

Principles of evolution (BIO 351)



University of
Zurich^{UZH}



Instructor: Pouria Dasmeh

Group Leader in Evolutionary Medicine, Institute for Human Genetics,
Marburg University.

Long-term research fellow, Shakhnovich Lab., Harvard University.



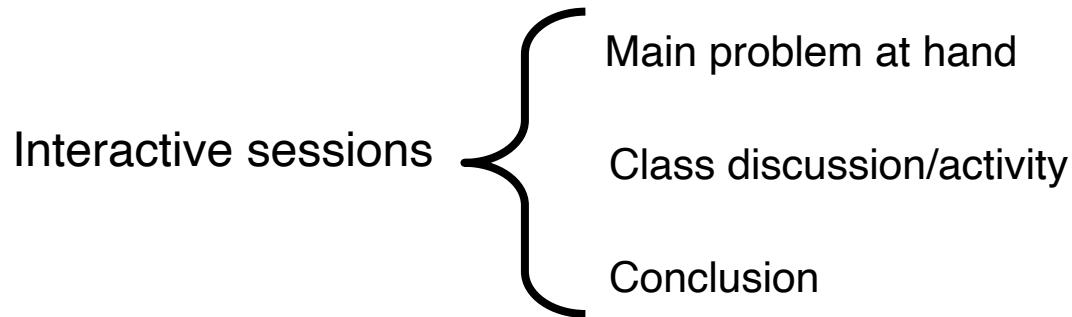
<https://github.com/dasmeh>



@PouriaDasmeh

Dasmeh@staff.uni-marburg.de

General structure of the course



We solve problems in groups!

Group 1:

Jonathan F.
Stefanie K.
Sarah R.

Group 2:

Dominik K.
Alex P.
Anna T.

Group 3:

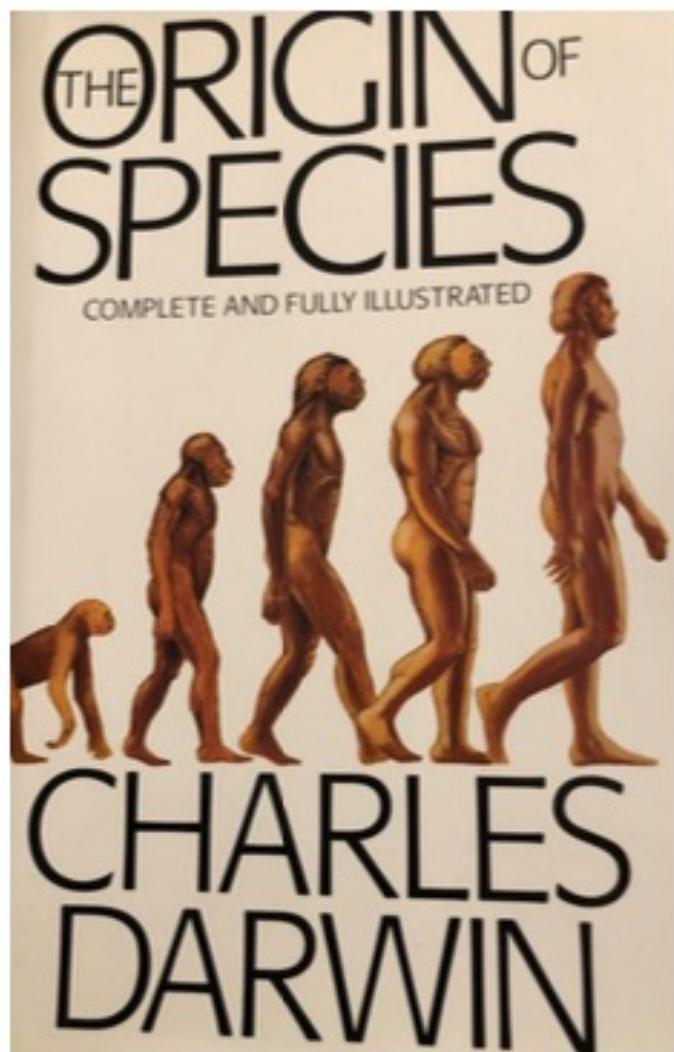
Gilles R.
Krystsina S.
Eric T.

Variation among individuals



Figure 5.1 Variation among individuals Top, whiskerbrush (*Linanthus ciliatus*). Photo by Eric Knapp. Center, bat stars (*Asterina miniata*). Bottom, variable ground snakes (*Sonora semiannulata*). Photo by Alison Davis Rabosky and Christian Cox.

Variation among individuals



CHAPTER I

VARIATION UNDER DOMESTICATION

Causes of Variability – Effects of Habit – Correlation of Growth
– Inheritance – Character of Domestic Varieties – Difficulty of
distinguishing between Varieties and Species – Origin of
Domestic Varieties from one or more Species – Domestic
Pigeons, their Differences and Origin – Principle of Selection
anciently followed, its Effects – Methodical and Unconscious
Selection – Unknown Origin of our Domestic Productions –
Circumstances favourable to Man's power of Selection

WHEN we look to the individuals of the same variety or sub-variety of our older cultivated plants and animals, one of the first points which strikes us, is, that they generally differ much more from each other, than do the individuals of any one species or

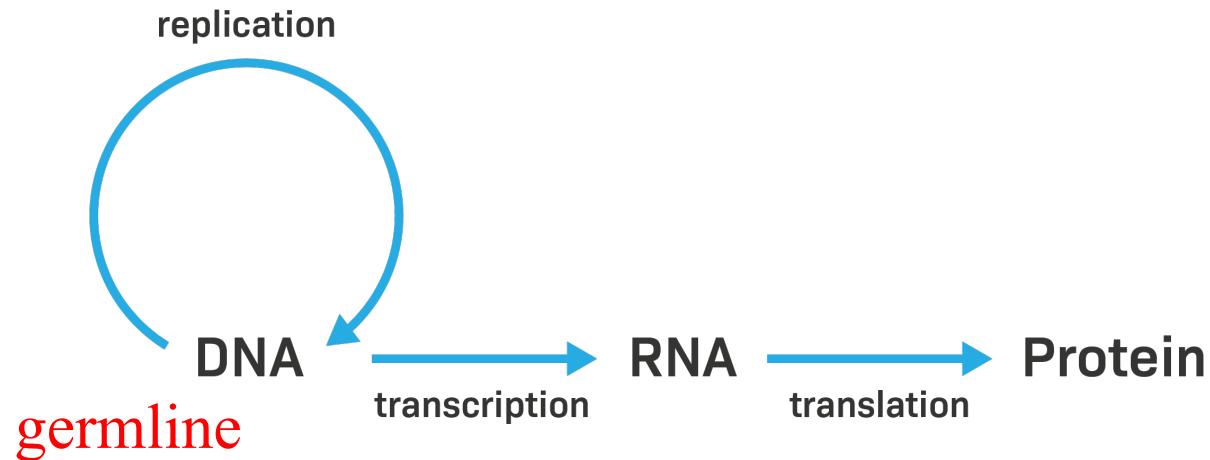
AN HISTORICAL SKETCH

OF THE PROGRESS OF OPINION ON THE ORIGIN OF SPECIES

PREVIOUSLY TO THE PUBLICATION OF THE FIRST EDITION OF THIS WORK

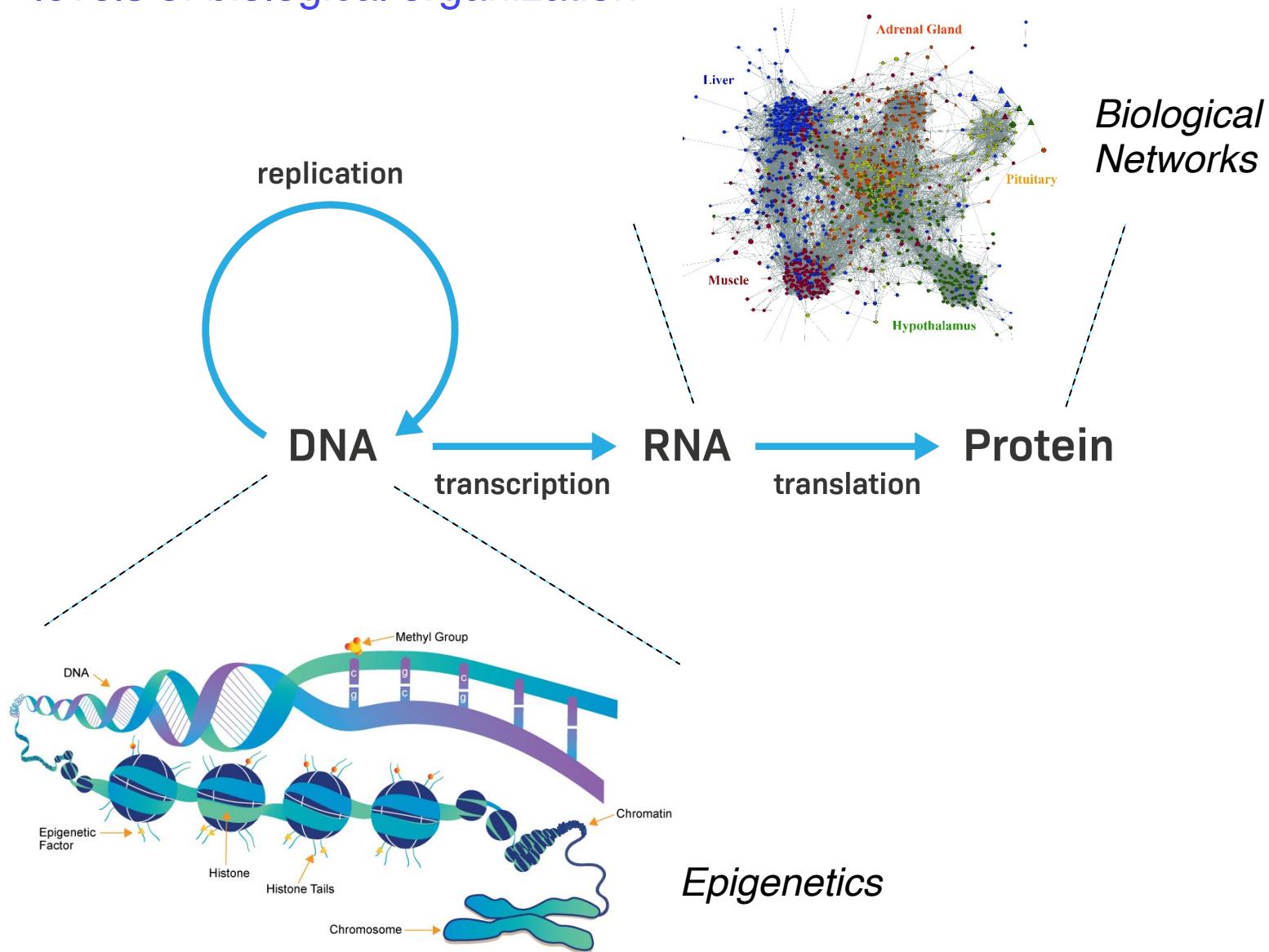
I WILL here give a brief sketch of the progress of opinion on the Origin of Species. Until recently the great majority of naturalists believed that species were immutable productions, and had been separately created. This view has been ably maintained by many authors. Some few naturalists, on the other hand, have believed that species undergo modification, and that the existing forms of life are the descendants by true generation of pre-existing forms. Passing over allusions to the subject in the classical writers,* the first author who in modern times has treated it in a scientific spirit was Buffon. But as his opinions fluctuated greatly at different periods, and as he does not enter on the causes

The sources of “heritable” variation

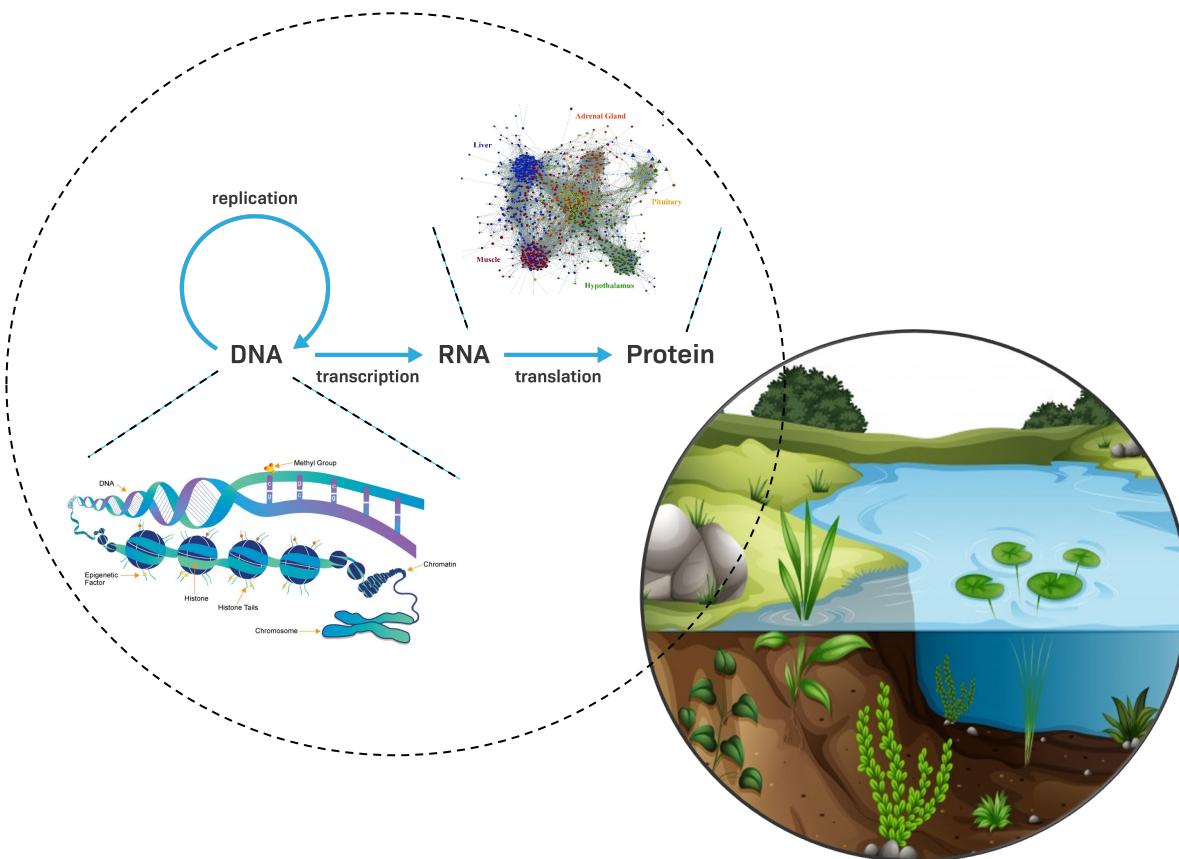


Central dogma of molecular biology

Mutations can influence different levels of biological organization



Genetic variation in interaction with the environment



Group Projects 1

(different types of variation)

Project 1: Genetic variation for bitter taste perception

Project 2: Inducible defenses in *Daphnia*

Project 3: Random variation in protein production

Presentation:

- 1) Quick intro
- 2) Observation(s)
- 3) Important concept(s)
- 4) Propose follow-up research

What types of variations provide the “raw material” for evolution?



Point mutations

Single base pair changes

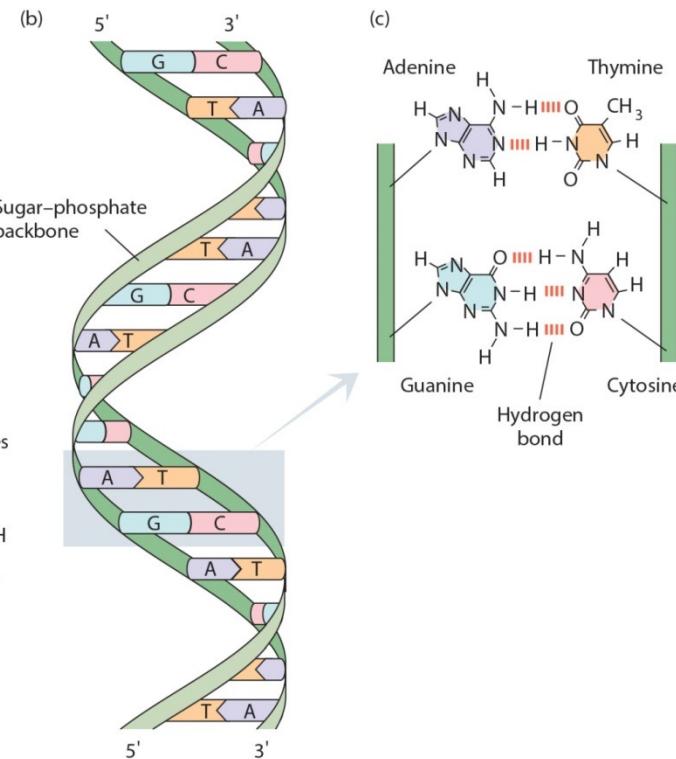
Historically the first characterized (1958):
mutation causing sickle-cell anemia

GAG → GTG
glutamic acid → valine
at amino acid 6 of β-globin

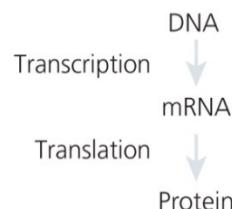
create new **alleles** (variants of a gene)

caused by
replication errors
chemical mutagens
radiation

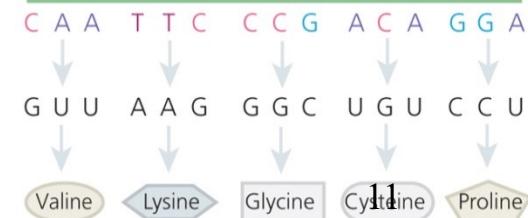
related small changes: **indels**
(insertions or deletions)



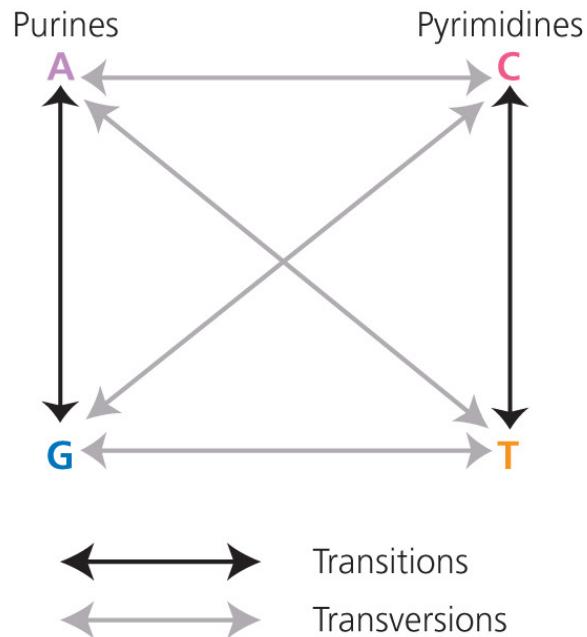
(a) The flow of genetic information



Example



Two main kinds of point mutations



Transitions are more frequent than transversions
e.g. 1.6:1 in the roundworm *C. elegans*

probably because transversions are more disruptive
and thus more likely to be detected and repaired

(The sickle-cell anemia mutation is a transversion.)

Redundancy of the genetic code

(b) The RNA genetic code

First base	Second base												Third base
	U			C			A			G			
U	UUU	Phenylalanine	UCU	Serine		UAU	Tyrosine		UGU	Cysteine			U C A G
	UUC	Phenylalanine	UCC	Serine		UAC	Tyrosine		UGC	Cysteine			
	UUA	Leucine	UCA	Serine		UAA	Stop		UGA	Stop			
	UUG	Leucine	UCG	Serine		UAG	Stop		UGG	Tryptophan			
C	CUU	Leucine	CCU	Proline		CAU	Histidine		CGU	Arginine			U C A G
	CUC	Leucine	CCC	Proline		CAC	Histidine		CGC	Arginine			
	CUA	Leucine	CCA	Proline		CAA	Glutamine		CGA	Arginine			
	CUG	Leucine	CCG	Proline		CAG	Glutamine		CGG	Arginine			
A	AUU	Isoleucine	ACU	Threonine		AAU	Asparagine		AGU	Serine			U C A G
	AUC	Isoleucine	ACC	Threonine		AAC	Asparagine		AGC	Serine			
	AUA	Isoleucine	ACA	Threonine		AAA	Lysine		AGA	Arginine			
	AUG	Start (Methionine)	ACG	Threonine		AAG	Lysine		AGG	Arginine			
G	GUU	Valine	GCU	Alanine		GAU	Aspartic acid		GGU	Glycine			U C A G
	GUC	Valine	GCC	Alanine		GAC	Aspartic acid		GGC	Glycine			
	GUA	Valine	GCA	Alanine		GAA	Glutamic acid		GGG	Glycine			
	GUG	Valine	GCG	Alanine		GAG	Glutamic acid						

Codon Amino acid nonpolar polar basic acidic

Replacement (nonsynonymous) mutation: causes amino acid change
Silent (synonymous) mutation: causes no such change

Most changes at 1st and 2nd codon positions are replacement changes

Many mutations at third positions are silent

Class activity

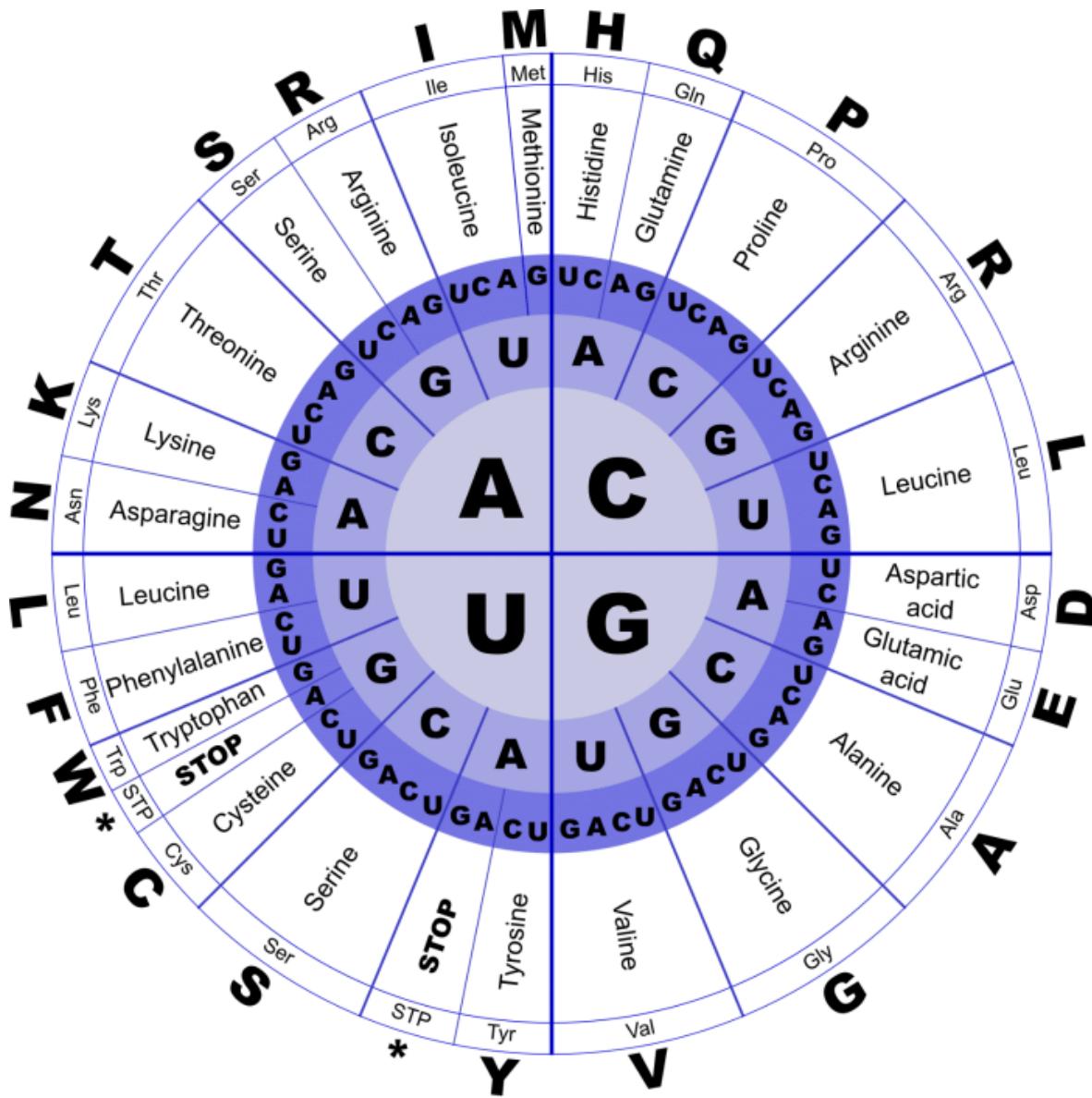
...CAGGAUGUUAAAC...

Mutation

...CAGG**G**UGUUAAAC...

Amino acid
RNA base
sequence

Bases
A: adenine
C: cytosine
G: guanine
U: uracil



Genetic code (RNA)

The case of D614G in SARS-CoV-2

nature

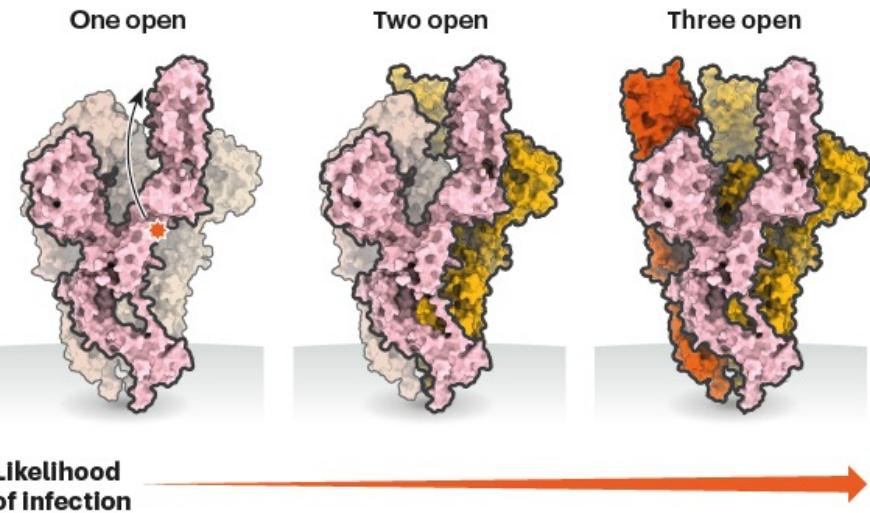
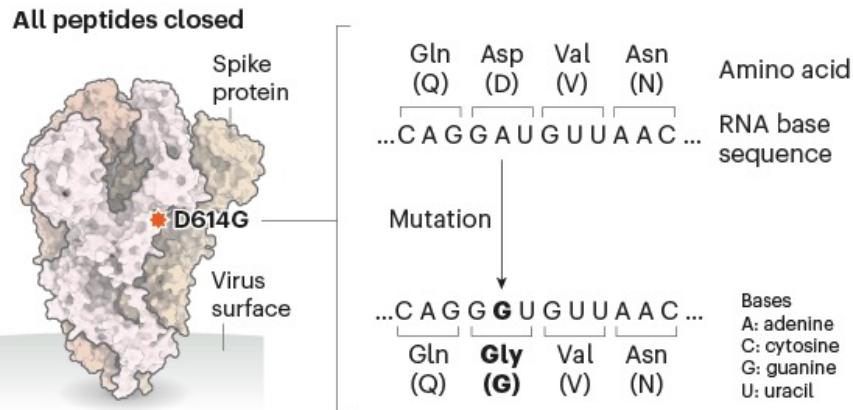
Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > news feature > article

NEWS FEATURE | 08 September 2020 | Correction [16 September 2020](#)

The coronavirus is mutating – does it matter?

Different SARS-CoV-2 strains haven't yet had a major impact on the course of the pandemic, but they might in future.



©nature

Group Projects 2 (D614G in SARS-CoV-2)

Group 1:
Global spread

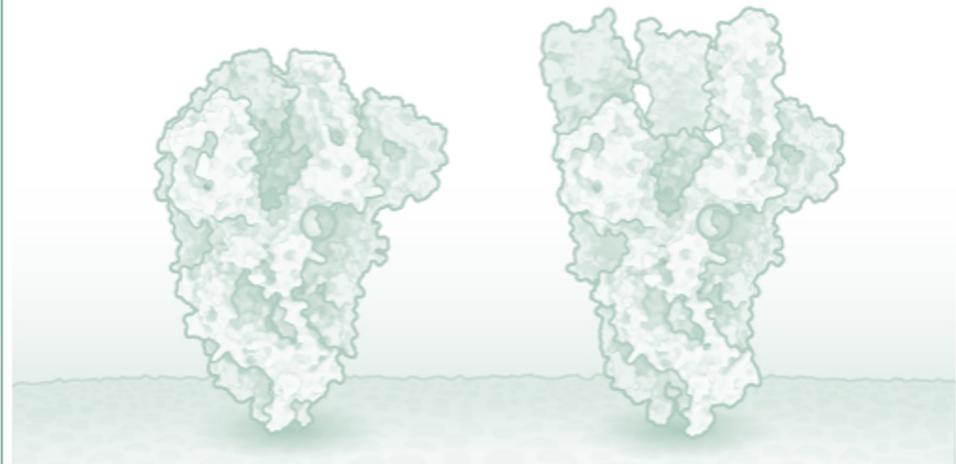
Group 2:
Faster spread

Group 3:
No escape from antibodies-yet

Presentation:

- 1) Quick intro
- 2) Observation(s)
- 3) Important concept(s)
- 4) Propose follow-up research

Feature



The spike protein of SARS-CoV-2 has a common mutation (circled) that seems to shift the protein from a closed (left) to an open (right) form.

SCIENCE STRUCTURAL BIOLOGY, SHOBHINI LUBAN

MAKING SENSE OF CORONAVIRUS MUTATIONS

Different SARS-CoV-2 strains haven't yet had a major impact on the course of the pandemic – but they might in future. By Ewen Callaway

When COVID-19 spread around the globe this year, David Montefiori wondered how the deadly virus behind the pandemic might be changing as it passed from person to person. Montefiori is a virologist who has spent much of his career studying how chance mutations in HIV help it to evade the immune system. The same thing might happen with SARS-CoV-2, he thought.

In March, Montefiori, who directs an AIDS-vaccine research laboratory at Duke University in Durham, North Carolina, contacted Bette Korber, an expert in HIV evolution and a long-time collaborator. Korber, a computational biologist at the Los Alamos National Laboratory (LANL) in New Mexico, had already

started scouring thousands of coronavirus genetic sequences for mutations that might have changed the virus's properties as it made its way around the world.

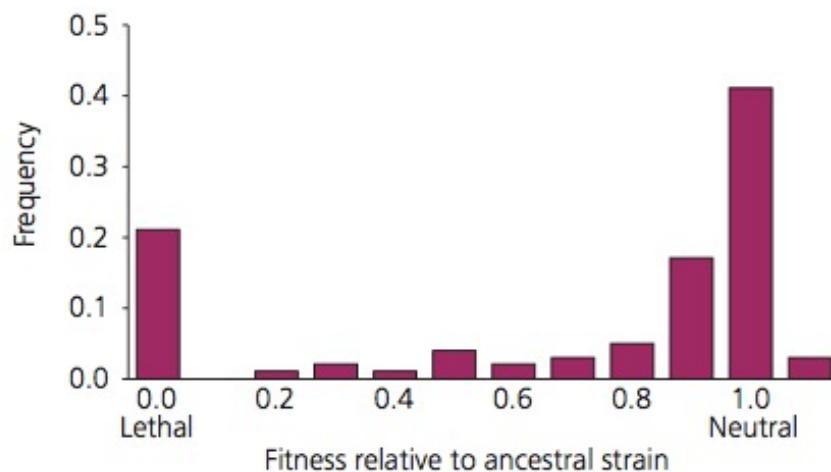
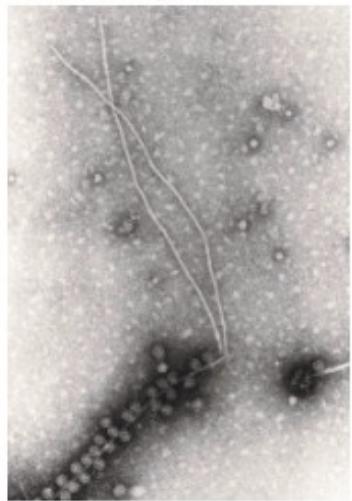
Compared with HIV, SARS-CoV-2 is changing much more slowly as it spreads. But one mutation stood out to Korber. It was in the gene encoding the spike protein, which helps virus particles to penetrate cells. Korber saw the mutation appearing again and again in samples from people with COVID-19. At the 614th amino-acid position of the spike protein, the amino acid aspartate (D, in biochemical shorthand) was regularly being replaced by glycine (G) because of a copying fault that altered a single nucleotide in the virus's 29,903-letter RNA code. Virologists were calling it the D614G mutation.

In April, Korber, Montefiori and others warned in a preprint posted to the bioRxiv server that "D614G is increasing in frequency at an alarming rate"¹. It had rapidly become the dominant SARS-CoV-2 lineage in Europe and had then taken hold in the United States, Canada and Australia. D614G represented a "more transmissible form of SARS-CoV-2", the paper declared, one that had emerged as a product of natural selection.

These assertions dismayed many scientists. It wasn't clear that the D614G viral lineage was more transmissible, or that its rise indicated anything unusual, they said. But alarm spread fast across the media. Although many news stories included researchers' caveats, some headlines declared that the virus was mutating to become more dangerous. In retrospect,

Distribution of fitness effects (DFE)

(a) Bacteriophage f1



(b) Brewer's yeast

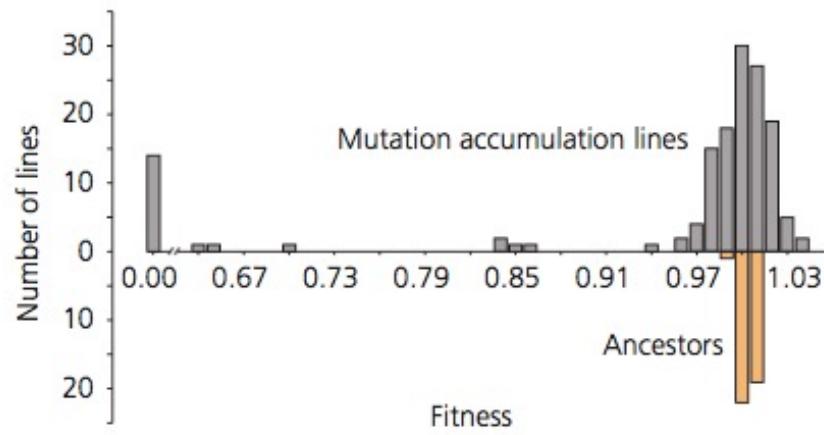
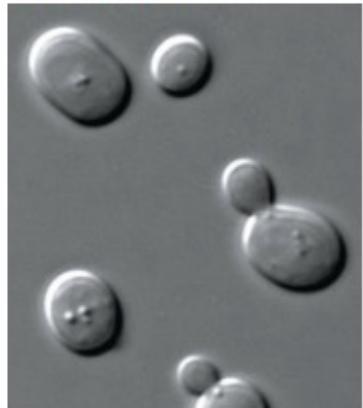


Figure 5.36 The distribution of fitness effects of new mutations (a) Fitness effects of 100 random mutations engineered into the genome of a virus. Redrawn from Peris et al. (2010). (b) Fitnesses of 144 mutation accumulation lineages in brewer's yeast compared to ancestral lineages. Redrawn from Hall and Joseph (2010).

Mutation-accumulation experiments

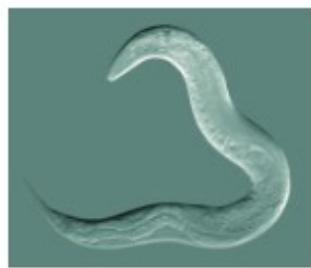
Direct estimation of mutation rates through mutation accumulation experiments

Principle: propagate an organism for many generations under conditions that minimize the effects of selection

ample food, little crowding, no predation
population bottlenecks (to maximize drift, see Ch. 7)

and then count the number of accumulated mutations by DNA sequencing

Mutation-accumulation experiments



Caenorhabditis elegans

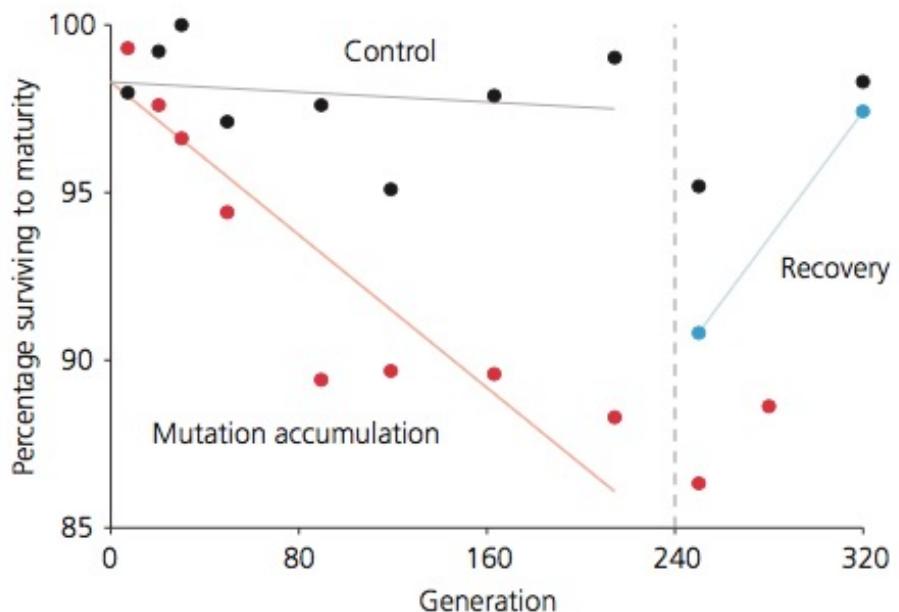


Figure 5.38 The balance between mutation and natural selection Nematode lineages insulated from natural selection declined in fitness relative to control lineages as they accumulated deleterious mutations. The lineages recovered upon exposure to natural selection due to the elimination of deleterious mutations and the preservation of favorable ones. Redrawn from Vassilieva et al. (2000), Estes and Lynch (2003). Photo by Bob Goldstein.

The importance of proofreading

LETTERS TO THE EDITOR

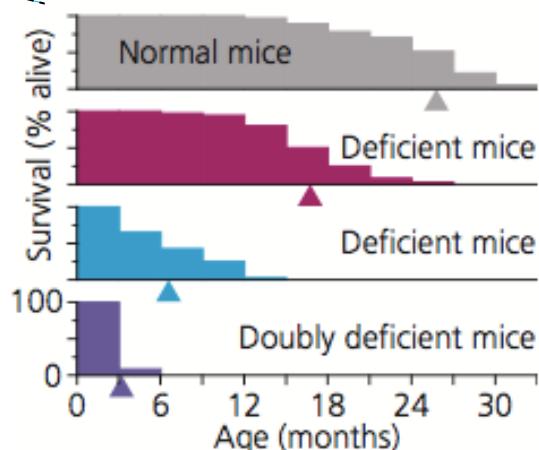
© 2001 Nature Publishing Group <http://medicine.nature.com>

Defective DNA polymerase- δ proofreading causes cancer susceptibility in mice

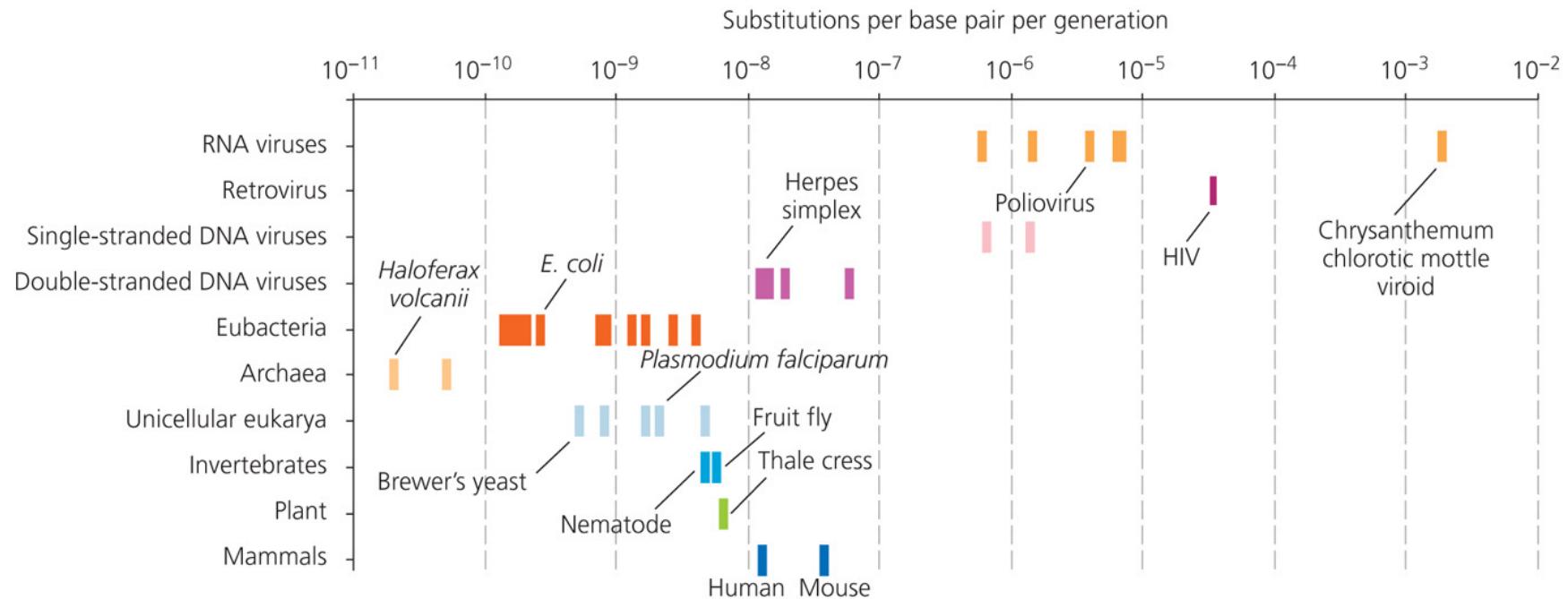
To the editor—Tumor development is a multistep process that requires an accumulation of mutations in genes regulating cell growth and metastasis¹. Spontaneous mutations occur rarely in normal cells, and thus it is argued that defects in pathways governing genetic

tant allele, we crossed a chimeric male with C57BL/6J females, and the resultant heterozygous F1 mice were interbred to generate a cohort of F2 animals. A total of 199 F2 mice were evaluated from 11 F1 breeding pairs (mean litter size, 6.8 pups). We obtained 53 wild-type, 97 *Pold1*^{+/-D400A}

23 tumors (Fig. 1b). The tumors were derived from two different cell lineages. Sixteen of the 23 tumors originated from mesenchymal tissue (12 thymic lymphomas, 2 diffuse lymphomas and 2 sarcomas) and 7 originated from epithelial tissue (6 tail-skin squamous-cell carcinomas and 1 mammary adenocarcinoma).



Mutation rates vary widely among different organisms



Class activity

Estimate the number of mutations that can accumulate in E.coli and human in one generation.

Human ~ 3.2 billion base-pairs

E.coli ~ 4.6 million base-pairs

The evolution of bacteria on a "mega-plate" Petri dish



HARVARD
MEDICAL SCHOOL



Technion
Israel Institute of
Technology

Different classes of mutations

point mutations and their rates

gene duplications and the origin of new genes

chromosome mutations

genome duplication (polyploidization)

Group projects (different classes of mutations)

Project 1: Immigration delay disease!

Project 2: Short leggedness in dogs!

Project 3: Inversion in *Drosophila subobscura*

Presentation:

- 1) Quick intro
- 2) Observation(s)
- 3) Important concept(s)
- 4) Propose follow-up research

Gene duplications can create new genes

Two main causes of gene duplication

1. Retroposition

reverse transcription of
a mRNA into DNA by reverse transcriptase
integration of DNA into the genome

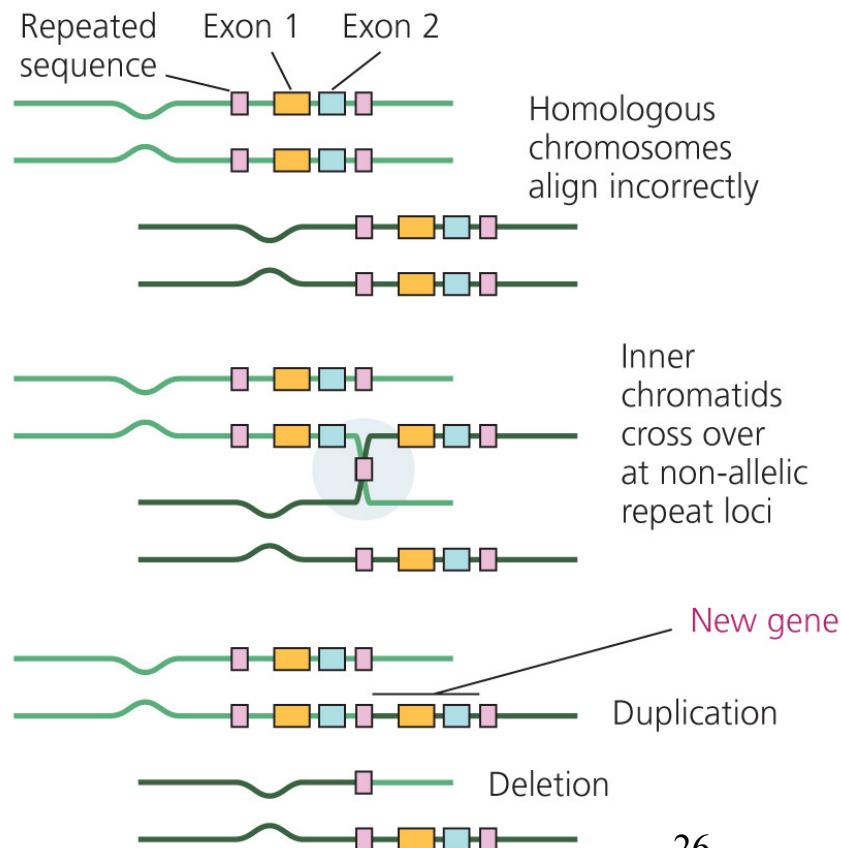
duplicates are often in new location,
lack introns, have poly A sequence

2. Unequal cross-over during meiosis

duplicates «in tandem»,

duplicates have introns,
lack poly A sequence

(a) Duplication by unequal crossing over



A retrotransposed duplicate gene is responsible for chondrodysplasia in dogs

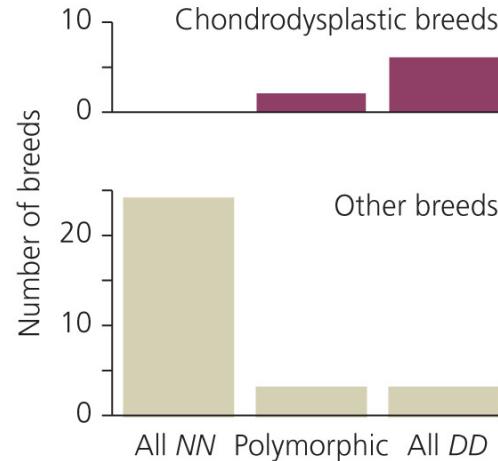
Chondrodysplasia is responsible for short-leggedness in some dog breeds

Fibroblast growth factor 4 (*fgf4*) is a protein important for bone development

In chondrodysplastic dogs the encoding gene is often duplicated

The second copy is intronless and occurs far from the original gene

(b) *Fgf4* retrogene genotypes in dogs of various breeds



(a) Who's a good girl?



NN... No known individual has two copies
DD... All known individuals have two copies

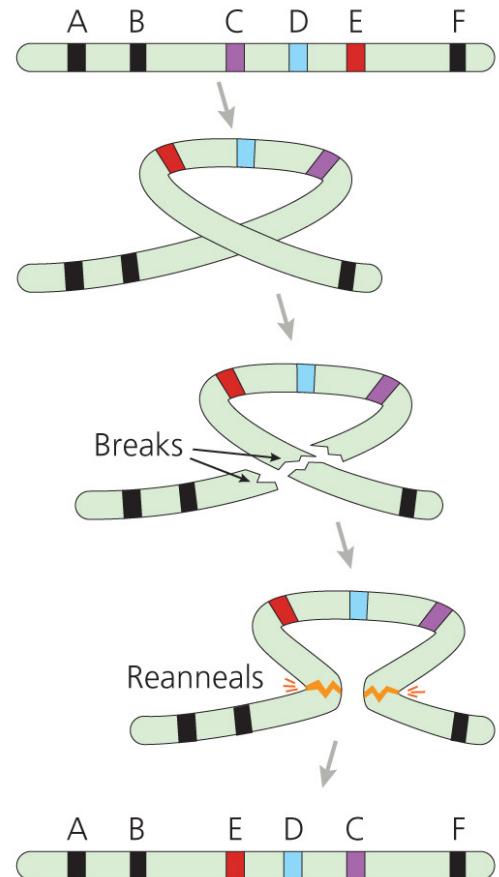
Chromosome mutations (inversions)

inversions affect **linkage**, the physical proximity of genes
e.g., they break up the linkage of B and C
and render C and F linked

inversion heterozygotes prevent alignment
of chromosomes during meiosis and thus
crossing over

alleles within an inversion are thus inherited
together

they form a **supergene**



Inversion polymorphisms and natural selection

Drosophila pseudoobscura is native to Europe, the Middle East, and North Africa

contains **inversion polymorphisms** on five of its six chromosomes

the incidence or **frequency** of these inversions varies with latitude and climate

such regular variation over a geographic area is also called a **cline**

is this variation a historical accident or caused by selection?

Inversion polymorphisms can be maintained by natural selection

in 1978, *D. pseudoobscura* invaded South America (Chile) and the US (Washington)
the invading populations contain the same inversion polymorphisms
they established the same clines, from wet/cold to desert/hot habitats

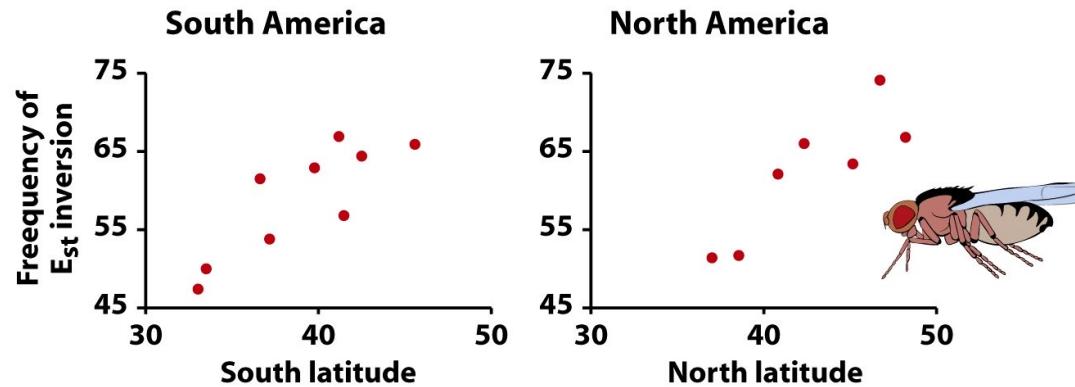


Figure 5-10 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Inversion polymorphisms can be maintained by natural selection

Small flies typically occur in dry, hot habitats

laboratory selection for small flies leads to inversion frequencies similar to those observed in dry, hot habitats in the wild

inversions likely contain genes affecting body size of *D.pseudoobscura*

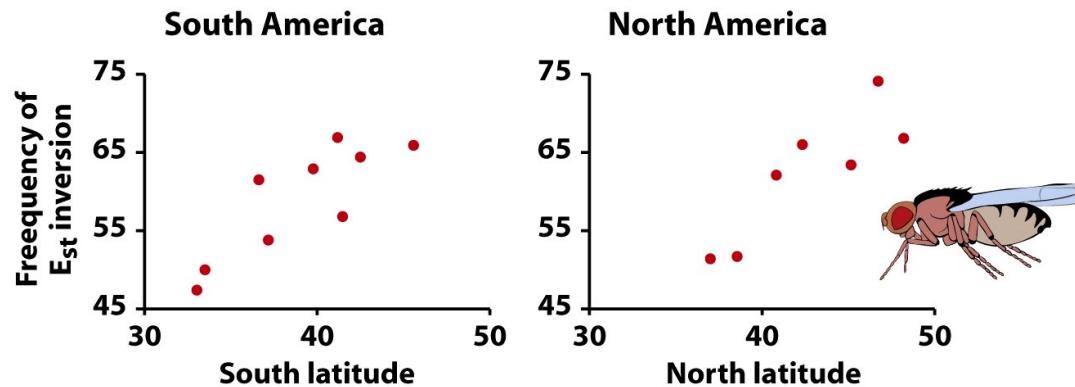


Figure 5-10 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Genome duplication (polyploidization)

polyploids are organisms with more than two chromosome sets
tetraploids ($4n$), hexaploids ($6n$), etc.

common in plants
many angiosperms (flowering plants)
estimated at 2 in every 100,000 offspring

less common in animals
but frequent in some groups, e.g., fish

facilitated by ability to self-fertilize and by parthenogenesis, a form of asexual reproduction

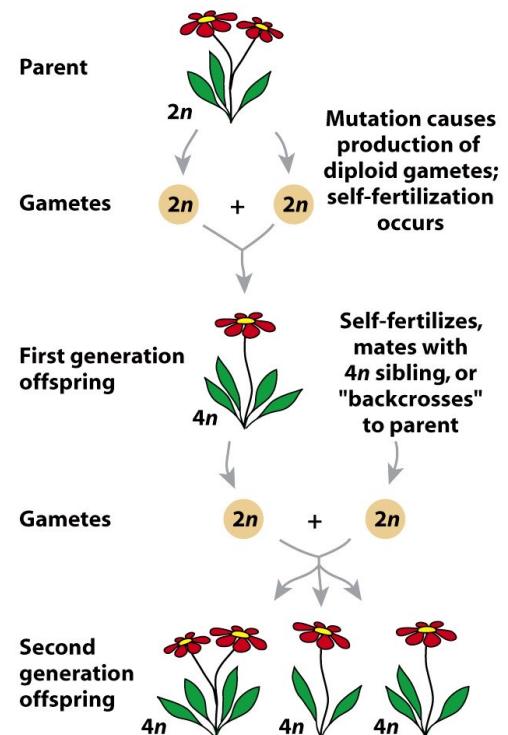


Figure 5-11 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

Two evolutionary consequences of polyploidization

Reproductive isolation

e.g., offspring of tetraploid with diploid will be triploid, which will lead to meiotic problems

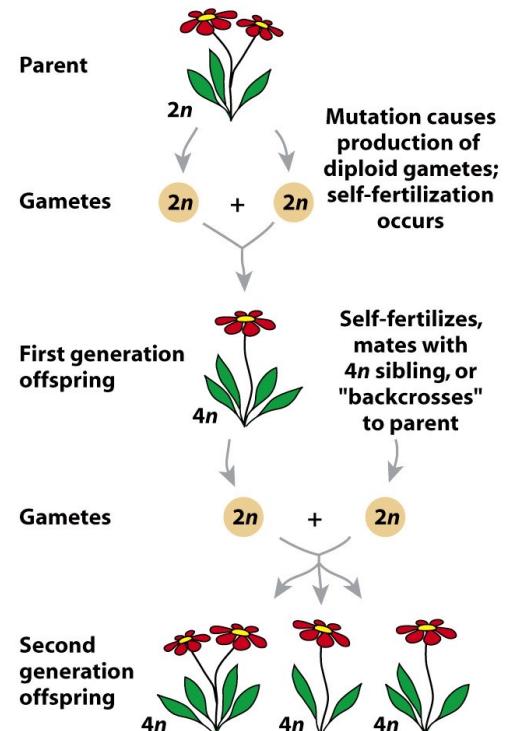
Evolutionary innovation and diversification

polyploidization creates many duplicate genes that can then diversify

it occurred early in the evolution of **ray-finned fish**, the most species-rich group of vertebrates

it also occurred in the evolution of **angiosperms**, which make up more than 80 percent of land plants

The connection between polyploidization and diversification is not proven



Genome duplications can help organisms adapt to new environments

Populations of yarrow («Schafgarbe») in California vary in ploidy

Tetraploids live in wetter habitats (coastal grasslands and forests)

Hexaploids live in drier habitats (sand dunes and oak woodlands)

Tetraploids occasionally produce hexaploid offspring

(a)



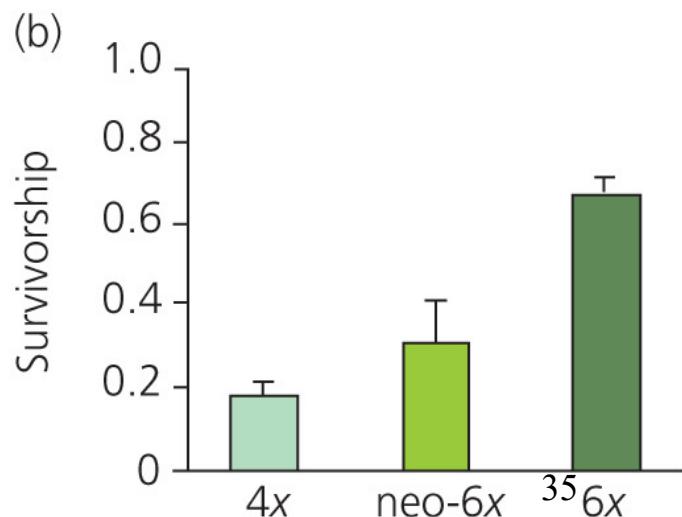
Genome duplications can help organisms adapt to new environments

Hypothesis: Hexaploids are better adapted to drier habitats

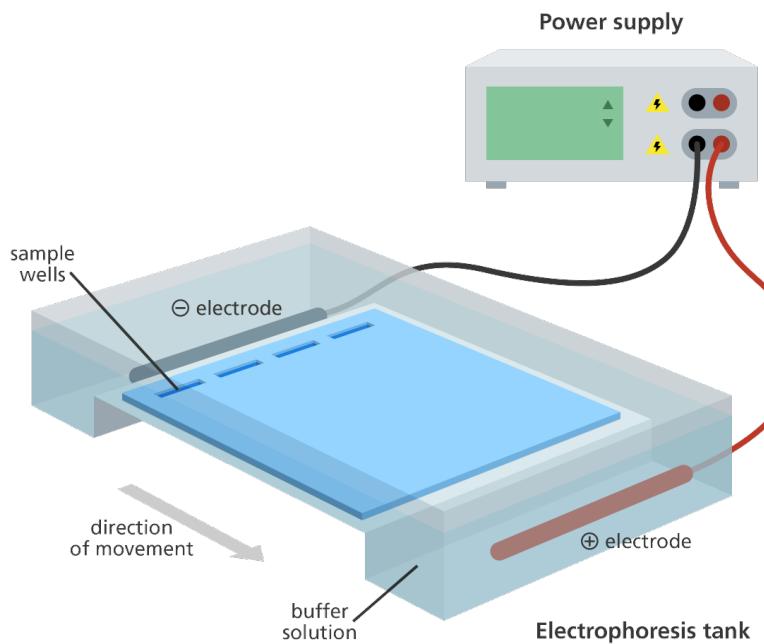
Test: Grow
tetraploids (4x)
first generation hexaploids (neo-6x)
hexaploids from wild populations (6x)
in sand-dunes (dry habitat)

Observation: Hexaploids are better adapted

Adaptation also occurs within hexaploids, because 6x plants do better than neo-6x plants



Detecting variants: Electrophoresis (review)



Experimental procedure:

<https://www.youtube.com/watch?v=vq759wKCCUQ>

Early work revealed abundant genetic variation through protein electrophoresis

33-50% of genes encoding enzymes are typically polymorphic

average individual is heterozygous at 4-15% of genes
thousands of heterozygous genes per individual

Neutralist-selectionist controversy:

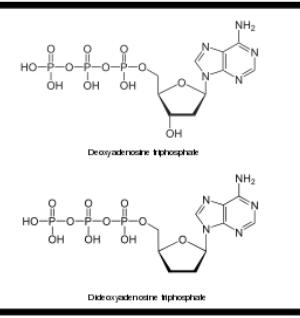
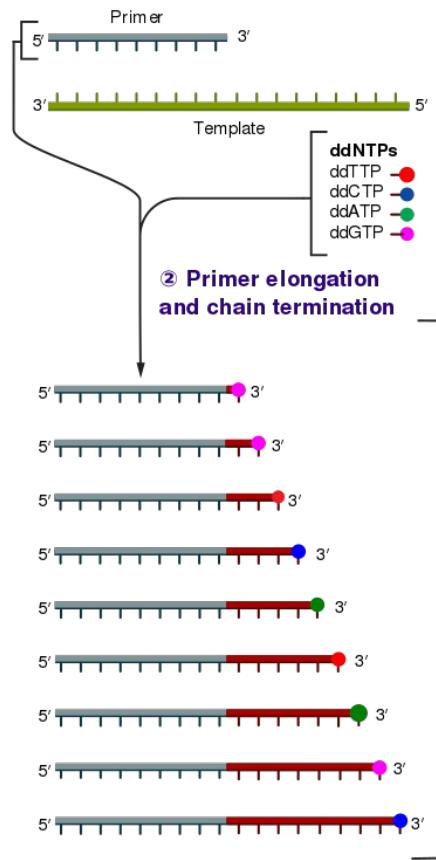
Are the most of observed variation
neutral or affected by selection?



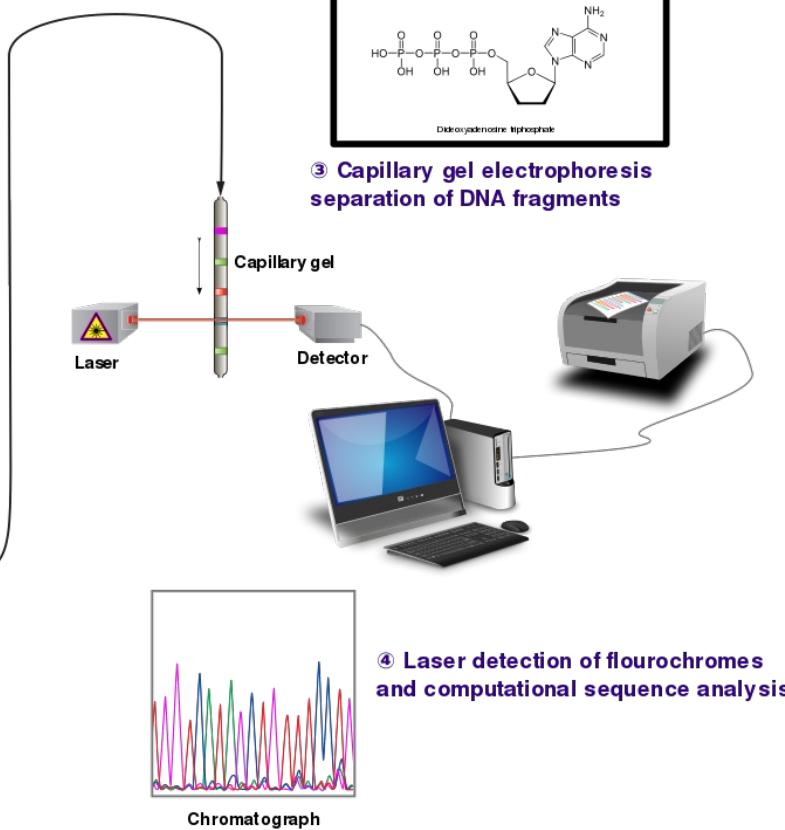
Detecting variants: DNA-sequencing (sanger sequencing)

① Reaction mixture

- Primer and DNA template ► DNA polymerase
- ddNTPs with flourochromes ► dNTPs (dATP, dCTP, dGTP, and dTTP)



③ Capillary gel electrophoresis separation of DNA fragments

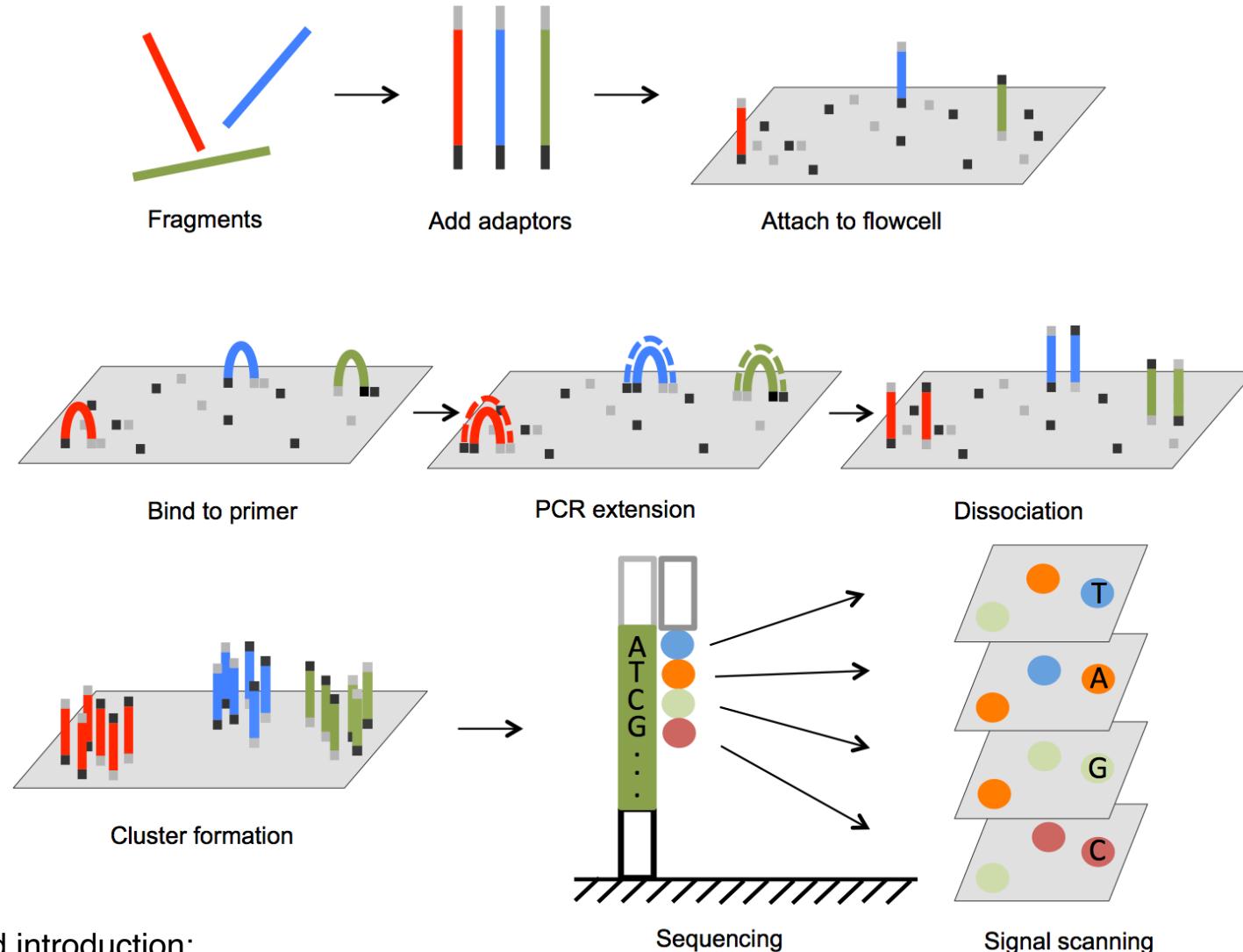


④ Laser detection of flourochromes and computational sequence analysis

A good introduction:

<https://www.youtube.com/watch?v=KTstRrDTmWI>

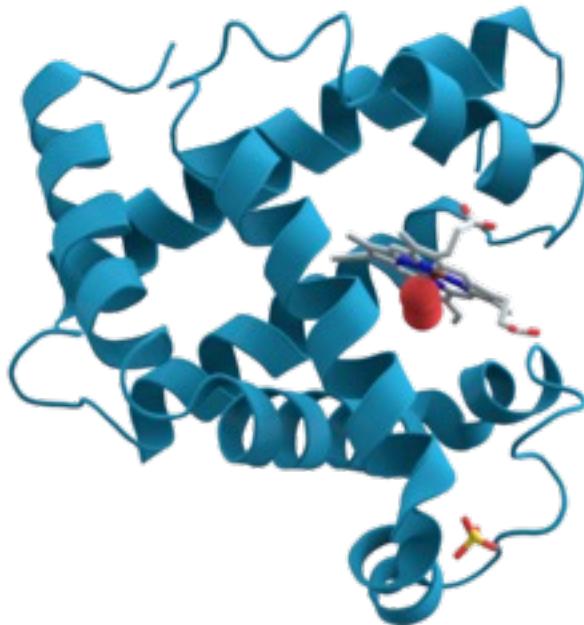
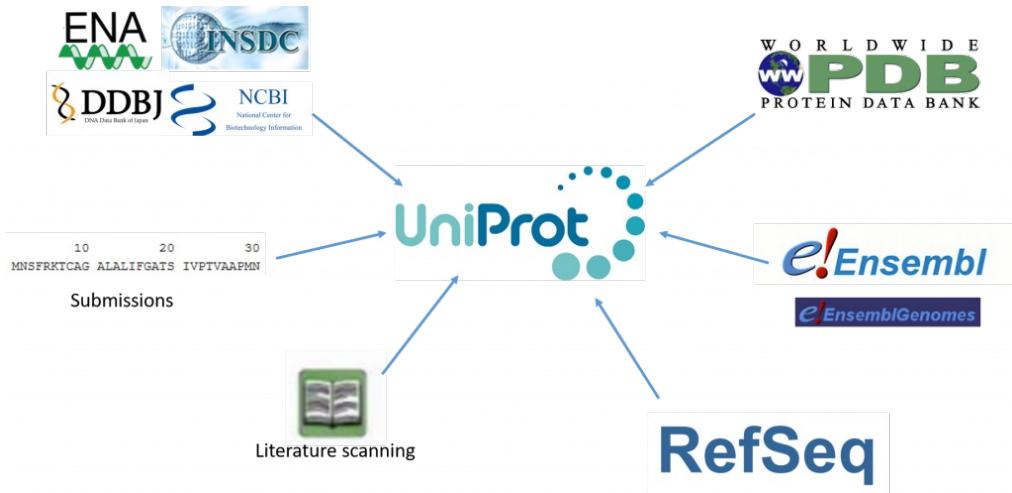
Detecting variants: DNA-sequencing (Next generation sequencing)



A good introduction:

<https://www.youtube.com/watch?v=fCd6B5HRaZ8>

Databases to find genetic variants (Example: UniProt)



Class activity:

Find the natural variants of Myoglobin
in UniProt database

Databases to find genetic variants (NextStrain: SARS-CoV-2)

