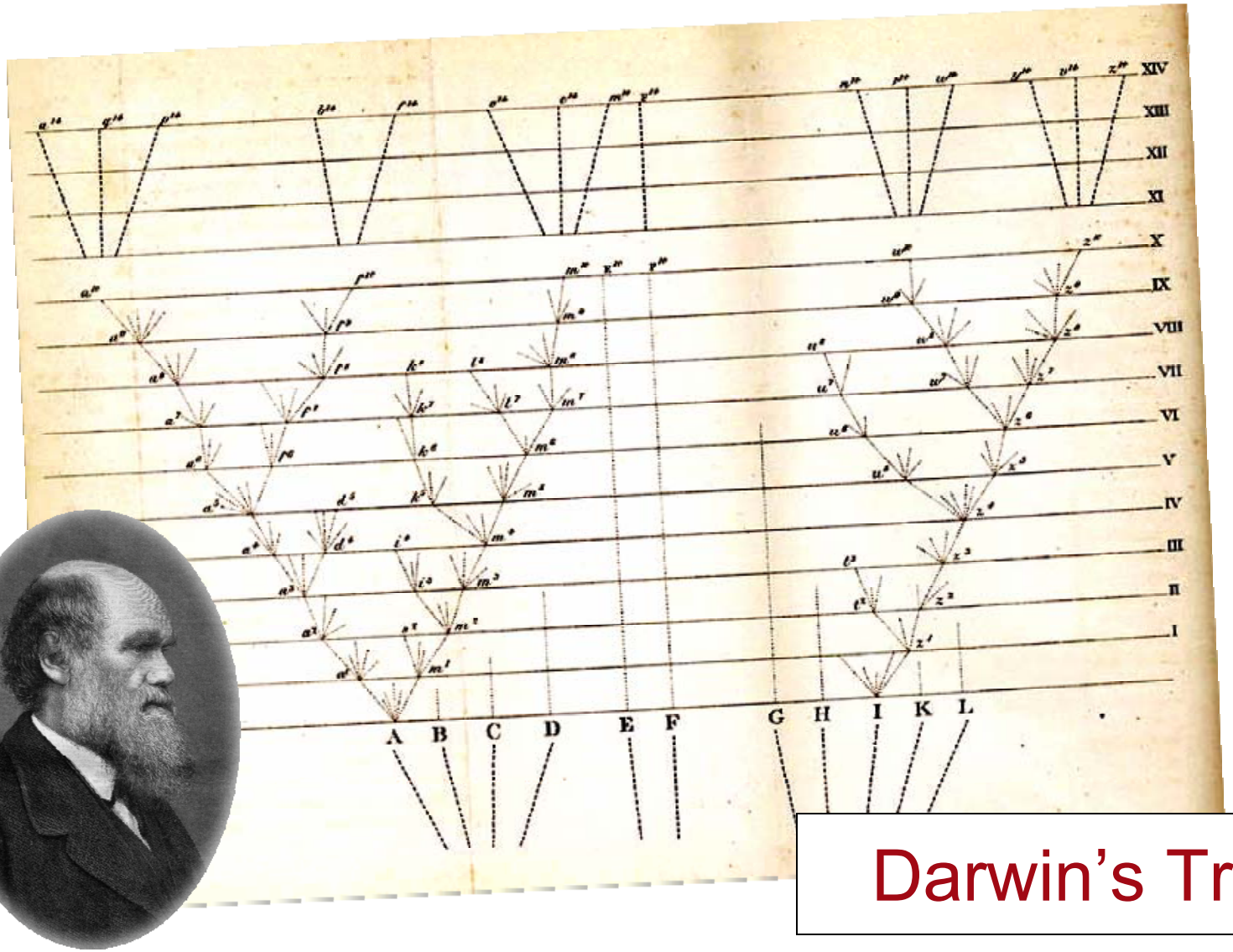
After Molecular Biology:

No fundamental units, no beads, and no string…

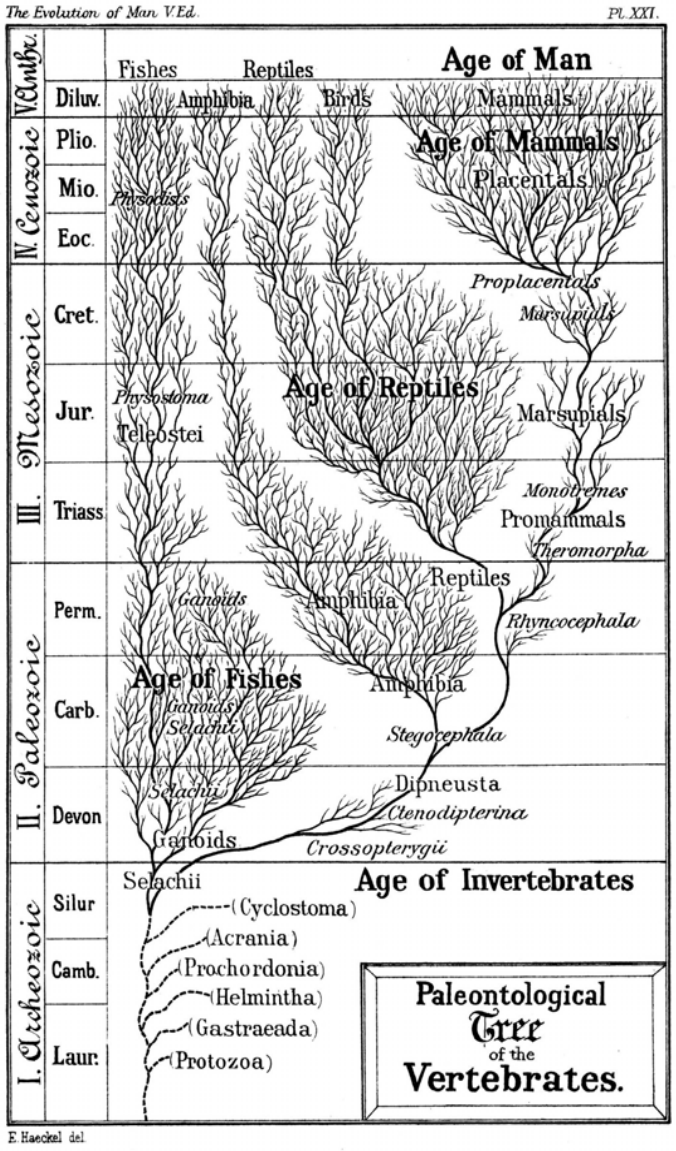
Reality is NOT Negotiable, II

Biodiversity

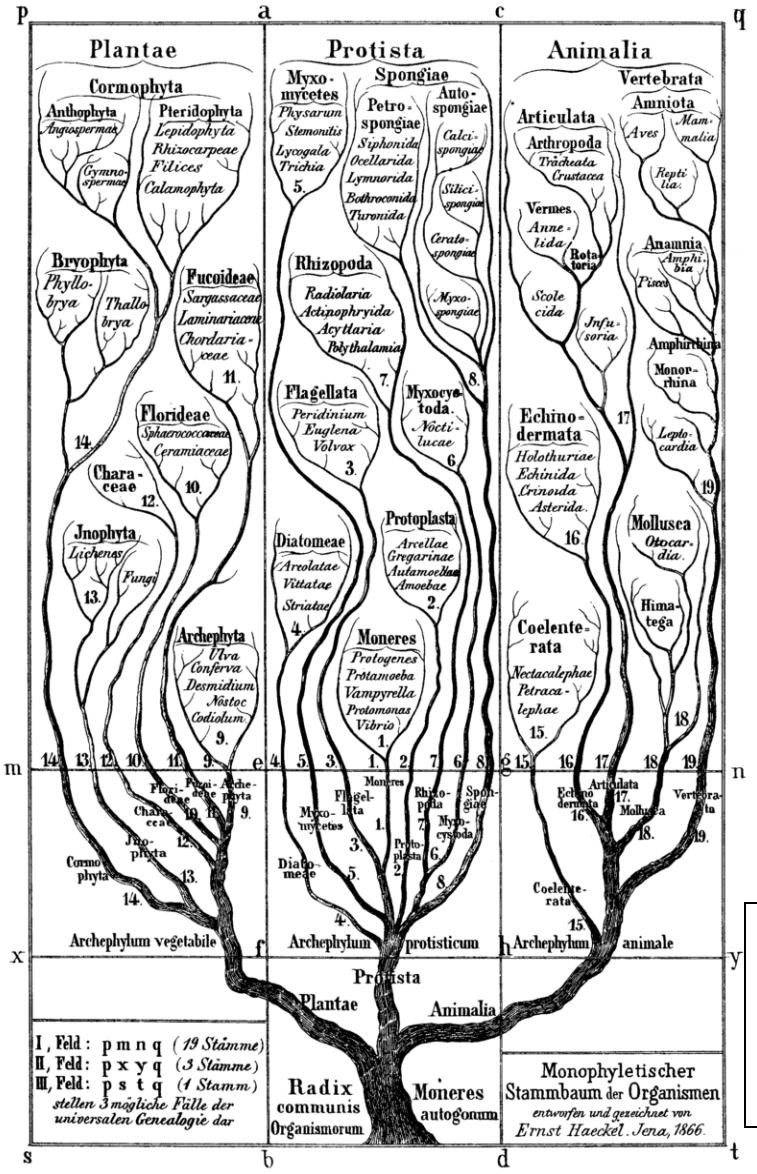
The Tree of Life



Darwin's Tree



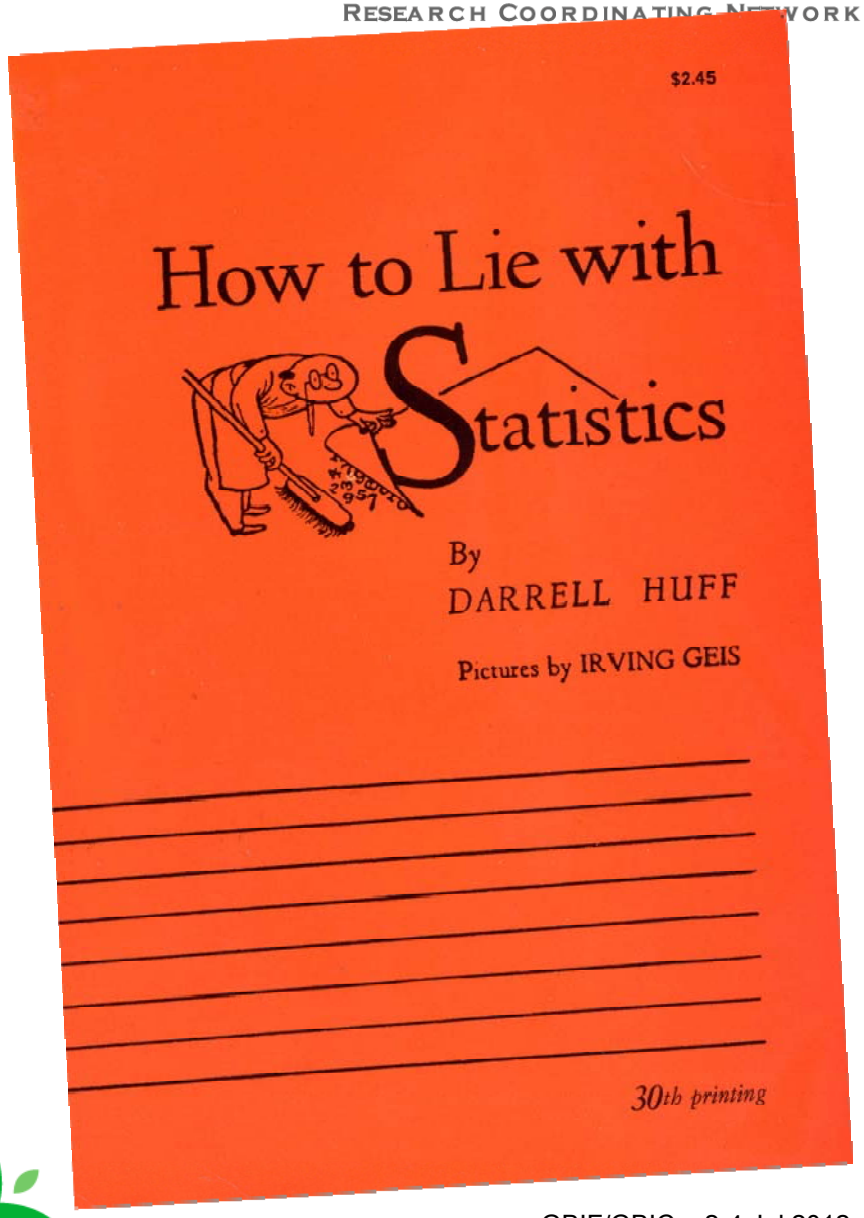
Haeckel's Paleontology Tree



Haeckel's Monophylogeny Tree

A Pretty Tree

Can Trees Lie?



Anyone of the right age probably remembers this book and how effectively Huff demonstrated the ease with which graphical devices can misrepresent quantitative information.

For example, …

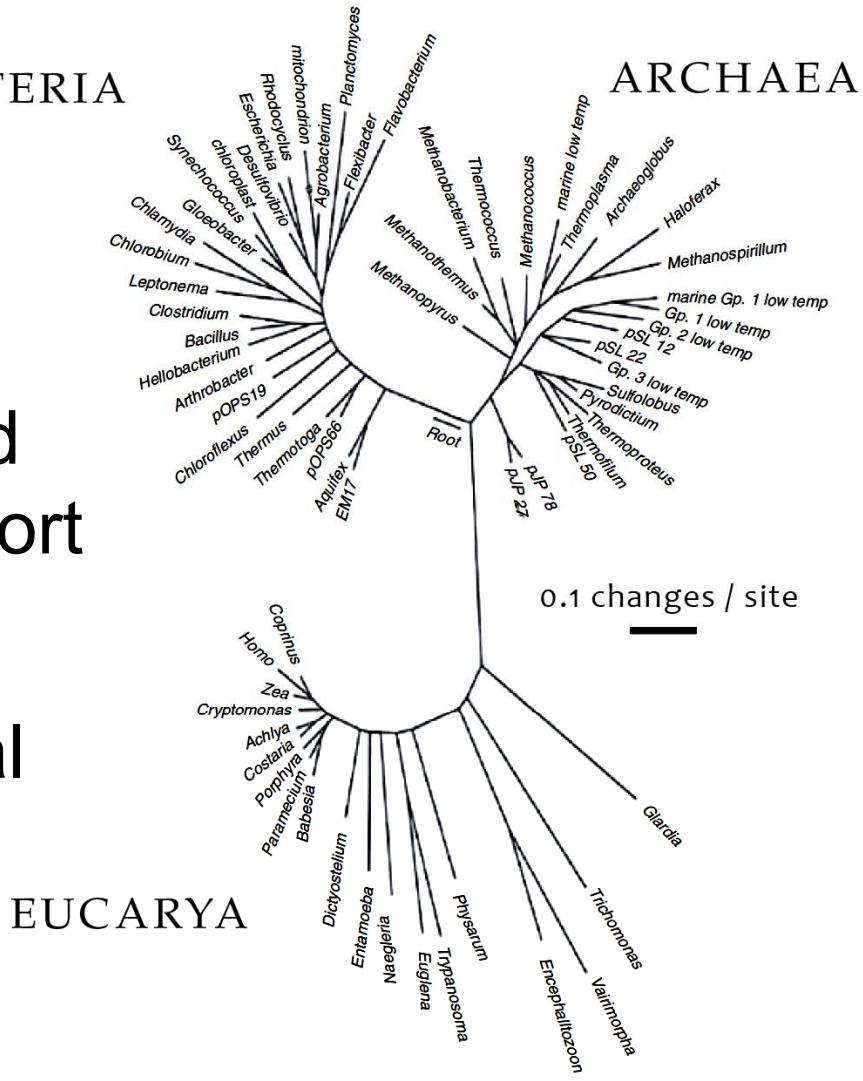
Lots of diversity up here

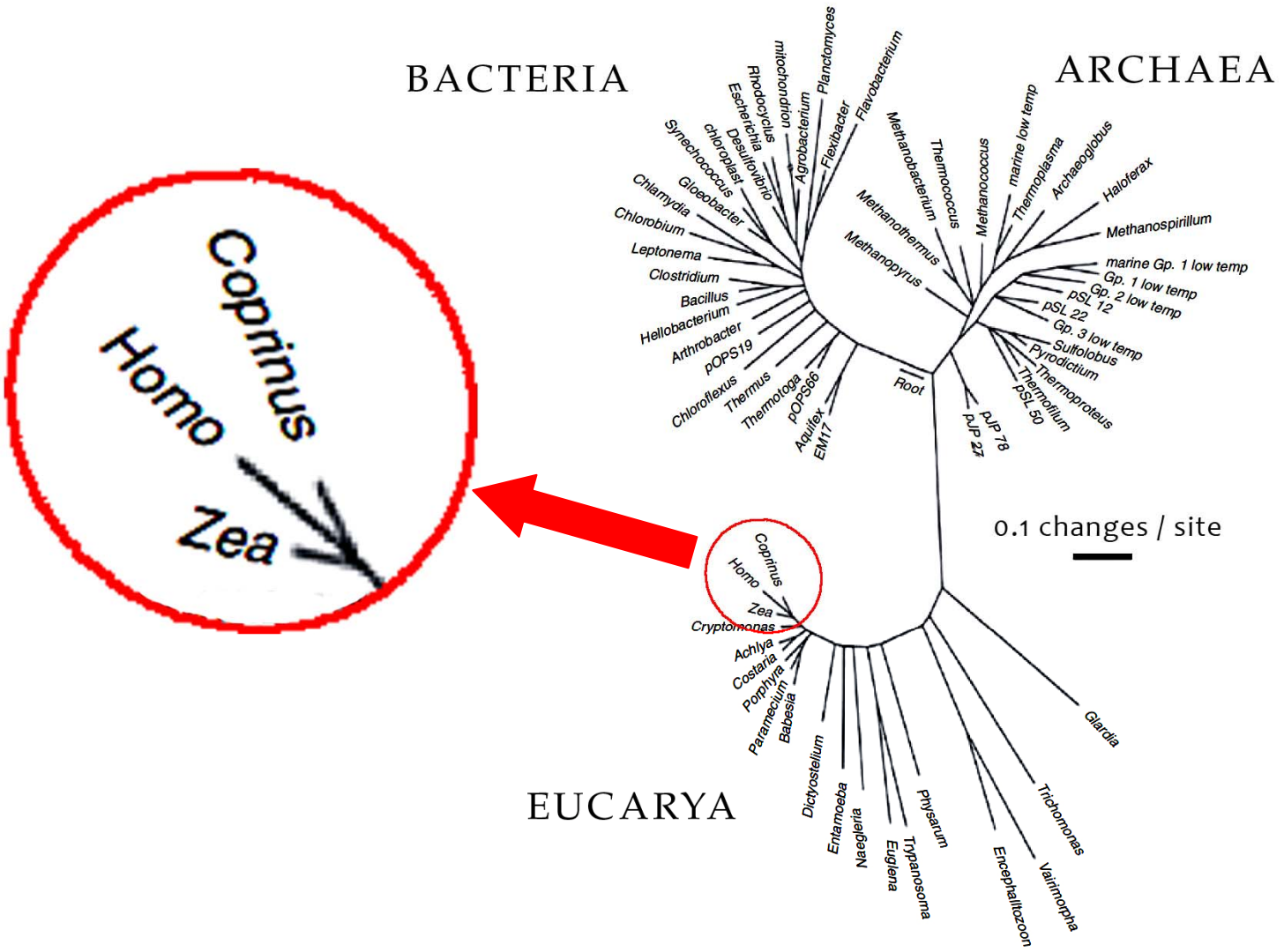
Lots of diversity up here

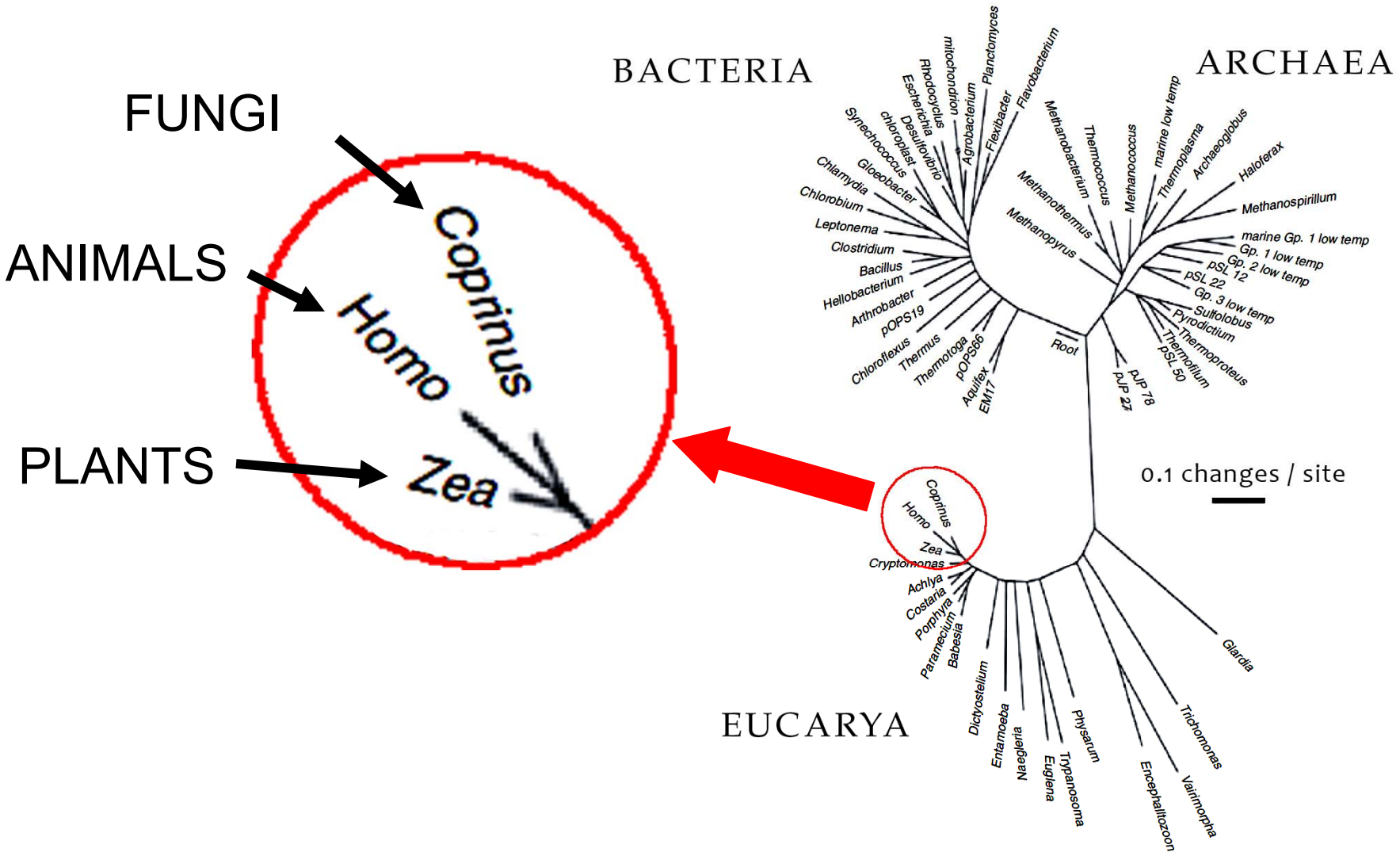
Not so much down here

Another tree, this one based on comparisons of rRNA short sub-unit sequences.

Branch length reflects actual divergence of sequence.




BACTERIA
ARCHAEA
EUCARYA
Coprinus
Homo
Zea
mitochondrion
Planctomyces
Agrobacterium
Flexibacter
Flavobacterium
Rhodocyclus
Escherichia
Desulfovibrio
chloroplast
Synechococcus
Gloeobacter
Chlamydia
Chlorobium
Leptonema
Clostridium
Bacillus
Hellobacterium
Arthrobacter
pOPS19
Chloroflexus
Thermus
Thermotoga
pOPS66
Aquifex
EM17
Root
pUP 27
pUP 78
pSL 50
Thermofilum
Thermoproteus
Pyrodictium
Sulfolobus
Gp. 3 low temp
pSL 22
pSL 12
Gp. 2 low temp
Gp. 1 low temp
marine Gp. 1 low temp
Methanospirillum
Haloferax
Archaeoglobus
Thermoplasma
marine low temp
Methanococcus
Thermococcus
Methanobacterium
Methanothermus
Methanopyrus
0.1 changes / site
Cryptomonas
Achlya
Costaria
Porphyra
Paramecium
Babesia
Dictyostelium
Entamoeba
Naegleria
Euglena
Trypanosoma
Physarum
Encephalitozoon
Vairimorpha
Trichomonas
Giardia


FUNGI
ANIMALS
PLANTS
Coprinus
Homo
Zea
BACTERIA
ARCHAEA
EUCARYA
mitochondrion
Planctomyces
Agrobacterium
Flexibacter
Flavobacterium
Rhodocyclus
Escherichia
chloroplast
Desulfovibrio
Synechococcus
Gloeobacter
Chlamydia
Chlorobium
Leptonema
Clostridium
Bacillus
Hellobacterium
Arthrobacter
pOPS19
Chloroflexus
Thermus
Thermotoga
pOPS66
Aquifex
EM17
Root
pUP 27
pUP 78
Thermofilum
pSL 50
Thermoproteus
Pyrodictium
Sulfolobus
Gp. 3 low temp
pSL 22
pSL 12
Gp. 2 low temp
Gp. 1 low temp
marine Gp. 1 low temp
Methanospirillum
Haloferax
Archaeoglobus
Thermoplasma
marine low temp
Methanococcus
Thermococcus
Methanobacterium
Methanothermus
Methanopyrus
0.1 changes / site
Cryptomonas
Achlya
Costaria
Porphyra
Paramecium
Babesia
Dictyostelium
Entamoeba
Naegleria
Euglena
Trypanosoma
Physarum
Encephalitozoon
Vairimorpha
Trichomonas
Giardia

If a tree is drawn to reflect **physiological** diversity, all of the differences among plants, animals, and fungi barely qualify as variations on a theme.

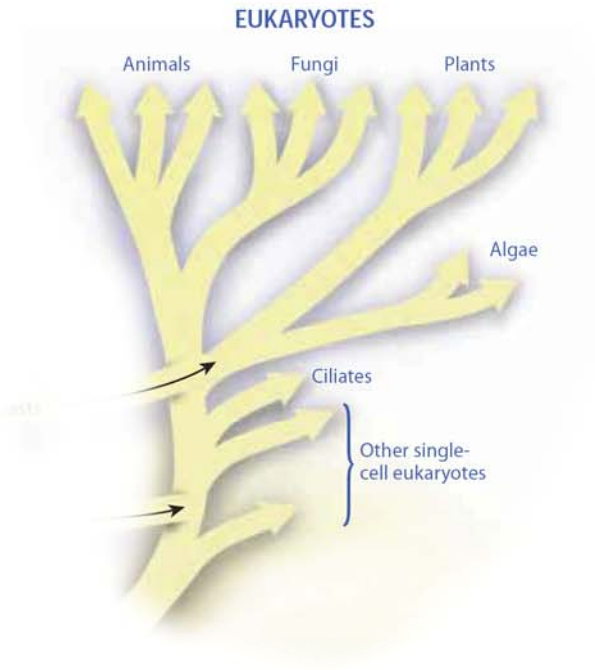
Mammals are essentially the same physiological trick, served up in a variety of different packages.

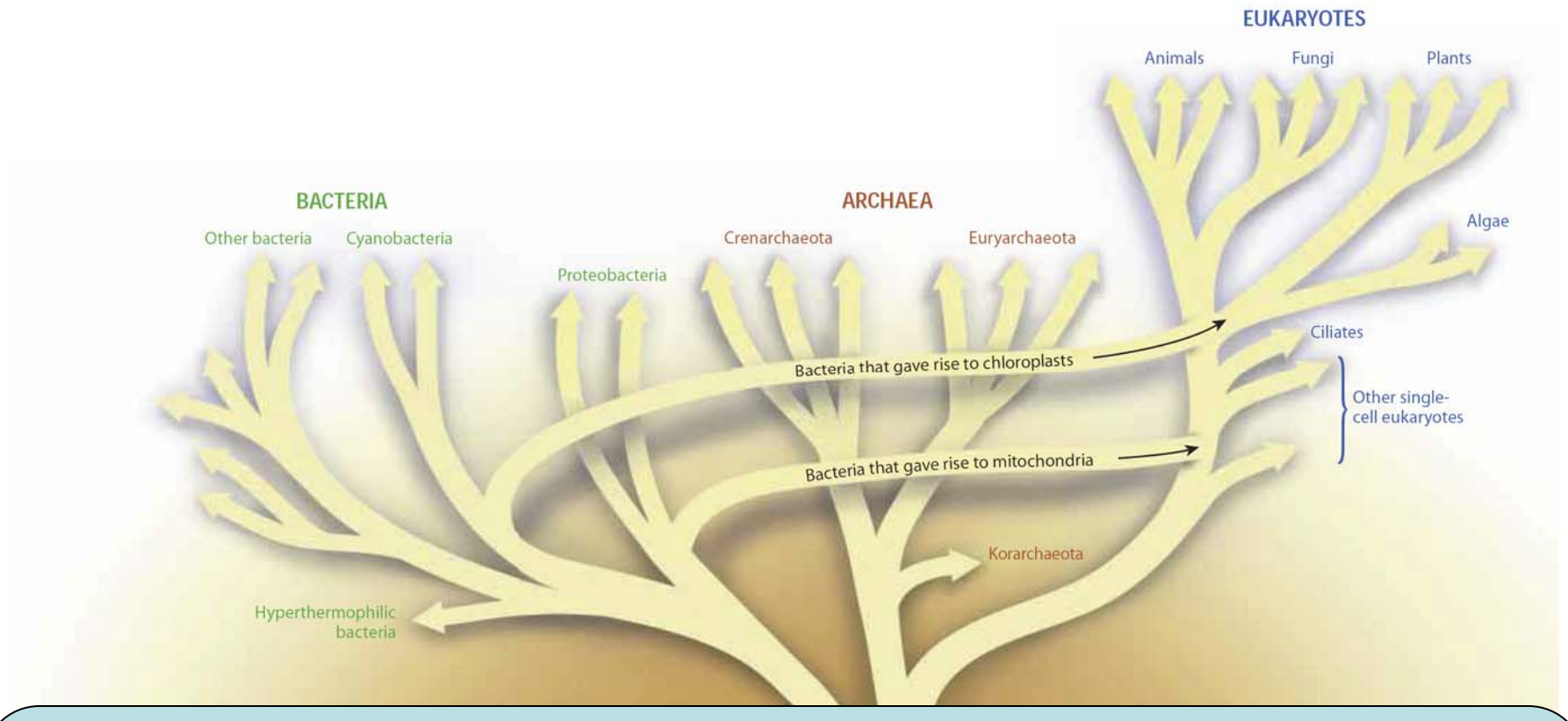
Average Mammalian Body Temperatures:

Human	37° C
Baboon	38° C
Fur seal	38° C
Humpback whale	36° C
Mouse (*Mus musculus*)	37° C
Elephant	36° C
Polar Bear	37° C

When is a Tree not a Tree?

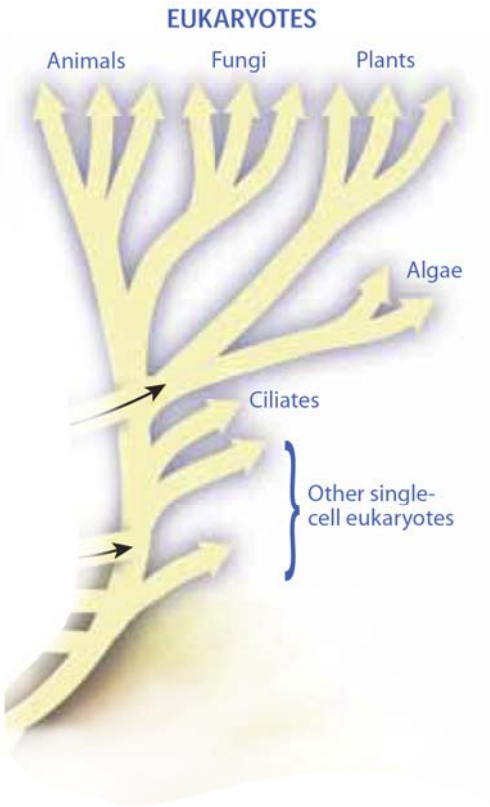
Most multicellular eukaryotic taxa can be arranged in a tree-like configuration, but when we include the origin of intra-cellular organelles things get more complicated…

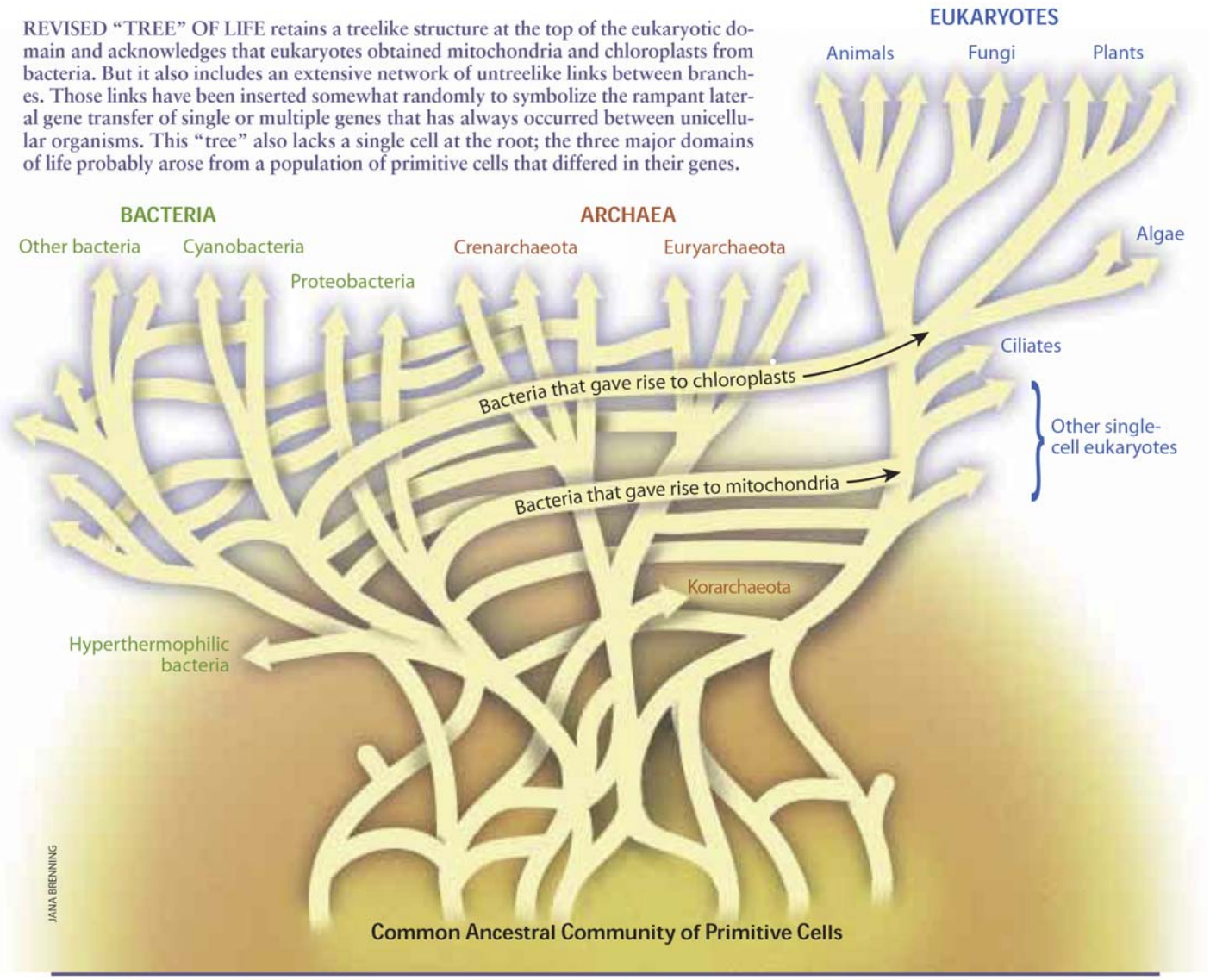




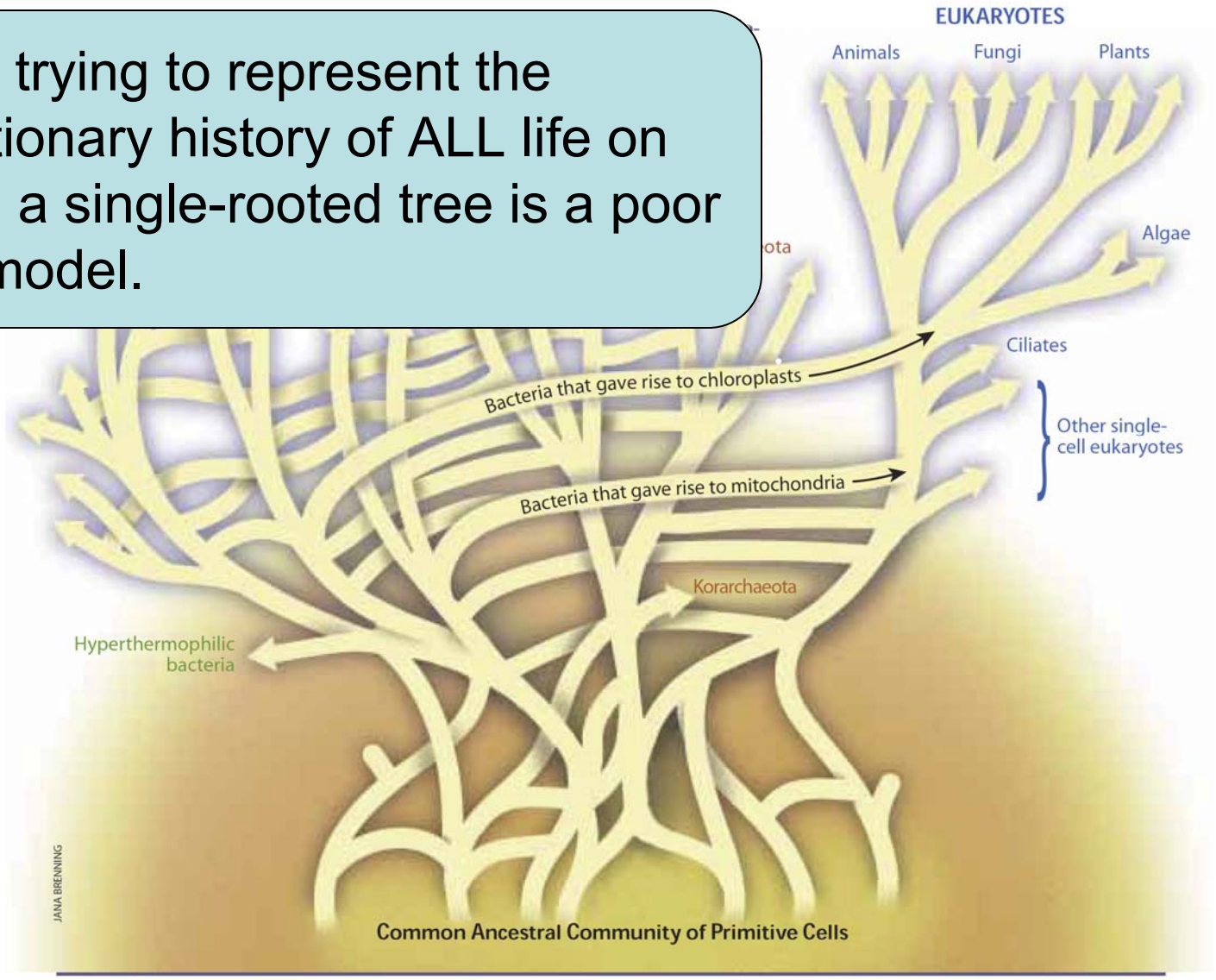
The origin of mitochondria, chloroplasts, and several other eukaryotic cell inclusions through endosymbiosis means that, technically speaking at least, not only are eukaryotic taxa polyphyletic, so are eukaryotic “individuals”.

We now know that horizontal gene transfer (HGT) occurs regularly among prokaryotes and most likely was the dominant form of inheritance during the early evolution of life on Earth.



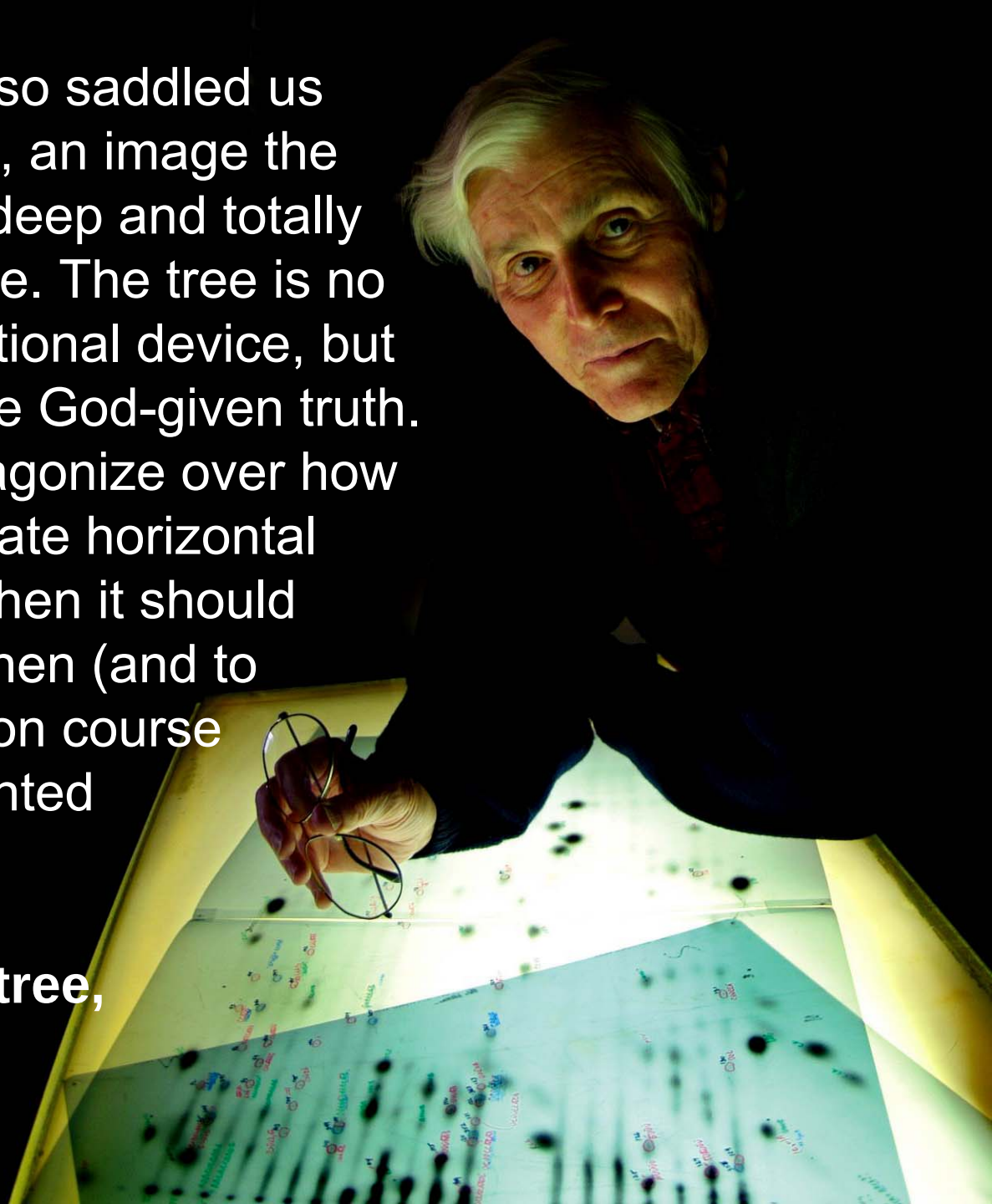

REVISED "TREE" OF LIFE retains a treelike structure at the top of the eukaryotic domain and acknowledges that eukaryotes obtained mitochondria and chloroplasts from bacteria. But it also includes an extensive network of untreelike links between branches. Those links have been inserted somewhat randomly to symbolize the rampant lateral gene transfer of single or multiple genes that has always occurred between unicellular organisms. This "tree" also lacks a single cell at the root; the three major domains of life probably arose from a population of primitive cells that differed in their genes.
EUKARYOTES
Animals
Fungi
Plants
BACTERIA
ARCHAEA
Other bacteria
Cyanobacteria
Crenarchaeota
Euryarchaeota
Proteobacteria
Algae
Ciliates
Bacteria that gave rise to chloroplasts
Other single-cell eukaryotes
Bacteria that gave rise to mitochondria
Korarchaeota
Hyperthermophilic bacteria
JANA BRENNING
Common Ancestral Community of Primitive Cells

When trying to represent the evolutionary history of ALL life on Earth, a single-rooted tree is a poor data model.



Classical biology has also saddled us with a phylogenetic tree, an image the biologist invests with a deep and totally unwarranted significance. The tree is no more than a representational device, but to the biologist it is some God-given truth. Thus, for example, we agonize over how the tree can accommodate horizontal gene transfer events, when it should simply be a matter of when (and to what extent) the evolution course can be usefully represented by a tree diagram:

Evolution defines the tree, not the reverse.



The Individual

The traditional biologist's view of life, first stated by Aristotle, starts by looking at individual organisms and asking what properties they have in common. (p 5)

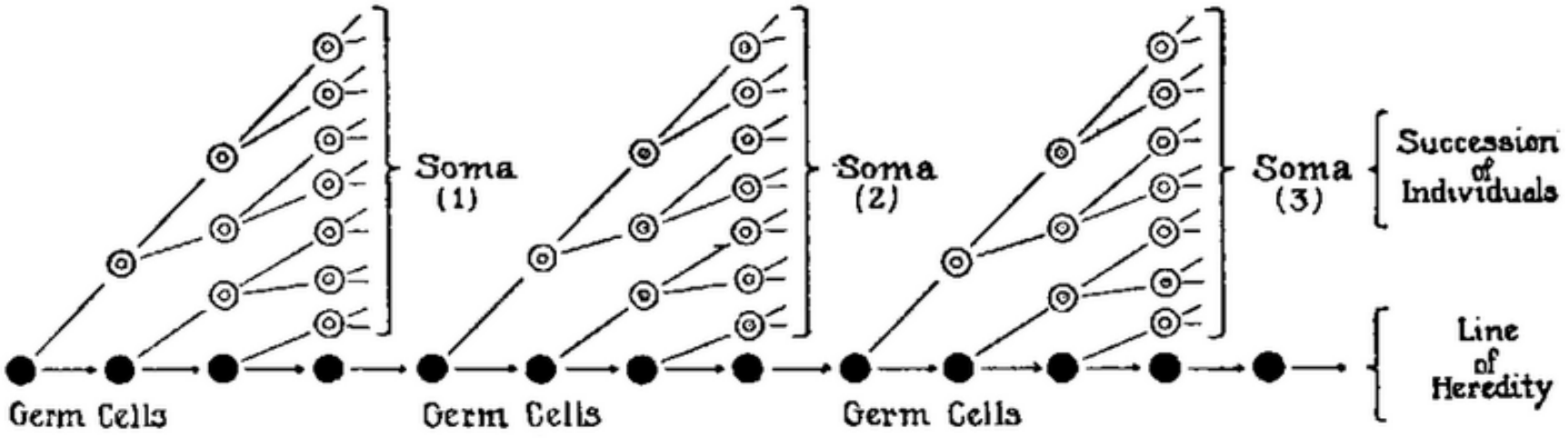
Morowitz, Harold J. 1992. *Beginnings of Cellular Life: Metabolism Recapitulates Biogenesis.* New Haven: Yale University Press.

The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)

It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)

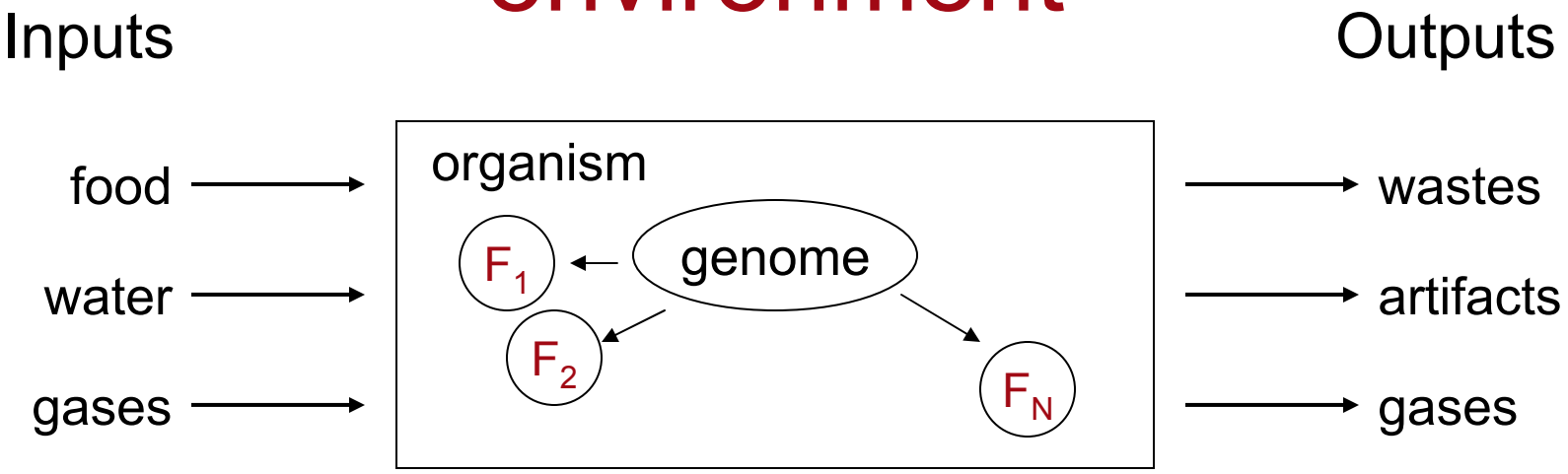
George Gaylord Simpson. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

Weismann's Germ Plasm Theory



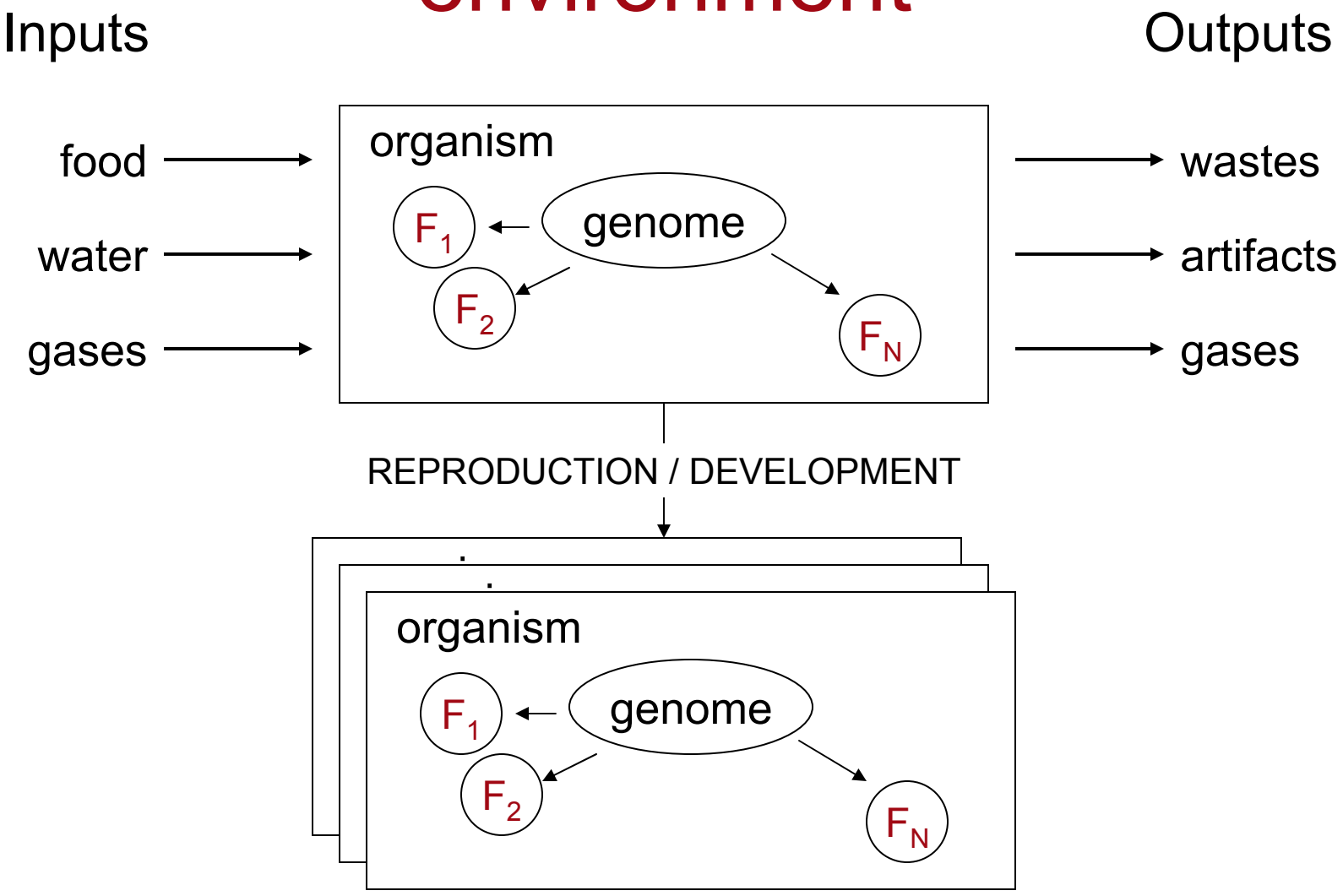
Organisms that best satisfy the notion of exhibiting "completely objective" individuals are animals that follow a Weismannian pattern of development – that is, an early sequestering of a separate germ line, with a complete logical and physical separation of somatic and germ tissue. Such animals begin as a zygote, then develop mitotically into a multicellular adult that, with luck, lives to adulthood and reproduces via the meiotic production of gametes.

environment



Such an individual can, with a certain amount of conceptual legerdemain, be envisioned as functioning autonomously in an environment, from which it acquires a few necessary inputs and into which it delivers certain outputs. It has a distinct body, with a clear external-internal boundary. Its phenotype is determined by its genotype, with internal physiological functions being carried out under the protein-mediated instructions of the genome. Reproduction involves “generations” in which new individuals come into existence, mature, and die.

environment



The subjects of classification are organisms and the subjects of taxonomy are classifications. (p11)

It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)

George Gaylord Simpson. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

The subjects of classification are organis onomy are cla

Not very many…

It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)

George Gaylord Simpson. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

The subjects of classification are organis onomy are cla

Not very important…

It seems obvious ... that the real unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)

George Gaylord Simpson. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

The su… organis… …onomy are cla…

Without doubt…

Indubitably…

Really ? ! ?

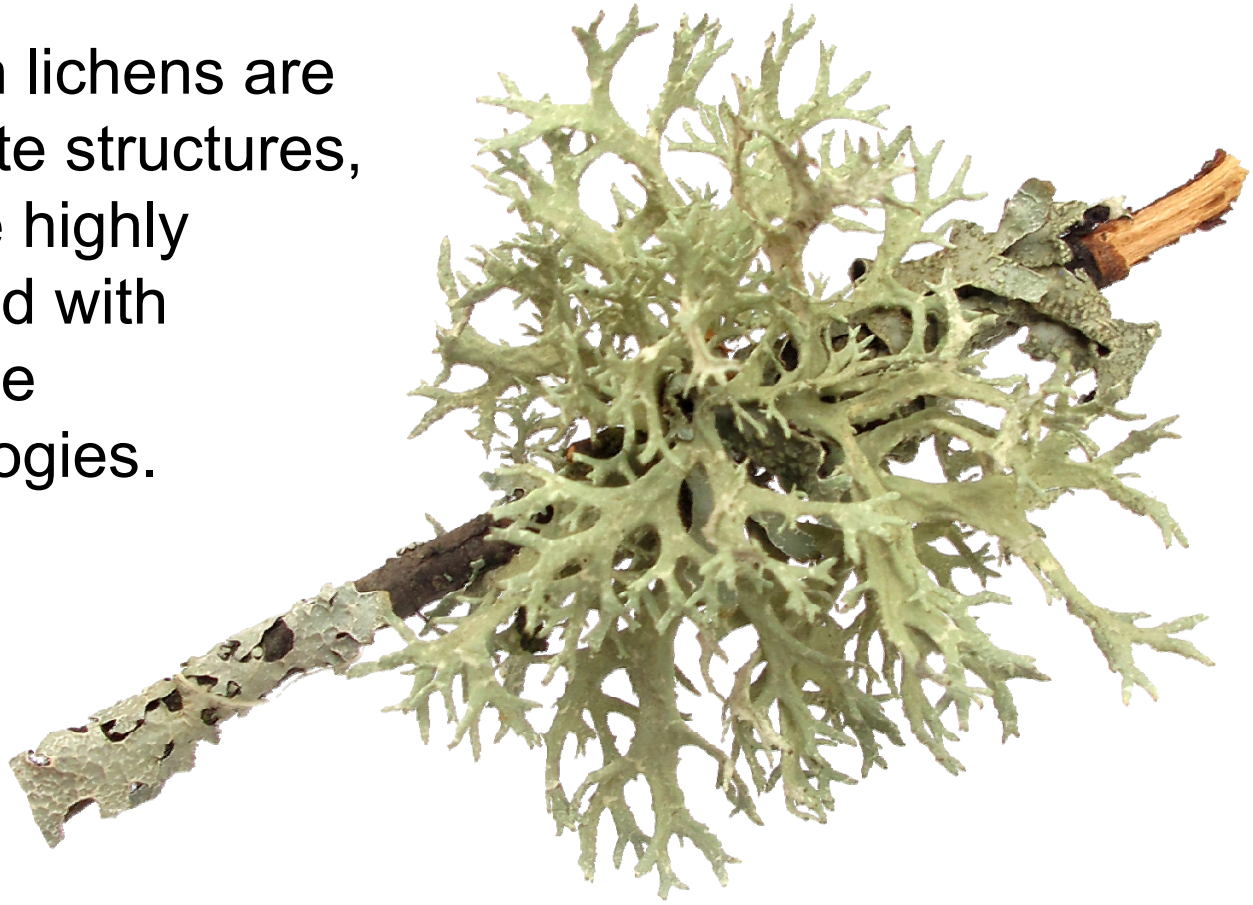
It seem… …unit in nature, the one thing that is usually completely objective in spite of some marginal cases, is the individual organism. (p. 18)

George Gaylord Simpson. 1961. *Principles of Animal Taxonomy*. New York: Columbia University Press.

What about lichens?

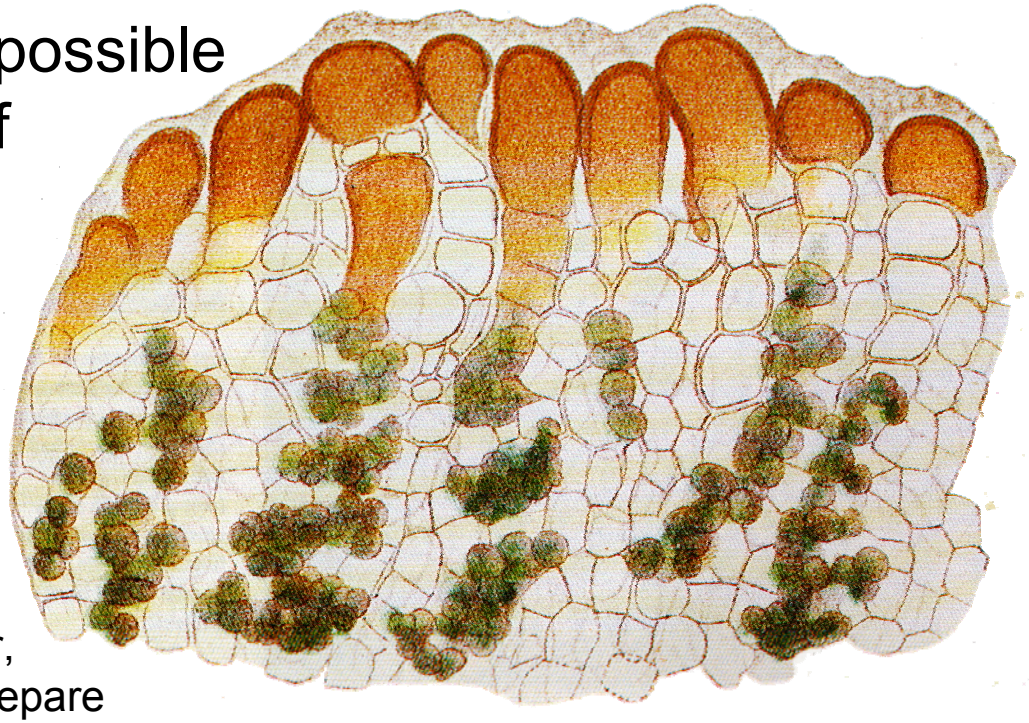
It is not easy to apply the concept of individual organism to a lichen.

Although lichens are composite structures, most are highly organized with distinctive morphologies.



Without a detailed microscopical examination, most lichens appear to be single entities.

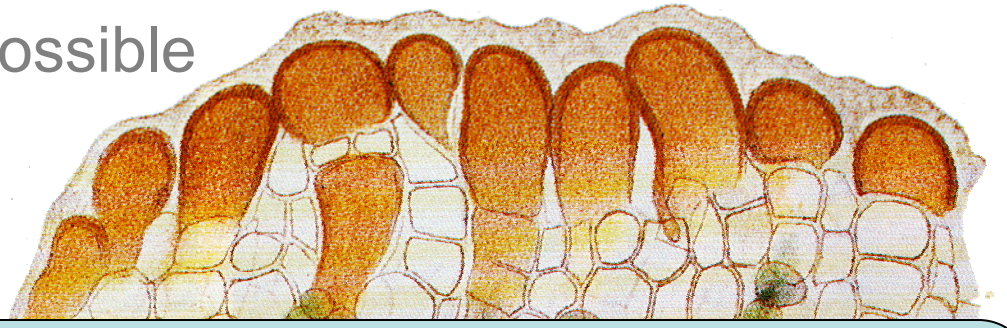
At the cellular level, it is possible to detect the presence of algae, embedded in the mycobiont tissue.



[L]ichens are not simple plants, not ordinary individuals in the ordinary sense of the word; they are, rather, colonies, which consist of hundreds of thousands of individuals, of which, however, one alone plays the master, while the rest, forever imprisoned, prepare the nutriment for themselves and their master. This fungus is a fungus of the class Ascomycetes, a parasite which is accustomed to live upon others' work. Its slaves are green algae, which it has sought out, or indeed caught hold of, and compelled into its service. It surrounds them as a spider its prey, with a fibrous net of narrow meshes, which is gradually converted into an impenetrable covering, but while the spider sucks its prey and leaves it dead, the fungus incites the algae found in its net to more rapid activity, even to more vigorous increase.

Simon Schwendener (1869)

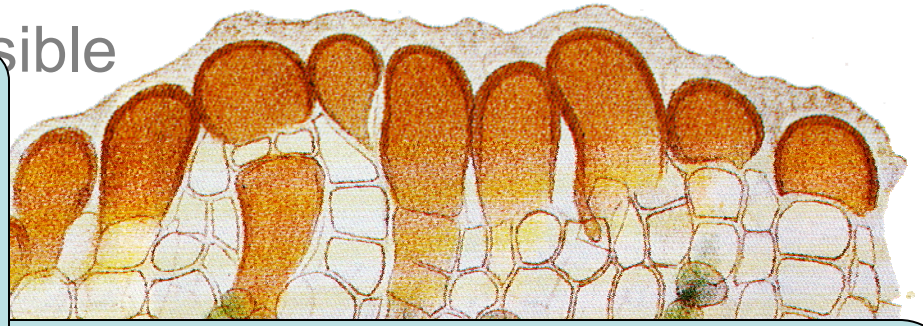
At the cellular level, it is possible to detect the presence of algae, embedded in the mycobiont tissue.



Schwendener's prose is dramatic, but makes the key point: a lichen is a living unit, but is not an "individual" as classically conceived. Nor can it be decomposed into individuals without giving up both its essence and its viability.

A lichen is a composite organism that cannot be subdivided into "individuals" and remain living. How does this square with the idea of the "individual" as the true "fundamental unit" of nature?

At the cellular level, it is possible



Maybe lichens are just a marginal case – the exception that proves the rule, and all that…

S atic, but makes the key point: a l t an individual. Nor can it be decomposed into individuals without giving up both its essence and its viability.

A lichen is a composite organism that cannot be subdivided into “individuals” and remain living. How does this square with the idea of the “individual” as the true “fundamental unit” of nature?

But they are certainly not rare…

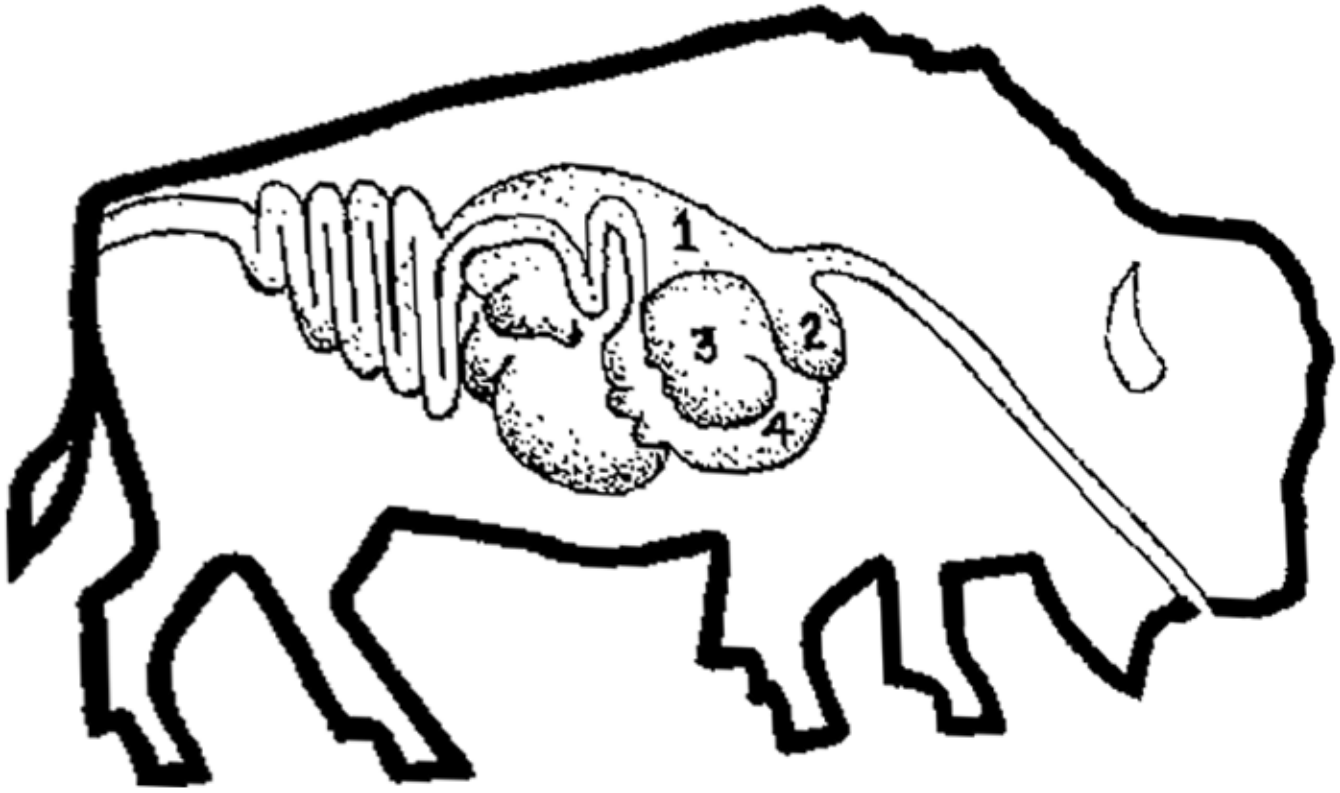
What about termites?

They are a critically important, sometimes dominant, species in many ecosystems, yet they cannot exist without their gut symbionts.

Can we really dismiss termites as a rare, marginal case?

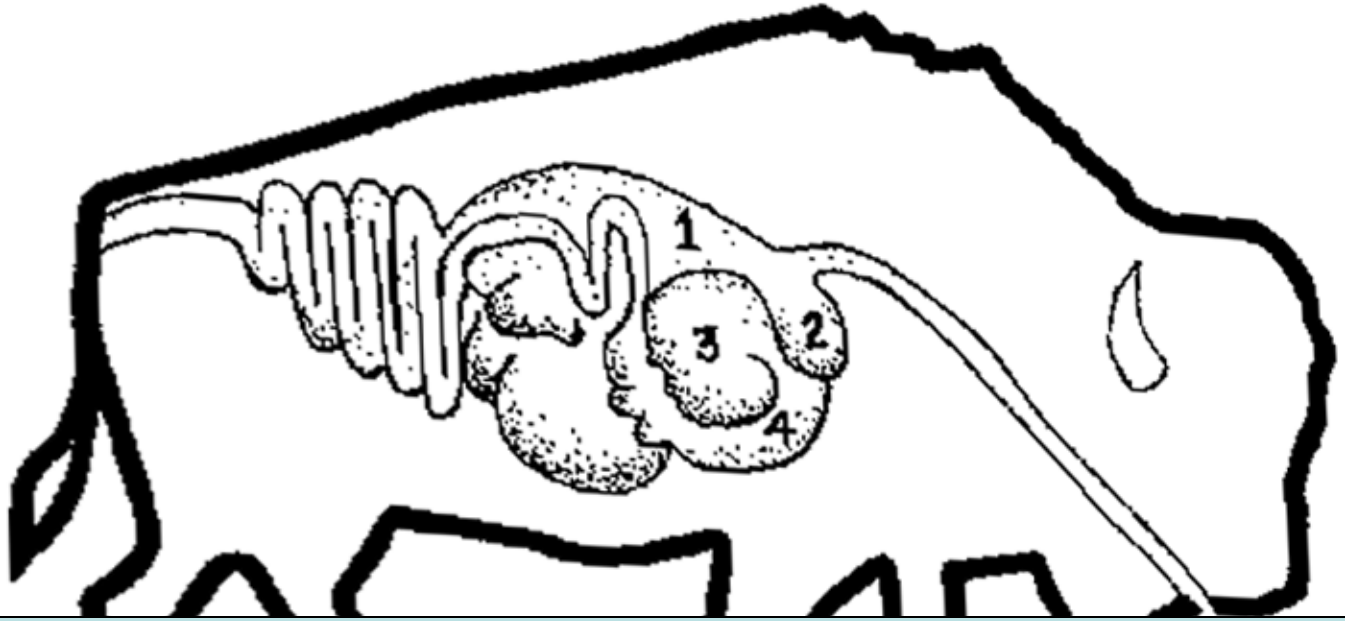
How about mammals? Surely they exemplify the idealized rugged individualism of autonomous organisms – the fundamental (and completely objective) unit of nature…

But wait, bison are ruminants



and cannot digest their food without the assistance of their gut flora.

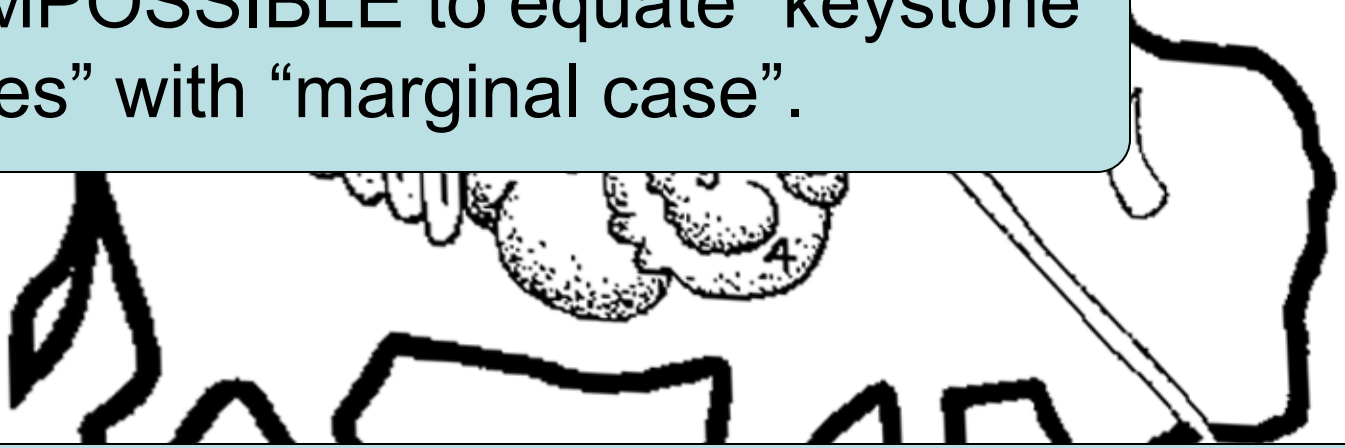
But wait, bison are ruminants



A ruminant is just as much a composite organism as a lichen. The alleged “individual” – whether mycobiont or buffalo – cannot obtain nutrients, and thus cannot live, without its microbial partners.

Ruminants are keystone species in many grassland ecosystems.

It is IMPOSSIBLE to equate "keystone species" with "marginal case".



A ruminant is just as much a composite organism as a lichen. The alleged "individual" – whether mycobiont or buffalo – cannot obtain nutrients, and thus cannot live, without its microbial partners.

OK, people then.

Surely human beings aren't just hopped up lichens…

8 June 2012

14 June 2012

NEWS & VIEWS

MICROBIOLOGY

Learning about who we are

Microbial inhabitants outnumber our body's own cells by about ten to one. These residents have become the subject of intensive research, which is beginning to elucidate their roles in health and disease. SEE ARTICLES P.207 & P.215

DAVID A. RELMAN

The dawn of the twenty-first century has seen the emergence of a major theme in biomedical research: the molecular and genetic basis of what it is to be human. Surprisingly, it turns out that we owe much of our biology and our individuality to the microbes that live on and in our bodies — a realization that promises to radically alter the principles and practice of medicine, public health and basic science. It is therefore appropriate that ever more research is focused on these microbes and their genes, which together are known as the human microbiome[1]. In this issue, the Human Microbiome Project Consortium[2,3] publishes the most extensive catalogue yet of organisms and genes pertaining to our microbiomes.

The first observations of indigenous human microbiota were published more than 300 years ago, soon after the invention of the microscope. Today's view of the microbial world has been radically improved by DNA-sequencing technology. In the wake of the Human Genome Project, calls were issued[1,4] for enhanced efforts to be made to characterize the 'second human genome' — the human microbiome. At the end of 2007, the US National Institutes of Health (NIH) launched the Human Microbiome Project (HMP) and, in early 2008, the European Commission and China initiated the Metagenomics of the Human Intestinal Tract (MetaHIT) project. Other countries have begun similar ventures, motivated in part by an interest in better defining their biological heritage.

Two studies, by Huttenhower *et al.*[2] (page 207) and Methé *et al.*[3] (page 215), together with 15 other papers[5,6] that are being published simultaneously elsewhere, comprise the first reports of the HMP Consortium research groups. The primary data, as described by Methé and colleagues[3], were derived from samples collected from 242 healthy adults in the United States, at 15 (for males) or 18 (for females) body sites — from the skin, nose, mouth, throat, vagina and faeces (to represent the distal gastrointestinal tract). Each person was sampled up to three times over 22 months, generating a total of 11,174 samples.

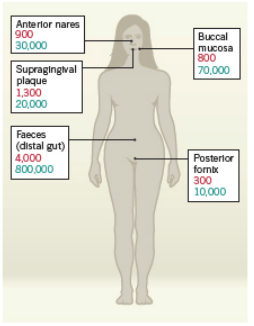


Figure 1 | Variation in diversity. Researchers of the Human Microbiome Project are studying the microbial inhabitants of the human body, using samples taken from 242 healthy adults at 15 (for males) or 18 (for females) body sites — from the skin (four sites), mouth and throat (nine sites), vagina (three sites), nostrils and faeces (to represent the distal gastrointestinal tract). Huttenhower *et al.*[2] and Methé *et al.*[3] have estimated the number of microbial species and their genes in these samples, and found substantial variation in microbial community composition at different body habitats. The two groups used different counting methodologies, and their numbers vary accordingly, such that exact figures are not available. However, crude estimations[3] of number of microbial species (red) and number of microbial genes (blue) are shown for examples of: sites containing high species diversity, such as the gastrointestinal tract and teeth (supragingival plaque); sites with intermediate diversity, such as the inside of the cheek (buccal mucosa) and nostrils (anterior nares); and sites with lower diversity, such as the vaginal posterior fornix. The authors also found substantial variation in both the diversity and the composition of the microbial communities at different sites within the same general body region.

The consortium researchers obtained the nucleotide sequence of the small-subunit ribosomal RNA — a molecule found in all cellular life — from microorganisms in 5,177 of these samples[3]. These sequences are commonly used to infer the genetic relationships between organisms. The researchers also surveyed the genomes of the microbes in 681 of the samples[3] using a shotgun sequencing approach, which generates random sequences (reads) from a complex pool of DNA molecules. The reads are then assembled on the basis of overlapping sequence similarity, allowing researchers to identify genes and to predict the functions of the proteins that they encode.

The investigators mapped their reads to all available microbial and viral genome sequences to assess community composition — the different types of microbes and their relative abundance — at the various body sites. The researchers also determined the whole-genome sequences of about 800 bacterial strains isolated from humans (from a planned total of 3,000); these sequences have been placed in public databases and can be used as reference genomes for comparative purposes. The consortium authors conclude[2,3] that they have identified the majority of the common microbial taxa and their genes present in these 242 healthy humans.

One of the great strengths of the HMP is that samples were collected simultaneously from multiple body habitats of the same individuals. This allowed Huttenhower *et al.*[2] to discover that taxonomic and genetic diversity were greatest in tooth and stool samples, intermediate in skin samples and on the inside surface of the cheek, and lowest in vaginal samples[2,3,7] (Fig. 1). The researchers report that each habitat is characterized by a small number of highly abundant 'signature' taxa, but that the relative representation of taxa and genes in each habitat varies considerably between individuals. In most samples, high-abundance taxa are accompanied by low-abundance taxa from the same genus, suggesting that within-community niche specialization occurs. These findings confirm those of an earlier study[8], which demonstrated that body habitat accounts for much of the variation in bacterial community composition. Although there is clear evidence for individuality in people's microbiome compositions, the limited temporal scope of the HMP data set prevents a robust analysis of how these communities change over time.

As shown previously for faecal samples[9],

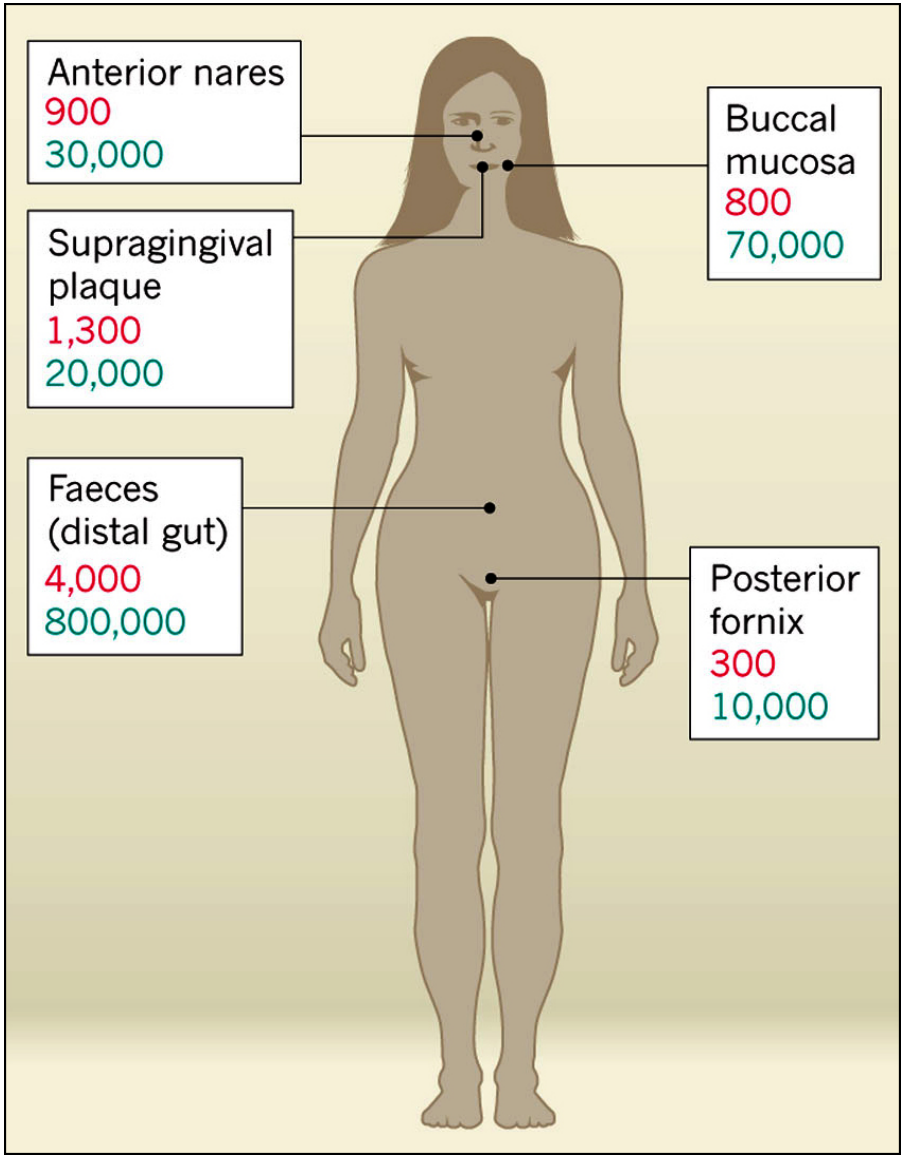


Figure 1 Variation in diversity. Researchers of the Human Microbiome Project are studying the microbial inhabitants of the human body, using samples taken from 242 healthy adults at 15 (for males) or 18 (for females) body sites — from the skin (four sites), mouth and throat (nine sites), vagina (three sites), nostrils and faeces (to represent the distal gastrointestinal tract). Huttenhower *et al.* and Methé *et al.* have estimated the number of microbial species and their genes in these samples, and found substantial variation in microbial community composition at different body habitats. The two groups used different counting methodologies, and their numbers vary accordingly, such that exact figures are not available. However, crude estimations of number of microbial species (red) and number of microbial genes (blue) are shown for examples of: sites containing high species diversity, such as the gastrointestinal tract and teeth (supragingival plaque); sites with intermediate diversity, such as the inside of the cheek (buccal mucosa) and nostrils (anterior nares); and sites with lower diversity, such as the vaginal posterior fornix. The authors also found substantial variation in both the diversity and the composition of the microbial communities at different sites within the same general body region.

Microbial World:

- In addition to being ubiquitous and abundant on and in every macroscale organism, prokaryotes occur in every imaginable environment (and maybe a few more).

Microbial World:

- In addition to being ubiquitous and abundant on and in every macroscale organism, prokaryotes occur in every imaginable environment (and maybe a few more).
- They are abundant (more bacteria in a bucket of seawater than there are mammals in Africa).

Microbial World:

- In addition to being ubiquitous and abundant on and in every macroscale organism, prokaryotes occur in every imaginable environment (and maybe a few more).
- They are abundant (more bacteria in a bucket of seawater than there are mammals in Africa).
- They are locally diverse (1 g of soil contains 10^7-10^9 prokaryotic cells, with 2,000–18,000 different genomes).

Microbial World:

- In addition to being ubiquitous and abundant

To repeat: That's 2,000 to 18,000 separate "species" in a teaspoon of soil…

of seawater than there are mammals in Africa).

- They are locally diverse (1 g of soil contains 10^7-10^9 prokaryotic cells, with 2,000–18,000 different genomes).

Bottom Line:

- This is fundamentally a microbial biosphere. Half of the biomass and most of the diversity occur in microbes.

Bottom Line:

- This is fundamentally a microbial biosphere. Half of the biomass and most of the diversity occur in microbes.
- Microbes occur in free-living communities and also in very tight, functional associated with ALL multi-cellular organisms.

Bottom Line:

- This is fundamentally a microbial biosphere. Half of the biomass and most of the diversity occur in microbes.
- Microbes occur in free-living communities and also in very tight, functional associated with ALL multi-cellular organisms.
- The notion of individual organisms, as fundamental units in nature, is not objective "truth" – instead it is, at best, a useful approximation.

Bottom Line:

From the perspective of community biology (which arguably is synonymous with "biology"), the "individual" is a reductionist abstraction.

It is useful in the way "assume a spherical cow" is useful in biophysics – it simplifies the analysis, but at some cost to its correspondence with reality.

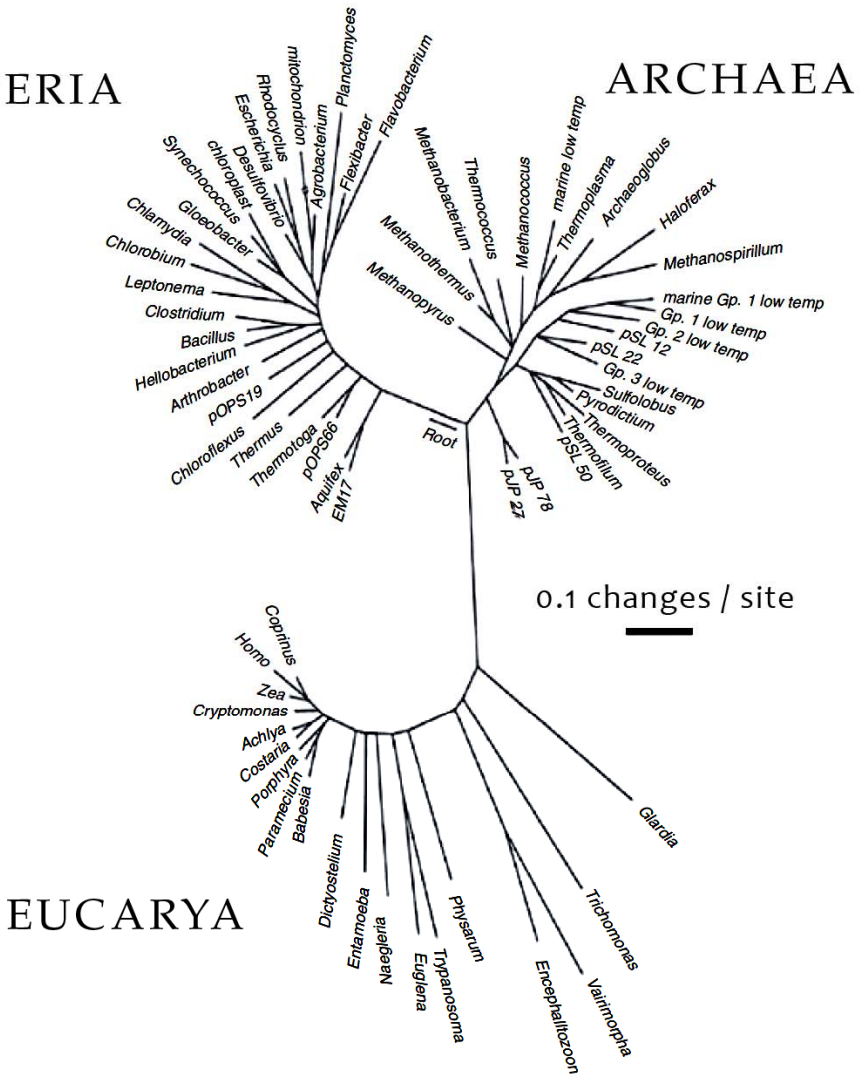
mental units in nature, is not objective "truth" – instead it is, at best, a useful approximation.

The idea that the whole may be understood by understanding all of its parts is the conceit of reductionism.

The idea that the whole may be understood by understanding all of its parts is the conceit of reductionism.

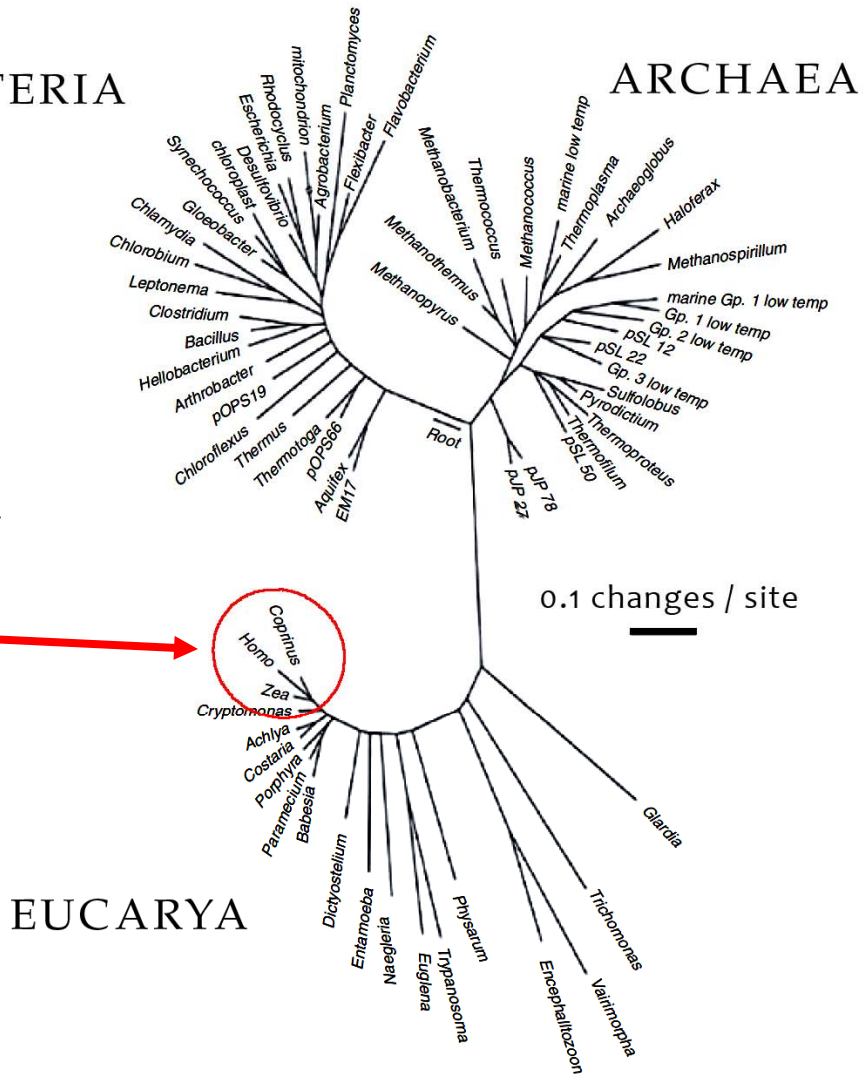
The idea that the whole may be understood by understanding a few (or only one) of its parts is simply wrong.

If the goal of biodiversity studies is to understand all of the diversity in the Earth's biodiversity…



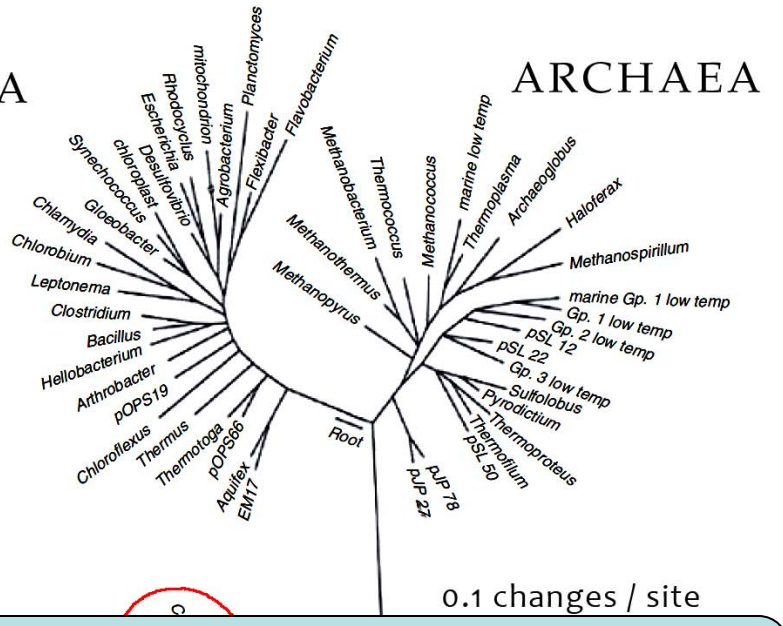
If the goal of biodiversity studies is to understand all of the diversity in the Earth's biodiversity…

Then the notion that we can accomplish that goal only by looking here is just plain wrong.



If the goal of biodiversity studies is to understand all of the diversity in the Earth's biodiversity…

Then the notion that we can accomplish that goal only by looking here



To maintain its relevancy, to deal with the non-negotiable aspects of nature, 21st century biodiversity studies MUST include a large and growing commitment to understanding microbial biodiversity.

Before Genomics:

Individual organisms are the fundamental units of biodiversity; their evolutionary history can be explained by arranging them into groups, with the groups composed into a single-rooted tree.

Before Genomics:

Individual organisms are the fundamental units of biodiversity; their evolutionary history can be explained by arranging them into groups, with the groups composed into a single-rooted tree.

After Genomics

No “completely objective” individuals, no one true tree.

END

(http://www.rj-robbins.com/slides/RJR-GBIC-2012.pdf)