# Introduction on genetics and the genotype-phenotype relationship

**Virginie Courtier-Orgogozo**
**Institut Jacques Monod, Paris**

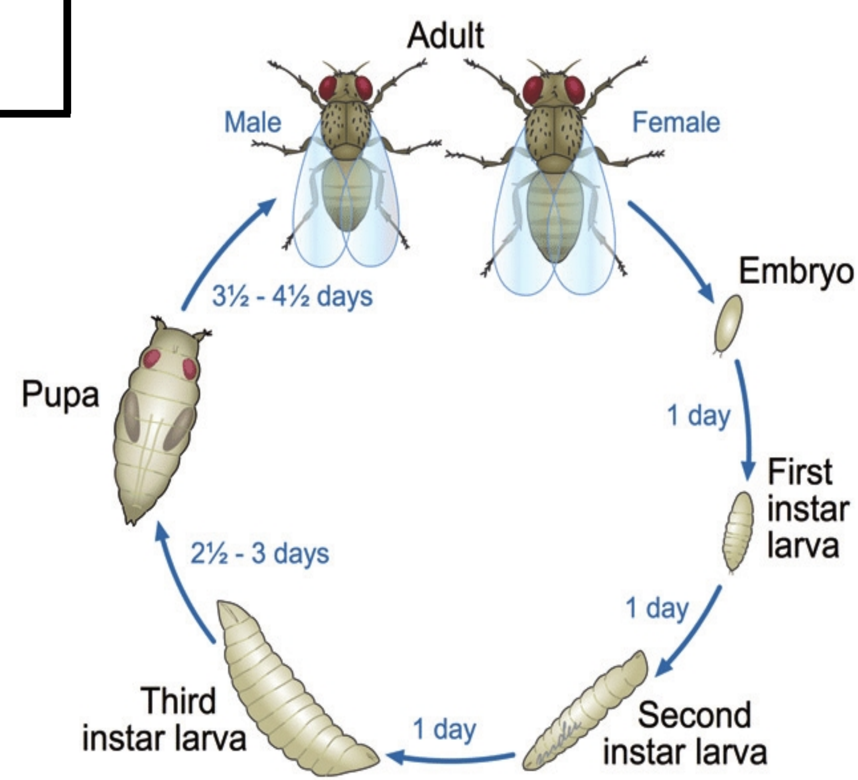
# My lab topic: How do species diverge with time ?

***D. pachea***

only senita cactus

***D. nannoptera***

many cactus species



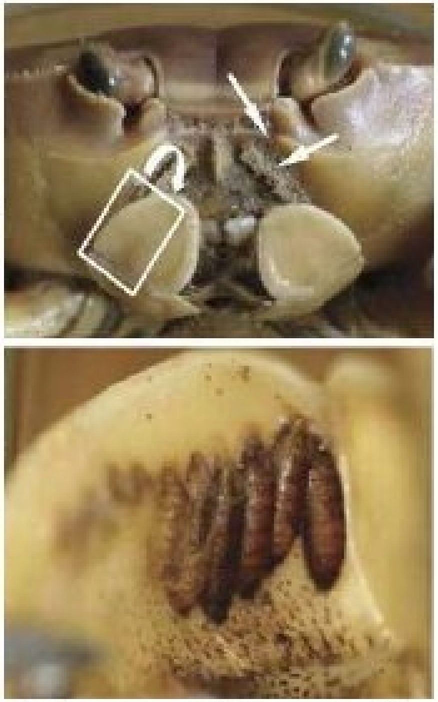
# Evolution of Drosophila glue

soil

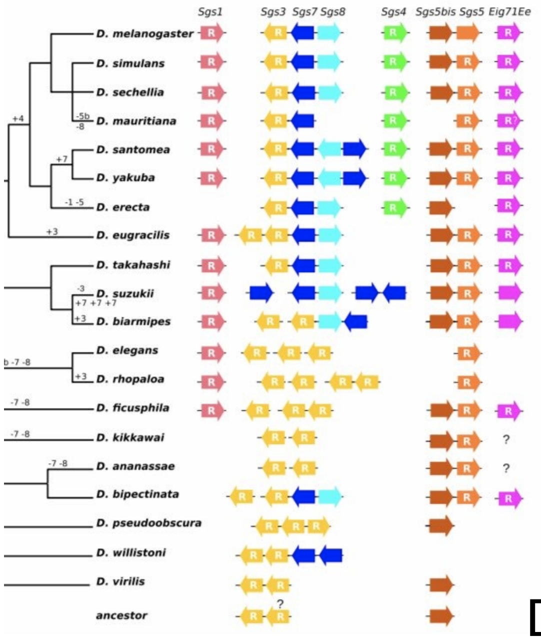
rotten strawberry

grape

potato

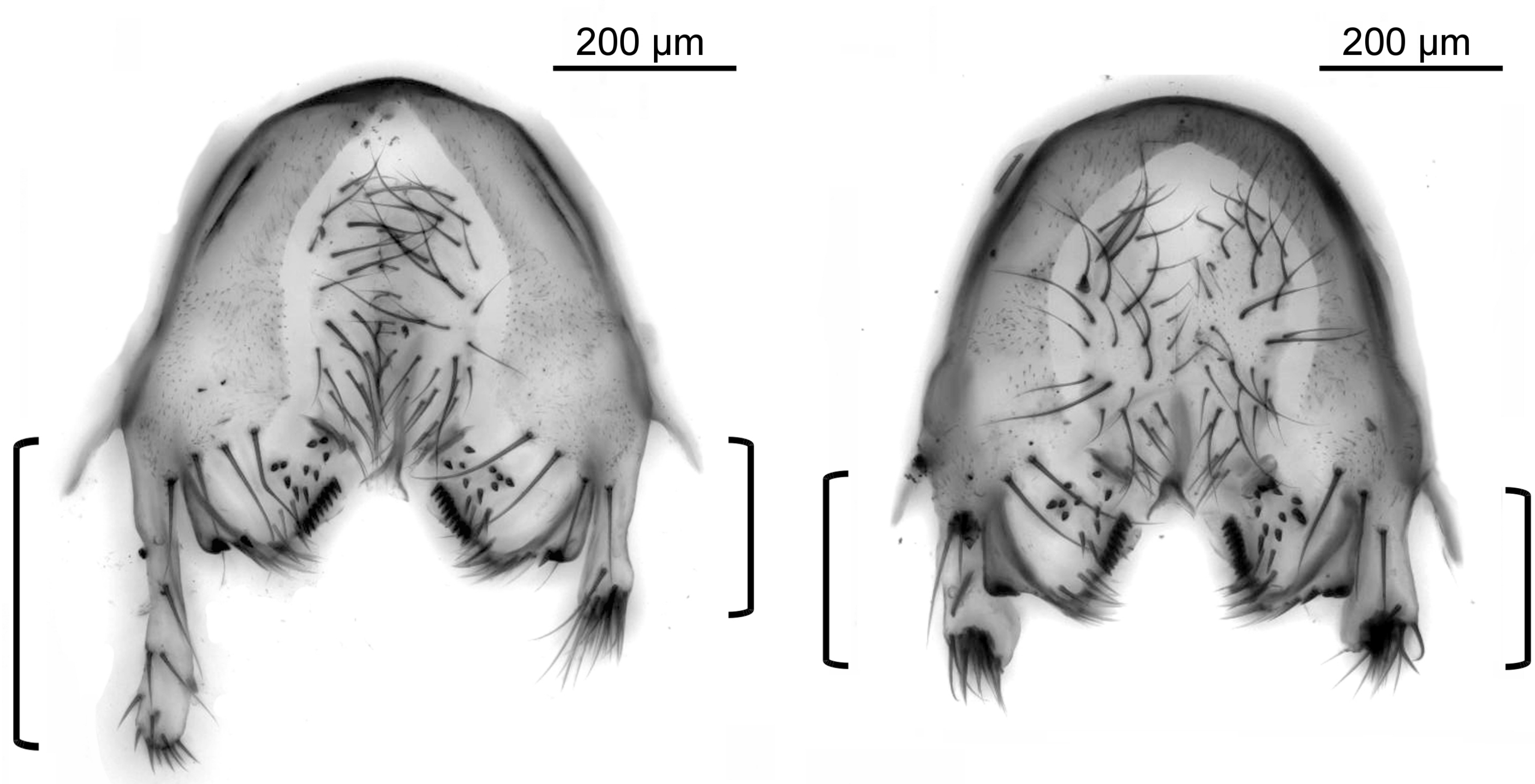


. *grimshawi* *D. suzukii* *D. melanogaster* *D. repleta* *D. carcinophila*




Da Lage et al. 2019 BMC Evol Biol

# Evolution of left-right asymmetry in *D. pachea*



# Introduction on genetics
# and
# the genotype-phenotype relationship

**Virginie Courtier-Orgogozo**
**Institut Jacques Monod, Paris**

# The distinction between genotype and phenotype is the basis of genetics

“The view of natural inheritance as realized by an act of transmission, viz., the transmission of the parent's (or ancestor's) personal qualities to the progeny, is the most naive and oldest conception of heredity.”

“All "types" of organisms, distinguishable by direct inspection or only by finer methods of measuring or description, may be characterized as "**phenotypes**.”

“ A "**genotype**"  is the sum of all the "genes" in a gamete or in a zygote.”

Johansen 1911

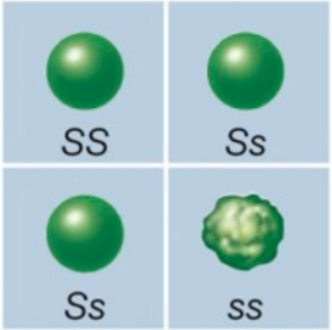
**Phenotype** = observable attributes of an individual

**Genotype** = inheritable genetic material
= DNA or RNA

**How do genotypes map onto phenotypes ?**

# A very brief history of genetics

*Transmission* + *Phenotypic expression*



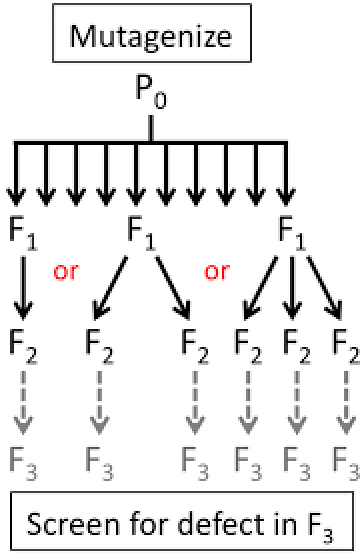

- **Mendel 1860s**
Controlled breeding, use of probabilities
Dominant/recessive/intermediate expression of traits
Random transmission of "factors"

- **Cytology from 1880-90s** Flemming, Boveri, Sutton
chromosomes as support of heredity, meiosis

- **Classical genetics 1900-1950**
distinction genotype-phenotype
segregation of characters & genes
genetic map, sex chromosomes
penetrance - expressivity

"Rediscovery" of Mendel's law
Extension to animals

- **Foundation of molecular biology  1940-1970**
  in part using bacteria and phage genetics
  DNA as the material support of heredity




**- Deciphering the role of the genes:   1970-**
Cellular, developmental, behavioral genetics: Screens
"High-throughput" versions   2000-   (e.g. deletion libraries)
Reverse genetics

**- Association mapping on natural populations: 2000-**

# Genetics of natural variation

**- Evolutionary genetics 1900-**

Biometricians versus Mendelians: continuous versus discrete hereditary variation

**- Neo-darwinian synthesis 1930**

mostly population genetics on allele frequencies, not phenotypes

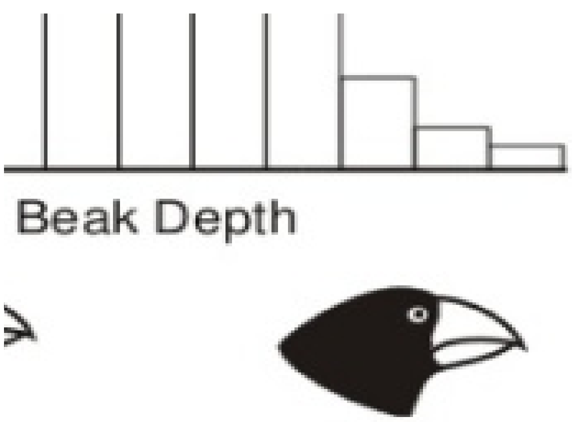
**- Human genetics**

***Now:***
***Bridging of "laboratory" and evolutionary genetics***
***Use of genomic data in model and "non-model" organisms***
***Role of the environment***

# Is natural variation discrete or continuous? Biometricians against Mendelians

*Karl Pearson*
*Walter Weldon*

- Continuous variation
- Pre-existing variation
- Gradual change



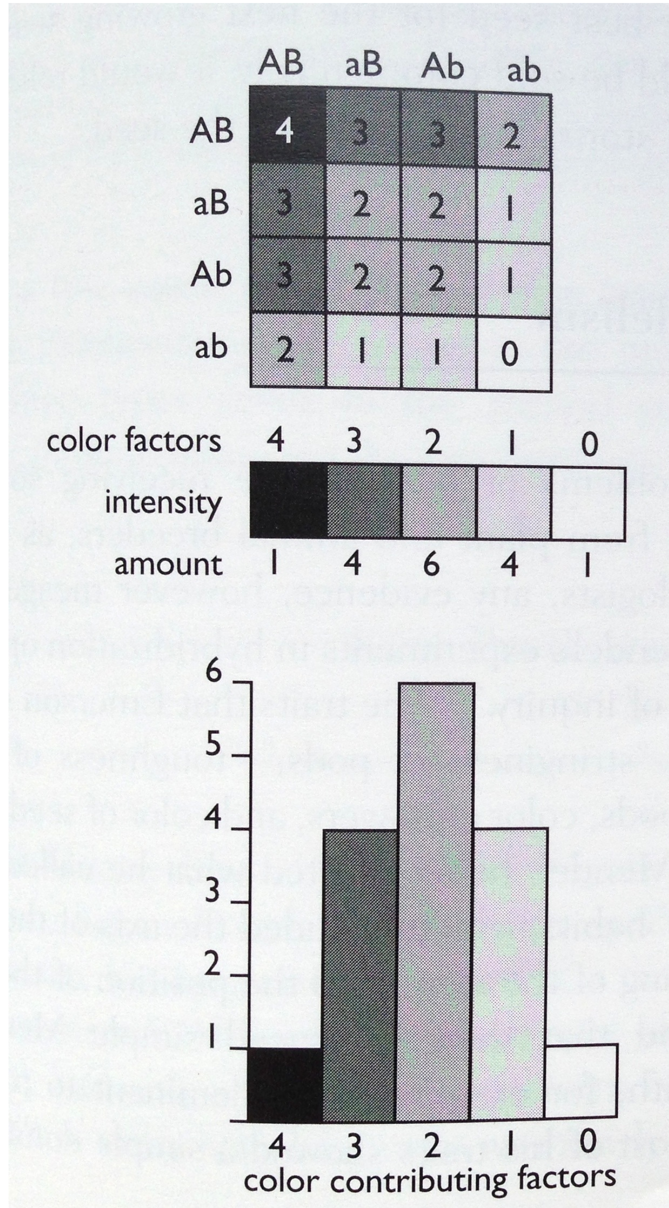

*William Bateson*
*Hugo de Vries*

- Discontinuous variation with discrete heredity factors
- Mutation
- Evolutionary jumps

# Reconciliation of Mendelian genetics and heredity of quantitative characters

Nilsson-Ehle (cereals)
East (corn)

example with only two factors with additive action:




Carlson 2004 book *Mendel's Legacy*

# Divergences and syntheses

Mendelian genetics
"Mendelians"

Statistical genetics
"Biometricians"

1900

agriculture

"Neodarwinian
synthesis"

1930

Laboratory
genetics

Population genetics
evolution
*genotype*

Quantitative
genetics
*phenotype*

1950 Molecular biology

Molecular genetics
cell biology,
development
physiology, etc.

1975

Molecular evolution

Molecular
markers

2000

Functional
genomics

Evo-
Devo

Evolutionary
genomics

Quantitative
genetics
**molecular basis**

**Present syntheses**

# Basic principles of genetics

# Genetic nomenclature

- Genetic nomenclature is different for each species

examples: *lacZ* in *Escherichia coli*
*CDC28* in *Saccharomyces cerevisiae*,
*cdc2* in *Schizosaccharomyces pombe*
*cdk-1* in *Caenorhabditis elegans*
*yellow* (*y*) in *Drosophila melanogaster*
*FUNNYNAME5* (*FNN5*) in *Arapidopsis thaliana*

Allele nomenclature...

- Wild-type allele "+" compared to mutation "*m*"

= for laboratory mutations
*no reference wild-type allele in natural populations*

# Aberration Types

SNP

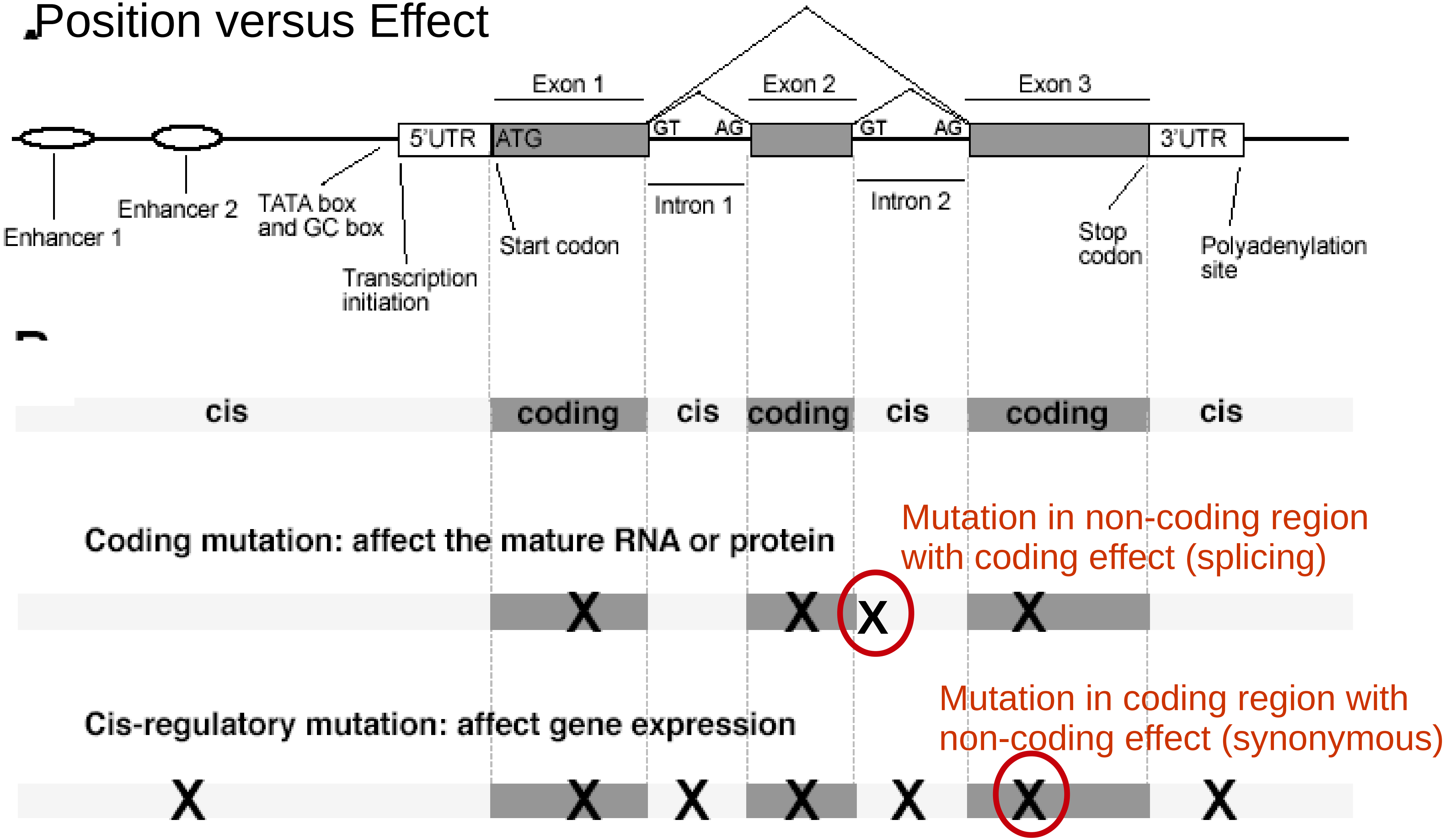
# Aberration Types

SNP
Insertion
Deletion
Indel
Inversion
Translocation
Complex change
(Epigenetic change)

**Estimation of mutation rates**
Mutation accumulation lines, sequencing family trio, across a phylogeny

# Coding versus cis-regulatory

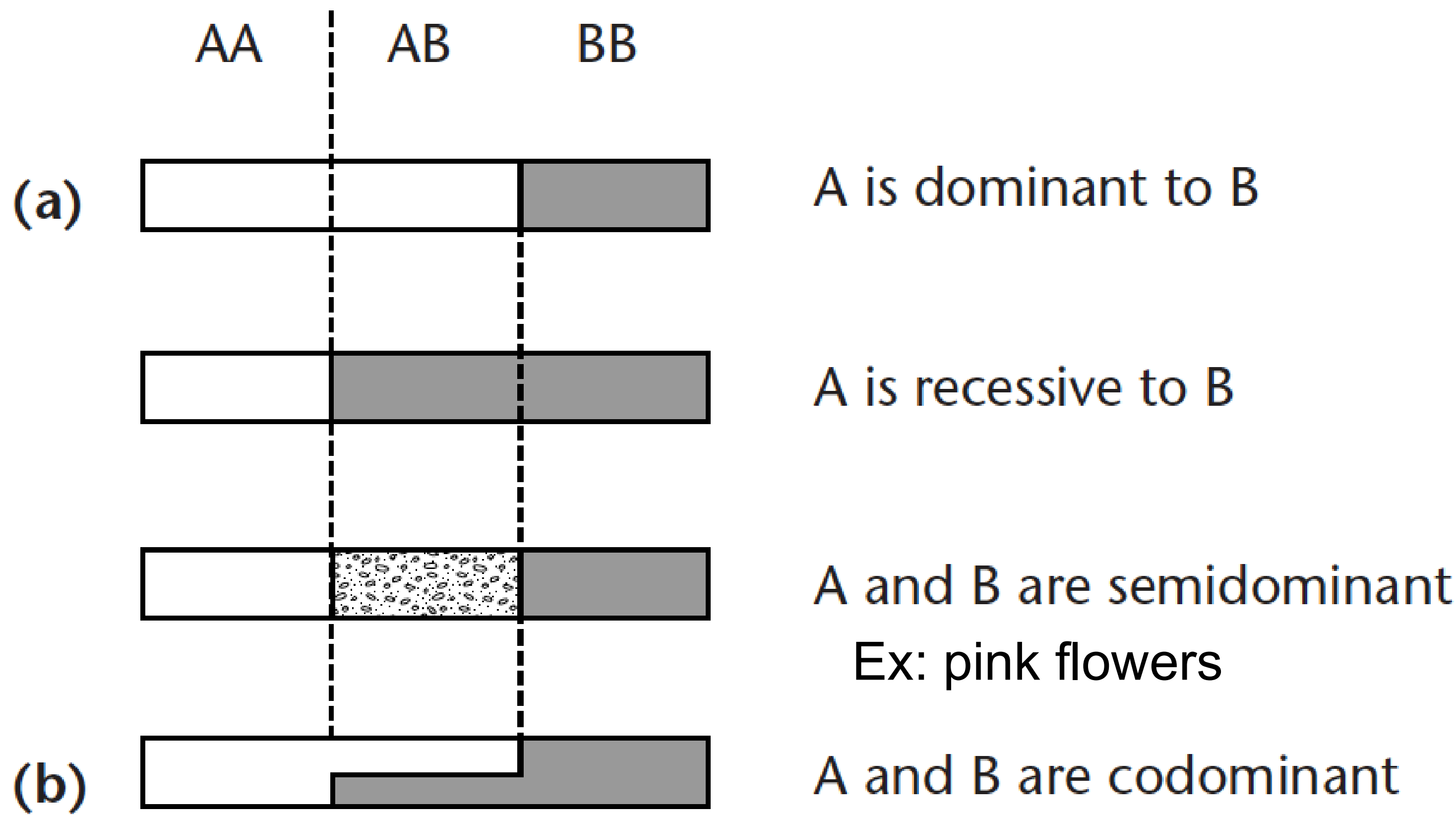
- Position versus Effect



# Coding versus cis-regulatory

Coding
Cis-regulatory
Gene loss
Gene amplification
(Gene rearrangement)

# Levels of dominance




Ex: pink flowers

Ex: AB blood groups, red and white flowers

Can be quantified as deviation from midpoint between parents

# Levels of activity of the various alleles

- **Null**: no activity, equivalent to gene deletion
- **Hypomorph ou loss-of-function**: reduced activity
- **Silent**: no influence on the phenotype
- **Hypermorph or gain-of-function**: increased activity
- **Neomorph**: new activity, can be ectopic expression

# Relationship with dominance

- Dominant (or semi-dominant) null allele:
  haplo-insufficiency – dose effect
- Dominant-negative via a poison effect
- Dominant gain-of-function/neomorph: common when gain-of-function, also with neomorph

# Dominance is not an intrinsic to an allele

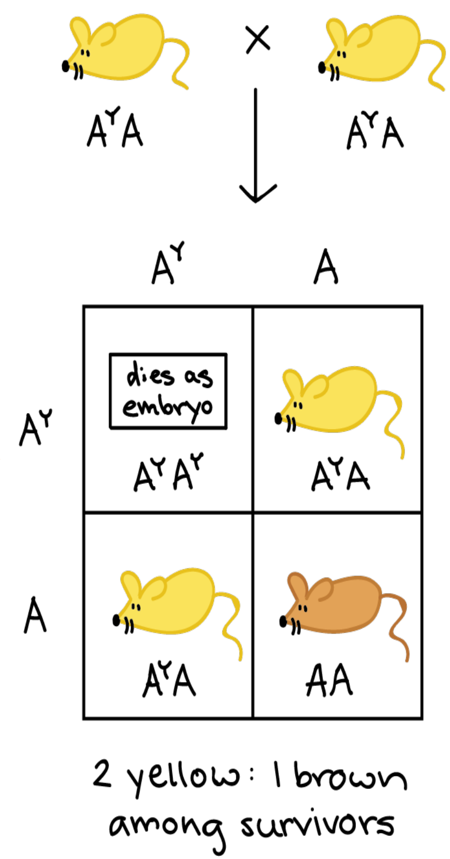
- It is relative to another allele, not to *all* other alleles

- It is a property of their effect on a given phenotypic trait

as in
"dominance of *a1* over *a2* for a particular trait"

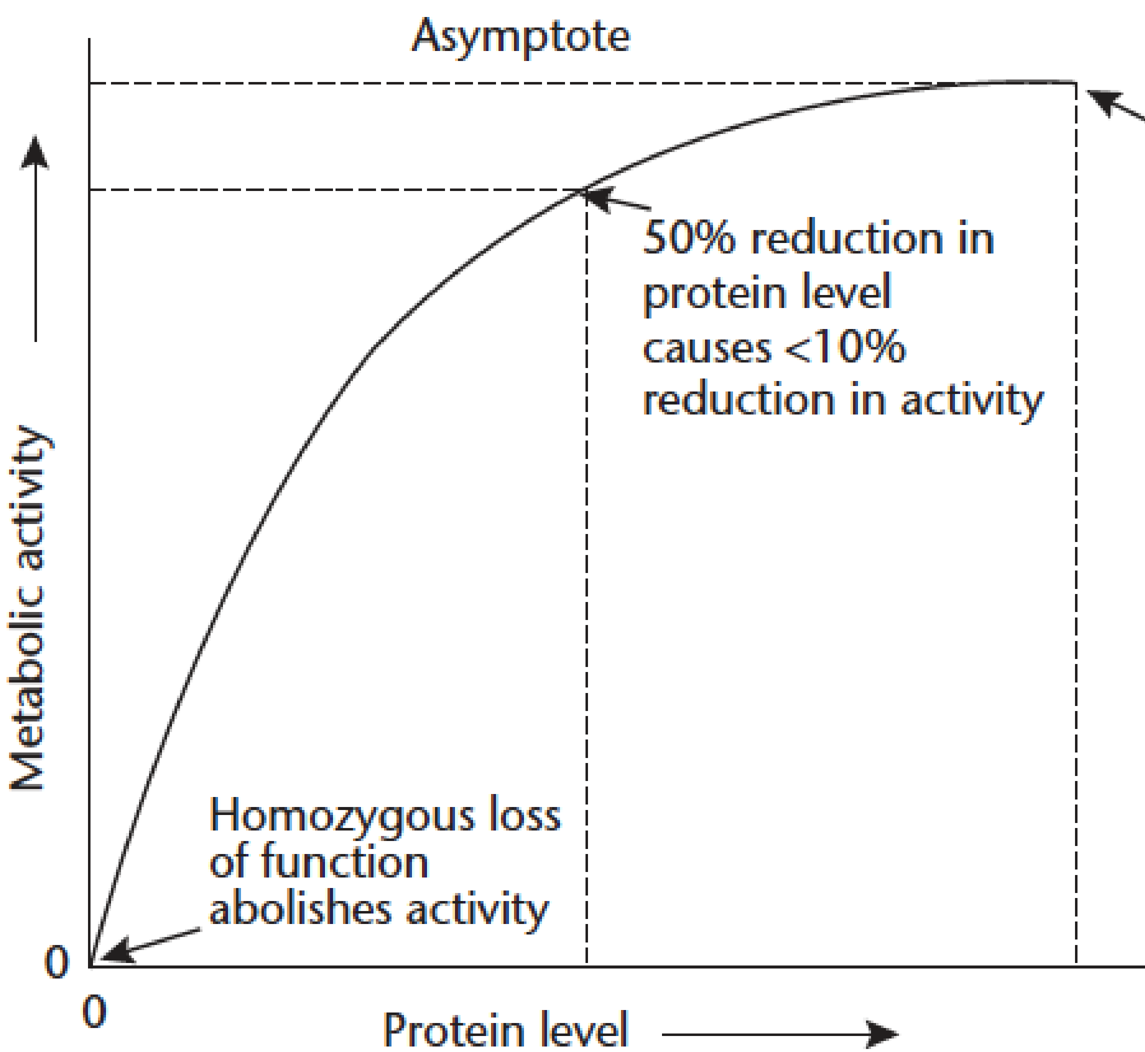
Example of the *agouti* locus in mouse

*yellow* allele is **dominant** over the + allele for coat color
*yellow* allele is **recessive** over the + allele for lethality



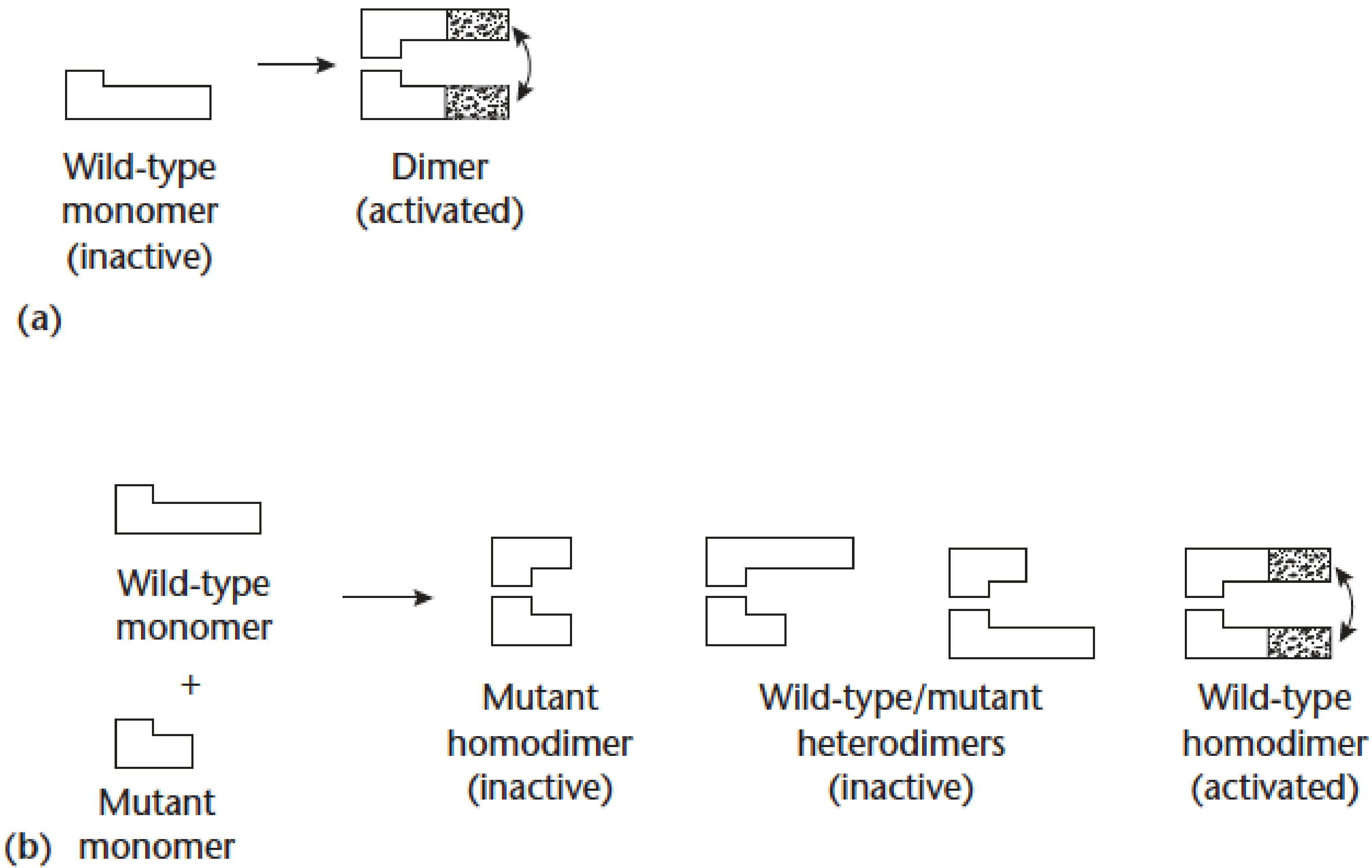
# Why are most wild alleles dominant?

Robustness to half-dose:




Wilkie 2005

# One type of dominant-negative mutation sequestration of wild-type in dimer




Wilkie 2005

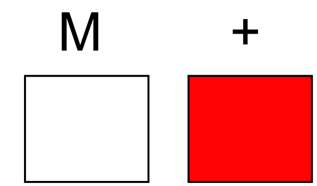
# Mechanisms of dominance

| Category of mutation | Mechanism | Affected gene/protein |
|---|---|---|
| **Loss-of-function** | | |
| Haploinsufficiency | Metabolic rate determining step | LDL receptor |
| | Developmental regulator | Transcription factors |
| **Dominant-negative** | | |
| Substrate sequestration | Binding by inactive monomer | Ligands, transcription factors |
| Dimer sequestration | Formation of inactive dimers | Receptors |
| Disruption of structure | Missense substitution | Collagens |
| **Dominant-positive** | | |
| Increased gene dosage | Duplication | *PMP22* |
| | Amplification | Oncogene products |
| Altered mRNA expression | Increased gene expression | $\gamma$ Hemoglobin |
| | Alternative splicing | *WT1* |
| | Toxic RNA inclusions | *DMPK* |
| Altered protein activity | Constitutive activity | Ion channels, receptors |
| | Increased binding affinity | Hemoglobin |
| | Formation of toxic proteins | Diverse |
| Novel protein activity | Altered substrate specificity | $\alpha$1 Antitrypsin |
| | Chimeric protein (translocation) | Transcription factors |

Wilkie 2005

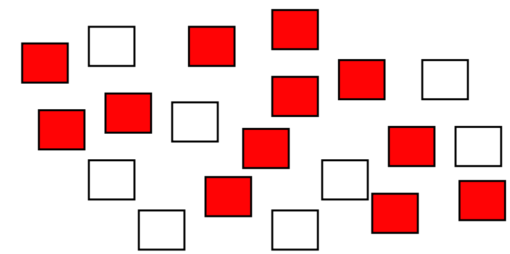
# Penetrance

*Discrete binary phenotype*



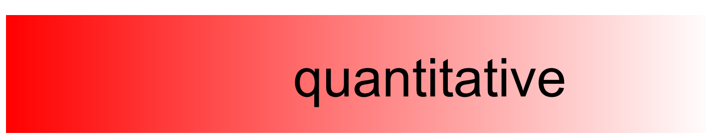

**% of individuals showing the phenotype**

ex: 40% of individuals have a white color
Partial penetrance

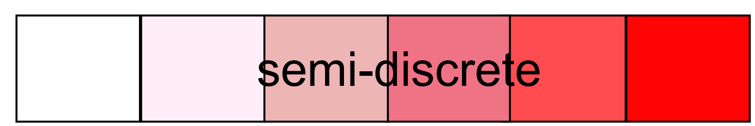


# Expressivity

*Phenotype with different degrees of severity*



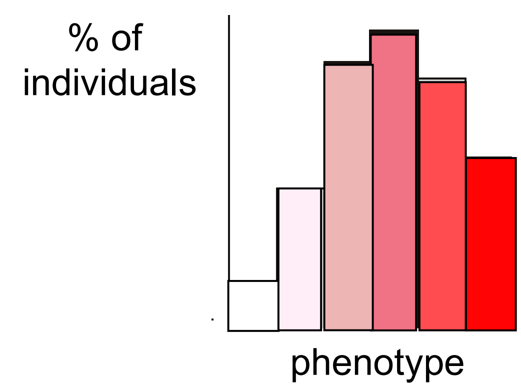

or




**Degree of severity of the phenotype**

ex: - number of affected ommatidia
- light pink color

often shown as a distribution of phenotypic values of individuals:



# Maternal (or paternal) effect

**The genotype of the parent matters,
not that of the individual itself.**

Frequent for mutations affecting early embryonic development

***m/m*** ♀ x *+/+* ♂

phenotype ? phenotype ?

F1 *m/+* x *+/+* ♂ or *m/m* or *m/+*

phenotype **M**

F2 *m/+* or *m/m* or *+/+*

phenotype **+**

**+/+** ♀ x *m/m* ♂

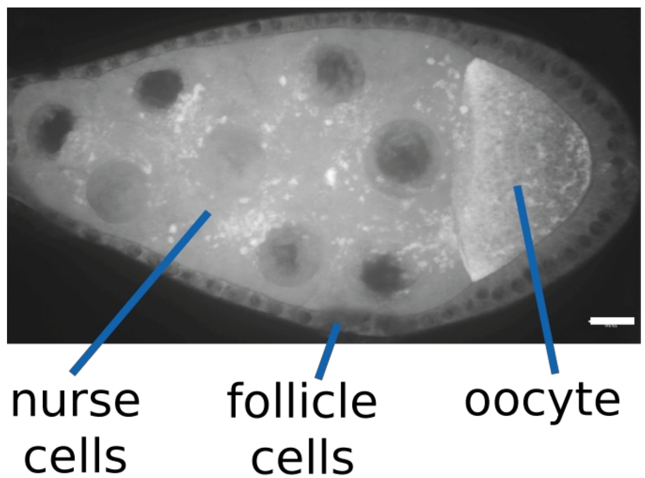
phenotype ? phenotype ?

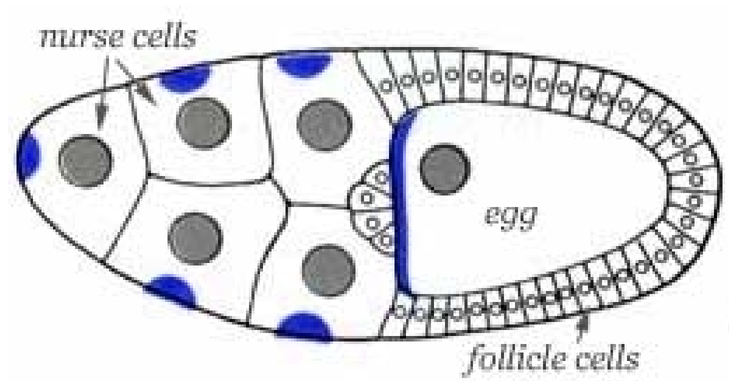
F1 *m/+*

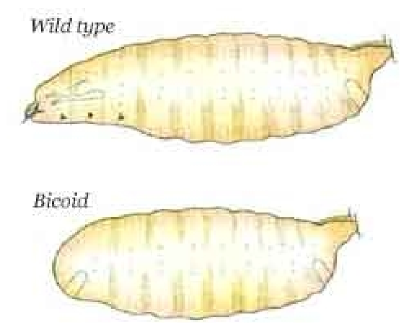
phenotype **+**

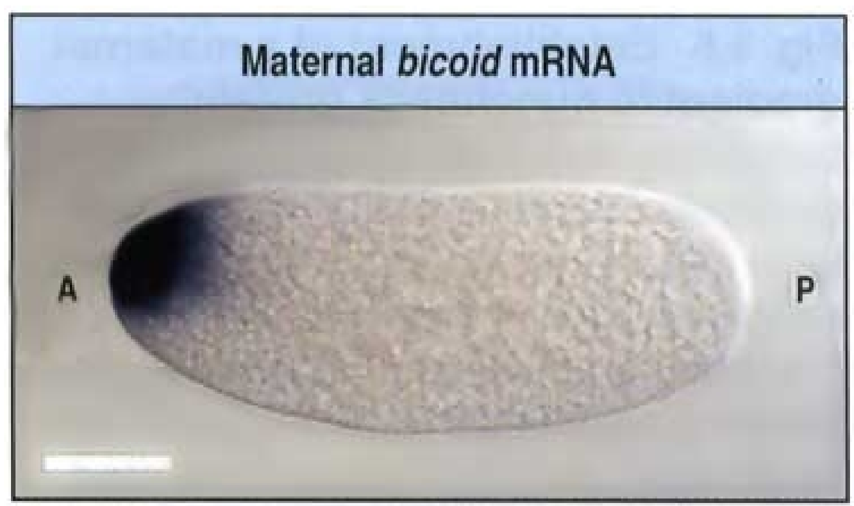
# An example of maternal effect

**deposition of mRNA or protein by maternal accessory cells into oocyte**



# Heredity with sex-linked transmission

Example: mutation on the X chromosome
in a species reproducing with XX ♀ x XY ♂

*m/m* ♀ x *+/Y* ♂

phenotype M phenotype +

↓

F1 *m/+* ♀ + *m/Y* ♂

females of phenotype + males of phenotype M

# Heredity with sex-linked expression

Example: mutation that affects the phenotype only in females

# Alleles

***m1* and *m2* are allelic if they are in the same gene**

in practice
- if the two mutations do not complement (at the phenotype level)
- or defined using sequence variations

# Complementation test

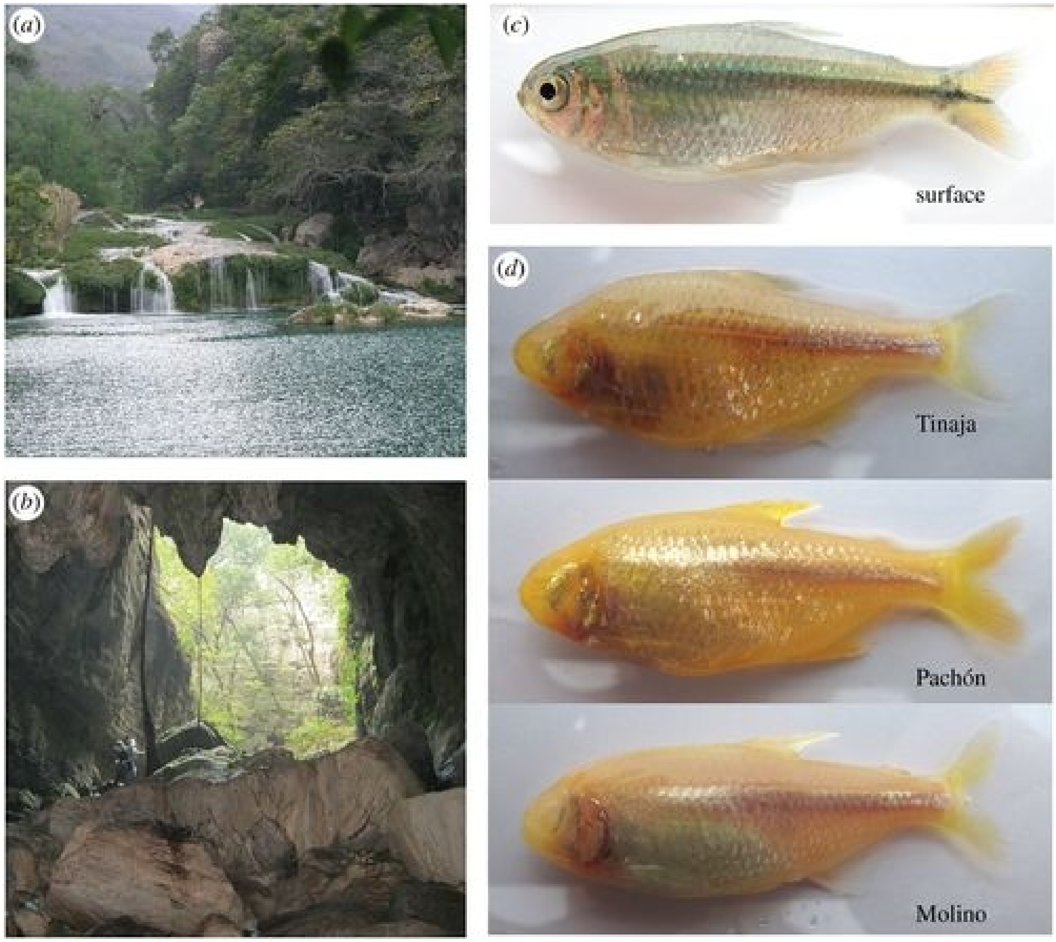
if *m1/m1* and *m2/m2* show the phenotype M
and act recessively

*m1/m1* x *m2/m2*

$\downarrow$

F1

| | | Interpretation | |
|---|---|---|---|
| • if Phenotype M | **non-complementation** | *m1/m2* | same gene (alleles) |
| • if Phenotype + | **complementation** | *m1/+; +/m2* | two genes (non allelic) |


(a)
(b)
(c)
surface
(d)
Tinaja
Pachón
Molino


Molino × Pachón

# The importance of DNA

# Why is DNA an important molecule in biology?

# *Newsweek*, May 23, 2005

Slide from
S. Gilbert

DNA still sells cars in the USA

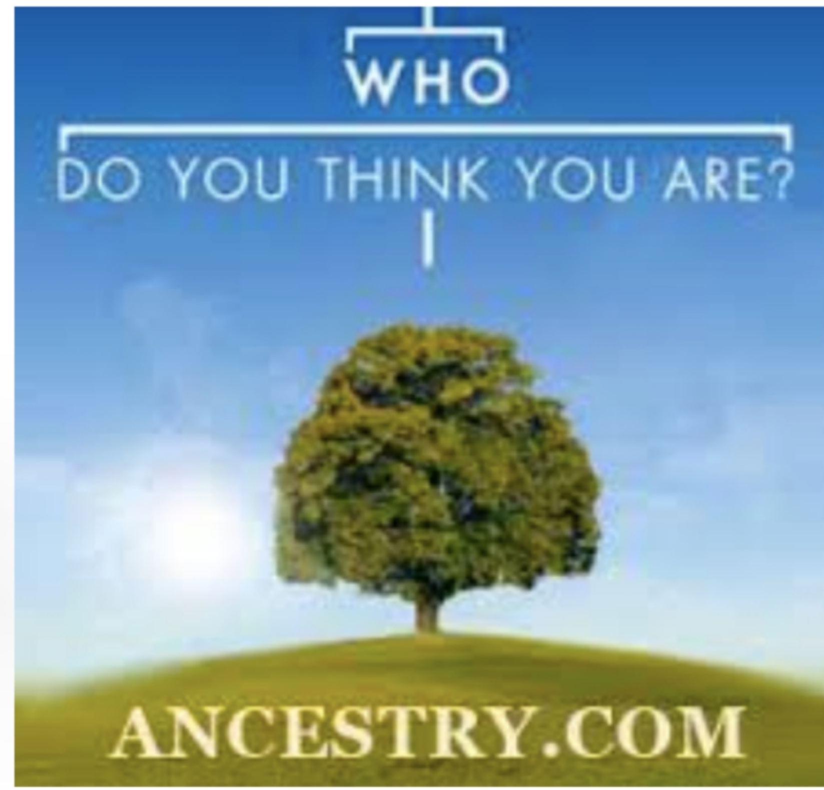
Subaru: “Genetic superstar”
Toyota: “Has a great set of genes


*Slide from S. Gilbert*

# GENETIC INDIVIDUALITY:

*Slide from S. Gilbert*

Each of us is a genetically unique individual, and the genes determine who we are.


"…revealing what it is that makes you, you."

-American television ad for ancestry.com 2015

# The importance of DNA in biology

Major basis of heritable variation (genotype-phenotype)

Transmitted (can help reconstruct history)

Present in all living entities (DNA/RNA)

Stable molecule (ancient DNA – oldest = horse in permafrost = 500 000 years, forensic)

String of letters, can be easily analyzed with computers (compared to anatomical traits for taxonomy)

# Genetic Individuality

*Slide from S. Gilbert*

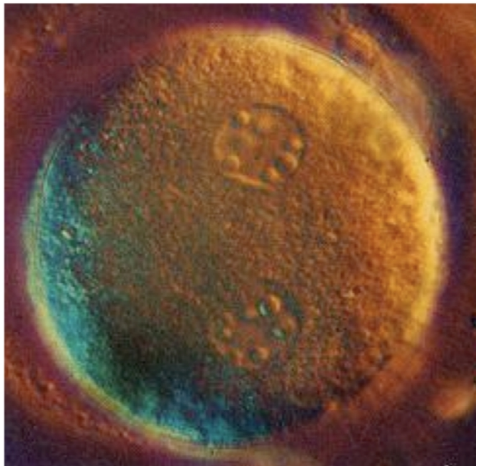
Genes determine who you are, and they act the same in each person.

LIFE Magazine, *First Days of Creation,* 1990:

"The result of fertilization is a single nucleus that contains an entire biological blueprint for a new individual, genetic information governing everything from the length of the nose to the diseases that will be inherited."

Standupgirl.com (anti-Choice website):

"And even more amazingly, intelligence and personality—the way you look and feel—were already in place in your genetic code. At the moment of conception you were essentially and uniquely you."



# Disclaimer:
# DNA is not the cause of everything

Monozygotic twins are not identical

Cardiovascular disease associates better with lifestyle than with DNA sequence (Mozaffarian 2008)

Lung cancer associated with smoking habits

Drug metabolism is mostly due to the microbiome

Several genes associated with autism, depression, etc. were “lost” in larger studies

Distilbene: anti-miscarriage drug, increases cancer risks in daughters and malformations in grand-daughters

## What the HGP Taught us with the first genome sequenced: Genes act differently and non-additively in different people

Cockayne syndrome: **Mutation in the DNA repair enzyme ERCC6 at position 5q12.1.** Homozygous recessive persons are characterized by growth failure, impaired neural development, premature aging, sensitivity to sunlight.

Usher Syndrome: **Mutation in the retinal and cochlear basement membrane myosin MYO7A at 11q13.** Homozygous recessive persons are characterized by congenital deafness and gradual loss of vision.

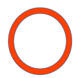
James Watson, presently 90 years old; not deaf, blind, nor stunted

*Slide from S. Gilbert*

# First “synthetic” cell developed by scientists

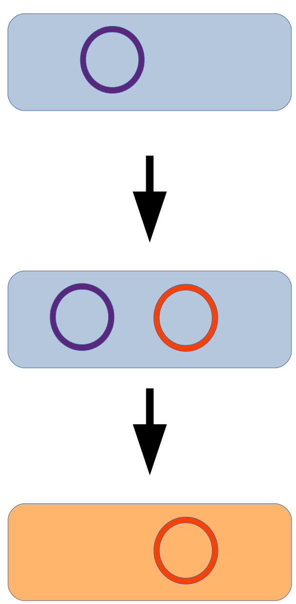
*M. mycoides* JCVI-syn1.0
1.08 million base pairs and 901 genes

single yeast artificial chromosome

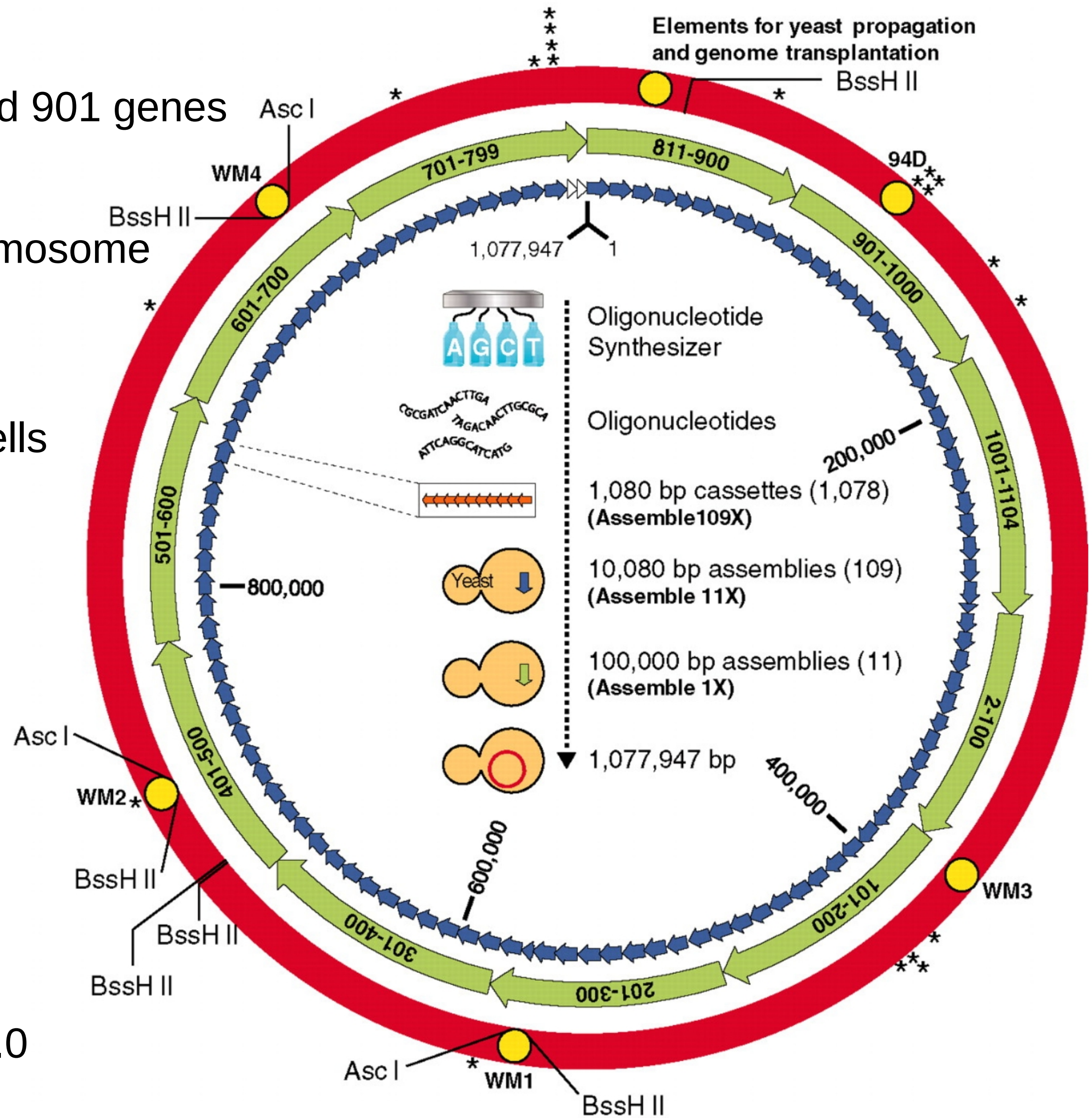


+

*M. capricolum* recipient cells



*M. mycoides* JCVI-syn1.0




Gibson et al. 2016

# First “synthetic” cell developed by scientists

*M. mycoides* JCVI-syn1.0 contains strings of bases that, in code, spell out:

a web address to send emails to if you can successfully crack the new code,

the names of 46 authors and other key contributors,

three famous quotations. One of which by James Joyce, perfectly encapsulates the ups and downs of a the 15 year project—“To live, to err, to fall, to triumph, to recreate life out of life.”

Gibson et al. 2016

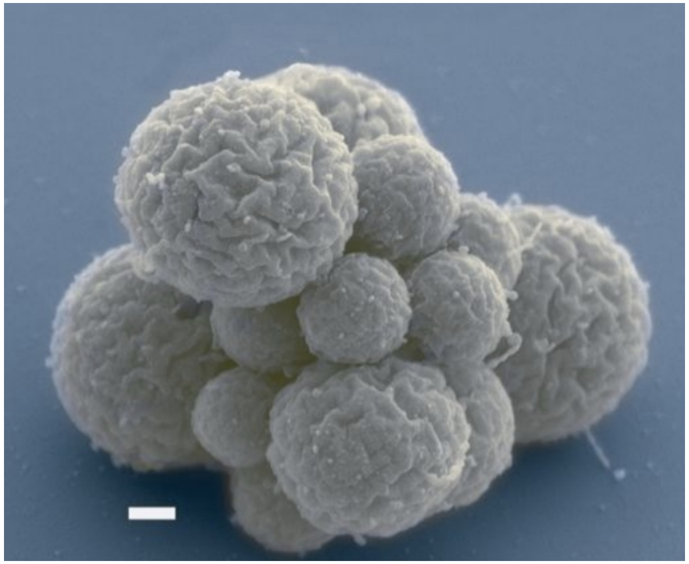
# Smallest “synthetic” cell

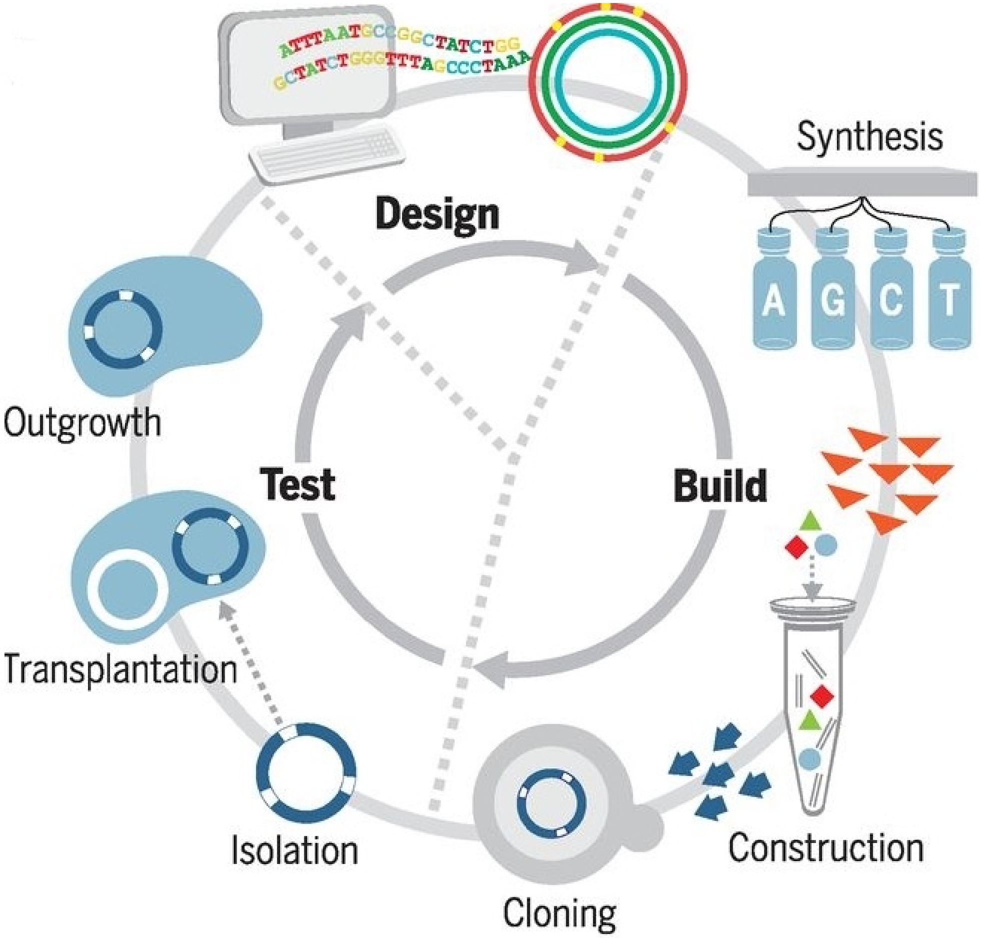
*M. mycoides* JCVI-syn3.0
531,560 base pairs and 473 genes

retains genes involved in key processes such as transcription and translation, but also contains 149 genes of unknown function

genome smaller than that of any autonomously replicating cell found in nature.

doubling time of ~180 min
colonies morphologically similar to those of JCVI-syn1.0






Hutchison III et al. 2016

# Theodore Schwann, 1839

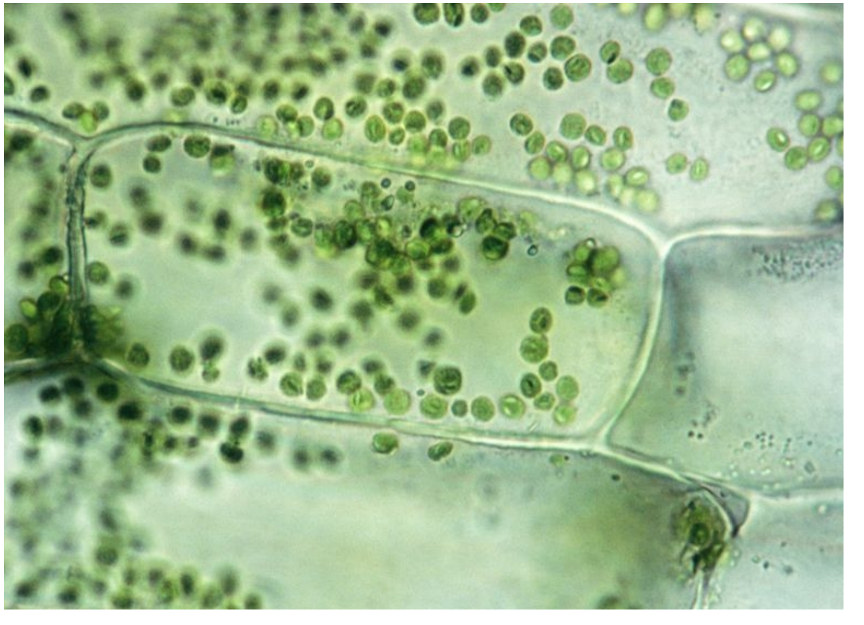
"*All living things are composed of cells and cell products*"

1) The cell is the unit of structure, physiology, and organization in living things.

2) The cell retains a dual existence as a distinct entity and a building block in the construction of organisms.

# Rudolf Virchow, 1857

*Every cell arises from another cell*

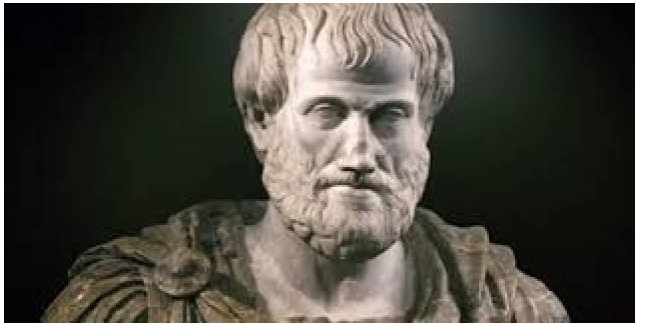


# The genotype-phenotype relationship

# Different kinds of phenotypes

## Morphology

- Color
- Size and shape
- Presence/absence
- Position



Aristote, Historia animalium, book I, 2, 300BC

## Physiology

## Behavior

# Genotype & Phenotype

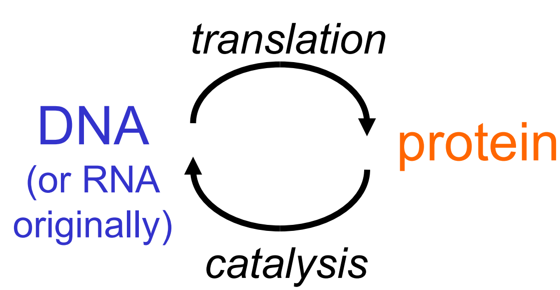
= what engenders = what is apparent

**DNA/RNA**

- **Regulation of gene expression**
- **Biochemical reactions**
- **Subcellular architecture**
- **Assembly of cells**
- **Organism morphology and behavior**

**etc.**

distinction appeared at the origin of life:



# Francis Crick Central Dogma
# A reductionist view of the GP relationship

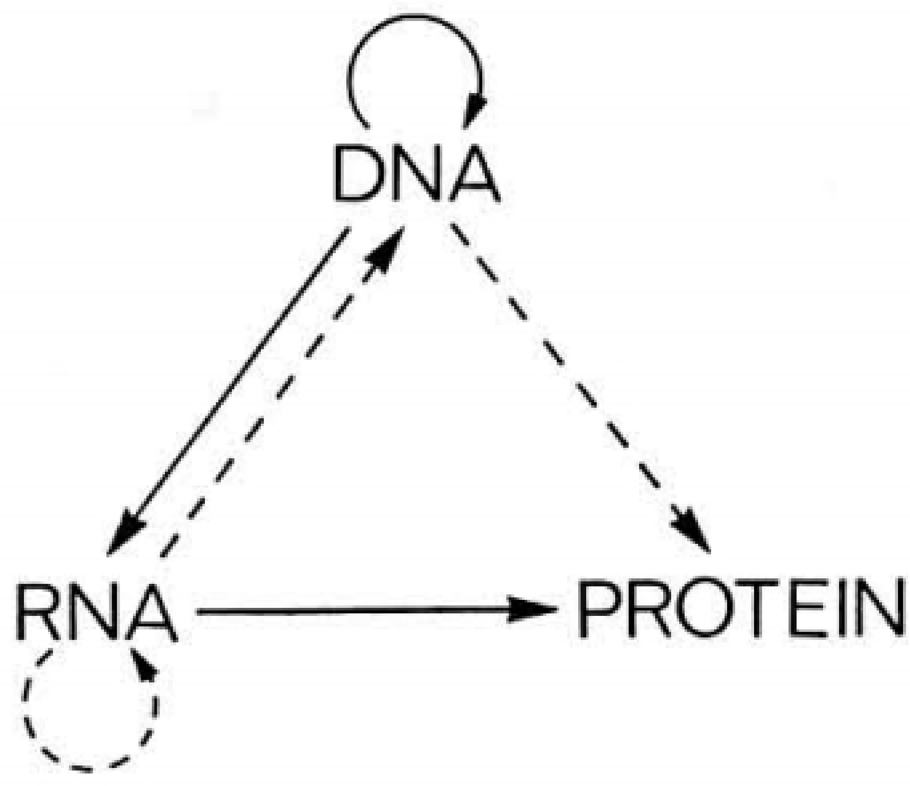



Fig. 3. A tentative classification for the present day. Solid arrows show general transfers; dotted arrows show special transfers. Again, the absent arrows are the undetected transfers specified by the central dogma.

Crick 1958
Crick 1970

# How do genotypes map onto phenotypes ?

## DEVELOPMENTAL BIOLOGY

## EVOLUTIONARY GENETICS

Both are direct descendants of Morgan's school. Emphasis on genes.

How does an organism form from a single cell?

What makes one organism different from another one?

*One of the central problems of biology is that of differentiation - how does an egg develop into a complex many-celled organism? That is, of course, the traditional major problem of embryology; but it also appears in genetics in the form of the question,*
***"How do genes produce their effects?***

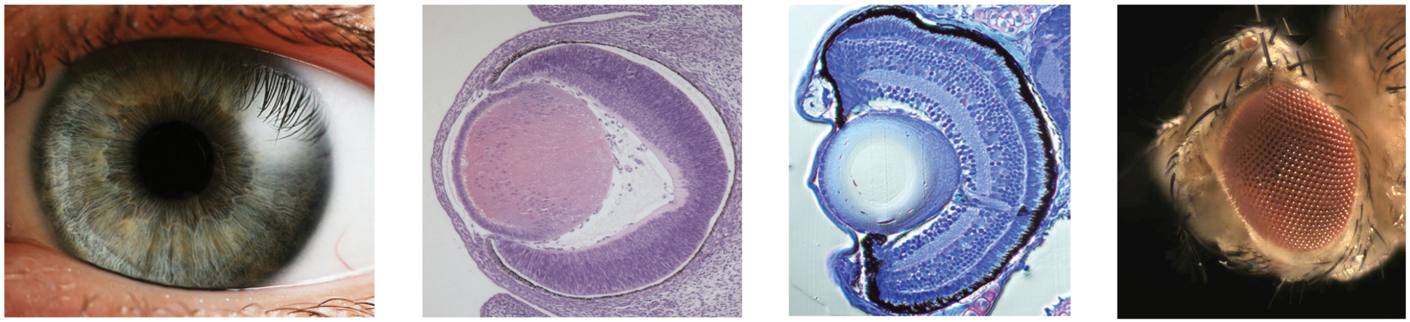
Sturtevant, 1932

How do genes produce observable traits?

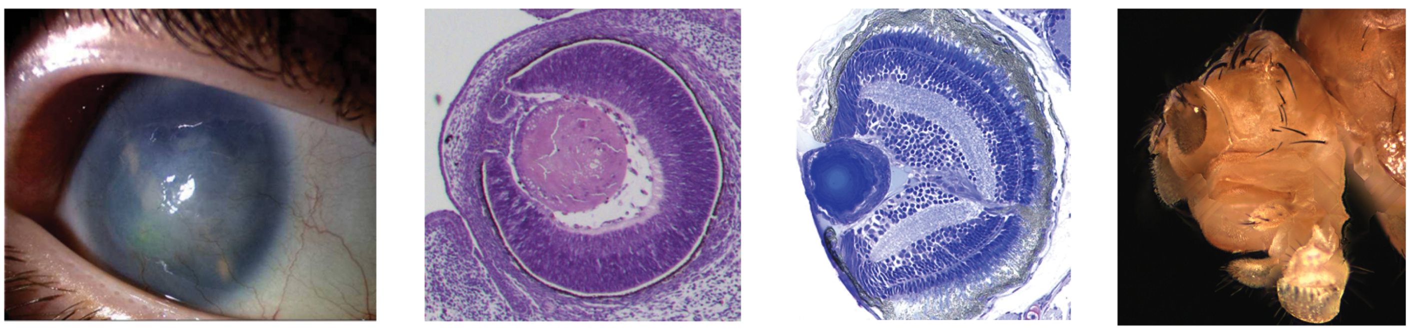
Gene → Observable character

# *Pax6* : an eye gene ?

| | Human | Mouse | Zebrafish | *Drosophila* |
|---|---|---|---|---|
| WT | | | | |



mut



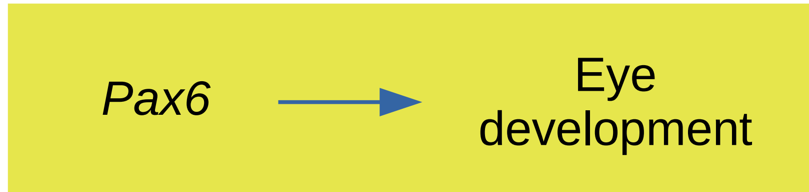
| | *PAX6*$^{+/-}$ | *Pax6*$^{-/-}$ | *pax6b*$^{-/-}$ | *ey*$^{-/-}$ |
|---|---|---|---|---|
| EQs | cornea opaque<br>iris absent<br>retina degenerate<br>lens opaque<br>aqueous humor of eyeball increased pressure | eye decreased size<br>lens fused_to cornea<br>iris morphology<br>anterior chamber absent | eye decreased size<br>lens decreased size<br>retina malformed | eye absent |

***Drosophila***

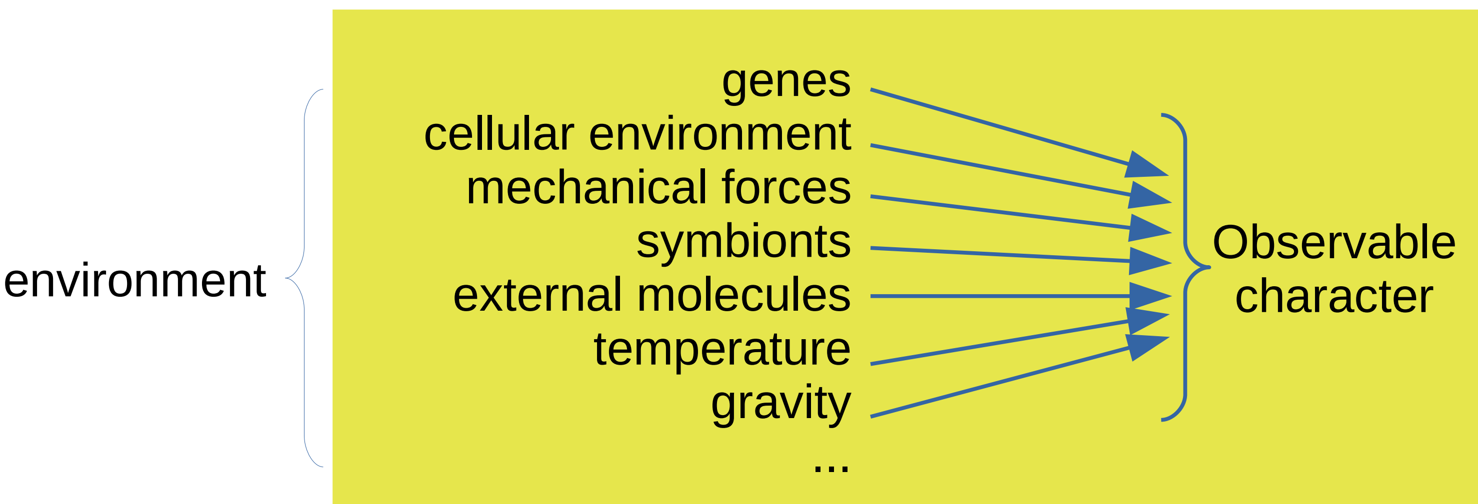

***Ectopic expression of Pax6***

Gene → Observable character

*Pax6* → Eye development

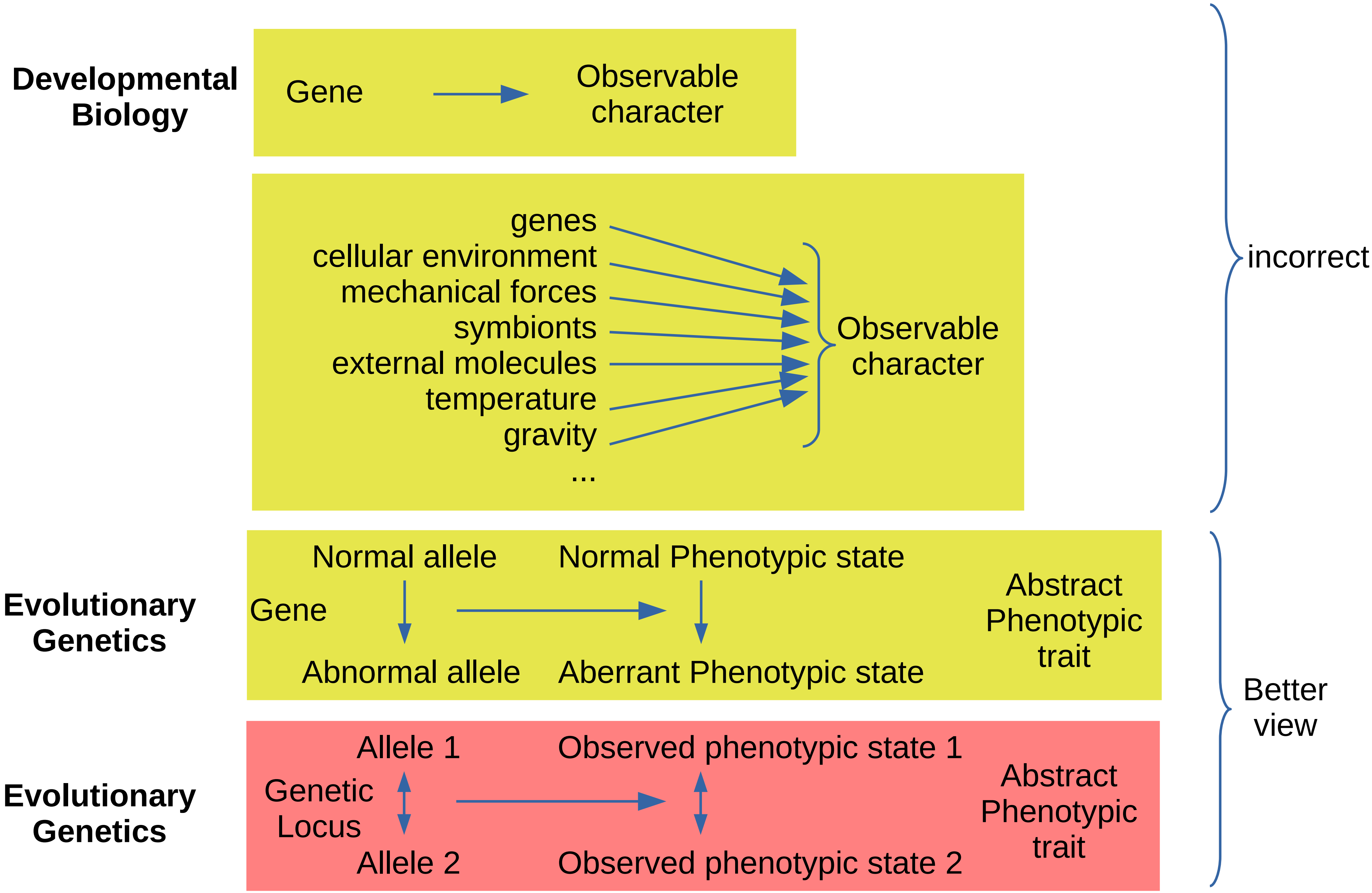



***Too simplistic***

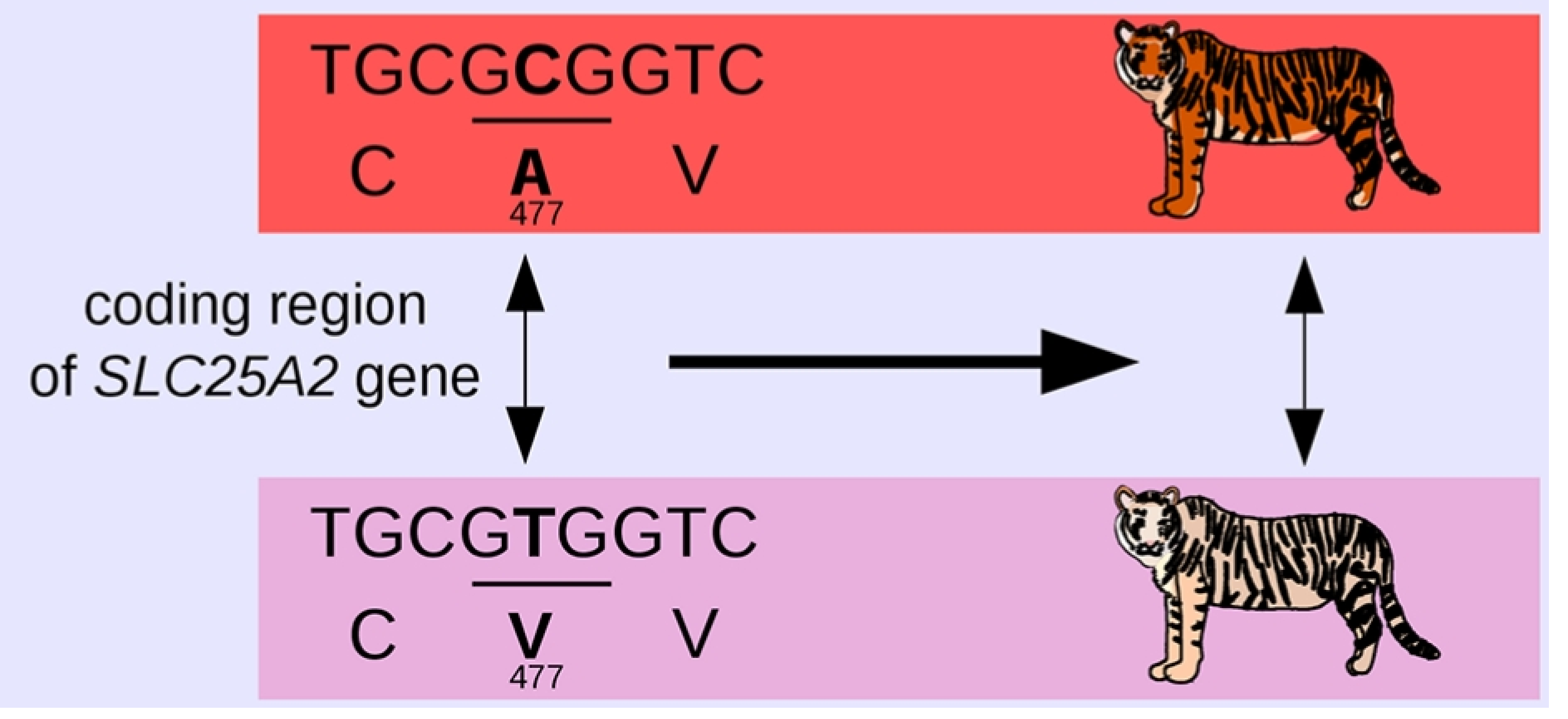



***Better, but difficult to disentangle the effects***

# The wrong and the right perspectives



# The genotype-phenotype connection




Xu et al 2013 Current Biology

Orgogozo et al 2015 Frontiers Genetics

*Genotype = "the genetic makeup of an organism that determines a specific phenotype (trait), from one generation to the next, and potentially throughout the population".*

*Report of the National Academy of Sciences on gene drive, 2016*

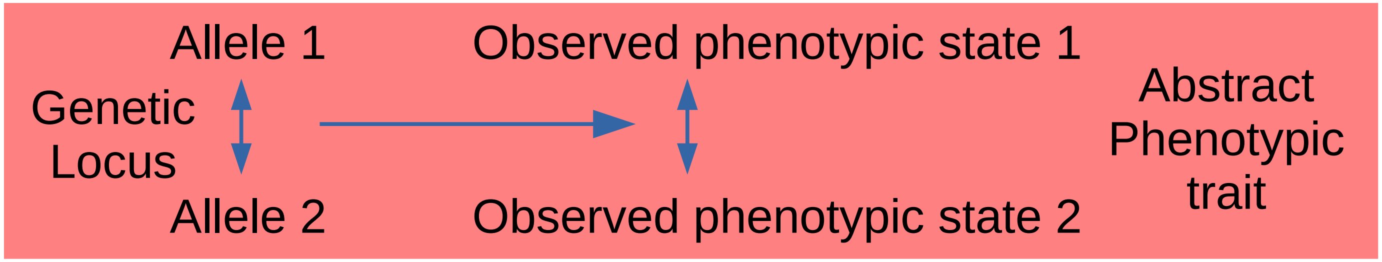
**NO!**

**The genotype-phenotype connection is about differences**

# Example: the causes of a difference in hair color

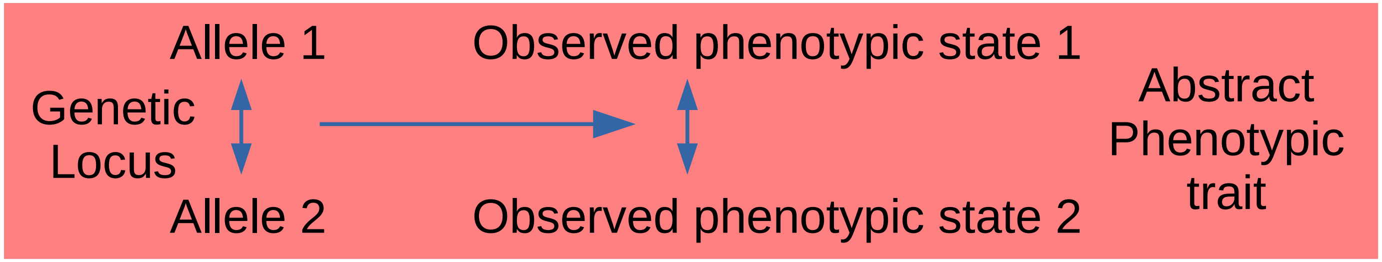
Genes coding for pigment synthesizing enzymes
Amount of received sun light
Hair dyeing
Age
...

# Gephe: a relationship between two levels of variation



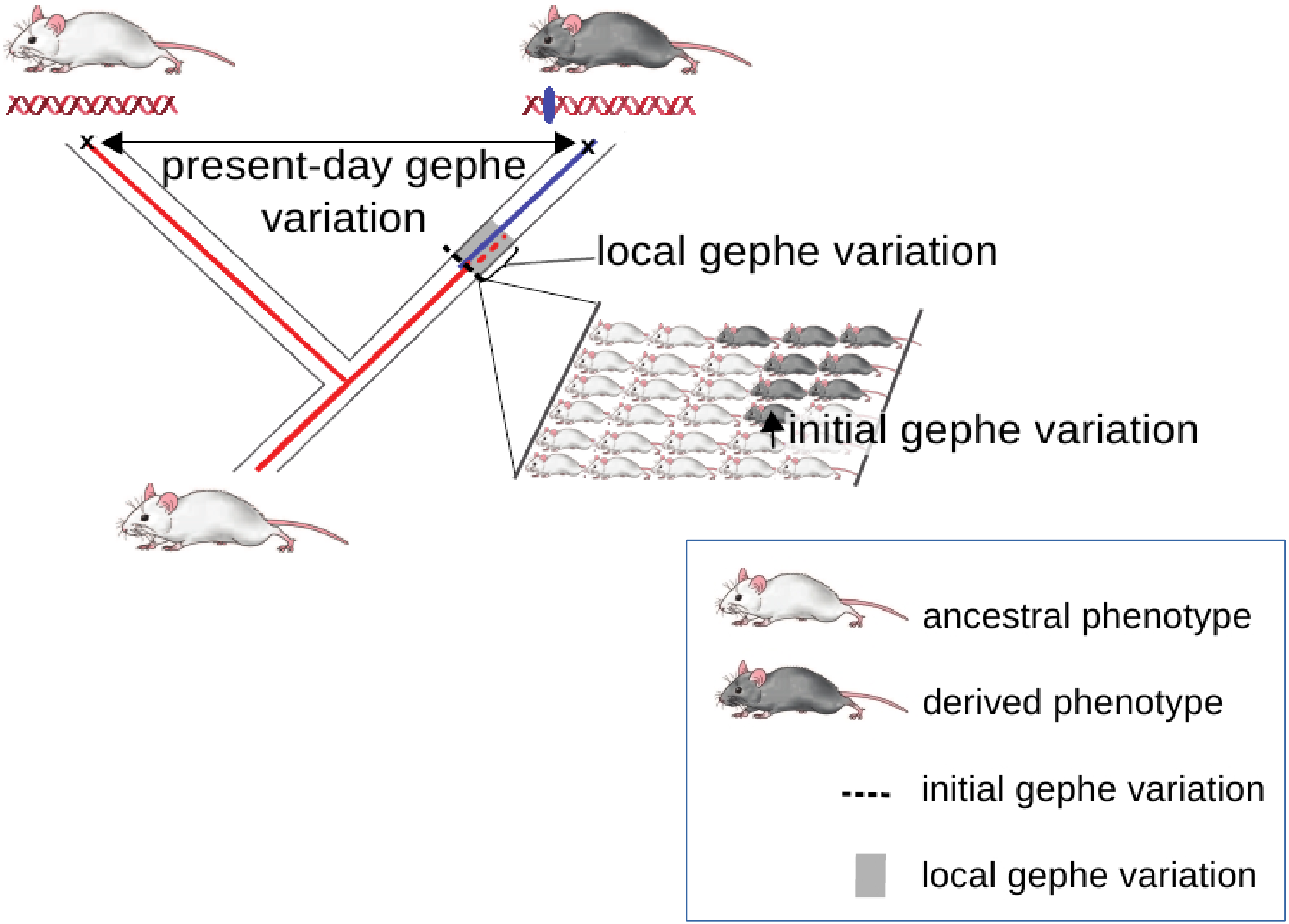

| **Genetic locus** | **Phenotypic trait** |
|---|---|
| One particular coding site in the *Nav1.4* gene | Resistance to tetrodotoxin or saxotoxin |
| Various coding sites in *opsin* genes | Color vision |
| *SLC45A2* coding region | Pigmentation of eye, hair and skin |
| *Mc1R* coding region | Pigmentation of hair and skin, but not eye |
| Cis-regulatory element in the *lactase* gene | Ability to digest milk |
| Cis-regulatory element in the *pitx1* gene | Pelvis morphology |
| Cis-regulatory elements in the *optix* gene | Red color pattern on butterfly wings |
| Cis-regulatory and coding regions of the *FRIGIDA* gene | Flowering time |

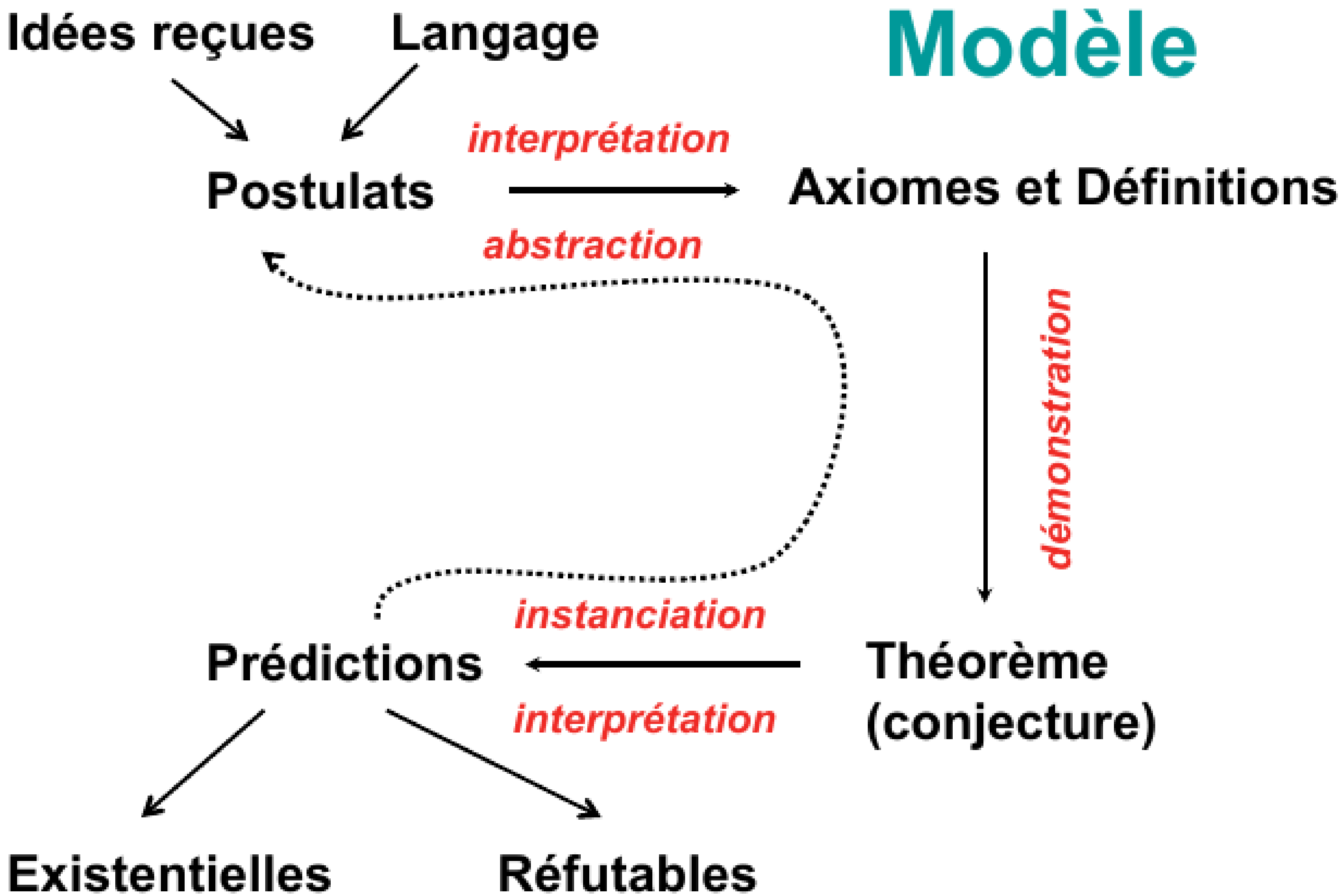
# Gephe: a relationship between two levels of variation



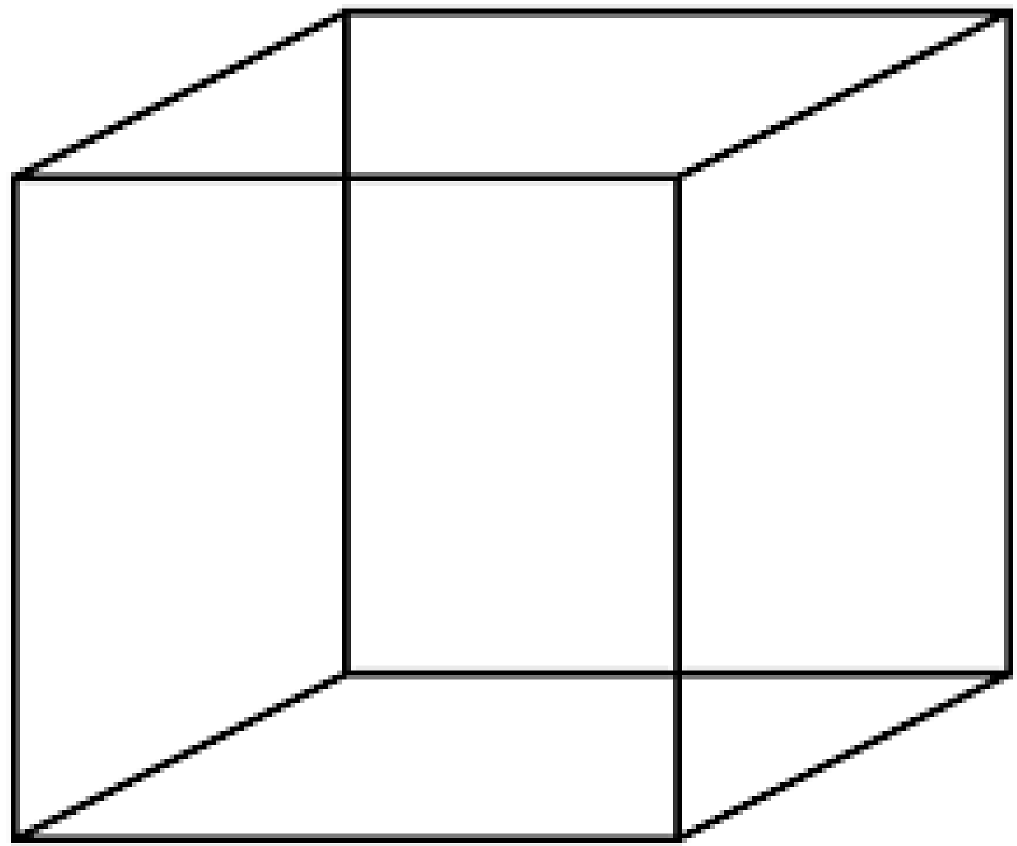

The genetic locus in the gephe concept is **not necessarily a gene.**
it can span a particular base-pair, a coding region, a cis-regulatory region, or extend to an entire gene or even a gene cluster.
This view rejoins previous reflections by Scott Gilbert (2000), David Stern (2000) and Lenny Moss (2003) on the distinct concepts of gene in developmental biology and in evolutionary biology.

# Gephe: a difference at various levels

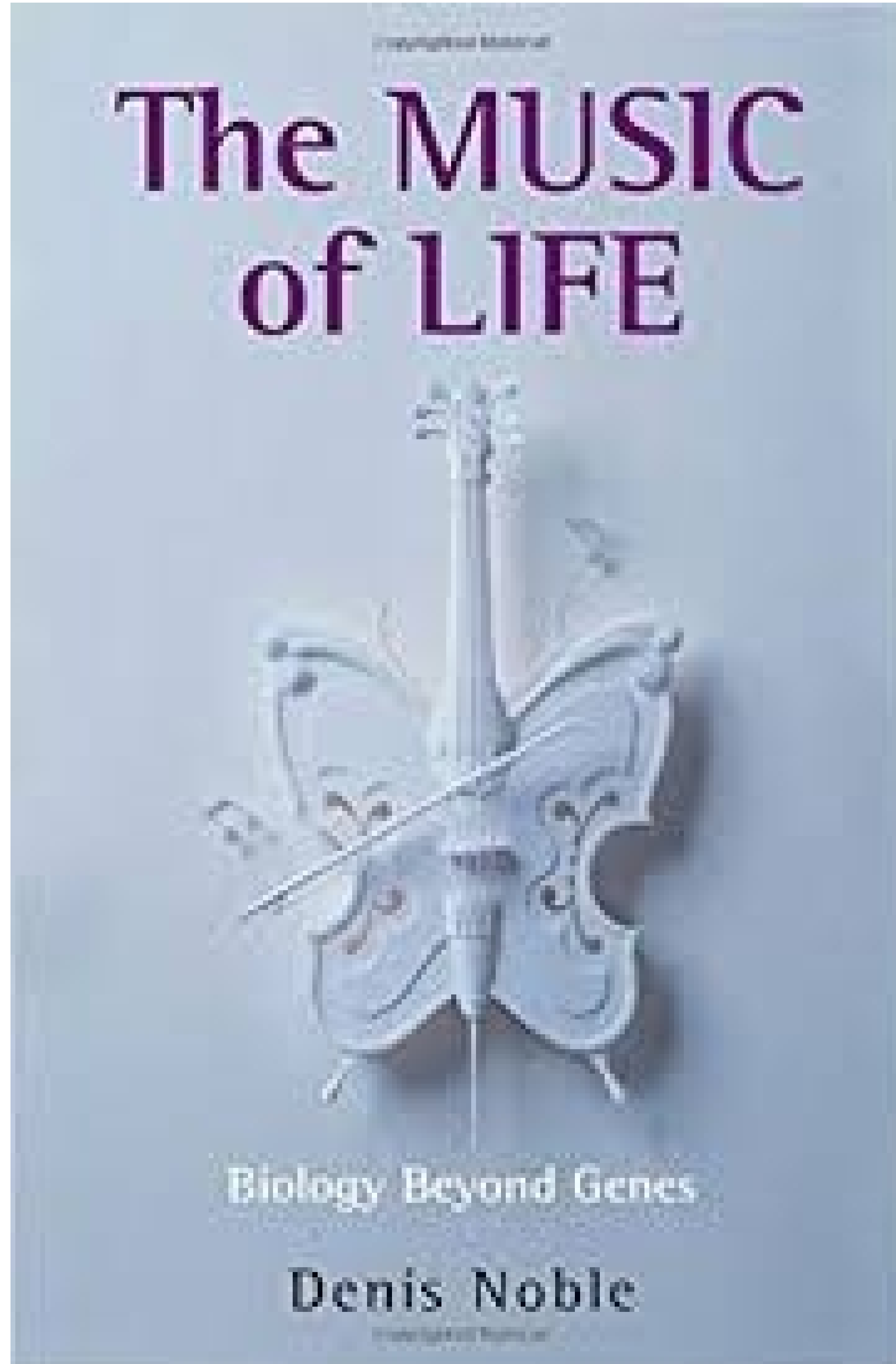





*We sometimes seem to have forgotten that the original question in genetics was not what makes a protein but rather 'what makes a dog a dog, a man a man'.*
*(D. Noble – The Music of Life)*




POINTS
Michel
Raymond
Cro-Magnon
toi-même !
Petit guide darwinien
de la vie quotidienne
SCIENCES


The MUSIC
of LIFE
Biology Beyond Genes
Denis Noble