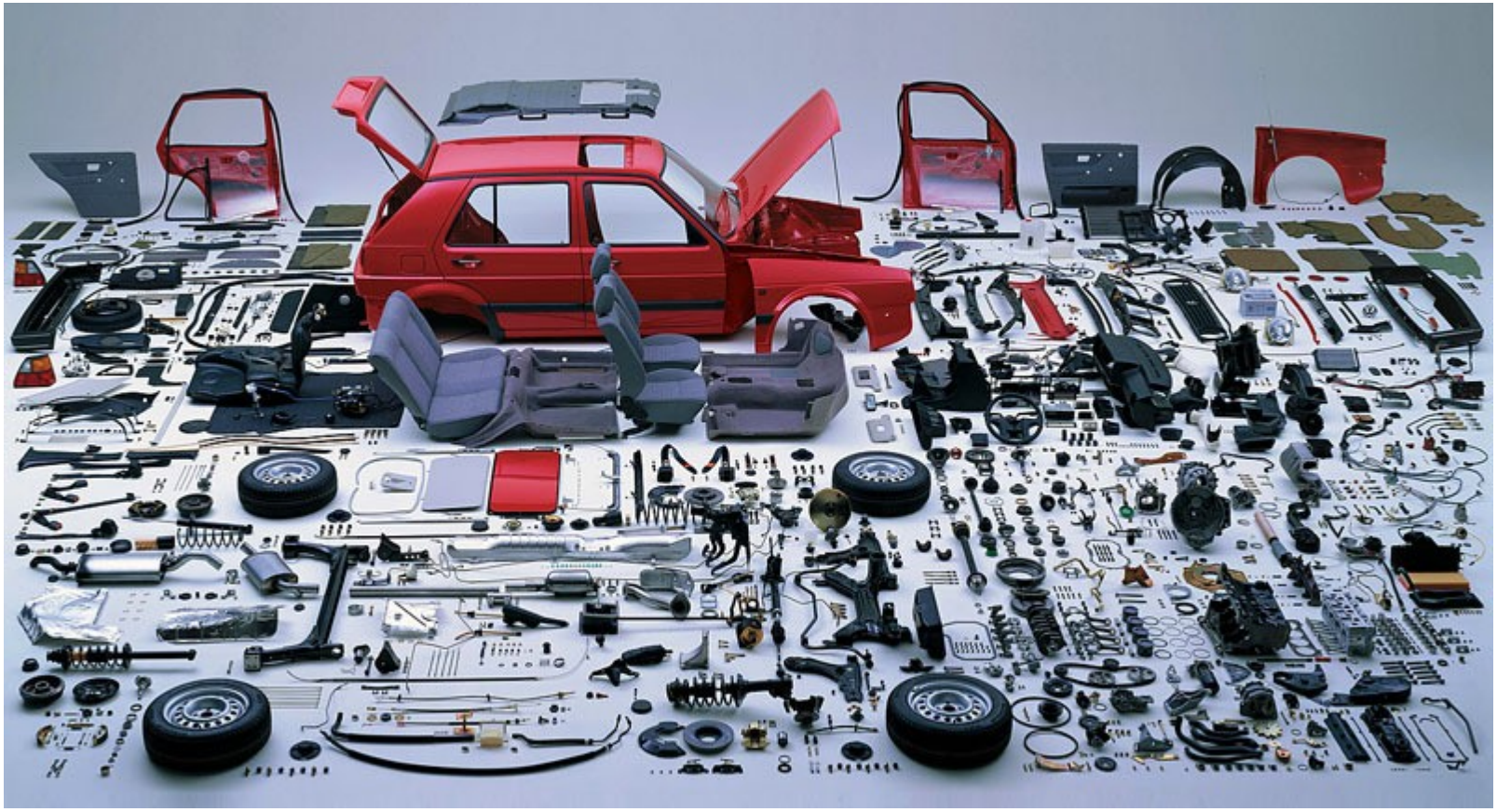
Genetic screens, Quantitative Genetics, Noise, Cryptic Variation, Robustness

Virginie Courtier-Orgogozo
Institut Jacques Monod, Paris

Biochemistry versus Genetics



Genetic screens

unbiased approach:
Biochemical screen (kinase assay, methylation assay, etc…)
Genetic screen (for genes)

A genetic screen

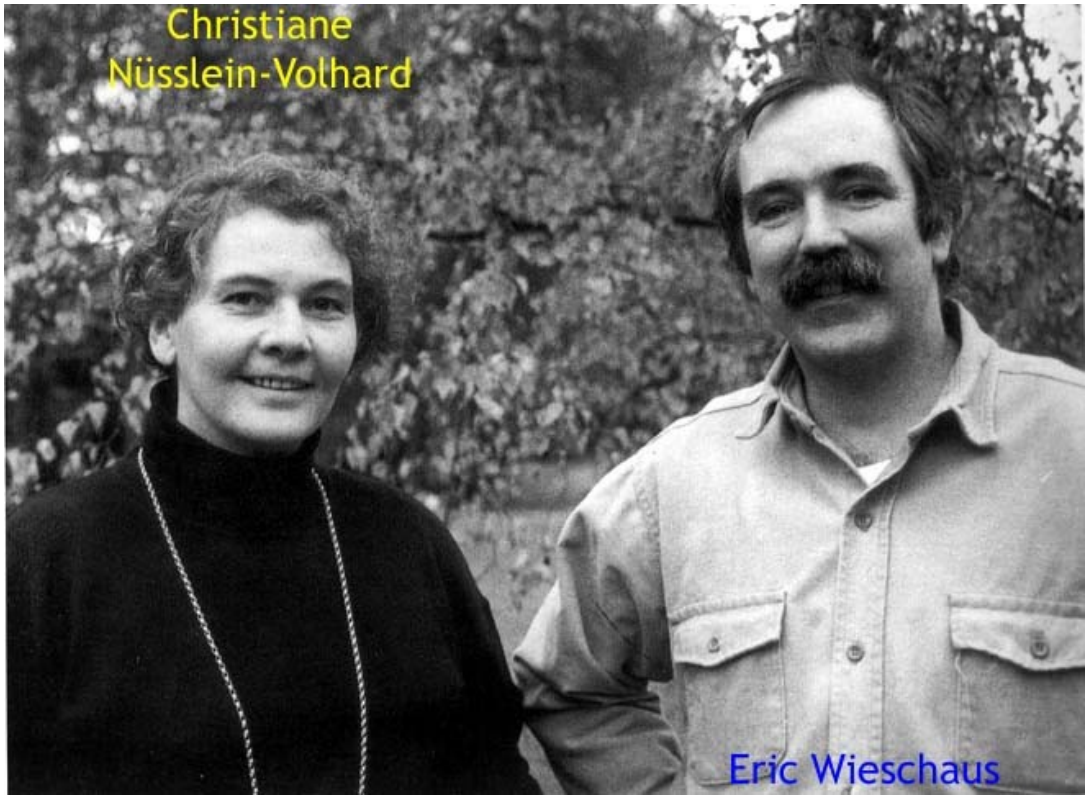
is a method used to find genes involved in a given phenotype
3 steps :
1. production of mutations
2. selection of individuals with the phenotype of interest
3. identification of the underlying genes

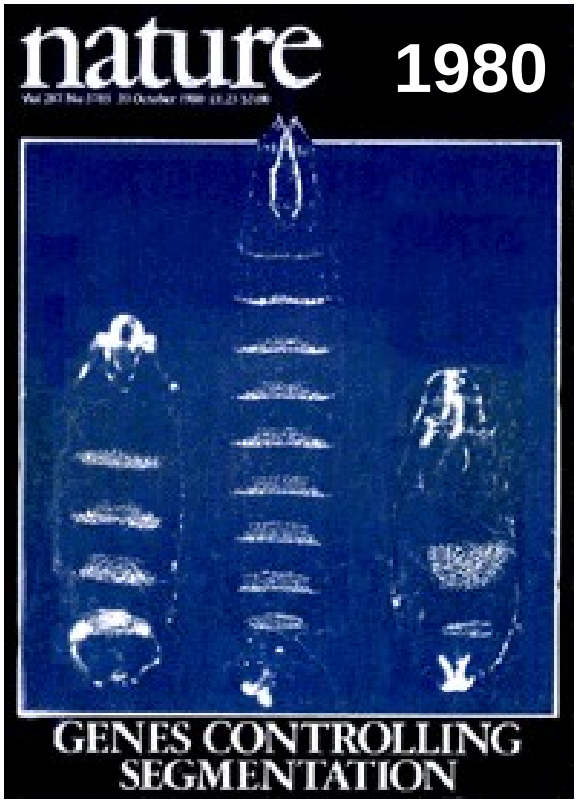
Historical screen
by C. Nüsslein-Wolhard et E. Wieschaus

General principles

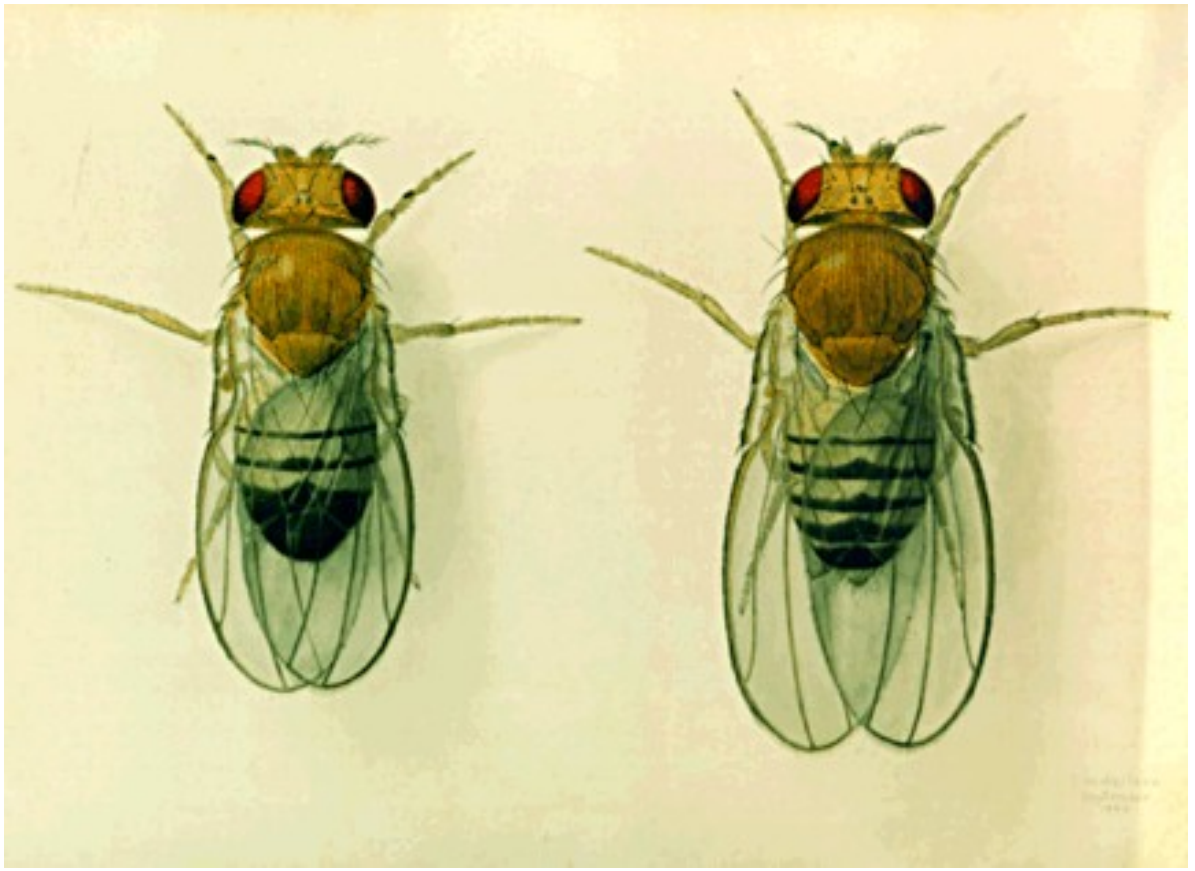
Other types of screen

Historical screen
by C. Nüsslein-Wolhard et E. Wieschaus





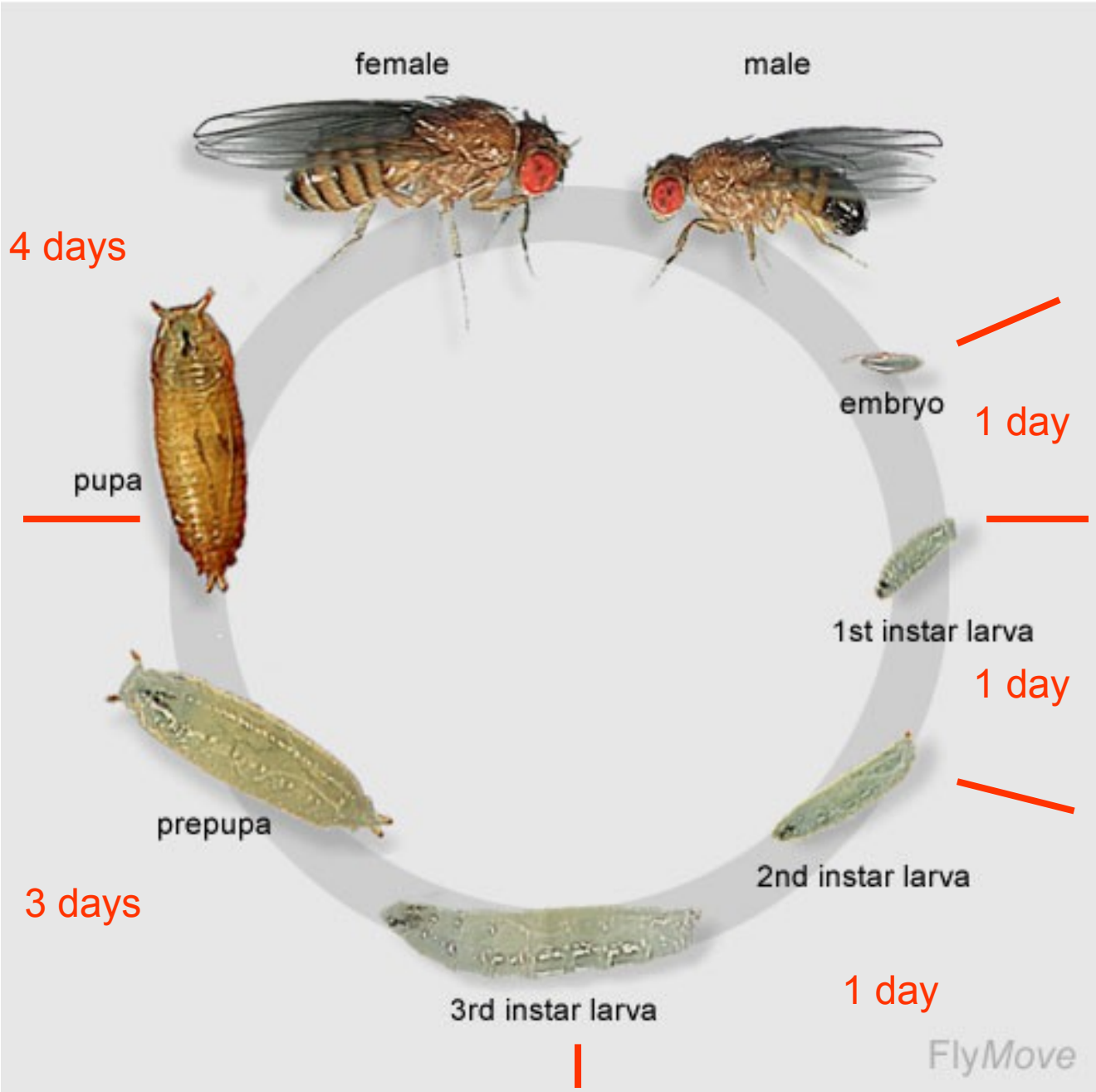
Nobel Price in Physiology/Medecine in 1995
with E. Lewis



♂ ♀


1521U
UAS GFP
813
FlyMove

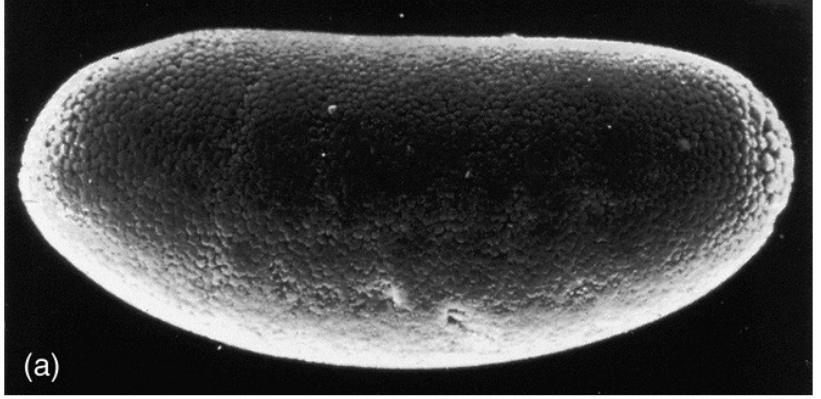
The life cycle of *Drosophila melanogaster*



One Generation
=
10 days

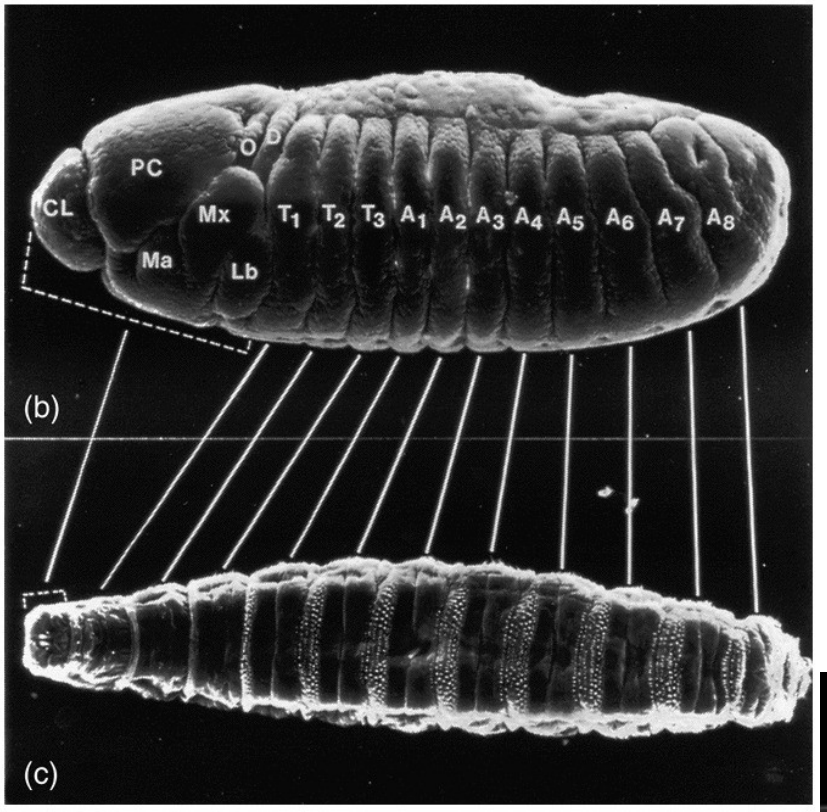
Cellular bastoderm stage

3 h



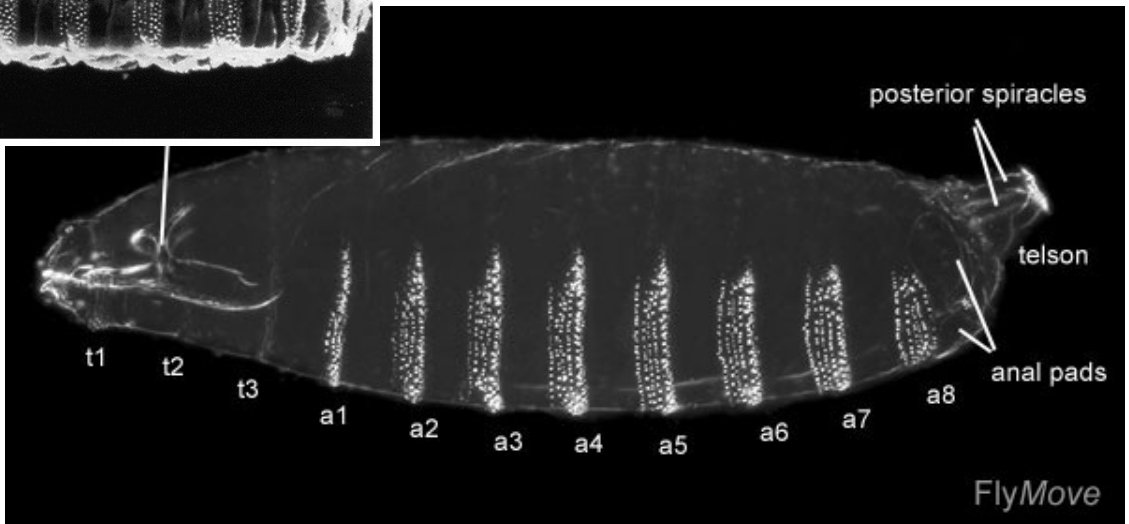
Segmented embryo

10 h



Larva

Wild-type larva



General Strategy

Random mutagenesis of the *Drosophila* genome and screen for mutant phenotypes

↓

Identification of the mutated gene

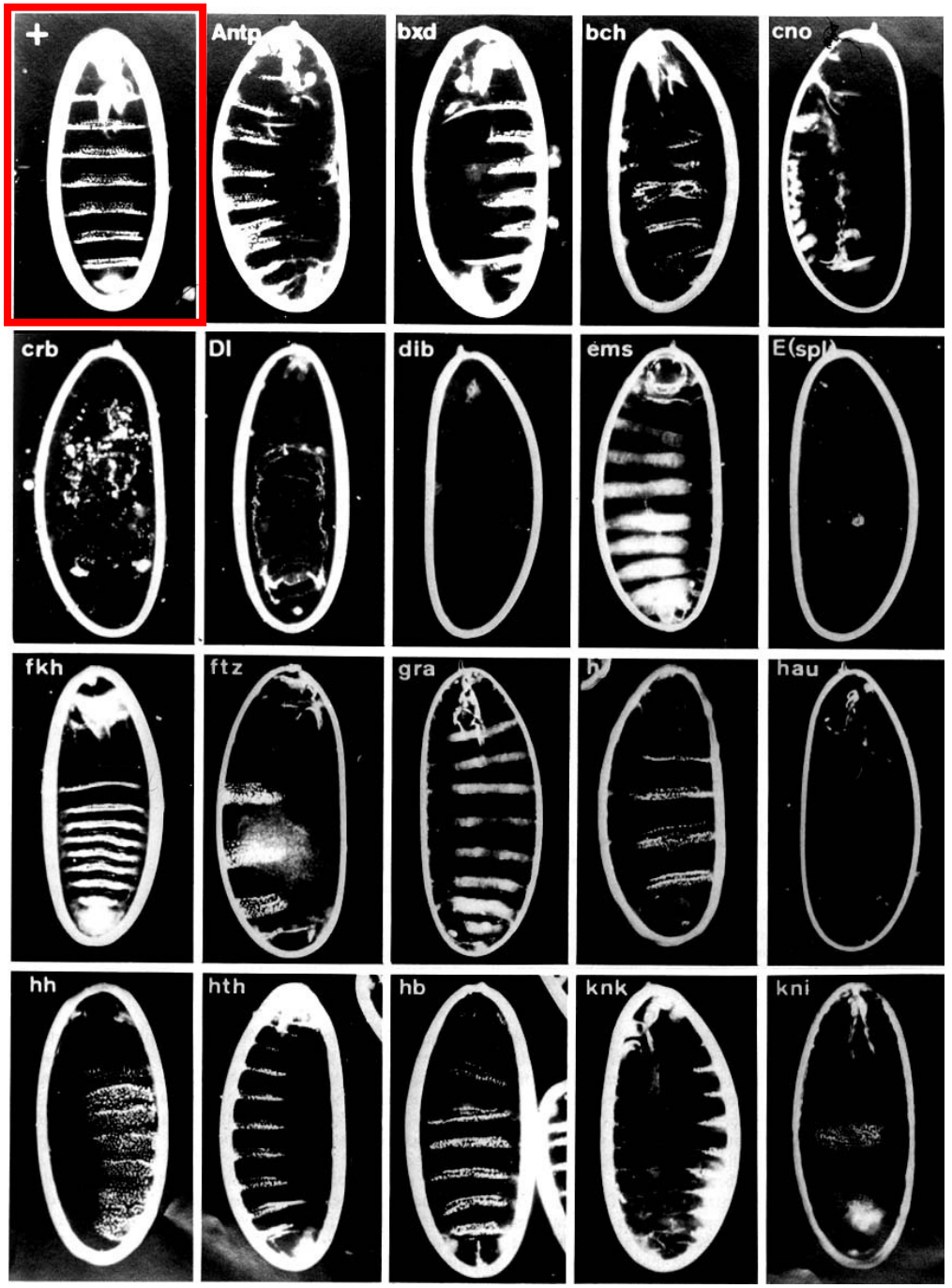
↓

Molecular analysis of the protein function

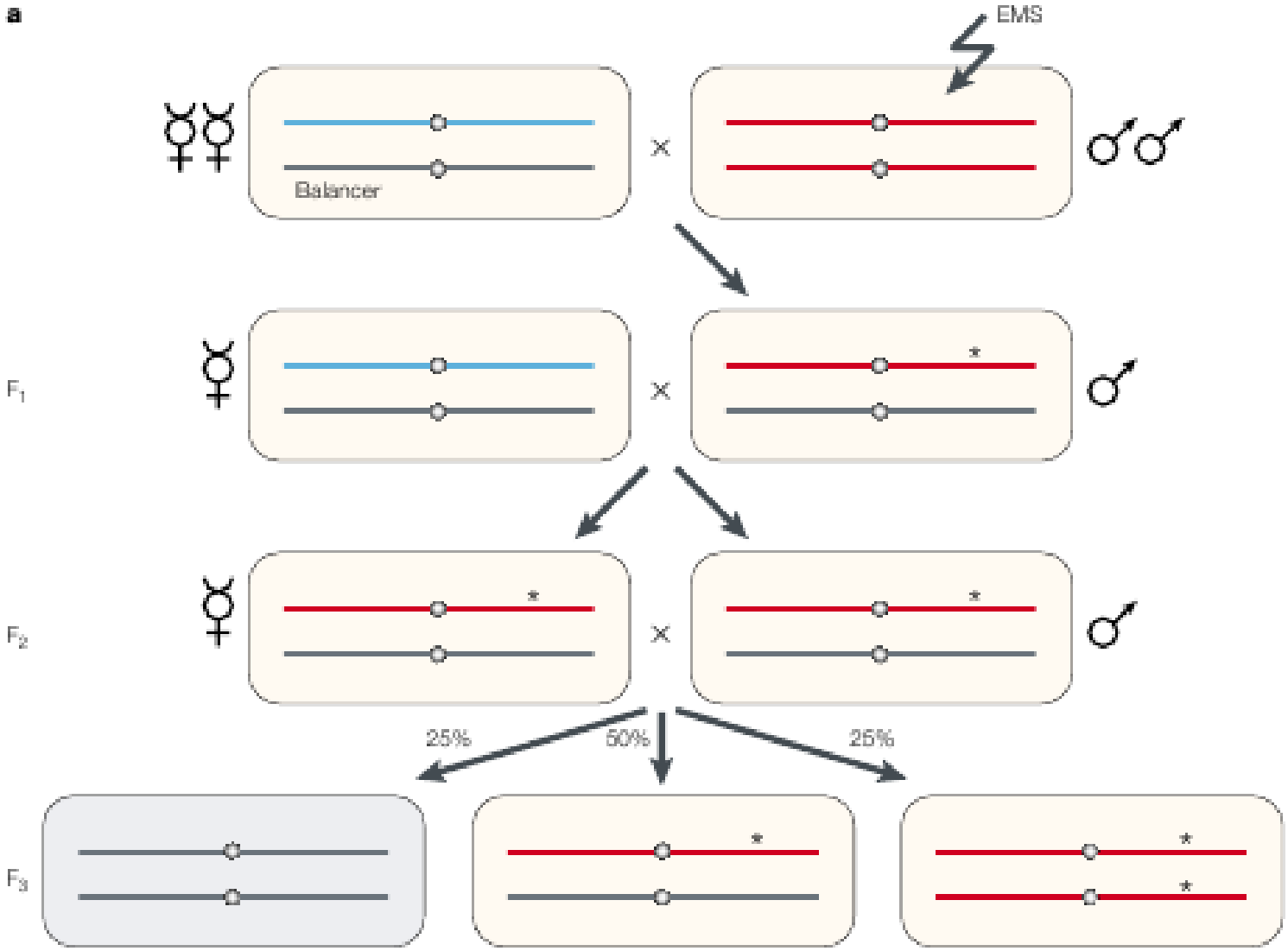
Wild-type

Others= mutants

(aberrant position /shape of trichomes)



Using balancers to screen recessive lethals



St Johnston, NRG 2002

Screen of chromosome 2

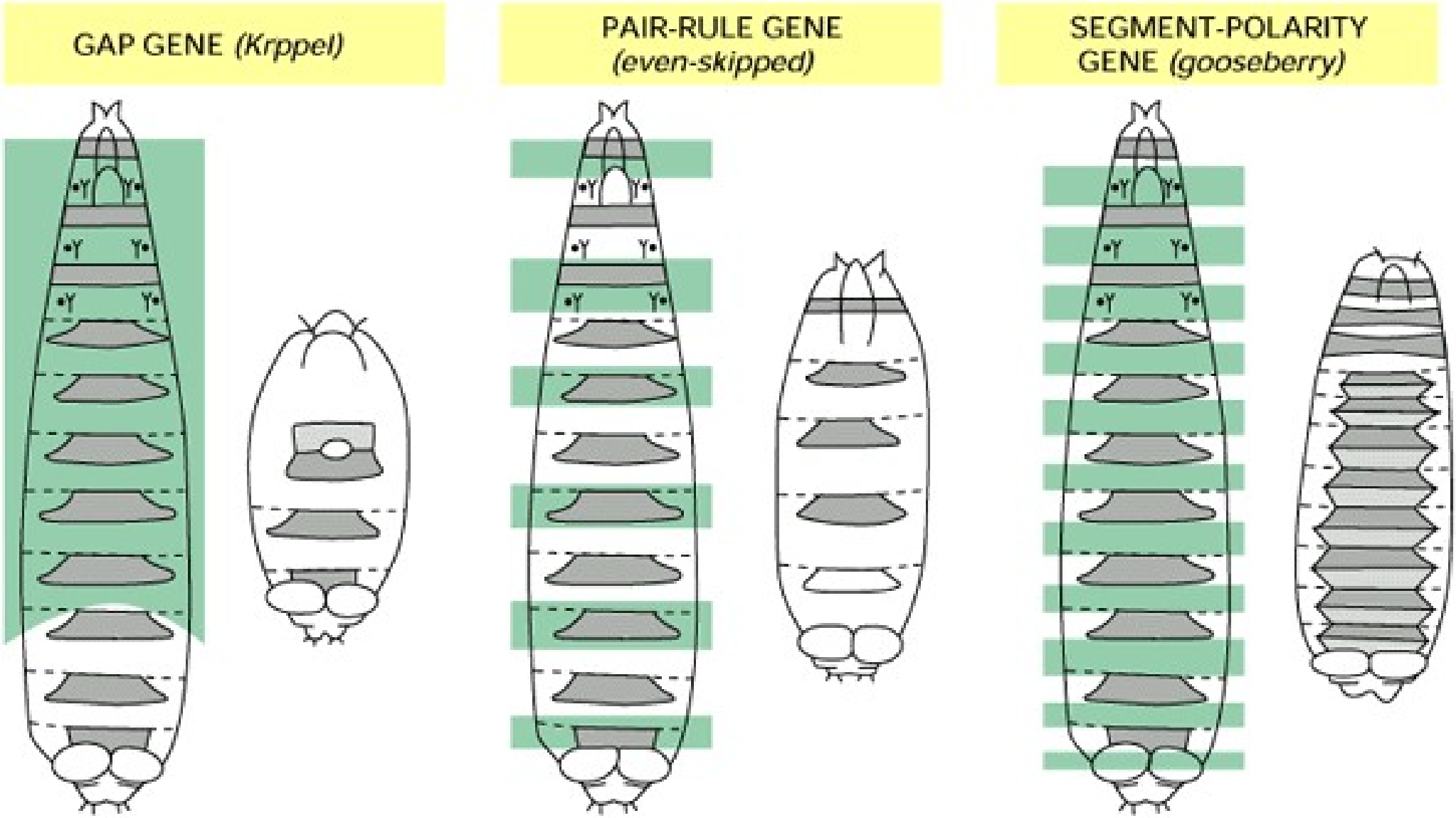
Production of 5.764 lines
    including 4.217 homozygote lethal lines

Identification of 7.600 lethal mutations
    including 2.843 mutations causing embryonic lethality
        and 272 mutations embryonic phenotypes

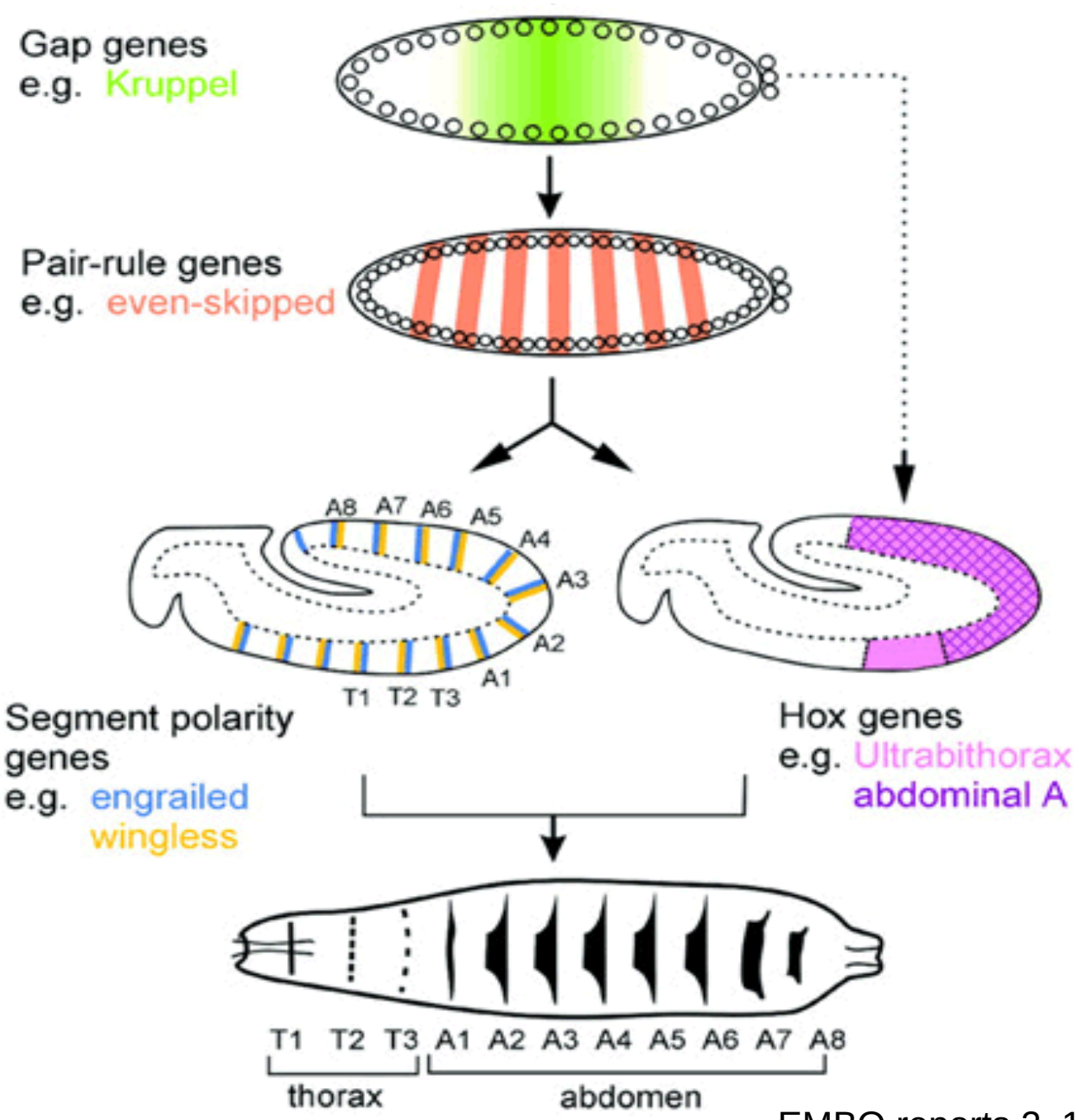
Complementation test for mutations with the same phenotype:
48 **complementation groups** containing on average 5.4 alleles
13 alleles are complemented by all other mutants

= 61 genes in total

Nusslein-Volhard, Wieschaus (1984) Roux's Arch Dev 193: 267

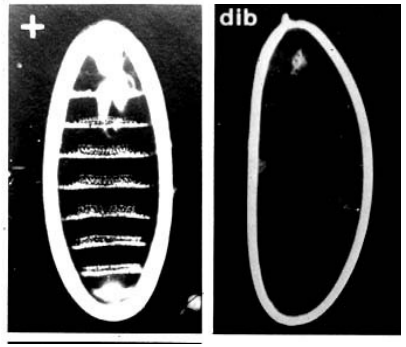

GAP GENE *(Krppel)*
PAIR-RULE GENE
(even-skipped)
SEGMENT-POLARITY
GENE *(gooseberry)*

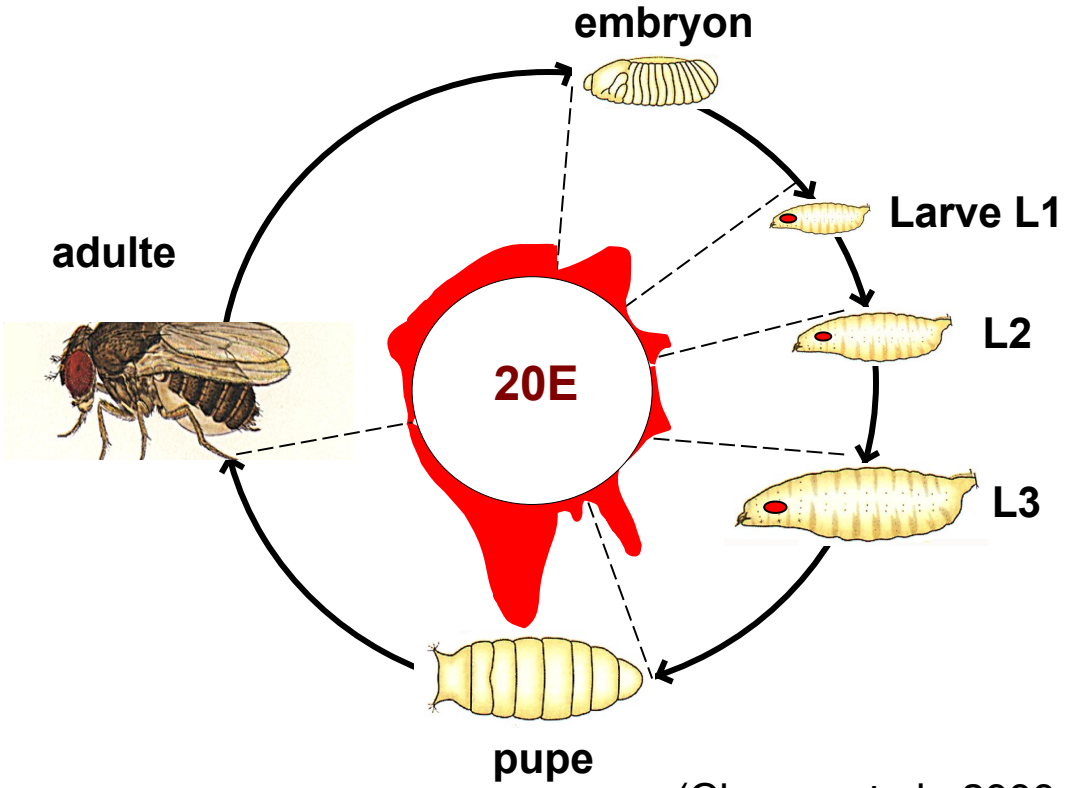
Embryonic development of Drosophila



EMBO reports 2, 12, 1083–1088 (2001)

Halloween mutants: steroid hormone biosynthesis





cholesterol

↓ *nvd*

7-dehydro-cholesterol

↓ *spo, sro*

ketodiol (2,22,25dE)

↓ *phm*

↓ *dib*

↓ *sad*

ecdysone (E)

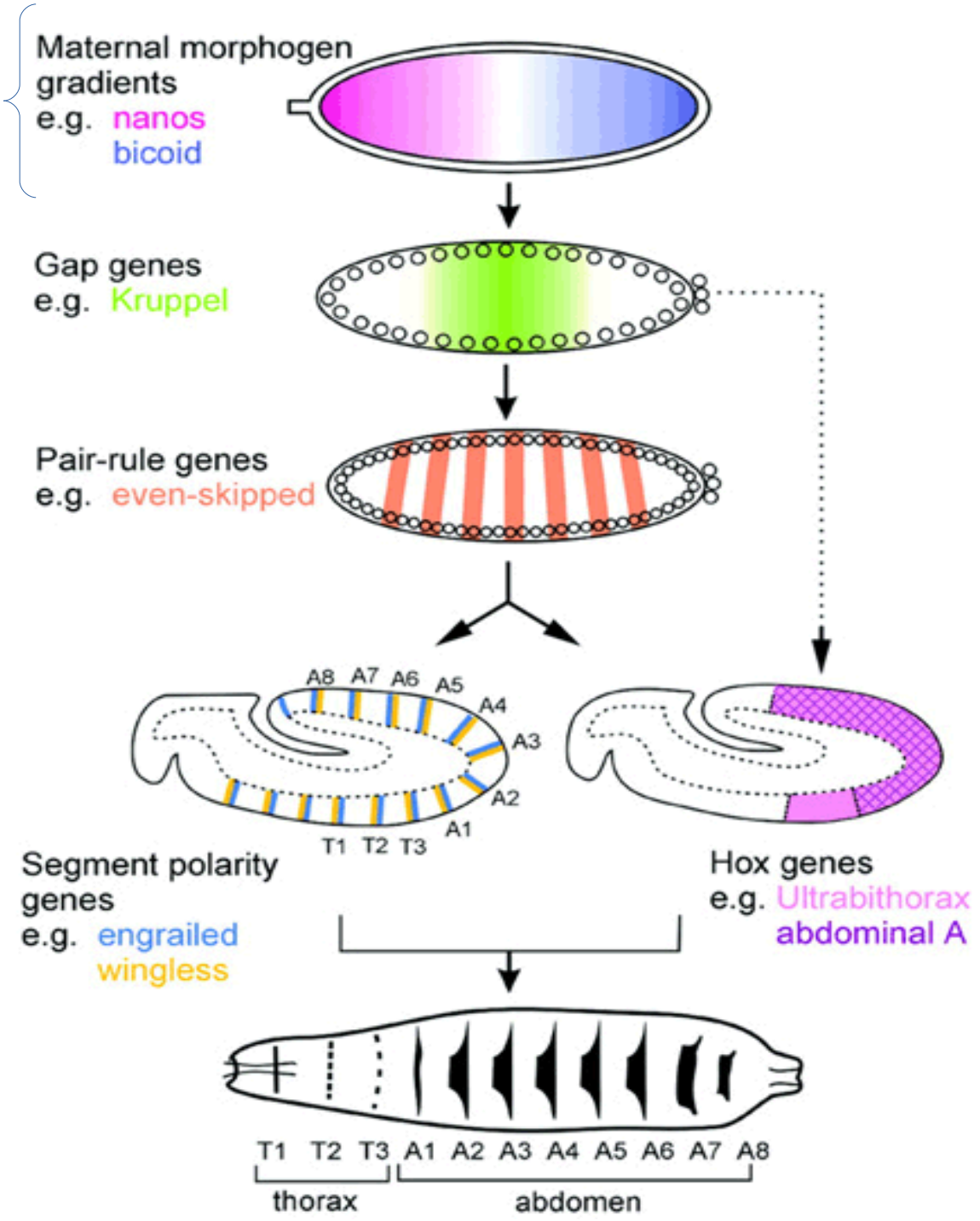
↓ *shd*

20-hydroxy-ecdysone (20E)

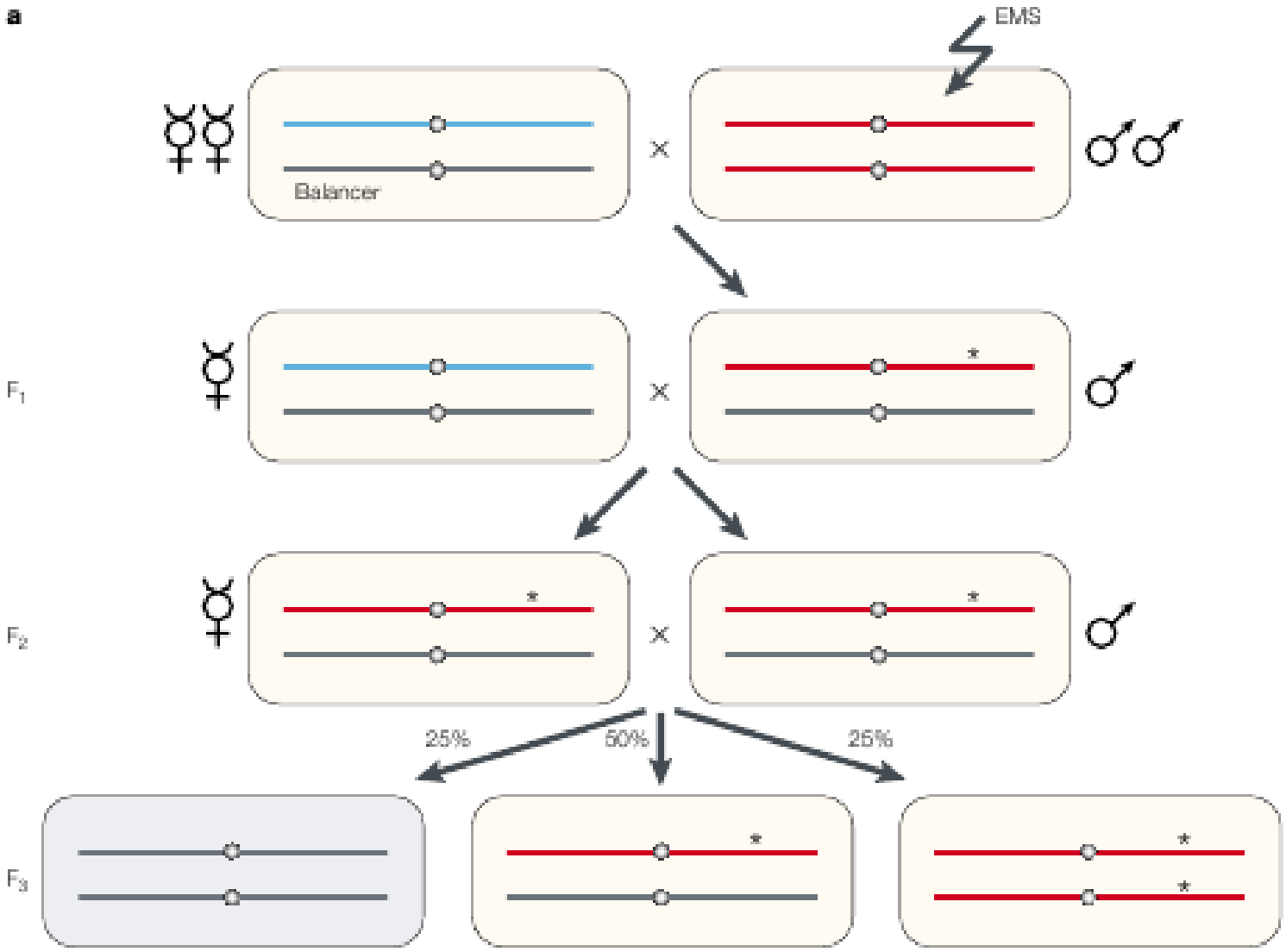
(Chavez et al., 2000; Warren et al., 2002; Petryk et al., 2003; Niwa et al., 2004; Warren et al., 2004; Namiki et al., 2005; Yoshiyama et al., 2006)

Maternal genes

ARNm deposited in the egg before fecundation



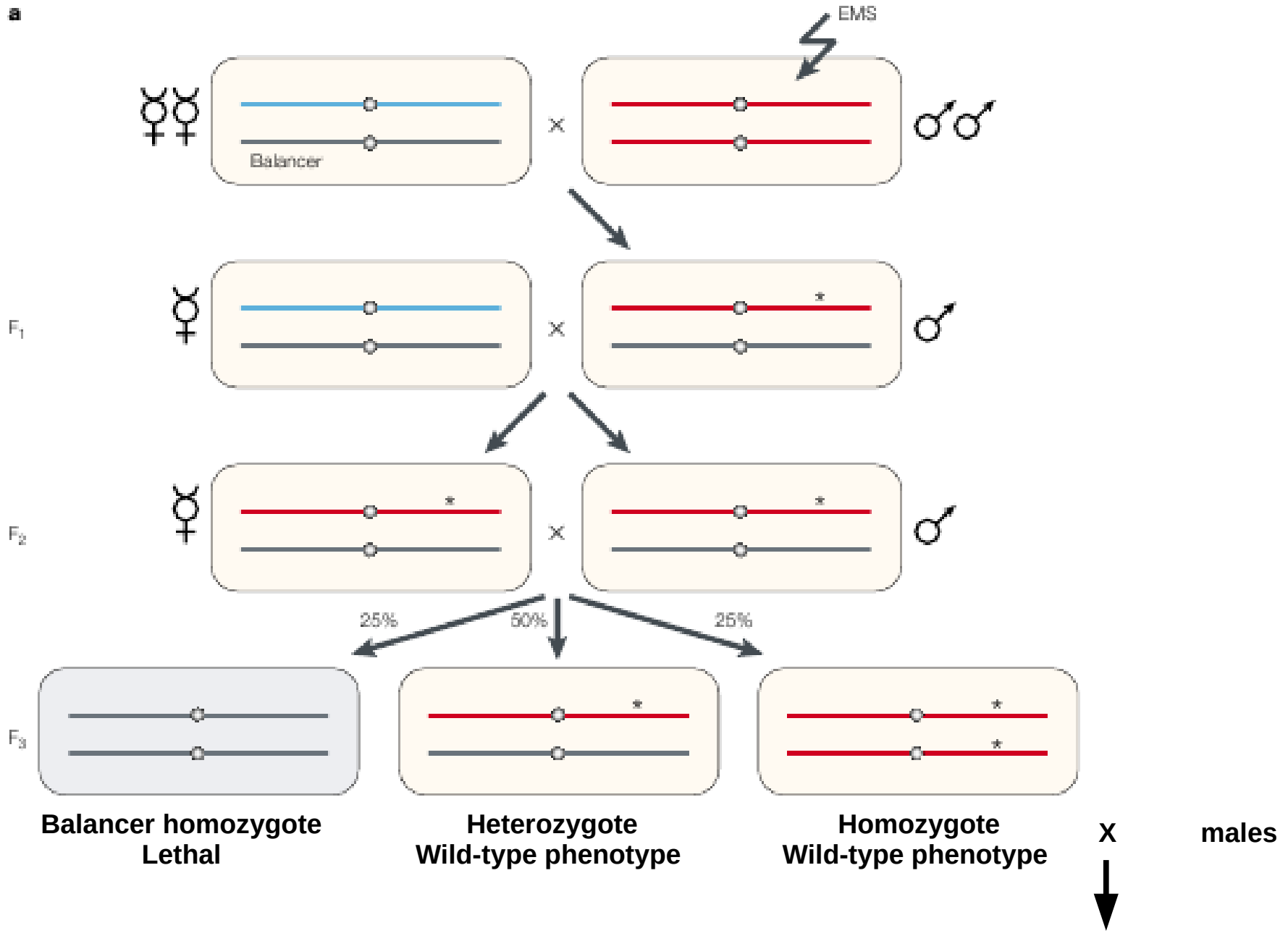
Screen for maternal effect genes



X males

Examination of progeny

Screen for maternal effect genes



Homozygous lethals cannot be identified
(ex : *Fz, Dsh, Apc, Nvd*)

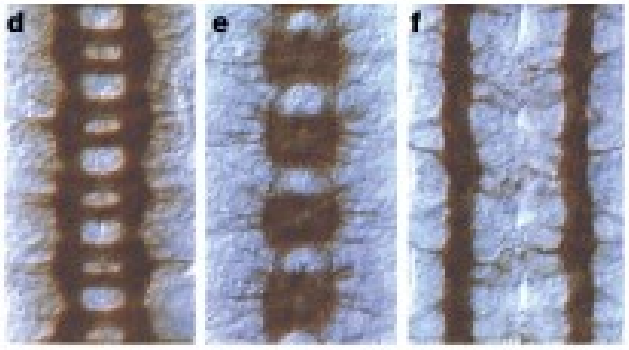
Certain genes involved in embryonic development were not identified with this screen

Maternal effect genes whose mutation is recessive lethal

Genes involved in the development of internal structures (brain, gut, etc.)

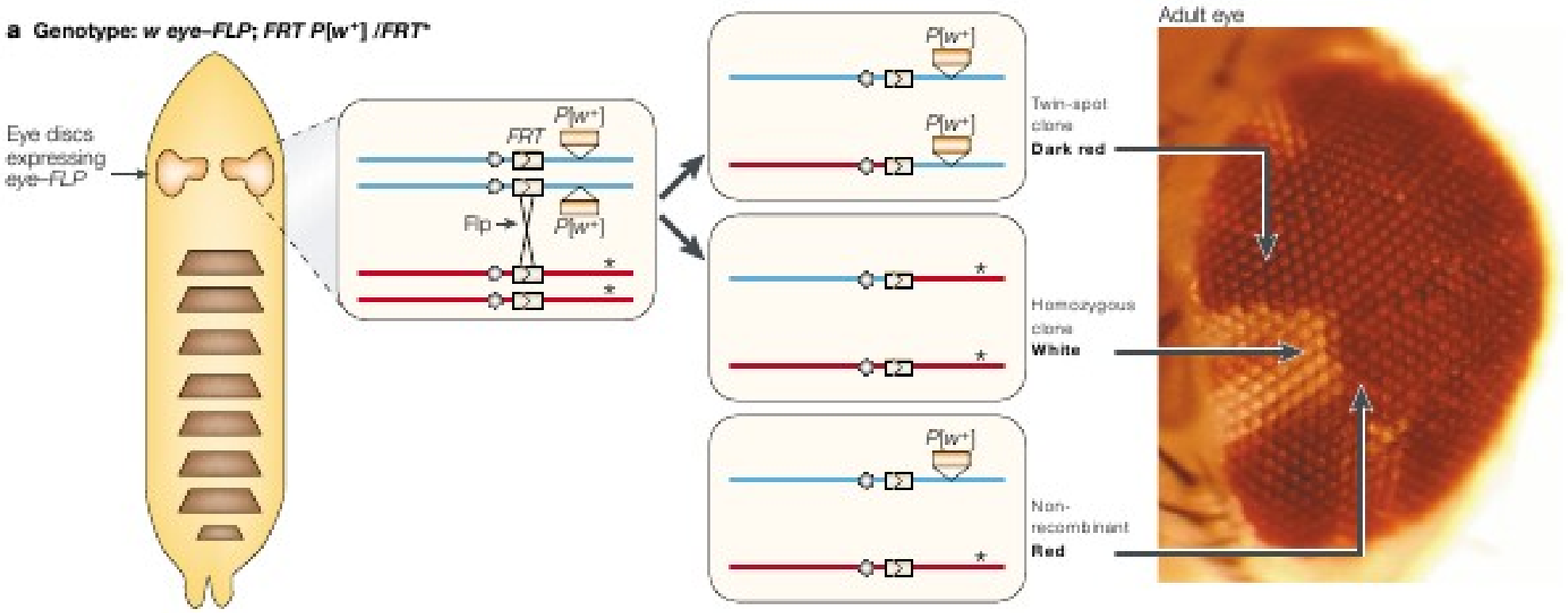
Redundant genes

WT *roundabout commissureless*



With such screens, only the first essential function of a gene can be identified.

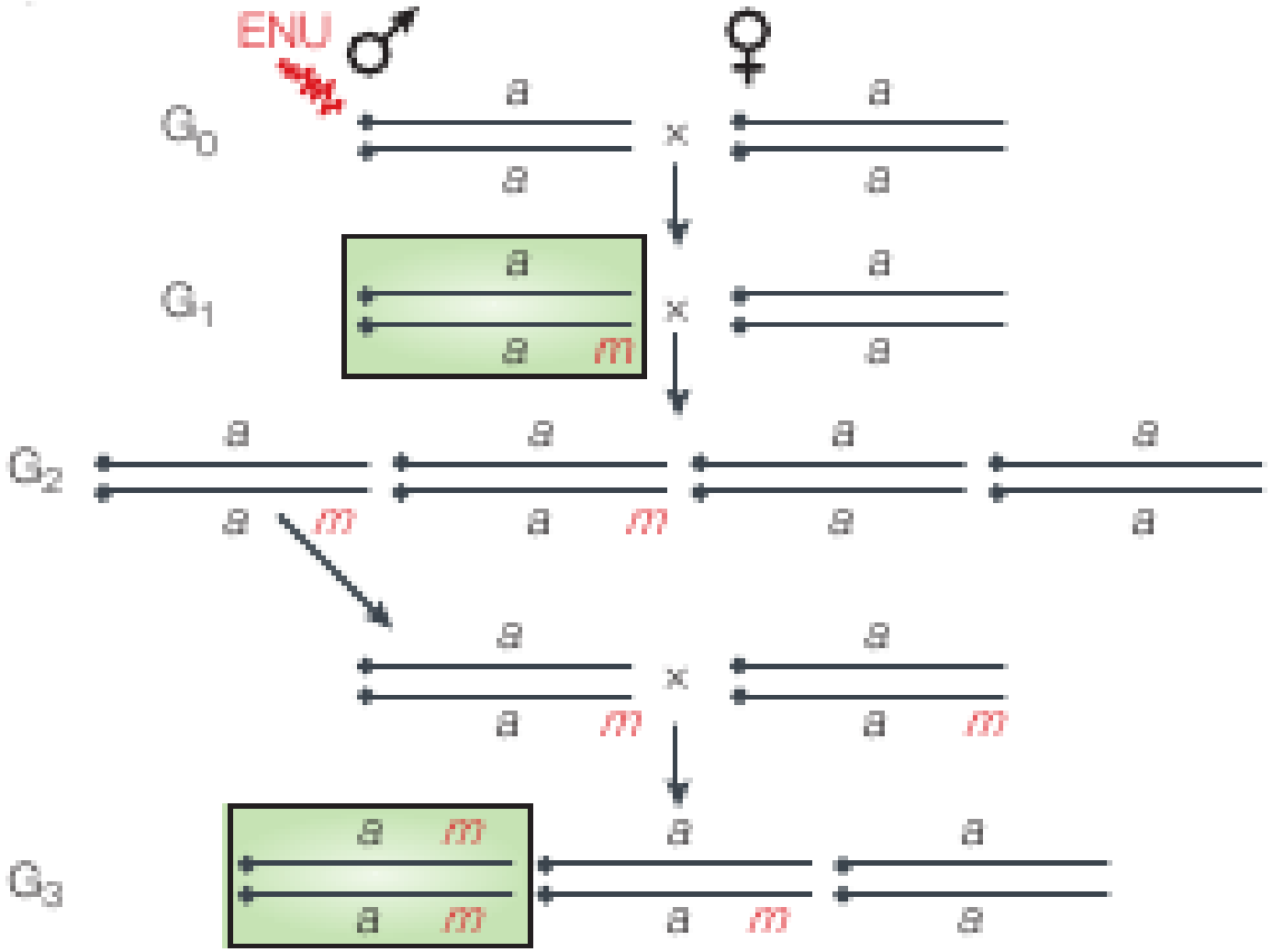
Mitotic clones



Flp = Flippase
FRT = Flippase Recombination Target

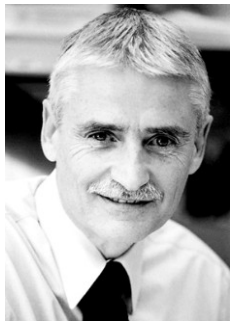
St Johnston NRG 2002

Screen for suppressors or enhancers



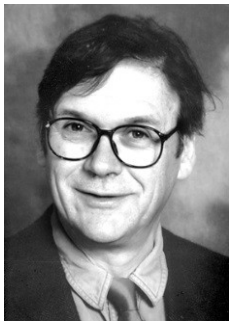
a/a individuals are viable and fertile.
Screen for enhancers or suppressors of the phenotype.

Three Nobel prizes associated with genetic screens



Leland Hartwell

Paul Nurse



Tim Hunt

Eukaryotic cell division
(Hartwell et al. 1974; Nurse et al. 1976)

Seymour Benzer

Circadian rhythms
(Konopka and Benzer 1971)

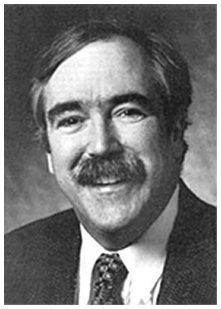
The Nobel Prize in Physiology or Medicine 2017

© Nobel Media. Ill. N. Elmehed
Jeffrey C. Hall

© Nobel Media. Ill. N. Elmehed
Michael Rosbash

© Nobel Media. Ill. N. Elmehed
Michael W. Young

C. Nusslein-Volhard



Eric Wieschaus

Development
(Lewis 1978; Nusslein-Volhard and Wieschaus 1980)

General principles

Mutagens

Crosses

Types of phenotype

Identification and validation of causal mutations

Mutagens (1)

X rays : breaks in double-stranded DNA, resulting in large deletions of pieces of chromosome or chromosomal re-arrangements. → good to map by cytological examination of chromosomes, but often not limited to single genes

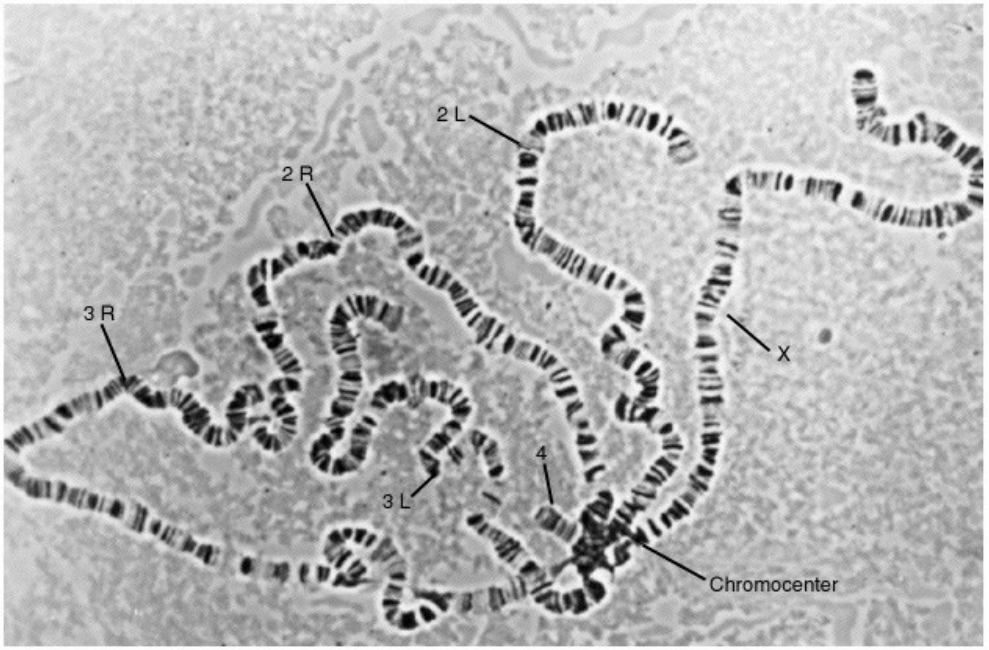
Chemical:
Ethylmethane sulfonate (EMS) : very efficient, alkylation agent (GC to AT), point mutations

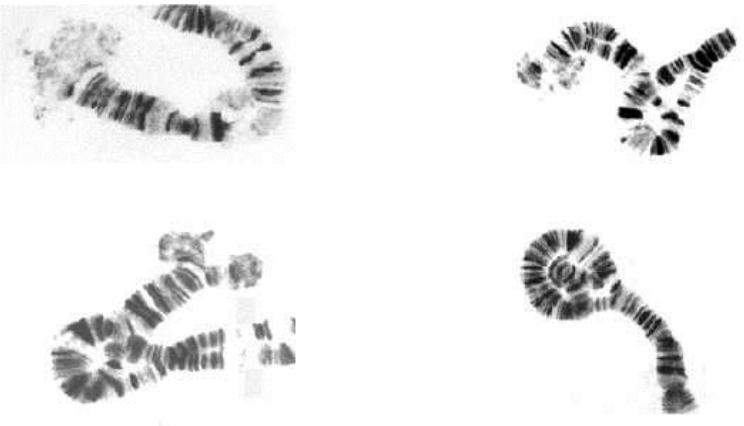
In Drosophila, EMS can produce ~10^{-3} mutations per gene

→ how many mutated genes on average on one chromosome containing 5000 genes?

→ if 6000 such EMS-treated chromosomes are generated, how many alleles per gene can be expected from the screen?

X rays





Photomicrographs of polytene chromosomes; name the mutation, if any

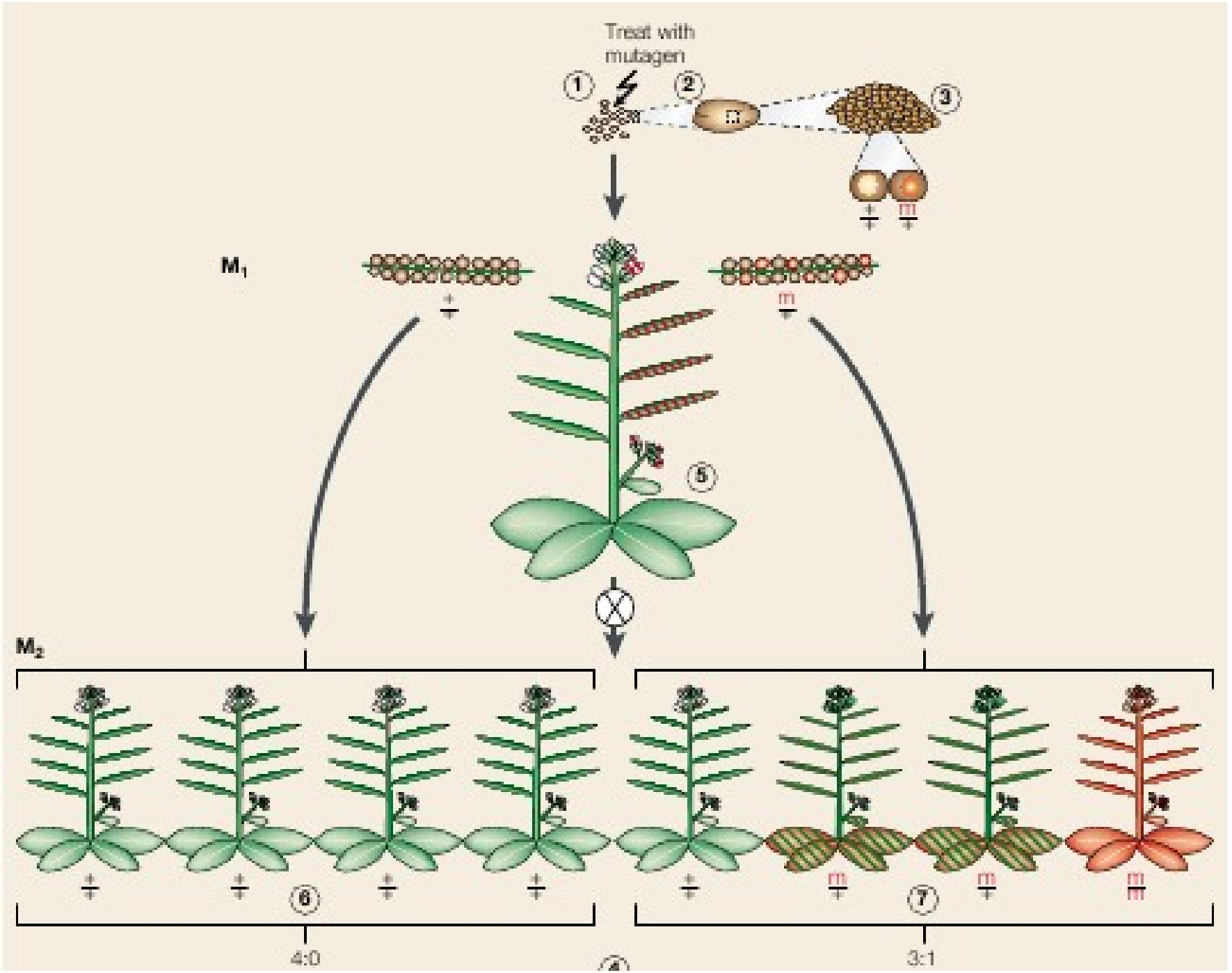
Mutagens (2)

<u>Chemical</u>:
Methylmethane sulfonate (MMS) : agent alkylant, moins efficace que l'EMS pour la drosophile, induit un peu plus de délétions que l'EMS.

N-nitroso-N-ethylurea (ENU) : ethyle oxygen atoms (O2 and O4 of T, AT to GC, O6 of G, GC to AT), fewer aberrations than EMS

Triethylmelanine (TEM) : deletions
Formaldehyde : deletions

Chemical mutagens create mosaic individuals



Mutagens (2)

<u>Chemical</u>:
Methylmethane sulfonate (MMS) : agent alkylant, moins efficace que l'EMS pour la drosophile, induit un peu plus de délétions que l'EMS.

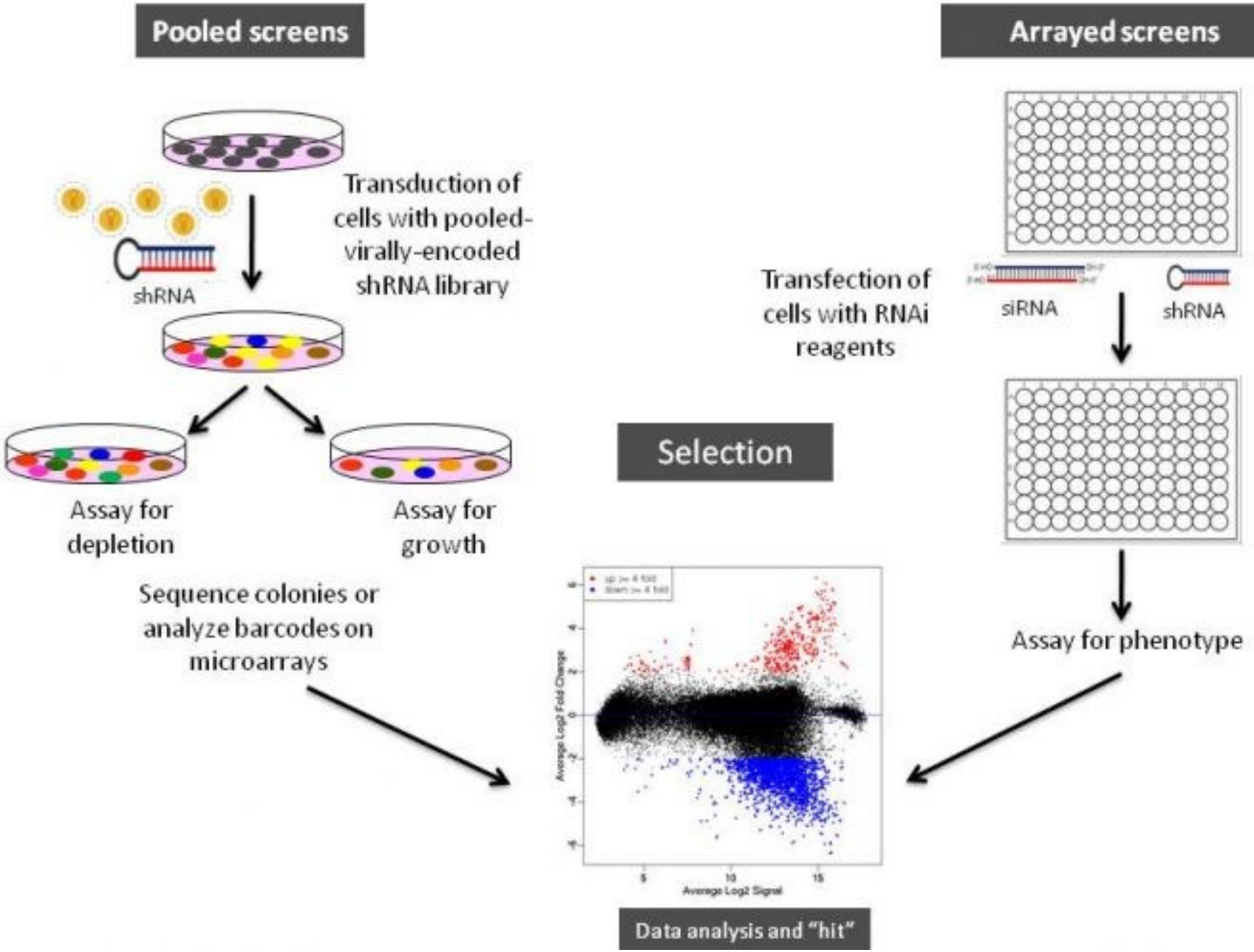
N-nitroso-N-ethylurea (ENU) : ethyle oxygen atoms (O2 and O4 of T, AT to GC, O6 of G, GC to AT), fewer aberrations than EMS

Triethylmelanine (TEM) : deletions
Formaldehyde : deletions

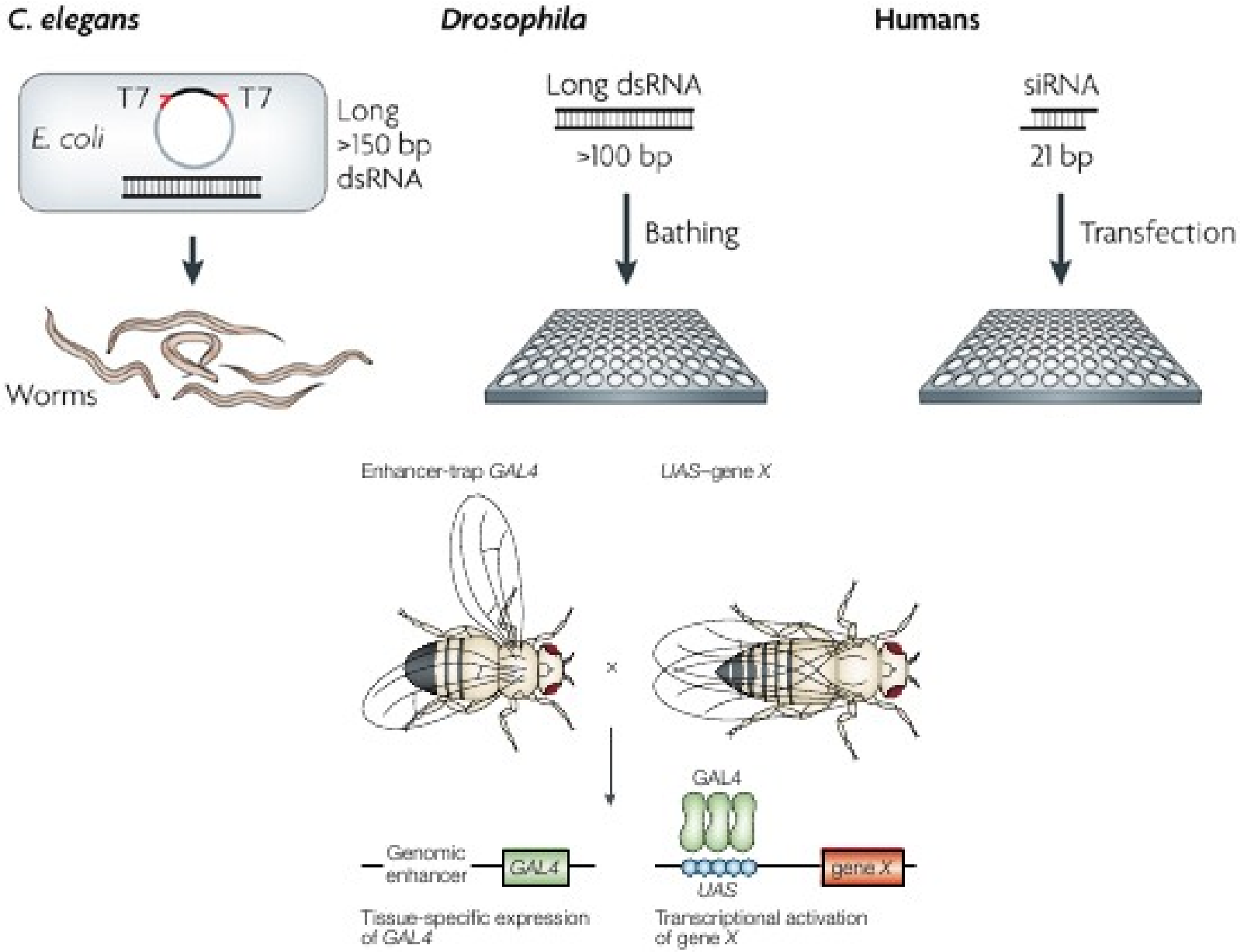
Insertational:
Transposable elements without transposase: integrate into the genome, facilitates identification of the mutation

RNAi

CRISPR-Cas-9


Pooled screens
Arrayed screens
Transduction of cells with pooled-virally-encoded shRNA library
shRNA
Transfection of cells with RNAi reagents
siRNA
shRNA
Assay for depletion
Assay for growth
Selection
Sequence colonies or analyze barcodes on microarrays
Assay for phenotype
Average Log2 Fold Change
Average Log2 Signal
Data analysis and "hit"

RNAi screens



Boutros et Ahringer (2008) Nat Rev Genetics 9, 554

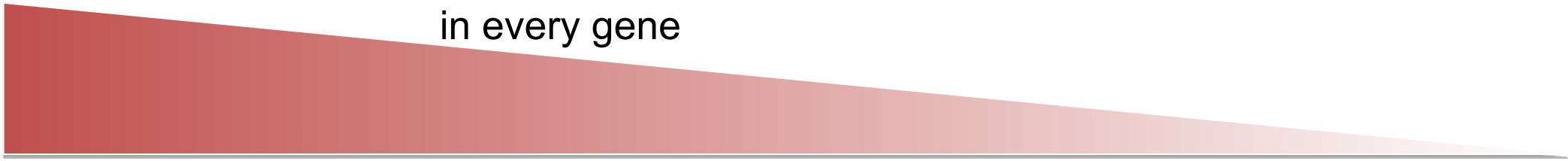
Forward genetics

Genetic screen for a
phenotype of interest,
identification of the mutated gene etc...

Reverse genetics

Start from a gene of interest
knockout, transgenesis, etc...

The distinction is fuzzy when one starts from a subset of a library of mutants



Crosses

Since chemical mutagens create mosaic individuals, the progeny must be screened

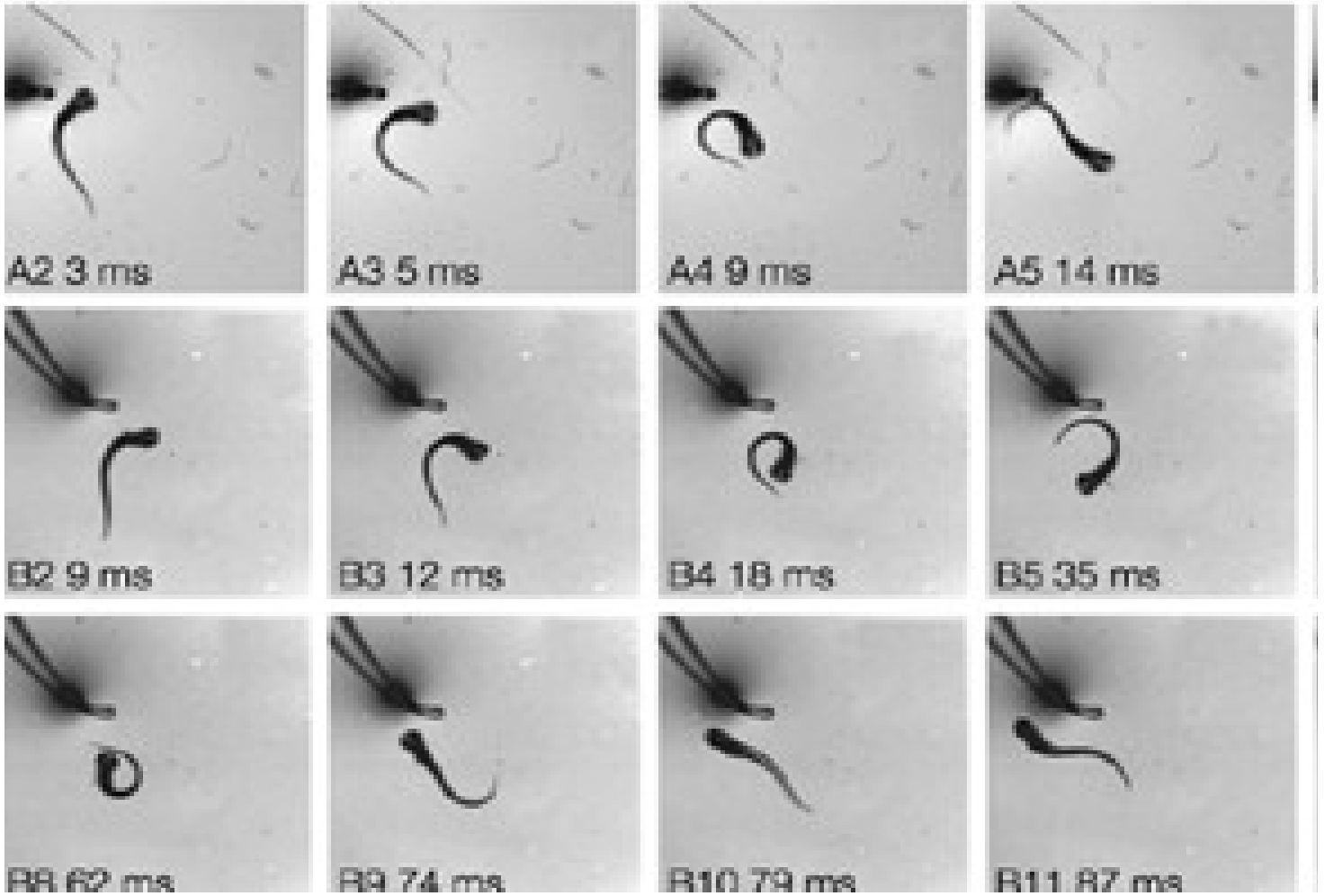
F1 screen: screen for suppressors and enhancers

F2 screen: screen for recessive mutations

F3 screen: screen for maternal effect genes

Phenotypes

Morphology, Physiology, Behavior



space cadet Zebrafish mutant

Phenotypes

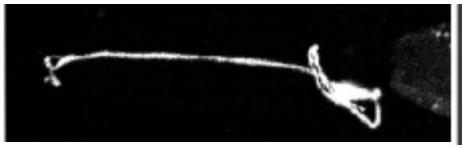
Direct observation

WT

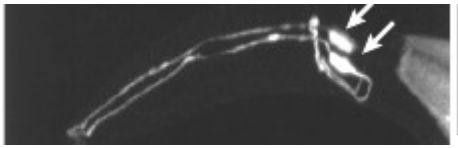
agamous-1

Staining (GFP, antibodies)

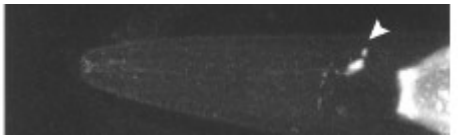
WT



mutant 1

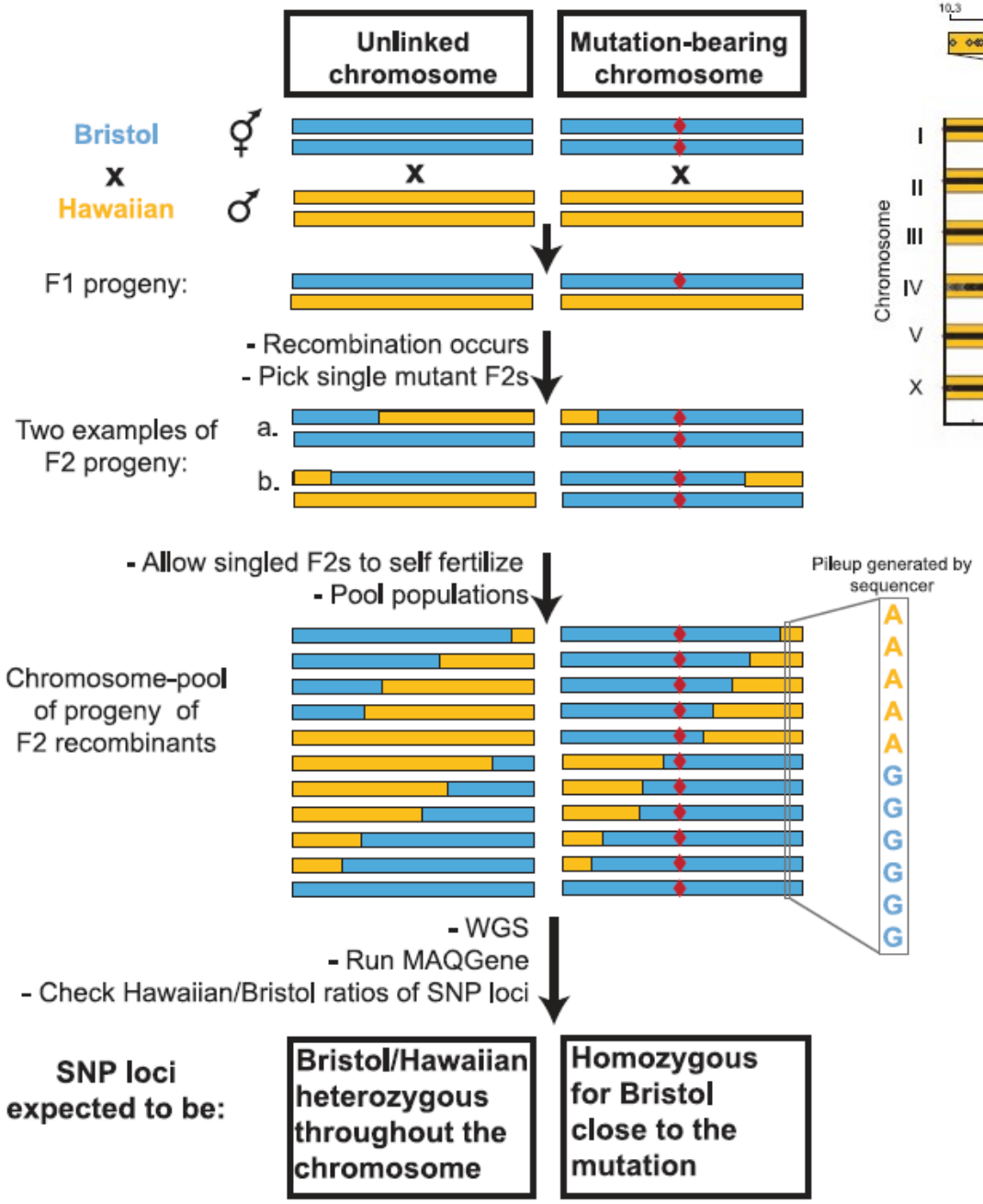


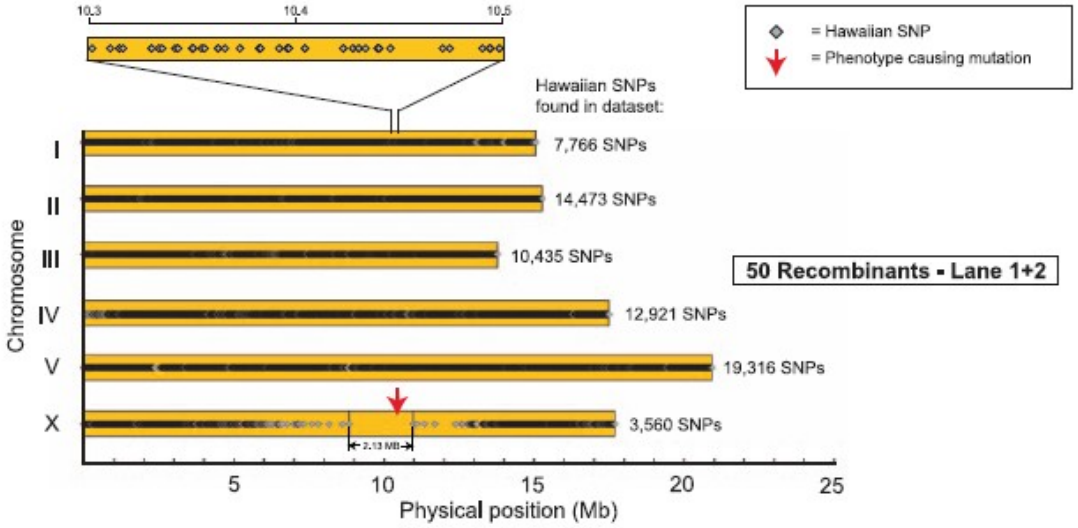
mutant 2



str2::GFP

Identification of the mutation





Requires nucleotide divergence and many recombinants

Oliver Hobert lab

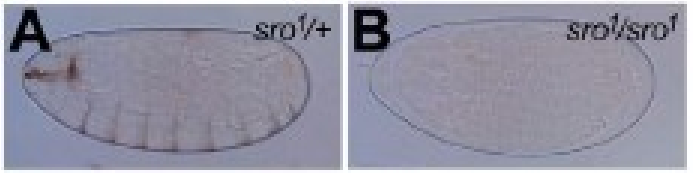
Identification of the mutation

Once a small region is identified

Complementation test with deletions/mutants already available

Analysis of candidate gene expression

Rescue of the mutant phenotype with transgenes



Niwa et al. (2010) Development 137, 1991

Table 1. *sro*^{1} lethality was rescued by *sro/nm-g* overexpression

Genotype	Number of adults
+/+; 2-286-GAL4 sro^{1}/+ sro^{1}	0 (308)
UAS-sro/+; sro^{1}/sro^{1}	0 (170)
UAS-nm-g/+; sro^{1}/sro^{1}	0 (138)
UAS-sro/+; 2-286-GAL4 sro^{1}/+ sro^{1}	128 (286)
UAS-nm-g/+; 2-286-GAL4 sro^{1}/+ sro^{1}	57 (270)

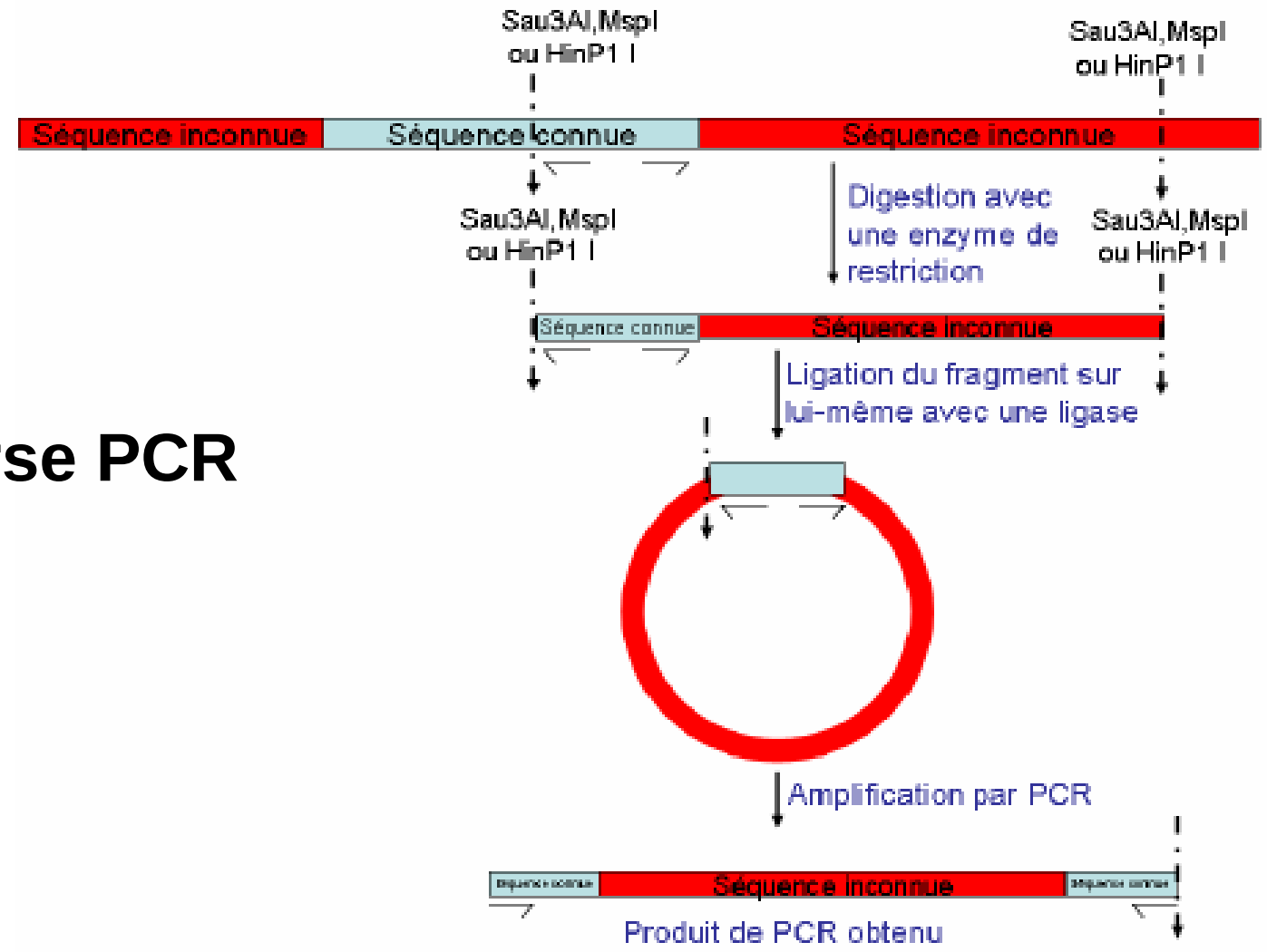
The numbers of viable adults were scored. Parentheses indicate the number of viable progeny with the presence of balancer markers from the parental strains.

Sequencing and search for mutations (nonsense, deletions, etc.)

Identification of the mutation

For transposable elements

Inverse PCR



Other types of screens

Gene expression screens

RNAseq

In situ hybridization of all genes

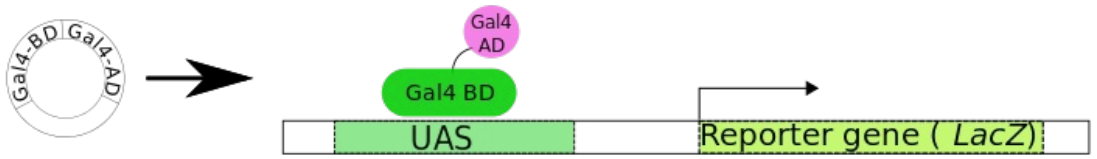
Screen of DNA sequences

Library with all the genes coding for transcription factors

Two-hybrid screen

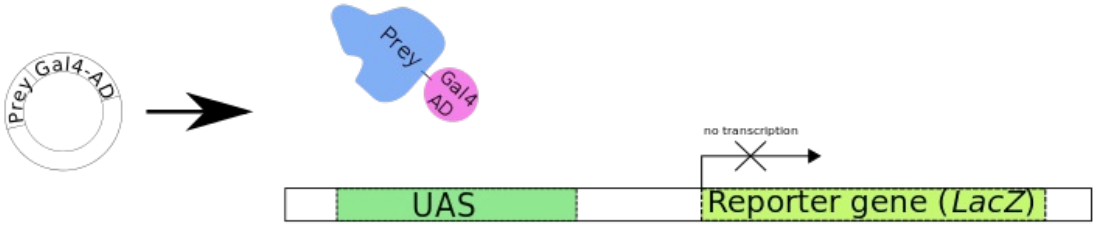
etc.

Yeast two-hybrid screen

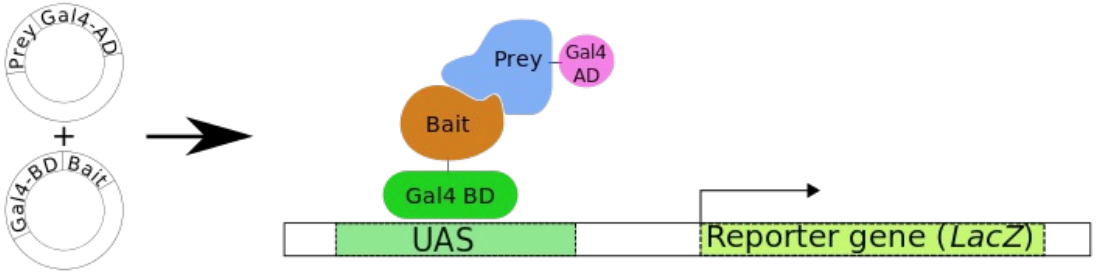


A. Regular transcription of the reporter gene

B. One fusion protein only (Gal4-BD + Bait) - no transcription



C. One fusion protein only (Gal4-AD + Prey) - no transcription



D. Two fusion proteins with interacting Bait and Prey

From laboratory to “real-life” data

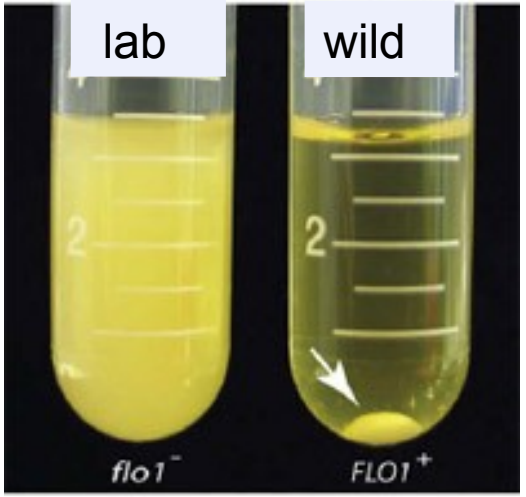
Knock out

Natural variation

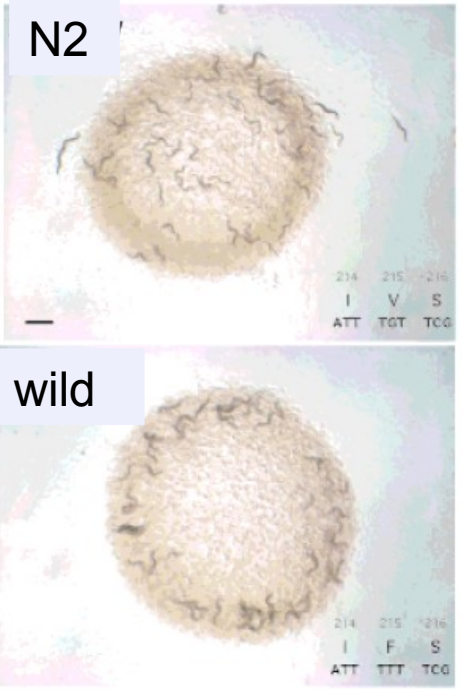
Domestication of laboratory strains

Arabidopsis thaliana

Saccharomyces cerevisiae



Caenorhabditis elegans



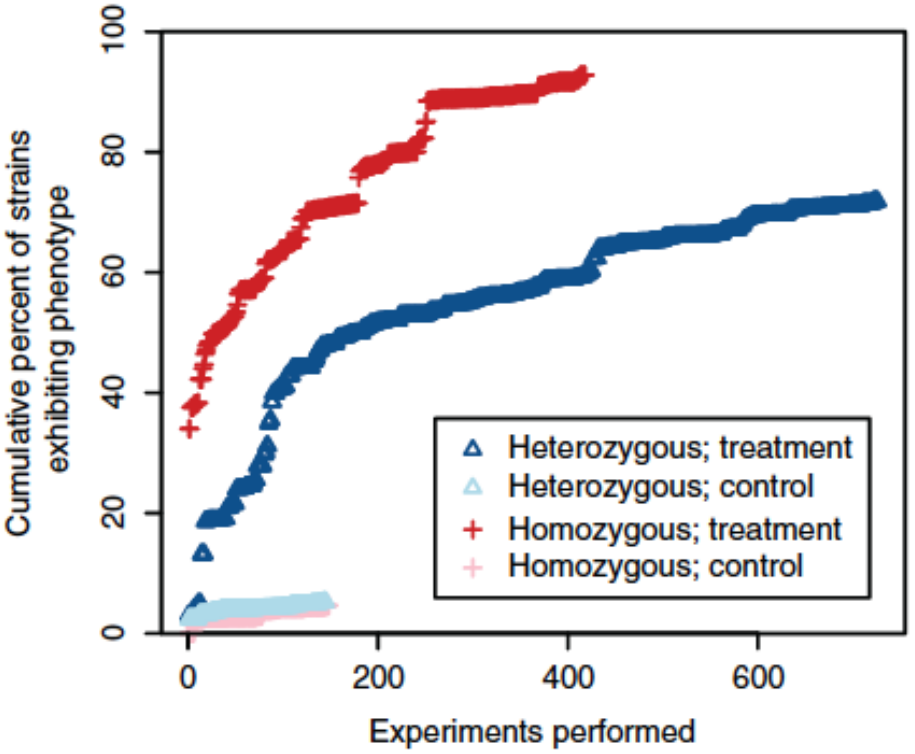
Domestication of laboratory strains
results in extreme phenotypic values
for many traits:
artificial selection and pleiotropy

Choice of laboratory environment

ca. 10-20 years ago: surprise at not finding phenotypes in gene knockouts

The Chemical Genomic Portrait of Yeast: Uncovering a Phenotype for All Genes
Maureen E. Hillenmeyer, *et al.*
Science **320**, 362 (2008);

1144 growth environments for *S. cerevisiae*



Genetic Screens
Laboratory mutations

- Not in nature
- Extreme effects
- Would likely be lost under selection
- Must be induced

- Interrogates (nearly) all regions
- Readily cloned
- Strong effects

Linkage/Asssociation mapping
Natural mutations

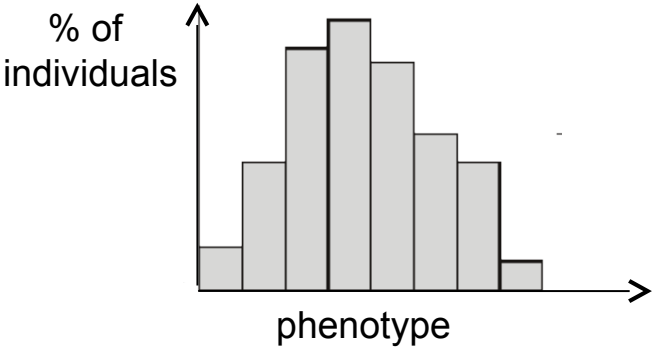
- Representative of nature
- Variants with small effects
- Sustained under selection
- Readily available

- Interrogates only variable regions
- Difficult to map
- Small effects

Quantitative genetics

Quantitative genetics

- If to each genotype corresponds a distribution of phenotypes
  = variable expressivity
  the character itself is quantitative

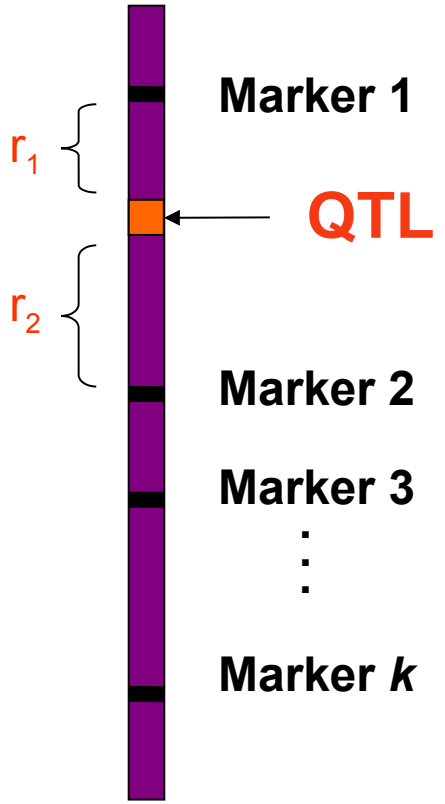


and/or

- If the variation of many genes is involved in the phenotypic difference between two strains/individuals
  the segregation of the character is quantitative

Quantitative Trait Loci (QTL) mapping

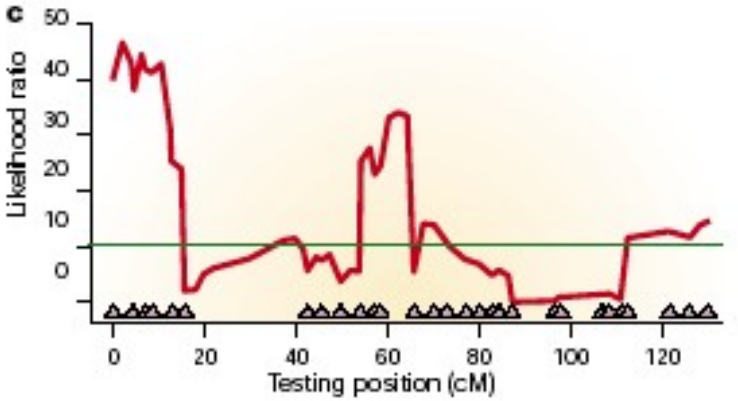
- QTL are specific genetic loci that affect quantitative traits.
- QTL can be detected by markers that are linked with it.



Two goals:

Identify the location of the QTL

Estimate the genetic effects of the QTL

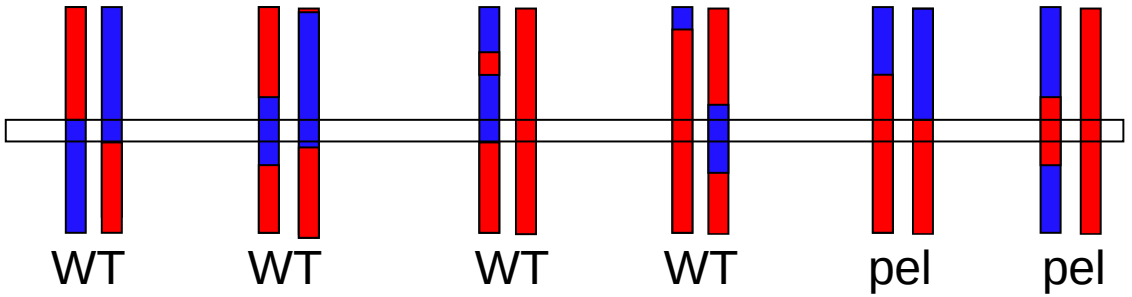


Epigenetics

Wild-type

Peloric

Linaria vulgaris

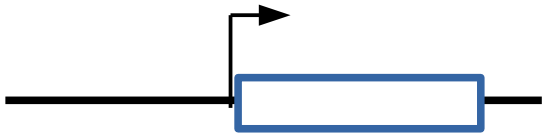


Cubas 1999 Nature

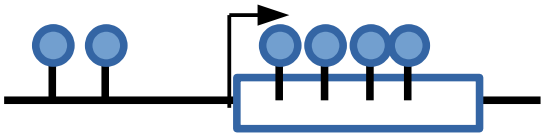
An epimutation

Wild-type

Peloric



CYCLOIDEA



CYCLOIDEA

Methylated DNA

Presence of CYCLOIDEA proteins

Absence of CYCLOIDEA proteins

Noise

Various concepts of chance/randomness in biology

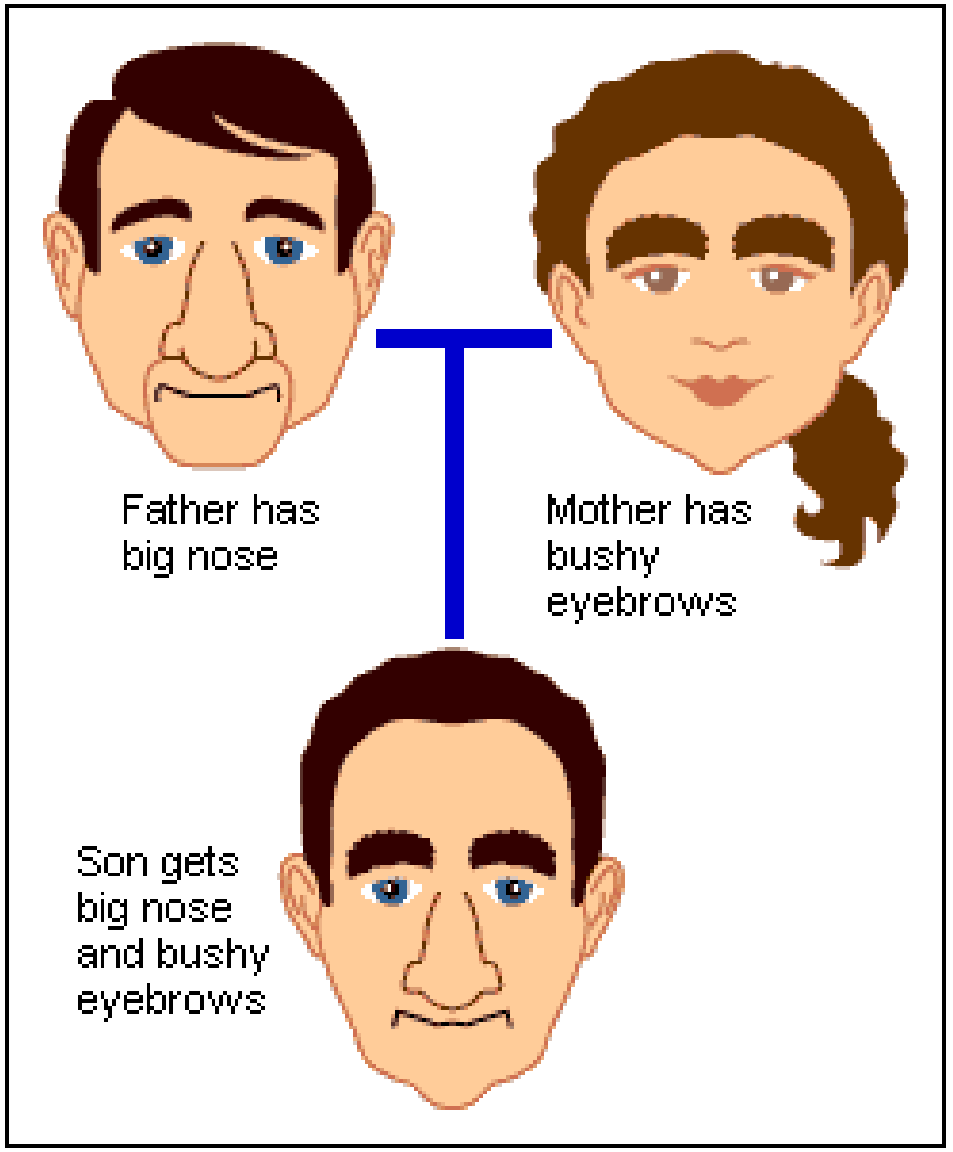
Are not explained within the framework of our current theory (no theory, initial conditions not known with sufficient precision, or because calculations are too complex)

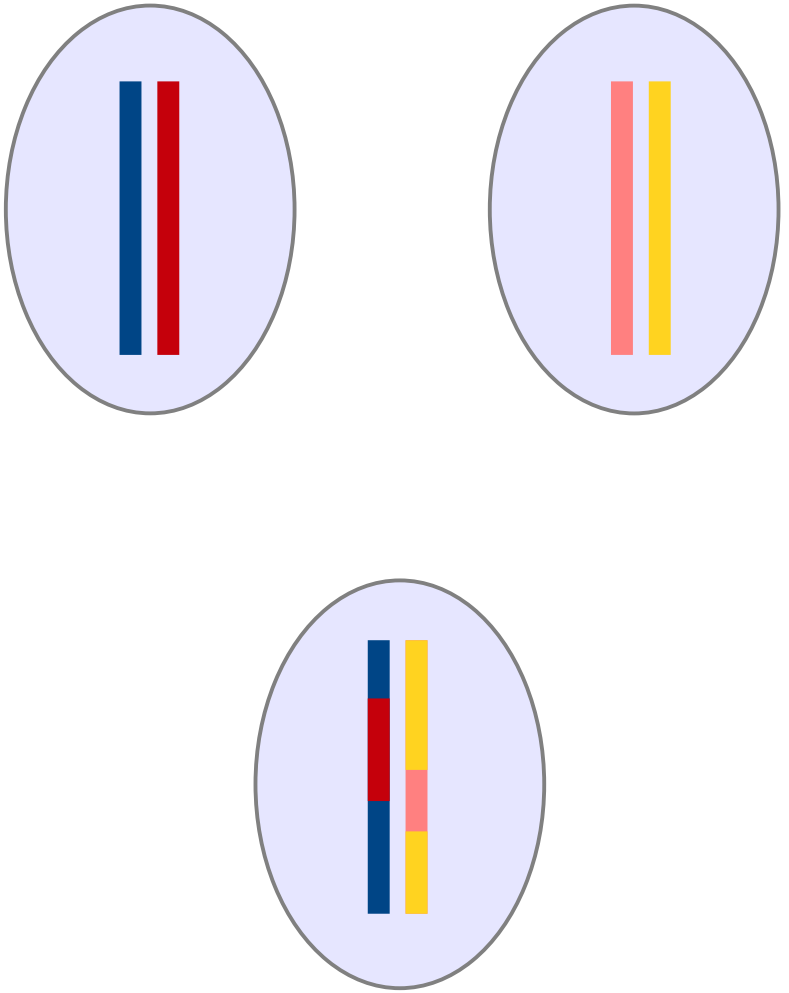
Cannot be predicted to occur: probabilistic events

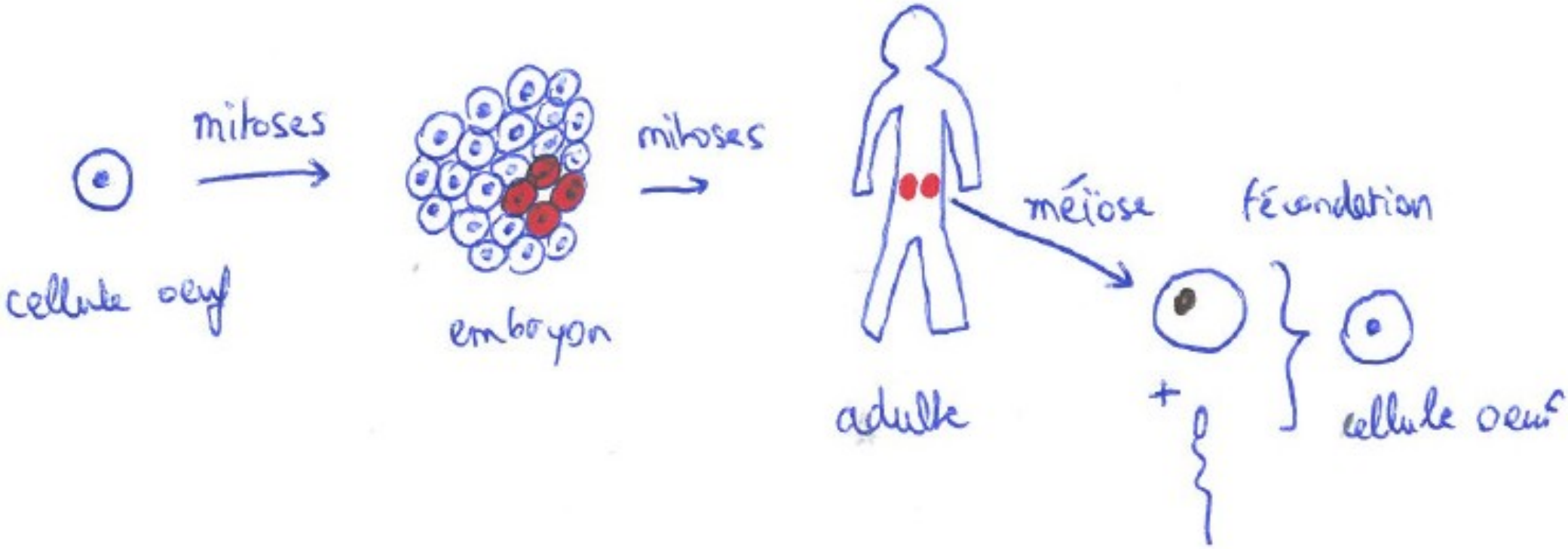
No finality/purpose: an end is achieved without having been the cause of the accomplishment of the effect

Gayon J. 2012 Evolution and chance

Assortment of chromosomes from father and mother

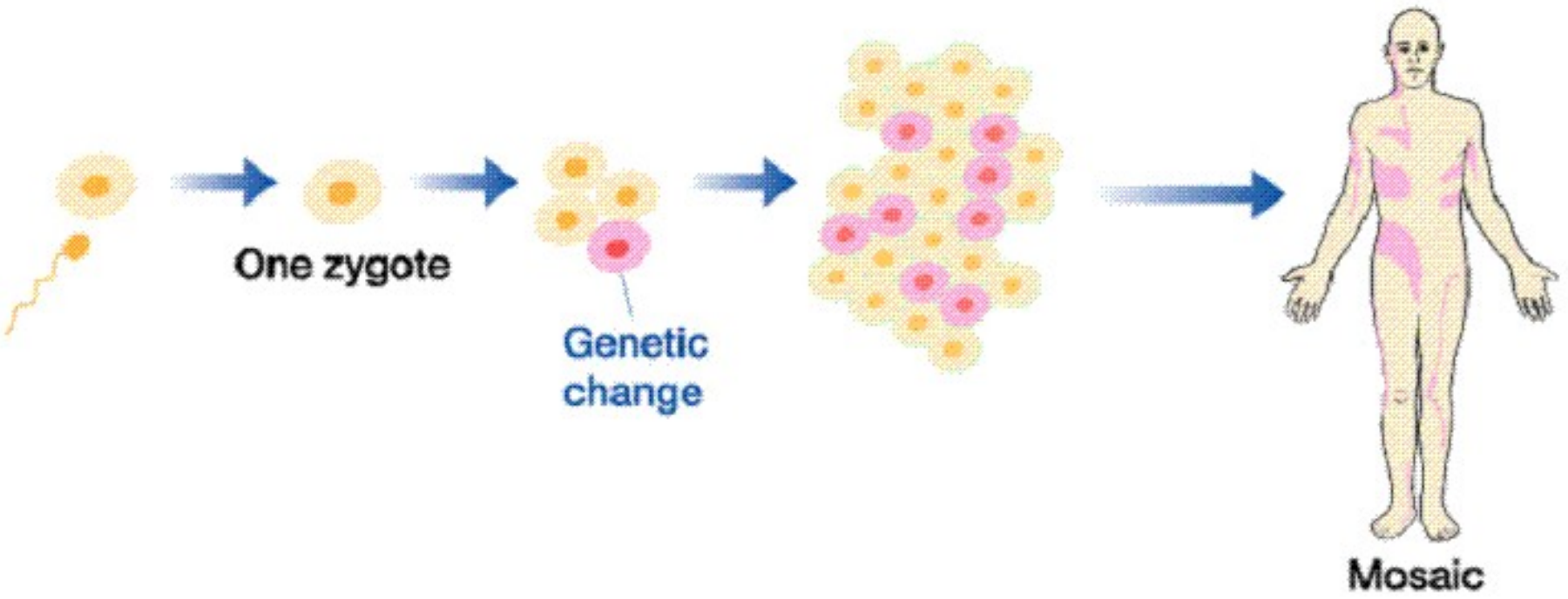




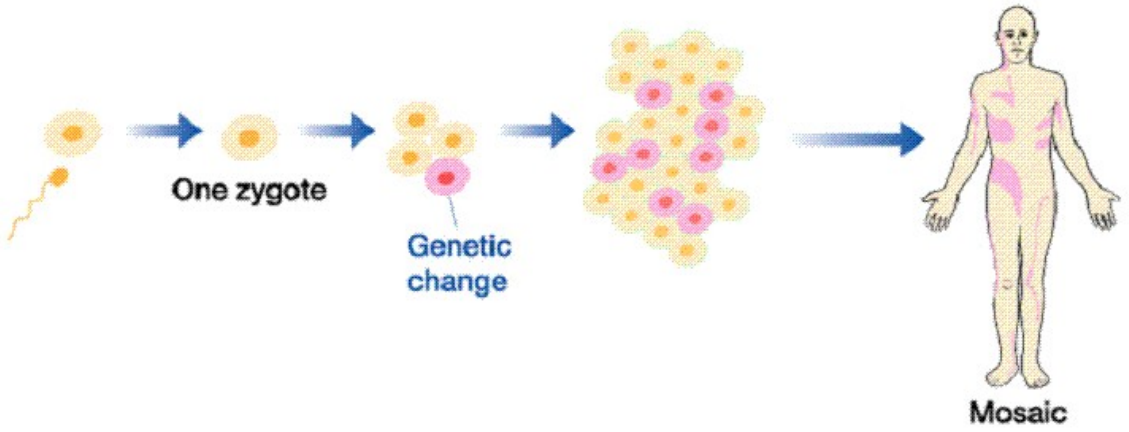


Germline

Somatic cells

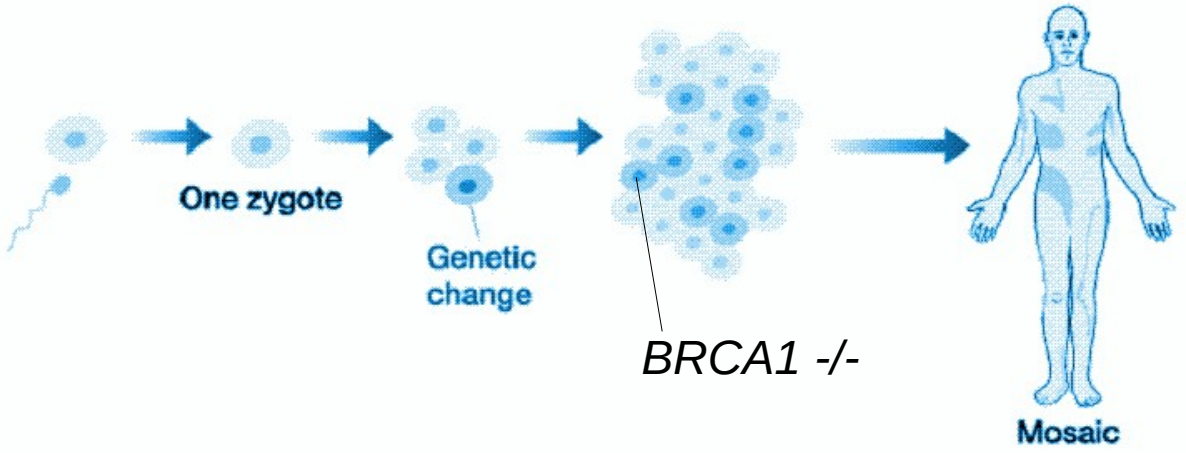

One zygote
Genetic change
Mosaic

woman
BRCA1 +/+



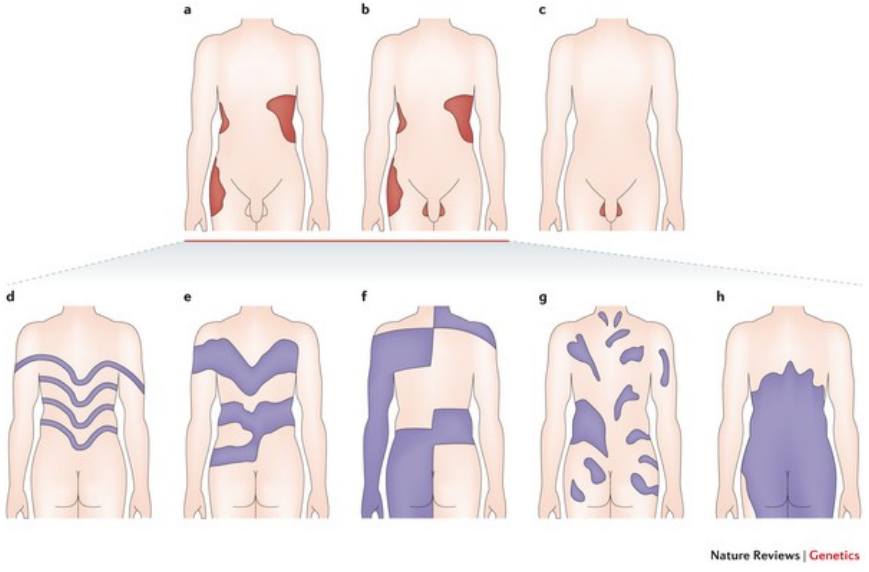
10% chance to develop breast cancer

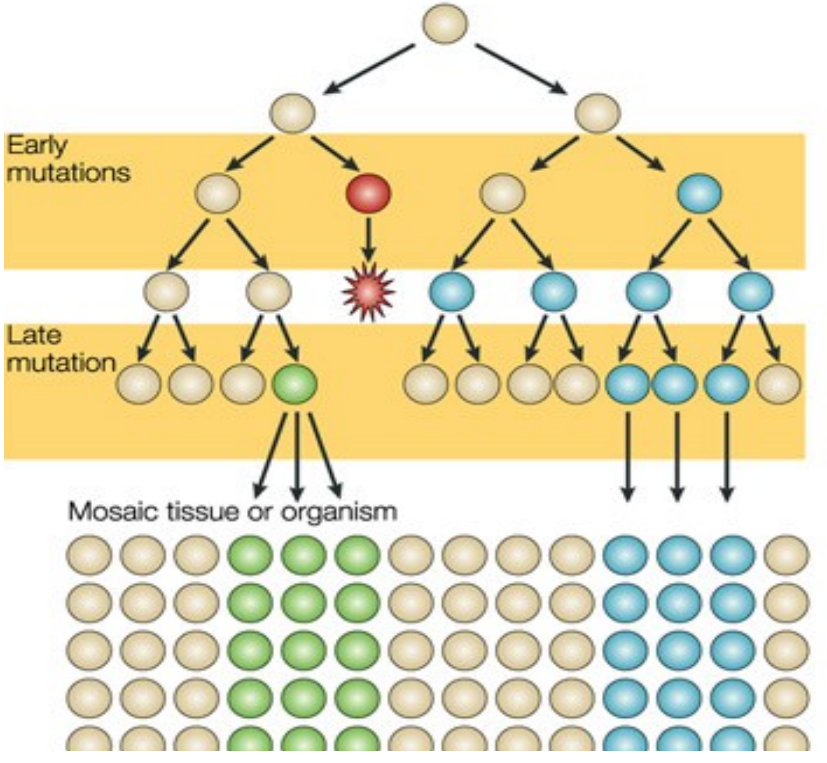
woman
BRCA1 +/-



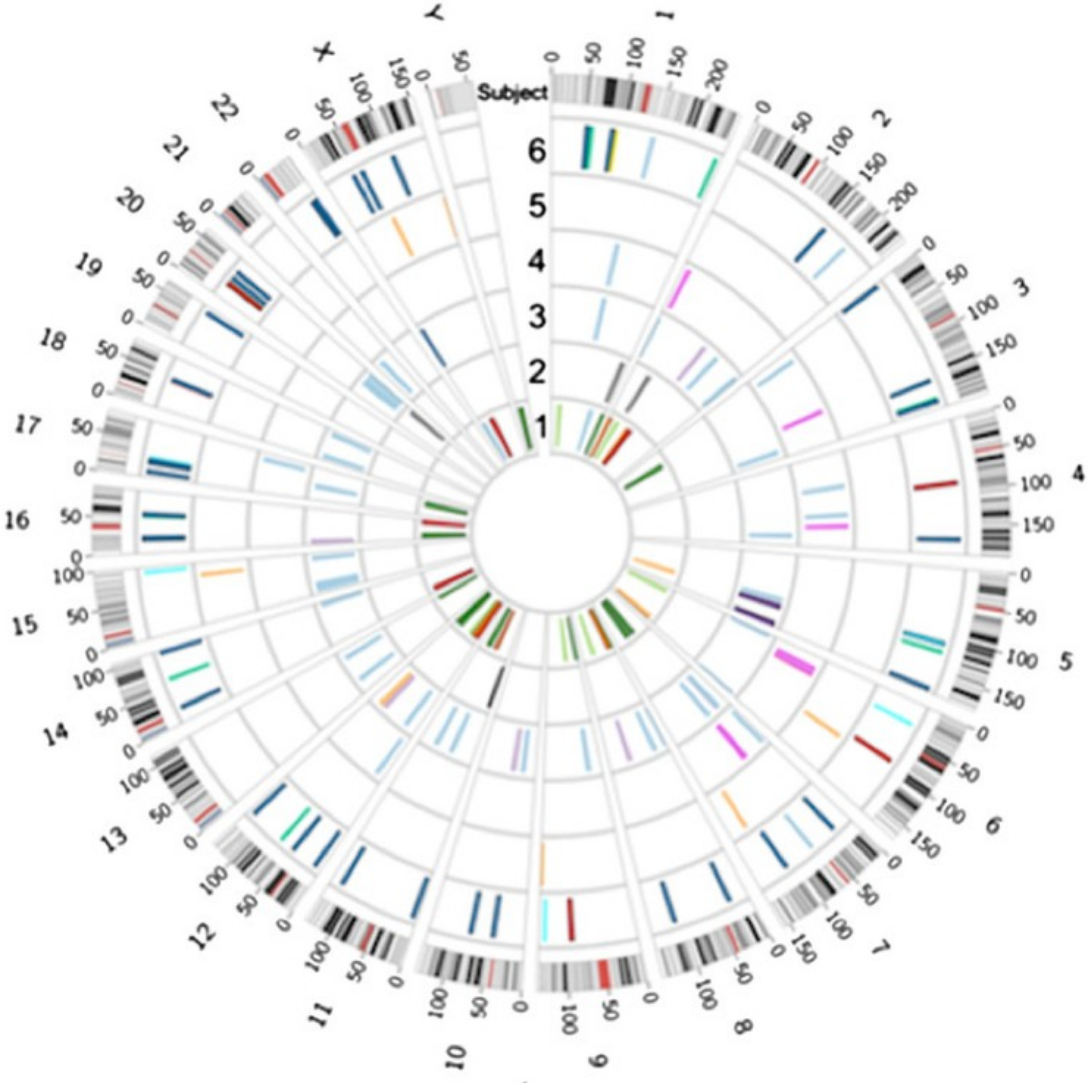
45% chance to develop breast cancer before 70 years old
Cancer cells will be *BRCA1 -/-*

Somatic mosaicism



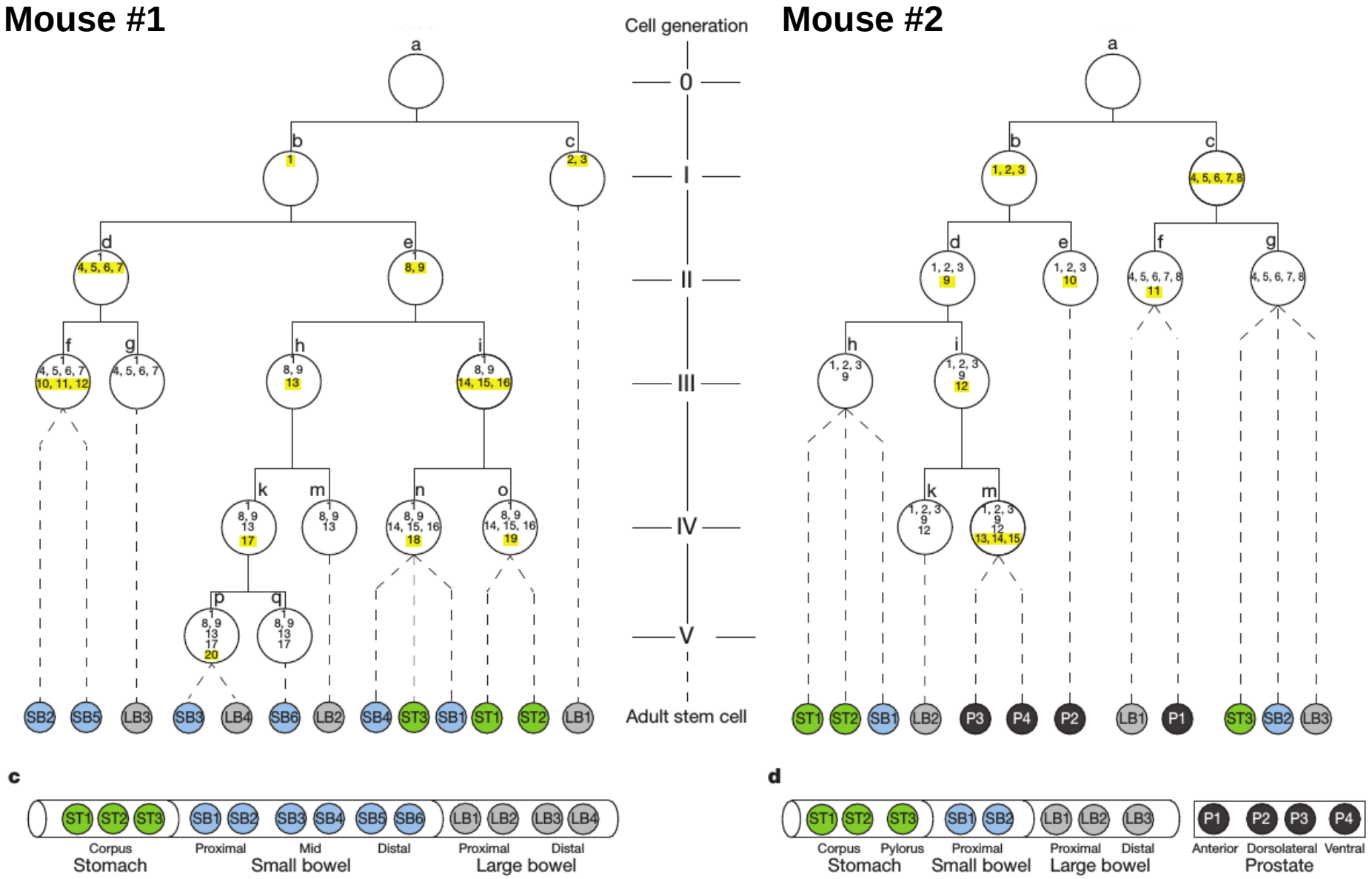


73 somatic CNVs in 11 tissues of six persons



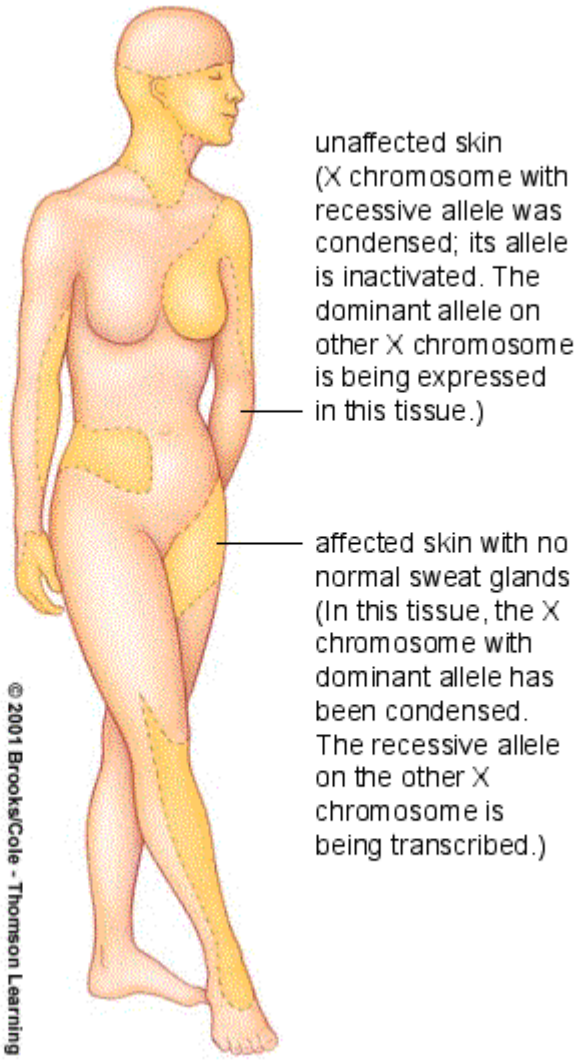
- O'Huallachain 2012 PNAS

Somatic mosaicism used to reconstruct cell lineages

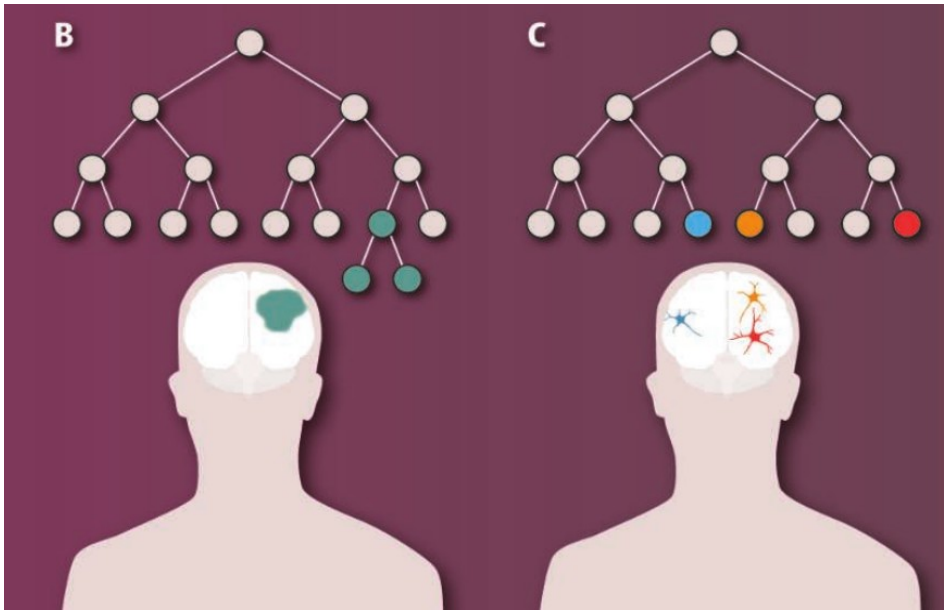


Behjati 2014 Nature

Female mosaicism
X inactivation pattern



Somatic transposition in human brain



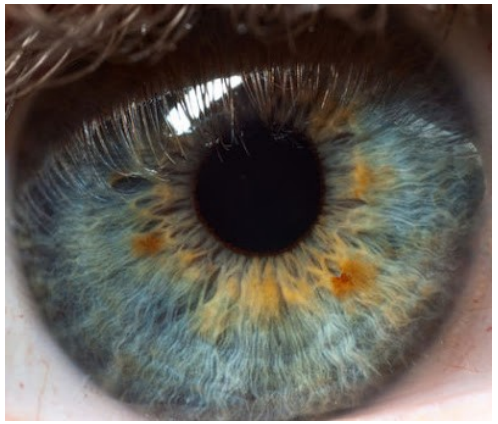
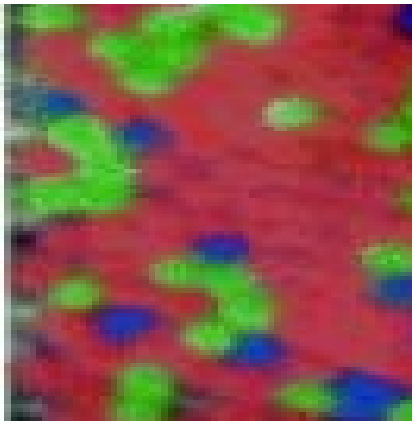
In three individuals:

in the hippocampus and caudate nucleus
7,743 somatic L1 insertions, 13,692 somatic Alu insertions and 1,350 SVA insertions

Baillie 2011 Nature

Developmental noise

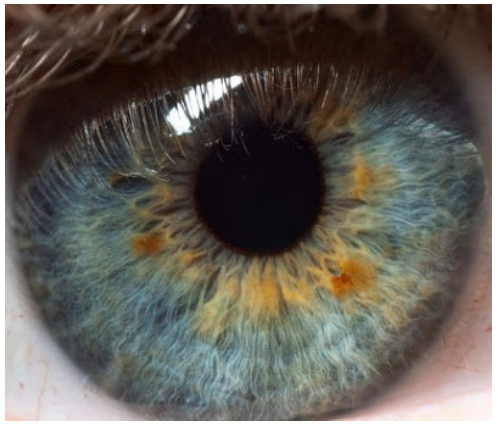
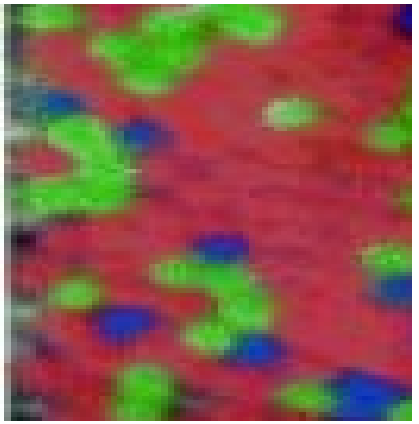
Differences between left and right sides of the body


ear shape, neuron connectivity, olfactory receptor gene expression, X inactivation pattern, organ cell number and size...

Developmental noise

Differences between left and right sides of the body


ear shape, neuron connectivity, olfactory receptor gene expression, X inactivation pattern, organ cell number and size...

Differences between twins

immune system cells, gait, arms crossing, voice, heart beat, brain waves...

Some can be attributed to variation in the number of determinant molecules

During terminal differentiation of mouse 3T3-L1 pre-adipocytes, individual TF abundance differs dramatically (from ~250 to >300,000 copies per nucleus) and the dynamic range can vary up to fivefold during differentiation.

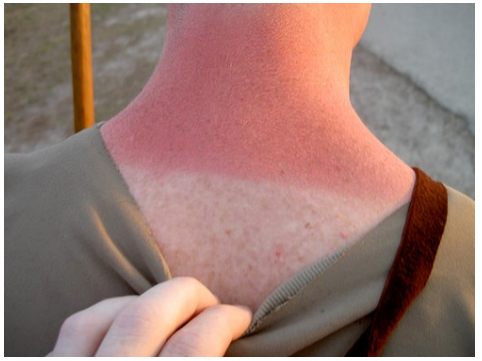
Simicevic 2013 Nature

Causes of phenotypic differences ?

Genetic

Epigenetic

Environment



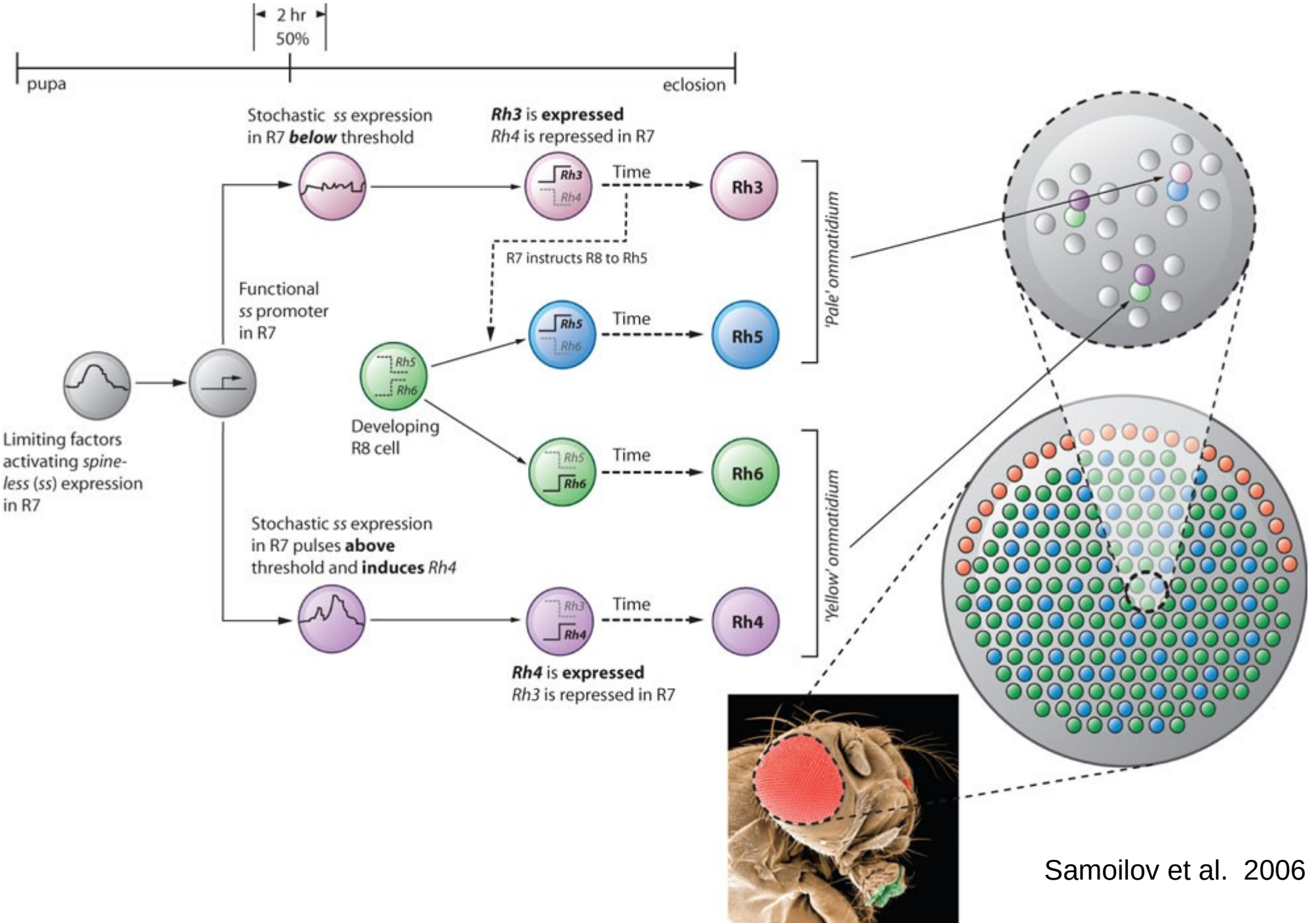
Stochasticity

Transmitted

Determistic

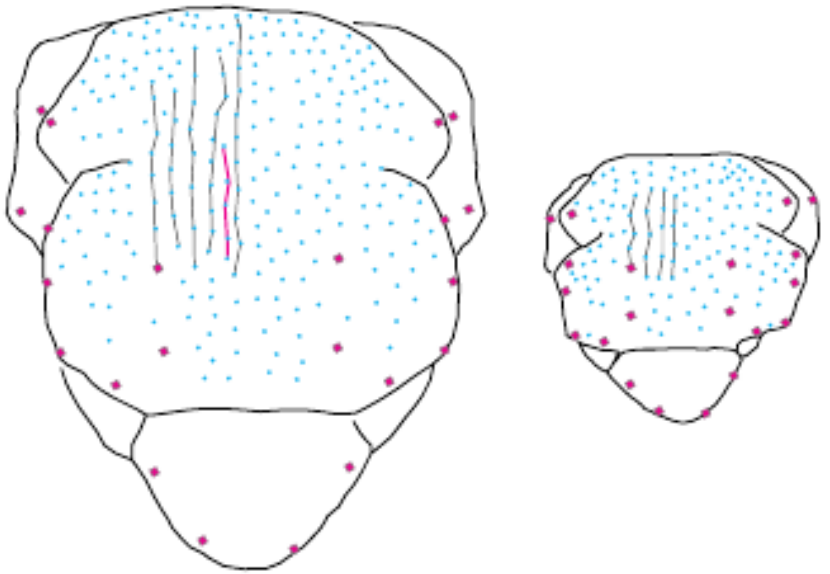
Interactions

Developmental noise can be “good”



Samoilov et al. 2006

Robustness



Well-fed Semi-starved

Robustness

Absence or low variation of a phenotype when faced with an incoming variation

1) Of what?

2) To what? To either:

- stochastic variation
- environmental variation: specify
- genetic variation: specify

3) How much?

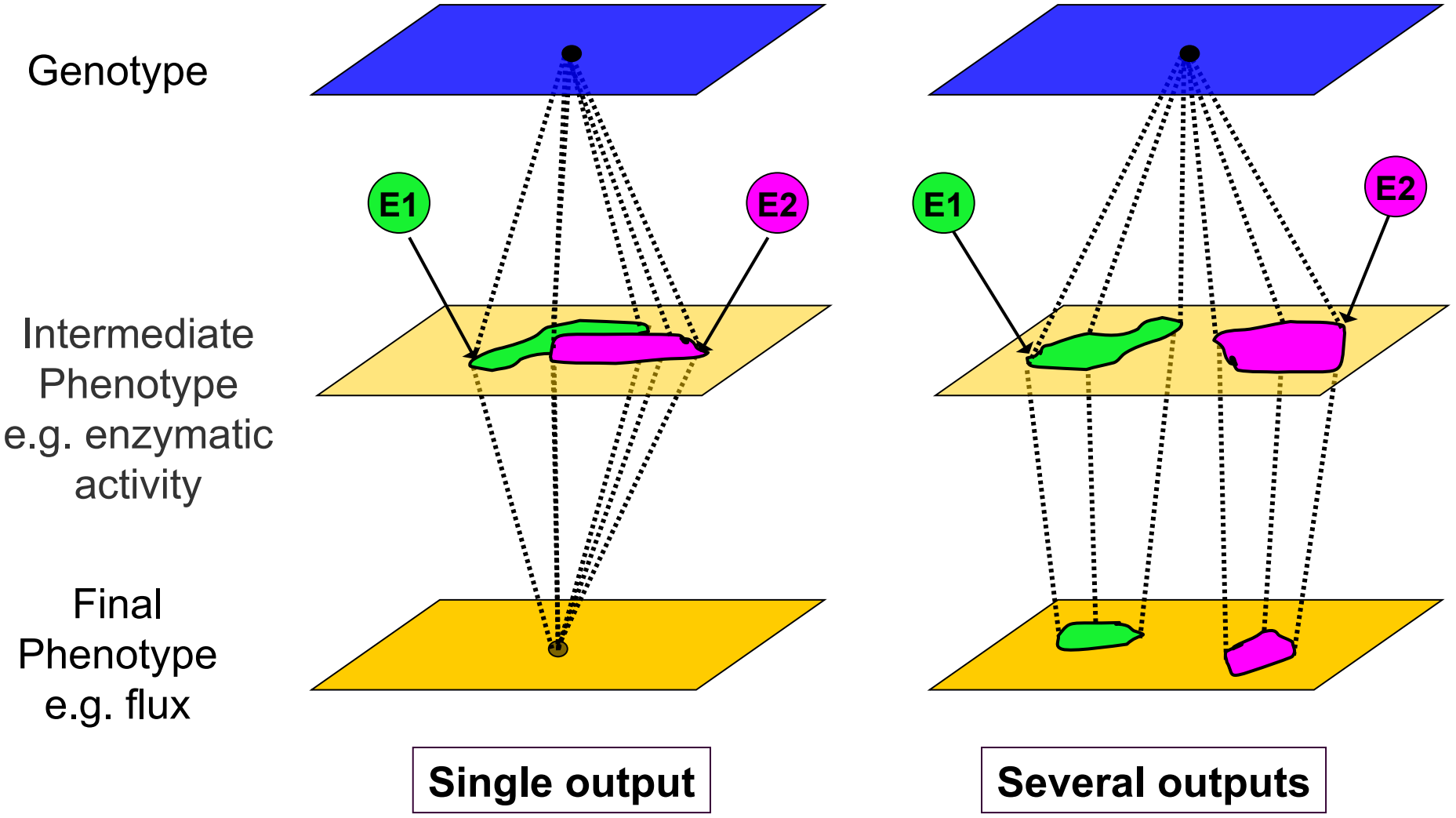
Different phenotypic metrics

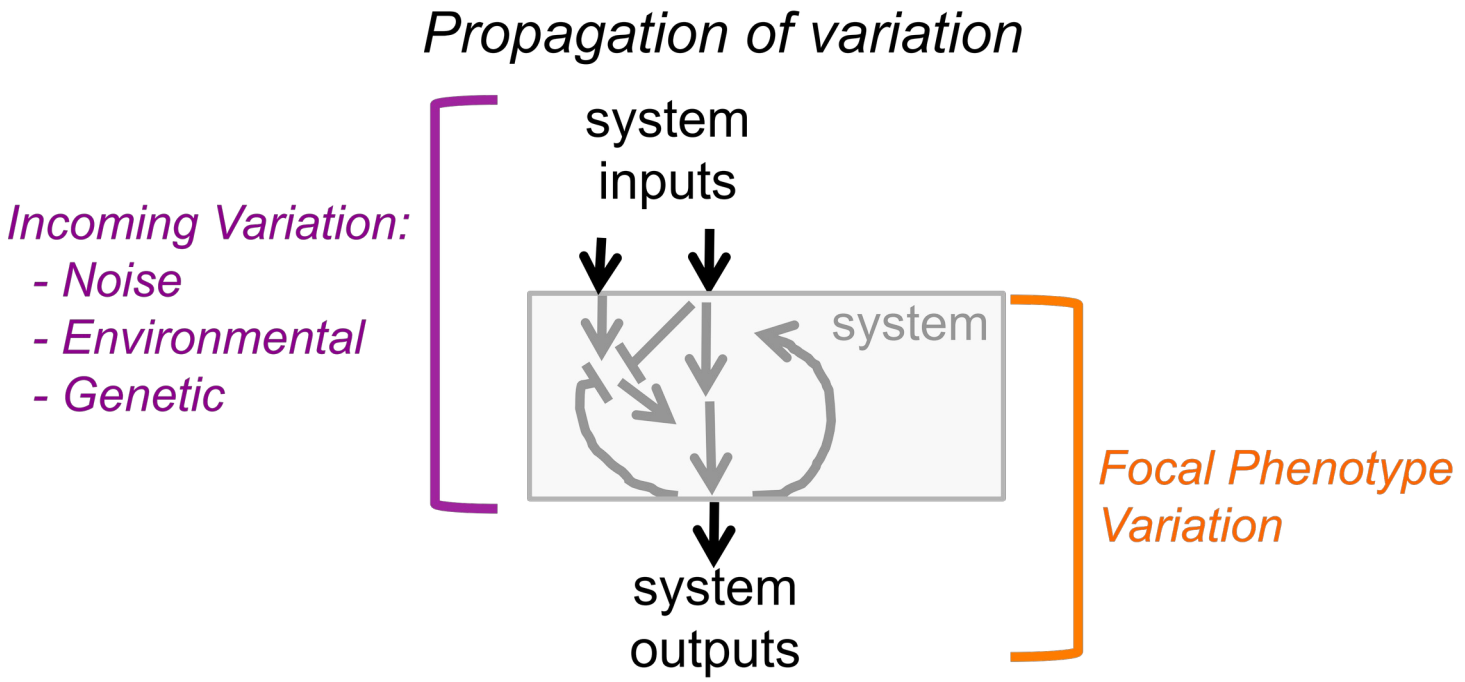
Coefficient of variation: standard deviation/mean

Historically:
quantitative genetics (low variance, canalization)
physics/chemistry/engineering (robustness, buffering)

Canalization: mechanisms that make the system follow a certain trajectory

Trait plasticity versus invariance (robustness) at different levels of the genotype-phenotype map

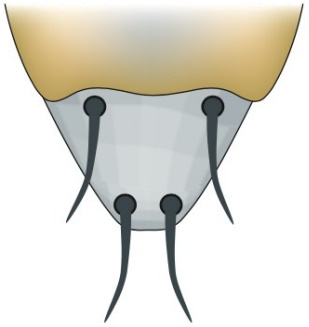


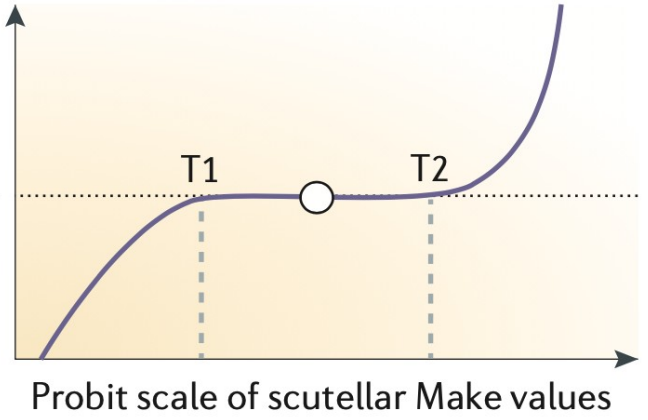


Felix & Barkoulas 2015

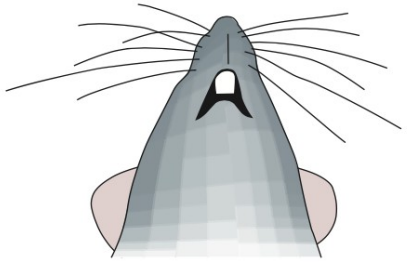
b Experiments

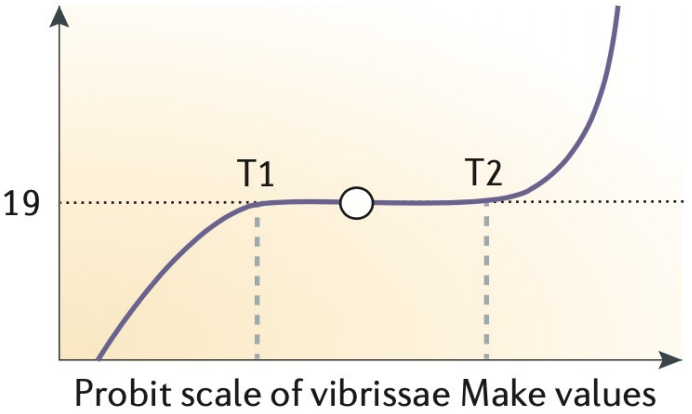
Number of scutellar bristles



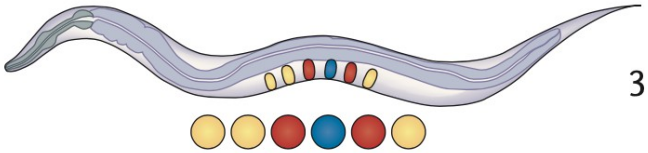


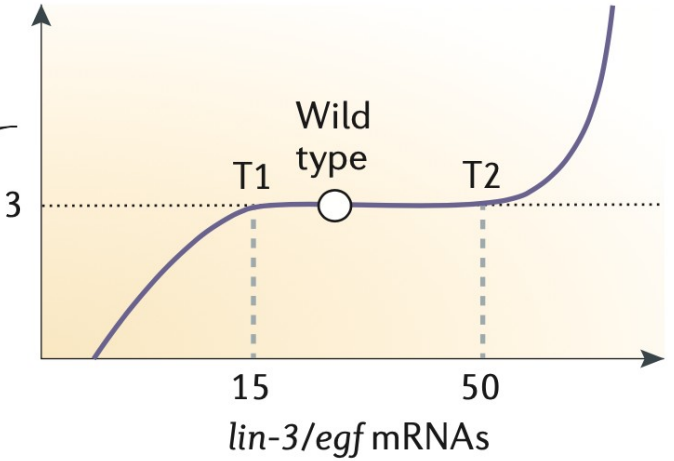
Number of vibrissae





Vulval induction index





Felix & Barkoulas 2015

Causes of robustness

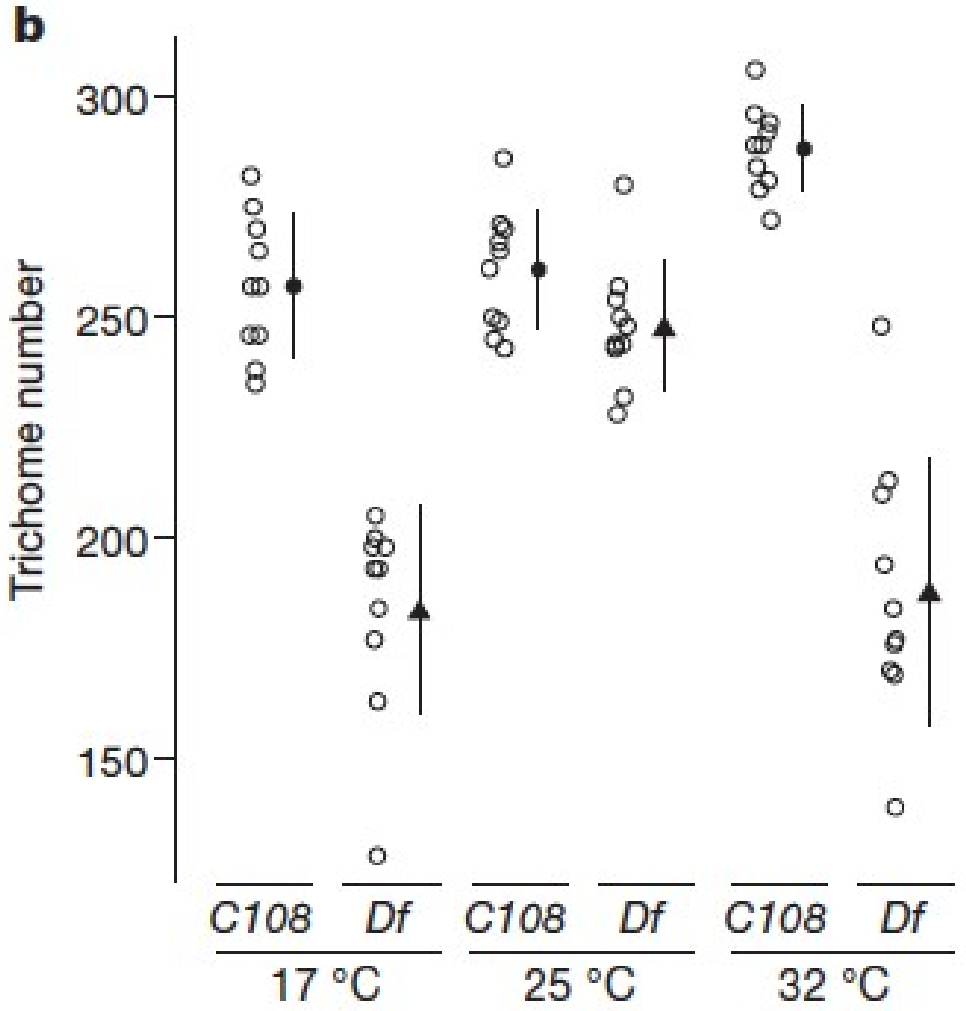
Non-linearity

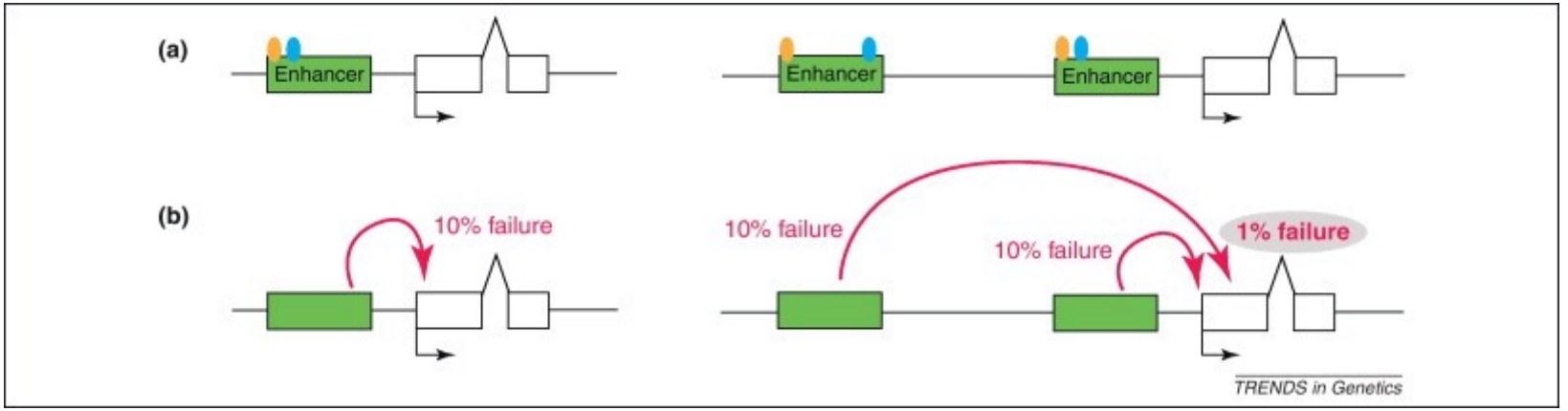
Redundancy

nature Vol 466|22 July 2010|doi:10.1038/nature09158

LETTERS

Phenotypic robustness conferred by apparently redundant transcriptional enhancers





Cryptic genetic variation

Cryptic genetic variation

First requires defining the *phenotype of interest*

Genetic variation that has no effect on phenotype of interest

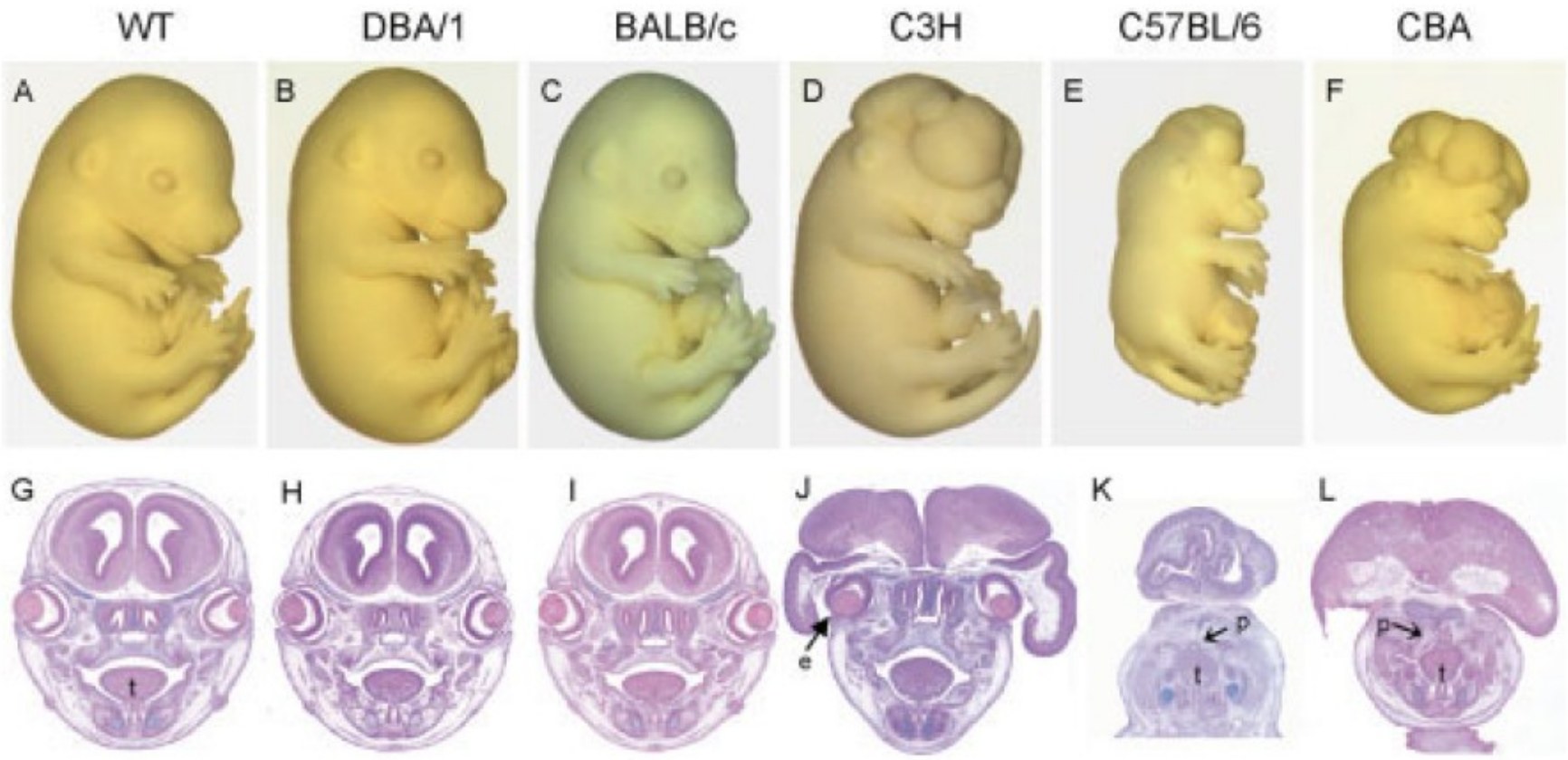
... but may be revealed *under some circumstances*
by its effect on this phenotype

Cryptic genetic variation (CGV) is defined as standing genetic variation that does not contribute to the normal range of phenotypes observed in a population, but that is available to modify a phenotype that arises after environmental change or the introduction of novel alleles.

Gibson & Dworkin *Nat Rev Gen* 2004

Expressivity of one mutation varies with wild genetic gackground

Tcof1/- heterozygote mice



Dixon & Dixon *Dev Dyn* 2004

Influence of contingency

Resistance
to carbamates and
organophosphates

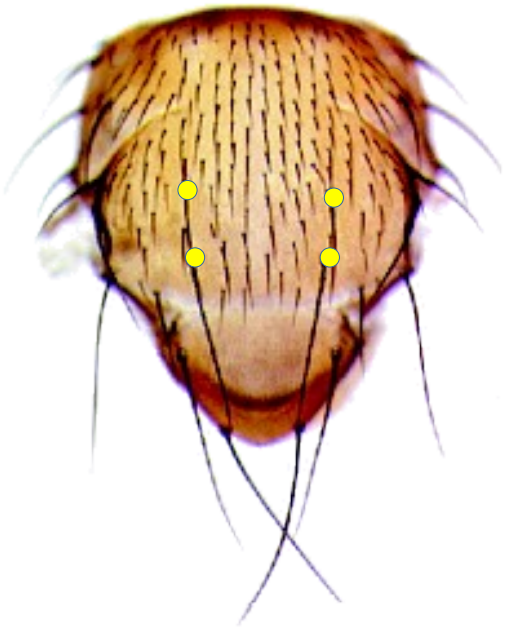
position 119
in *AchE1* gene

	sensitive		highly resistant
Anopheles albimanus *Anopheles gambiae* *Culex pipiens pipiens* *Culex pipiens* *Culex pipiens molestus*	GGY Gly	1 mutation →	**A**GY Ser
Evolved high resistance			
Aedes aegypti *Aedes albopictus* *Aedes taeniorynchus* *Anopheles sacharovi* *Anopheles stephensi* *Culex cinereus* *Culex pereexiguus* *Culex tarsalis* *Culex thelieri* *Culex tritaeniorhynchus* *Ochlerotatus caspius* *Ochlerotatus detritus*	GGR Gly	2 mutations → →	**AGY** Ser
Did not evolve high resistance			

.. But might if more time is allowed

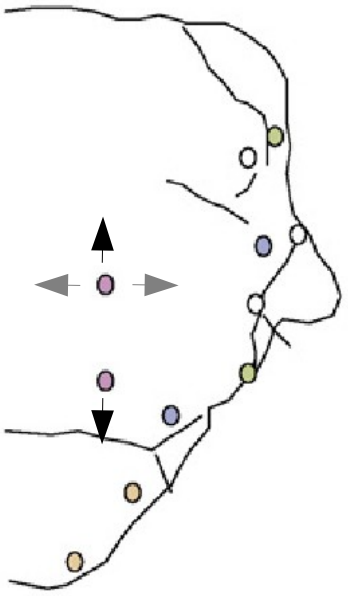
Weill et al. 2004

The genome constrains evolution



Standing variation

D. melanogaster

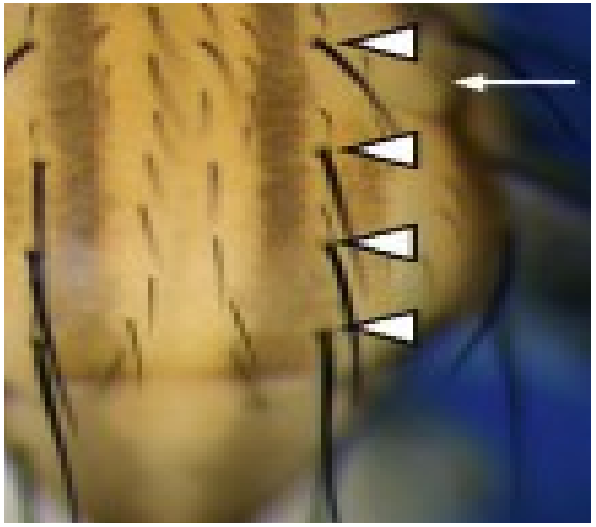


variation
no variation

Garcia-Vázquez 1988 J. Heredity

Natural evolution

D. quadrilineata

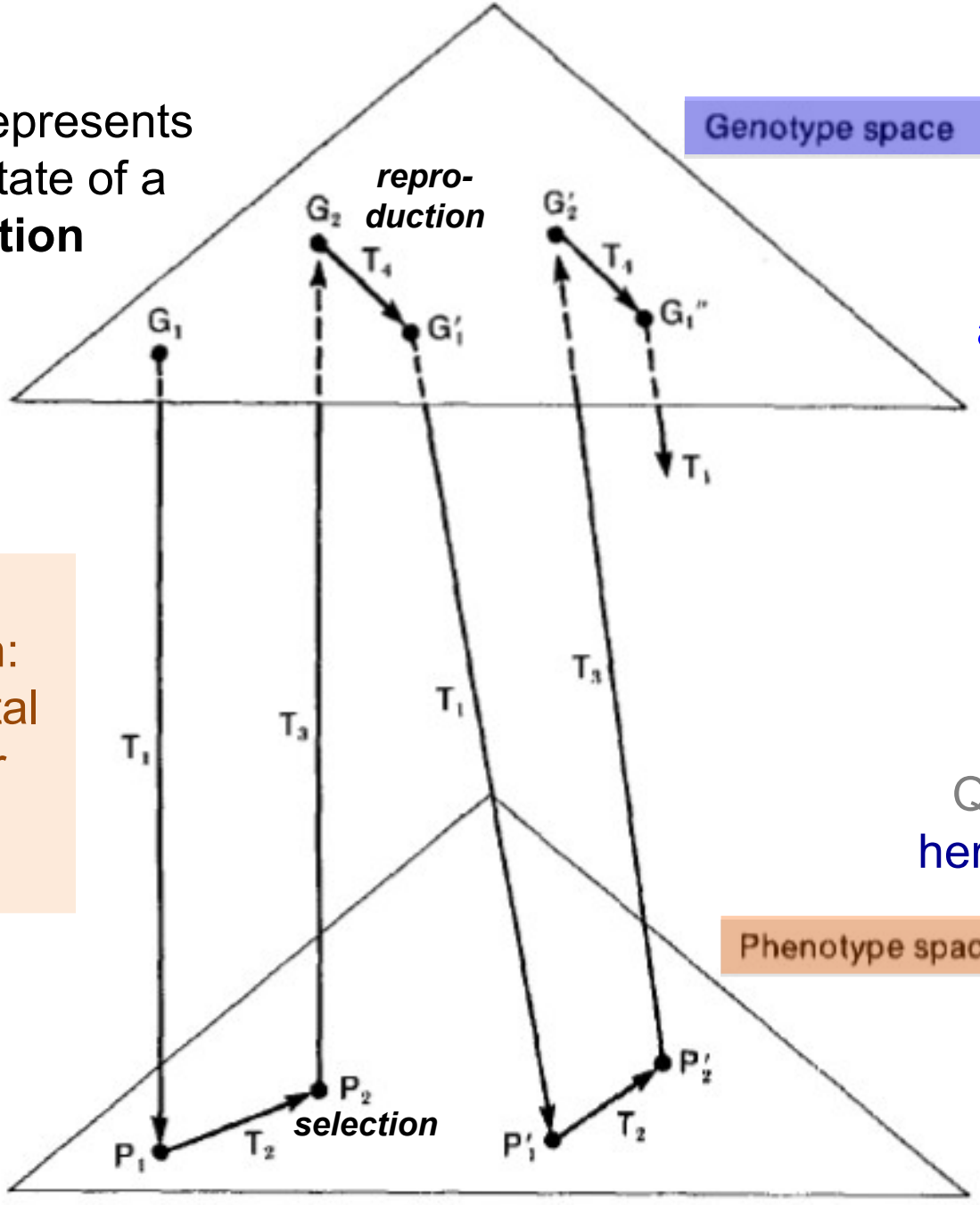


Marcellini et al 2006 PloS Biol

The Genotype-Phenotype Map

The first genotype-phenotype map

Here a dot represents the mean state of a **population**



Phenotype construction: developmental and cellular biology physiology

Population genetics: stochastic processes and selection coefficient

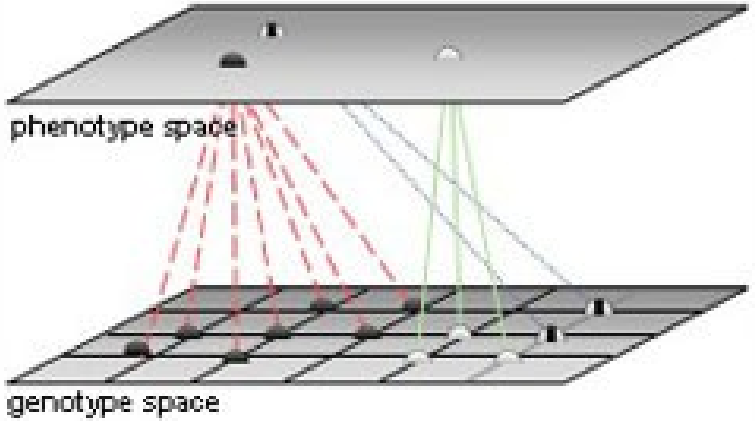
Quantitative genetics: heritability of phenotypes

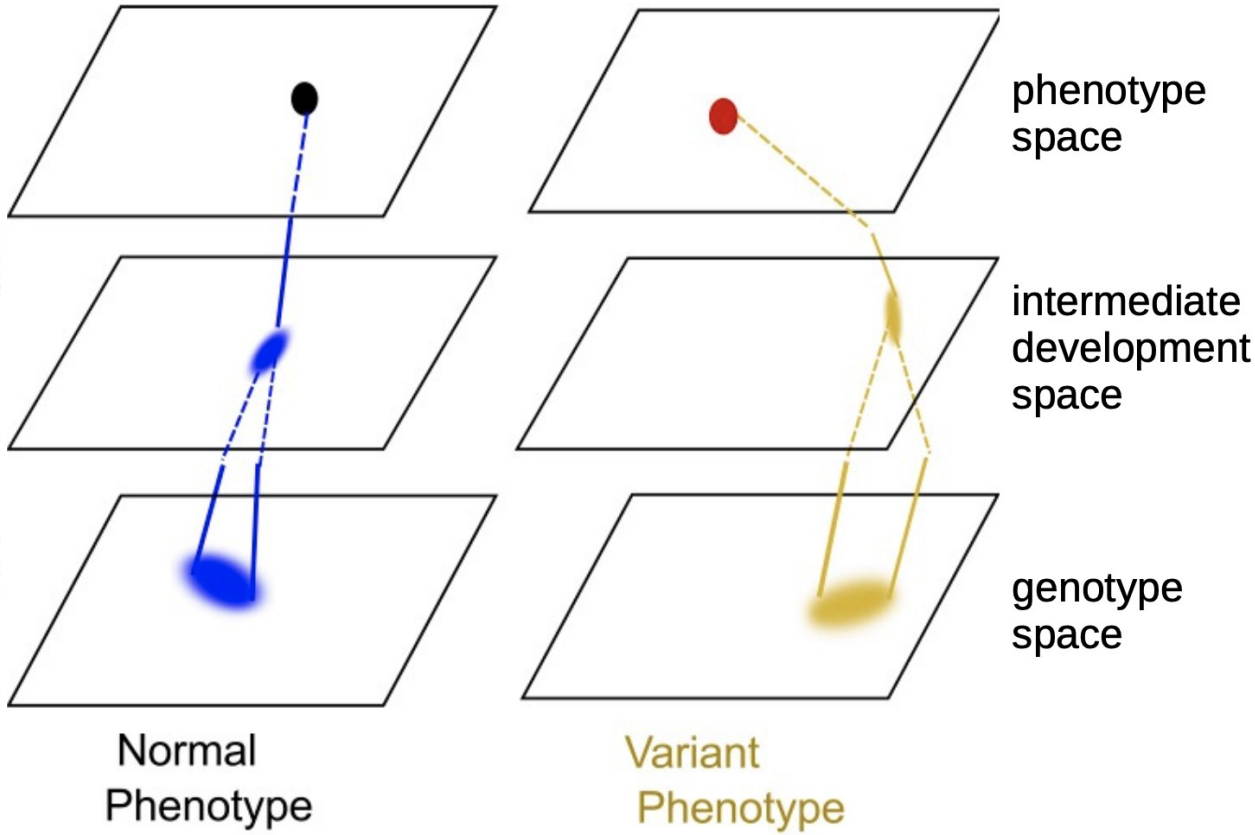
Evolutionary biology of phenotypes, evolutionary ecology

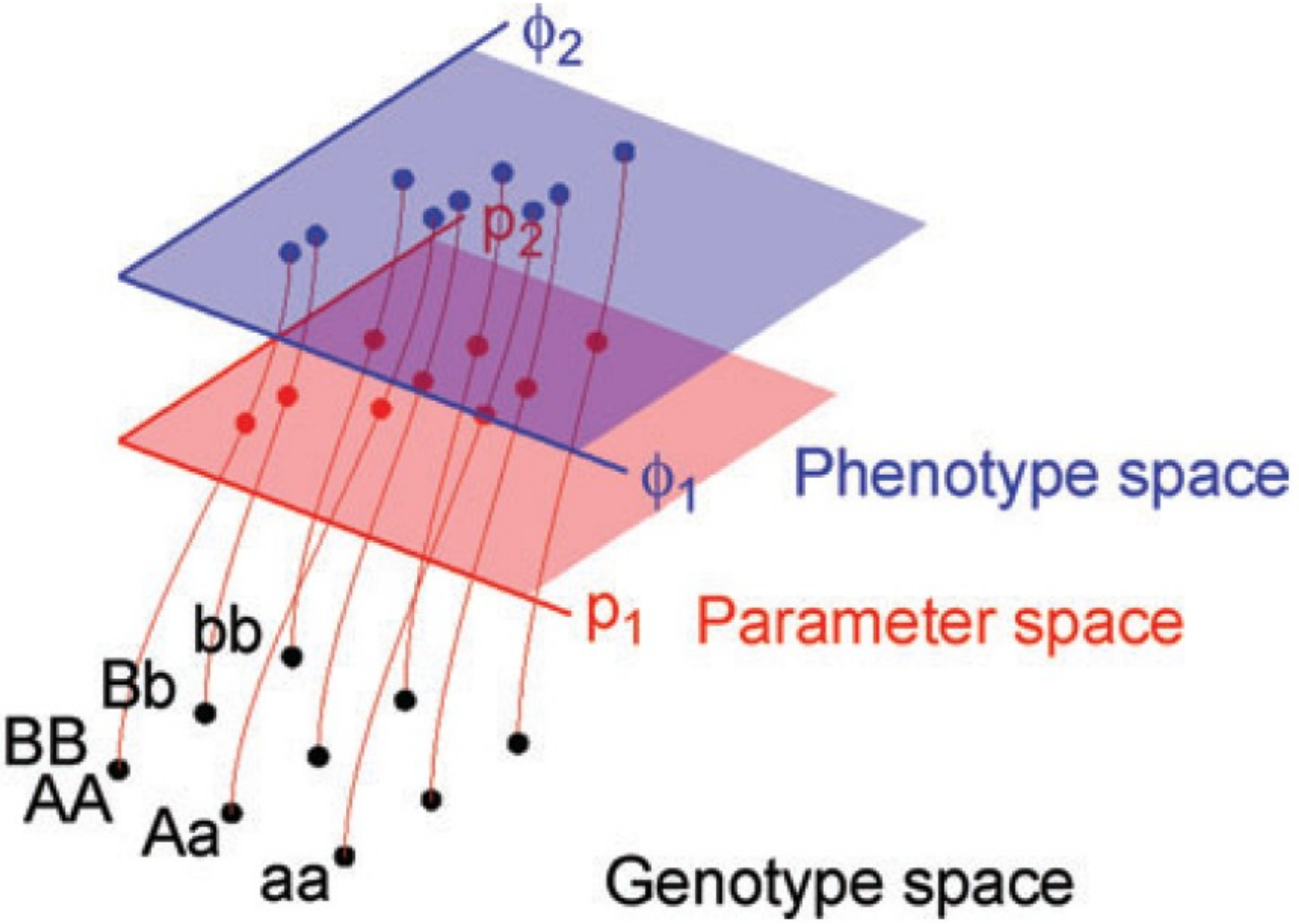
Lewontin 1974

Intermediate steps in the genotype-phenotype map

Here a dot represents one **individual**







Gjuvsland et al. 2013

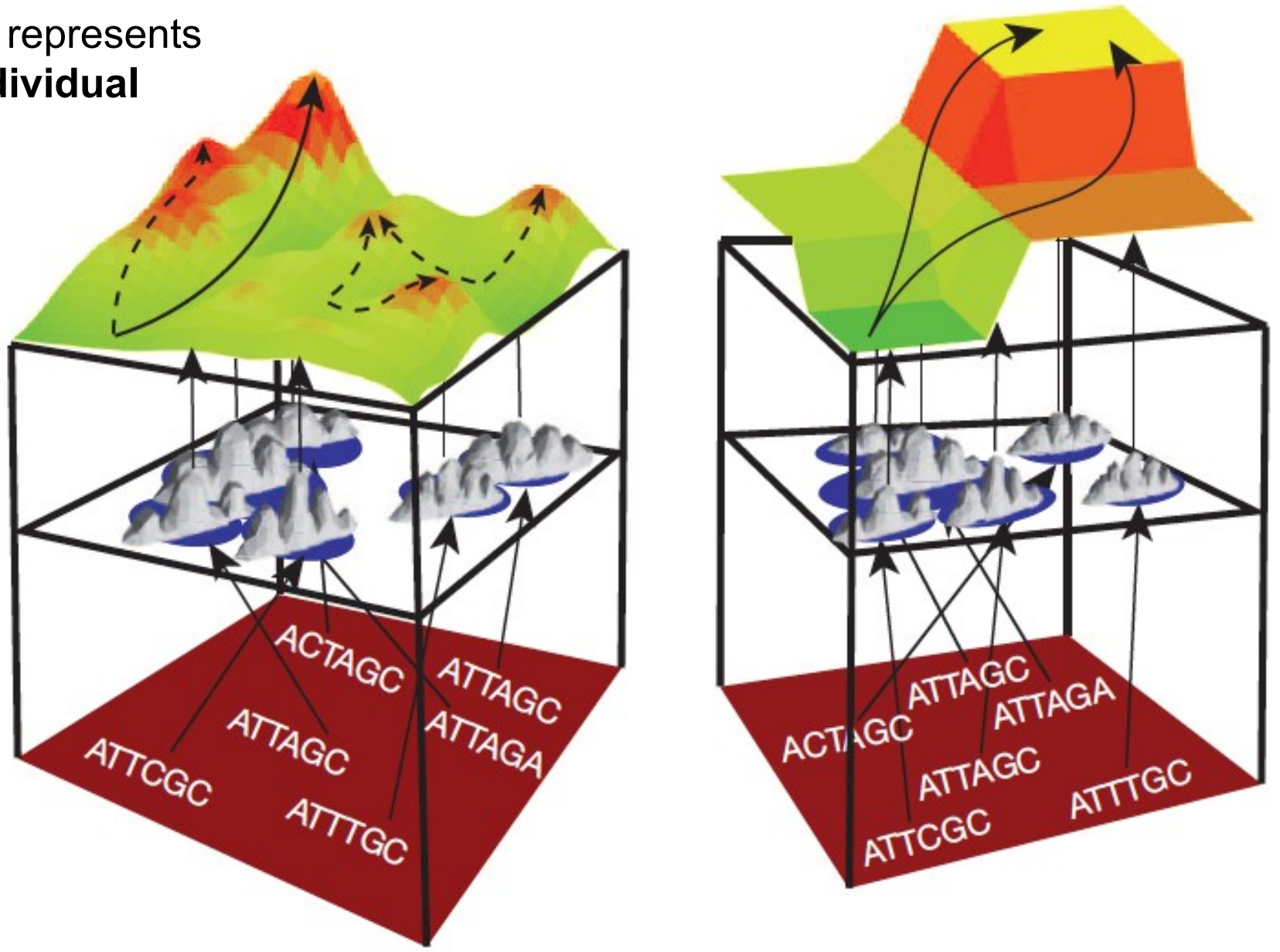
The genotype-phenotype-fitness map

Here a dot represents one **individual**

Fitness

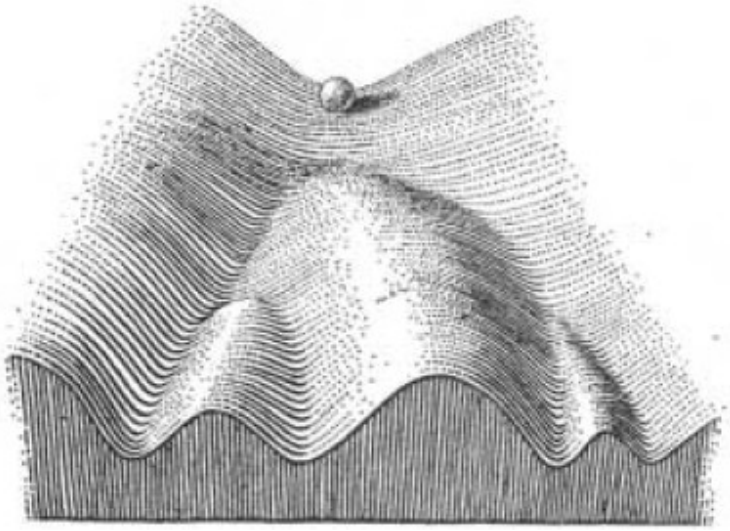
Phenotype
Tooth
morphology

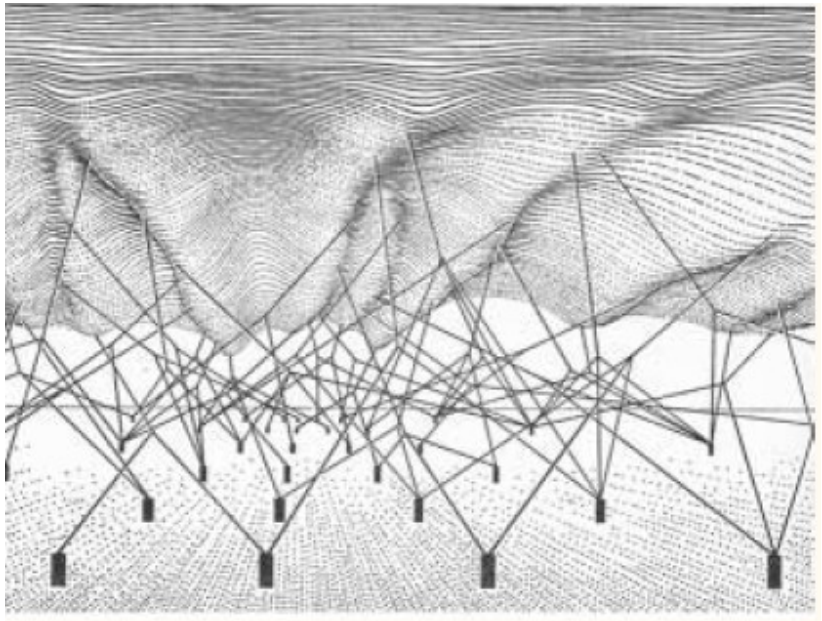
Genotype



Salazar-Cuidad & Martin-Riera 2013

The Epigenetic Landscape
A metaphor for the G-P relationship



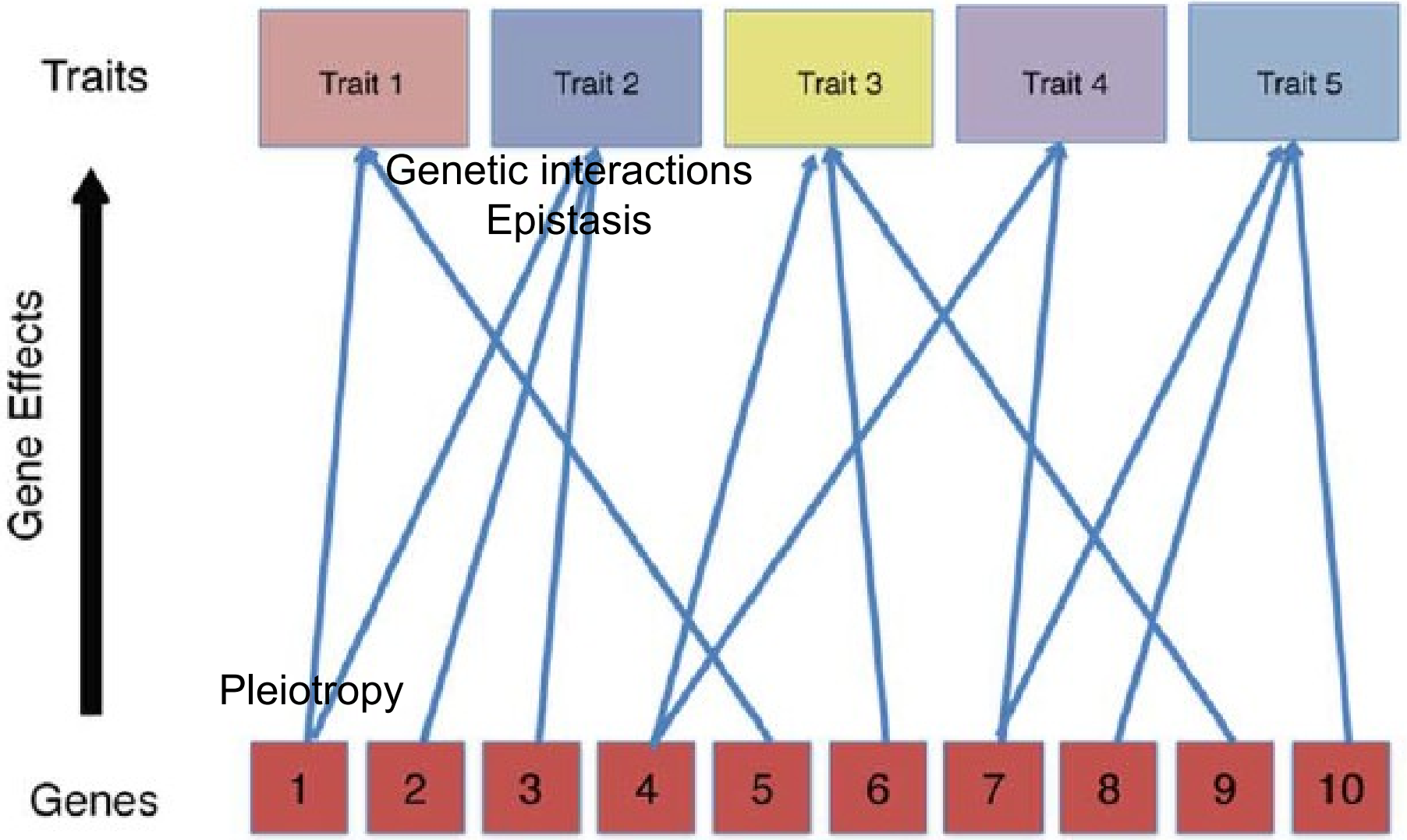


Development

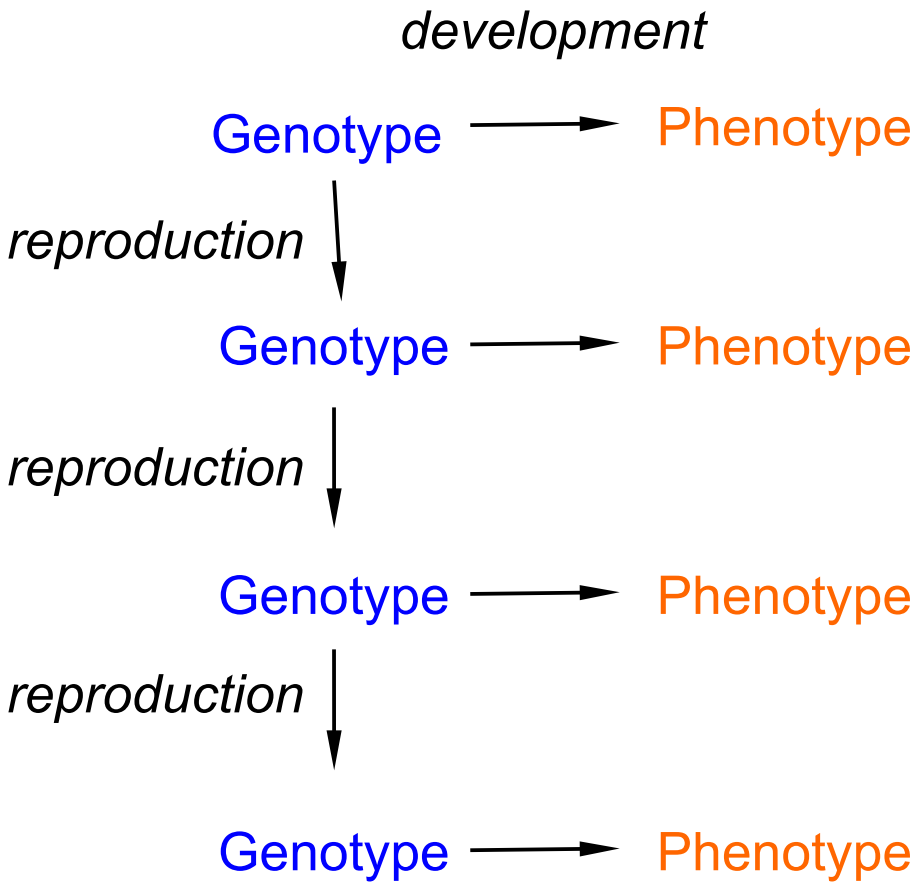
Canalization

Genes underlying
the landscape

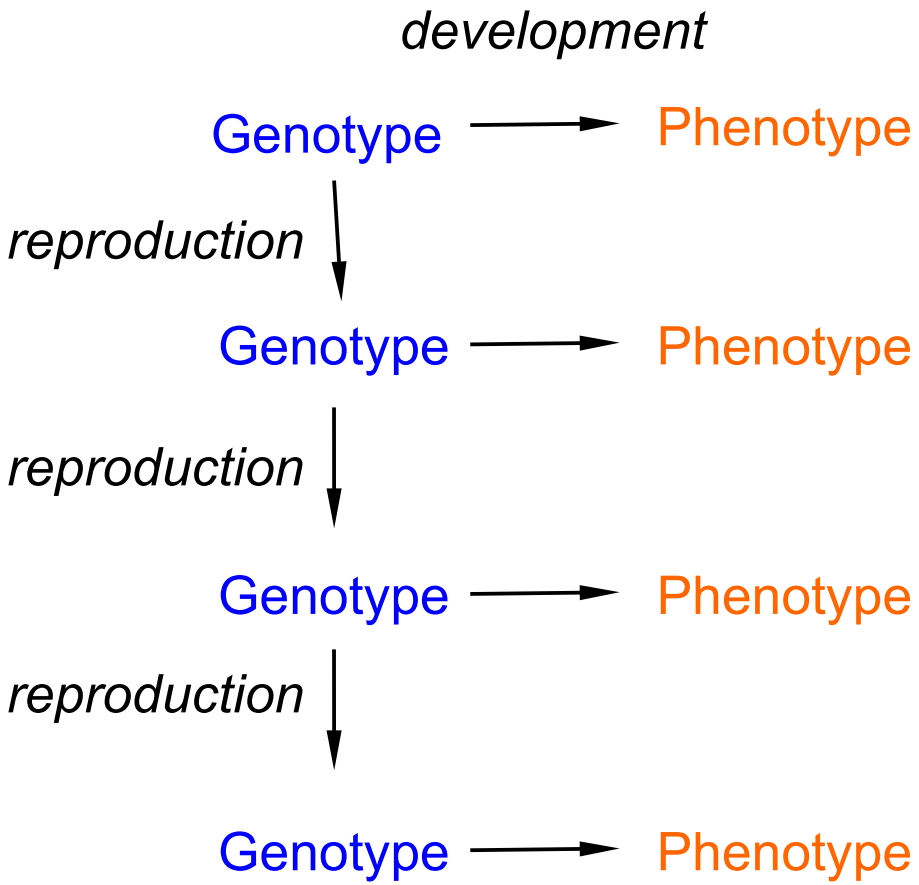
Waddington 1957



Hallgrimsson et al. 2014


development
Genotype
Phenotype
reproduction
Genotype
Phenotype
reproduction
Genotype
Phenotype
reproduction
Genotype
Phenotype

A simplistic view



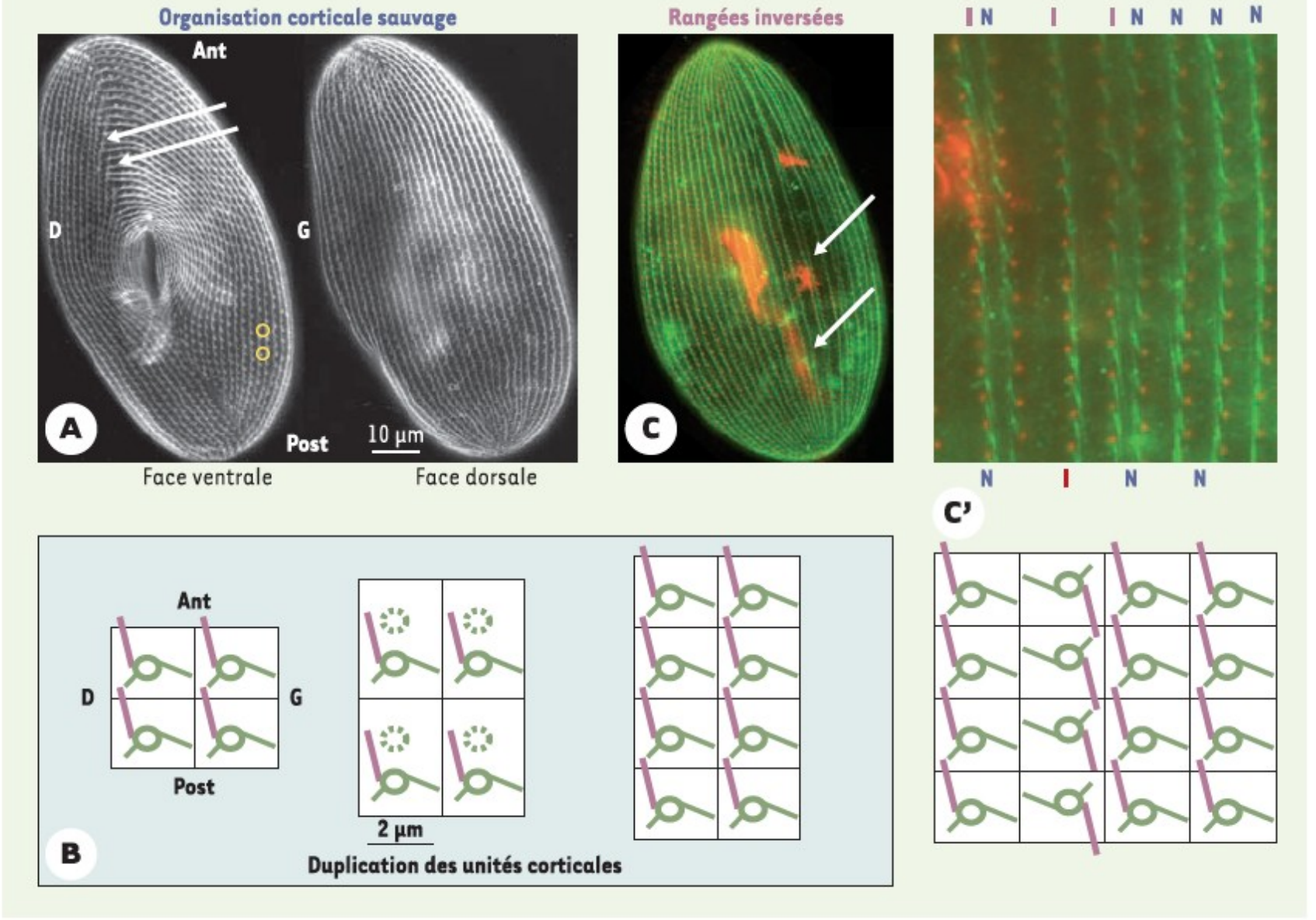
Heritable traits are not always due to genes

The genotype does not determine entirely the phenotype

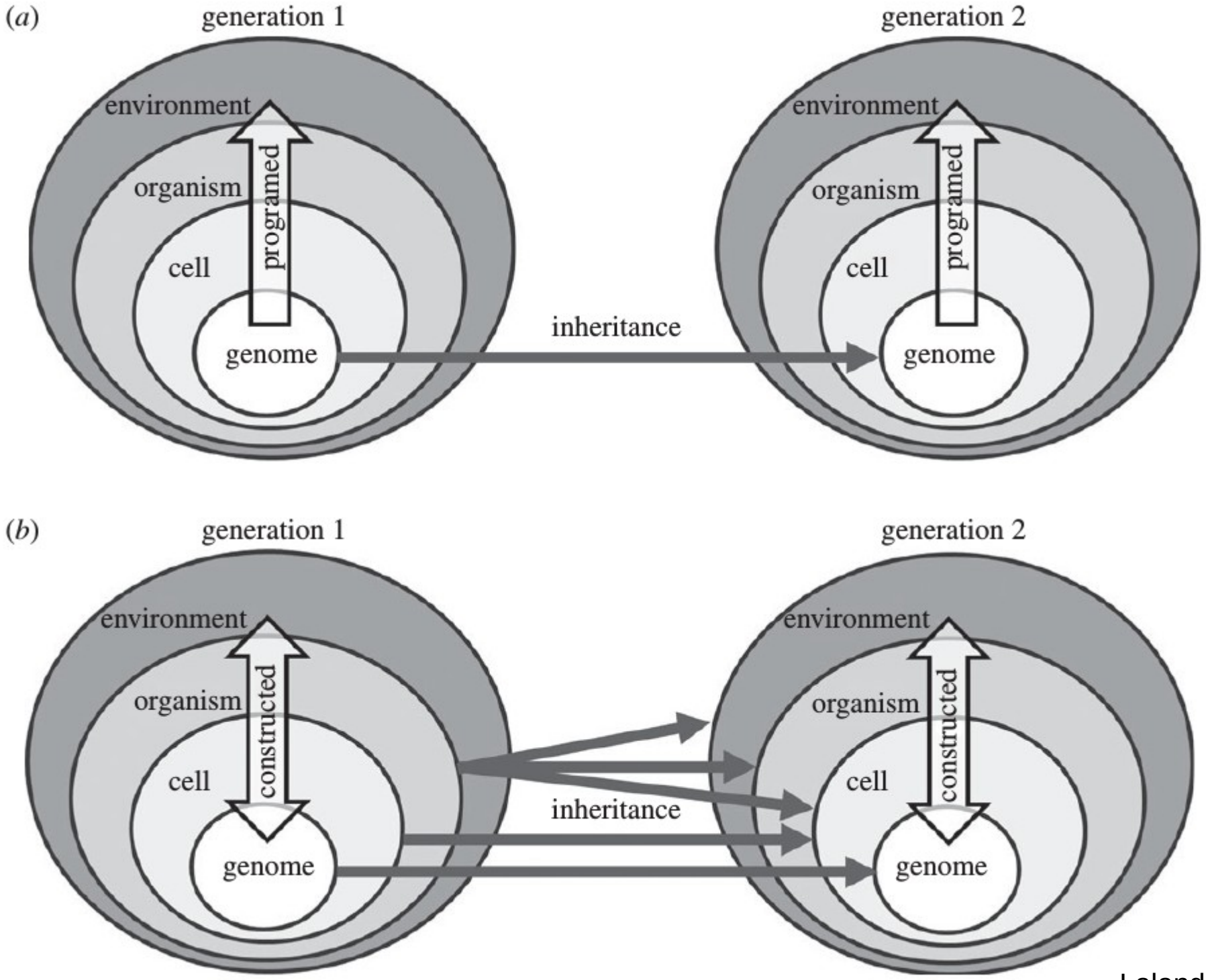
The genotype cannot replicate by itself

Genotype and phenotype imply variation

Cortical heredity in *Paramecium*



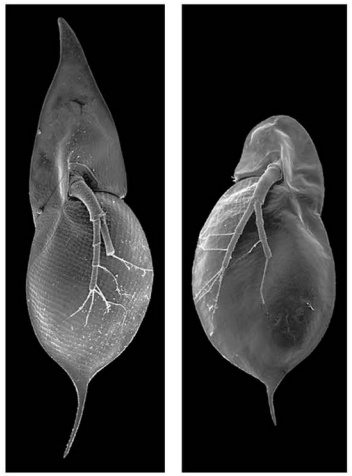
Meyer & Beisson M/S 2005


(a)
generation 1
generation 2
environment
organism
cell
genome
programed
inheritance
environment
organism
cell
genome
programed
(b)
generation 1
generation 2
environment
organism
cell
genome
constructed
inheritance
environment
organism
cell
genome
constructed

Laland 2015

Plasticity: one genotype → several phenotypes

Daphnia



with helmet | without helmet

Nemoria arizonaria caterillars

spring: caterpillars feed on catkins

summer: caterpillars feed on leaves

Water crowfoot plant

leaves growing above water | leaves growing below water

Desert locusts

solitary

gregarious

Commodore butterfly

winter

summer

Commodore butterly: Michael Wild, CC-BY-SA-3.0 (winter), Svdmolen, CC-BY-SA-3.0 (summer)

Daphnia: Agrawal *et al* (1999)

Nemoria arizonaria caterillars: Sadava *et al* (2014)

Water crowfoot plant: J R Crellin, CC BY-NC-ND 3.0

Complexifications of the G-P map

Genetic Linkage

Epistasis

Supergene

Pleiotropy

GxE
Plasticity

Large number of alleles

Noise

Robustness

Cryptic genetic variation

Epigenetics