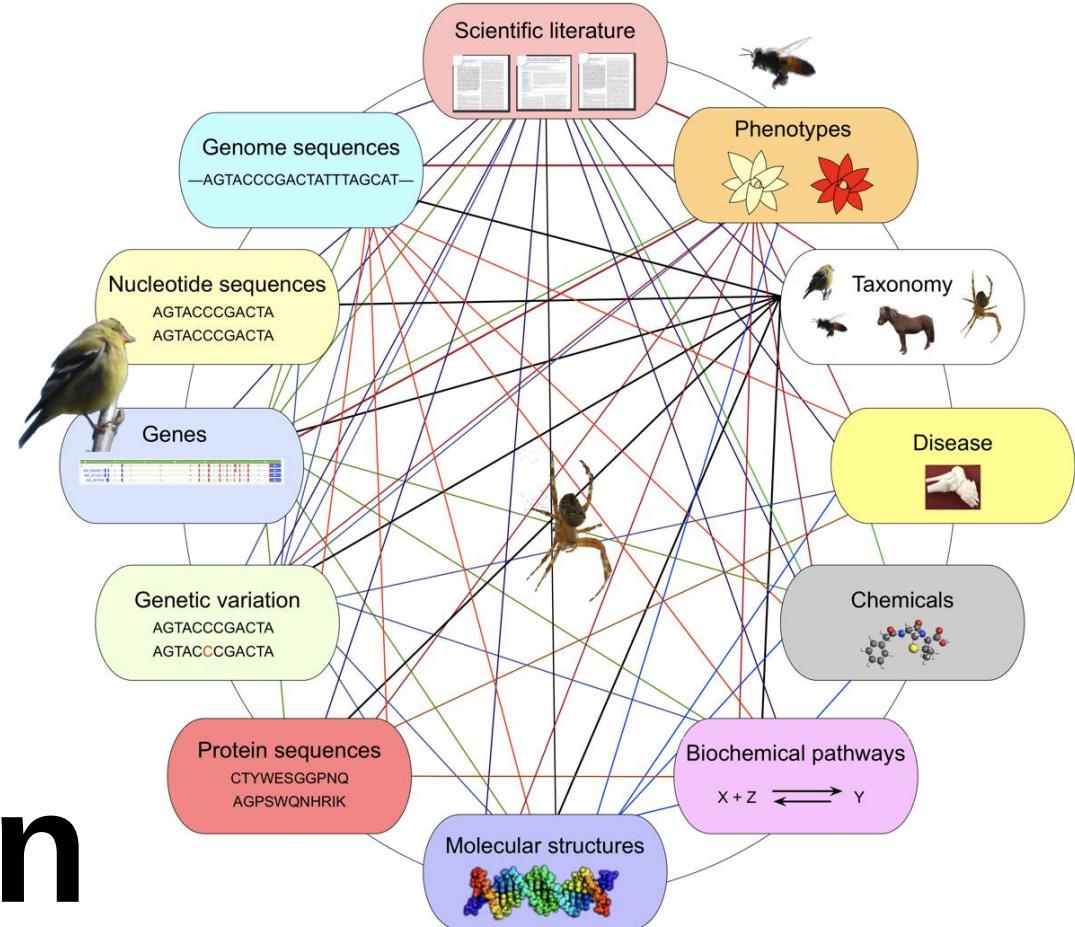


Biology Databases And Data Manipulation

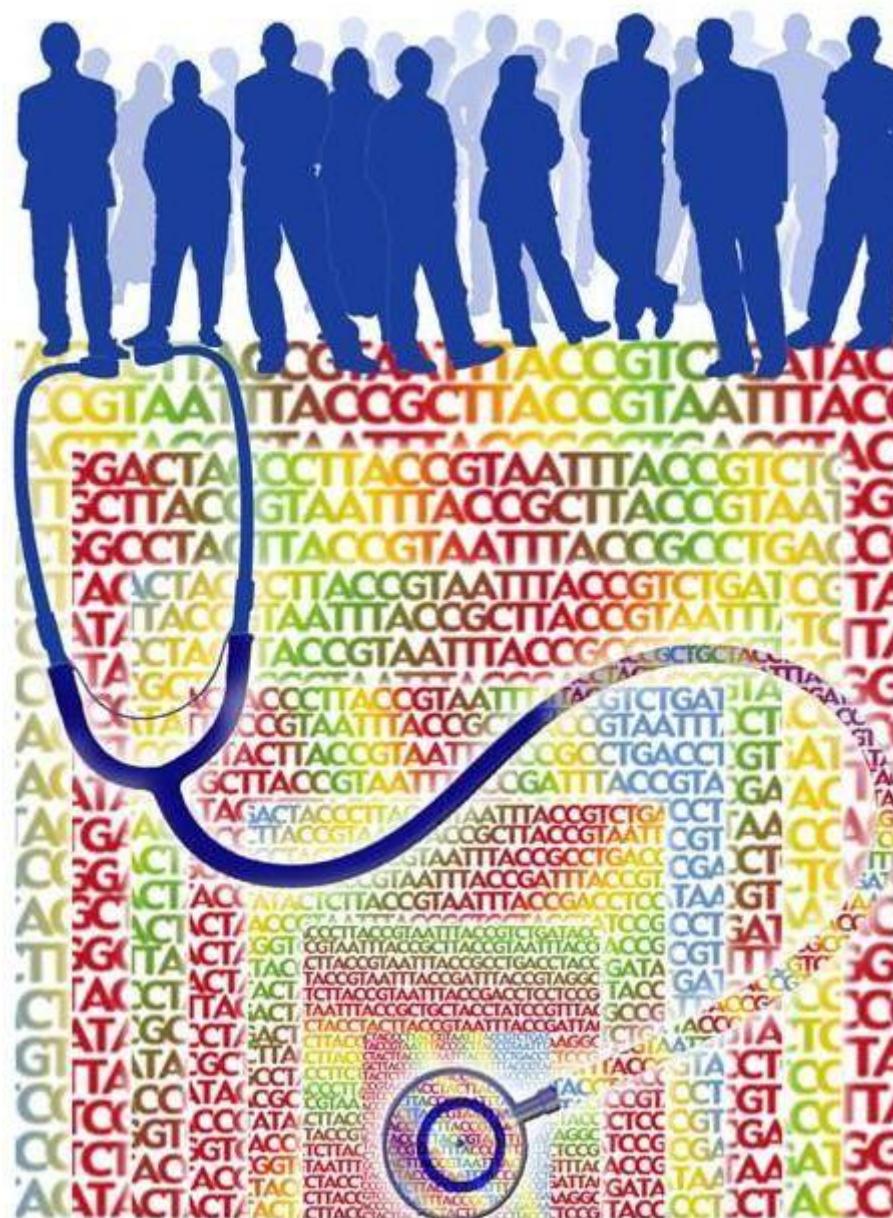


Sandra Porter, PhD

Digital World Biology LLC

V. Courtier-Orgogozo

Human genetic diversity



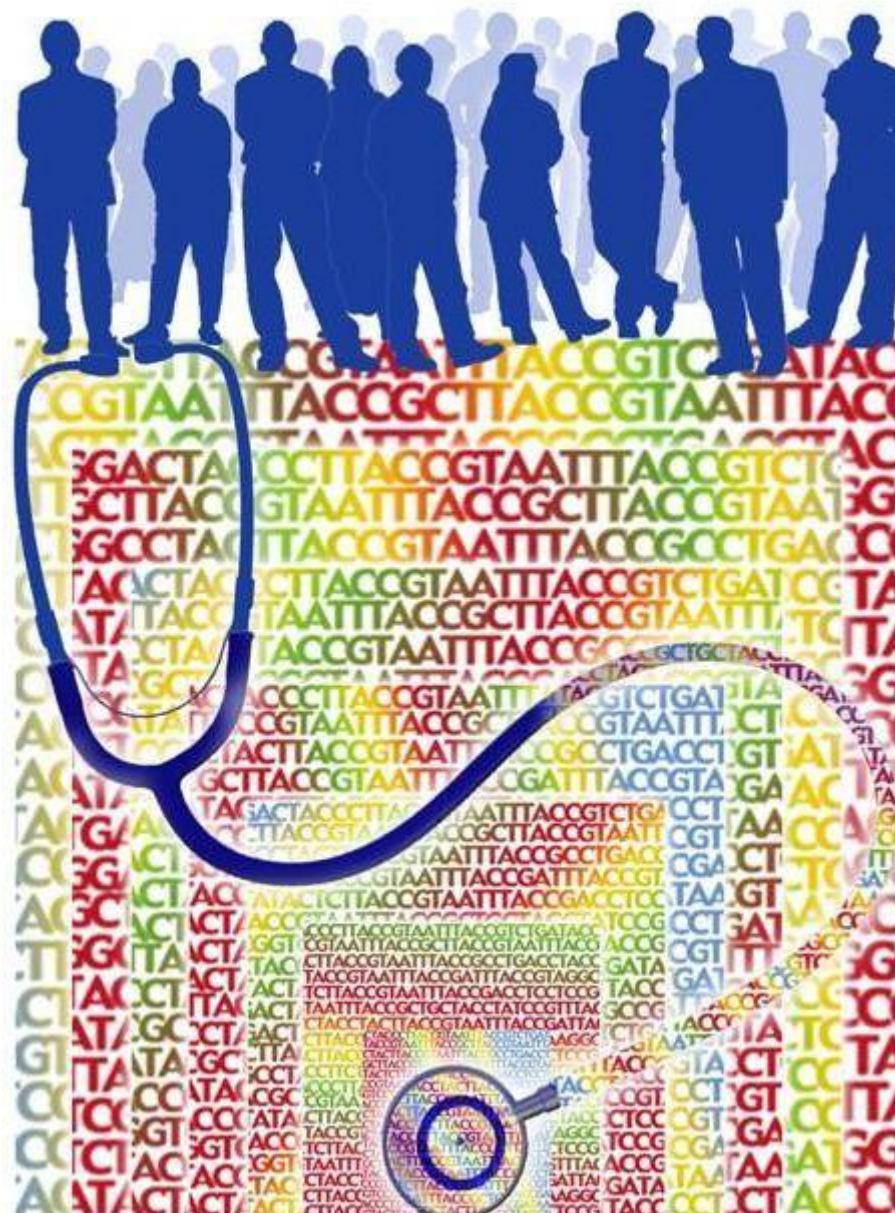
Genome size: 2.9 Gb
Gene number: 25 000
(1% of coding sequences)

In one individual:
~100 new mutations compared to his parents
~20 lethal mutations (heterozygous)

Genetic difference between two humans?

Genetic differences between humans and chimps?

Human genetic diversity



Genome size: 2.9 Gb
Gene number: 25 000
(1% of coding sequences)

In one individual:
~100 new mutations compared to his parents
~20 lethal mutations (heterozygous)

Genetic difference between two humans?

~0.1%

Genetic differences between humans and chimps?

~4% (<1% for coding sequences)



99.4% human?

Banners by www.zephyr-tvc.com

Manipulate DNA sequences in silico

Manipulate DNA sequences in silico

Store

Store/retrieve data
Annotate

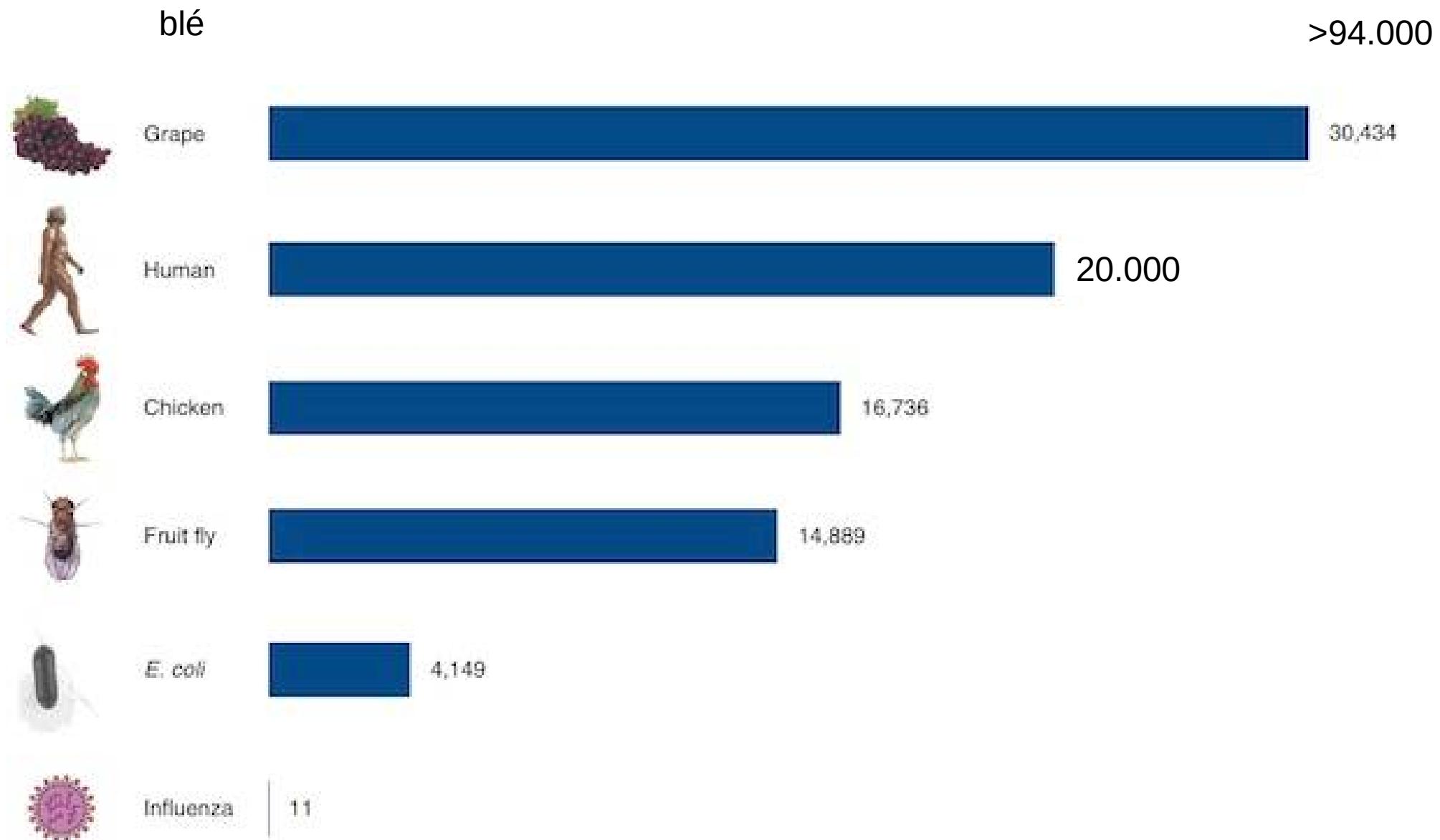
Assemble

Compare

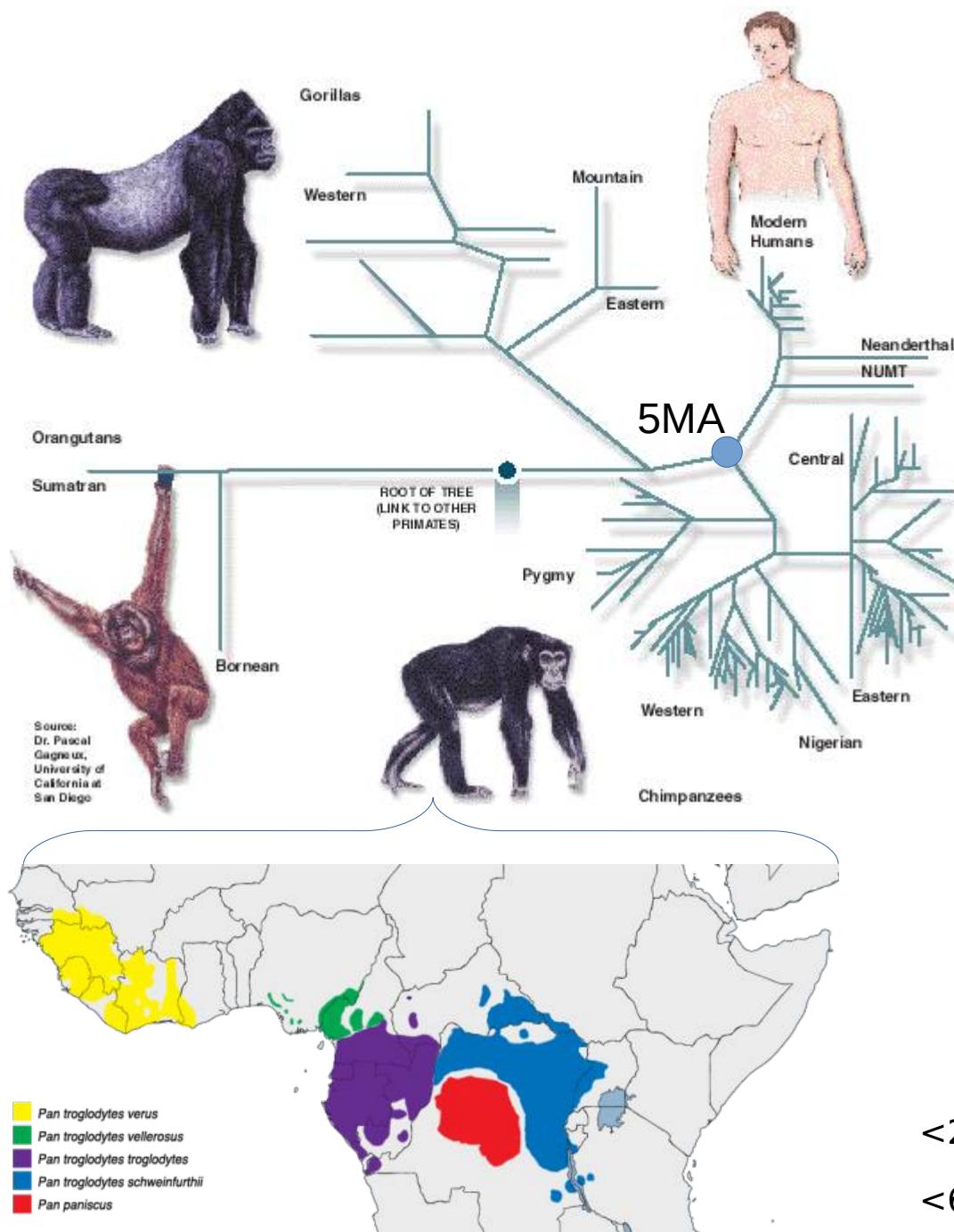
BLAST
Align
Phylogenetic tree

etc.

Nombre de gènes et complexité phénotypique



L'homme et les autres primates



Taille du génome : $2,9 \text{ Gb} = 2,9 \cdot 10^9 \text{ pb}$

Différence entre deux humains

$\sim 0.1\%$

Différence entre homme et chimpanzé

$\sim 5\% (\sim 1\% \text{ for coding sequences})$

Différence entre homme et Néanderthal

$\sim 0.13\%$

Différence entre deux chimpanzés

<250,000 chimpanzées, <110,000 gorillas,
<60,000 orangutans, and <50,000 bonobos

Peau, cheveux, santé...

Voici ce que Neandertal nous a vraiment légué

Si notre génome a conservé la trace de notre cousin, comment s'incarne cet héritage au niveau de nos gènes ? Deux études américaines viennent de répondre. Avec quelques surprises à la clé...

Par Emilie Rauscher

En 2010, les rumeurs de leur union avaient défrayé la chronique scientifique... Le scandale était venu des laboratoires de paléogénétique du Max-Planck Institut (Allemagne), qui avaient fait d'une pierre deux coups : ils lisaient pour la première fois l'ADN de notre cousin perdu, Neandertal, et, par ricochet, en découvraient des traces dans notre propre génome, pourtant estampillé *sapiens*. La conséquence d'une relation forcément sexuelle. D'autres travaux avaient suivi, pour approfondir le sujet : il y

avait eu relations, certes, mais il y a 47 000 à 65 000 ans, à notre sortie d'Afrique, et l'hypothèse d'une orgie généralisée était écartée puisque seul un faible pourcentage d'ADN néandertalien survit en nous.

Assez toutefois pour aiguiser la curiosité des biologistes : de lui à nous, qu'est-ce qui était passé ? "Jusqu'ici, nous cherchions plutôt à confirmer qu'il y avait bien eu échange... Maintenant, nous traquons ces fragments à travers tout notre génome !", s'enthousiasme Joshua Akey, généticien à l'université de Washington.

Chronologie

Le dernier ancêtre commun à Neandertal et *Homo sapiens* est vieux de 500 000 ans. Puis, ces deux groupes ont évolué chacun de leur côté : en Europe et en Asie pour le premier, en Afrique pour le second. Si bien que quand ils se sont "retrouvés" il y a 40 000 ans au Moyen-Orient, ils avaient accumulé de nombreuses différences.

Le cas particulier du chromosome X

Il compte 5 fois moins d'apports néandertaliens que les autres chromosomes. Ces fragments ont disparu sous l'effet de la sélection naturelle, sans doute parce qu'ils diminuaient la fertilité des hybrides mâles Neandertal/*Sapiens*.

ADN *sapiens*
Séquences d'ADN
néandertalien

Un héritage réparti sur nos 23 chromosomes

Les populations européennes et asiatiques ont conservé des fragments différents de l'ADN de Neandertal. En les rassemblant sur les 23 chromosomes d'un être humain type, il apparaît que près de 40 % de l'ADN de notre cousin est passé dans celui de notre espèce. Et de premières séquences de gènes ont été identifiées, dévoilant les fonctions sous influence de Neandertal.

Formation des cellules de la peau

Des séquences de régulation néandertaliennes ont été gardées par les Asiatiques de l'Est (66 %) pour le gène POU2F3, qui dirige la multiplication des principales cellules de l'épiderme.



Dégradation des lipides
Chez les Européens, 38 gènes sur 19 de nos chromosomes, impliqués dans l'utilisation des lipides par le cerveau, possèdent des apports néandertaliens.

Constitution de la peau, des ongles, des cheveux
Plusieurs gènes KRT liés à la production de protéines fibroses (les kératines) ont un fort apport néandertalien.

Développement des maladies auto-immunes
Des variants néandertaliens sur les chromosomes 7, 9, 10, 11 pourraient être associés à un risque de maladies auto-immunes (lupus, cirrhose biliaire, maladie de Crohn...).

Pigmentation de l'épiderme
Certaines séquences néandertaliennes sont conservées par les Européens (70 %) pour réguler l'expression du gène BNC2, qui influe sur le niveau de pigmentation de la peau.

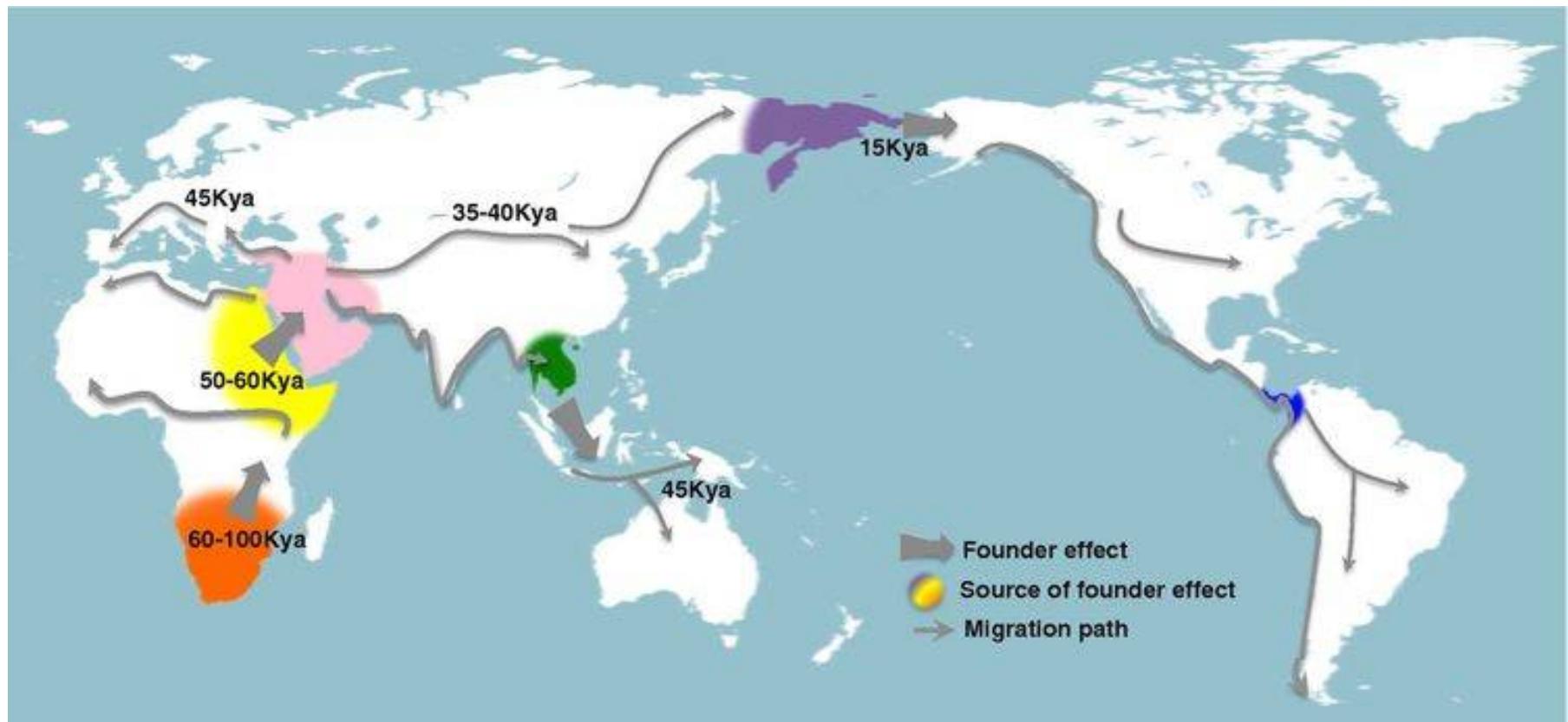
Protection contre les UV

Dix-huit gènes importants dans la protection contre les UV, dont HYAL2, sont touchés chez les Asiatiques. Il y a 49 % de séquences d'origine néandertaliennes chez les Japonais et 66 % chez les Chinois du Sud.

Renforcement du système immunitaire
Les variations apportées par Neandertal y ont été conservées par les 200 gènes HLA de notre système immunitaire. Cette diversification est particulièrement utile pour reconnaître un maximum de pathogènes.

Modification du métabolisme

Certains processus métaboliques cellulaires, comme ceux dirigés par le gène SIPA1L2, sont modifiés par des variants néandertaliens.



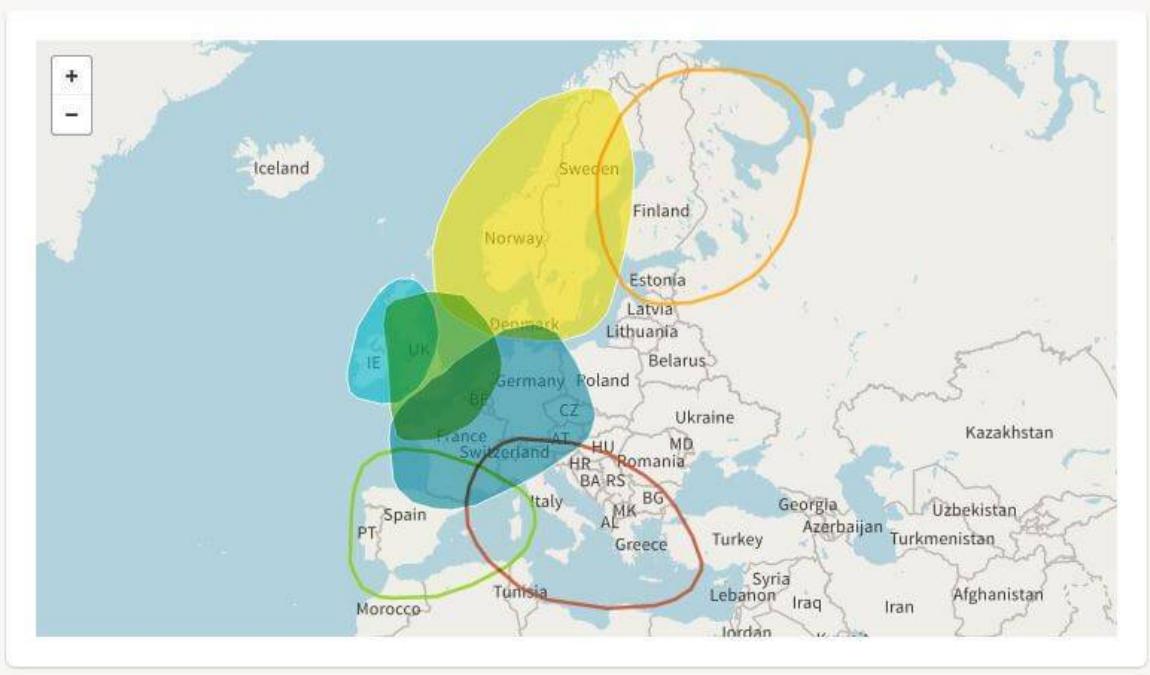
Afrique:
Plus grande diversité génétique

Séquences de Néanderthal: ~1-4% de notre génome, partout sauf en Afrique



Ethnicity estimate for Elaine Clark

REGION	APPROXIMATE AMOUNT
Asia	< 1%
Trace Regions ?	< 1%
Europe	99%
Europe West	40%
Ireland	21%
Great Britain	17%
Scandinavia	11%
Trace Regions ?	10%
Finland/Northwest Russia	5%
Italy/Greece	4%
Iberian Peninsula	< 1%



04-ancestry-com-o0mVUu2kRcs.mp4

Creating a phylogenetic tree

Model based approaches (computer intensive)

Assume a specific model (potentially very complicated) for the probability of evolutionary events occurring and then compare trees using these models.

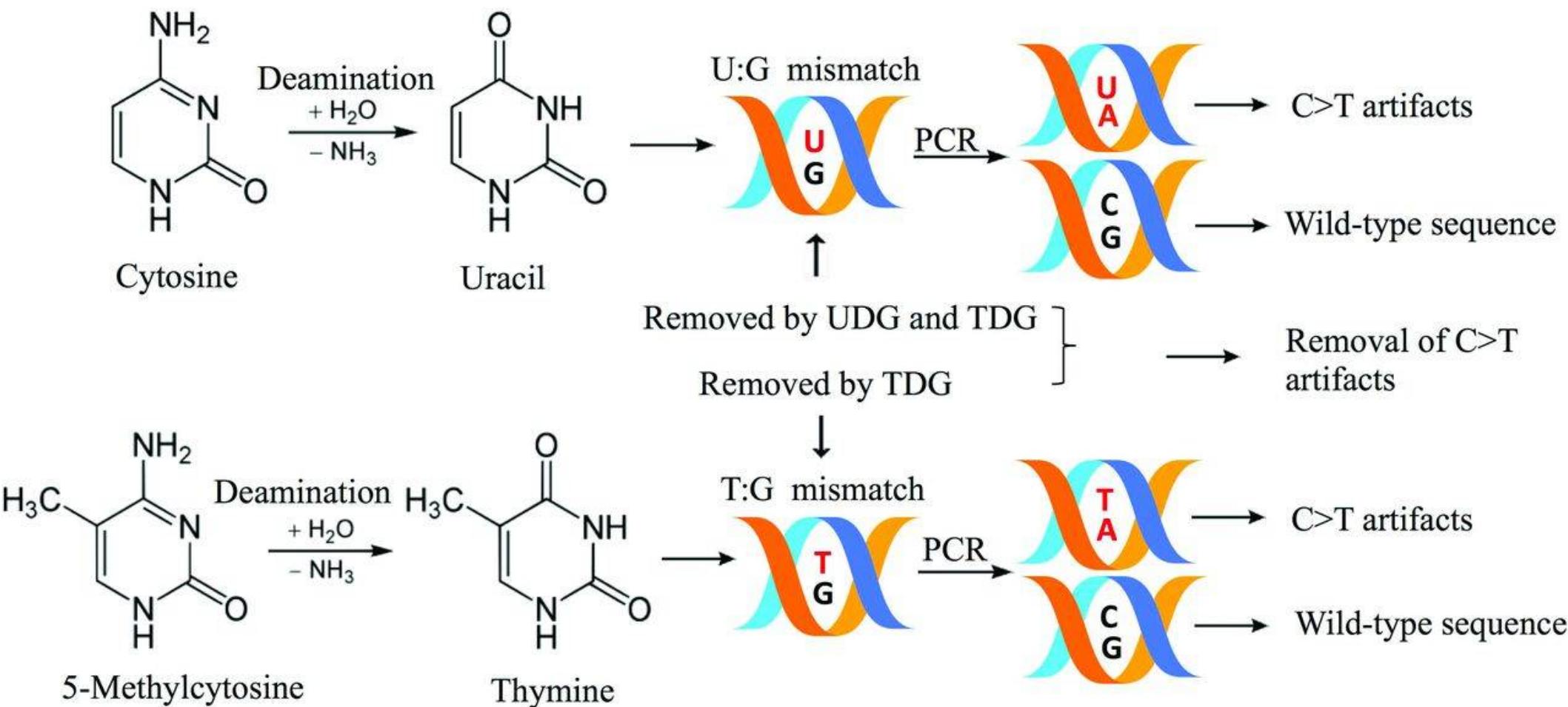


Maximum likelihood: compute the probability of seeing the data we observe given the tree we are assuming. Pick tree with highest probability.

$$P(\text{data}|\text{tree})$$

Bayesian: compute the probability of getting the tree we assume given the data we have. Pick tree with highest probability.

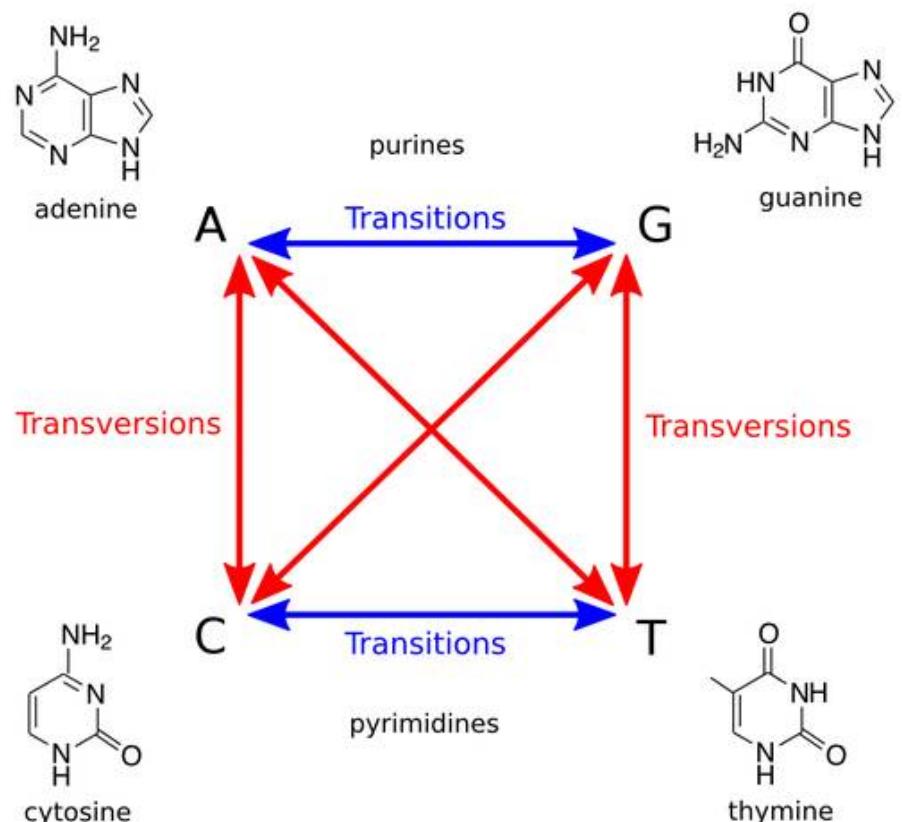
The most common mutation is C>T



UDG (uracile-ADN glycosylase) removes the uracil from U:G mismatches.

TDG (thymine ADN glycosylase) removes the uracil from U:G and thymine from T:G mismatches.

IUPAC nucleotide code



Symbol	Bases	Description
A	A	Adenine
C	C	Cytosine
G	G	Guanine
T (or U)	T (or U)	Thymine (or Uracil)
W	A or T	Weak
S	C or G	Strong
M	A or C	aMino
K	G or T	Keto
R	A or G	puRine
Y	C or T	pYrimidine
B	C or G or T	not A (B comes after A)
D	A or G or T	not C (D comes after C)
H	A or C or T	not G (H comes after G)
V	A or C or G	not T (V comes after T and U)
N	any base	any Nucleotide (not a gap)

The standard statistical technique to determine the validity of a tree involves resampling our data to see how robust the tree is.

Bootstrapping is resampling (with replacement) to create a new data set.

- Pick n characters from the set (may duplicate some and omit others)
- Conceptually simulates robustness to gathering even more data



Jacknifing is subsampling (without replacement) to create a smaller data set.

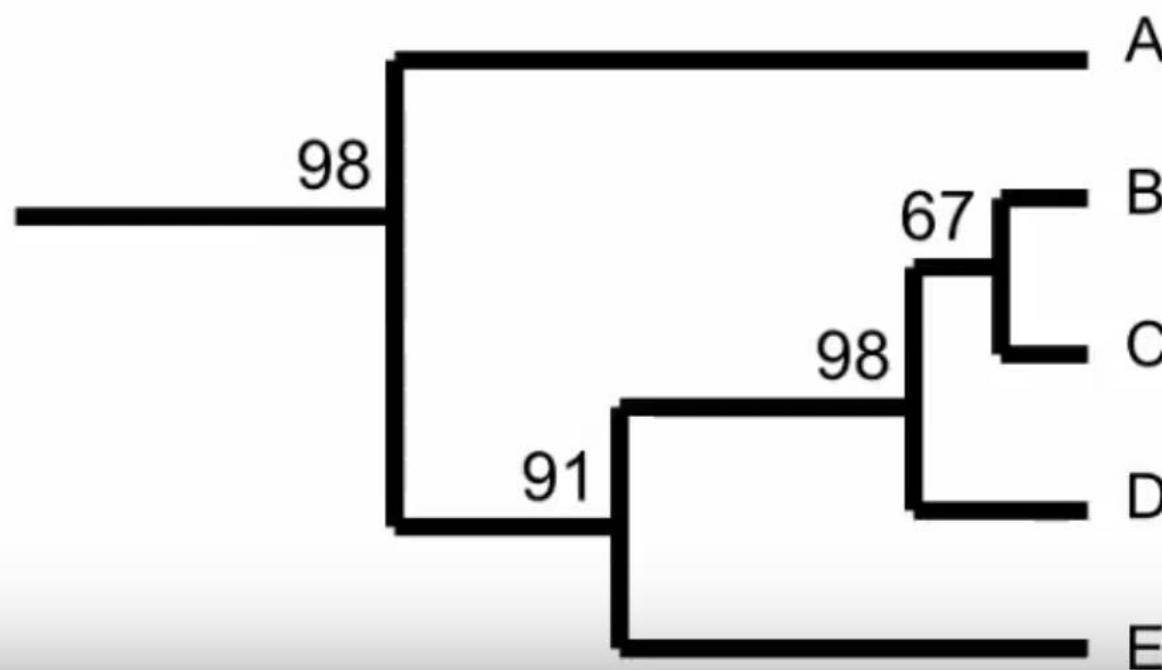
- Pick n characters from the set (will omit some)
- Conceptually simulates robustness to having gathered less data

Bootstrapping is more common.

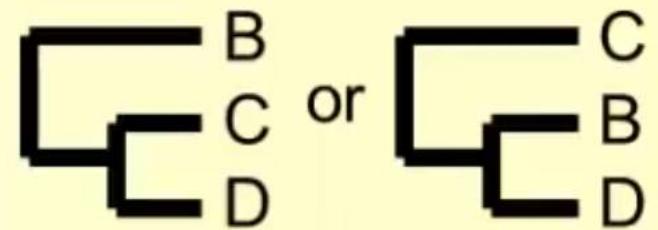
Bootstrap values for maximum likelihood trees

In practice the tree with the highest criteria score is shown and bootstrapping values are show for each node.

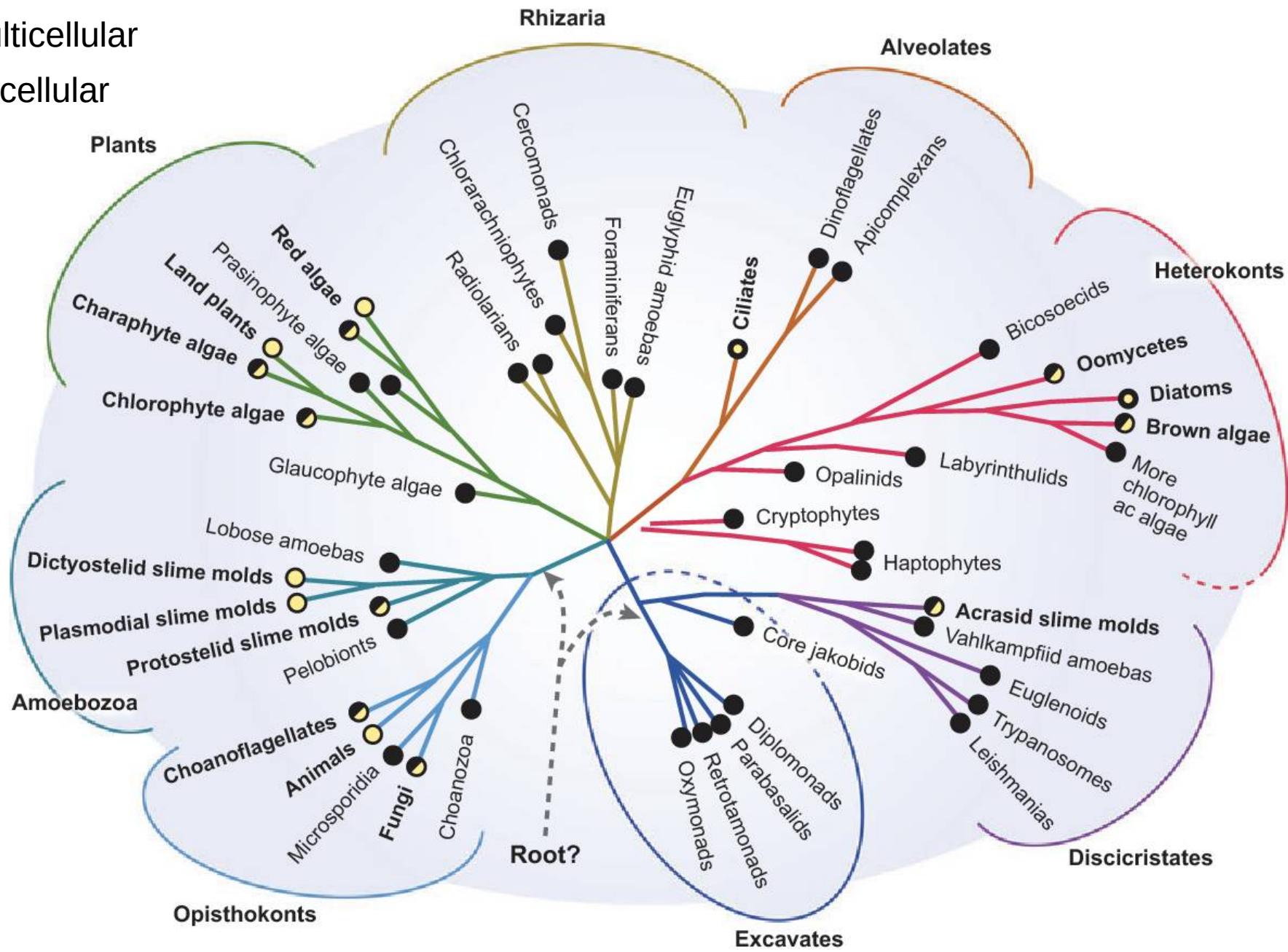
These values typically indicate the number of times out of 100 bootstraps that the branch was the same way in the best tree.



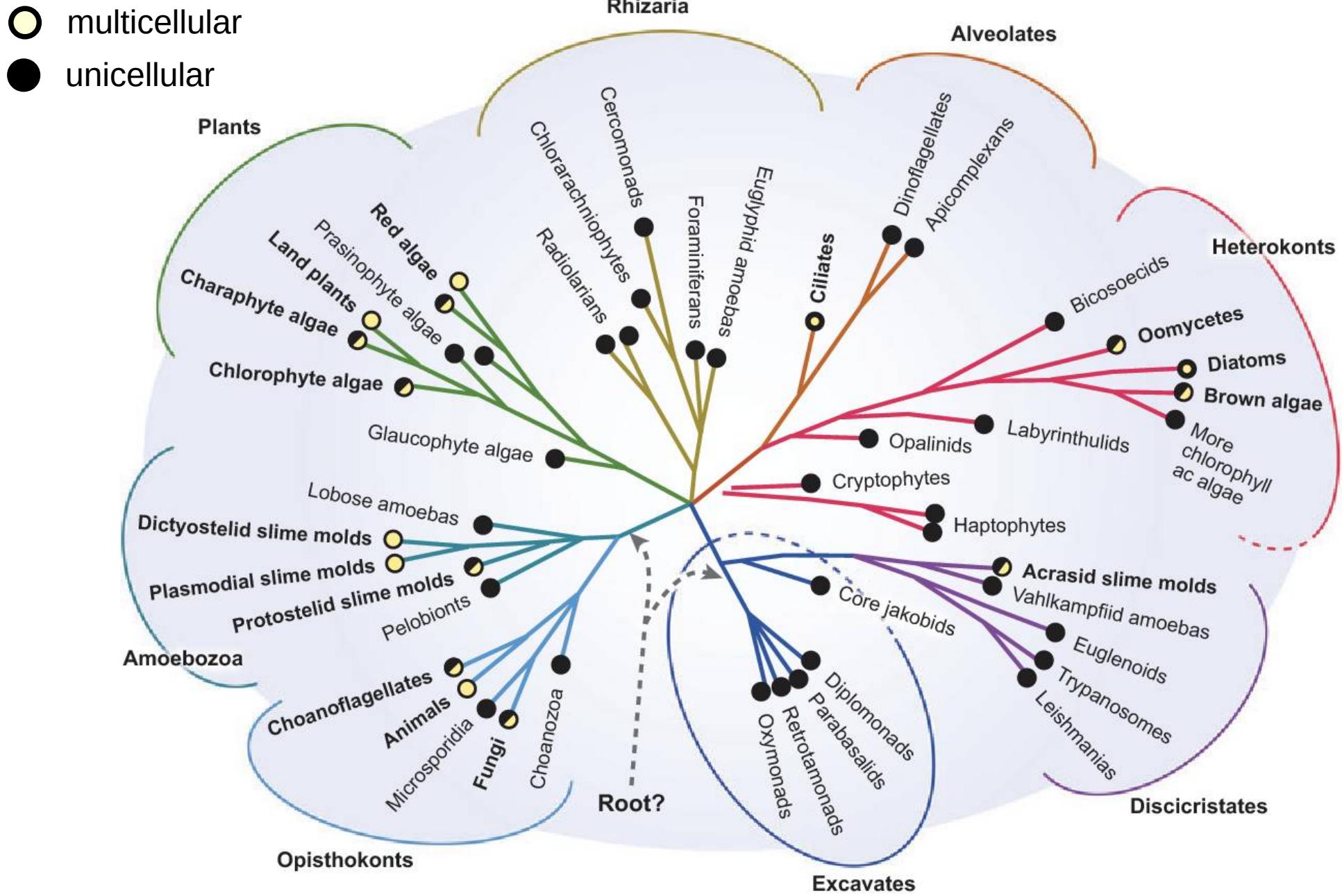
The 67 means that about 1/3 of the time



- multicellular
- unicellular



Multicellularity evolved multiple times in Eukaryotes



Taxonomy

Kingdom: Animalia
Phylum: Chordata
Class: Mammalia
Order: Primates
Suborder: Haplorhini
Infraorder: Simiiformes
Family: Hominidae
Subfamily: Homininae
Tribe: Hominini
Genus: Homo
Species: *Homo sapiens*

Linnaeus binomial name

<https://www.ncbi.nlm.nih.gov/Taxonomy/>

The screenshot shows the NCBI Taxonomy Browser interface. At the top, there's a navigation bar with links like Entrez, PubMed, Nucleotide, Protein, Genome, Structure, PMC, Taxonomy, and BioCollections. Below the navigation bar is a search bar with the placeholder "Search for" and a dropdown menu set to "complete name". There are also buttons for "lock", "Go", and "Clear". Under the search bar, there's a "Display" section where "3 levels using filter: none" is selected. A large grid of checkboxes follows, categorized by type: Nucleotide, Gene, Bio Project, PubChem BioAssay; Protein, HomoloGene, Bio Sample, Bio Systems; Structure, SRA Experiments, LinkOut, Assembly; Genome, BLAST, dbVar; Popset, SNP, GEO Profiles, Genetic Testing Registry; Conserved Domains, Protein Clusters, Host; GEO Datasets, Identical Protein Groups, Viral Host; Chordata, GEO Profiles, Host; and PubMed Central, SPARCLE, Probe. Below the filter section, the "Lineage (full)" is listed, showing the taxonomic hierarchy from cellular organisms up to Homo. Under the Homo entry, two sub-entries are shown: Homo sapiens (human) and Homo sapiens neanderthalensis (Neandertal).
Lineage (full): cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Euarchontoglires; Primates; Haplorrhini; Simiiformes; Catarrhini; Hominoidea; Hominidae; Homininae; **Homo**
o **Homo sapiens** (human) Click on organism name to get more information.
▪ **Homo sapiens neanderthalensis** (Neandertal)
▪ **Homo sapiens subsp. 'Denisova'** (Denisova hominin)

Disclaimer: The NCBI taxonomy database is not an authoritative source for nomenclature or classification - please consult the relevant scientific literature for the most reliable information.

Comments and questions to info@ncbi.nlm.nih.gov

[Help] [Search] [NLM NIH] [Disclaimer]



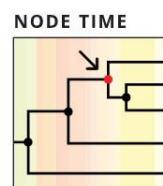
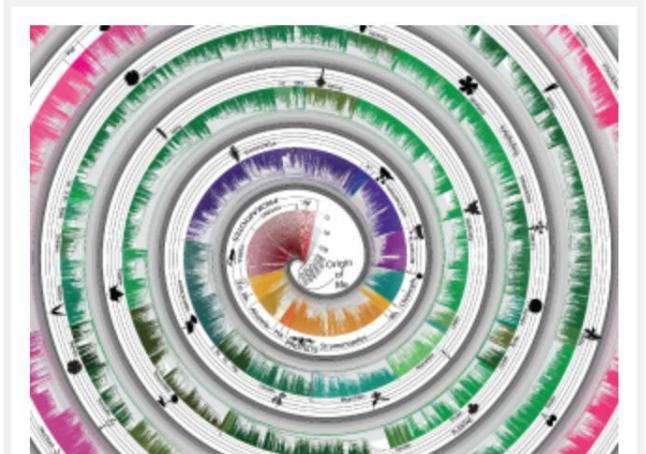
Search

www.timetree.org

TimeTree is a public knowledge-base for information on the evolutionary timescale of life. Data from thousands of published studies are assembled into a searchable tree of life scaled to time. Three search modes are possible:

- **NODE TIME** - to find the divergence time of two species or higher taxa
- **TIMELINE** - to drill back through time and find evolutionary branches from the perspective of a single species
- **TIMETREE** - to build a timetree of a group of species or custom list

TIMEPANELS showing events in geological time and astronomical history are provided for comparison with timelines and timetrees. Results can be exported in different formats for additional analyses and publication.



GET DIVERGENCE TIME FOR A PAIR OF TAXA

Specify 2 Taxon Names ?

Taxon 1:

Taxon 2:

Clear

Search

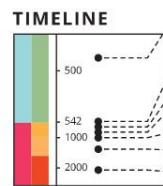


Resolve Ambiguity ?

Taxon 1:

Taxon 2:

Show Time



GET AN EVOLUTIONARY TIMELINE FOR A TAXON

Specify a Taxon Name ?

Taxon:

Clear

Search

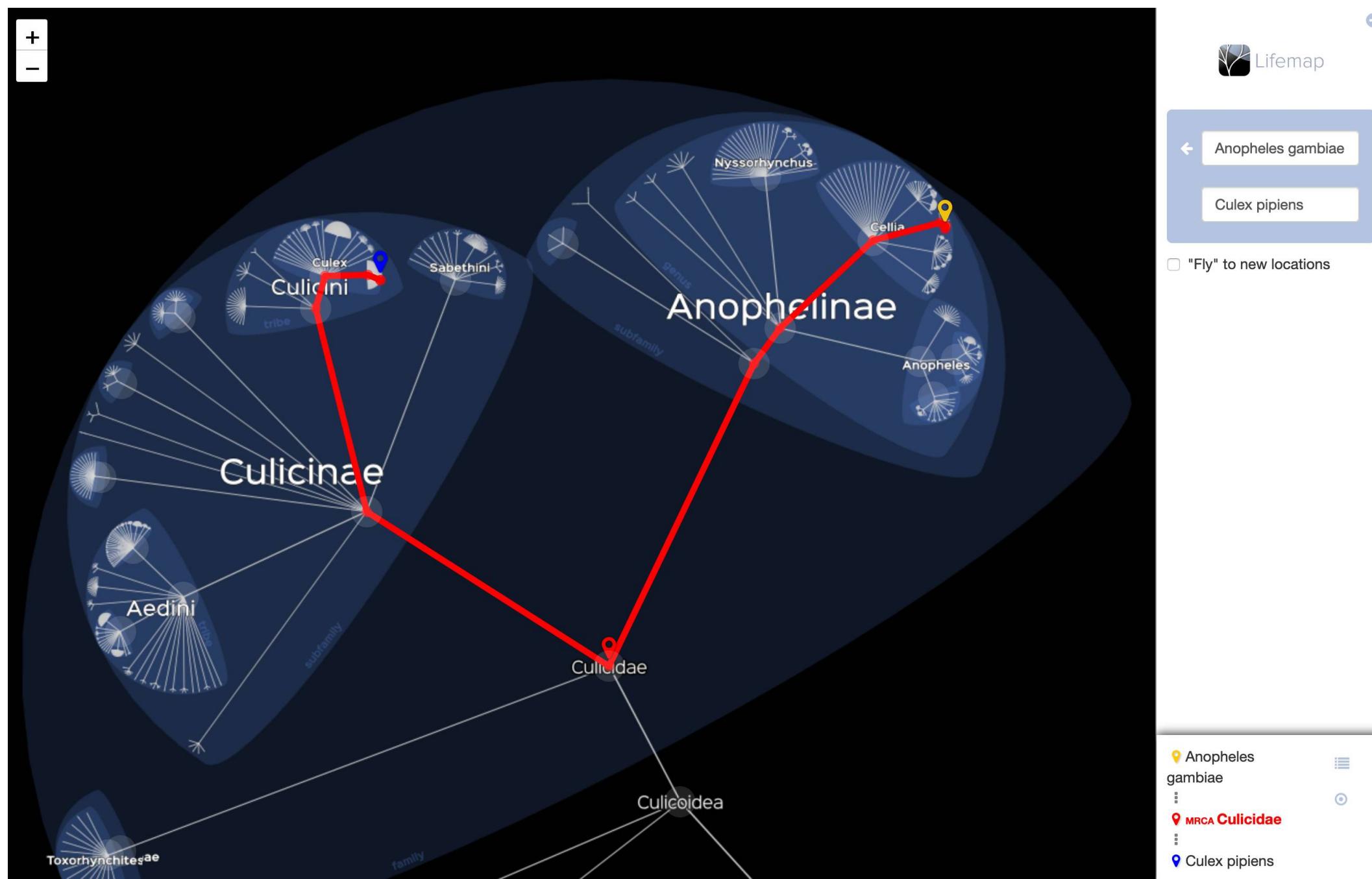


Resolve Ambiguity ?

Taxon

Show Timeline

<http://lifemap.univ-lyon1.fr>





Hominidae

Humans, great apes, and their extinct relatives



temporary page

Hominidae

- ▶ page content
- ▶ articles & notes
- ▶ treehouses
- ▶ collections
- ▶ people
- ▶ options



Explore Other Groups

- ▶ other Catarrhini
- ▶ containing groups
- ▶ subgroups

[random page](#)



Phylogeny in part from Purvis (1995).

Containing group:  [Catarrhini](#)

Other Names for Hominidae

Humans, great apes, and their extinct relatives

References

<https://blast.ncbi.nlm.nih.gov/Blast.cgi>

NIH U.S. National Library of Medicine NCBI National Center for Biotechnology Information

BLAST® Home Recent Results Saved S

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](#)

N
E
W
S

End of updates for BLAST+ version 4 databases (dbV4)

Start moving to the new version 5 databases!

Fri, 27 Sep 2019 16:00:00 EST

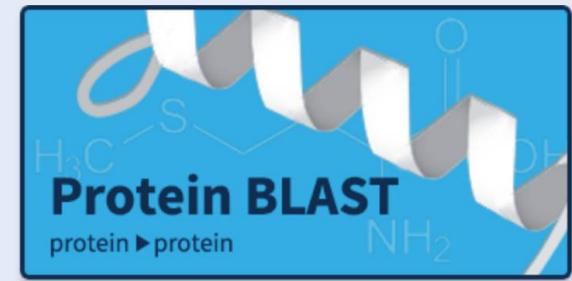
[More BLAST news...](#)

Web BLAST



blastx
translated nucleotide ▶ protein

tblastn
protein ▶ translated nucleotide



BLAST Genomes

Enter organism common name, scientific name, or tax id

Search

Human

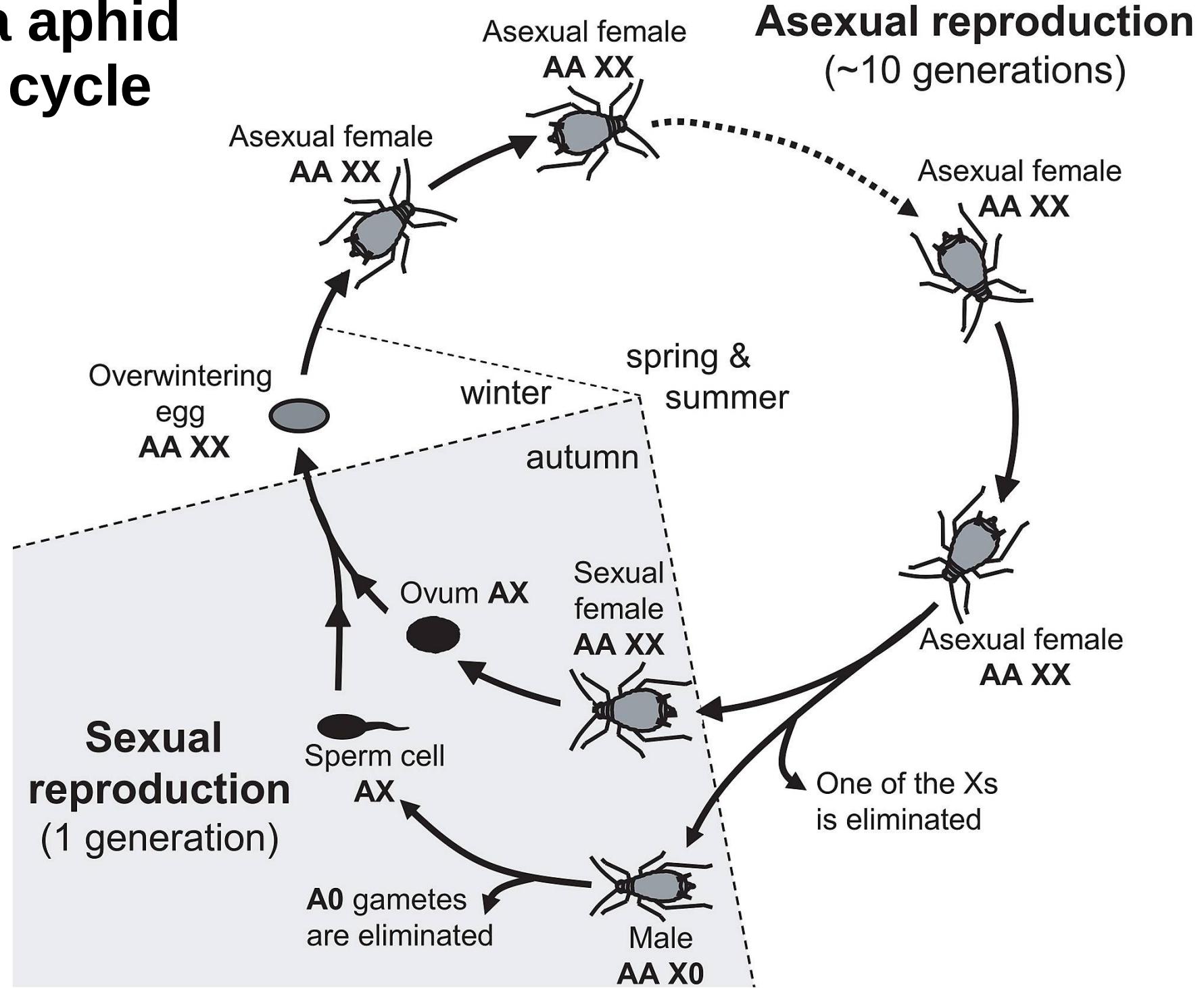
Mouse

Rat

Microbes

À vous de jouer !

Pea aphid life cycle



Program	TBLASTN ?	Citation ▾
Database	nt	See details ▾
Query ID	Icl Query_96735	
Description	None	
Molecule type	amino acid	
Query Length	602	
Other reports		?

Type common name, binomial, taxid or group name

+ Add organism

Percent Identity E value

[] to [] [] to []

[Filter](#) [Reset](#)

[Descriptions](#)

[Graphic Summary](#)

[Alignments](#)

[Taxonomy](#)

Sequences producing significant alignments

[Download ▾](#)

[Manage Columns ▾](#)

Show

100 ▾



select all 7 sequences selected

[GenBank](#) [Graphics](#)

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase (LOC100161104), transcript variant X1, mRNA	249	249	98%	2e-71	28.43%	XM_001943135.5
<input checked="" type="checkbox"/>	PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase (LOC100161104), transcript variant X2, mRNA	249	249	98%	6e-71	28.43%	XM_016801018.2
<input checked="" type="checkbox"/>	PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase (LOC100574964), mRNA	245	245	99%	1e-70	28.87%	XM_003241620.4
<input checked="" type="checkbox"/>	PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase-like (LOC100571777), mRNA	213	213	83%	2e-61	30.53%	XM_008185904.3
<input checked="" type="checkbox"/>	PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase (LOC100159332), mRNA	206	206	99%	2e-56	26.70%	XM_001950752.5
<input checked="" type="checkbox"/>	Acyrthosiphon pisum ACYPI005179 mRNA, clone: 3815, complete cds, full-insert cDNA sequence based on the ESTs (5'-EST:EX621862_3)	102	188	53%	6e-41	34.84%	AK340189.1
<input checked="" type="checkbox"/>	PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase-like (LOC103308018), mRNA	86.3	86.3	55%	1e-17	25.85%	XM_008180624.2

[Descriptions](#)[Graphic Summary](#)[Alignments](#)[Taxonomy](#)

Alignment view

Pairwise

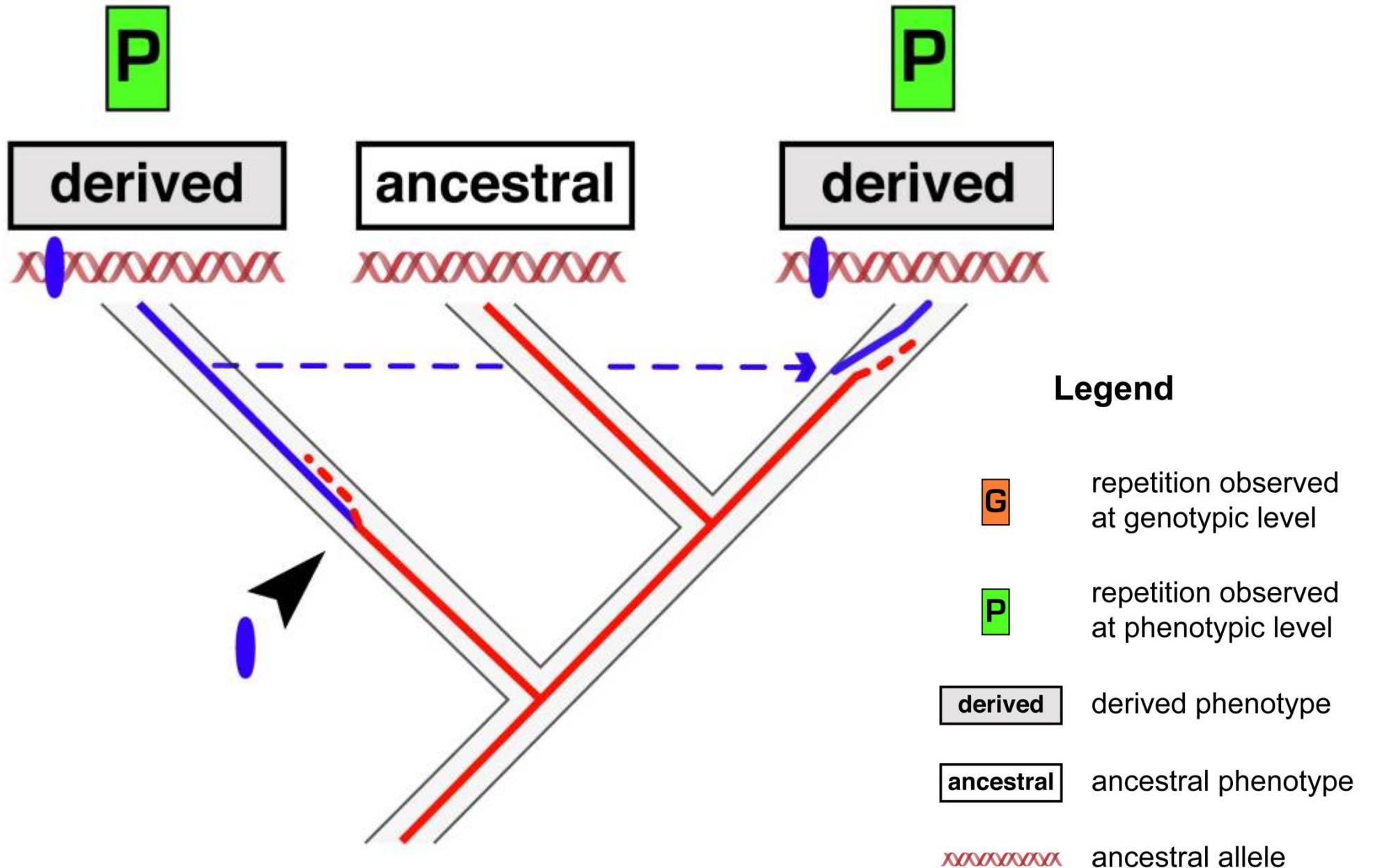
[Download](#) ▾7 sequences selected [?](#)[Download](#) ▾ [GenBank](#) [Graphics](#)▼ [Next](#) ▲ [Previous](#) [◀ Descriptions](#)**PREDICTED: Acyrthosiphon pisum bifunctional lycopene cyclase/phytoene synthase (LOC100161104), transcript variant X1, mRNA**Sequence ID: [XM_001943135.5](#) Length: 2952 Number of Matches: 1Range 1: 191 to 1924 [GenBank](#) [Graphics](#)▼ [Next Match](#) ▲ [Previous Match](#)

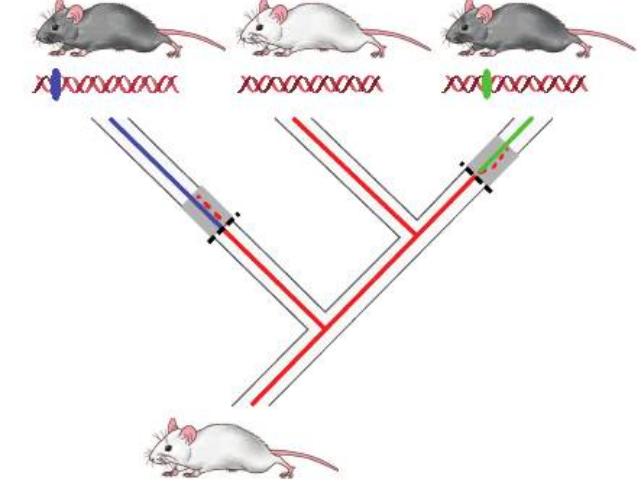
Score	Expect	Method	Identities	Positives	Gaps	Frame
249 bits(635)	2e-71	Compositional matrix adjust.	174/612(28%)	300/612(49%)	54/612(8%)	+2

Query	1	MYDYAFVHLKFTVPAAVLLTAIAYPILNRHILQQTGFLVVVAFTAALPWDAYLIKHKVWS	60
Sbjct	191	M L T Y I D V H F I Y T L P V V A V L A L I T W P F I S R L E L F K I G F V C T M A F V Y T T P W D N Y I I F H N A W M	370
Query	61	Y P P E A I V G P R L L G - I P F E E L F F F V I Q T Y I T A L V Y I L F N K P V L H A L H L N N Q Q N P P A W M R V V	119
Sbjct	371	Y K P K N I L A -- V I G Y V P V E E Y M F F V I Q T L M T S L W A L V F T R W S P A C F N F N K T S Y T L I R W I	544
Query	120	K V T G Q V V L V A L S V W G W N A A Q V H Q E T S Y L G L I L V W A C P F L L A I W T I L A G R F I L S L P W Y A T V L	179
Sbjct	545	P I --- L A L V M T T I Q G Y N I A V P G K N T F Y L G C I M W W S C P V I M F L W Y G A G N Y F V K K S - T S S A I	712
Query	180	P M F L P T F Y L W A V D E F A L H R G T W S I G S G T K L D F C L F G K L D I E E A T F F L V T N M L I V G G M A A F	239
Sbjct	713	A V I V P T L Y L C W V D R I A L K D D V W H I N E K T S L N I F V V D D L P F E E C L F F L I T N V I I V L G G M A F	892
Query	240	D Q Y L A V I Y A F P T L F P K V N R Y P T T H M L L Q S R L I N T S R Y D L E R - - - I E G L R E A V E R L R L K S	295
Sbjct	893	D K S Y G L A D T Y T F E F P L - - - R Y S S S W K Y Y S Q Q M Q Q F V R A E C D M S P S P V N D I R Q C L N V L K R A S	1066
Query	296	R S F Y L A N S L F S G R L R I D L L I L L Y S F C R L A D D L V D D A K S R R E V L S W T A K L - N H F L D L H Y K D A	354
Sbjct	1067	K S F N V A S L V F P A G V R L H L I I L Y A F C R V T D D M I - D S E P K V G V K K Q K L K L I E T F I D E L F A D R	1243
Query	355	D A T E D - - - - - P K K K A E R I D A Y - - - I K T A F P P C A Y Q A L H L L P T H I L P P K P L Y D L I K G F E M D	406
Sbjct	1244	S A D Y D V K T S M T P R K P E V K W E O Y R L D L T D E E L S C F R A I S R I S F Y - L P R K P F Y E L L D G Y R W D	1420

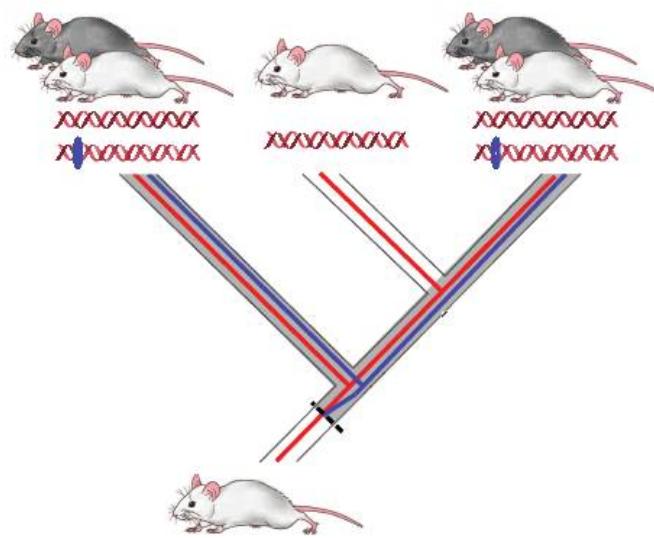
Related Information[Gene](#) - associated gene details[Genome Data Viewer](#) - aligned genomic context

Lateral transfer

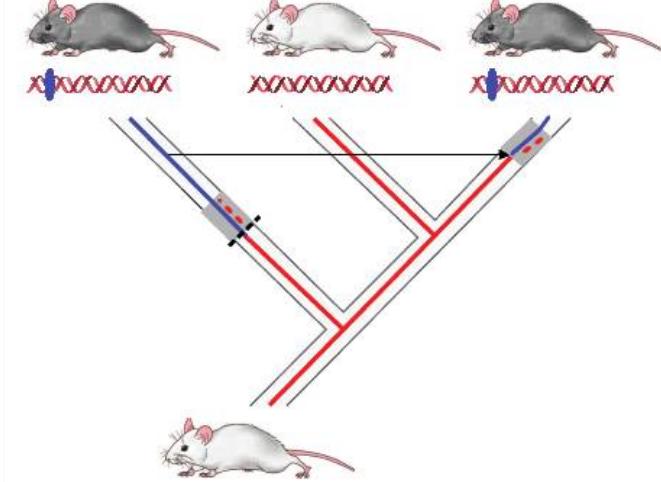




Genetic convergence



Ancestral polymorphism



Lateral transfer

OÙ ILS SE SONT CROISÉS

- Pelage : **brun**
- Taille : **1,70 à 2,80 m**
(tête et corps)
- Membres : **griffes non rétractiles**
- Régime alimentaire : **omnivore**
- Lieu de vie : **forêts, zones côtières, montagnes**
- Période d'accouplement : **mai à juillet**



- Pelage : **blanc**
- Taille : **1,80 à 3 m**
(tête et corps)
- Membres : **doigts partiellement palmés**
- Régime alimentaire : **carnivore**
- Lieu de vie : **banquise**
- Période d'accouplement : **avril à juin**



Grizzly ou ours brun

Ours polaire

Pizzly ou grolar
Le pizzly, encore appelé grolar ou prizzly, a été trouvé sur l'île de Banks, au nord-ouest du Canada, en 2006.



Brown Bears

1. Mainland Alaska ♀
2. ABC Islands ♀

Polar bears

3. West Hudson Bay ♀
4. West Hudson Bay ♂
5. North Beaufort Sea ♂
6. Lancaster Sound ♂
7. South Beaufort Sea ♂
8. Wrangel Island ♀
9. Chukchi Sea ♂

The enigmatic ABC Islands brown bears likely derive from a population of polar bears likely stranded by the receding ice at the end of the last glacial period. Since then, male brown bear migration onto the island has gradually converted these bears into an admixed population whose phenotype and genotype are principally brown bear. This process of genome erosion and conversion may be a common outcome when climate change or other forces cause a population to become isolated and then overrun by species with which it can hybridize.

The genes and the mutations responsible for phenotypic evolution

Evolution repeats itself

www.Gephebase.org

>2000 entries



@gephebase

Gephebase compiles genotype-phenotype relationships, i.e. associations between a mutation and a phenotypic variation. Gephebase consolidates data from the scientific literature about the genes and the mutations responsible for phenotypic variation in Eukaryotes (mostly animals, yeasts and plants). We plan to include non Eukaryote species in the future. For now, genes responsible for human disease and for aberrant mutant phenotypes in laboratory model organisms are excluded and can be found in other databases ([OMIM](#), [OMIA](#), [FlyBase](#), etc.). QTL mapping studies that did not identify single genes are not included in Gephebase.

If you use Gephebase for your publication, please cite: Martin, A., & Orgogozo, V. (2013). The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution*, 67(5), 1235- 1250.

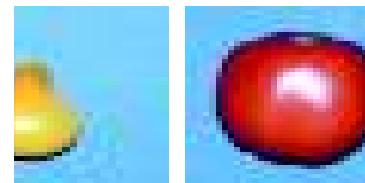
[Conference on Gephebase and the loci of evolution \(Paris, 2016\)](#)

You can retrieve data via HTTP requests through APIs. Below is the list of available APIs. By default, the response data is sent in xml format. For each field, it is possible to only enter a subset of a keyword and still be able to successfully retrieve the desired data. (example: "Bir" for "Birds" will display all data that have the characters "Bir").

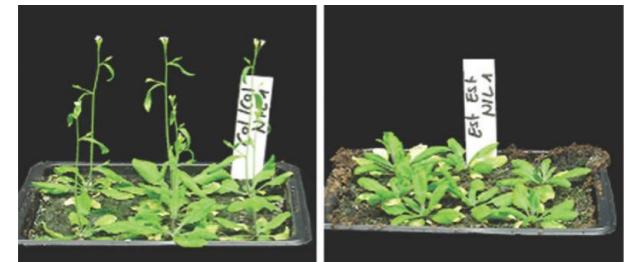
Wrinkled seed: TE insertion
(Bhattacharyya 1990)



OVATE coding region
(Liu 2002)



FRI coding region
(Johanson 2000)



myostatin coding region
(Grobet 1997)



luciferase coding region
(Stolz 2003)



Mc1r coding region
(Eizirik 2003)



anthocyanin-2 coding region
(Quattrocchio 1999)



THREE APPROACHES to FIND the GOLDEN LOCI of EVOLUTION



CANDIDATE GENE

REVERSE GENETICS

From genes to traits

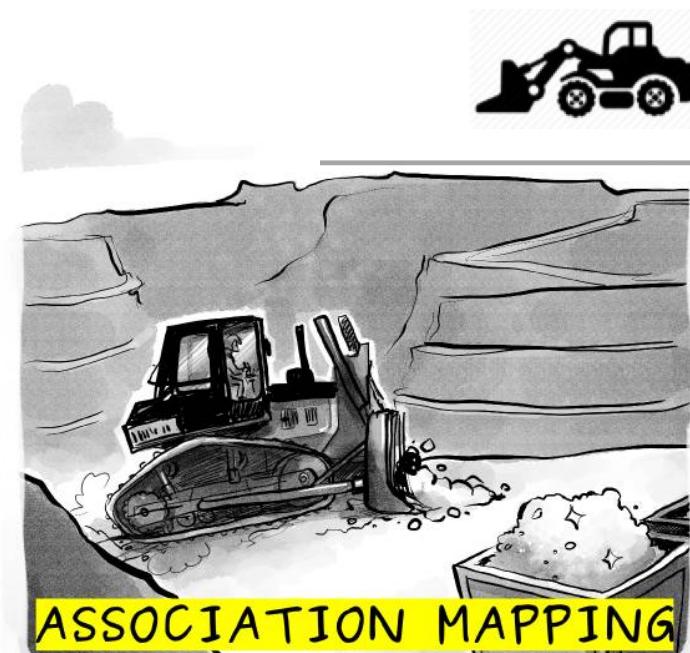


LINKAGE MAPPING

FORWARD GENETICS

From traits to genes

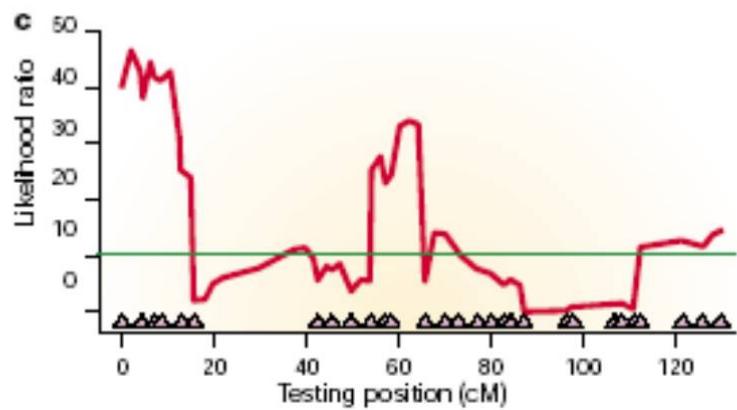
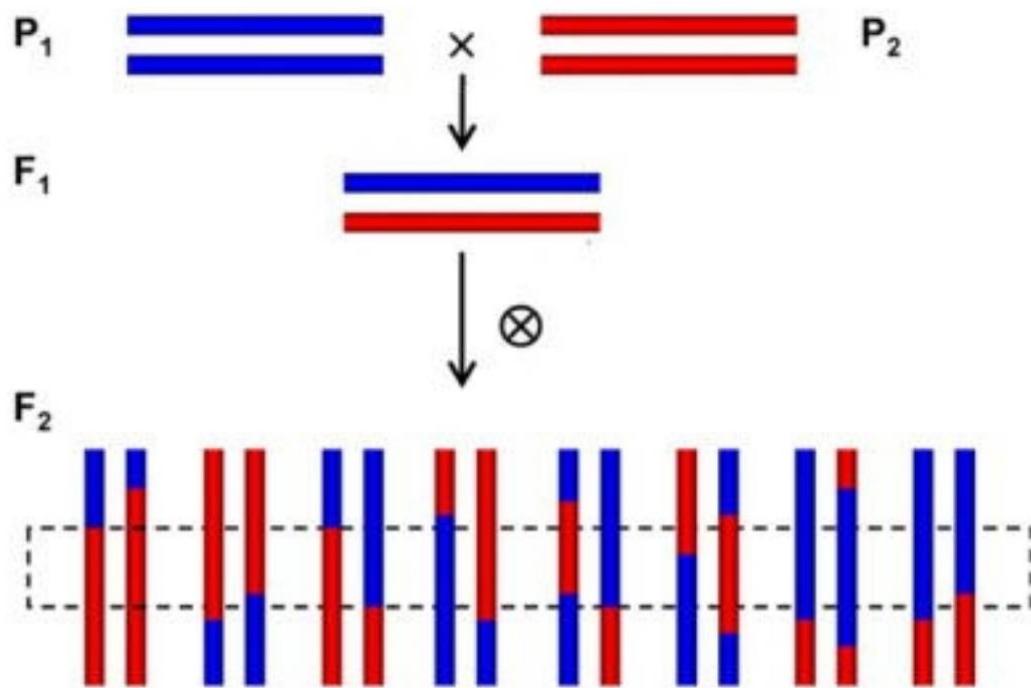
Little Ascertainment Bias, but Micro-Evolution only



ASSOCIATION MAPPING

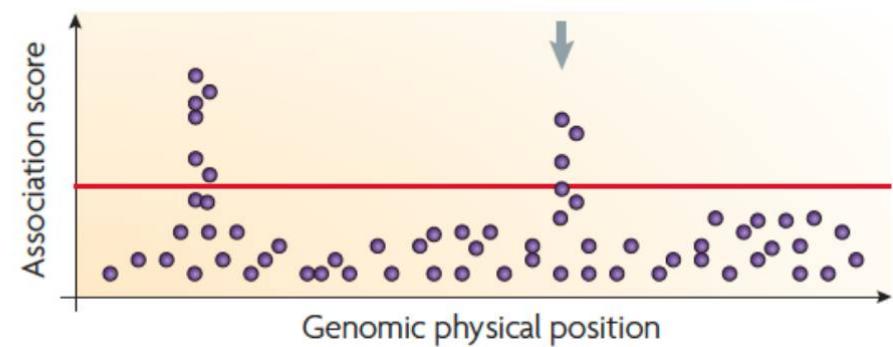
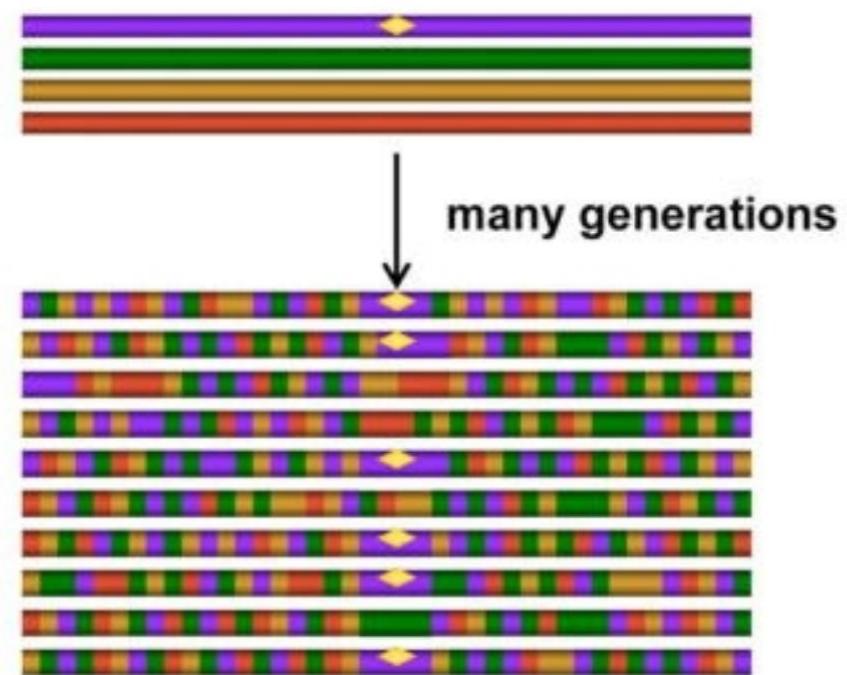
Linkage Mapping

Crosses in the lab

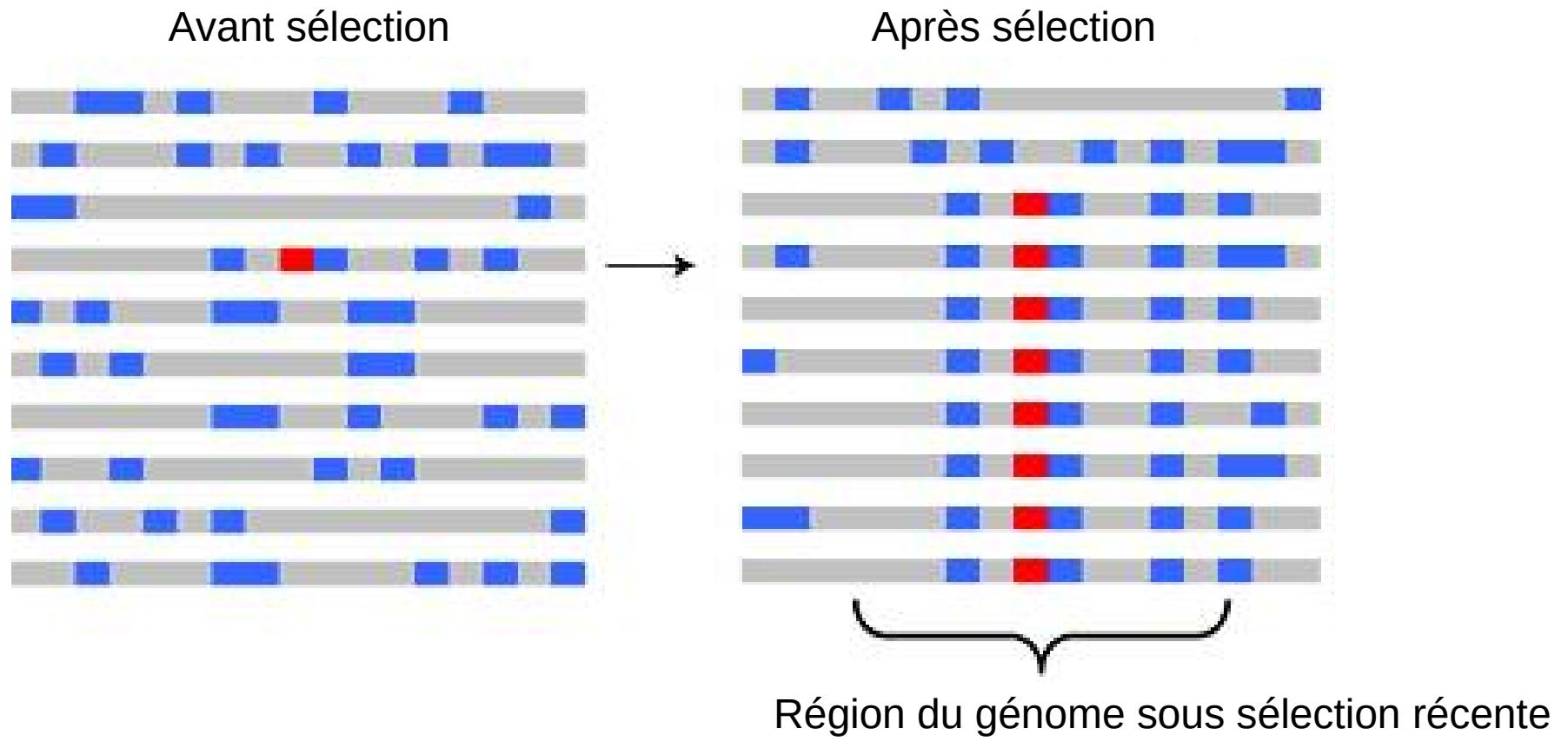


Association Mapping

Past crosses in natural populations

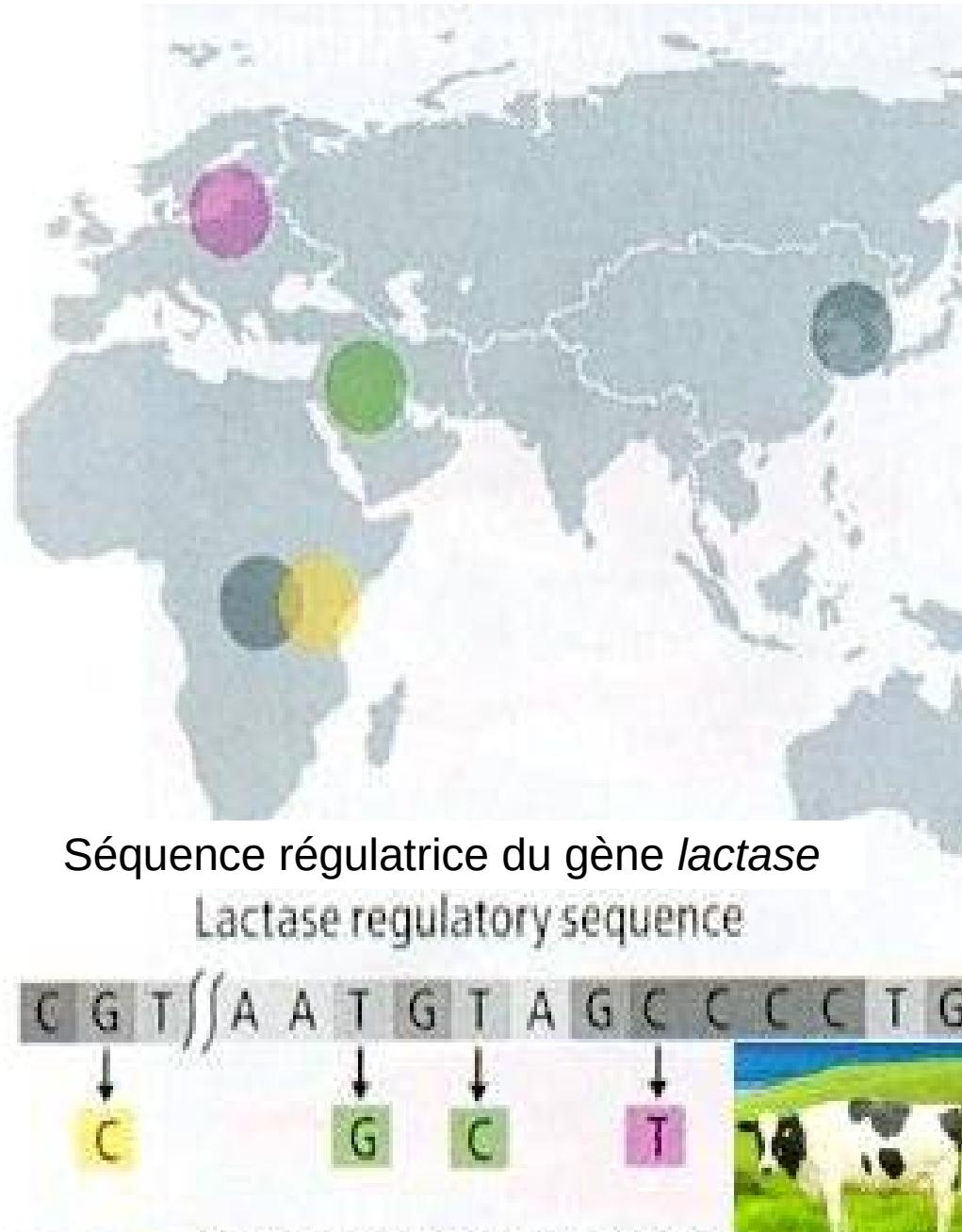


Des événements de sélection récente laissent des traces dans les génomes



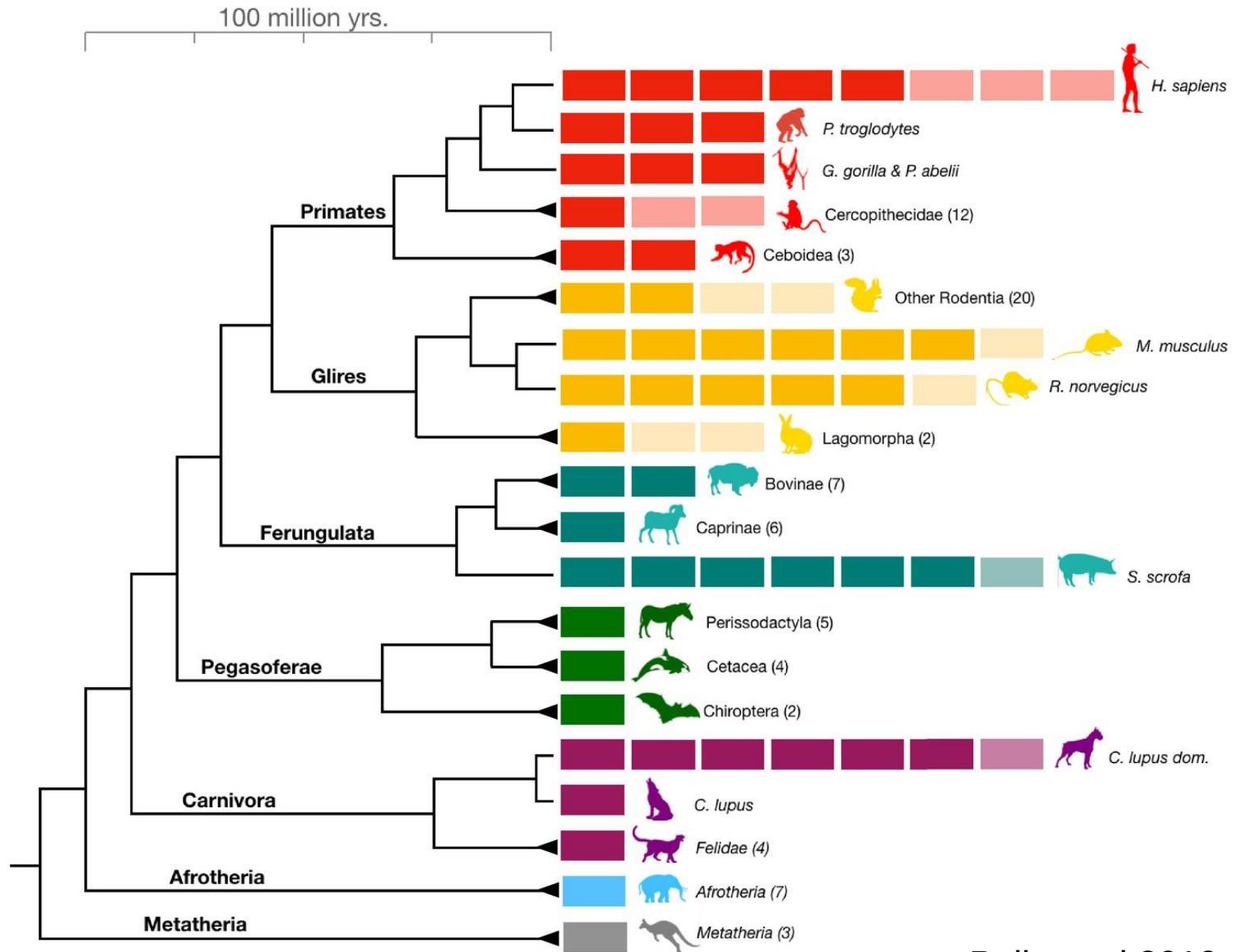
Ex: *lactase*, *CCR5-Δ32* et résistance au VIH

Repeated evolution in the *lactase* gene

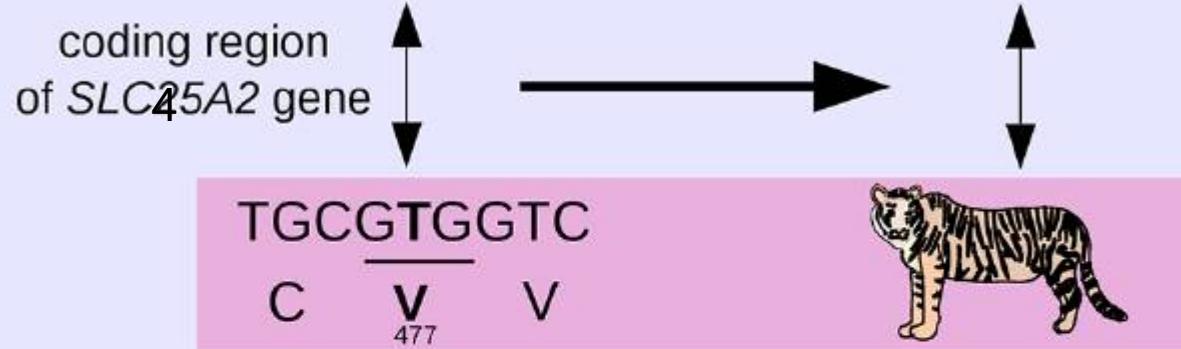


Digestion de l'amidon

Duplications répétées du gène amylose



Repeated evolution



Also in:
Humans
Horses
Quails
Chickens
Mice
Pigeons



Repeated evolution via the same amino acid change

clam



↑ Nav1.4
X E945D

Saxitoxin



toxic plancton

garter snake



↑ Nav1.4 sodium channel
X E945D

Tetrodotoxin



toxic newt

↑ Nav1.4
E945D



pufferfish (fugu)

Repeats in..



.. the genes responsible for natural evolution

Ex : *hemoglobin* in dogs and humans in Tibet
(Wang et al 2014 GBE)



.. the genes responsible for experimental evolution

Ex : *sulfate transporter SUL1* in yeasts in low sulfate
(Gresham et al 2008 PloS Genetics)



.. the phenotypes (convergent evolution)

Ex : thylacines and dogs, water chevrotain

From random processes can emerge predictability

Many unpredictable processes
at a low level



Predictable Evolution
at the genetic level

- Mutations in DNA
- Chromosome segregation during meiosis
- Assortative mating
- Gamete competition during fecondation
- Life history traits
- Genetic linkage
- Environmental changes (meteorite, etc.)

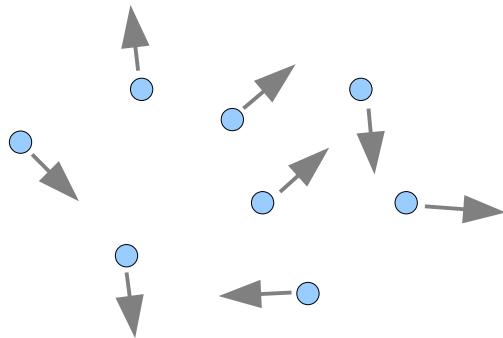
...



coding mutation in
VKORC1 gene

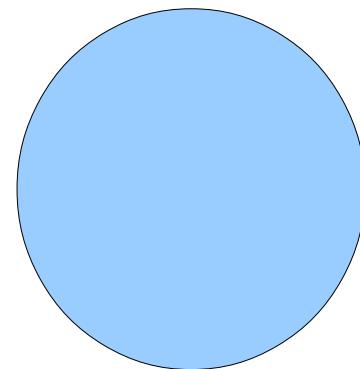
From random processes can emerge predictability

Microscopic world

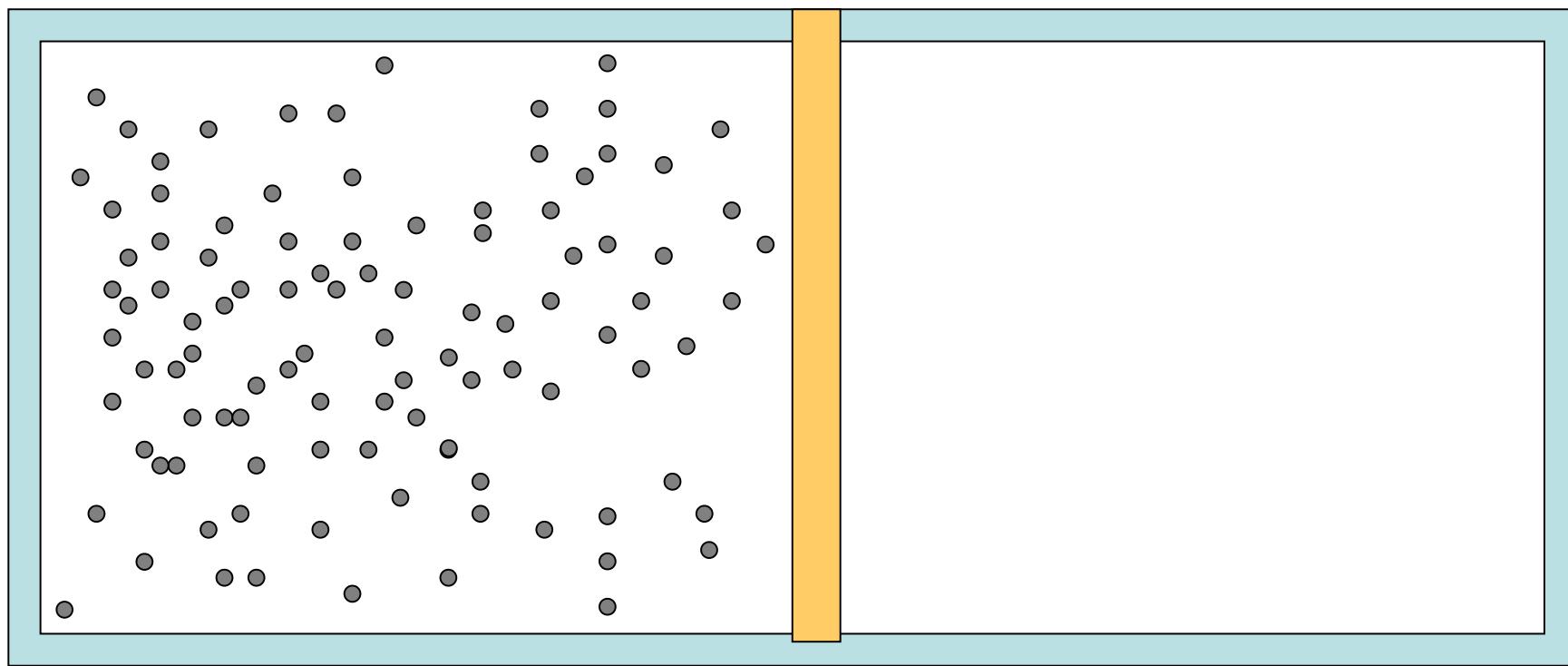


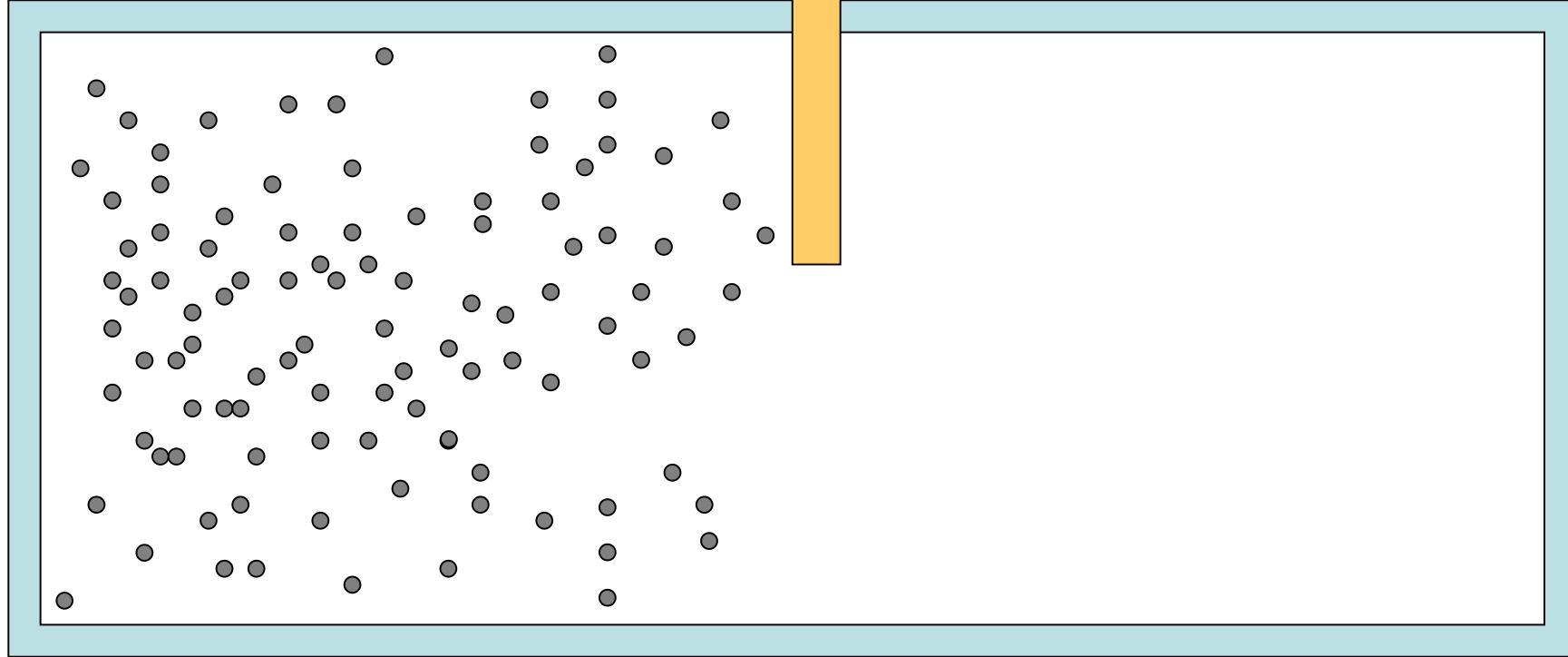
Position, mass, velocity of each particle

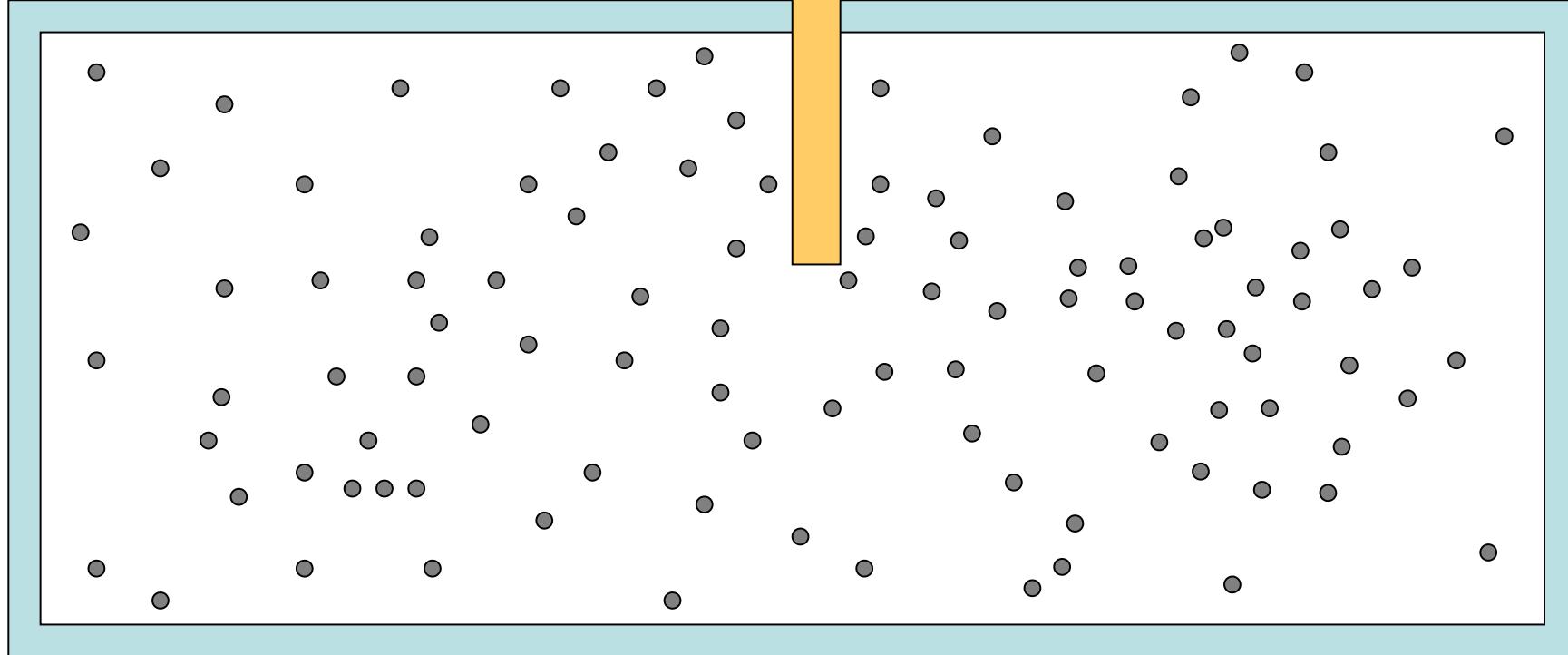
Macroscopic world



Pressure, Volume, Temperature,
Number of moles







After a few seconds

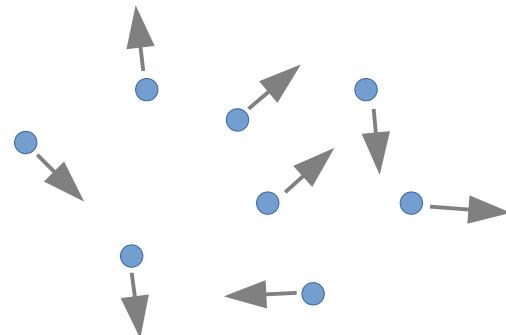
Predictable evolution at a higher level?

Many unpredictable processes
at a low level



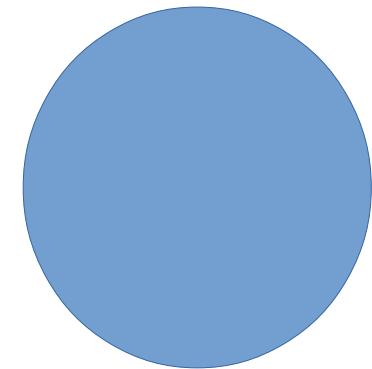
Predictable Evolution
at higher level?

Microscopic world

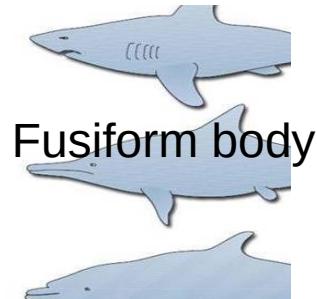


Position, mass, velocity of each particle

Macroscopic world

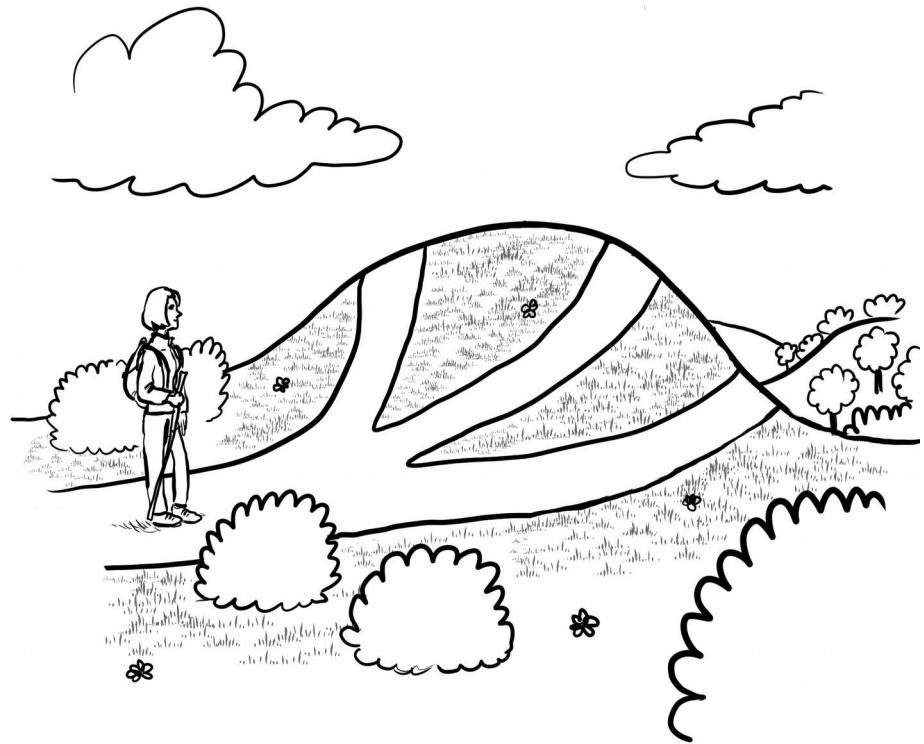
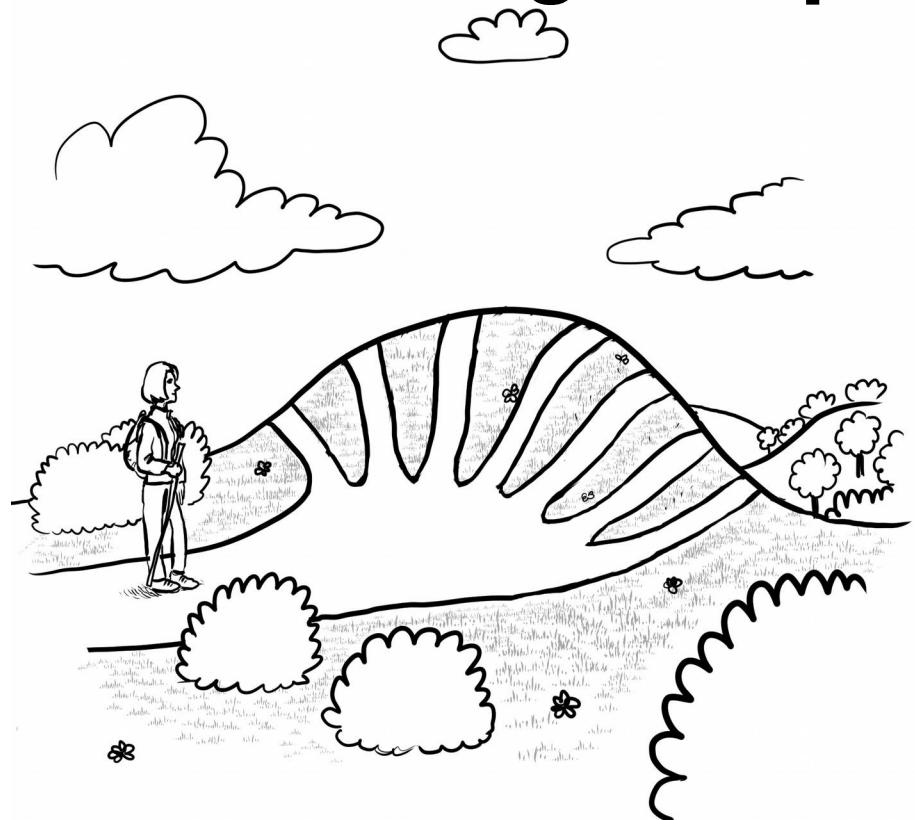


Pressure, Volume, Temperature,
Number of moles



**Key: finding the relevant
concepts for predictive
evolution**

A small number of genetic solutions for a given phenotypic change



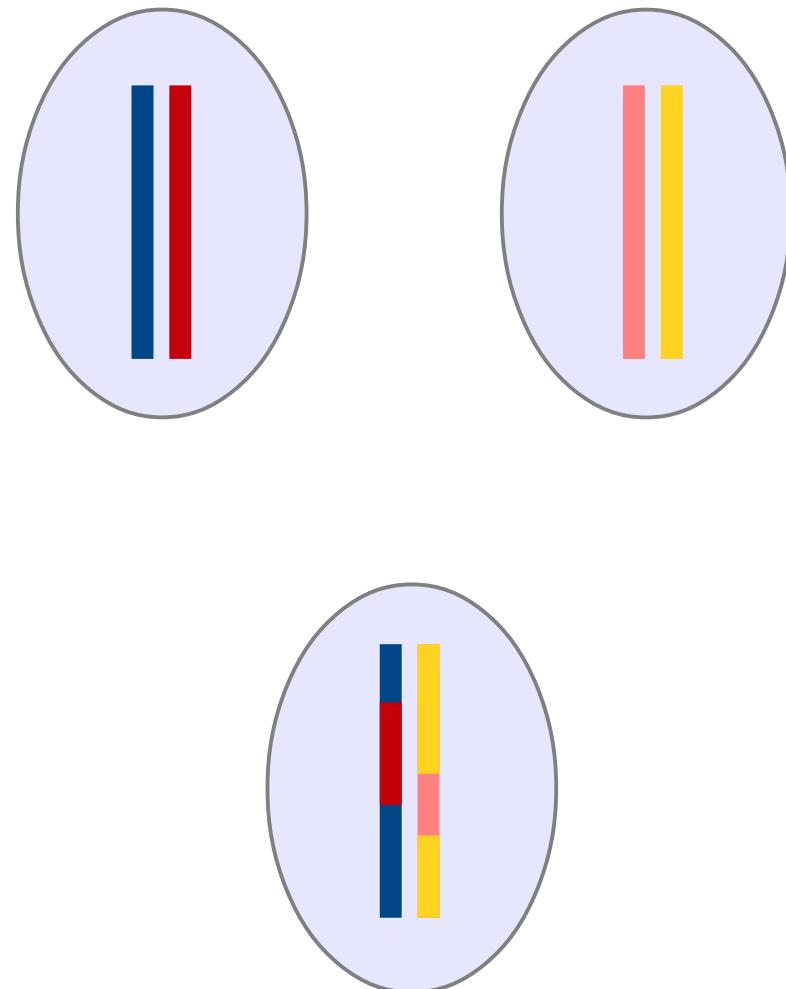
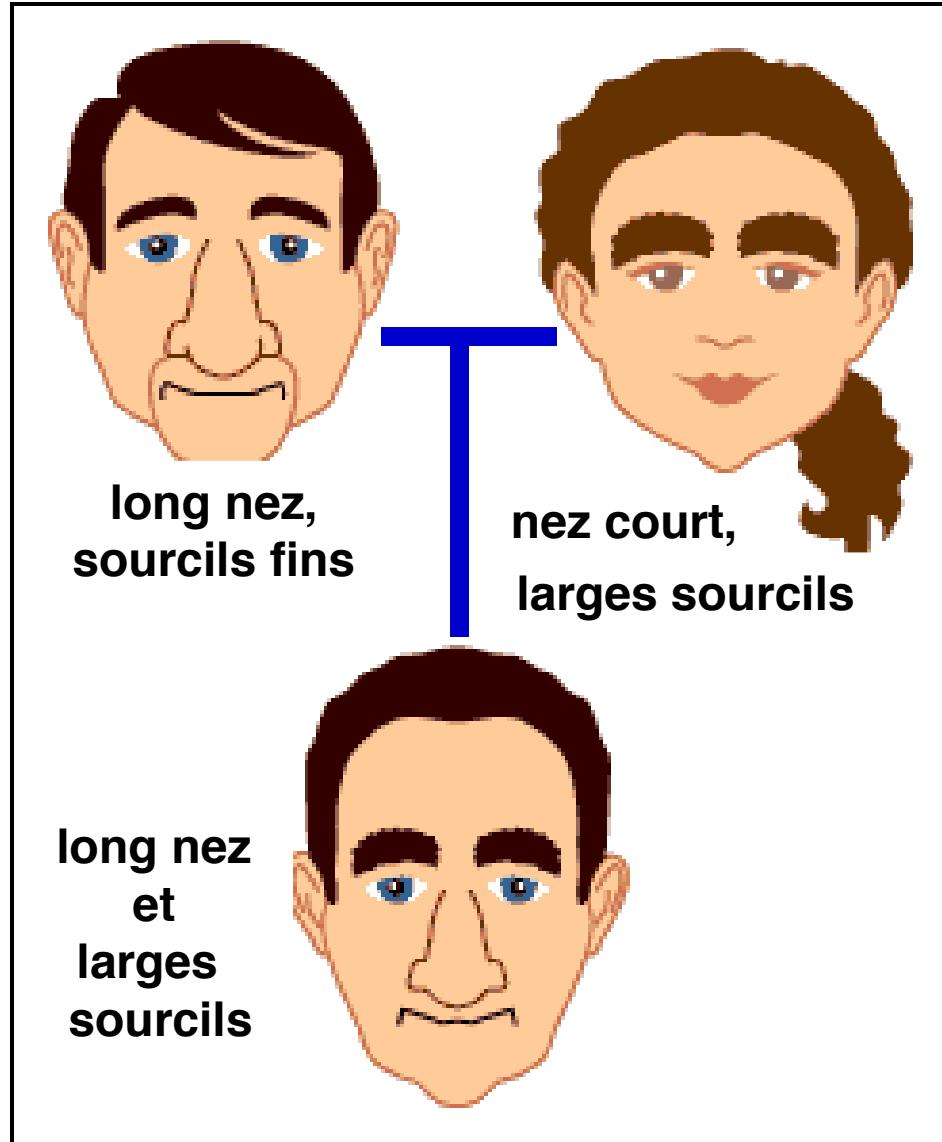
Conclusion: on ne peut pas prévoir dans quelle direction se fera l'évolution, il y a trop d'événements impossibles à prévoir qui interviennent.

Michael DuBow

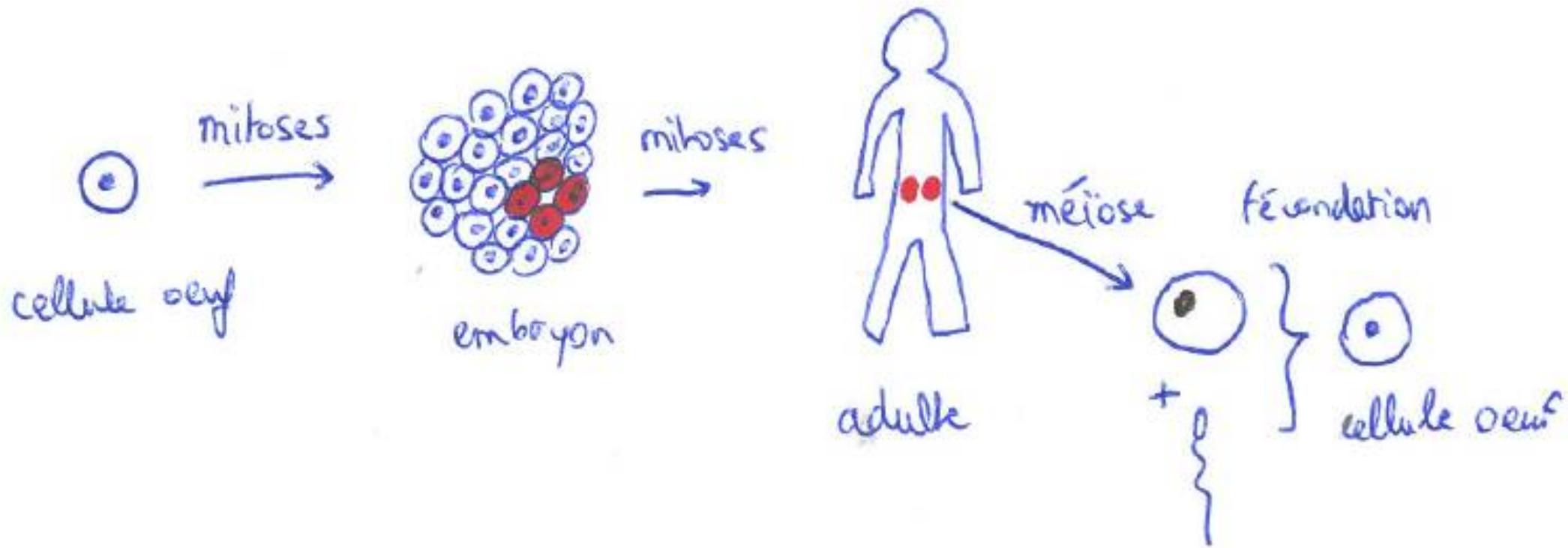


We are mosaics

Assortiment des chromosomes du père et de la mère



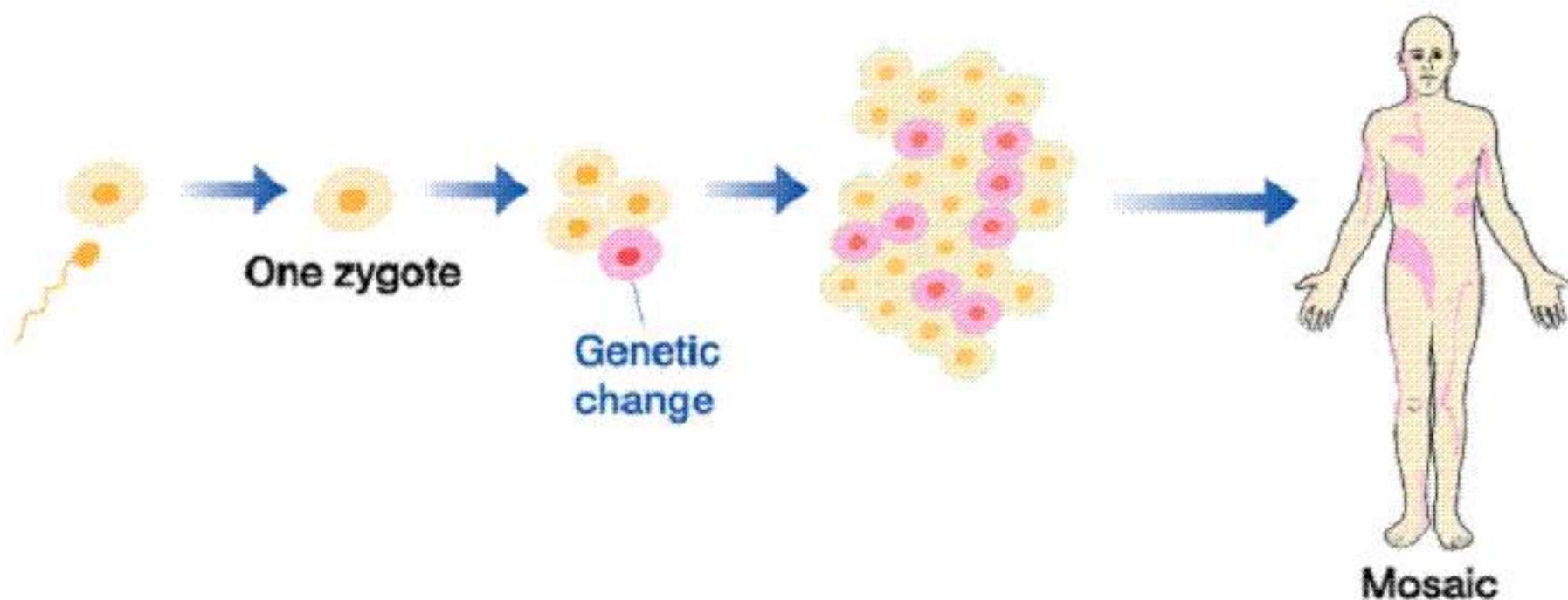
Le génome au cours des générations

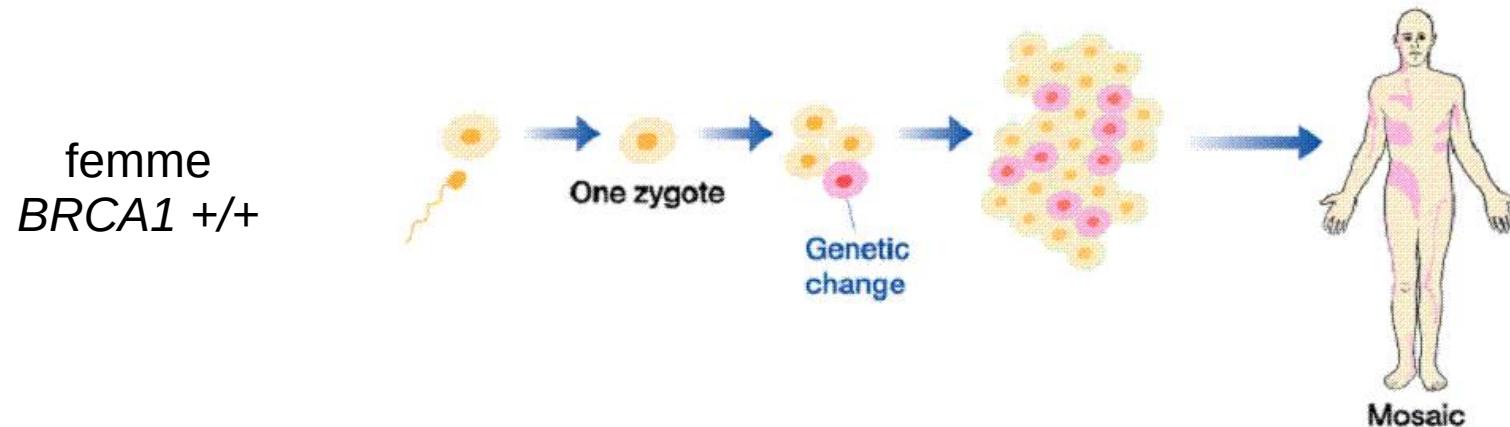


Lignée germinale

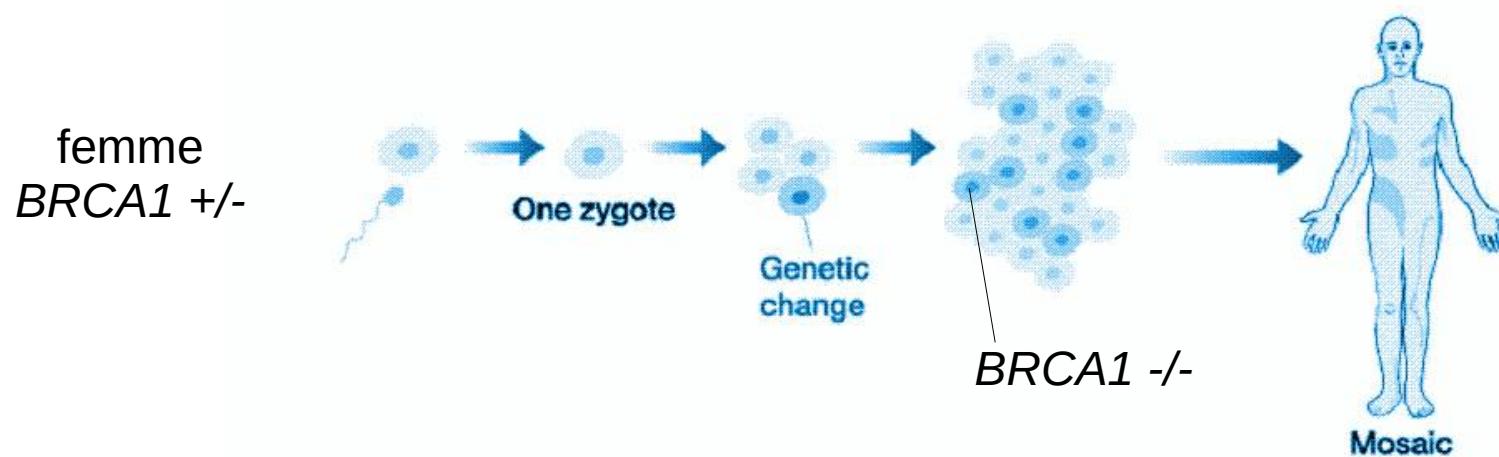


Lignée somatique





10% de chance de développer un cancer du sein durant sa vie

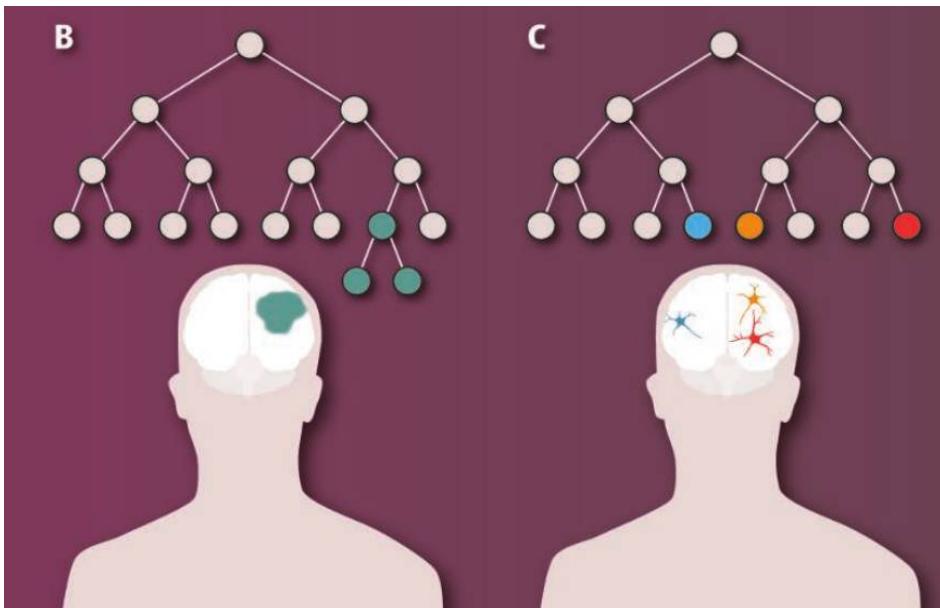


45% de chance de développer un cancer du sein avant 70 ans
Les cellules de ce cancer sont *BRCA1* -/-

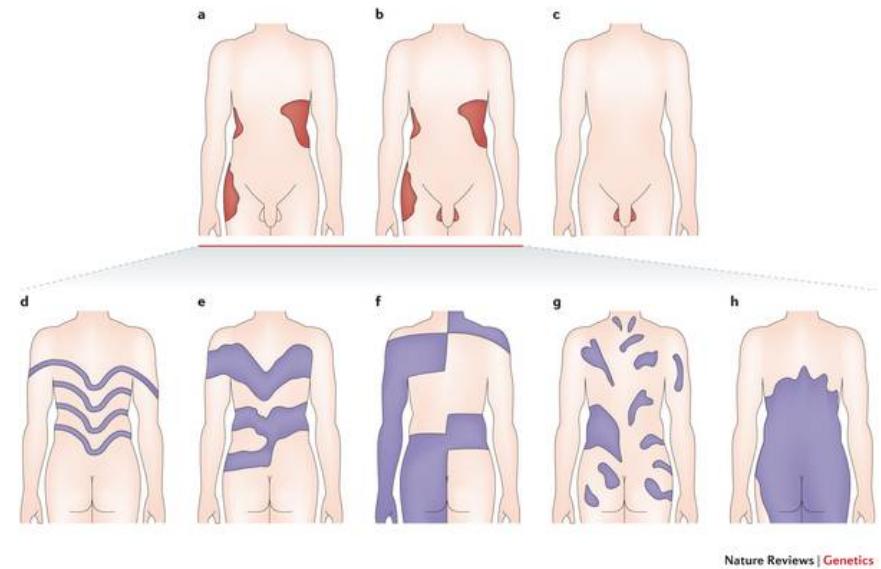
Mosaïque somatique

Dans l'hippocampe et le noyau caudé de 3 individus :

7,743 somatic L1 insertions, 13,692 somatic Alu insertions and 1,350 SVA insertions

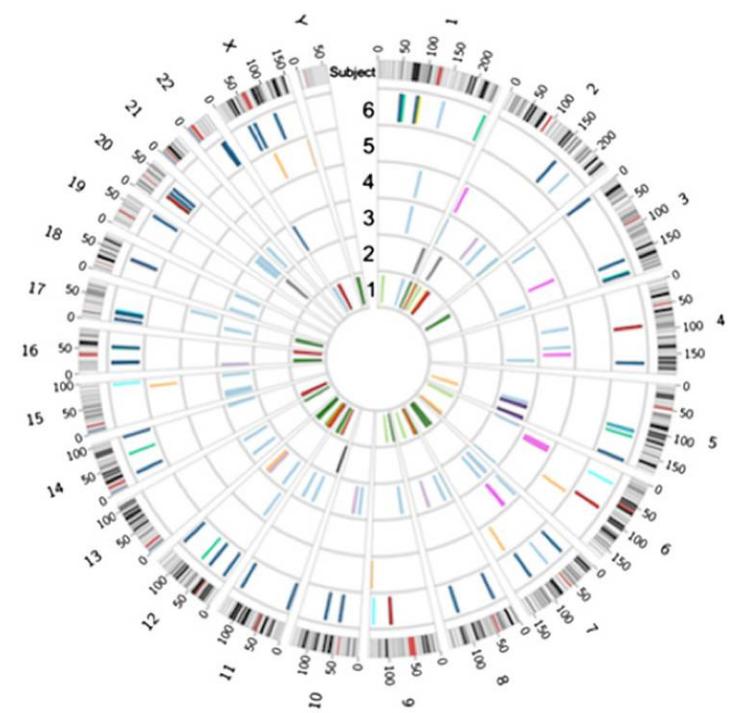


Baillie 2011 Nature



Nature Reviews | Genetics

73 somatic CNVs in 11 tissues of six persons

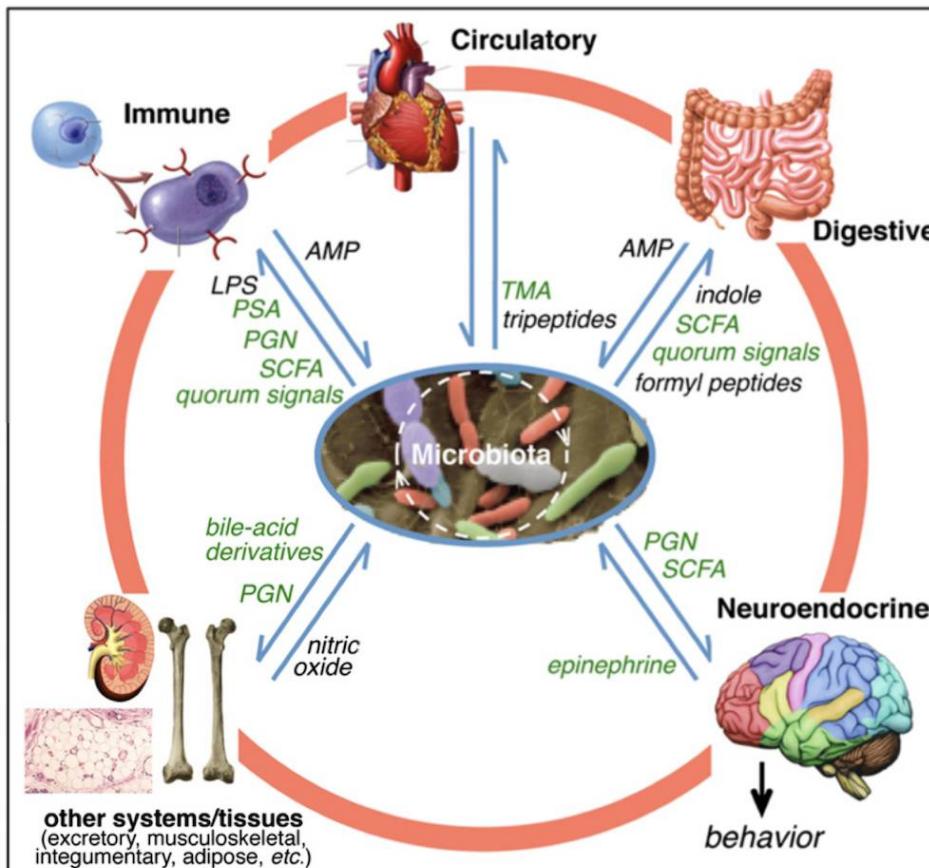


O'Huallachain 2012 PNAS

Holobiont perspective: Physiology

Slide from
S. Gilbert

Symbionts are intimately involved in regulating our normal physiological functions

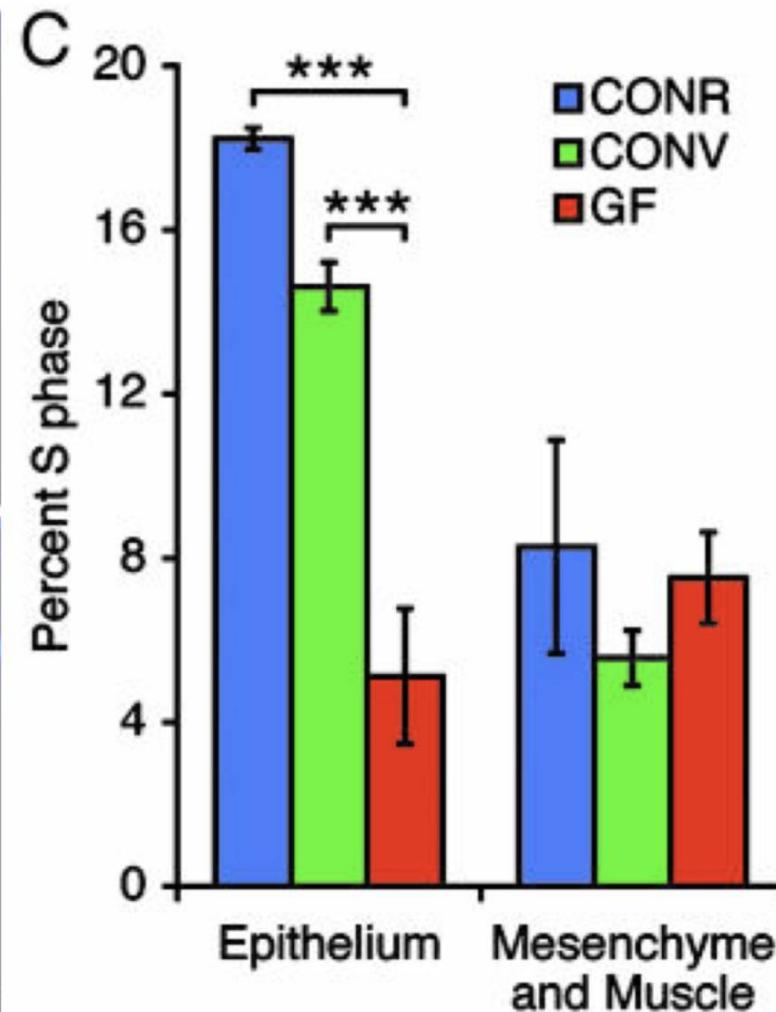
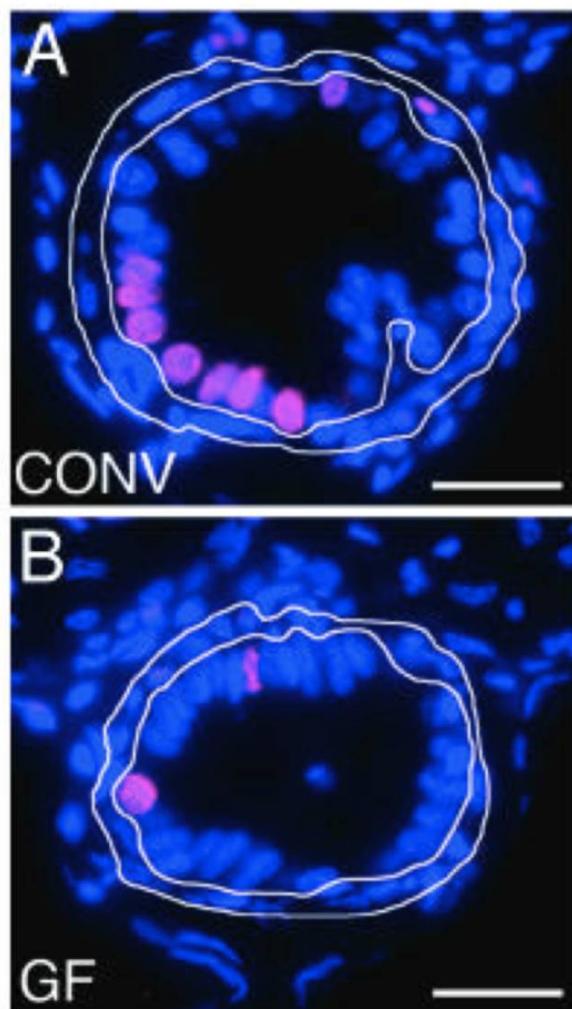


McFall-Ngai et al . 2013. PNAS 110: 3229 -3236

- Approximately 70% of human resting body temperature is the result of gut microbe metabolism (Henry 2005; Rosenberg and Zilberman-Rosenberg, 2016).

Microbial Symbionts Induce Epithelial Cell Formation in Zebrafish

Slide from
S. Gilbert



A, B. 6d conventional (A) and germ-free (B) intestine stained for cells (blue) and dividing cells (magenta). Mesenchyme and muscles outlined in white
Rawls et al. 2004. PNAS 101: 4596 – 4601. (Symbionts work through beta-catenin stabilization)

The “Mendelian Gene” and the “Molecular Gene”: Two Relevant Concepts of Genetic Units

2016

V. Orgogozo^{*,†}, A.E. Peluffo*, B. Morizot[†]



A SYMBIOTIC VIEW OF LIFE: WE HAVE NEVER BEEN INDIVIDUALS

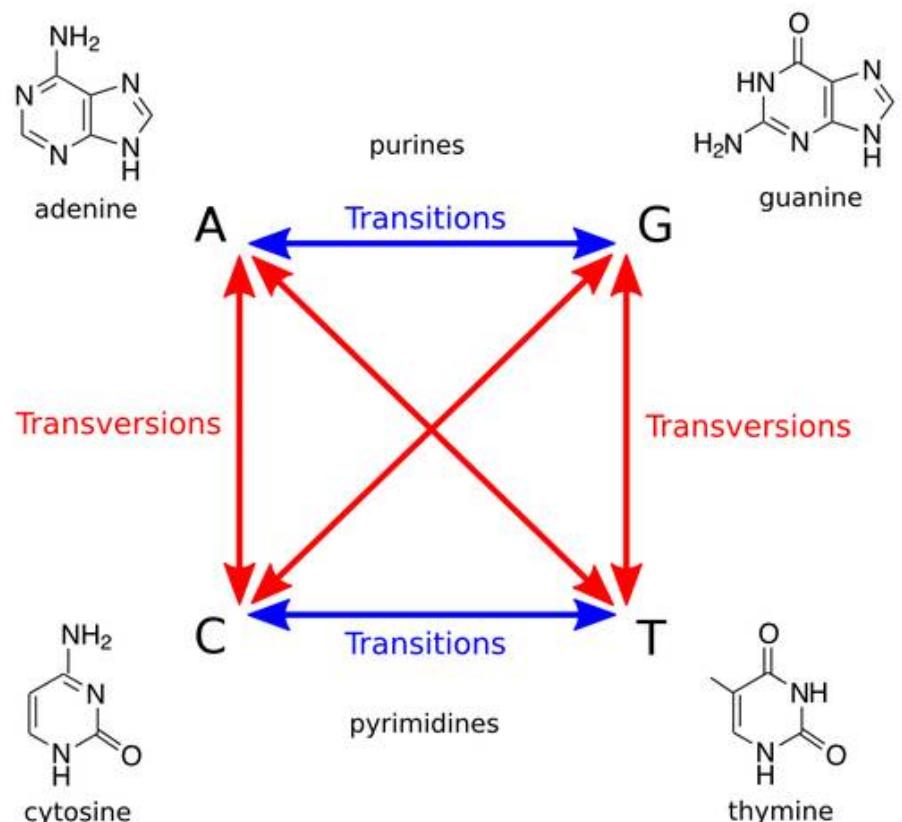
2012

SCOTT F. GILBERT, J. SAPP and A. TAUBER

À vous de jouer !

	U	C	A	G	
U	UUU } Phe - F UUC } UUA } Leu - L UUG }	UCU } UCC } Ser - S UCA }	UAU } Tyr - Y UAC } UAA stop UAG stop	UGU } Cys - C UGC } UGA stop UGG Trp - W	U C A G
C	CUU } CUC } Leu - L CUA } CUG }	CCU } CCC } CCA } Pro - P CCG }	CAU } His - H CAC } CAA } Gln - Q CAG }	CGU } CGC } CGA } Arg - R CGG }	U C A G
A	AUU } AUC } Ile - I AUA } AUG Met - M start	ACU } ACC } ACA } Thr - T ACG }	AAU } Asn - N AAC } AAA } Lys - K AAG }	AGU } Ser - S AGC } AGA } Arg - R AGG }	U C A G
G	GUU } GUC } Val - V GUA } GUG }	GCU } GCC } GCA } Ala - A GCG }	GAU } Asp - D GAC } GAA } Glu - E GAG }	GGU } GGC } GGA } Gly - G GGG }	U C A G

IUPAC nucleotide code



Symbol	Bases	Description
A	A	Adenine
C	C	Cytosine
G	G	Guanine
T (or U)	T (or U)	Thymine (or Uracil)
W	A or T	Weak
S	C or G	Strong
M	A or C	aMino
K	G or T	Keto
R	A or G	puRine
Y	C or T	pYrimidine
B	C or G or T	not A (B comes after A)
D	A or G or T	not C (D comes after C)
H	A or C or T	not G (H comes after G)
V	A or C or G	not T (V comes after T and U)
N	any base	any Nucleotide (not a gap)

Specialized loci in the genome

Proteins that interact with external molecules

oxygen, photons, insecticide, cholesterol...

Specialized loci in the genome

Proteins that interact with external molecules

oxygen, photons, insecticide, cholesterol...

Cis-regulatory elements of “developmental switch genes”

