

The genetic basis of phenotypic diversity

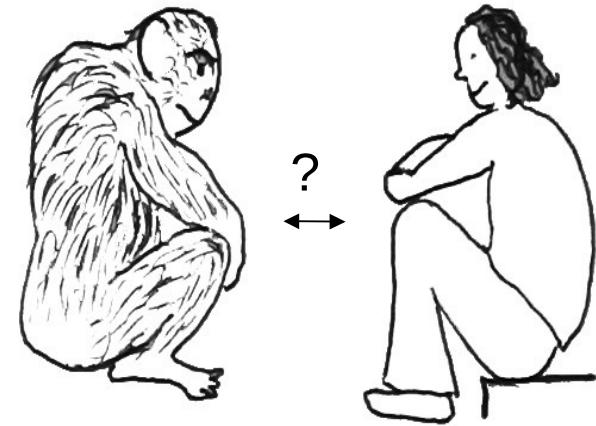
**Virginie Orgogozo
Institut Jacques Monod
Paris**

What makes us different?

between individuals



between species



How do genotypes map onto phenotypes ?

Phenotype = observable attributes of an individual

Genotype = inheritable genetic material = DNA or RNA

1) Identifying the mutations responsible for phenotypic differences

2) Gephé : thinking in terms of differences

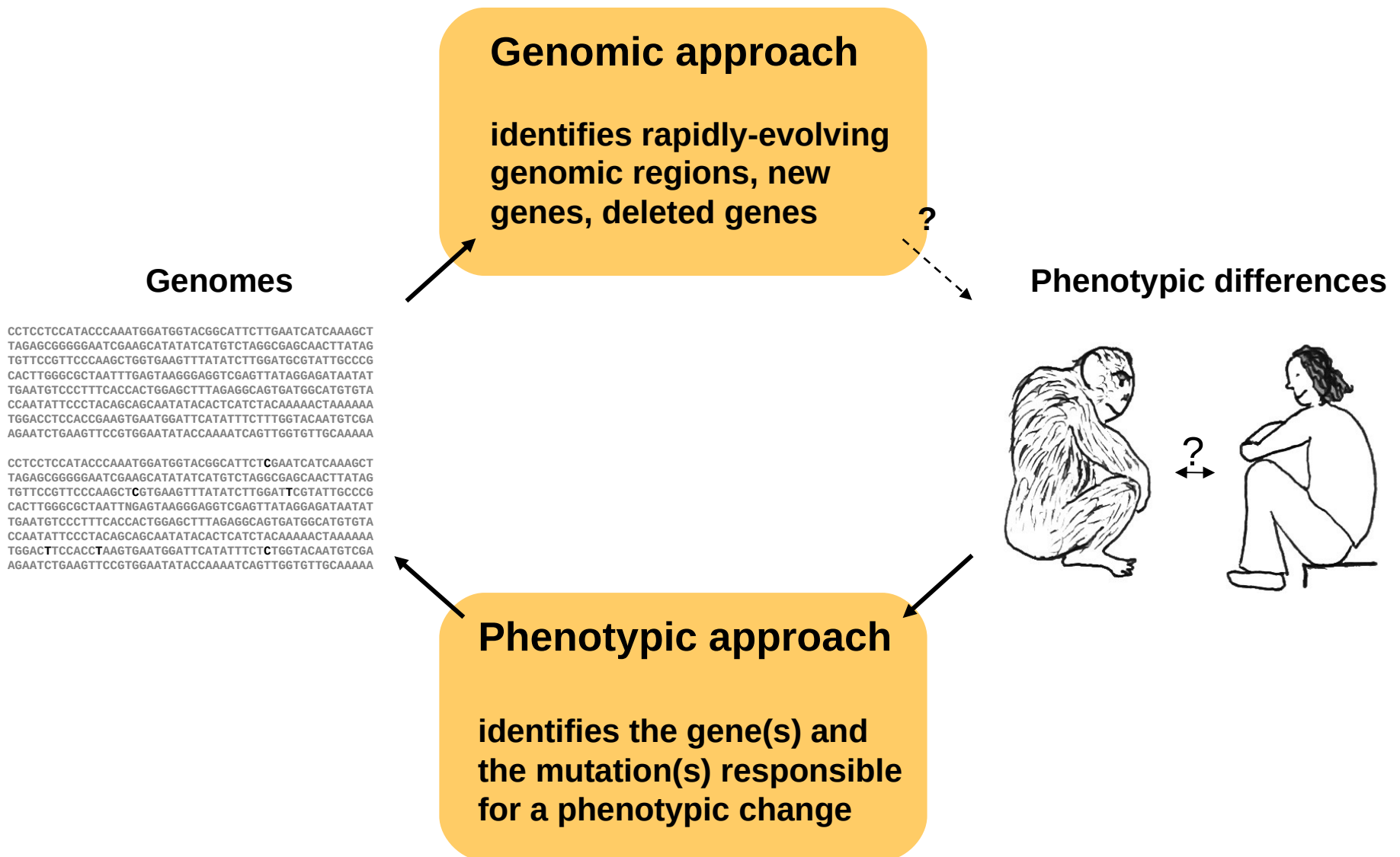
3) Genetic hotspots of evolution

1) Identifying the mutations responsible for phenotypic differences

2) Gephé : thinking in terms of differences

3) Genetic hotspots of evolution

What are the mutations responsible for phenotypic differences?



Pelvic reduction in sticklebacks



marine



Paxton Lake, Canada

Gasterosteus aculeatus

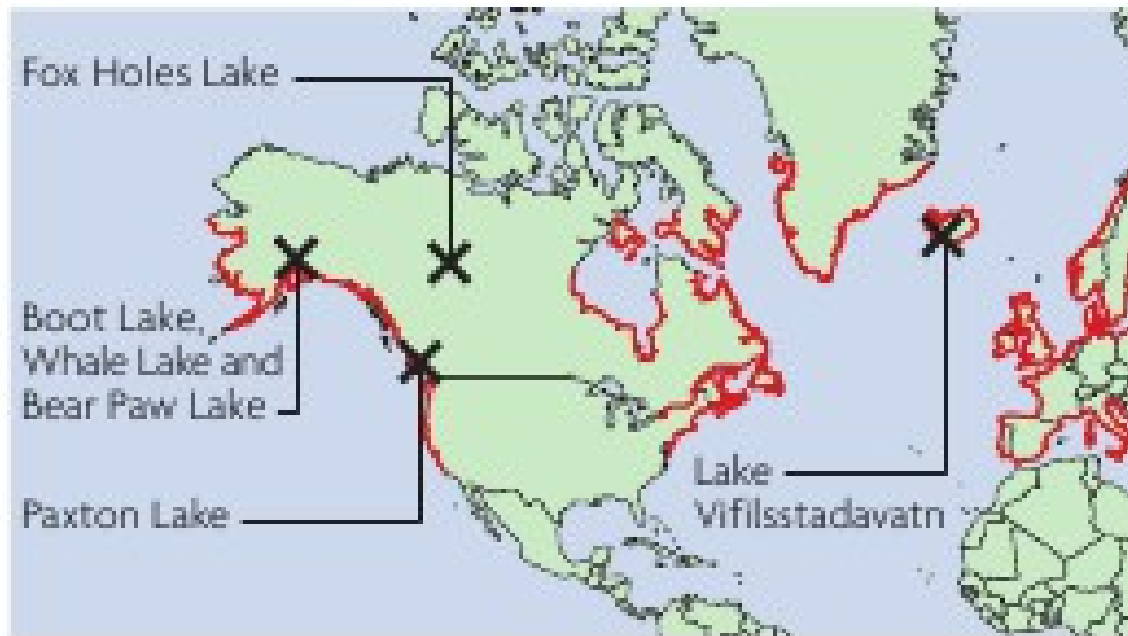
(Peichel et al., 2001 ; Shapiro et al, 2004 ; Chan et al. 2010)

Marine fishes with robust pelvis = ancestral

**Freshwater fishes with reduced pelvic structures = derived,
independently at least 20 times**

- limited calcium availability
- absence of gape-limited predatory fishes
- predation by grasping insects

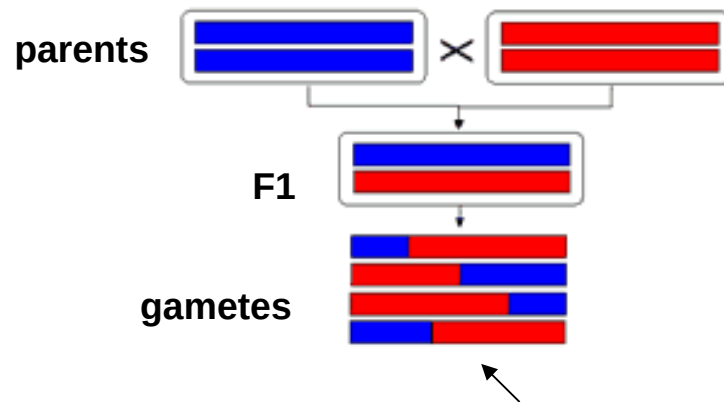
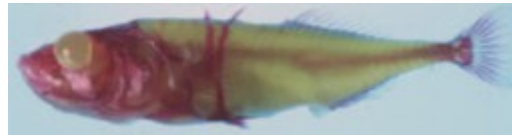
Last glacier retreat = 10 000 – 20 000 years ago



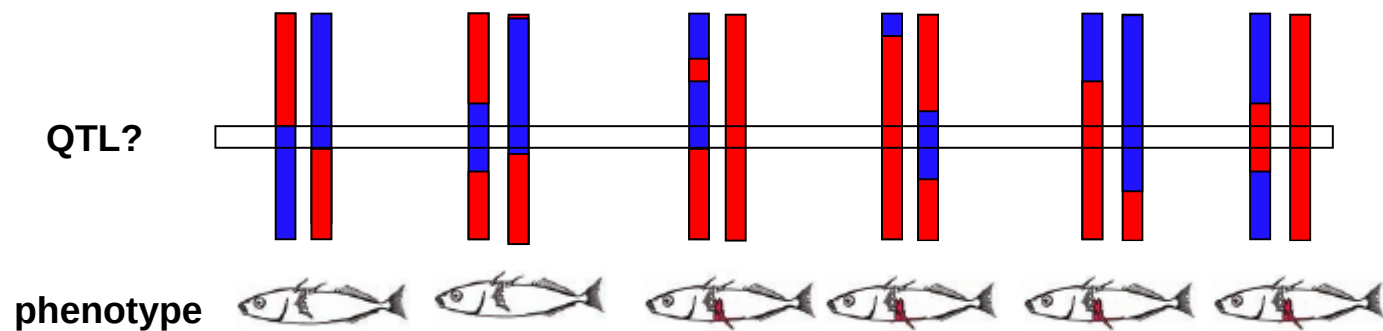
QTL mapping

lake

marine

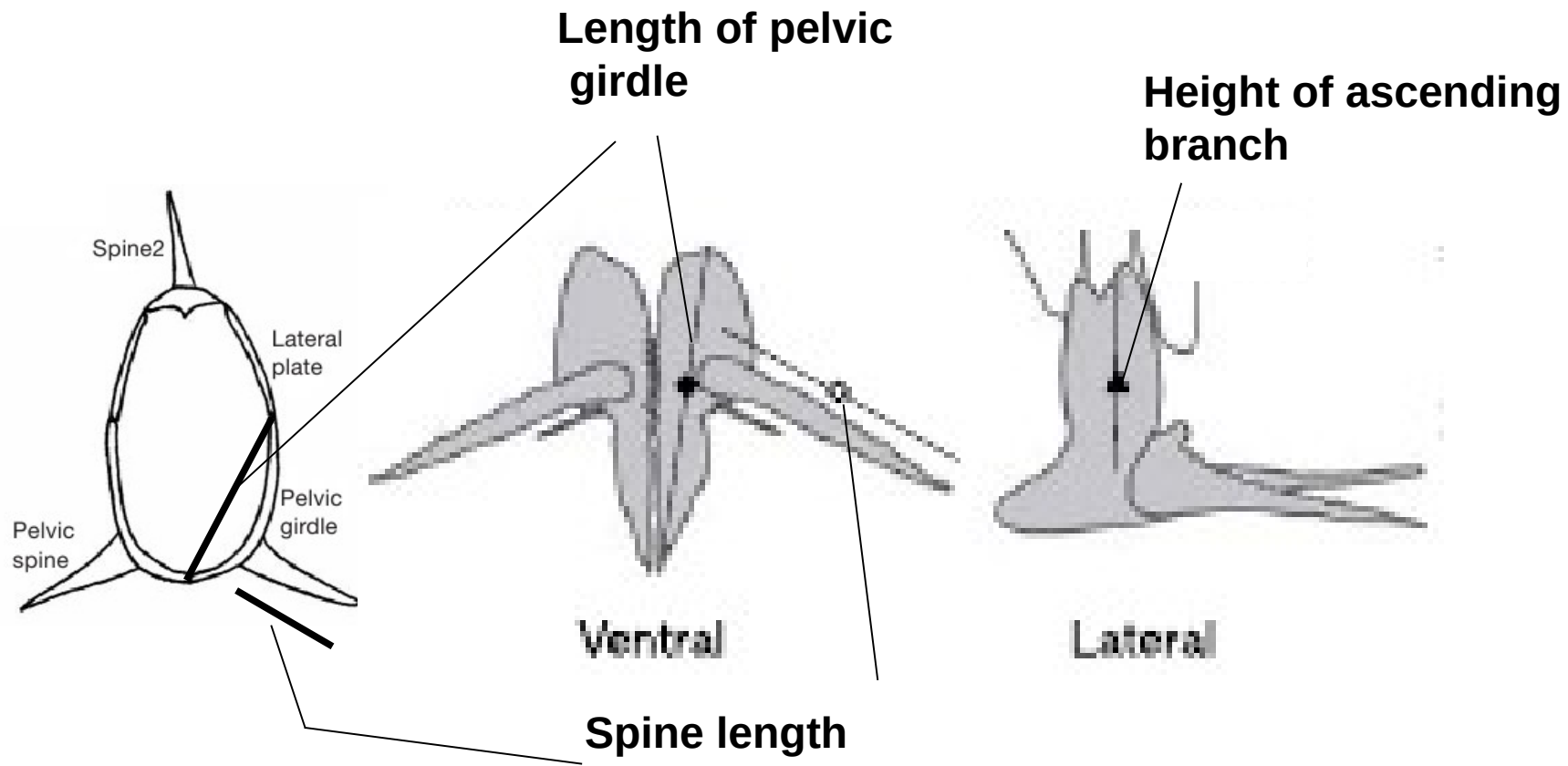


375 F2 individuals



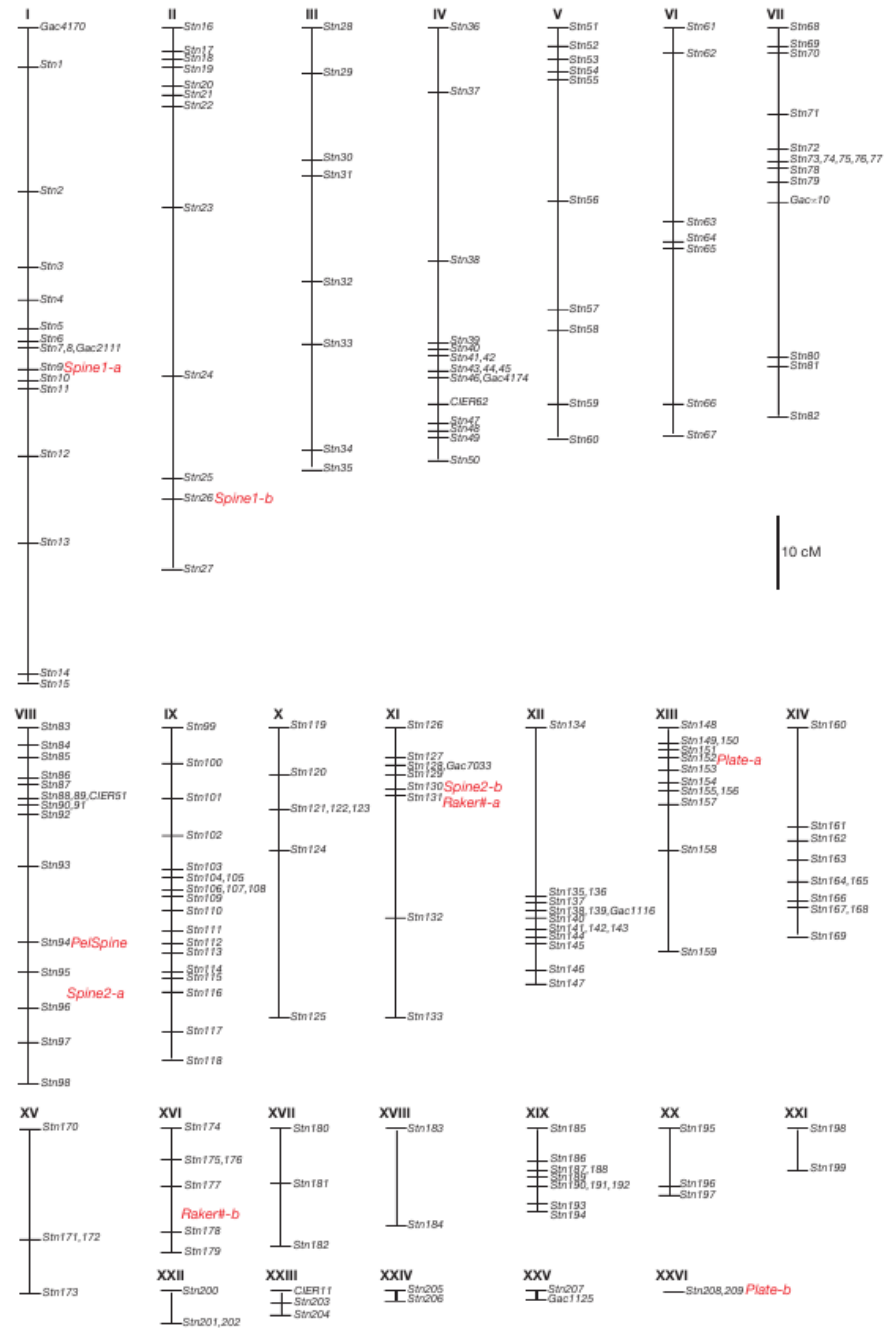
(Shapiro et al., 2004)

Quantitative measurement of the phenotype



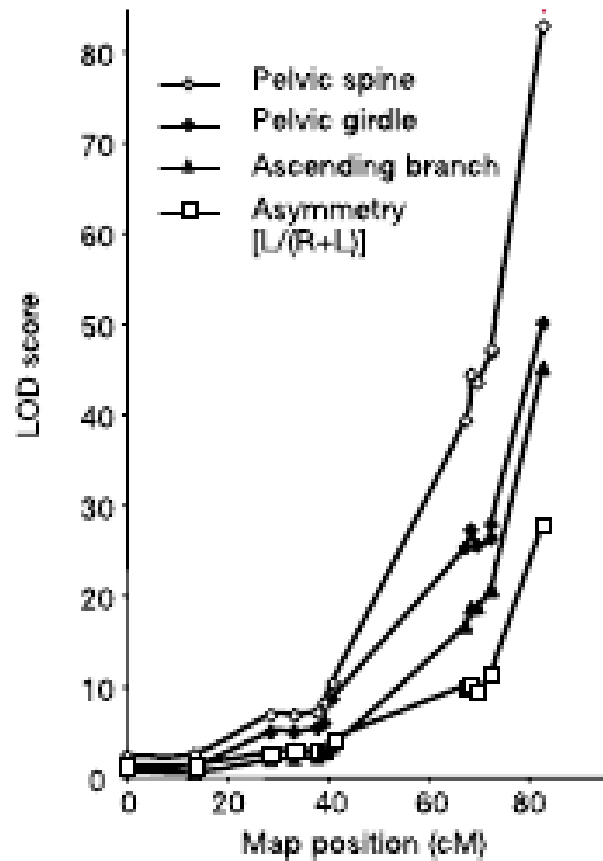
1000 microsatellite markers

26 linkage groups



(Peichel et al., 2001)

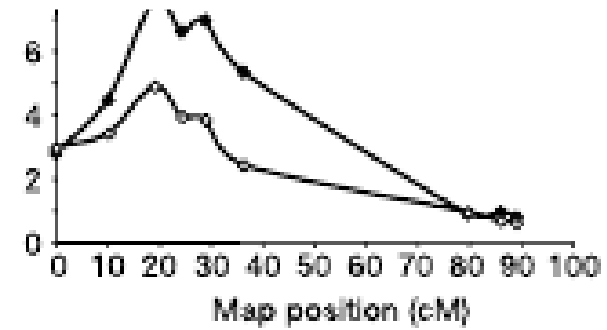
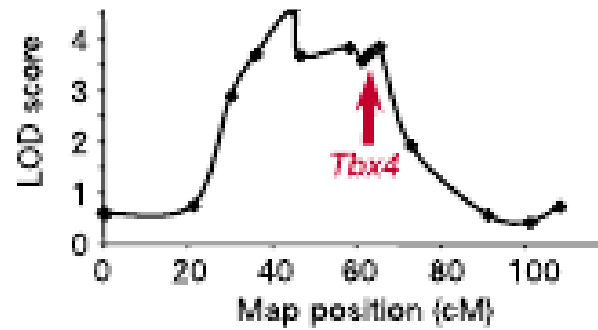
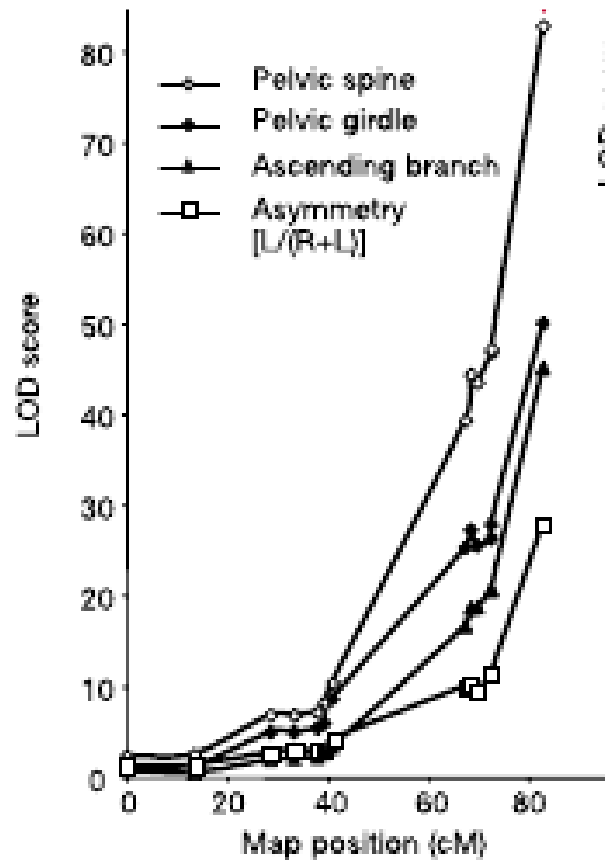
One major locus at the end of linkage group 7



Major locus responsible for 65% of the variance

One major locus at the end of linkage group 7

A few minor loci



Major locus responsible for 65% of the variance

Three candidate genes: Pitx1, Pitx2, Tbx4



Screen of a BAC library (Bacterial Artificial Chromosomes)
containing 100-350-kb fragments of stickleback genomic DNA



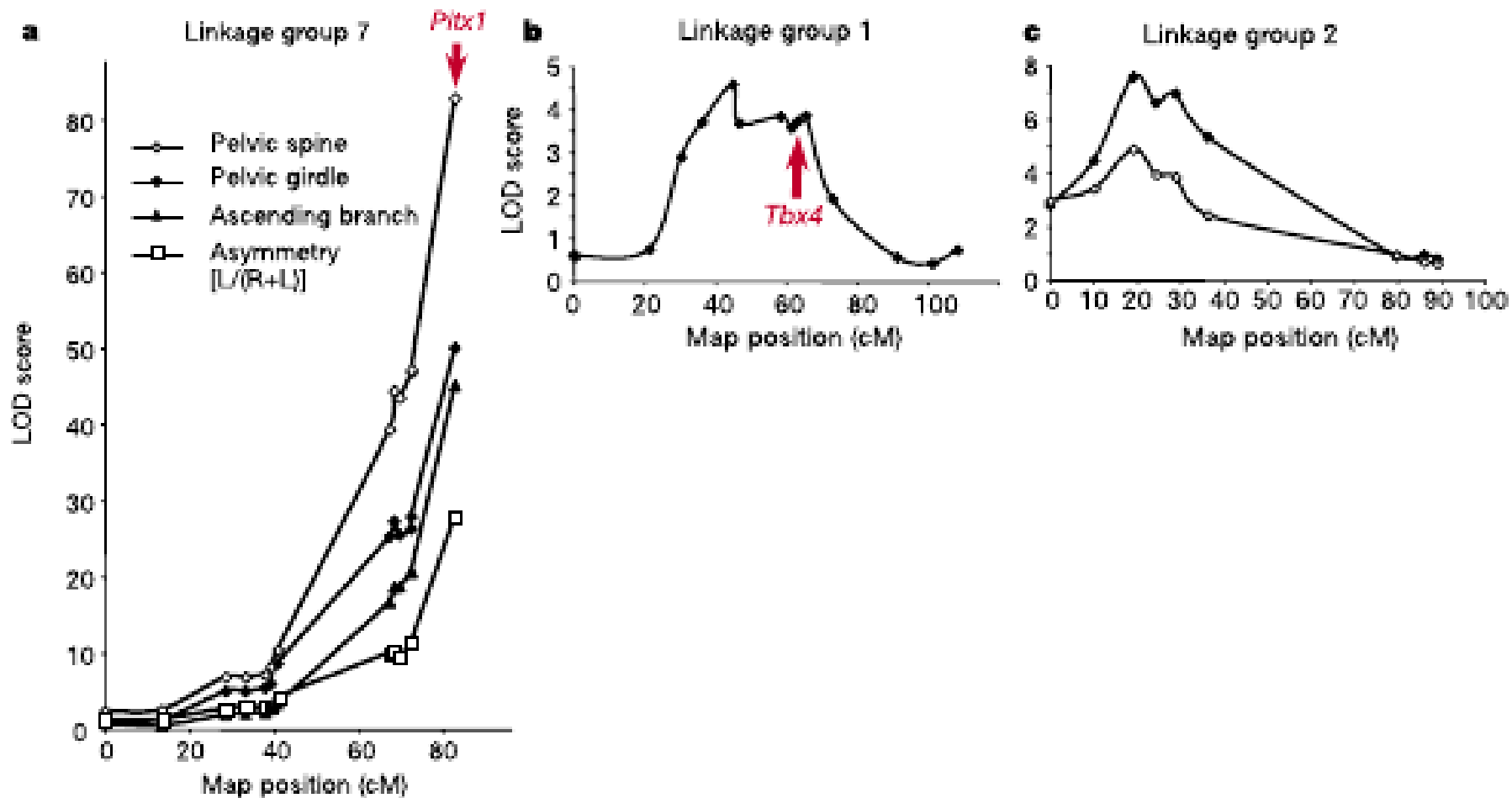
Partial sequencing and generation of new markers



Genotyping of 375 F2 individuals

One major locus at the end of linkage group 7

A few minor loci



***Pitx1*, responsible for the phenotypic change?**

***Pitx1* null mutations in mice
(pelvis reduction, stronger on right side)**

QTL mapping

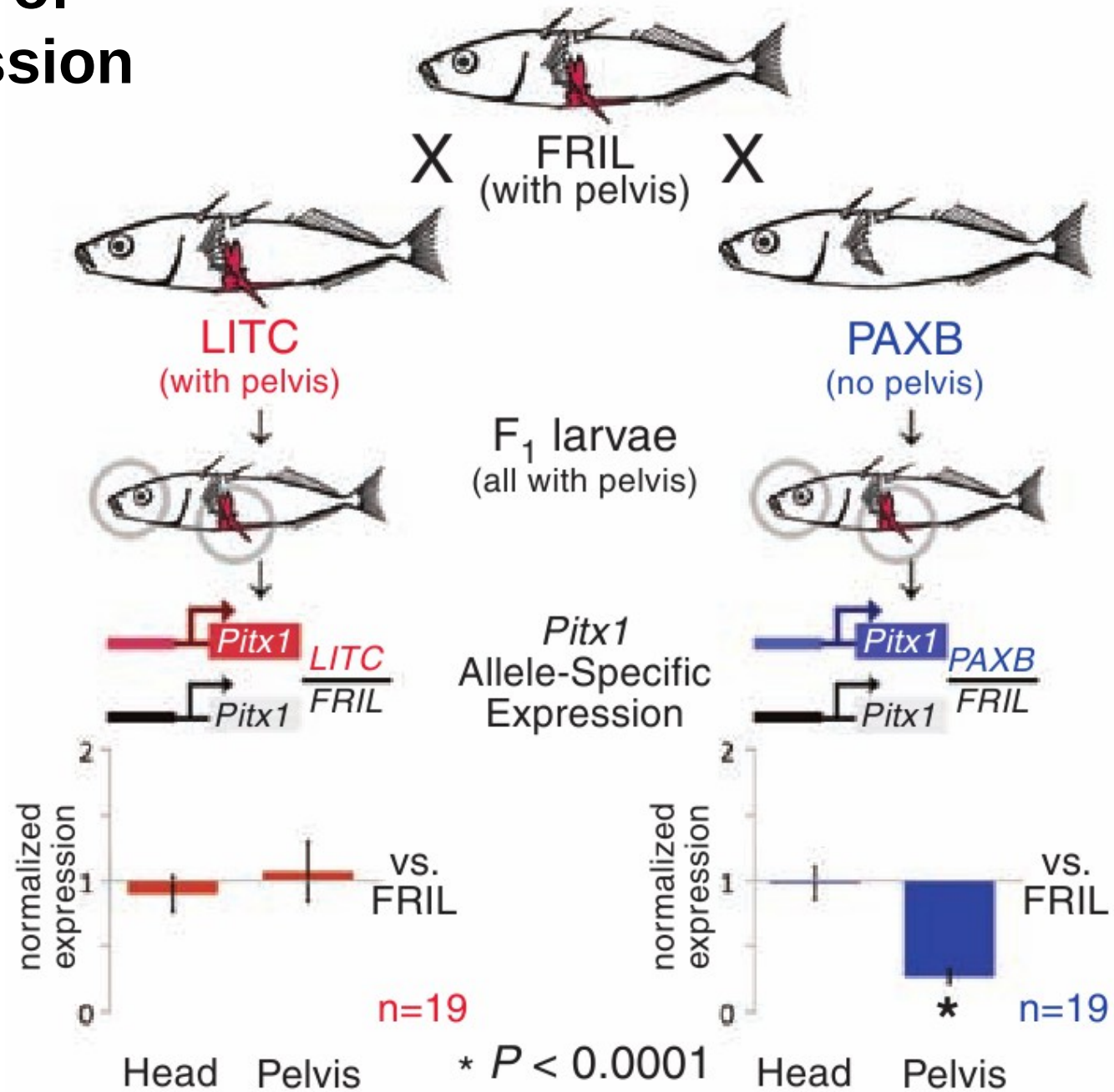
Same coding sequence in lake and marine forms

***Pitx1* expressed at stage 29 in marine individuals but not in
lake individuals**

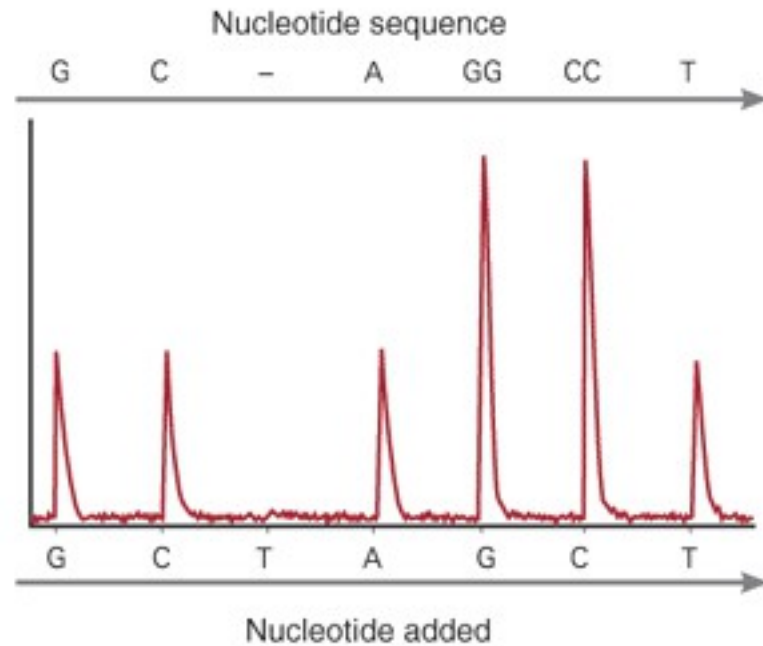
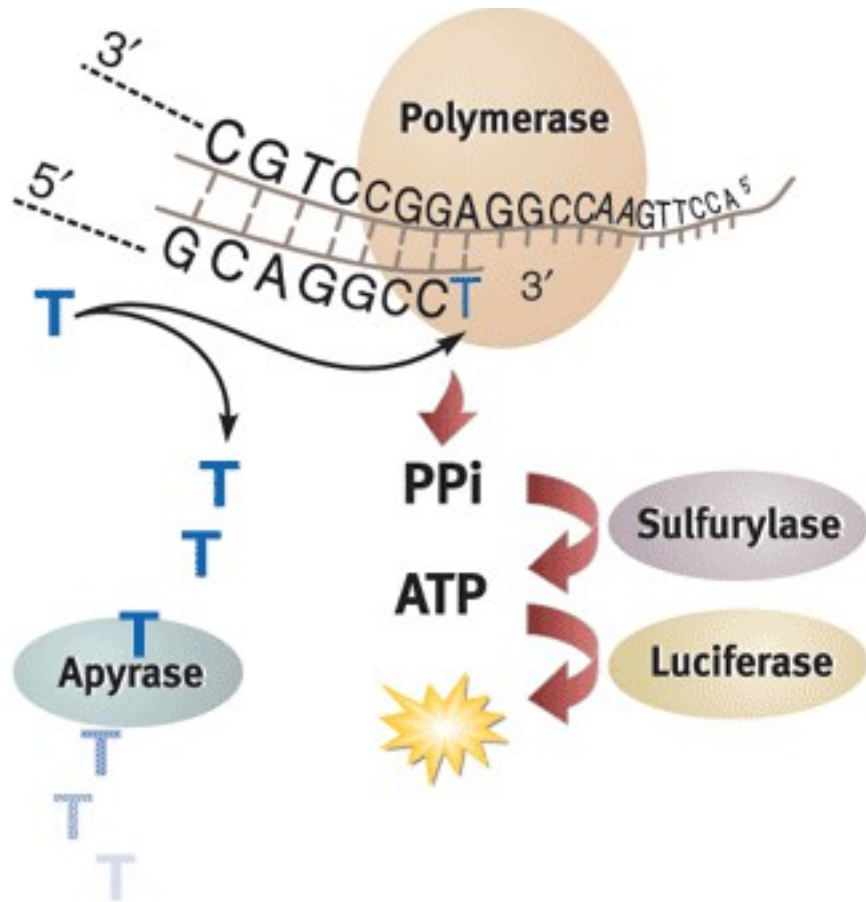
BUT

**The decrease in *Pitx1* expression levels
might have evolved due to mutations in
an upstream regulatory gene**

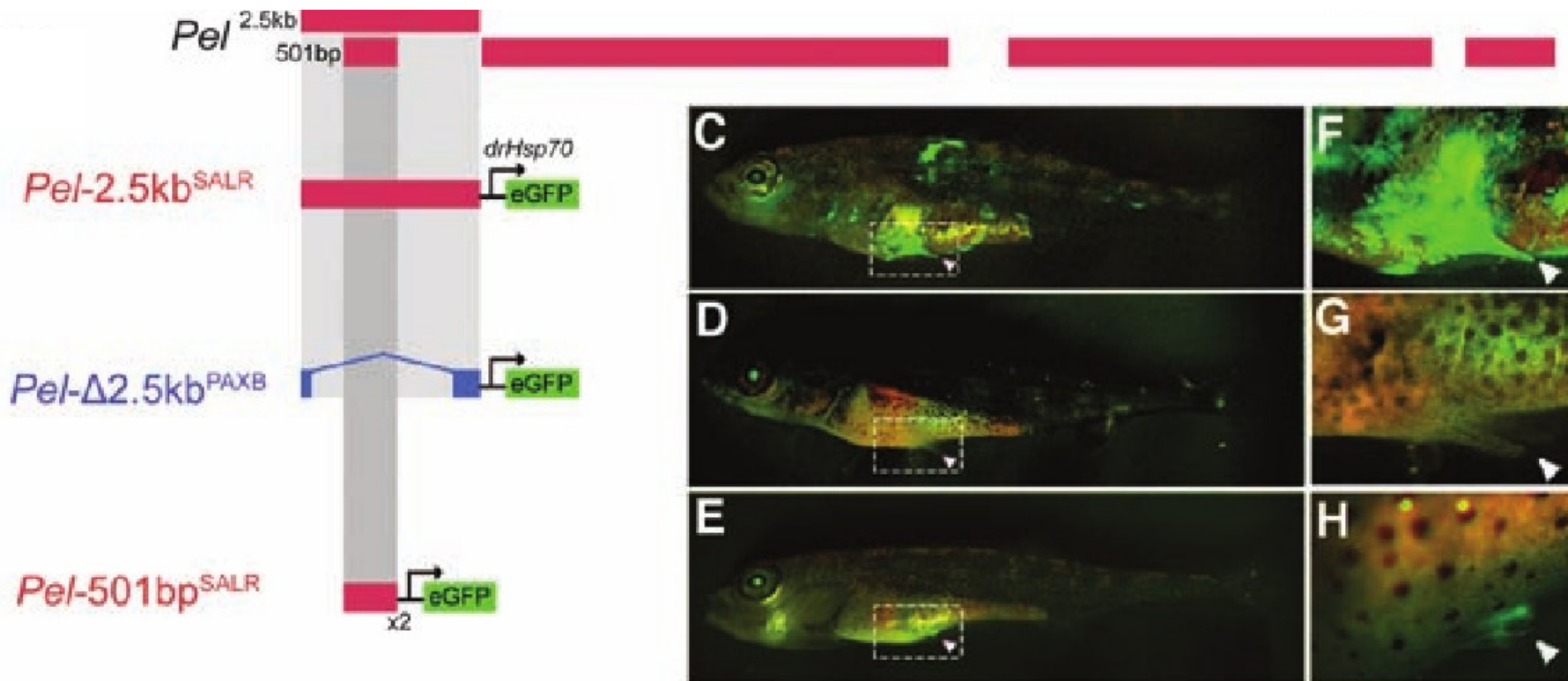
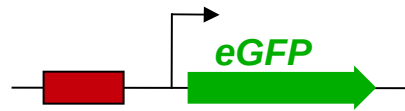
Comparison of allele expression in hybrids



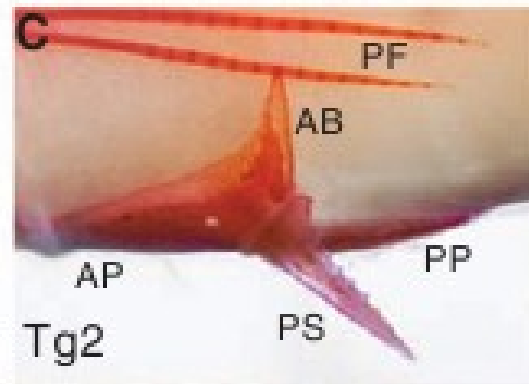
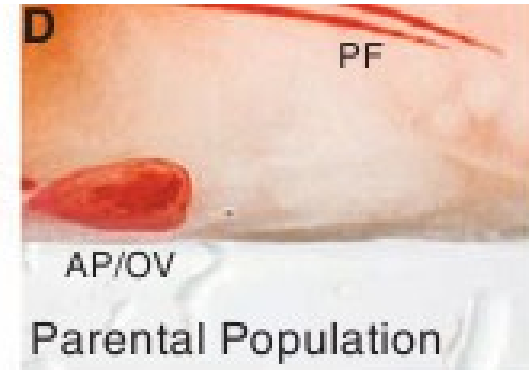
What is pyrosequencing?



Test of *Pitx1* cis-regulatory regions



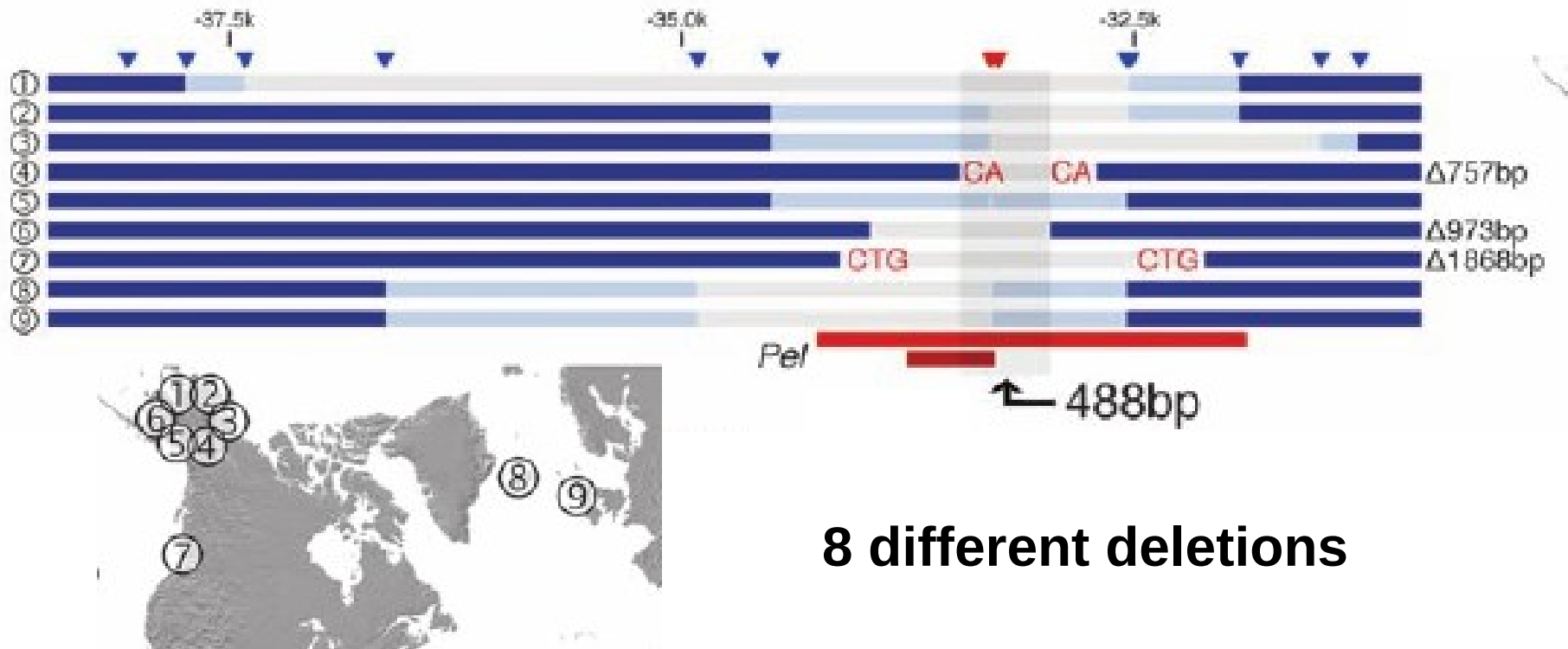
Rescue of a pelvis in freshwater individuals



Several independent deletions in the cis-regulatory region of *Pitx1*

Region sequenced in two lake pop[ulations]: a 2-kb deletion in one and a 757-bp deletion in the other one

SNP genotyping in 13 populations with reduced pelvis and in 21 populations with complete pelvis



8 different deletions

How a *Drosophila* species adapted to its cactus host





Picture: E. Matzkin



senita cactus

Drosophila pachea





**toxic
alcaloids**

resistance

senita cactus

Drosophila pachea





**toxic
alcaloids**

resistance

dependance

senita cactus

Drosophila pachea



How did *D. pachea* become dependent on a cactus ?

generalist



specialist



obligate specialist





**toxic
alcaloids**

resistance

Drosophila pachea

Δ^7 -sterols

dependance

senita cactus



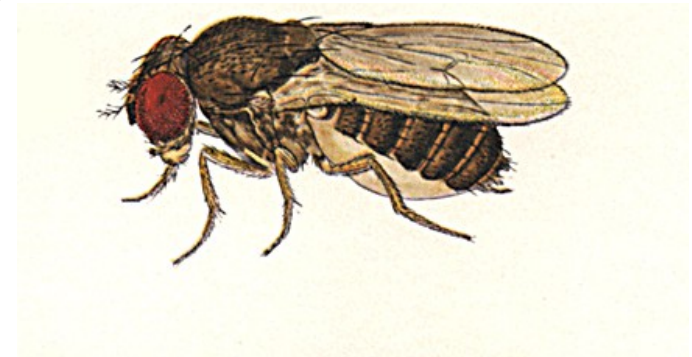
Picture: L. Mazkin

Ineluctable degradation of metabolic activities during evolution

arginine
histidine
isoleucine
leucine
lysine
methionine
phenylalanine
threonine
valine
ascorbic acid (vitamin C)
biotin (vitamin H)
folic acid (vitamin M)
riboflavin (vitamin B2)
thiamine (vitamin B1)
cobalamin (vitamin B12)



Δ^7 -sterols



Heed 1965 Science

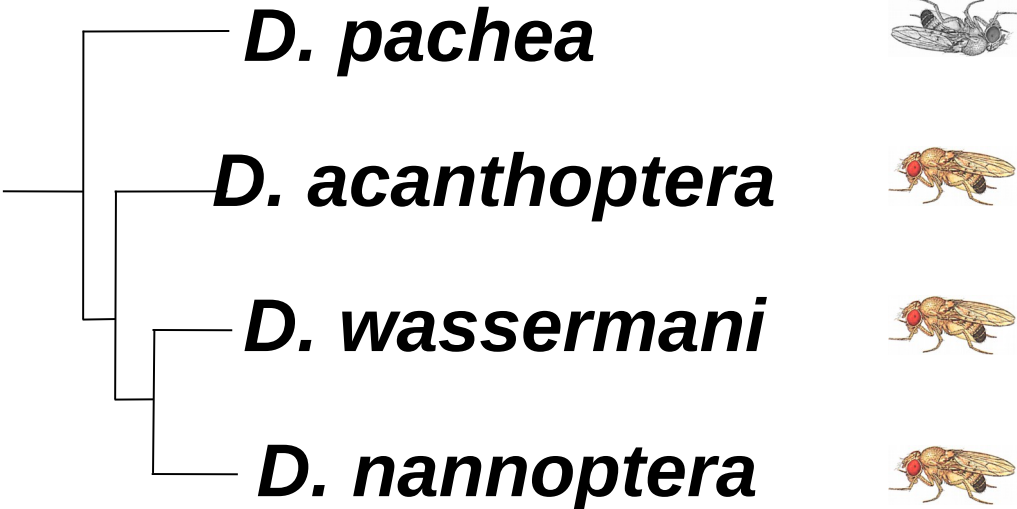
Romero 2005 Genome Biology

Why have these metabolic activities been lost?

***< selective advantage?
neutral?***

The *nannoptera* group

on normal food



(Ward and Heed, 1970)

How did *D. pachea* become dependent on a cactus ?



Drosophila pachea

Δ^7 -sterols

cactus senita

dependance



D. melanogaster



cholesterol

D. pachea



cholesterol



steroid hormone
(ecdysone)

D. melanogaster



cholesterol



7-dehydro-
cholesterol



steroid hormone
(ecdysone)

D. pachea



cholesterol



7-dehydro-
cholesterol

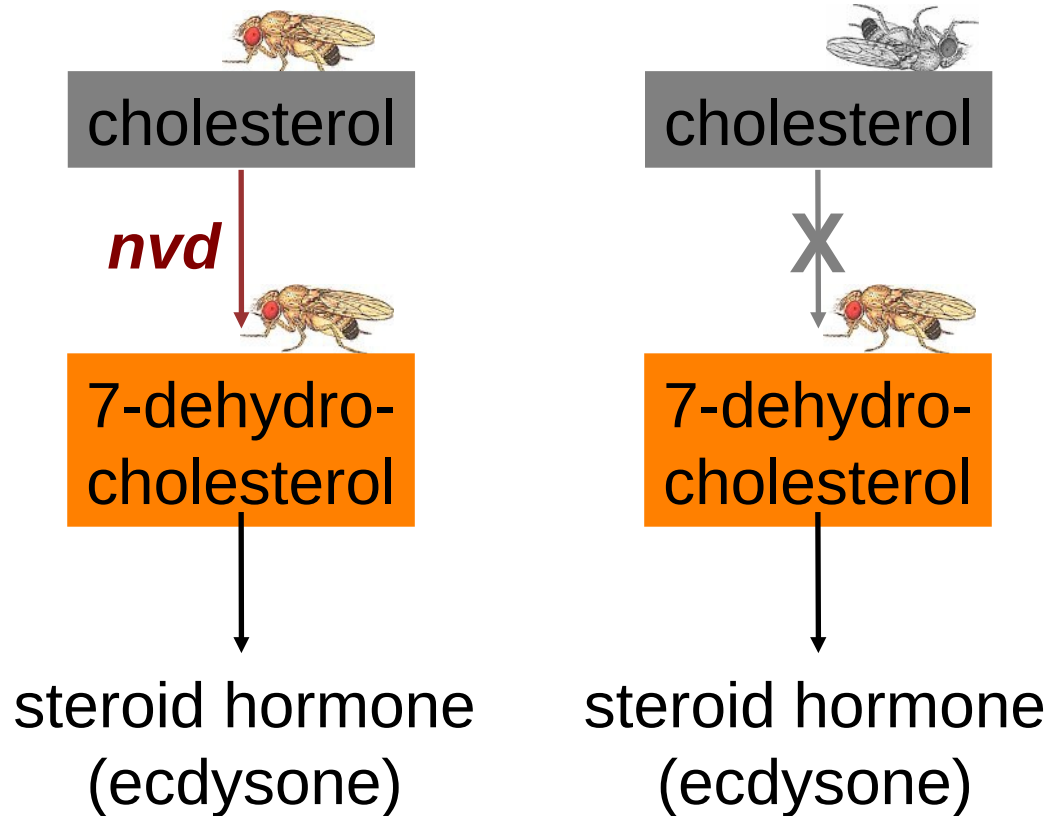


steroid hormone
(ecdysone)

(Clayton, 1964 ;
Heed and Kircher, 1965;
Chu et al., 1970;
Warren et al., 2001)

D. melanogaster

D. pachea



**Mutation(s) in
the *nvd* gene in *D. pachea* ?**

(Yoshiyama et al., 2006)

Several important amino acid changes in *D. pachea* NVD

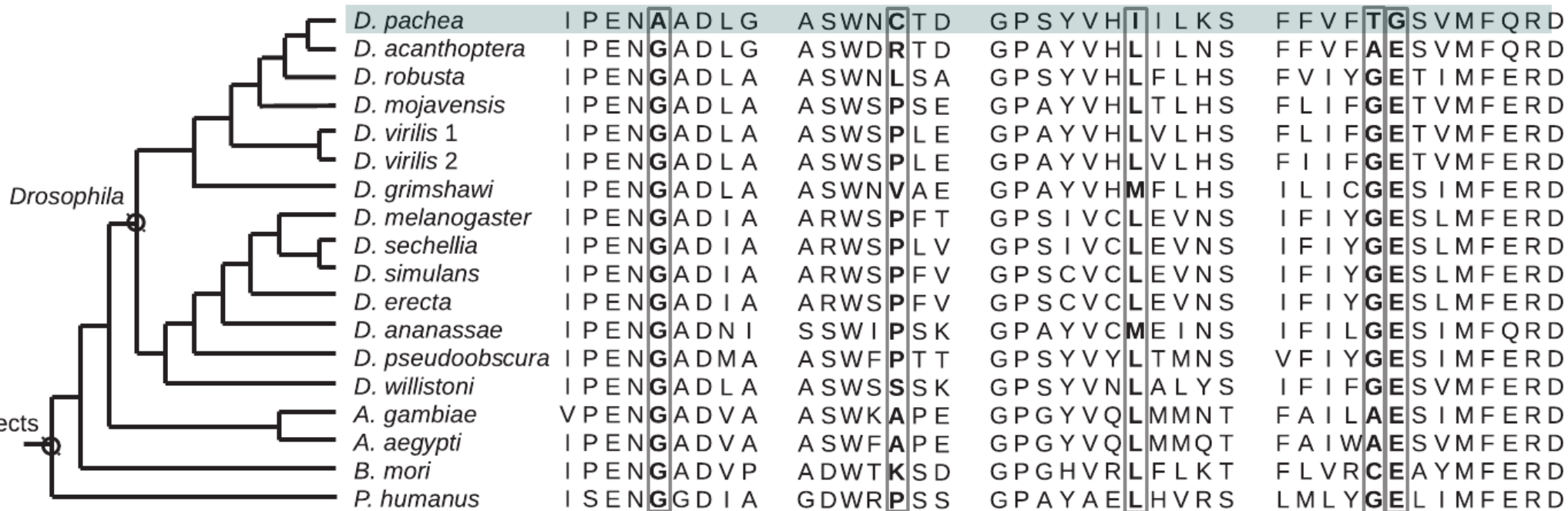


G250A

R290C

L330I

**A376T
E377G**



E → G

polar, acidic → non polar, neutral

In *D. pachea* *nvd* is still expressed in prothoracic glands

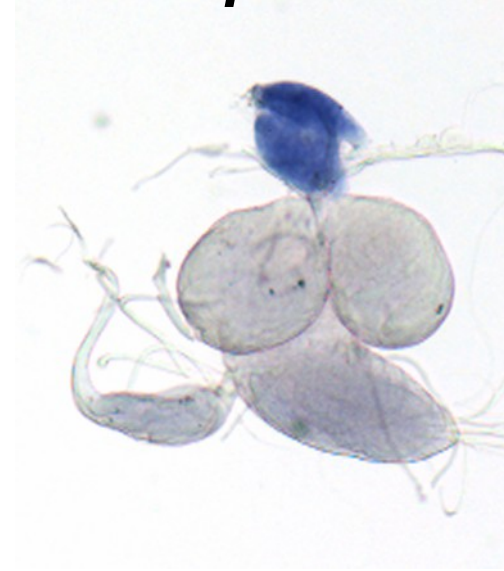
D. melanogaster



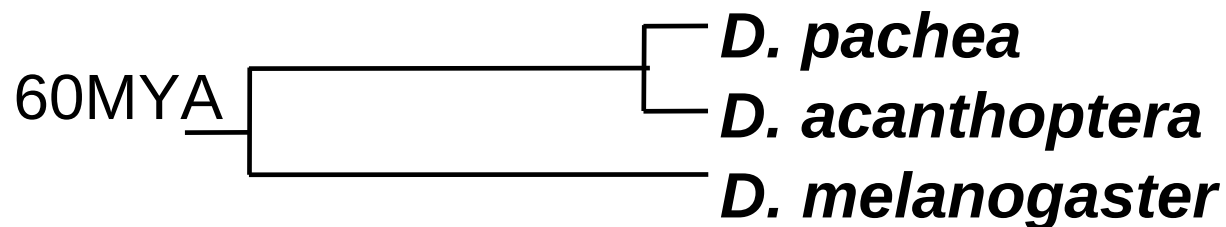
D. acanthoptera



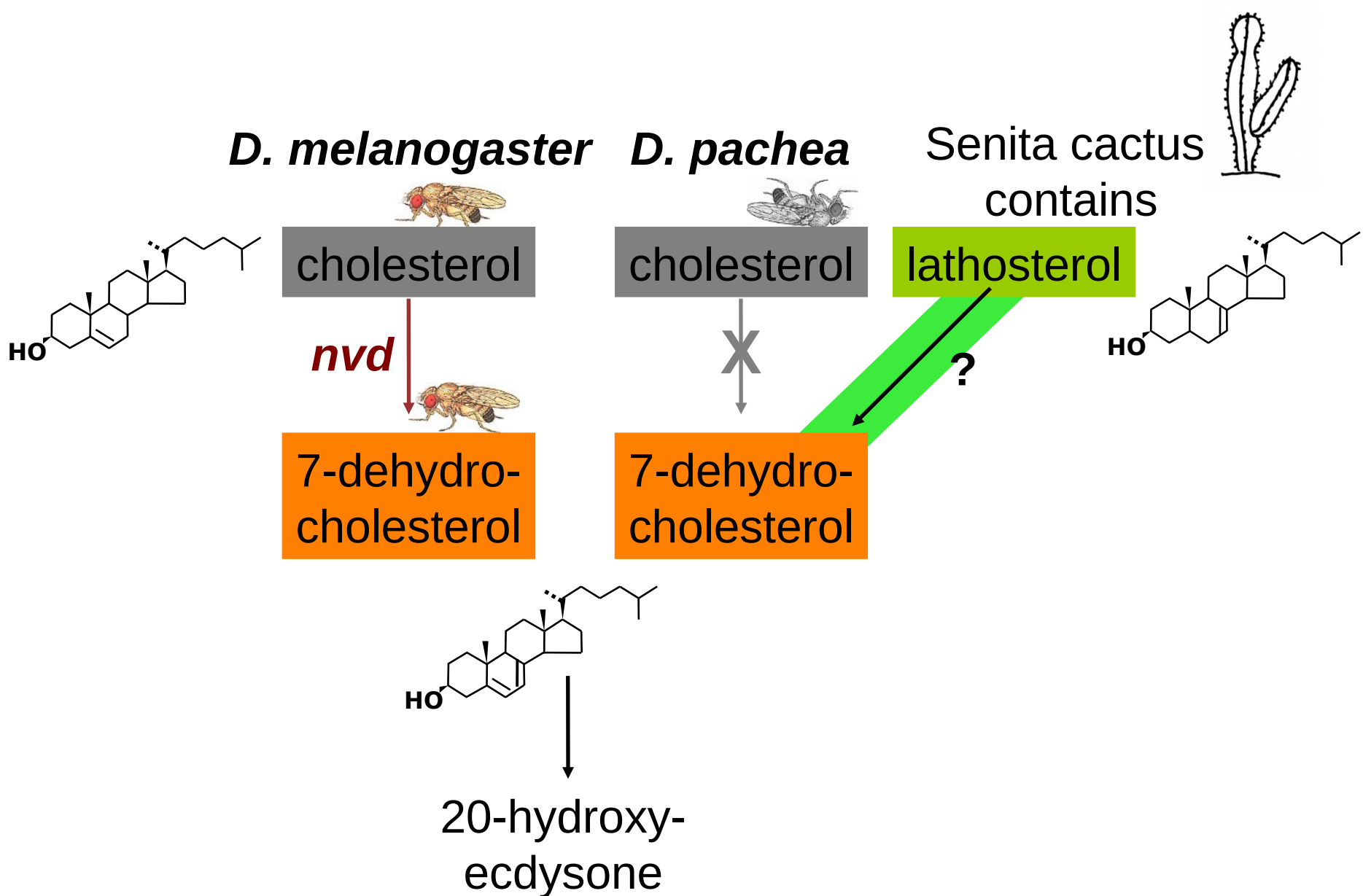
D. pachea



100μm



Novel metabolic pathway in *D. pachea*?



In vivo tests in *D. melanogaster*

Fly
survival

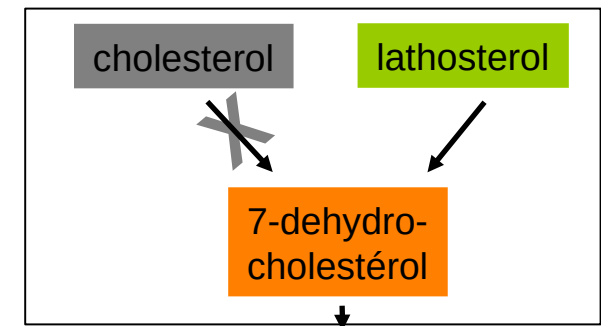
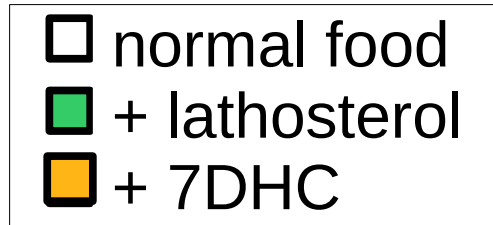
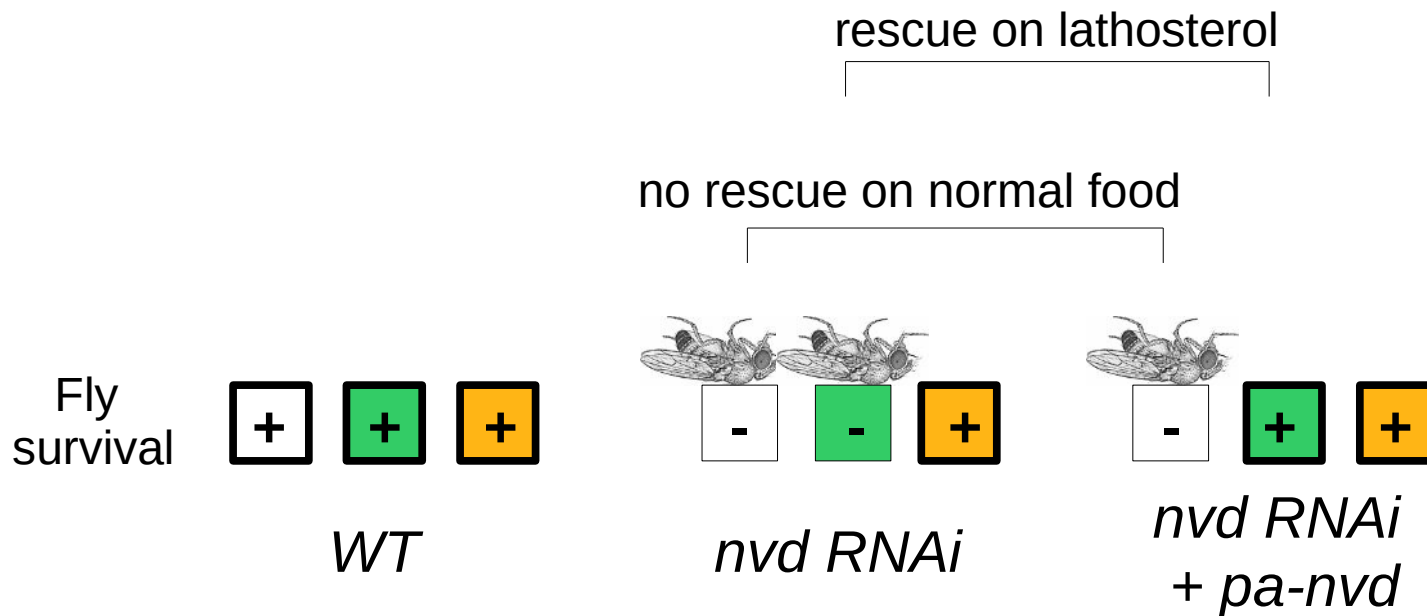
WT

nvd RNAi

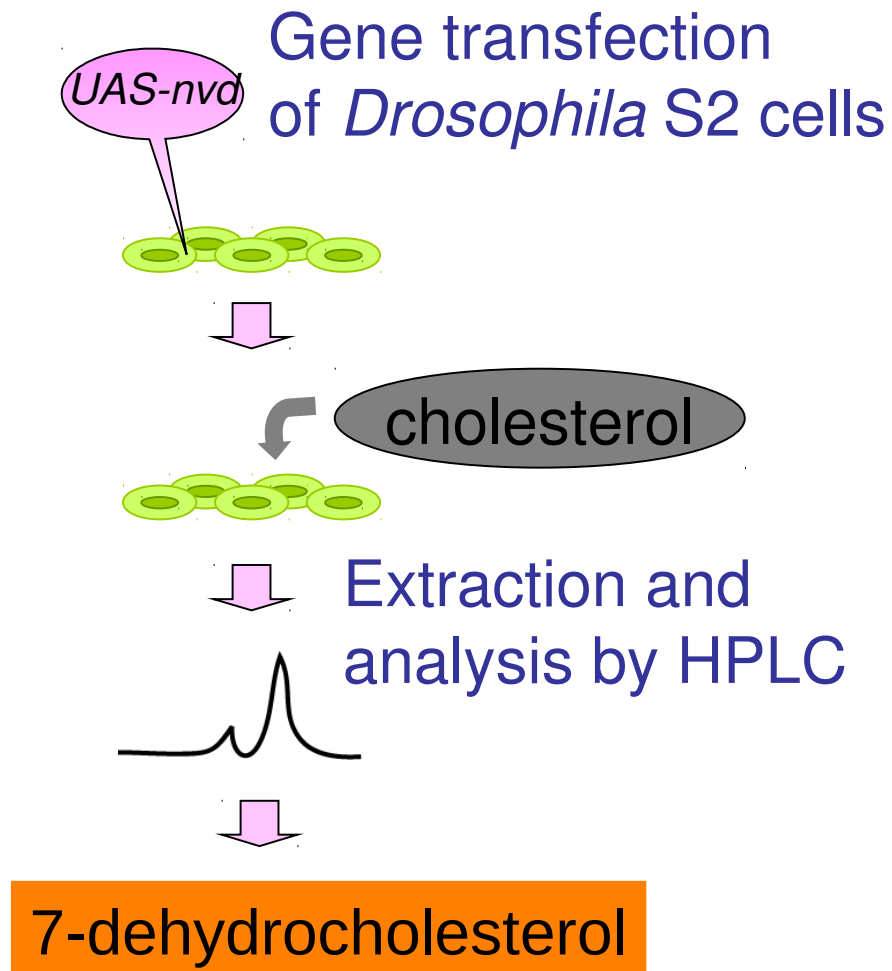
nvd RNAi
+ *pa-nvd*

- normal food
- + lathosterol
- + 7DHC

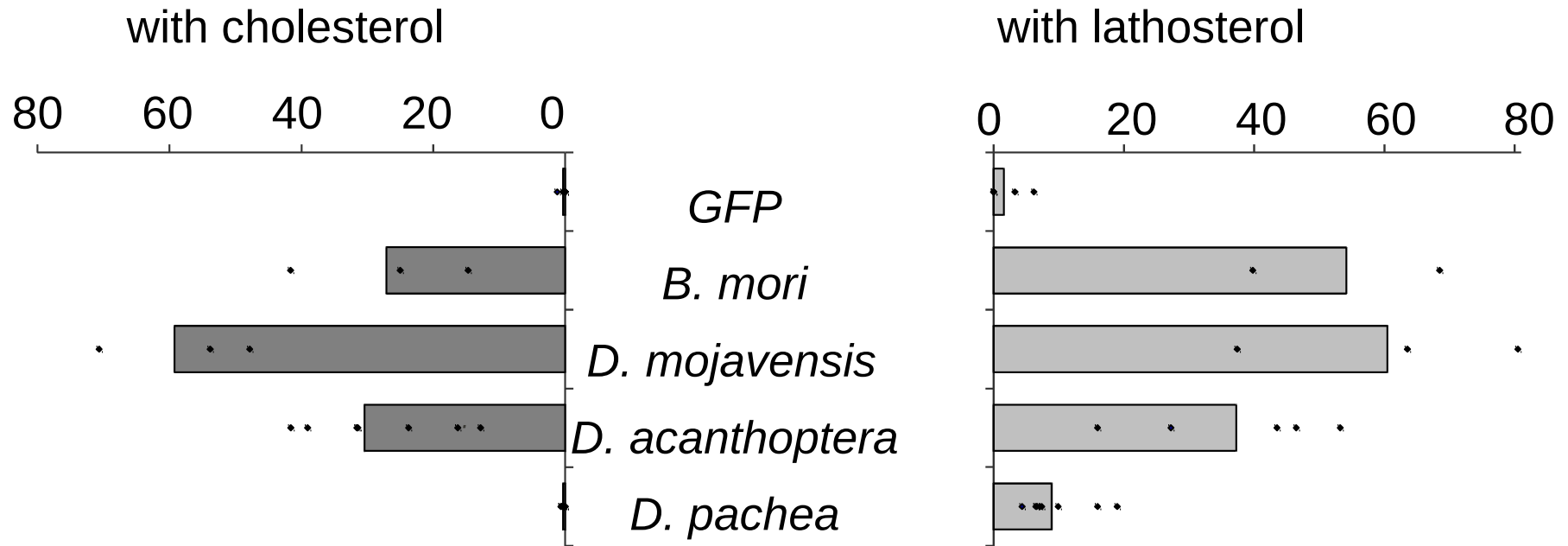
D. pachea NVD converts lathosterol and not cholesterol

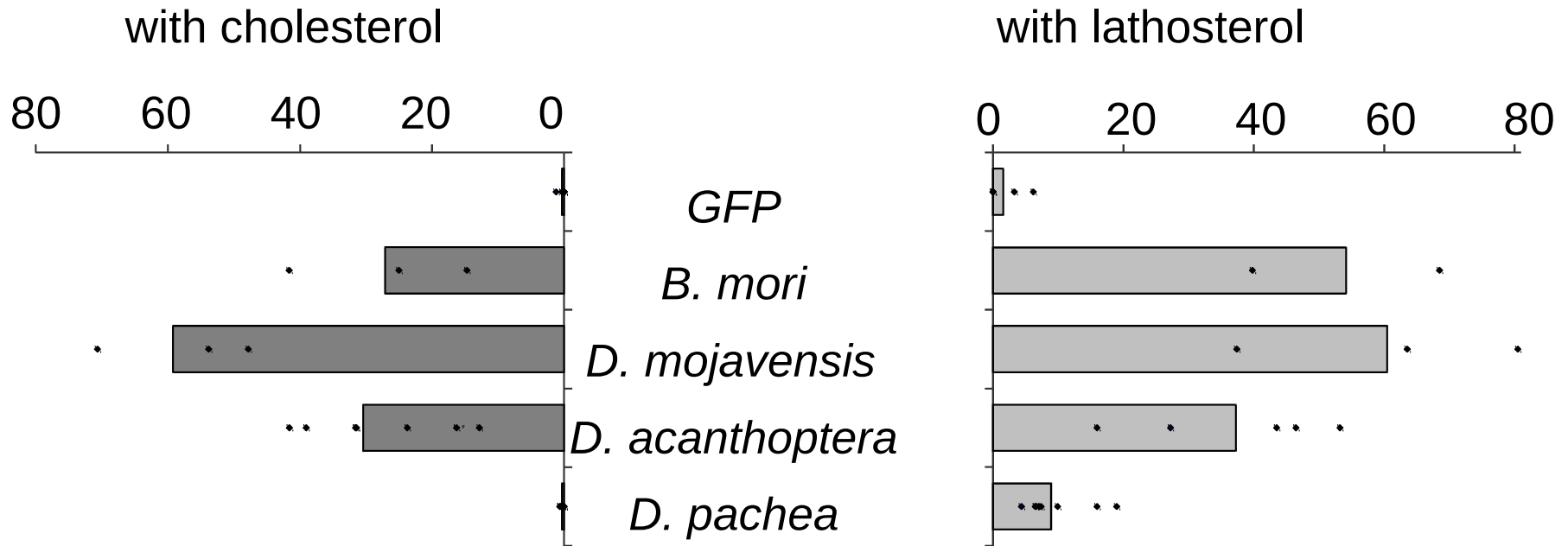


In vitro assay of Nvd activity



As expected, *D. pachea* NVD converts lathosterol but not cholesterol

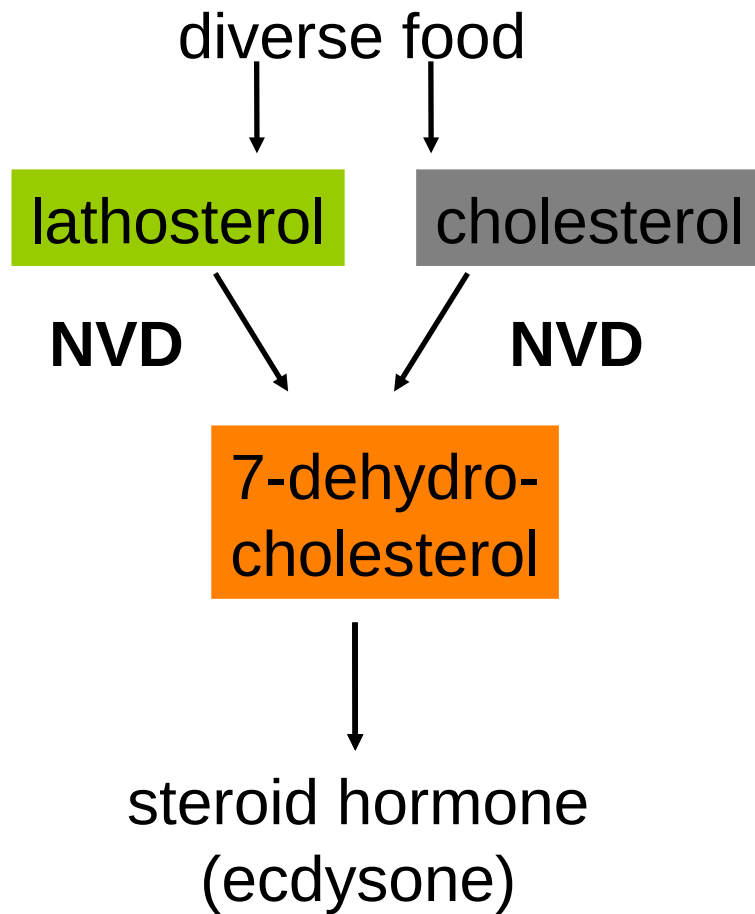




Unexpectedly, all the tested enzymes convert lathosterol

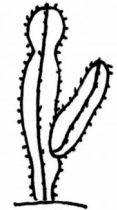
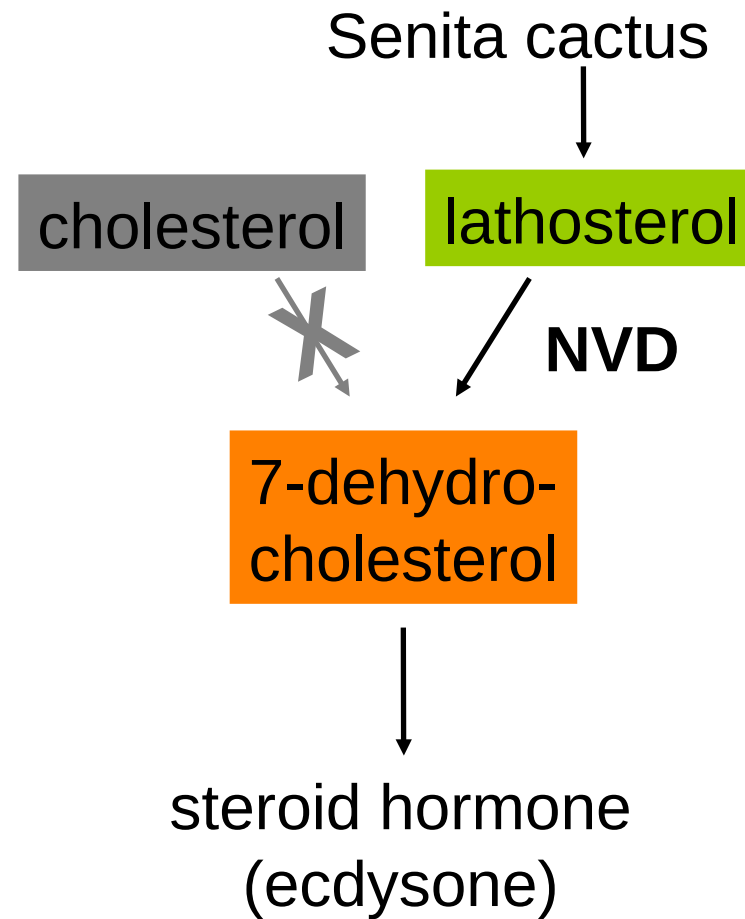
ANCESTRAL STATE

D. mojavensis and
D. acanthoptera

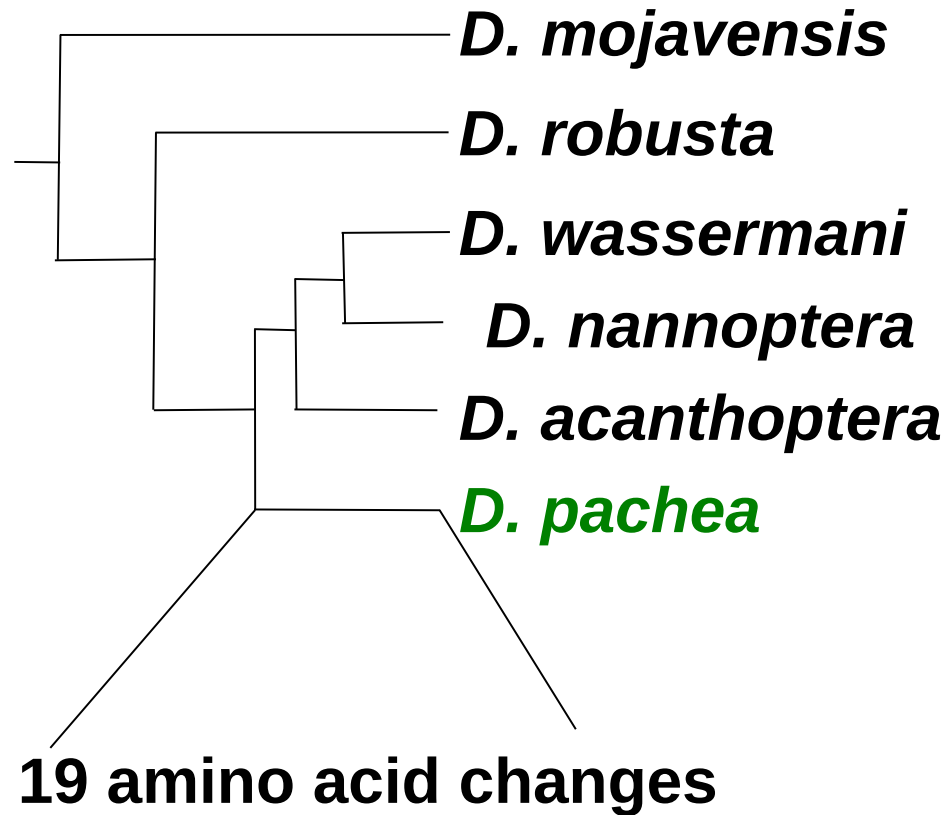


DERIVED STATE

D. pachea



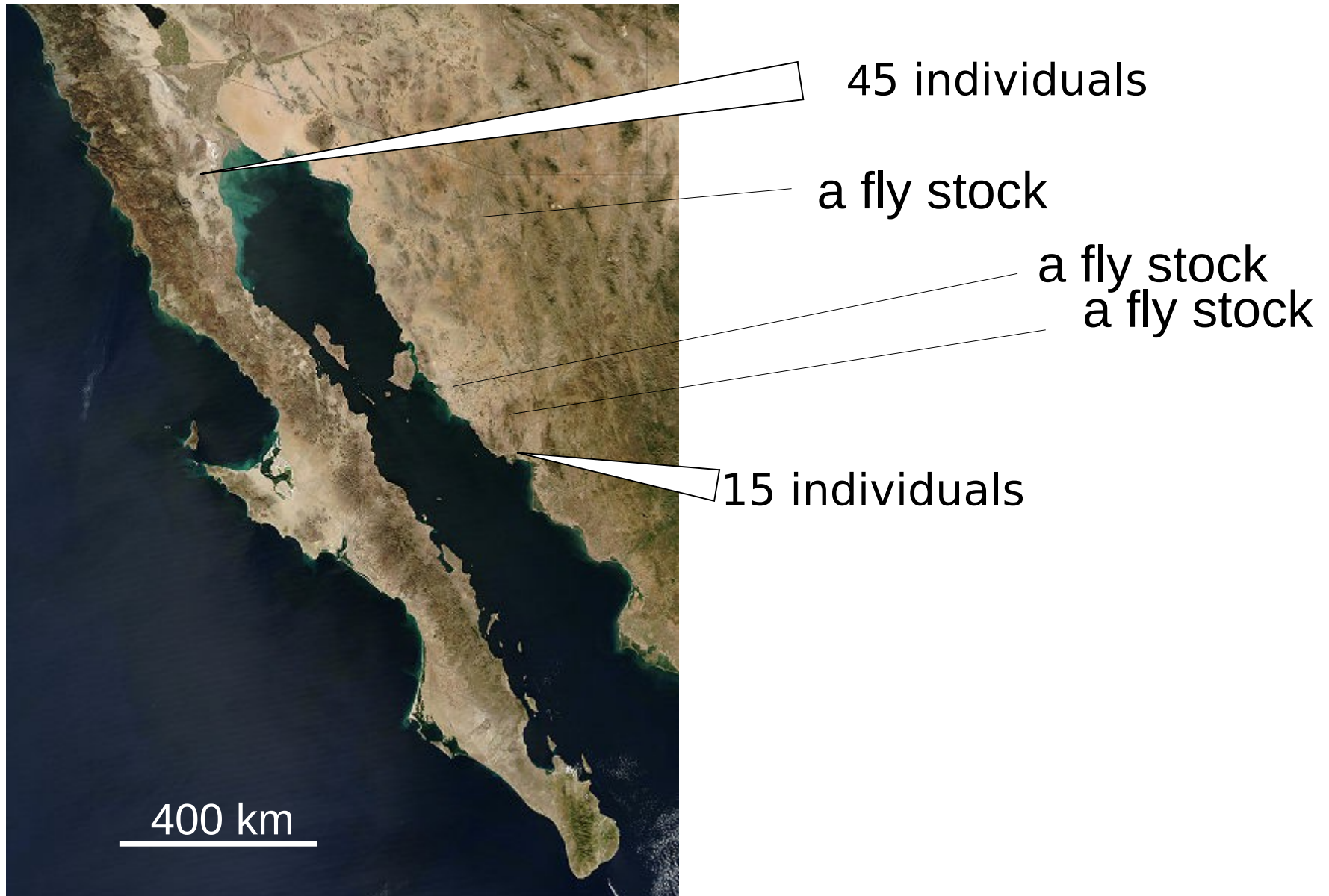
Which amino acid changes?



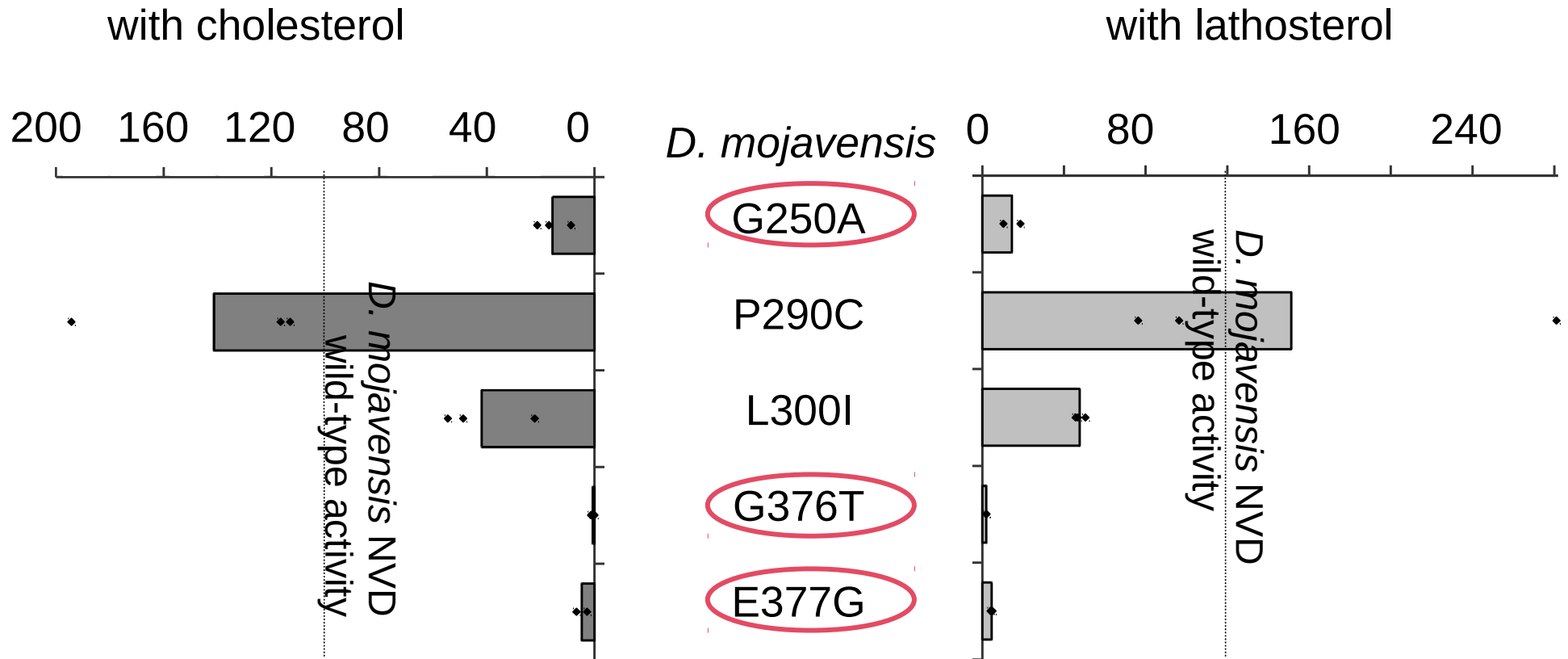
**5 are predicted to
be deleterious**

SIFT (Ng et al. 2003)

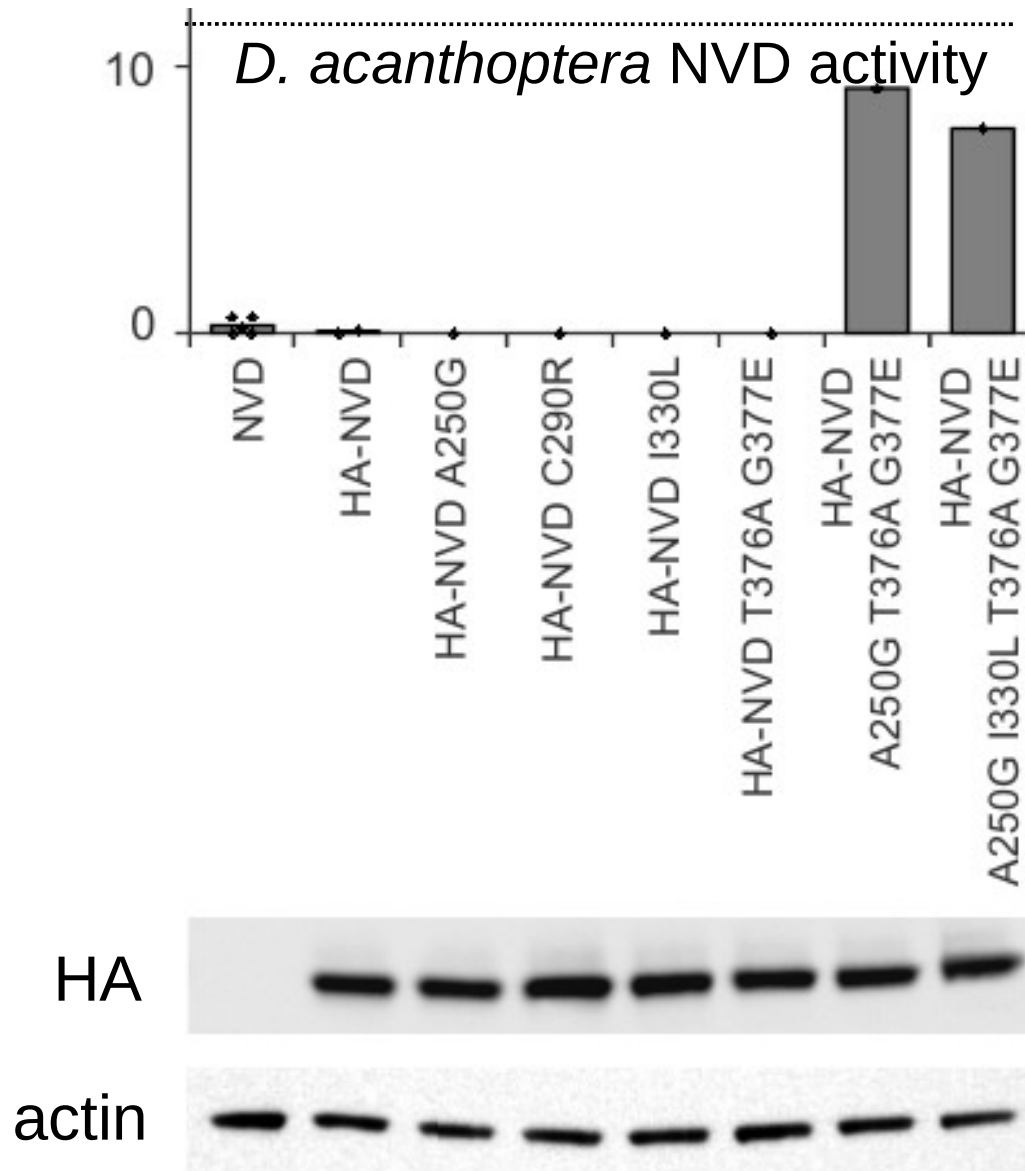
These five amino acids are present in all the sampled *D. pachea* individuals



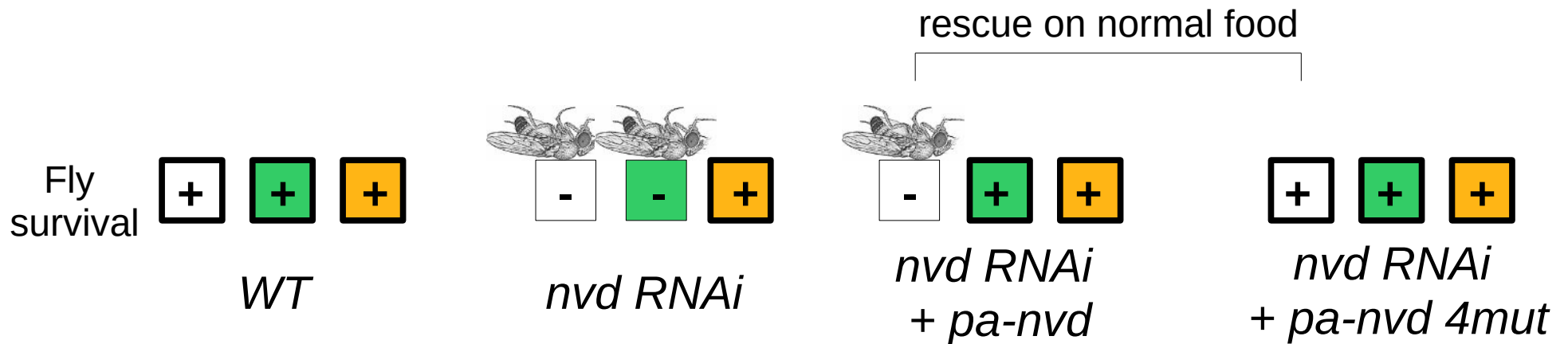
Three mutations reduce NVD activity to less than 15%



At least 2 mutations are required to restore NVD activity on cholesterol

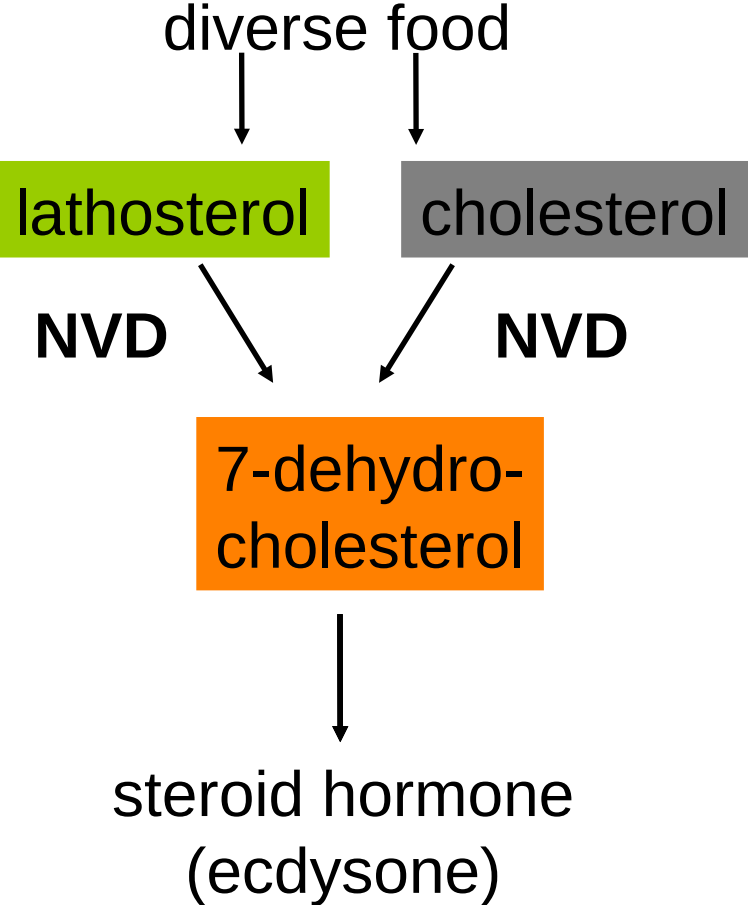


Reverting 4 amino acids restores NVD activity in vivo

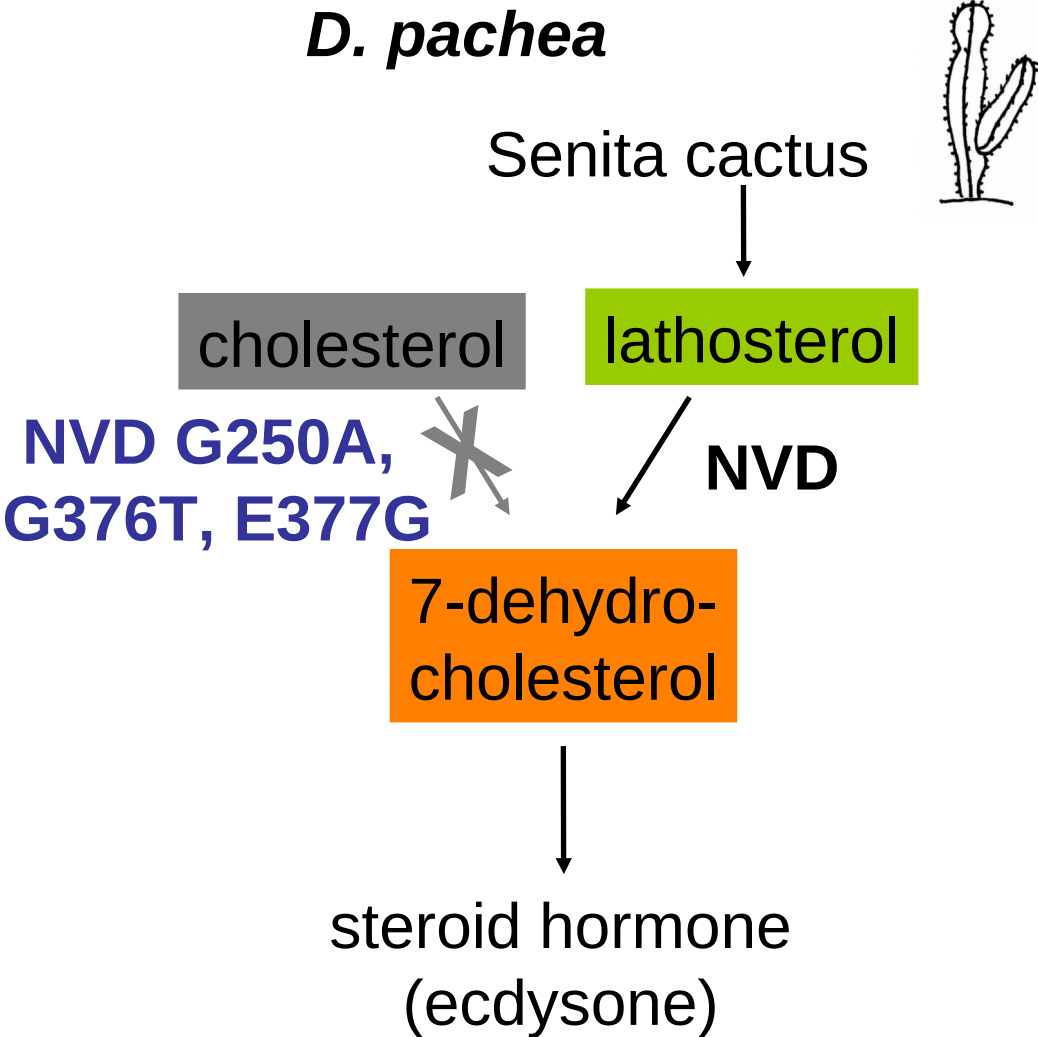


How a few mutations can restrict the ecological niche of a species

Most insects



D. pachea



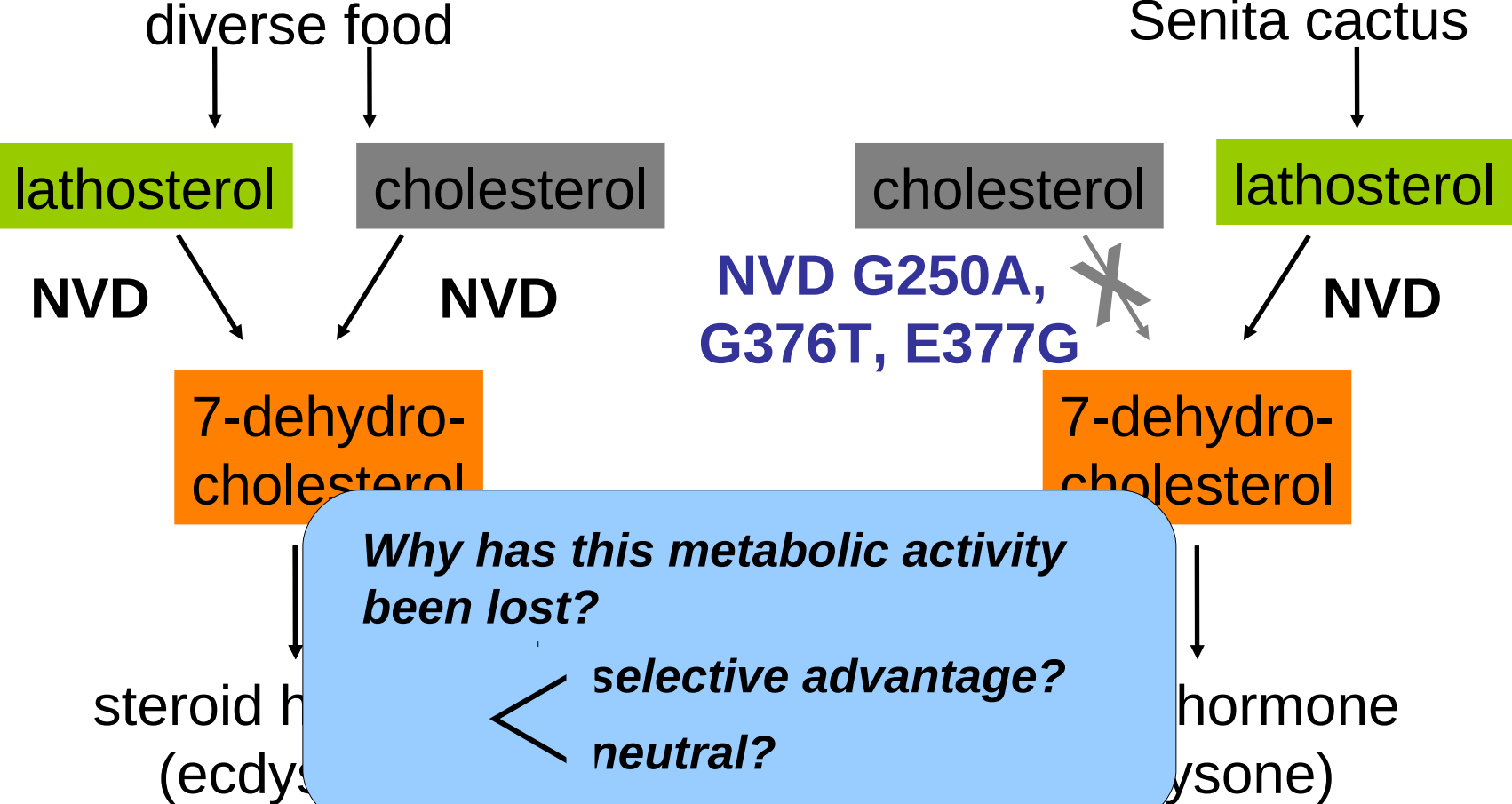
How a few mutations can restrict the ecological niche of a species

Most insects

D. pachea



Senita cactus



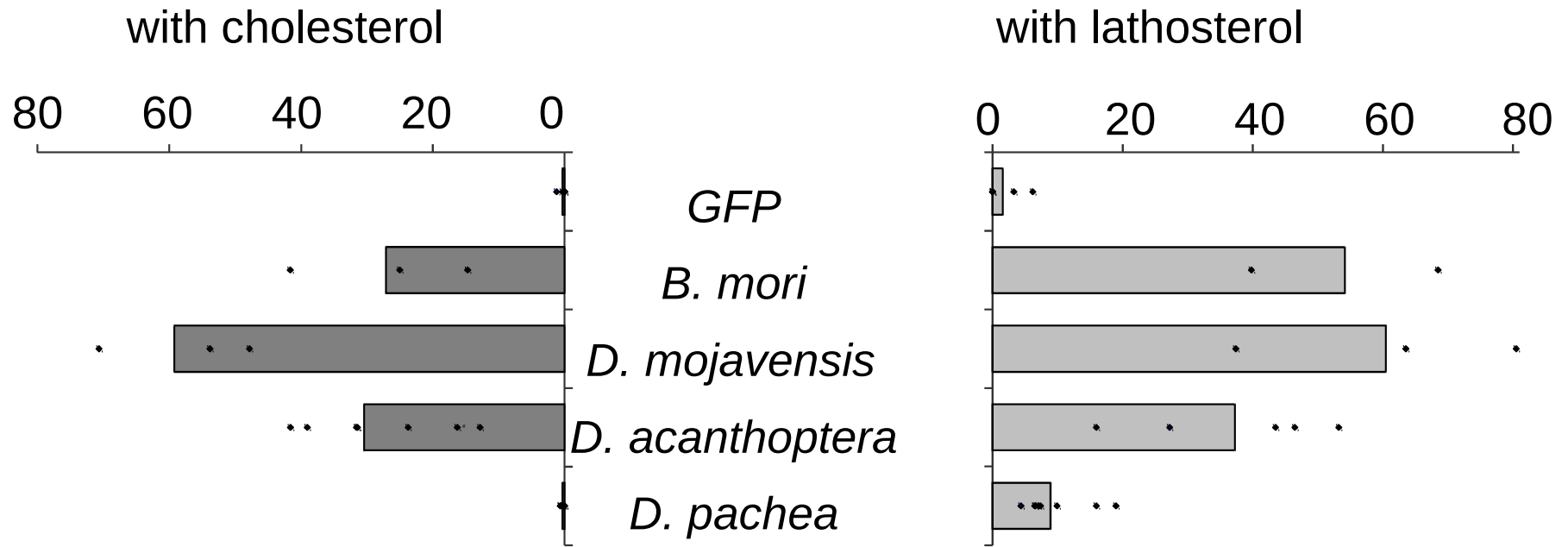
Why has this metabolic activity been lost?

< *selective advantage?*

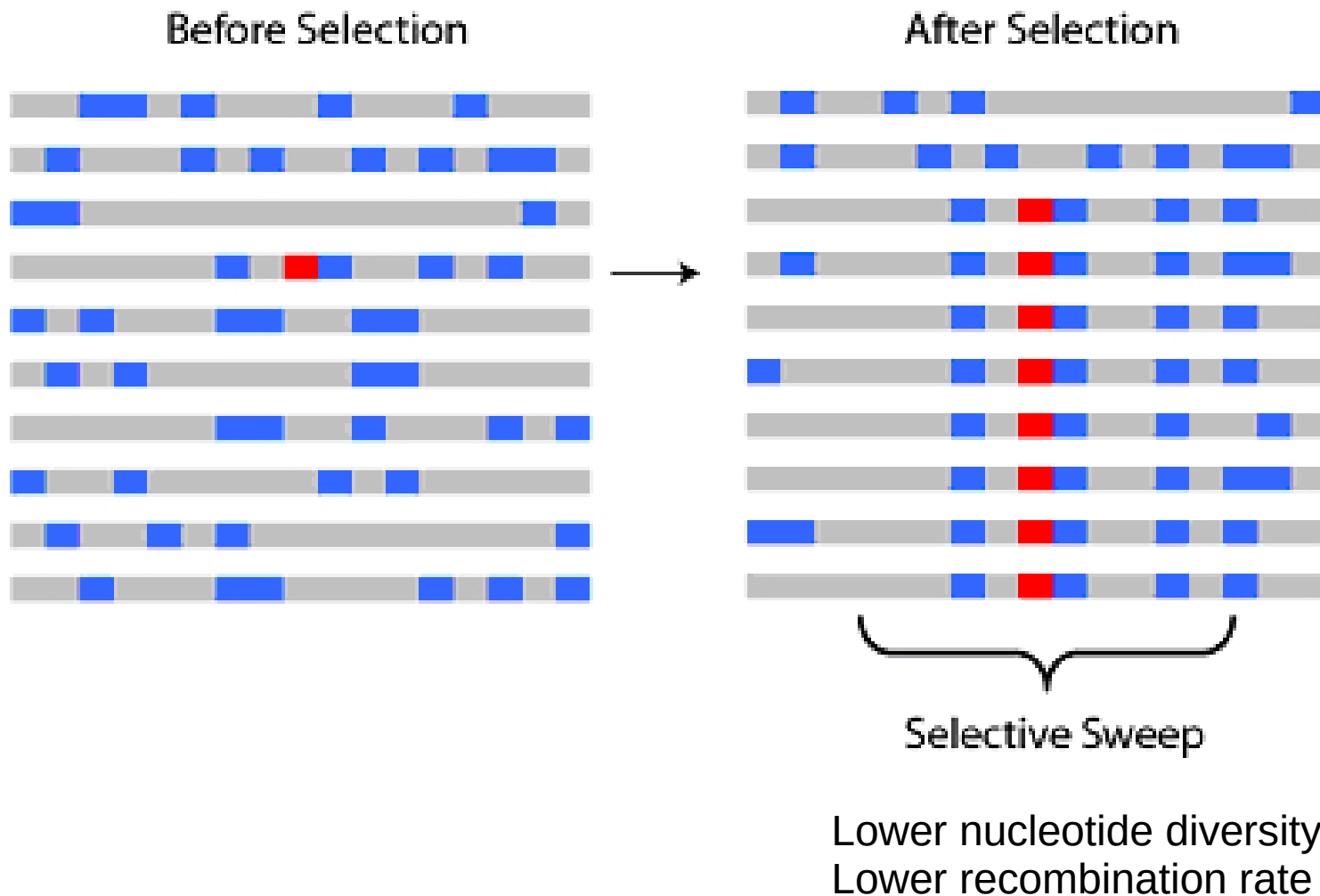
< *neutral?*

In vitro:

No advantage



Recent selection produces a selective sweep



Collecting flies in Mexico



Sequencing the *D. pachea* genome

5 000 €

12 kb



nvd

Sequencing the *D. pachea* genome

5 000 €

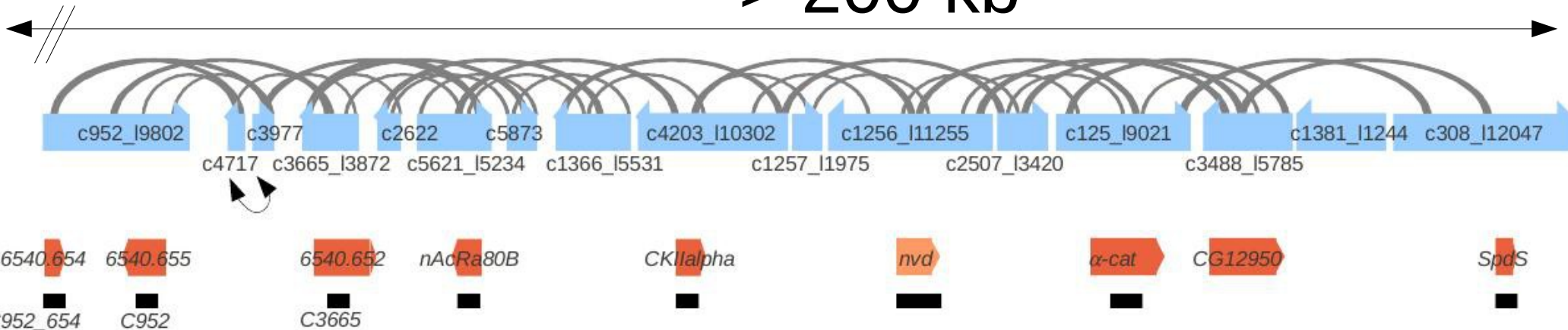
12 kb



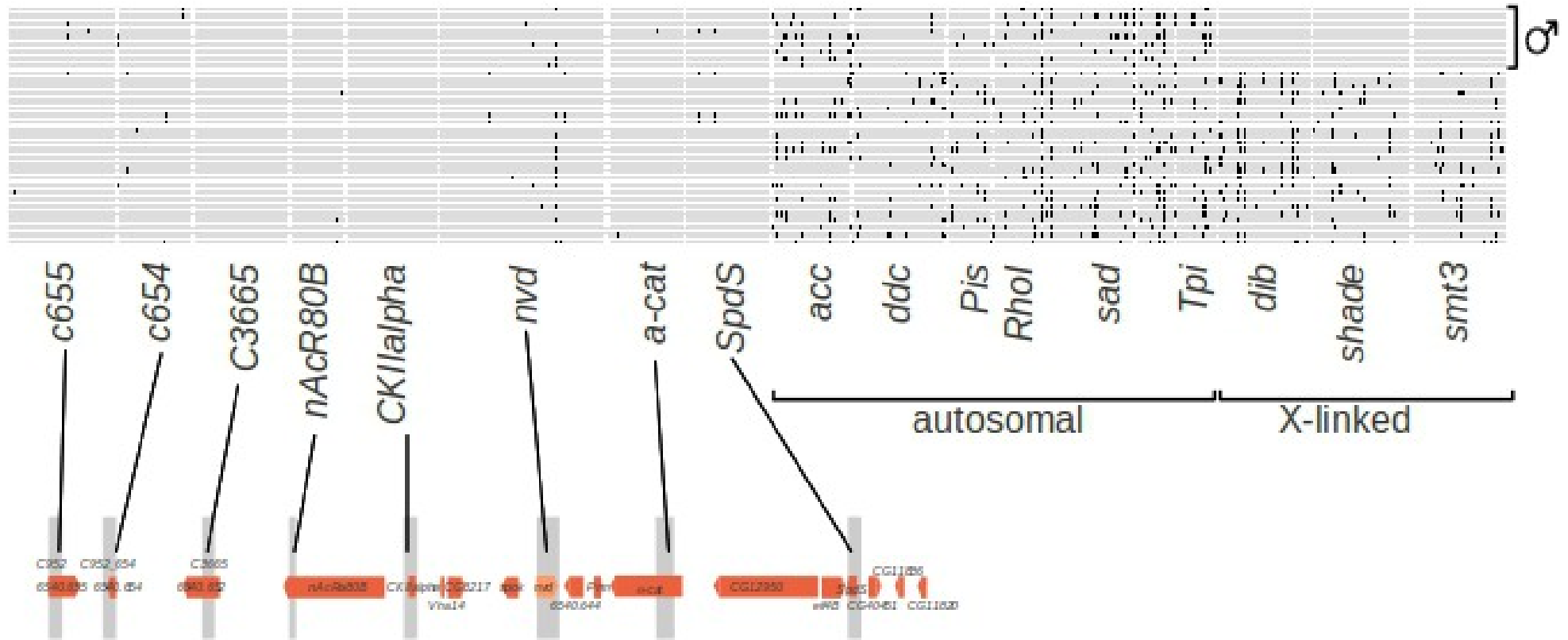
nvd

+ 9 000 €

> 200 kb



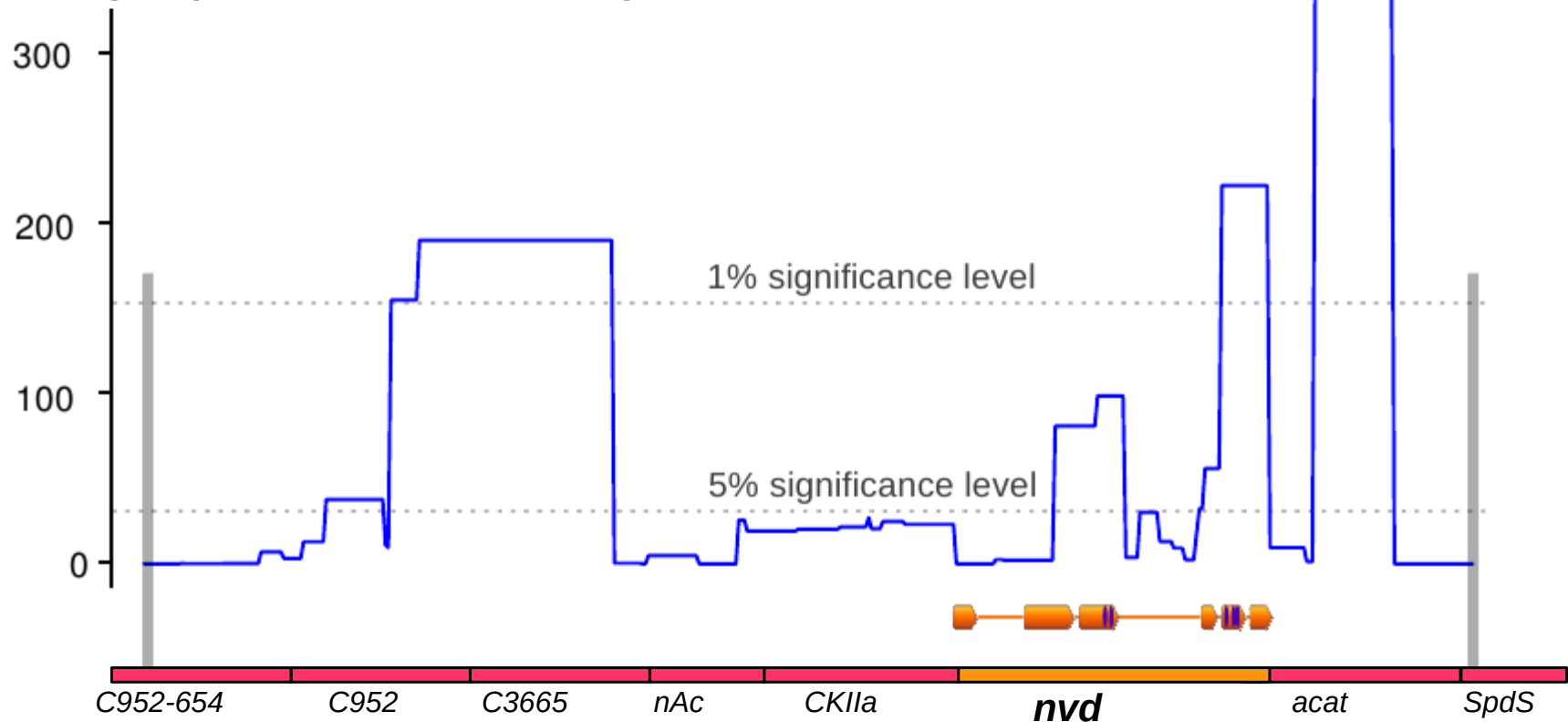
nvd is in a low-nucleotide-diversity region



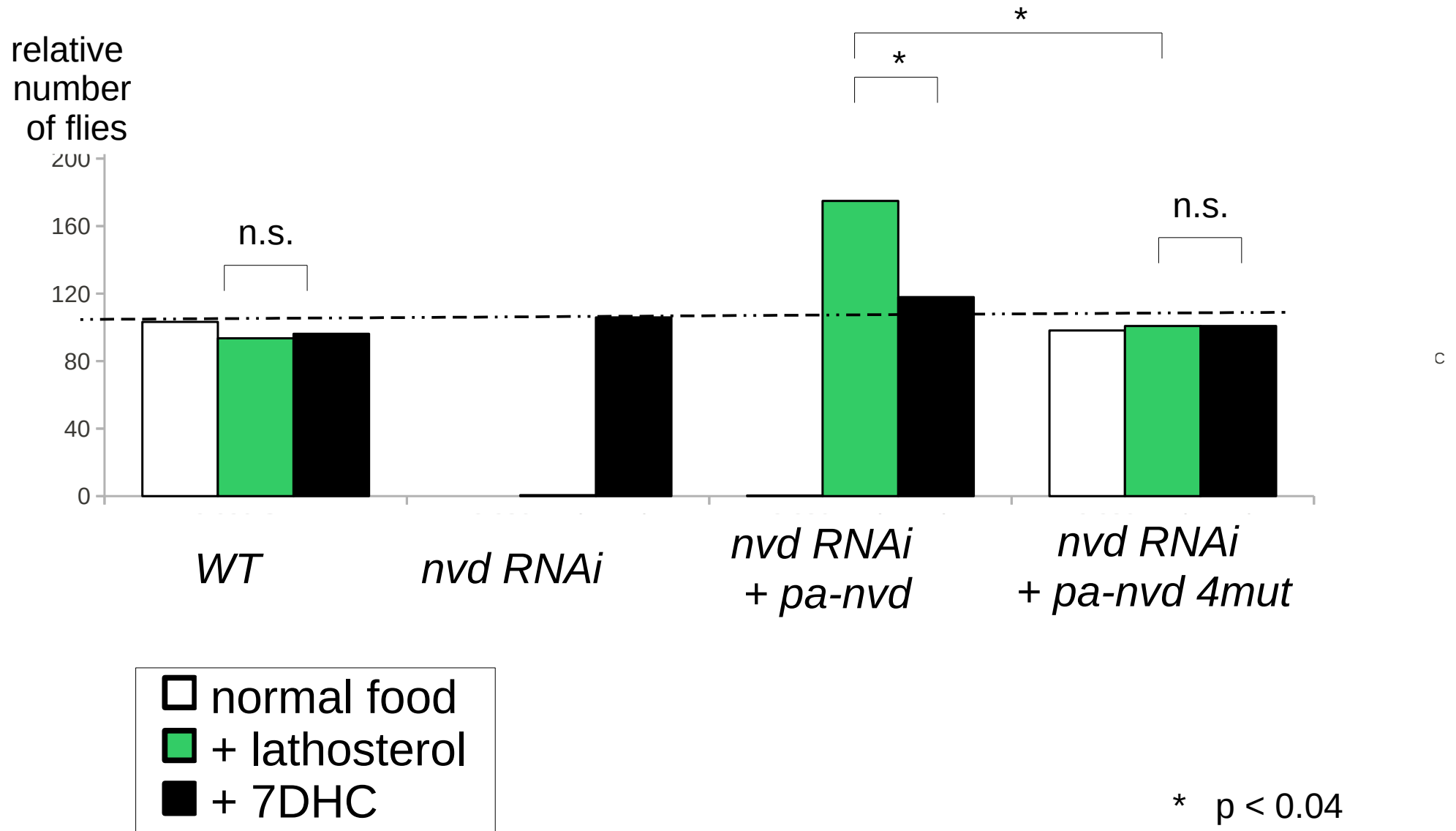
	<i>nvd</i> region	control loci
Size of the region (bp)	12 212	11 763
Inferred number of recombination events	4	53

The *nvd* region has faced recent positive selection

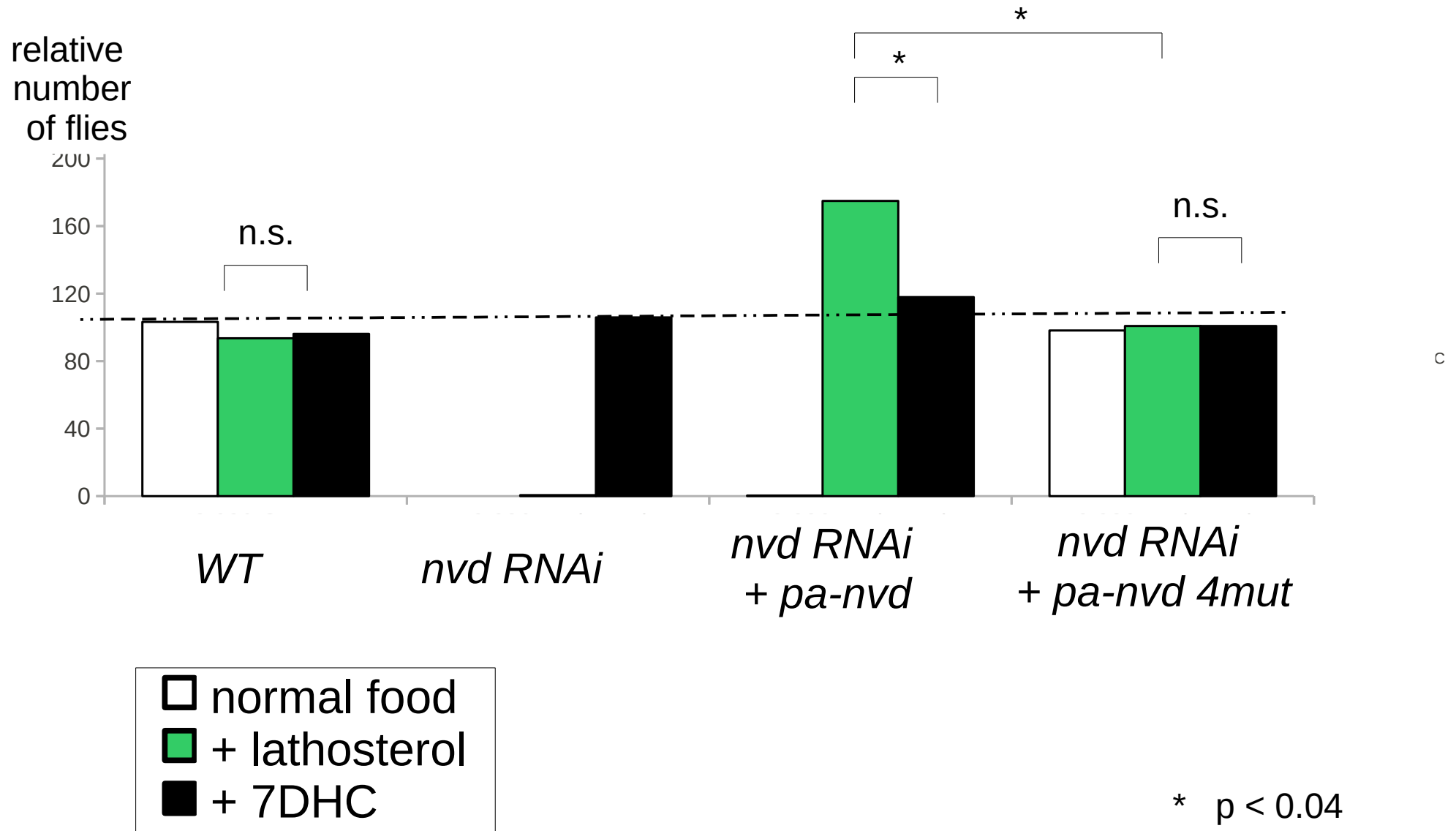
Omega (Kim and Nielsen)



The *D. pachea nvd* allele is advantageous on lathosterol



The *D. pachea nvd* allele is advantageous on lathosterol





D. pachea became dependent on a cactus

How?

via several mutations in a single gene
+ maybe other genes

Why?

These mutations are advantageous on lathosterol

***Losses of metabolic activities can be beneficial.
Short-term beneficial mutations can be selected
despite long-term detrimental effects.***

Michael Lang, Sophie Murat, Géraldine Gouppil (Institut Jacques Monod)

Catherine Blais, Émilie Guittard, René Lafont, Chantal Dauphin-Villemant (UPMC)

Andrew Clark (Cornell U)

Luciano Matzkin (U Alabama)

Takuji Yoshiyama, Hiroshi Kataoka, Ryusuke Niwa (Japon)



cactus generalist ancestor

1) evolved resistance to senita toxic compounds
still feeding on multiple cactus species

2) adaptation to senita
together with inability to live on other cactus species

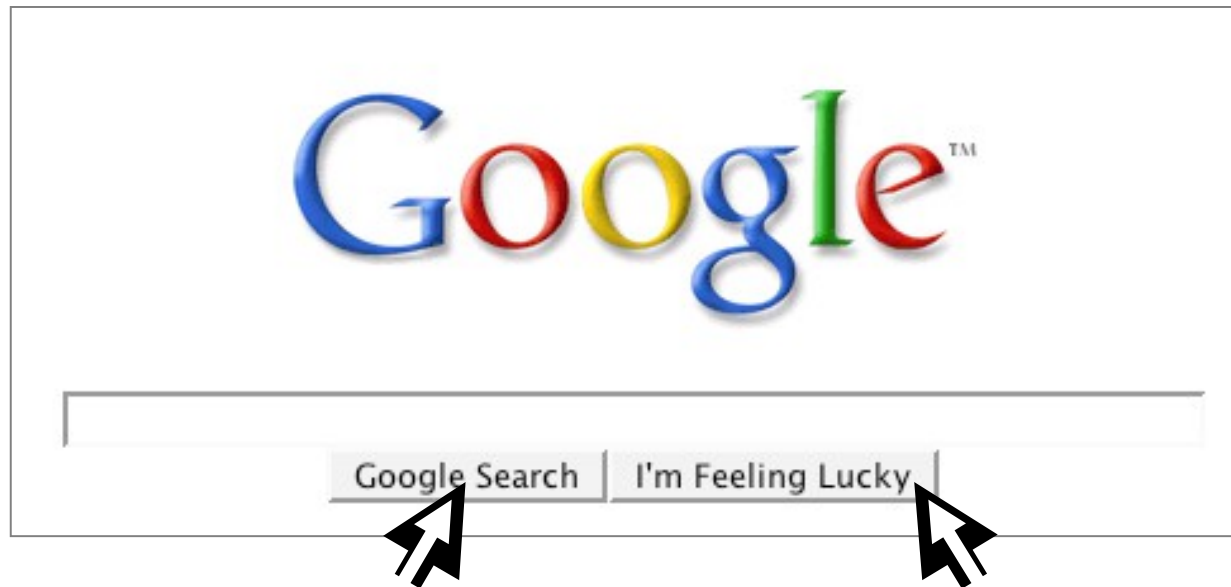
cactus obligate specialist



Methods to identify the genes and the mutations responsible for phenotypic evolution



Two types of approaches



Genetic mapping

Candidate gene

no a priori, fewer biases
long and tedious
rarely ends with identification of the gene

only with strains/species which produce fertile hybrids

Based on an a priori idea
can be fast and efficient

will only find known genes

In both cases, genes with small effect are more difficult to identify

Various methods

Genetic

which chromosome (ex: autosomal versus sex)

QTL mapping

Genetic association studies

Complementation tests

General biology

General knowledge of the genes involved in the phenotype

Similarity with a known phenotype

Correlation with a change in gene expression level/pattern

Final test of protein activity

in vitro in *E. coli*, by transgenesis in the studied species or the closest model organism (ex: *beta-defensin* of dogs tested in mouse)

Final test of cis-regulatory regions

- with reporter constructs, transgenesis, comparison of both regions
- comparison of allele expression levels in hybrids (pyrosequencing)