

BIG DATA *BIOINFORMATICS*

or why medical doctors
need computers



Wojciech Makałowski
Institute of Bioinformatics
University of Münster, Germany



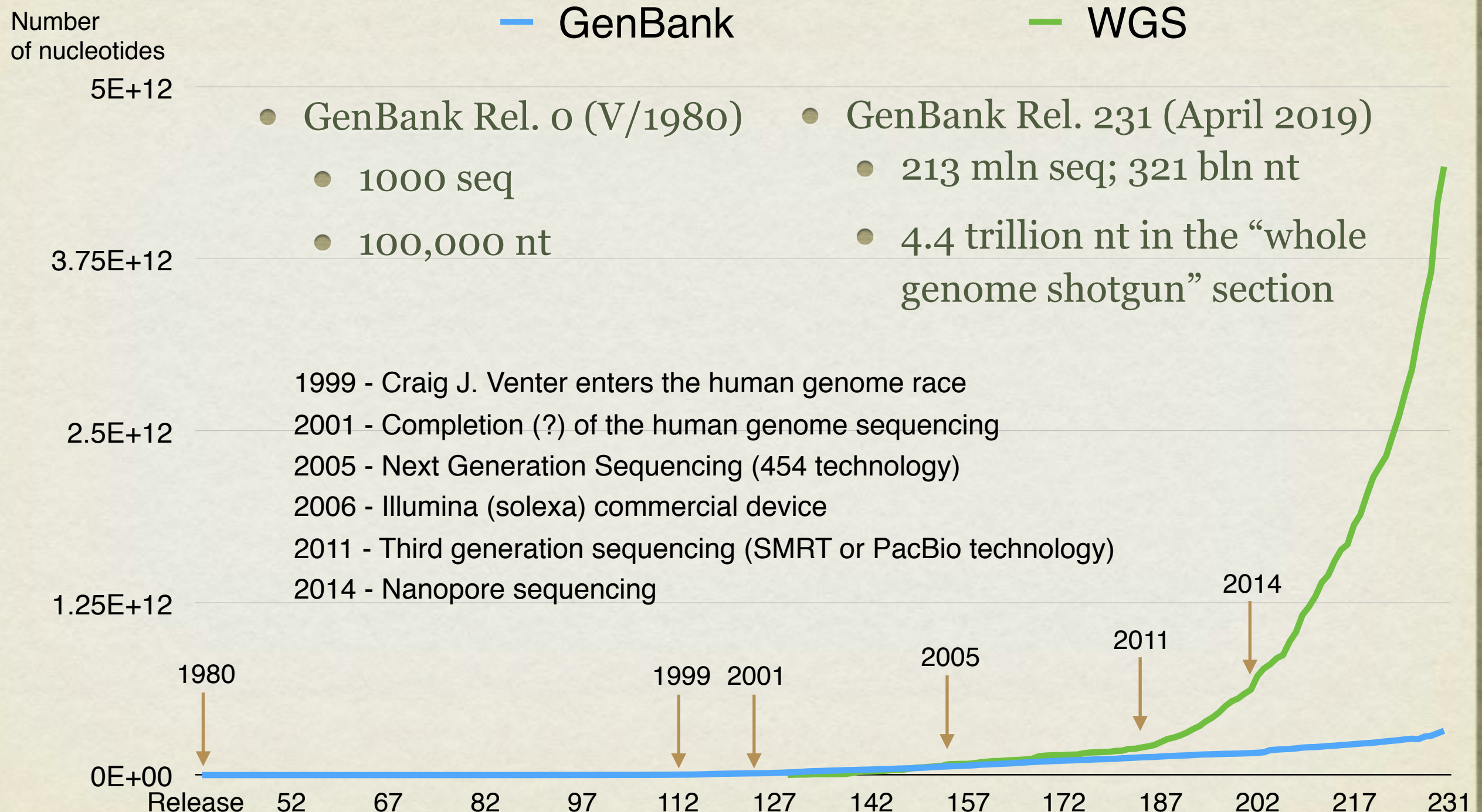
**It's sink or swim as a tidal
wave of data approaches**

Nature 399:517 10 June 1999

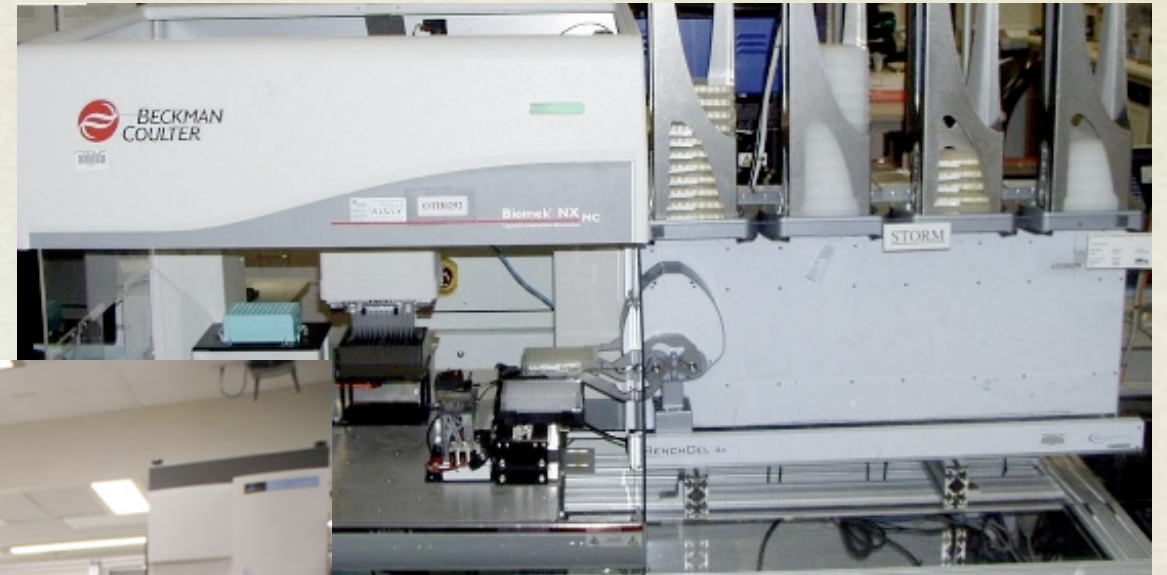
Unfortunately, it's not a tidal wave,
it's a tsunami!



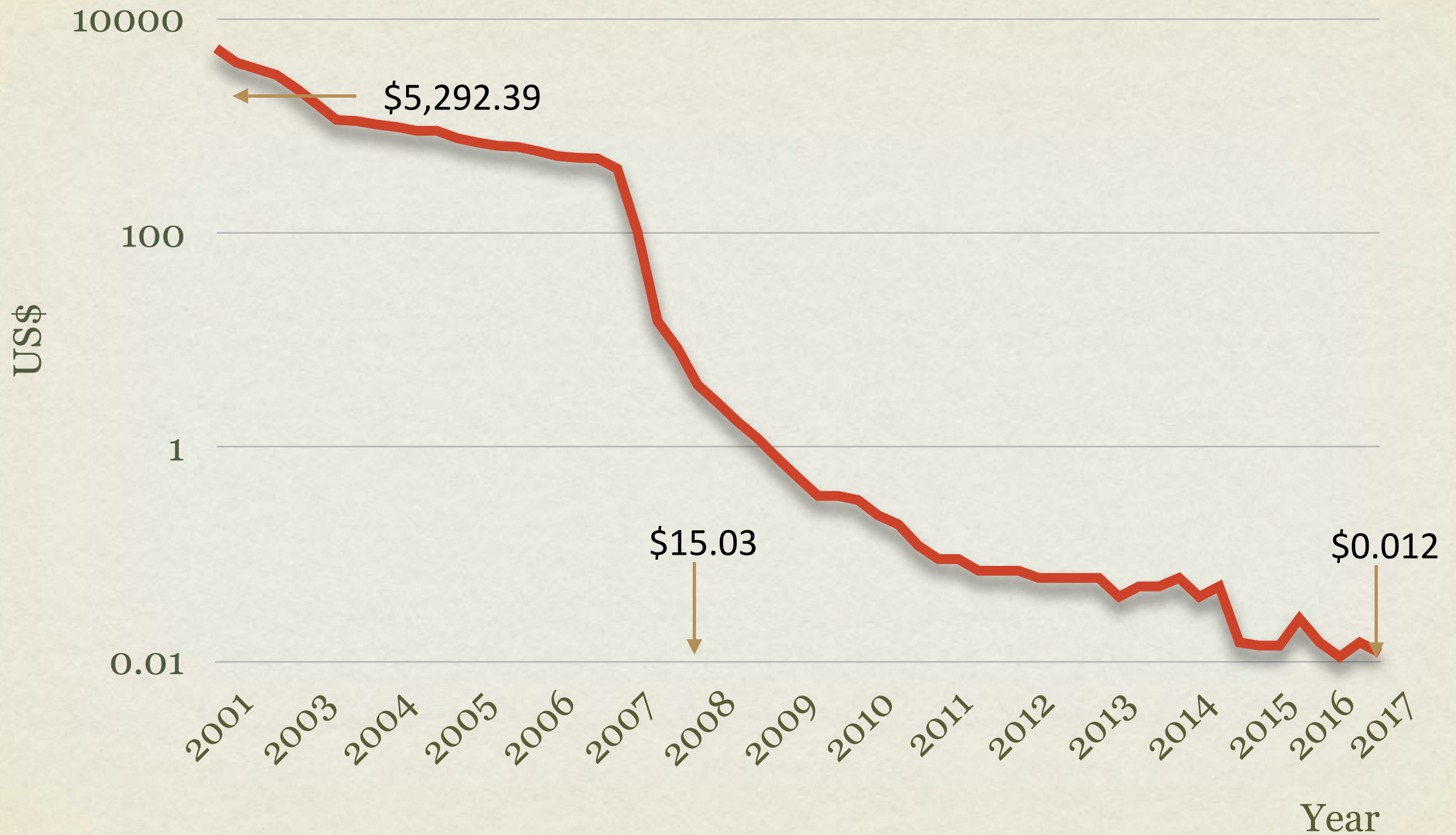
GROWTH OF BIOMEDICAL INFORMATION - GENBANK



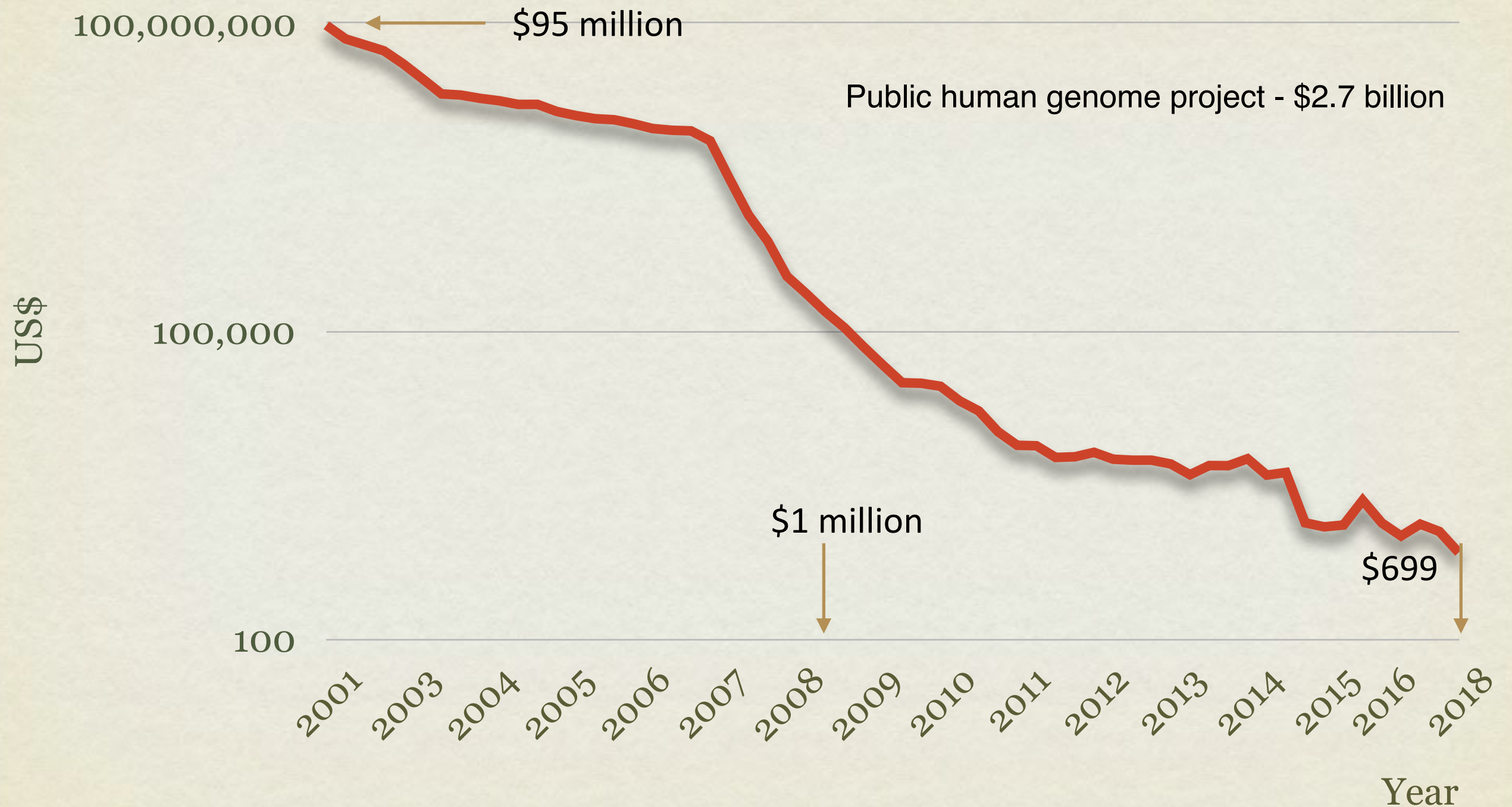
TECHNOLOGY MEETS BIOLOGY



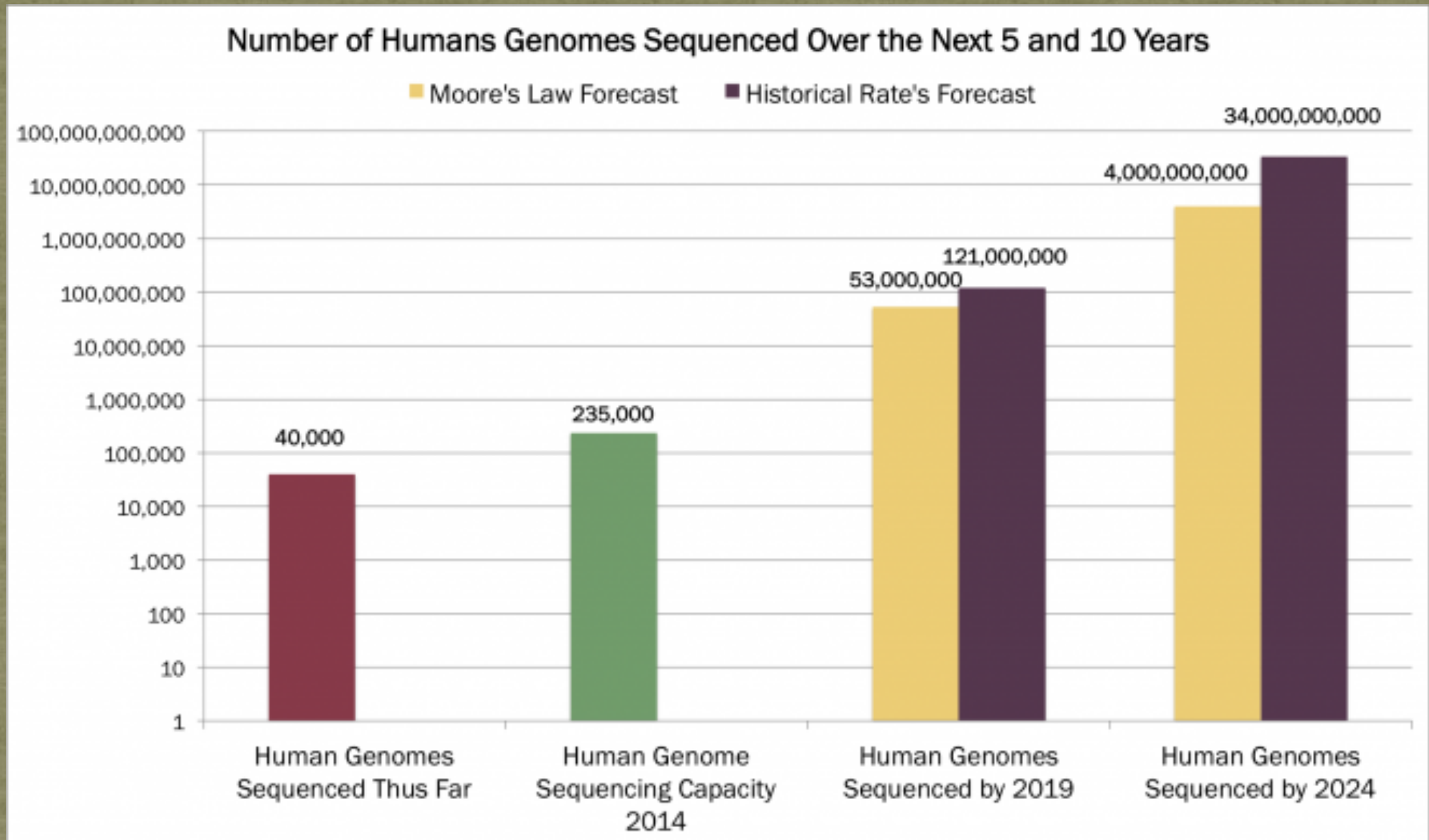
SEQUENCING COST PER MB



SEQUENCING COST PER GENOME



IMPROVING TECHNOLOGY



GETTING SEQUENCES

TGCATCGATCGTAGCTAGCTAGCGCATGCTAGCTAGCTAGCTAGCTACGATGCATCG
TGCATCGATCGATGCATGCTAGCTAGCTAGCTAGCATGCTAGCTAGCTAGCTATTGG
CGCTAGCTAGCATGCATGCATGCATCGATGCATCGATTATAAGCGCGATGACGTCAG
CGCGCGCATTATGCCGCGGCATGCTGCGCACACACAGTACTATAGCATTAGTAAAAA
GGCCGCGTATATTTTACACGATAGTGCGGGCGCGGGCGCGTAGCTAGTGCTAGCTAGTC
TCCGGTTACACAGGTAGCTAGCTAGCTGCTAGCTAGCTGCTGCGATGCATGCATTAGT
AGCTAGTGCTAGCTAGCTAGCATGCTGCTAGCATGCAGCATGCATCGGGCGCGGATGCT
GCTAGCGCTGCTAGCTAGCTAGCTAGCTAGCTAGGCGCTAATTATTTTGGGGGGTTA
AAAAAAAAAAATTCGCTGCTTATACCCCCCCCCCACATGATGATCGTTAGTAGCTACT
AGCTCTCATCGCGCGGGGGGATGCTTAGCGTGGTGTGTGTGTGTGGTGTGTGTGGTC
CTATAATTAGTGCATCGGCGCATCGATGGCTAGTCGATCGATCGATTTTATATATCT
AAAGACCCCATCTCTCTCTTTTCCCTTCTCTCGCTAGCGGGCGGTACGATTTACC
GGCCGCGTATATTTTACACGATAGTGCGGGCGCGGGCGCGTAGCTAGTGCTAGCTAGTC
AGCTCTCATCGCGCGGGGGGATGCTTAGCGTGGTGTGTGTGTGTGGTGTGTGTGGTC
TGCATCGATCGATGCATGCTAGCTAGCTAGCTAGCATGCTAGCTAGCTAGCTATTGG
CTATAATTAGTGCATCGGCGCATCGATGGCTAGTCGATCGATCGATTTTATATATCT
CGCTAGCTAGCATGCATGCATGCATCGATGCATCGATTATAAGCGCGATGACGTCAG
TCCGGTTACACAGGTAGCTAGCTAGCTGCTAGCTAGCTGCTGCGATGCATGCATTAGT

READING ≠ UNDERSTANDING

Carmina qui quondam studio
florente peregi, flebilis heu maestos
cogor inire modos.

Ecce mihi lacerae dictant scribenda
Camenae et ueris elegi fletibus ora
rigant.

READING $\neq$ UNDERSTANDING

We shall best understand the probable course of natural selection by taking the case of a country undergoing some physical change. If the country were open were open on its borders, new forms would certainly immigrate, and this also would bla, bla bla become extinct inhabitants.

Charles Darwin - *The Origin of Species*

READING ≠ UNDERSTANDING

We shall best understand the probable course of [redacted] by taking the case of a country undergoing some physical change. If the country were open on its borders, new forms would certainly immigrate, and this also would [redacted] become extinct inhabitants.

Charles Darwin - *The Origin of Species*

CHALLENGE: HOW FROM THIS...

TGCATCGATCGTAGCTAGCTAGCGCATGCTAGCTAGCTAGCTAGCTACGATGCATCG
TGCATCGATCGATGCATGCTAGCTAGCTAGCTAGCATGCTAGCTAGCTAGCTATTGG
CGCTAGCTAGCATGCATGCATGCATCGATGCATCGATTATAAGCGCGATGACGTCAG
CGCGCGCATTATGCCGCGGCATGCTGCGCACACACAGTACTATAGCATTAGTAAAAA
GGCCGCGTATATTTTACACGATAGTGCGGCGCGGCGCGGTAGCTAGTGCTAGCTAGTC
TCCGGTTACACAGGTAGCTAGCTAGCTGCTAGCTAGCTGCTGCATGCATGCATTAGT
AGCTAGTGCTAGCTAGCTAGCATGCTGCTAGCATGCAGCATGCATCGGGCGCGATGCT
GCTAGCGCTGCTAGCTAGCTAGCTAGCTAGCTAGGGCGCTAATTATTTTATTTTGGGGGGTTA
AAAAAAAAAAAAATTCGCTGCTTATACCCCCCCCCCACATGATGATCGTTAGTAGCTACT
AGCTCTCATCGCGCGGGGGGGATGCTTAGCGTGGTGTGTGTGTGTGGTGTGTGTGGTC
CTATAATTAGTGCATCGGCGCATCGATGGCTAGTCGATCGATCGATTTTATATATCT
AAAGACCCCATCTCTCTCTCTTTTCCCTTCTCTCGCTAGCGGGCGGTACGATTTACC
GGCCGCGTATATTTTACACGATAGTGCGGCGCGGCGCGGTAGCTAGTGCTAGCTAGTC
AGCTCTCATCGCGCGGGGGGGATGCTTAGCGTGGTGTGTGTGTGTGGTGTGTGTGGTC
TGCATCGATCGATGCATGCTAGCTAGCTAGCTAGCATGCTAGCTAGCTAGCTATTGG
CTATAATTAGTGCATCGGCGCATCGATGGCTAGTCGATCGATCGATTTTATATATCT
CGCTAGCTAGCATGCATGCATGCATCGATGCATCGATTATAAGCGCGATGACGTCAG
TCCGGTTACACAGGTAGCTAGCTAGCTGCTAGCTAGCTGCTGCATGCATGCATTAGT

Infer this



HOW TO SOLVE THE PROBLEM - A HUMAN OR A COMPUTER?



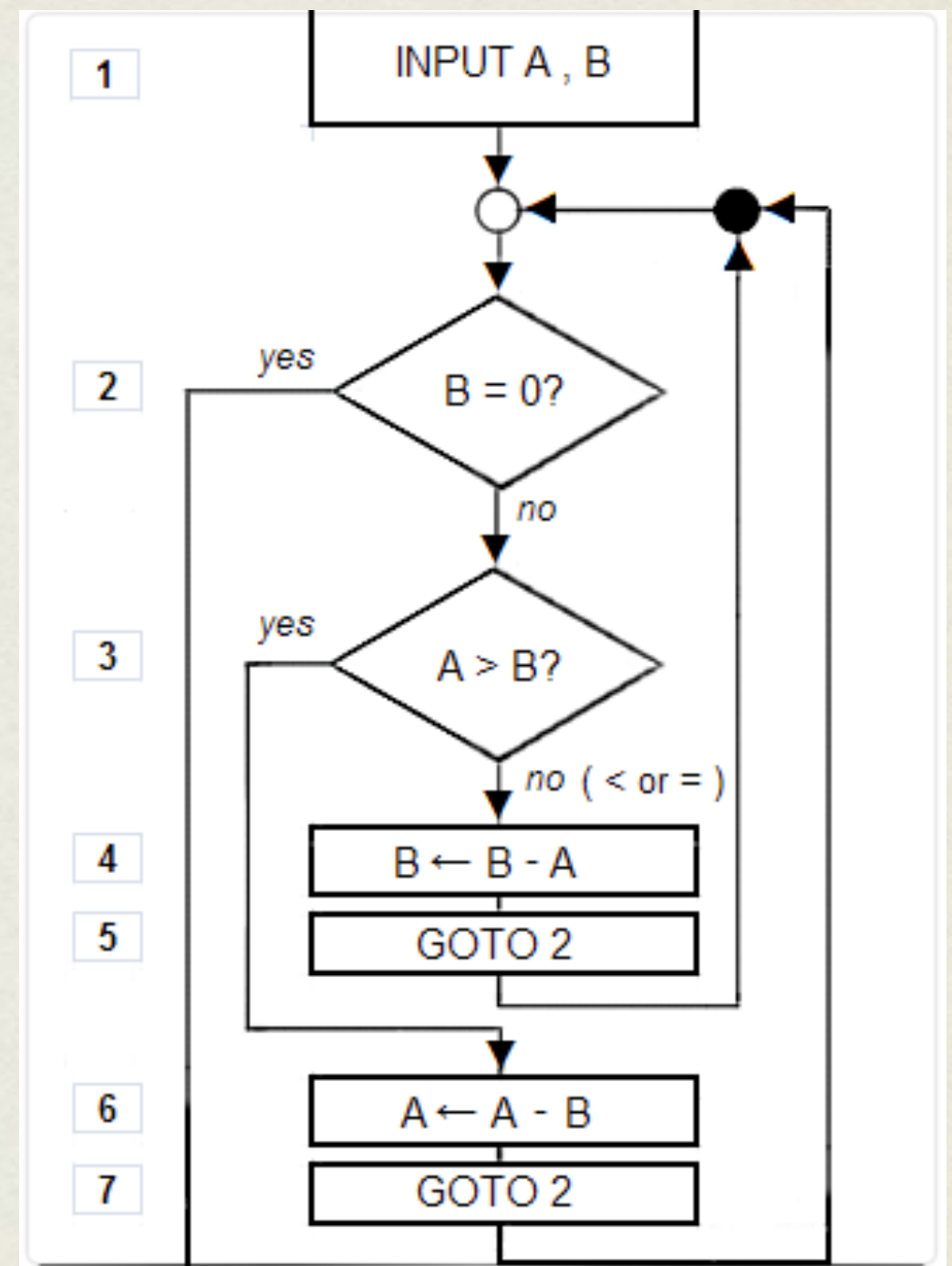
- very smart
- slow
- error prone
- doesn't like repetitive tasks

- not so smart (stupid)
- extremely fast
- very accurate
- doesn't understand human languages;
needs instruction provided in a special way



ALGORITHM

A step-by-step problem-solving procedure, especially an established, recursive computational procedure for solving a problem in a finite number of steps.

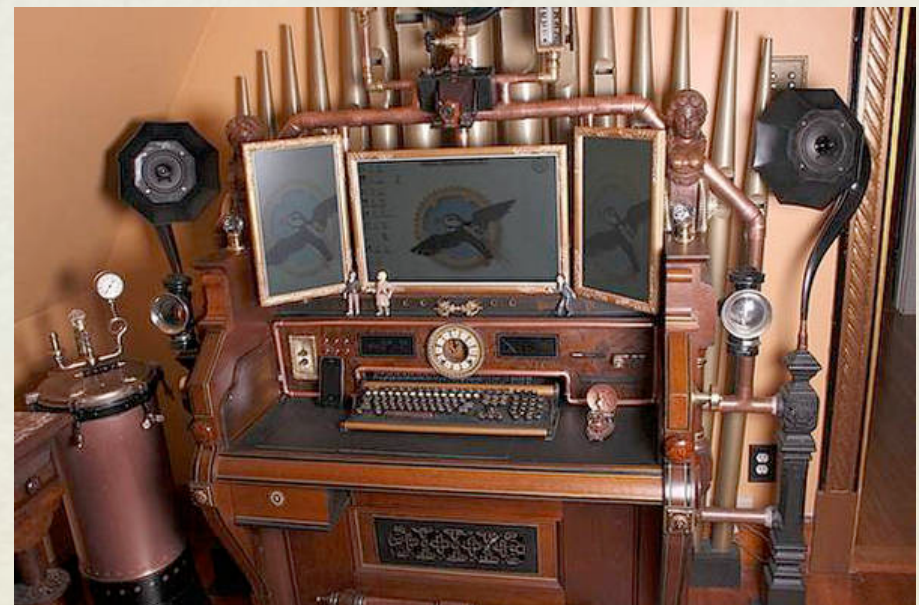


EXAMPLE TASK: PUT SHOES ON!



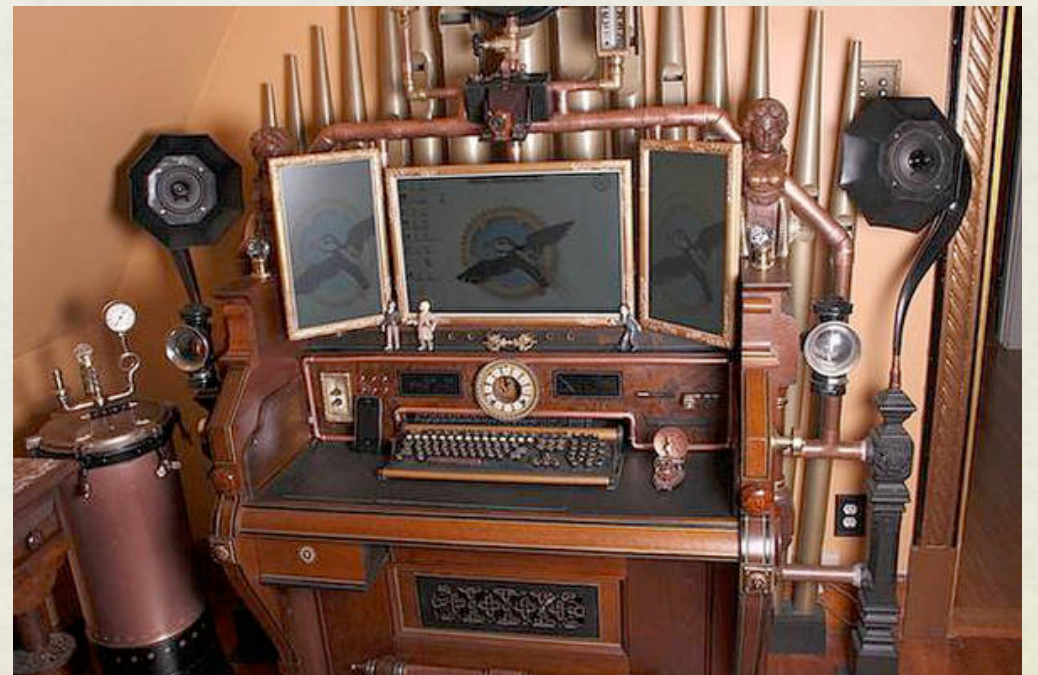
A human just understands an order
and often executes it automatically
even without thinking

A computer needs detailed
instruction (an algorithm)



PUT SHOES ON! INSTRUCTION FOR A COMPUTER

1. Find two the same shoes
2. Check if you have left and right shoe
3. Check if they are of the same size
4. Check if this is the right size
5. Put the left shoe on
6. Put the right shoe on
7. Tie the laces



THE ORIGIN OF THE FIELD



Paulien Hogeweg coined the term *bioinformatica* to define “the study of informatic processes in biotic systems”.

Hesper B, Hogeweg P (1970) Bioinformatica: een werkconcept. Kameleon 1(6): 28–29. (In Dutch.) Leiden: Leidse Biologen Club.

... but its origin can be tracked back many decades earlier.



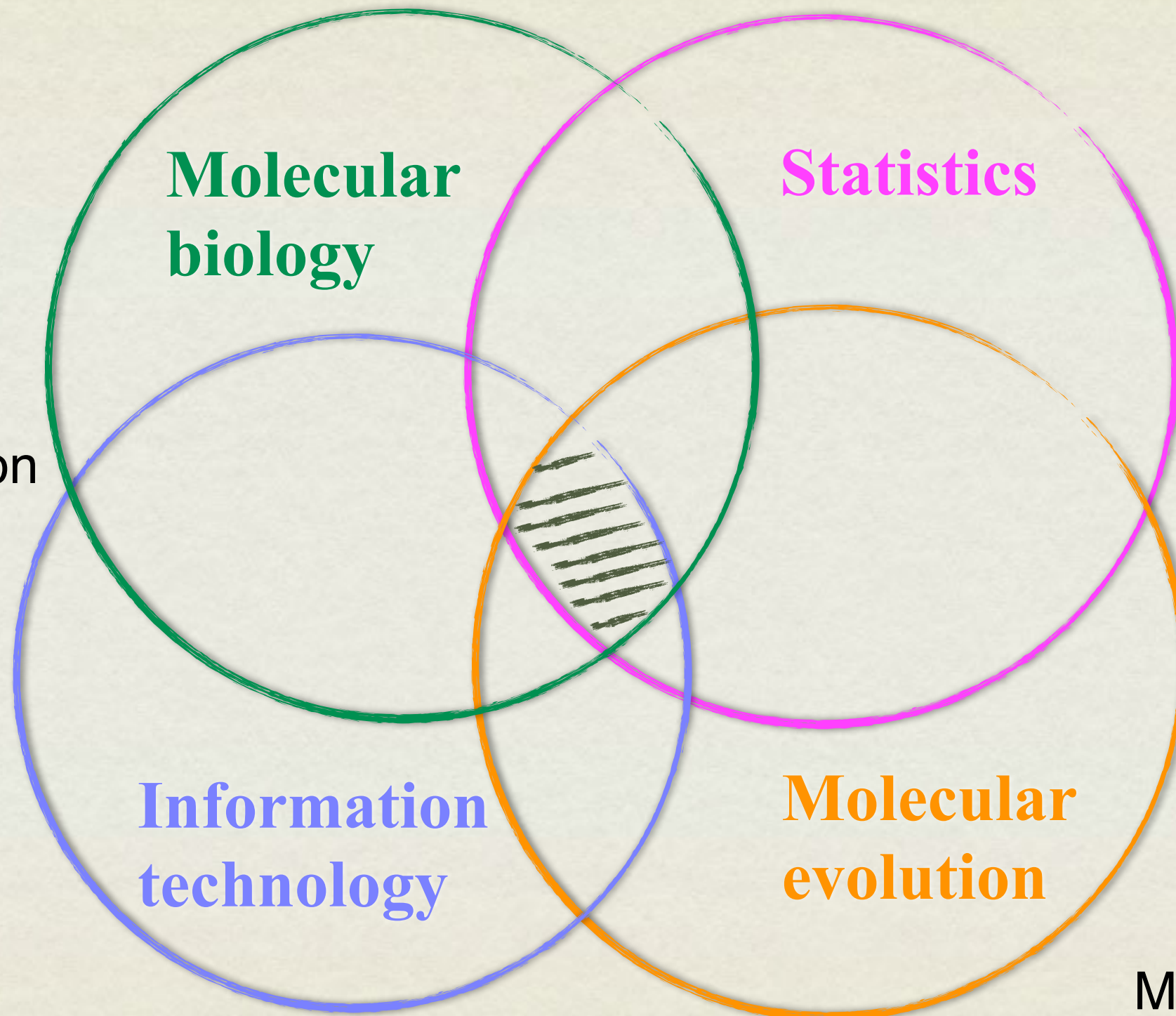
BIOINFORMATICS EMERGED AS AN INTERSECTION BETWEEN DIFFERENT DISCIPLINES



James Watson



Alan Turing



Thomas Bayes



Motoo Kimura

BIOINFORMATICS - DEFINITION

- Research, development, or application of computational tools and approaches for expanding the use of biological data, including those to acquire, store, organize, archive, analyze, or visualize such data.
- Its goal is to enable biological discovery based on existing information or in other words transform biological data into information and eventually into knowledge.

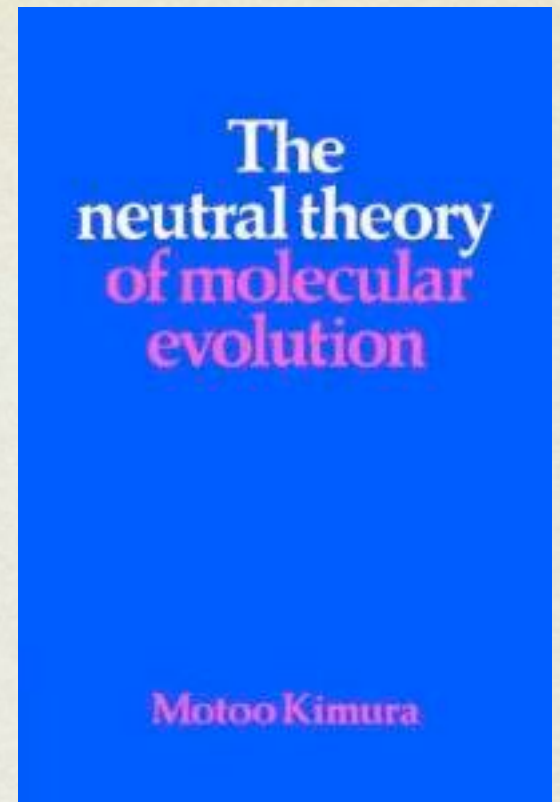
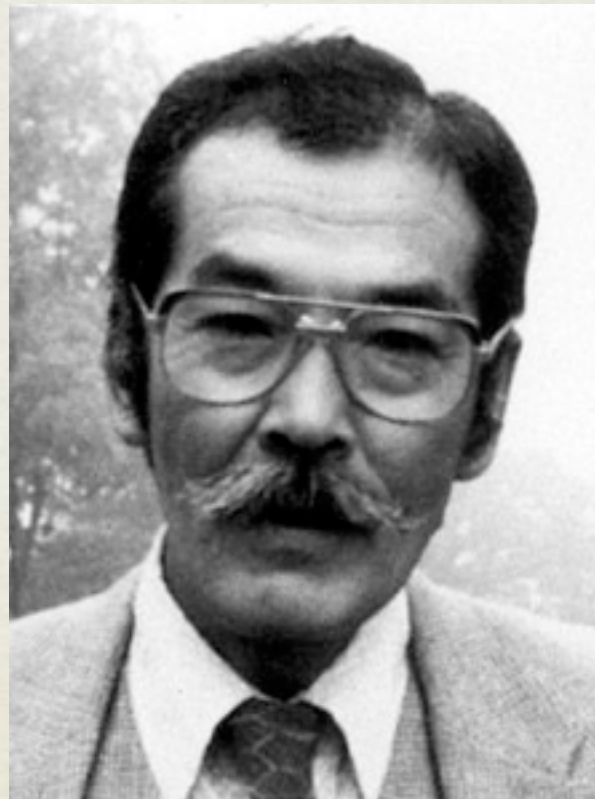
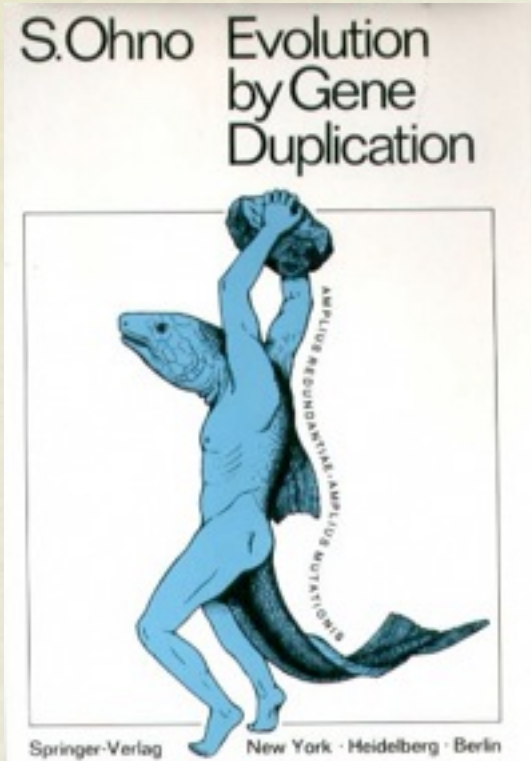


ROLE OF BIOINFORMATICS IN MODERN LIFE SCIENCES

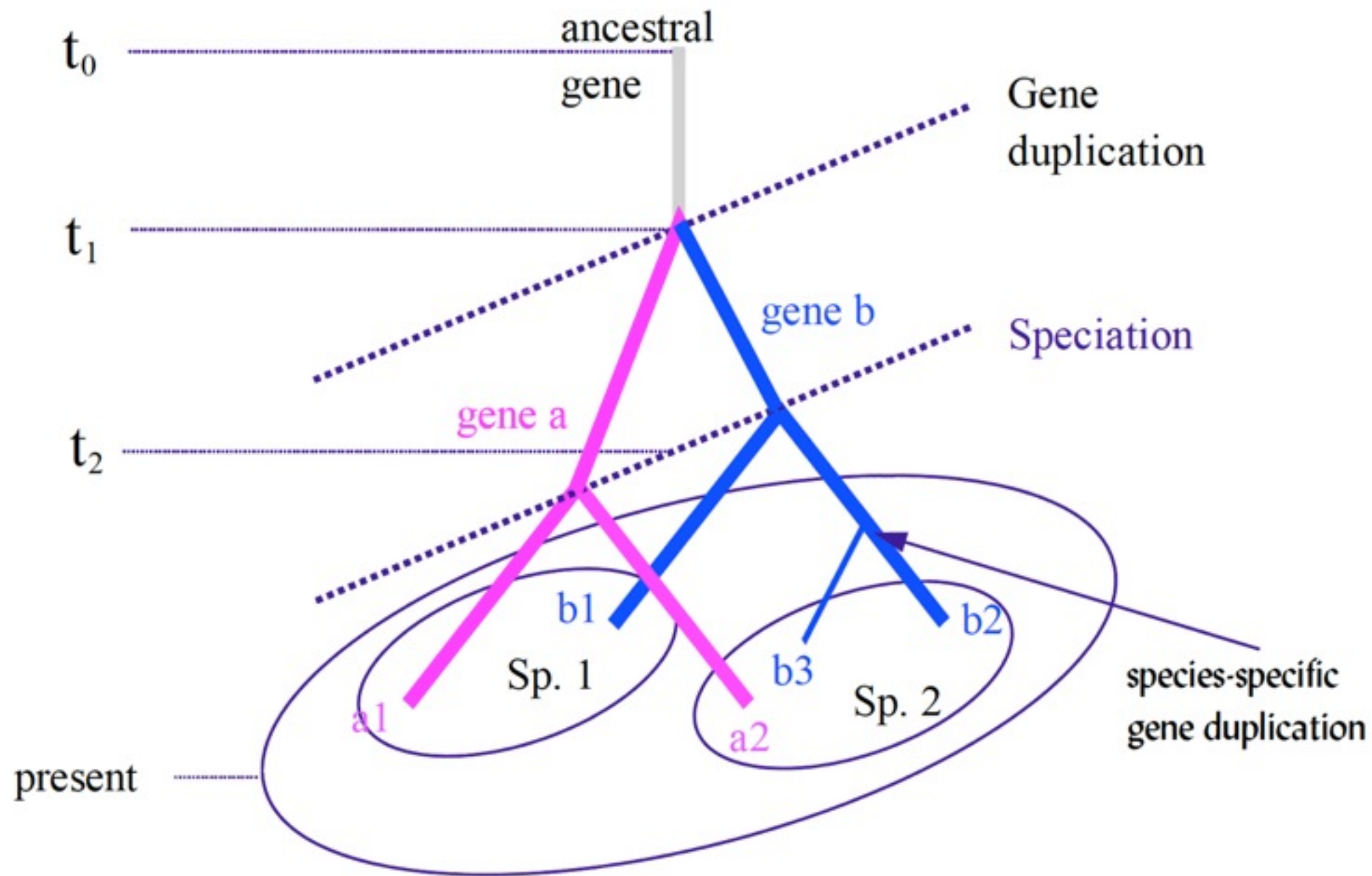
- molecular biology
- molecular evolution
- genomics
- system biology
- protein engineering
- drug design
- human genetics
- personalized medicine



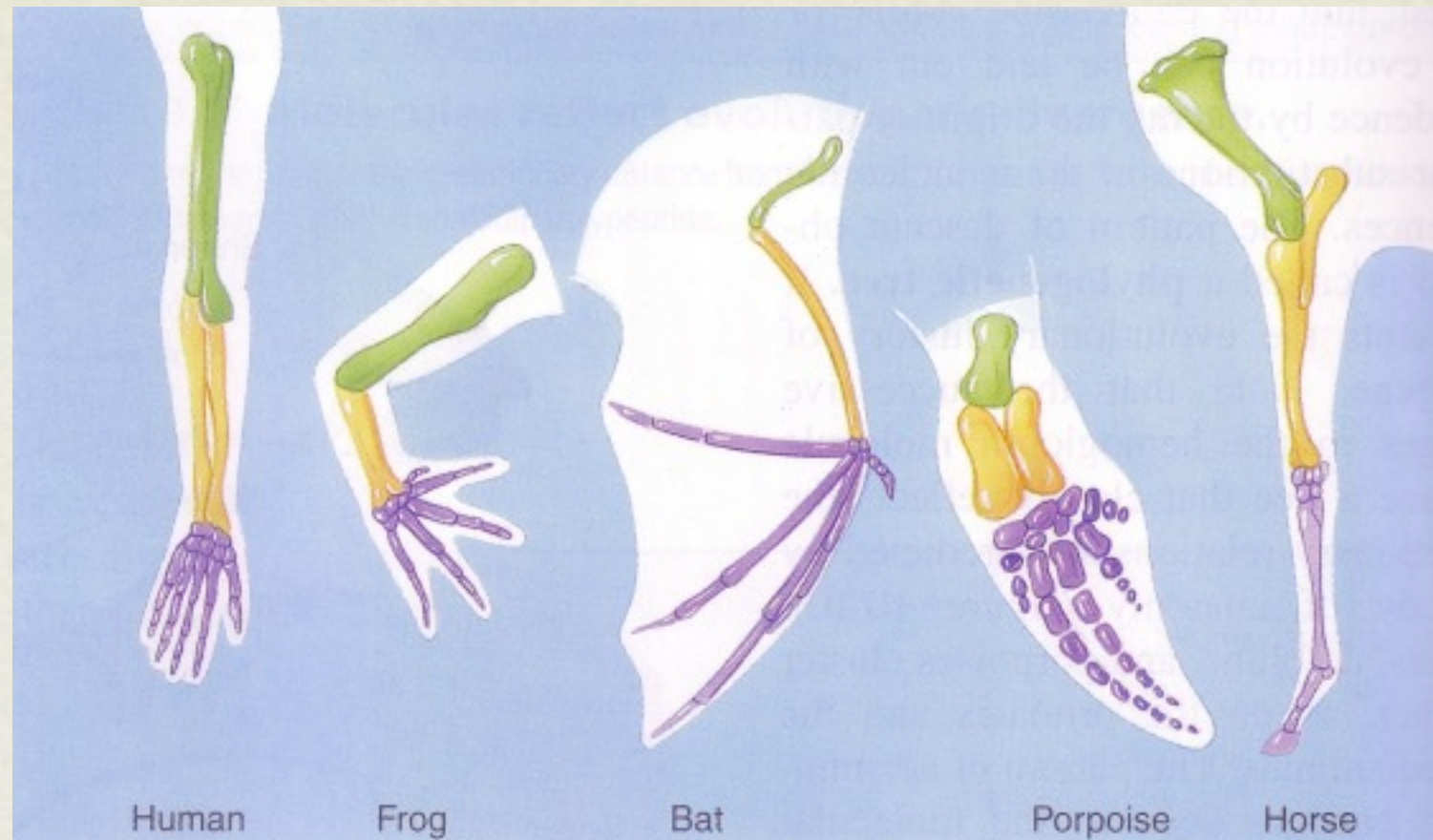
EVOLUTIONARY BASIS OF BIOINFORMATICS



EVOLUTIONARY BASIS OF BIOINFORMATICS



HOMOLOGS



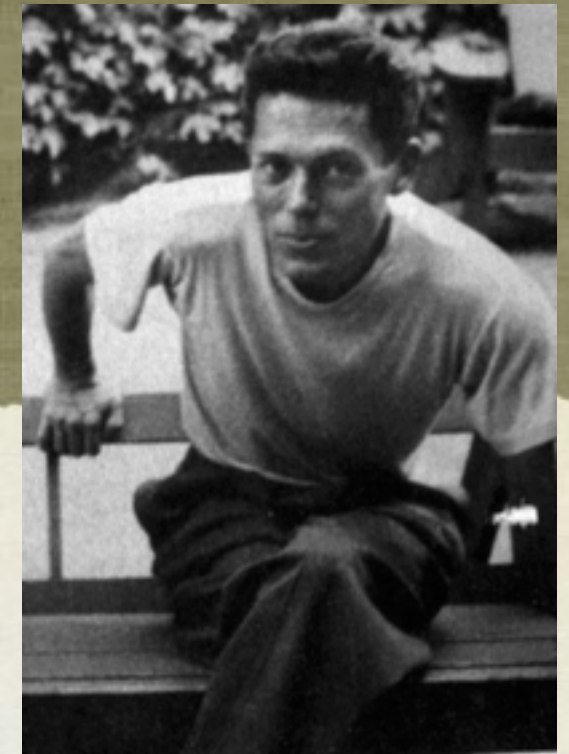
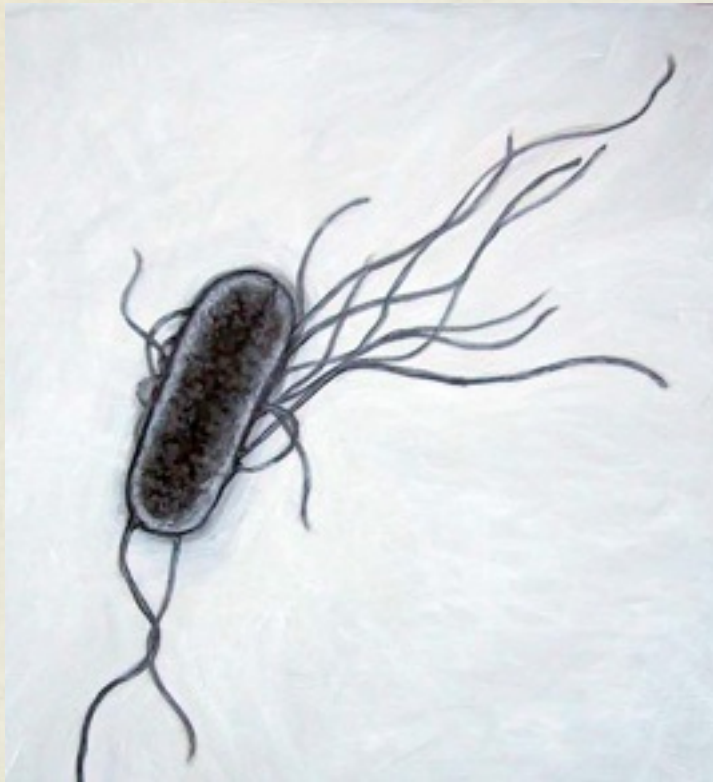
Two anatomical structures or behavioral traits within different organisms which originated from a structure or trait of their common ancestral organism. The structures or traits in their current forms may not necessarily perform the same functions in each organism, nor perform the functions it did in the common ancestor. An example: the wing of a bat, the fin of a whale and the arm of a man are homologous structures.

HOMOLOGS AT THE MOLECULAR LEVEL

cow	ATG---ACTAACATTTCGAAAGTCCCAACCCACTAATAAAAAATTGTAAAC
sheep	ATG---ATCAACATCCGAAAAACCCACCCACTAATAAAAAATTGTAAAC
goat	ATG---ACCAACATCCGAAAGACCCACCCATTATAATAAAAAATTGTAAAC
horse	ATG---ACAAACATCCGGAAATCTCACCCACTAATTAAAATCATCAAT
donkey	ATG---ACAAACATCCGAAATCCCAACCCGCTAATTAAAATCATCAAT
ostrich	ATGGCCCCAACATTTCGAAATCGCACCCCTTGCTCAAAATTATCAAC
emu	ATGGCCCCTAACATCCGAAATCCCAACCTCTACTCAAAATCATCAAC
turkey	ATGGCACCCAATATCCGAAATCACACCCCTATTAAAAACAATCAAC

Two sequences that share common ancestry. Significant sequence similarity usually suggests homology, however sequence similarity may occur also by chance and some homologous sequences may diverge beyond detectable similarity.

COMPARATIVE GENOMICS

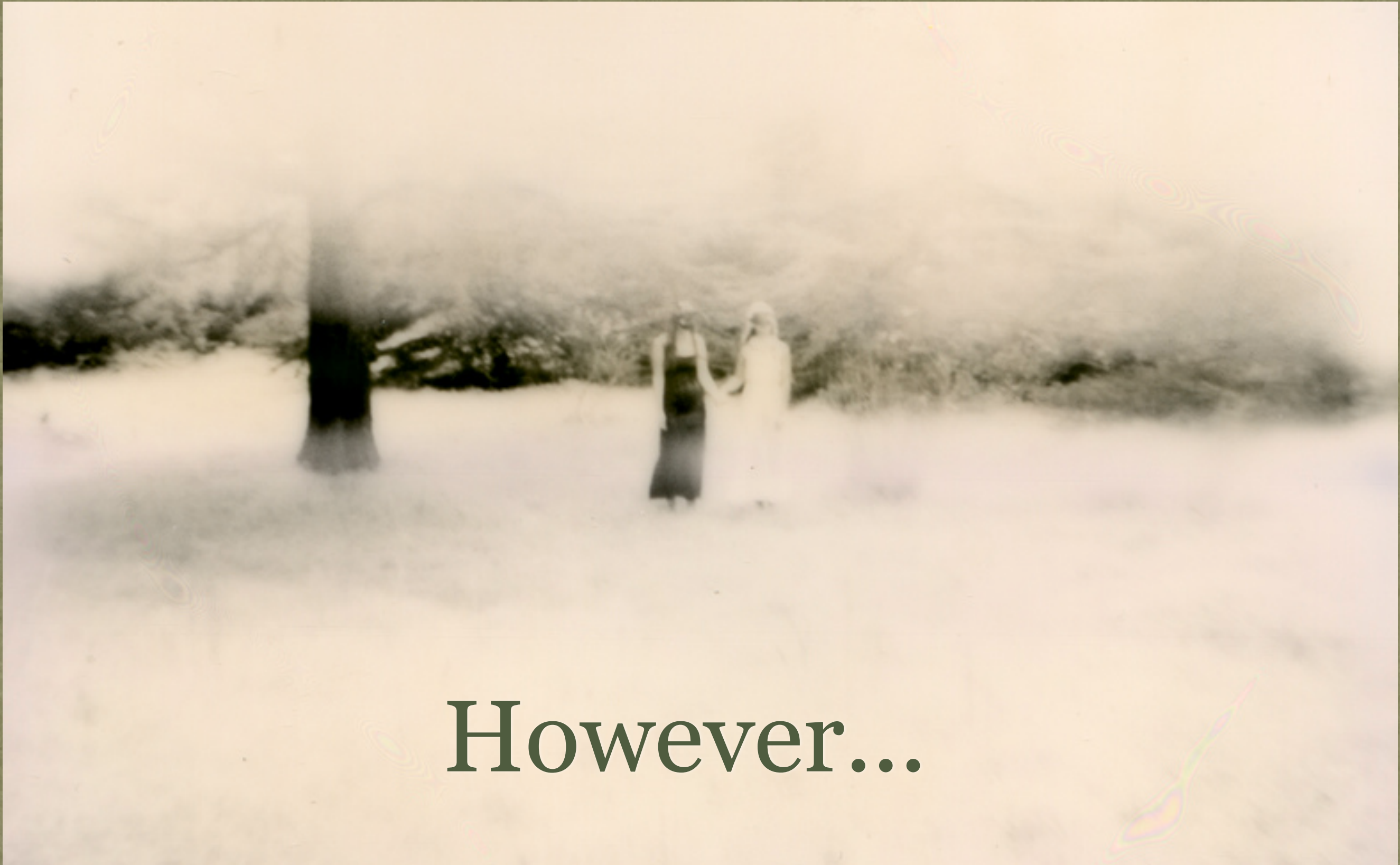


What is true for *E. coli* is
also true for elephant.

J. Monod, c. 1961

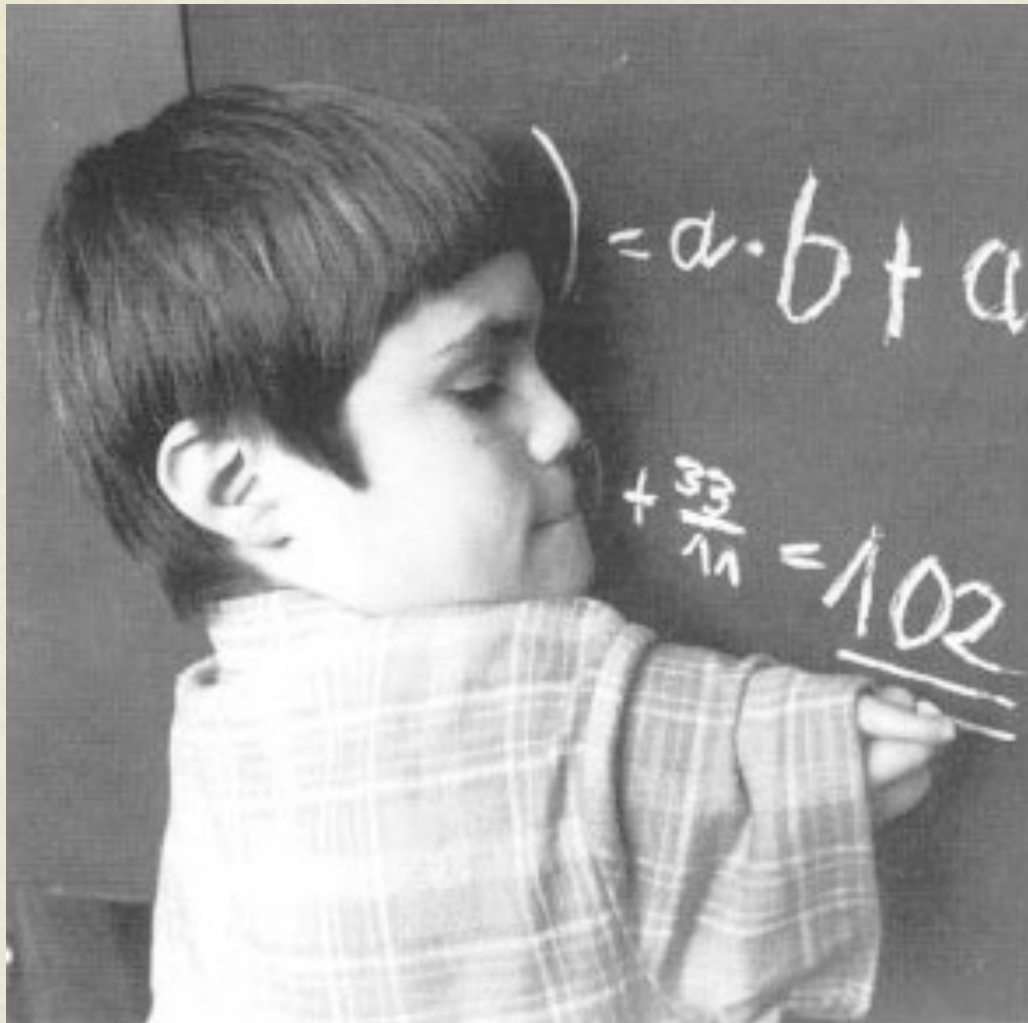


COMPARATIVE GENOMICS



However...

COMPARATIVE GENOMICS



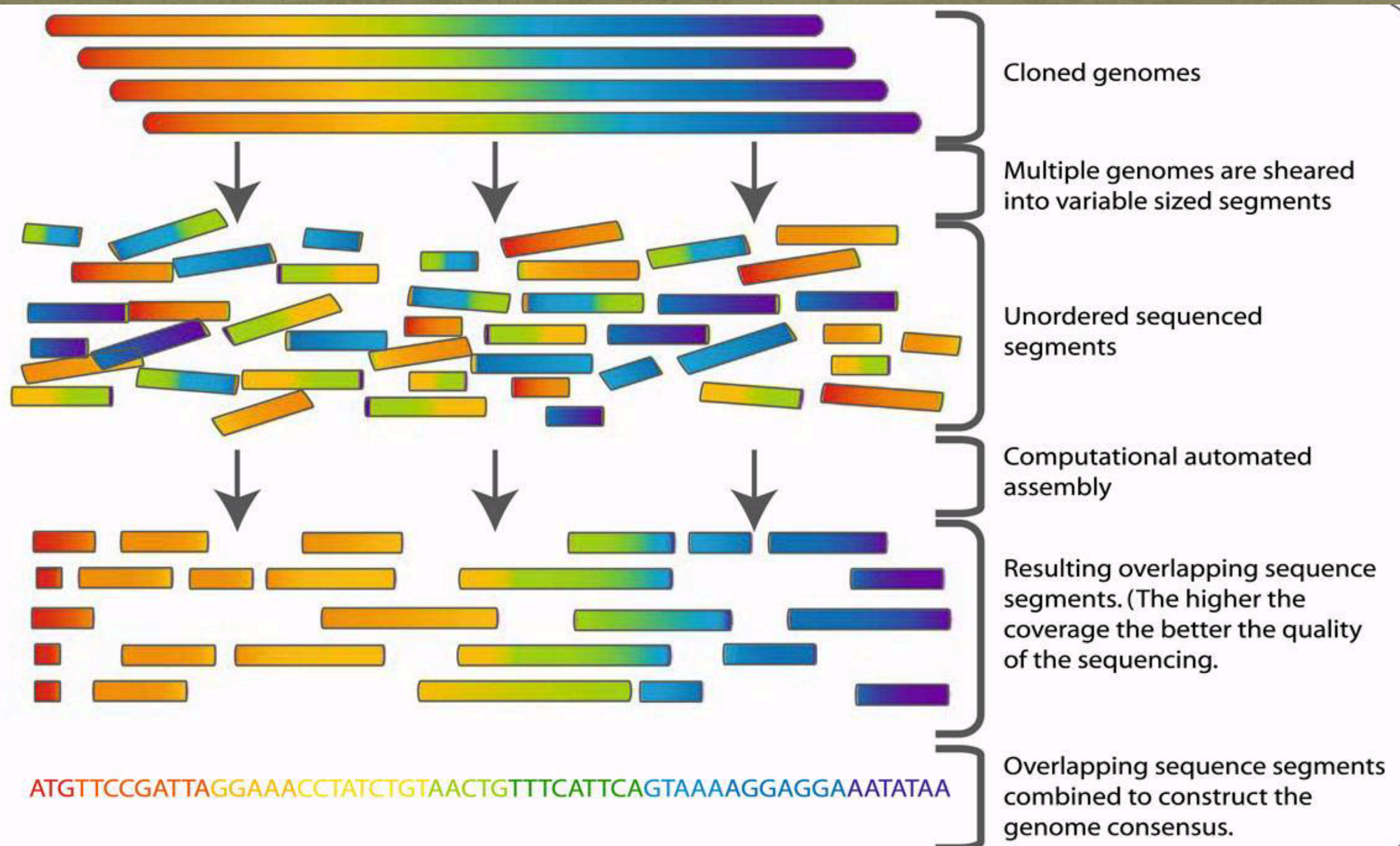
15 000 victims of thalidomide

What is true for mouse is not necessarily true for human...

Nucleotide Sequence Assembly



NUCLEOTIDE SEQUENCE ASSEMBLY



A group of seven children, four girls and three boys, are posing on a blue-tiled pier. They are all smiling and making various hand gestures. The background features a calm blue sea, distant mountains, and a cloudy sky. The children are dressed in casual summer clothing like t-shirts and shorts. The text 'Similarity Search' is overlaid on the right side of the image in a large, bold, red serif font.

Similarity Search

Basic Local Alignment Search Tool

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance. [Learn more](#)

NEWS

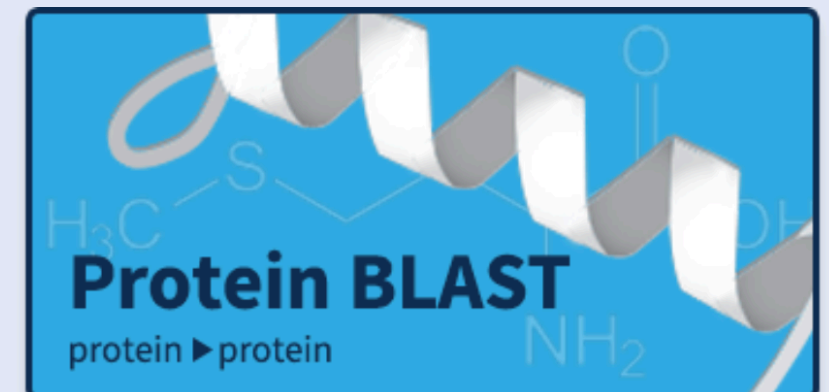
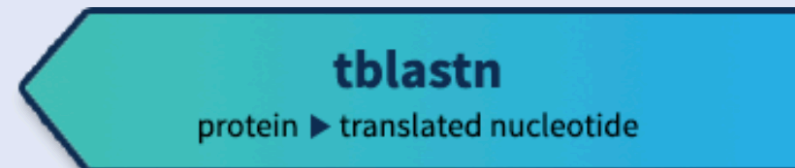
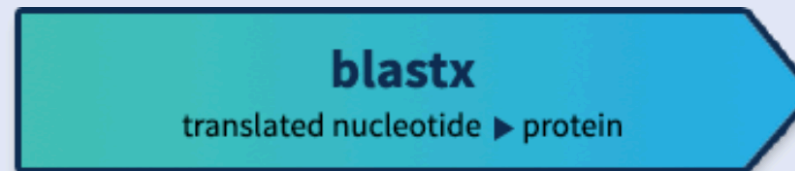
BLAST+ 2.9.0 is here!

The latest version has enhanced support for the new database format.

Tue, 02 Apr 2019 17:00:00 EST

[More BLAST news...](#)

Web BLAST



BLAST Genomes

Enter organism common name, scientific name, or tax id

Search

[Human](#)

[Mouse](#)

[Rat](#)

[Microbes](#)

FUN WITH BLAST

Job title: Wojciech Makalowski

RID [BXXS14F8014](#) (Expires on 04-24 19:29 pm)

Query ID lcllQuery_303143

Description Wojciech Makalowski

Molecule type amino acid

Query Length 18

Database Name nr

Description All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+PRF
excluding environmental samples from WGS projects

Program BLASTP 2.9.0+ [▶ Citation](#)

Other reports: [▶ Search Summary](#) [\[Taxonomy reports\]](#) [\[Distance tree of results\]](#) [\[Multiple alignment\]](#) [\[MSA viewer\]](#)

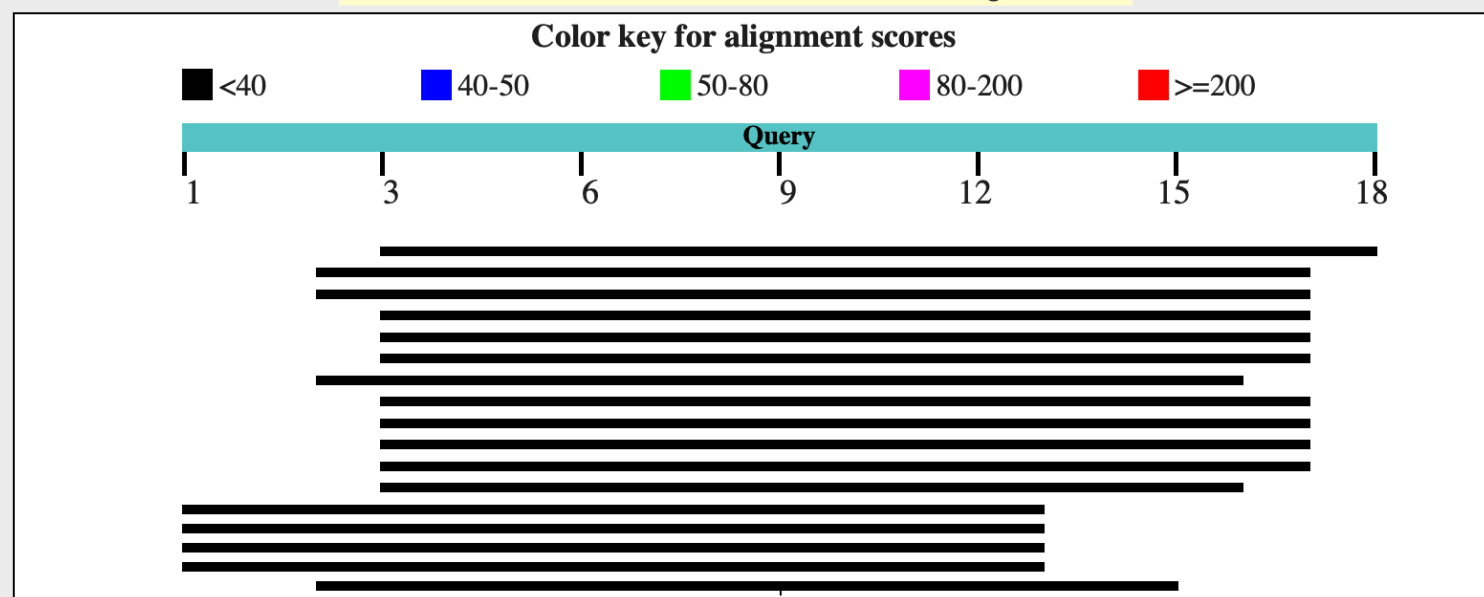
[Graphic Summary](#)

 [Show Conserved Domains](#)

No putative conserved domains have been detected

Distribution of the top 104 Blast Hits on 100 subject sequences 

Mouse over to see the title, click to show alignments



FUN WITH BLAST

 Download [GenPept](#) [Graphics](#)

HNH endonuclease [Klebsiella michiganensis]

Sequence ID: [WP_122117762.1](#) Length: 122 Number of Matches: 1

[▶ See 1 more title\(s\)](#)

Range 1: 81 to 91 [GenPept](#) [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Positives	Gaps
33.3 bits(71)	7.5	8/11(73%)	8/11(72%)	0/11(0%)

Query	1	WOJCIECHMAK	11
		W CIECH AK	
Sbjct	81	WTLCIECHSAK	91

 Download [GenPept](#) [Graphics](#)

hypothetical protein [Desulforhopalus singaporensis]

Sequence ID: [WP_092225757.1](#) Length: 351 Number of Matches: 1

[▶ See 1 more title\(s\)](#)

Range 1: 267 to 277 [GenPept](#) [Graphics](#)

[▼ Next Match](#) [▲ Previous Match](#)

Score	Expect	Identities	Positives	Gaps
32.0 bits(68)	21	10/14(71%)	10/14(71%)	3/14(21%)

Query	4	CIECHMAKALOWSK	17
		CIECHM KA SK	
Sbjct	267	CIECHMPKA---SK	277

Gene Prediction

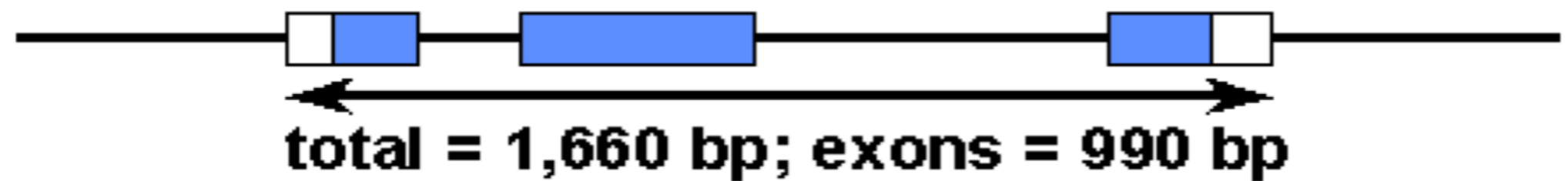


(exon-intron-exon)_n structure of various genes

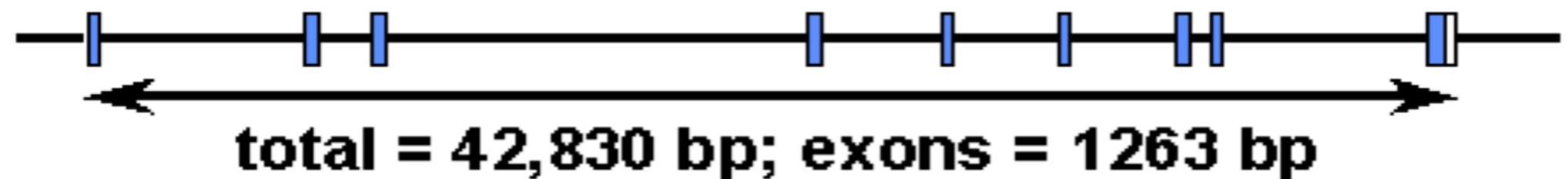
histone



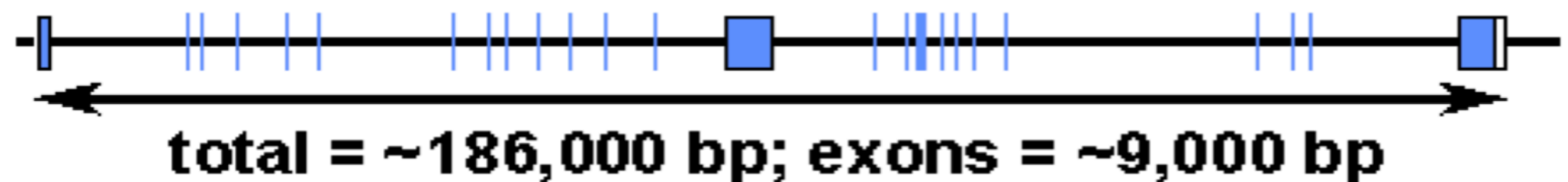
β-globin



**HGPRT
(HPR T)**



factor VIII



GENE FINDING METHODS

coding/non-
coding sequence
discrimination

homology based

model based

based on
similarity to
known genes

multi-genome
approach

sequence
composition

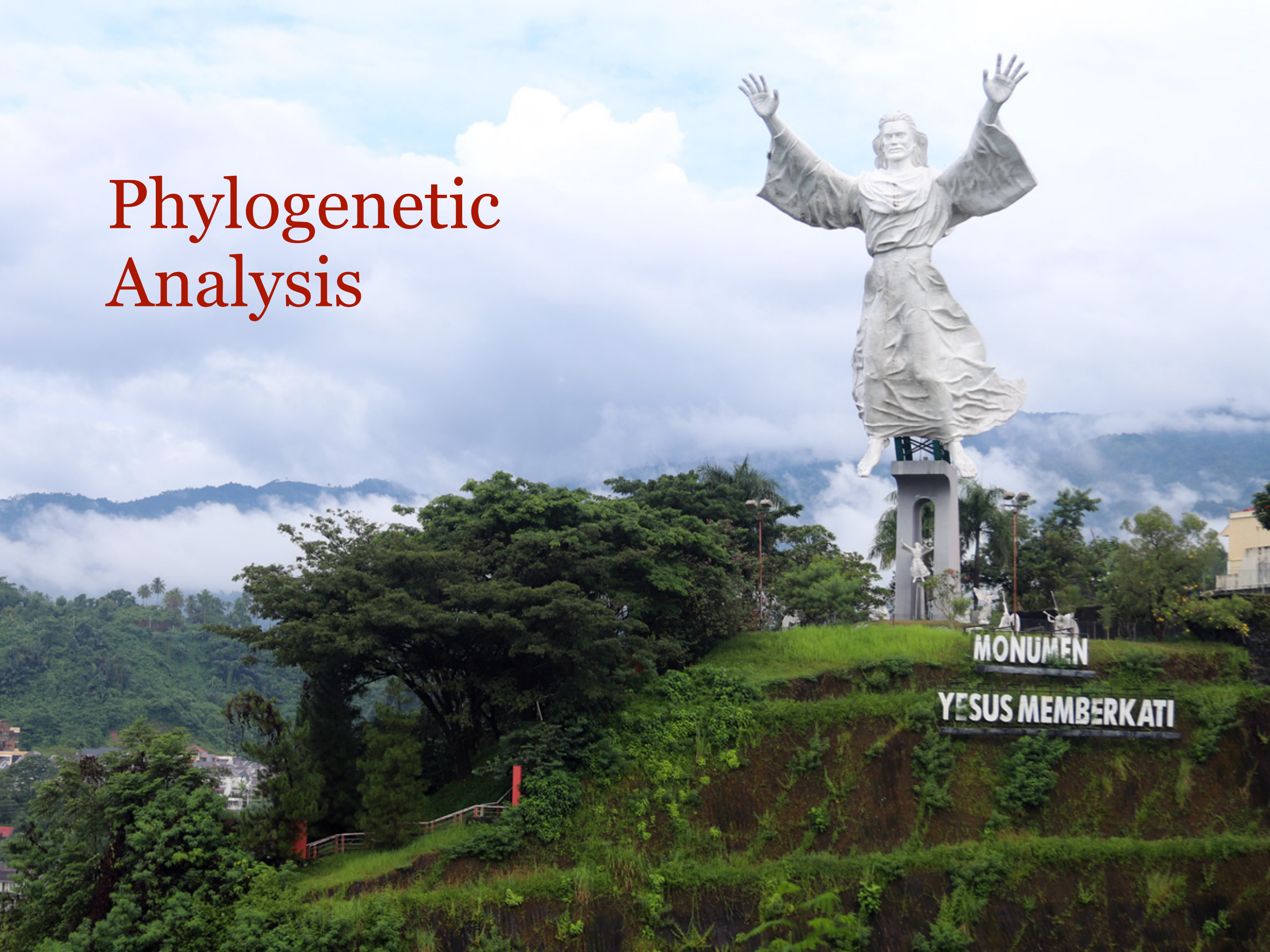
signals

transcripts

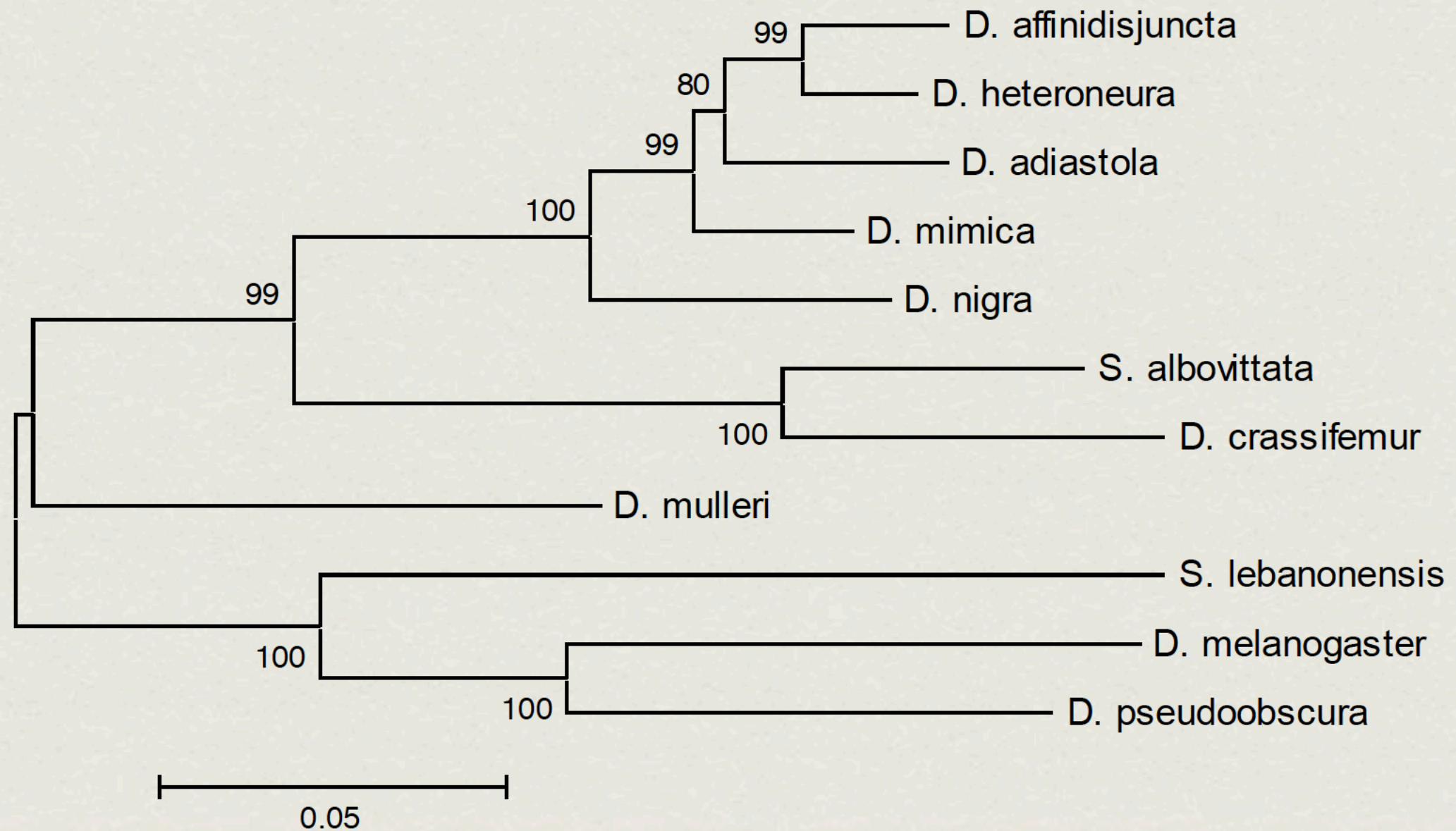
proteins

conservation

Phylogenetic Analysis

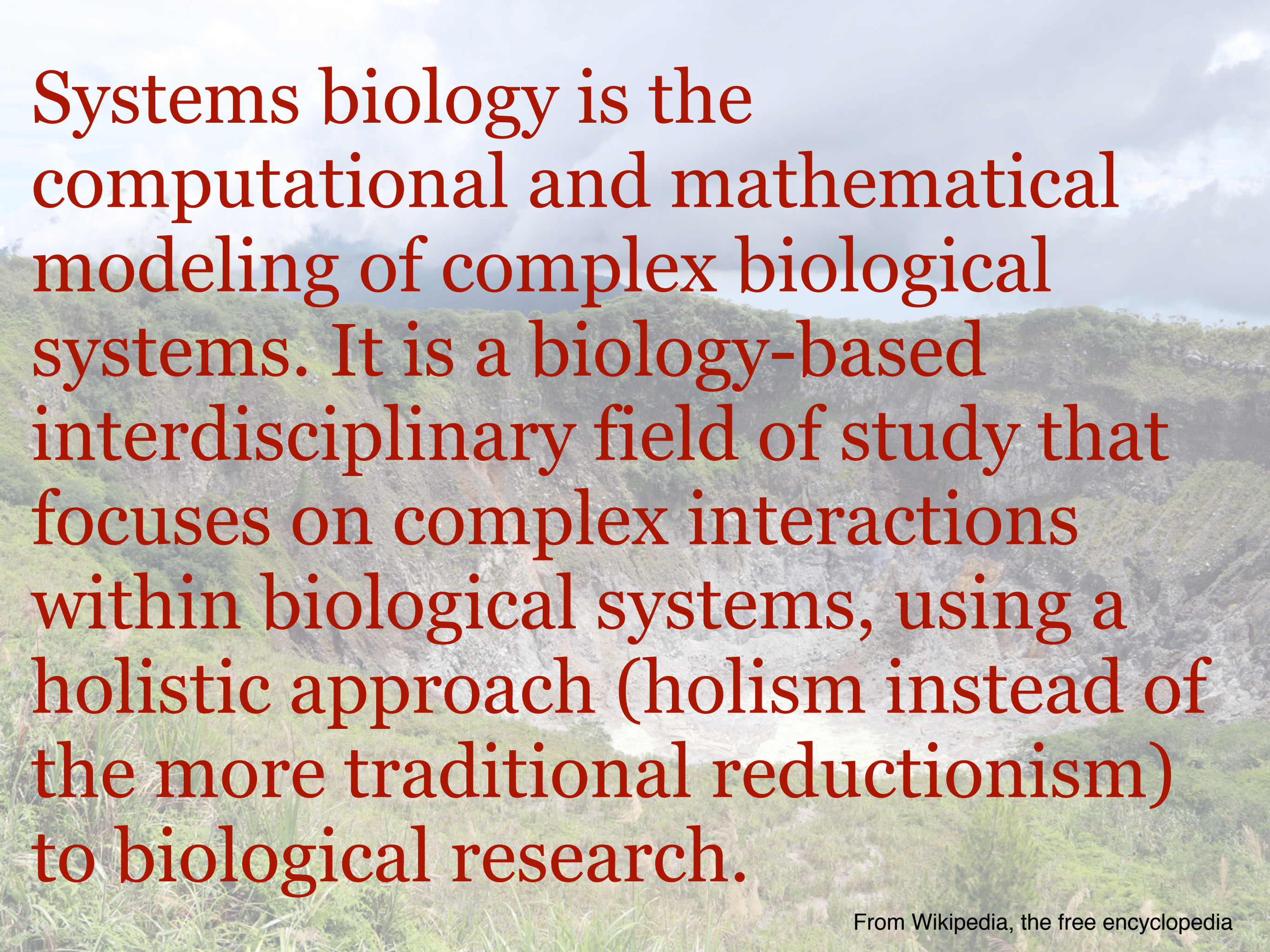


Phylogenetic Analysis



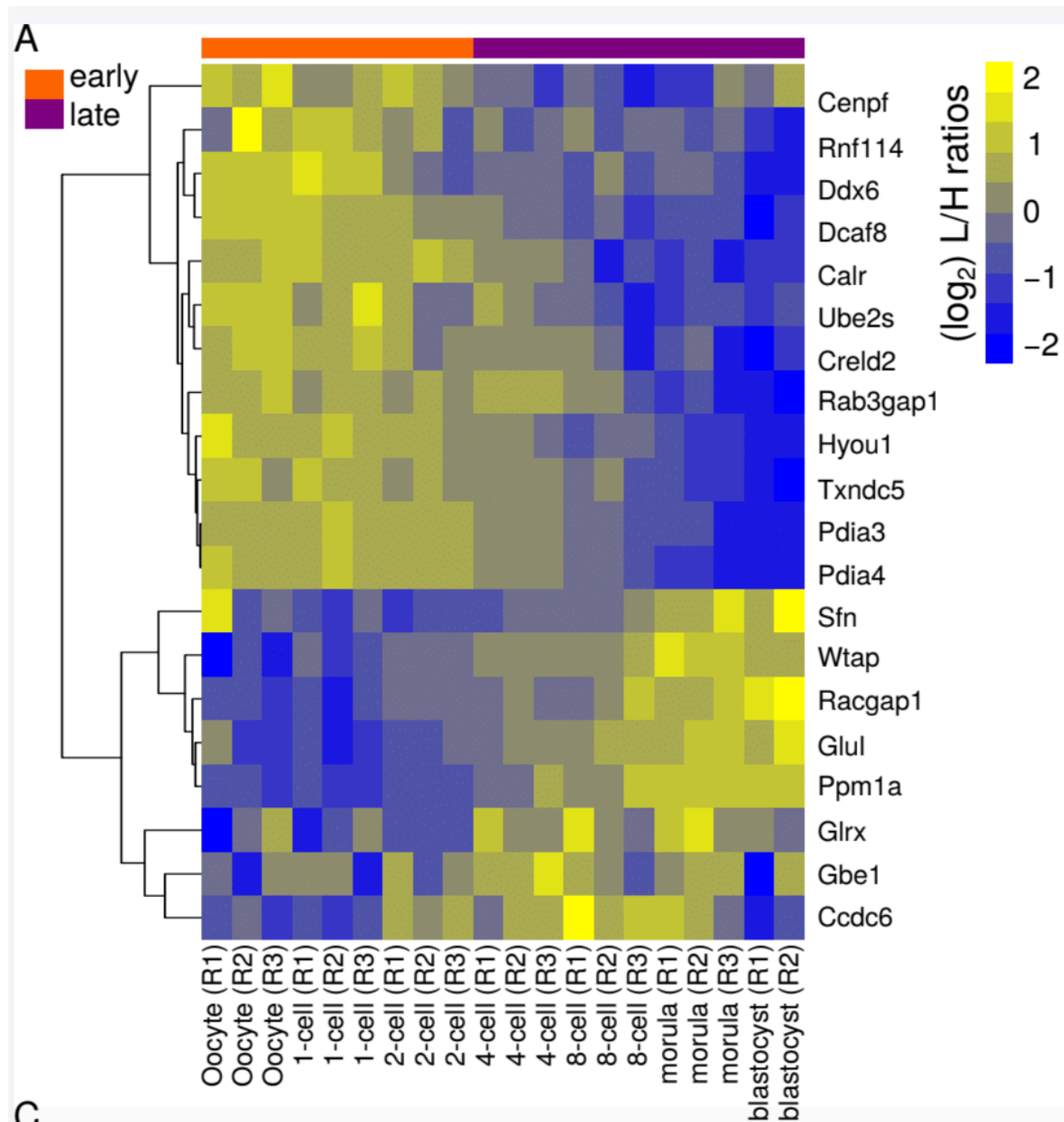
Systems Biology



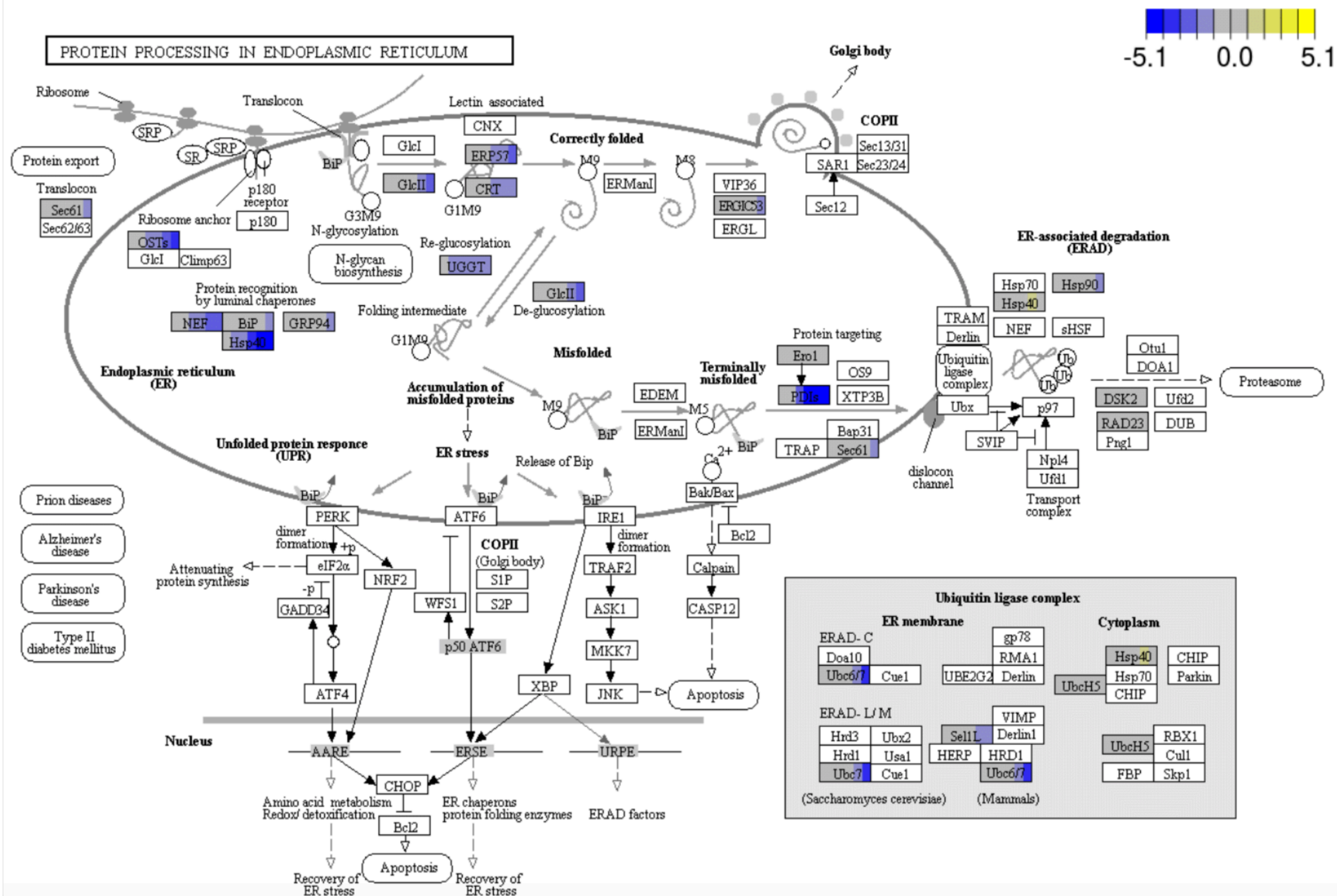


Systems biology is the computational and mathematical modeling of complex biological systems. It is a biology-based interdisciplinary field of study that focuses on complex interactions within biological systems, using a holistic approach (holism instead of the more traditional reductionism) to biological research.

Differential gene expression during mouse early embryogenesis



C



Data on KEGG graph
Rendered by Pathview

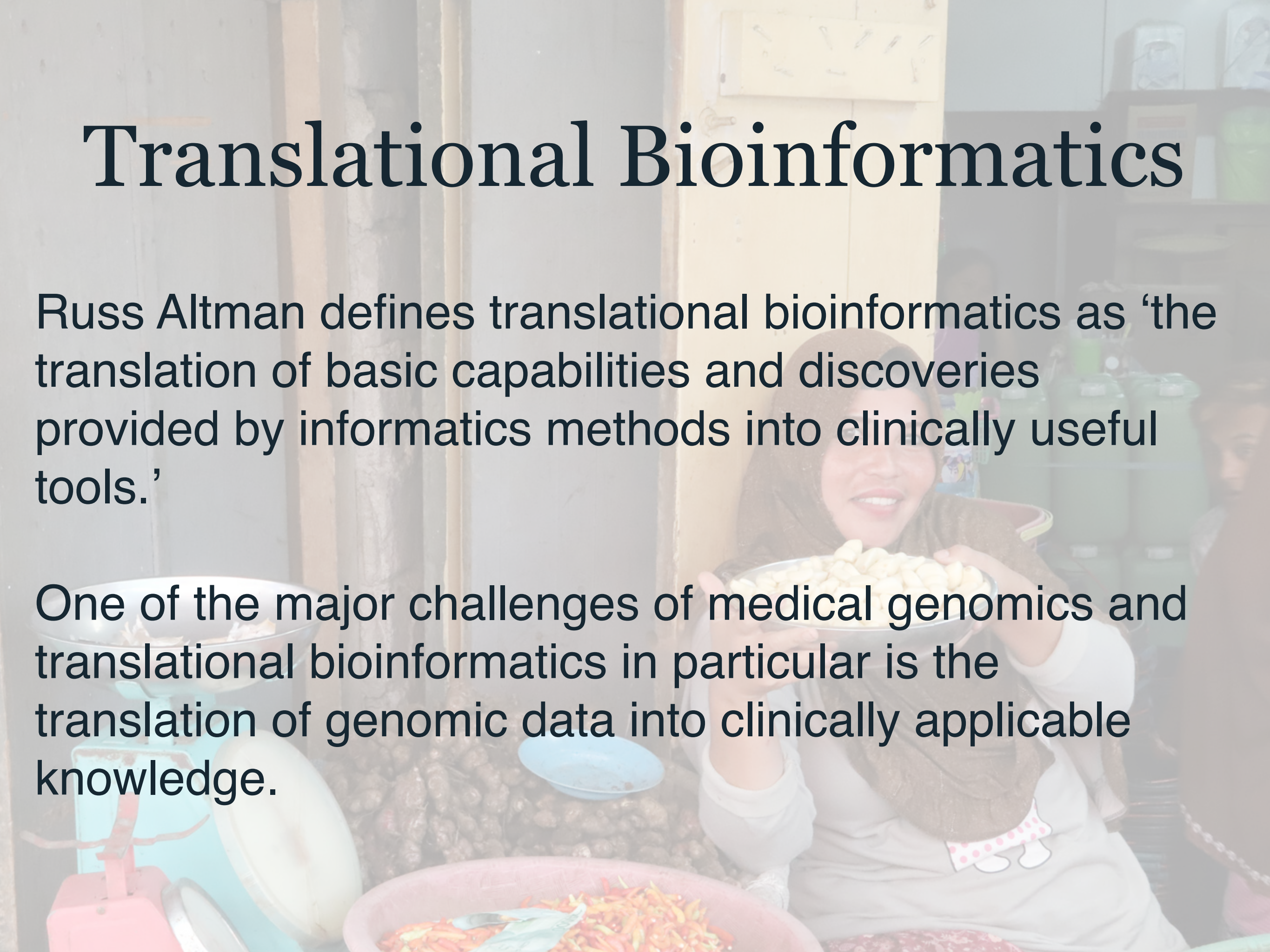
Translational Bioinformatics



Translational Bioinformatics

Russ Altman defines translational bioinformatics as ‘the translation of basic capabilities and discoveries provided by informatics methods into clinically useful tools.’

One of the major challenges of medical genomics and translational bioinformatics in particular is the translation of genomic data into clinically applicable knowledge.

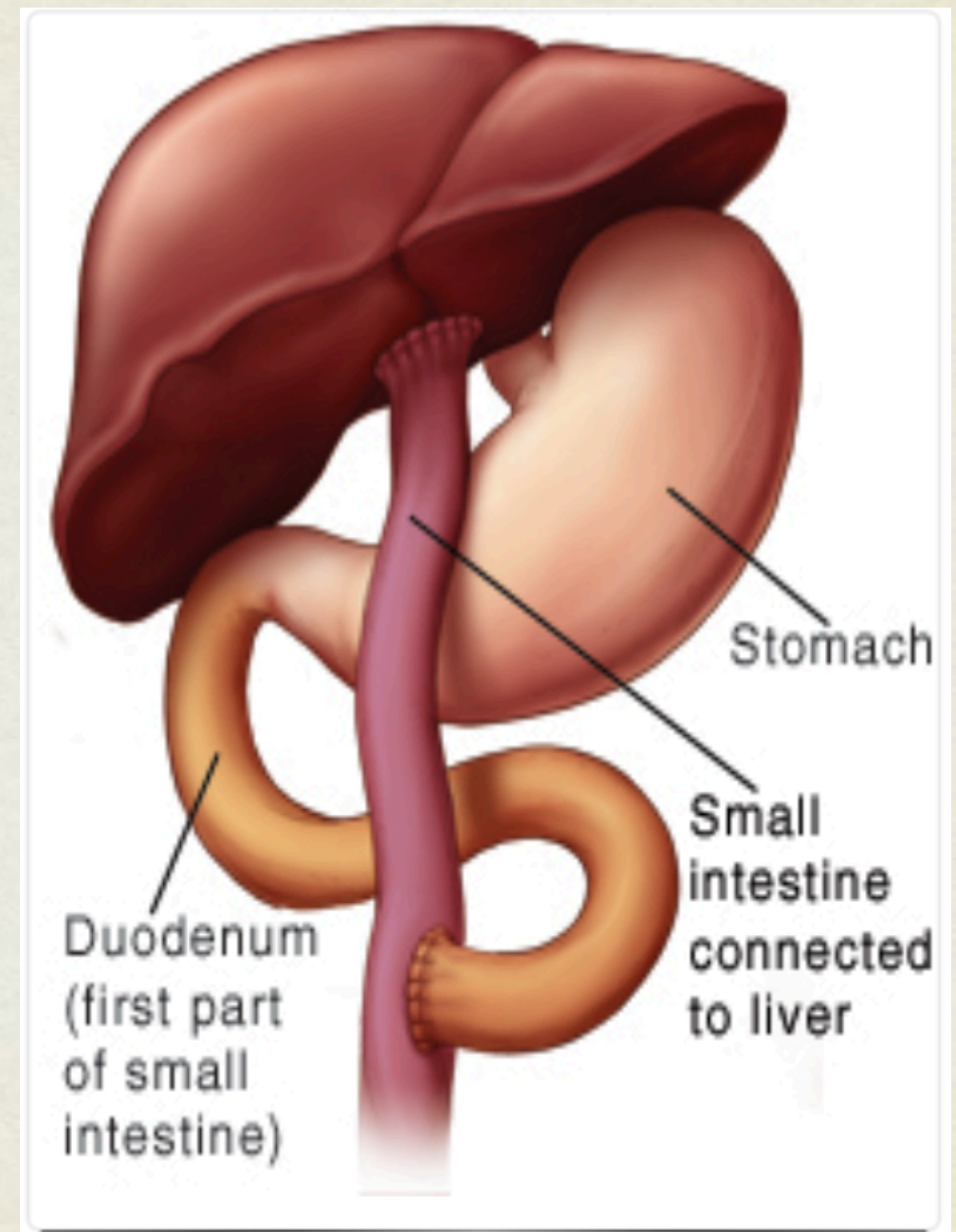


CLINICAL SUCCESS STORY

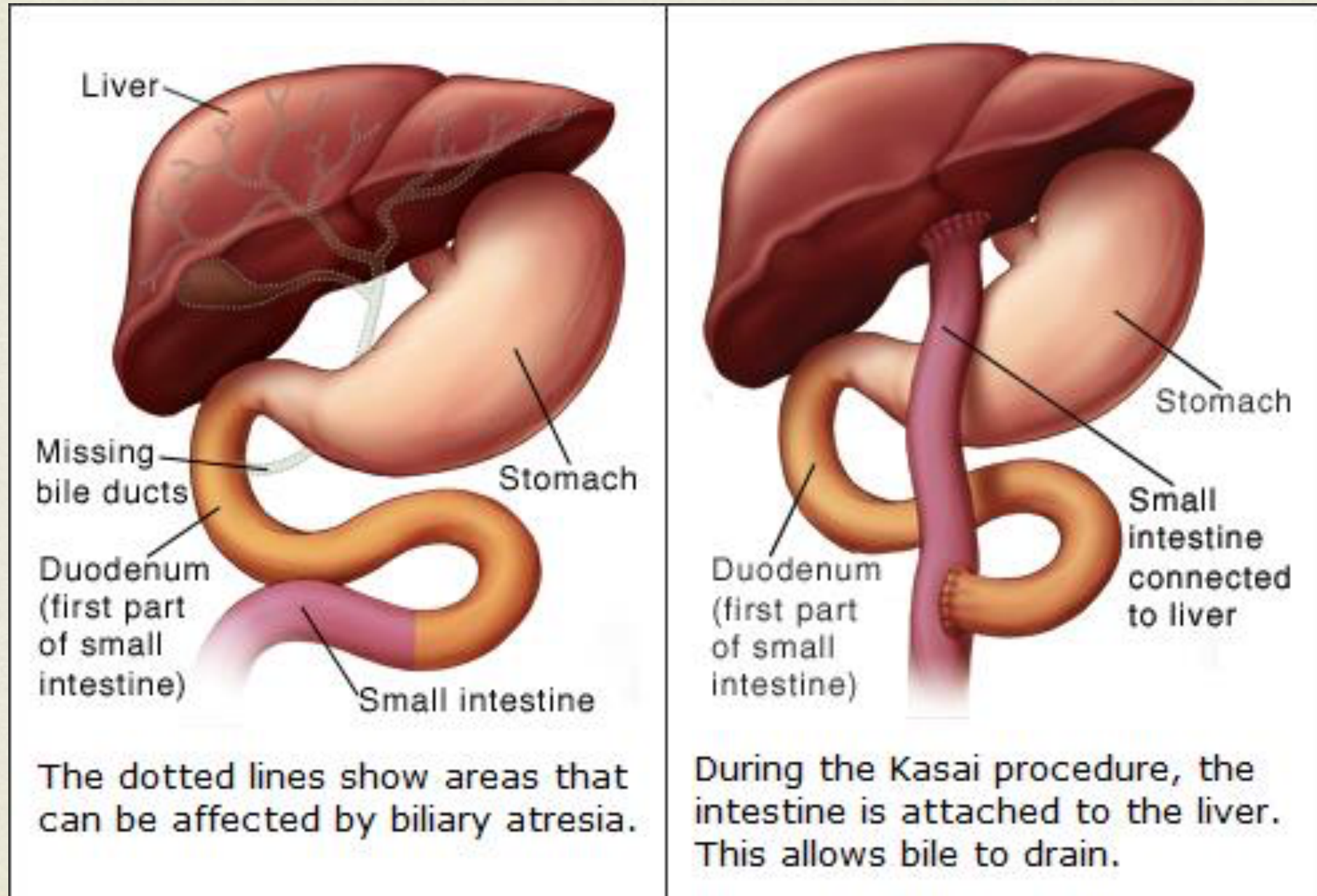


RADY CHILDREN'S HOSPITAL BABY 6026

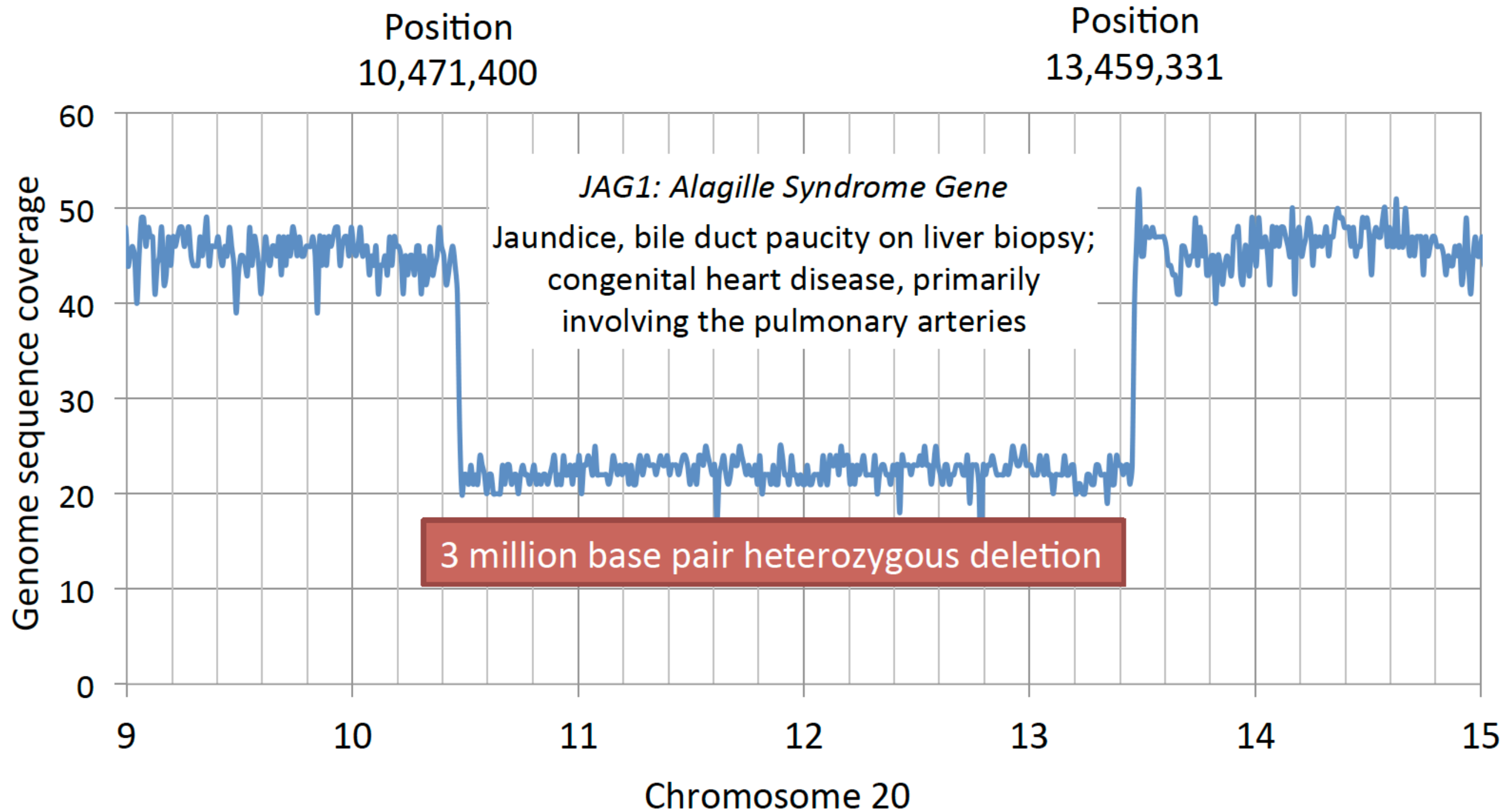
- Two month old child admitted to PICU with severe jaundice & poor weight gain for one month
- Echo: Congenital heart disease, underdeveloped pulmonary arteries
- Clinical diagnosis: biliary atresia
 - one incidence in ten thousand
- Empiric treatment: Kasai procedure



KASAI PROCEDURE



43 HOURS LATER: PROVISIONAL DIAGNOSIS



CLINICAL IMPACT & OUTCOME

- Kasai procedure scheduled for 11:00 am
- Genetic diagnosis communicated to clinical team just before surgery – procedure cancelled
- Infants with Alagille syndrome are occasionally misdiagnosed as biliary atresia and subsequently undergo Kasai operation during infancy
- Among 15 children with Alagille syndrome, mortality was 60% after Kasai procedure, and only 10% among those without Kasai procedure. Liver transplantation was performed in 100% of the Kasai group, and 20% of the non-Kasai group.

Data Volume Problem

Type of cancer	Number of whole genome	Number of whole exome	Data volume (Tb)	Time to download
Colon Adenocarcinoma (COAD)	302	443	33.04	24 days
Lung	134	582	40.95	30 days
Breast	248	1050	69.82	50 days
Prostate Adenocarcinoma (PRAD)	272	1049	26.53	10 days

OCTOBER 22, 1990 • \$1.95

People

weekly



THE DENTIST AND THE PATIENT: AN AIDS MYSTERY

Two years
after routine dental
surgery, college
student Kim Bergalis
developed AIDS.
Now her dentist is
dead of the disease,
and she charges that
he infected her.
"It's hard to believe,"
she says, "but it
happened to me."



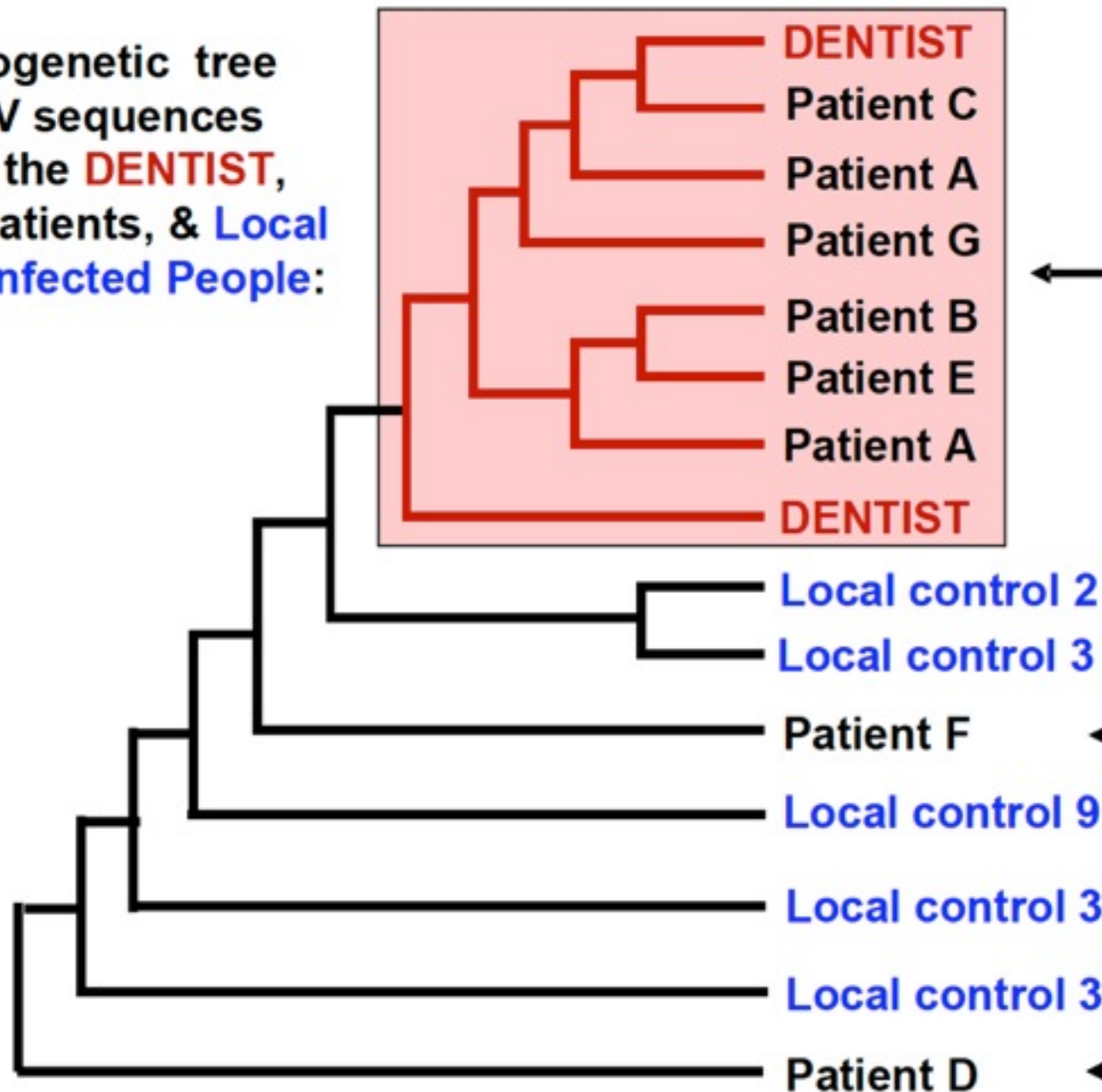
Did the Florida
Dentist infect his
patients with HIV?

Kimberly Bergalis
(1968-1991)

David J. Acer
(1940-1990)

DID THE FLORIDA DENTIST INFECT HIS PATIENTS WITH HIV?

Phylogenetic tree of HIV sequences from the **DENTIST**, his Patients, & **Local HIV-infected People**:



Yes:

The HIV sequences from these patients fall within the clade of HIV sequences found in the dentist.

No

No

From Ou et al. (1992) Molecular epidemiology of HIV transmission in a dental practice. Science. 256:1165-71.

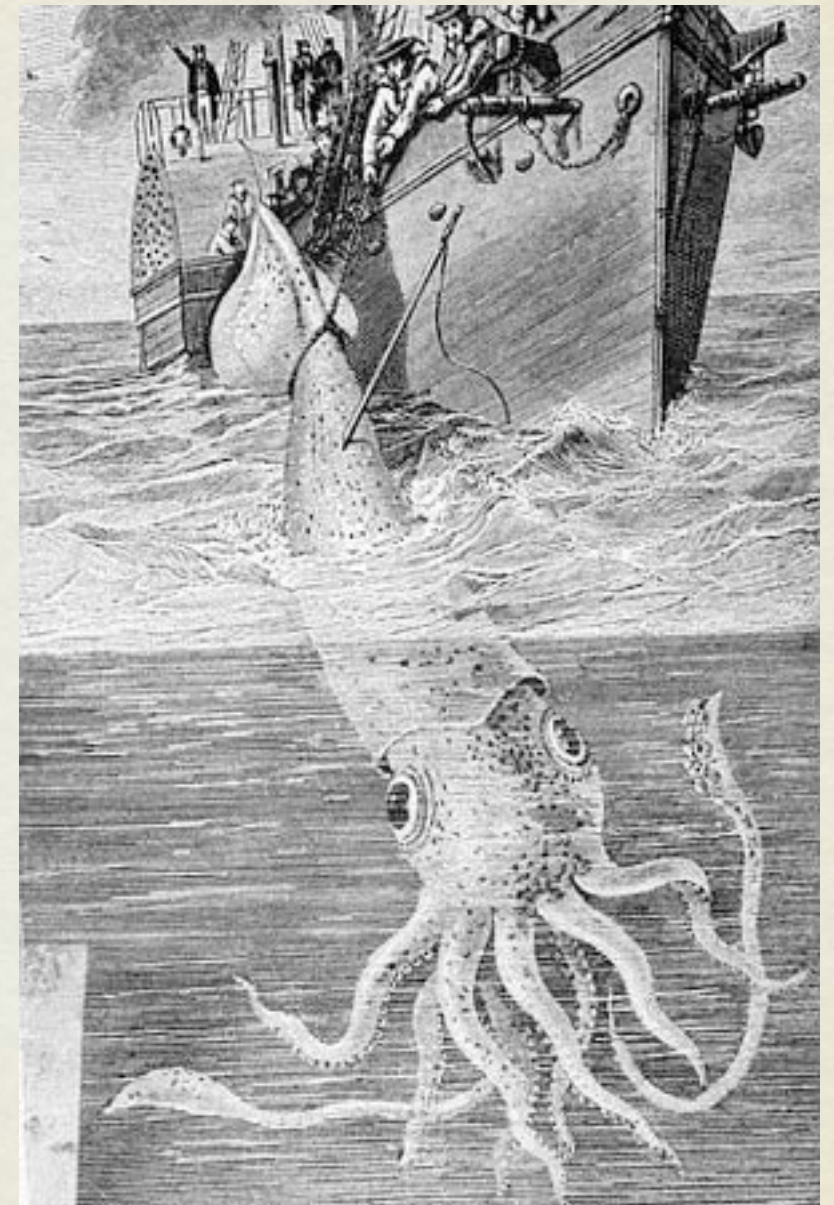
THE MYSTERY OF THE CHILEAN BLOB



THE MYSTERY OF THE CHILEAN BLOB

>Chilean_Blob

TAATACTAACTATATCCCTACTCTCCATTCTCATCGGGG
GTTGAGGAGGACTAAACCAGACTCAACTCCGAAAAATTA
TAGCTTACTCATCAATCGCCCACATAGGATGAATAACCA
CAATCCTACCCTACAATACAACCATAACCCTACTAAACC
TACTAATCTATGTCACAATAACCTTCACCATATTCATAC
TATTTATCCAAAACCTCAACCACAACCACACTATCTCTGT
CCCAGACATGAAACAAAACACCCATTACCACAACCCTTA
CCATACTTACCCTACTTTCCATAGGGGGGCCTCCCACCAC
TCTCGGGCTTTATCCCCAAATGAATAATTATTCAAGAAC
TAACAAAAAACGAAACCCTCATCATACCAACCTTCATAG
CCACCACAGCATTACTCAACCTCTACTTCTATATACGCC
TCACCTACTCAACAGCACTAACCCTATTCCCCTCCACAA
ATAACATAAAAATAAAAATGACAATTCTACCCCACAAAAC
GAATAACCCTCCTGCCAACAGCAATTGTAATATCAACAA
TACTCCTACCCCTTACACCAATACTCTCCACCCTATTAT
AG



THE MYSTERY OF THE CHILEAN BLOB

Lineage Report

Cetacea	[whales & dolphins]				
. Odontoceti	[whales & dolphins]				
. . Physeteridae	[whales & dolphins]				
. . . Physeter catodon	-----	1085	3 hits	[whales & dolphins]	Physeter catodon NADH dehydrogenase subunit 2 (nad2) gene,
. . . Kogia breviceps	-----	638	1 hit	[whales & dolphins]	Kogia breviceps complete mitochondrial genome
. . Orcaella brevirostris	-----	593	1 hit	[whales & dolphins]	Orcaella brevirostris isolate 97 mitochondrion, complete ge
. . Grampus griseus	-----	593	1 hit	[whales & dolphins]	Grampus griseus mitochondrion, complete genome
. . Feresa attenuata	-----	592	2 hits	[whales & dolphins]	Feresa attenuata isolate 36 mitochondrion, complete genome
. . Tursiops truncatus (bottle-nosed dolphin)	...	592	1 hit	[whales & dolphins]	Tursiops truncatus mitochondrion, complete genome
. . Globicephala melas	-----	586	3 hits	[whales & dolphins]	Globicephala melas isolate GlomelG42 mitochondrion, partial
. . Peponocephala electra	-----	580	2 hits	[whales & dolphins]	Peponocephala electra isolate M6 mitochondrion, complete ge
. . Globicephala macrorhynchus	-----	580	4 hits	[whales & dolphins]	Globicephala macrorhynchus isolate Glomac65 mitochondrion,
. . Pseudorca crassidens	-----	577	3 hits	[whales & dolphins]	Pseudorca crassidens mitochondrion, complete genome
. . Orcinus orca (Orca)	-----	569	54 hits	[whales & dolphins]	Orcinus orca isolate ENPTGA2 mitochondrion, complete genome
. . Sotalia fluviatilis	-----	569	2 hits	[whales & dolphins]	Sotalia fluviatilis haplotype 10 NADH dehydrogenase subunit
. . Platanista minor	-----	569	1 hit	[whales & dolphins]	Platanista minor complete mitochondrial genome
. . Steno bredanensis	-----	566	2 hits	[whales & dolphins]	Steno bredanensis isolate Stebres9 mitochondrion, partial g
. Megaptera novaeangliae	-----	636	5 hits	[whales & dolphins]	Megaptera novaeangliae voucher GOM9049 NADH dehydrogenase s
. Balaenoptera bonaerensis	-----	630	1 hit	[whales & dolphins]	Balaenoptera bonaerensis mitochondrial DNA, complete genome
. Eubalaena japonica	-----	619	1 hit	[whales & dolphins]	Eubalaena japonica mitochondrial DNA, complete genome
. Balaenoptera brydei	-----	614	2 hits	[whales & dolphins]	Balaenoptera brydei mitochondrial DNA, complete genome, iso
. Balaena mysticetus (Greenland right whale)		614	2 hits	[whales & dolphins]	Balaena mysticetus mitochondrial DNA, complete genome
. Balaenoptera musculus	-----				
. Balaenoptera edeni	-----				
. Balaenoptera omurai	-----				
. Eschrichtius robustus (California gray whale)	.				
. Balaenoptera borealis	-----				
. Caperea marginata	-----				
. Balaenoptera physalus (finback whale)	-----				



THE MYSTERY OF THE CHILEAN BLOB

>[emb|AJ277029.2](#) D *Physeter macrocephalus* mitochondrial genome
Length=16428

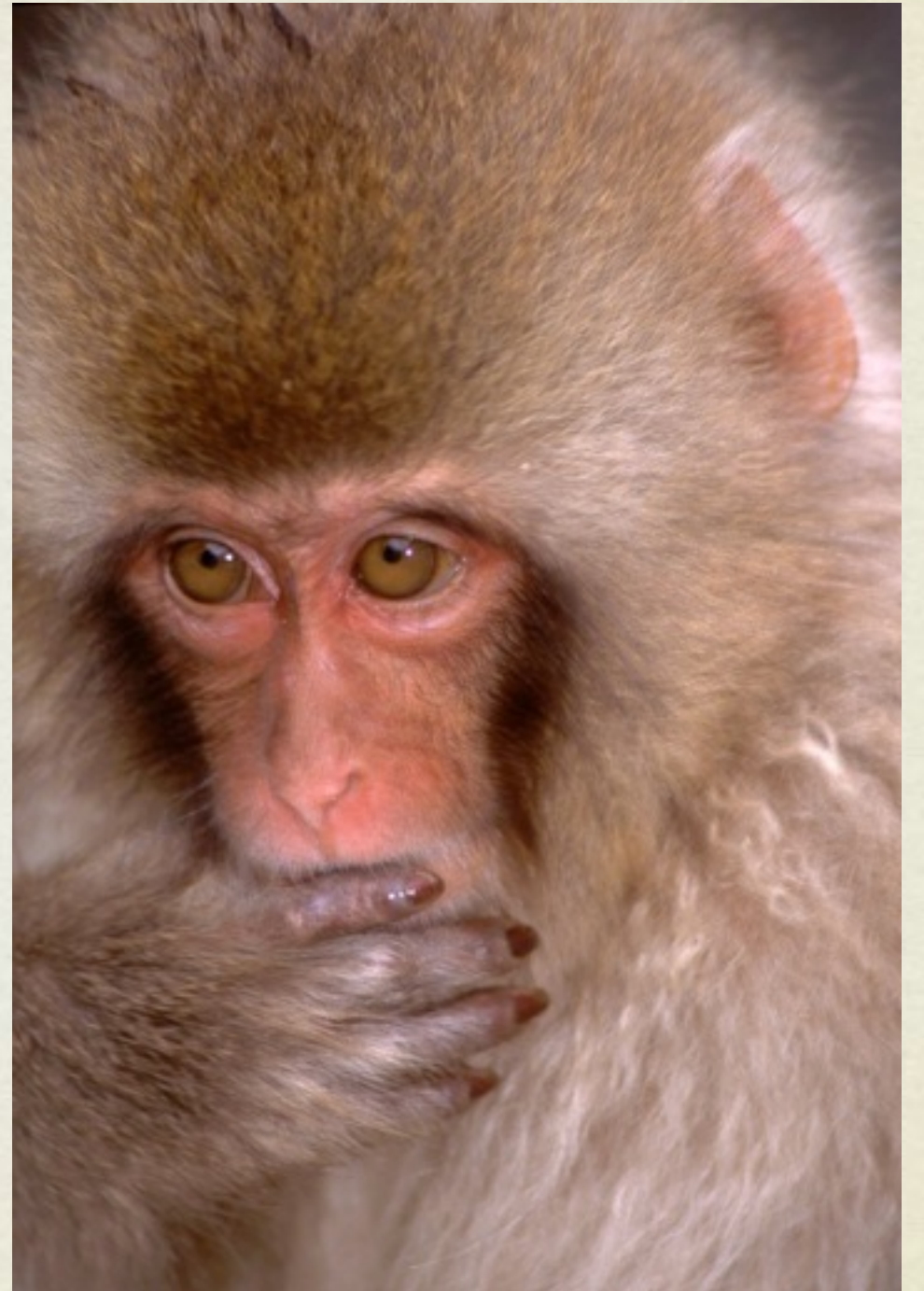
Score = 1074 bits (581), Expect = 0.0
Identities = 585/587 (99%), Gaps = 0/587 (0%)
Strand=Plus/Plus

Query	1	TAATACTAACTATATCCCTACTCTCCATTCTCATCGGGGGTTGAGGAGGACTAAACCAGA	60
Sbjct	4400	TAATACTAACTATATCCCTACTCTCCATTCTCATCGGGGGTTGAGGAGGACTAAACCAGA	4459
Query	61	CTCAACTCCGAAAAATTATAGCTTACTCATCAATCGCCACATAGGATGAATAACCACAA	120
Sbjct	4460	CTCAACTCCGAAAAATTATAGCTTACTCATCAATCGCCACATAGGATGAATAACCACAA	4519
Query	121	TCCTACCCTACAATACAACCATAACCCTACTAAACCTACTAATCTATGTCACAATAACCT	180
Sbjct	4520	TCCTACCCTACAATACAACCATAACCCTACTAAACCTACTAATCTATGTCACAATAACCT	4579
Query	181	TCACCATATTCATACTATTTATCCAAAACCTCAACCACAACCACACTATCTCTGTCCCAGA	240
Sbjct	4580	TCACCATATTCACACTATTTATCCAAAACCTCAACCACAACCACACTATCTCTGTCCCAGA	4639
Query	241	CATGAAACAAAACACCCATTACCACAACCCTTACCATACTTACCCTACTTTCCATAGGGG	300
Sbjct	4640	CATGAAACAAAACACCCATTACCACAACCCTTACCATACTTACCCTACTTTCCATAGGGG	4699
Query	301	GCCTCCCACCACTCTCGGGCTTTATCCCCAAATGAATAATTATTCAAGAACTAACAAAA	360
Sbjct	4700	GCCTCCCACCACTCTCGGGCTTTATCCCCAAATGAATAATTATTCAAGAACTAACAAAA	4759
Query	361	ACGAAACCCTCATCATACCAACCTTCATAGCCACCACAGCATTACTCAACCTCTACTTCT	420
Sbjct	4760	ACGAAGCCCTCATCATACCAACCTTCATAGCCACCACAGCATTACTCAACCTCTACTTCT	4819
Query	421	ATATACGCCTCACCTACTCAACAGCACTAACCCCTATTCCCCTCCACAAATAACATAAAAA	480
Sbjct	4820	ATATACGCCTCACCTACTCAACAGCACTAACCCCTATTCCCCTCCACAAATAACATAAAAA	4879
Query	481	TAAAATGACAATTCTACCCACAAAACGAATAACCCTCCTGCCAACAGCAATTGTAATAT	540
Sbjct	4880	TAAAATGACAATTCTACCCACAAAACGAATAACCCTCCTGCCAACAGCAATTGTAATAT	4939
Query	541	CAACAATACTCCTACCCCTTACACCAATACTCTCCACCCTATTATAG	587
Sbjct	4940	CAACAATACTCCTACCCCTTACACCAATACTCTCCACCCTATTATAG	4986



BIOINFORMATICS IN MEDICINE CHALLENGES

- Computational skills for in-depth analyses
- Data interpretation
- Research translation
- Data volume!!!



<http://bioinformatics.uni-muenster.de>

