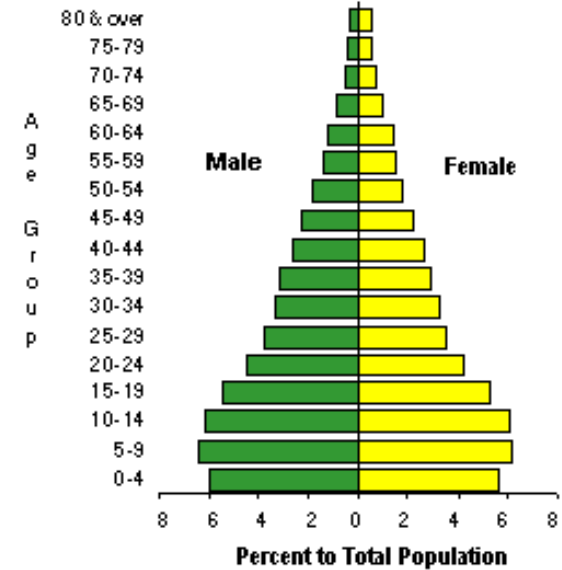
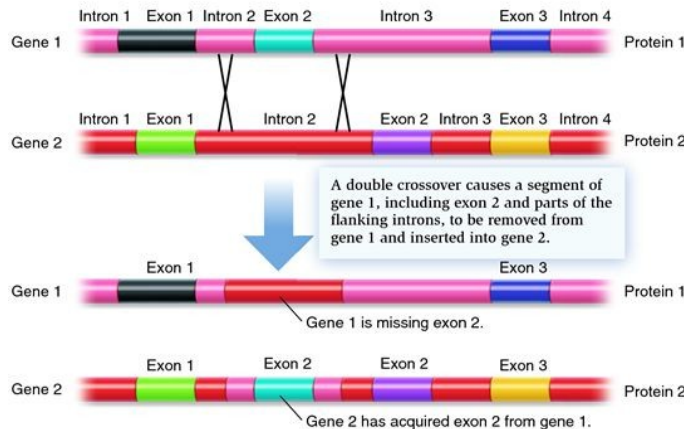
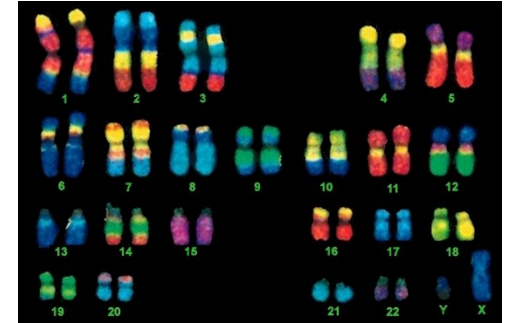
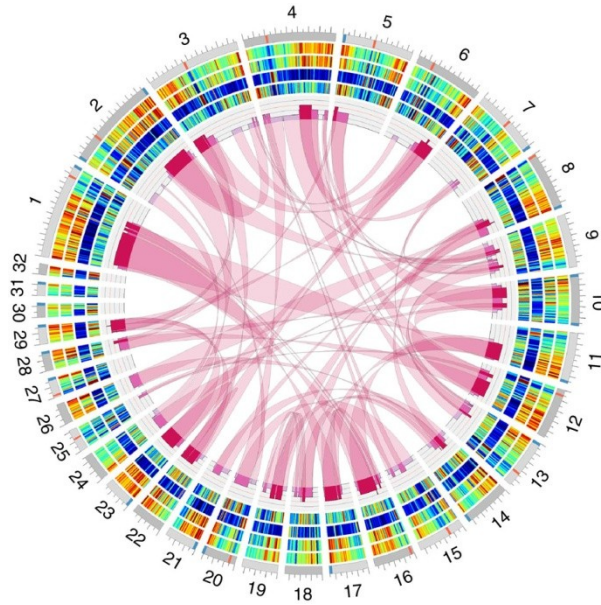


# EVOLUTION OF GENETIC SYSTEMS



# EVOLUTION OF THE GENOME

Genome size and C-value:

C-value = amount of DNA in haploid genome (pg, bp)

Prokaryotes:

$6 \times 10^5 - 10^7$  bp (20-fold span)

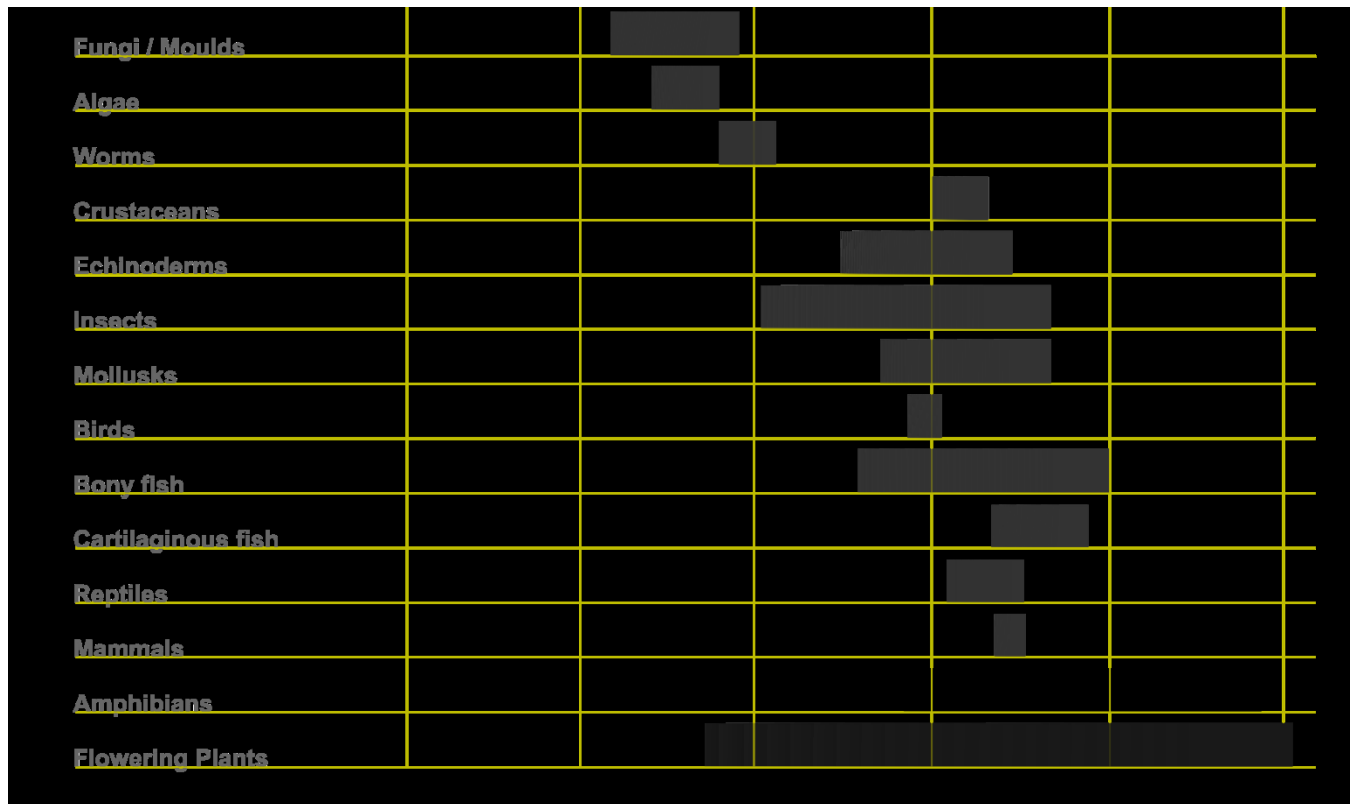
smallest: *Mycoplasma genitalium* 525 genes, the smallest functional  
artificially synthesized genome  $\approx 473$  genes (modified from *M. mycoides*)

largest: some G+ bacteria, cyanobacteria

					in bp
<i>Mycoplasma</i>					
Gram positive bacteria					
Gram negative bacteria					

# Eukaryotes:

$8,8 \times 10^6$  –  $6,9 \times 10^{11}$  bp (80 000-fold span!)



no relation between genome size and organismal complexity or number of genes

large differences even in related organisms:

*Paramecium caudatum* (8 600 000 kb) × *P. aurelia* (190 000 kb)

human: ca.  $6 \times 10^9$  bp (~ 6,5 pg DNA)

× *Amoeba proteus*:  $2,9 \times 10^{11}$  bp

*Polychaos dubium* (*Amoeba dubia*):  $6,7 \times 10^{11}$  bp

⇒ C-value paradox (C-value enigma)

**Table 4. C values from eukaryotic organisms ranked by genome size.**

Species	C value (kb)
<i>Navicola pelliculosa</i> (diatom)	35,000
<i>Drosophila melanogaster</i> (fruitfly)	180,000
<i>Paramecium aurelia</i> (ciliate)	190,000
<i>Gallus domesticus</i> (chicken)	1,200,000
<i>Erysiphe cichoracearum</i> (fungus)	1,500,000
<i>Cyprinus carpio</i> (carp)	1,700,000
<i>Lamprreta planeri</i> (lamprey)	1,900,000
<i>Boa constrictor</i> (snake)	2,100,000
<i>Parascaris equorum</i> (roundworm)	2,500,000
<i>Carcarias obscurus</i> (shark)	2,700,000
<i>Rattus norvegicus</i> (rat)	2,900,000
<i>Xenopus laevis</i> (toad)	3,100,000
<b><i>Homo sapiens</i> (human)</b>	<b>3,400,000</b>
<i>Nicotiana tabaccum</i> (tobacco)	3,800,000
<i>Paramecium caudatum</i> (ciliate)	8,600,000
<i>Schistocerca gregaria</i> (locust)	9,300,000
<i>Allium cepa</i> (onion)	18,000,000
<i>Coscinodiscus asteromphalus</i> (diatom)	25,000,000
<i>Lilium formosanum</i> (lily)	36,000,000
<i>Amphiuma means</i> (newt)	84,000,000
<i>Pinus resinosa</i> (pine)	68,000,000
<i>Protopterus aethiopicus</i> (lungfish)	140,000,000
<i>Ophioglossum petiolatum</i> (fern)	160,000,000
<i>Amoeba proteus</i> (amoeba)	290,000,000
<i>Amoeba dubia</i> (amoeba)	670,000,000

closely related species!

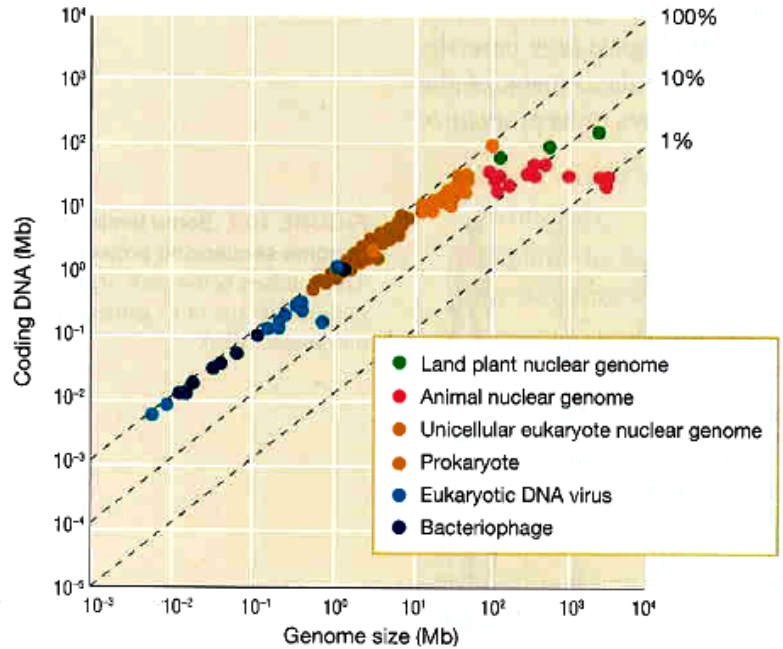
marbled lungfish:  
> 40× larger than human

almost 200× larger than human

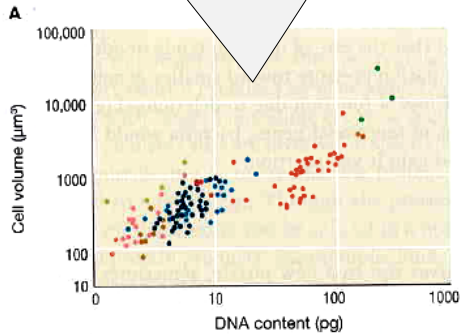
Data from Cavalier-Smith (1985), Sparrow et al. (1972), and other references.

# C-value paradox:

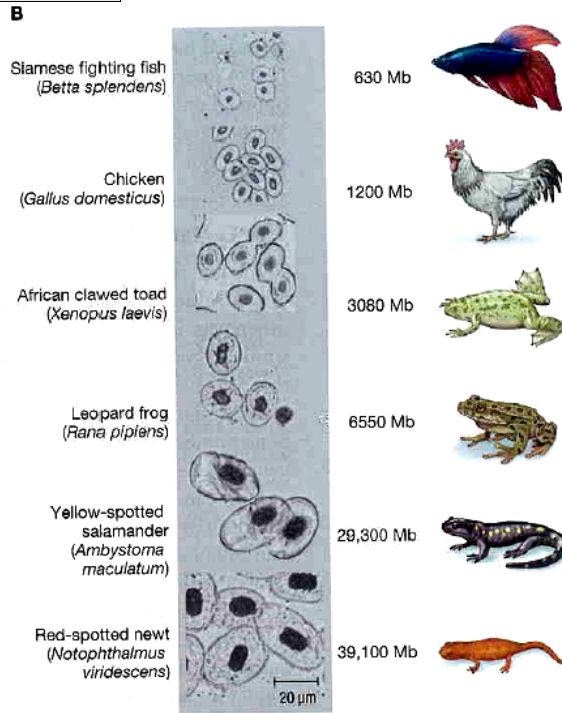
large genomes include large amount of non-coding DNA



the larger the genome,  
the larger the cell



- Teleost fishes
- Reptiles
- Birds
- Cartilaginous fishes
- Anuran amphibians
- Urodele amphibians
- Jawless fishes
- Lungfishes



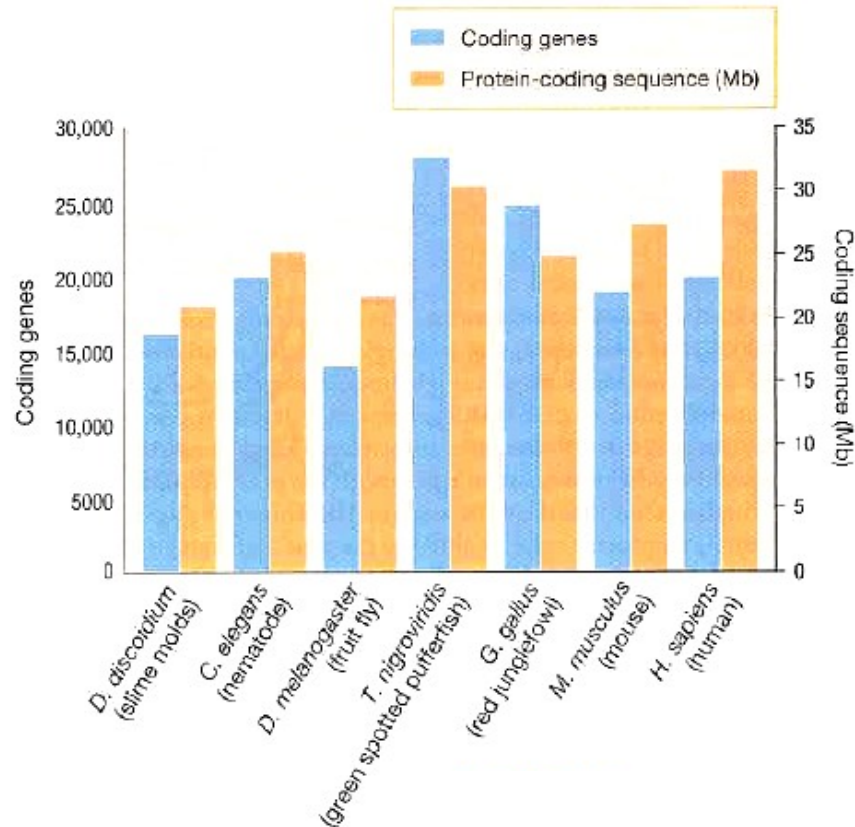
**FIGURE 10.4** Cell size increases with genome size in vertebrates. (A) Cell volume in cubic micrometers ( $\mu\text{m}^3$ ) is plotted against DNA content in picograms (pg) of a diploid cell. Both axes are plotted on a logarithmic scale. Adapted from Gregory (2001). (B) Micrographs of individual cells reveal the same trend. The dark-stained material in the center of each cell is DNA.

large genomes  $\Rightarrow$   
large cells  $\Rightarrow$  influence on:

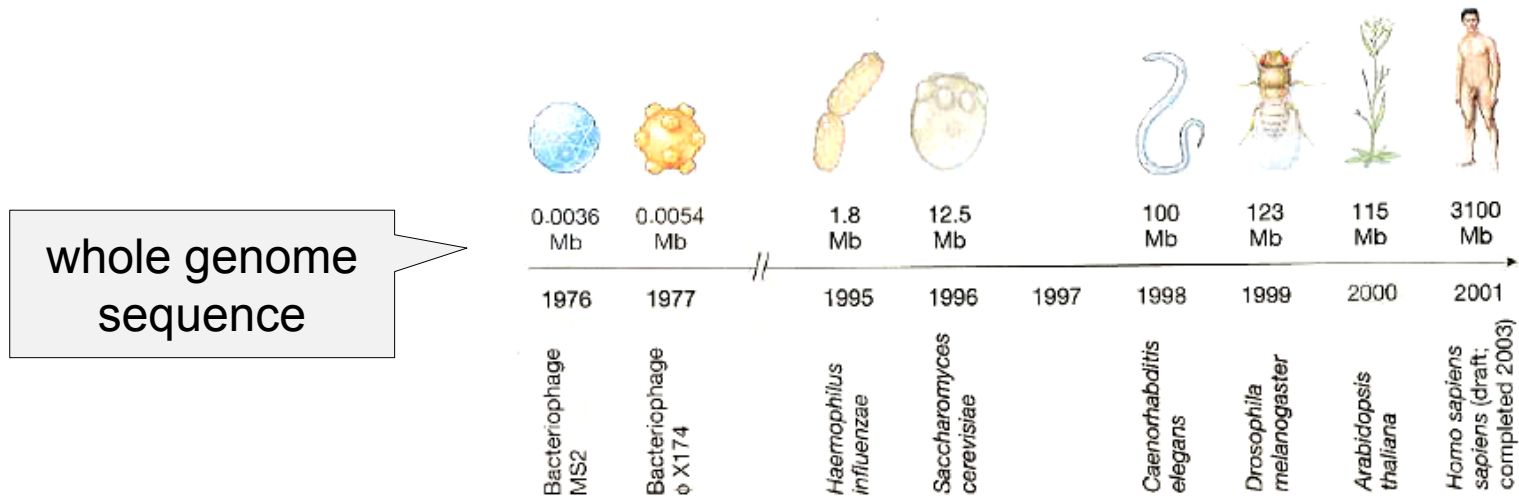
- cell division speed
- metabolism efficiency
- rate of ions/proteins exchange
- body size

## G-value paradox:

despite diversity of organismal complexity, metazoans tend to have similar numbers of protein-coding genes (G-value)



No dependency on the total number of genes but on complexity of gene regulation networks – organisms with similar number of genes may have very different patterns of gene regulation networks



## How many coding genes are in the human genome?

before 2001 (draft version of the genome) estimates from 50 000 till  
> 140 000 (max. 212 278) genes

Int. Human Genome Sequencing Consortium (IHGSC) 2001:  
30 000–40 000 protein coding genes

IHGSC 2004: 20 000–25 000 protein coding genes

Ensembl – May 2012: 21 065 coding genes

Ensembl – January 2013: 20 848 coding genes

Ensembl – February 2014: 20 805 coding genes

Ensembl – December 2014: **20 364 coding genes**



## Repetitive DNA:

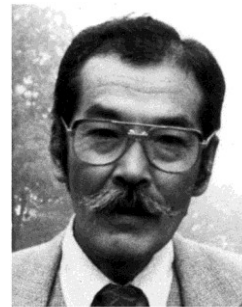
1. Highly repetitive = satellite
2. Moderately repetitive = minisatellites, microsatellites
3. Transposable elements, retroelements (SINE, LINE)

## Why does repetitive DNA exist?

Cavalier-Smith (1978): there must be some function

Doolittle and Sapienza, Orgel and Crick (1980): repetitive DNA is „selfish“

Susumu Ohno (1972): „junk DNA“



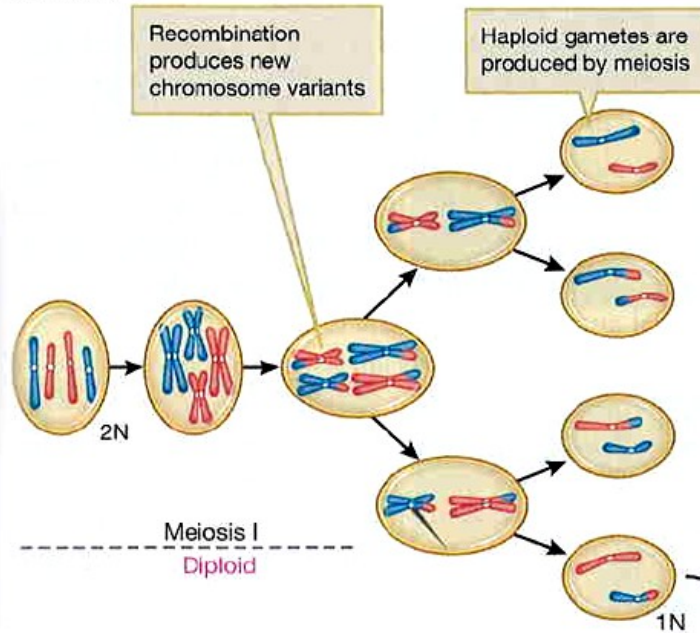
„junk“ ≠ „garbage“ ⇒ in future it may gain some function

# EVOLUTION OF SEX

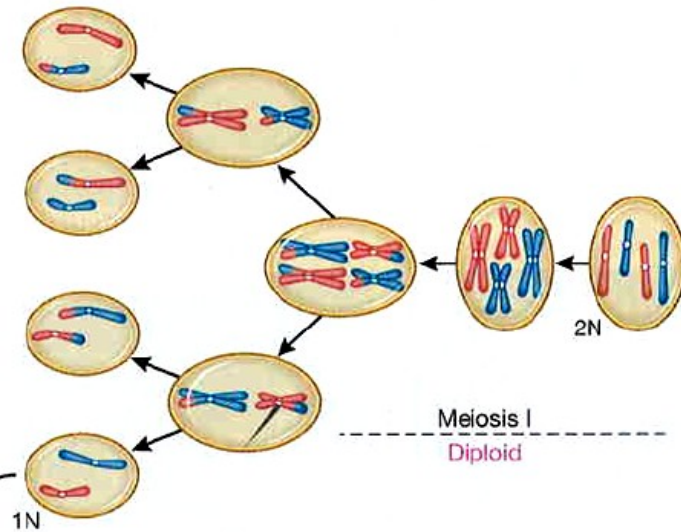
sex = meiosis, recombination

amphimixis

Individual 1



Individual 2



2N  
Zygote  
Diploid

Gametes fuse, and diploidy is restored in the zygote

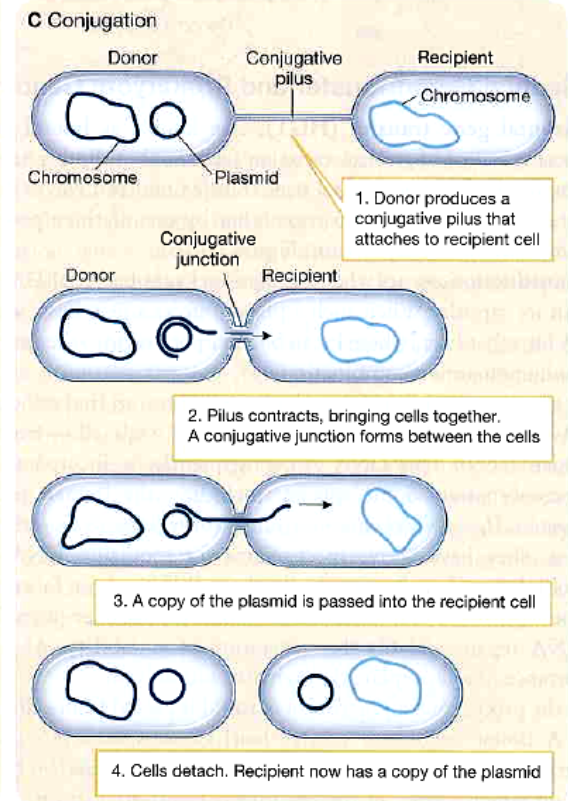
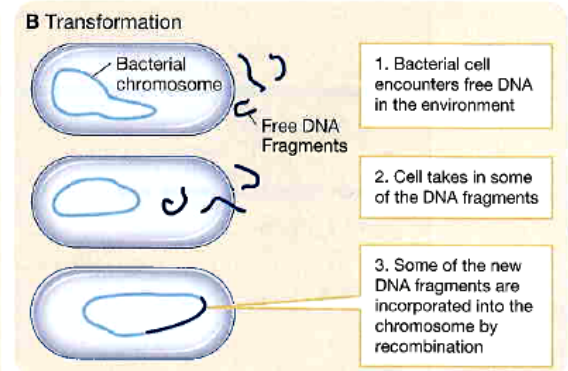
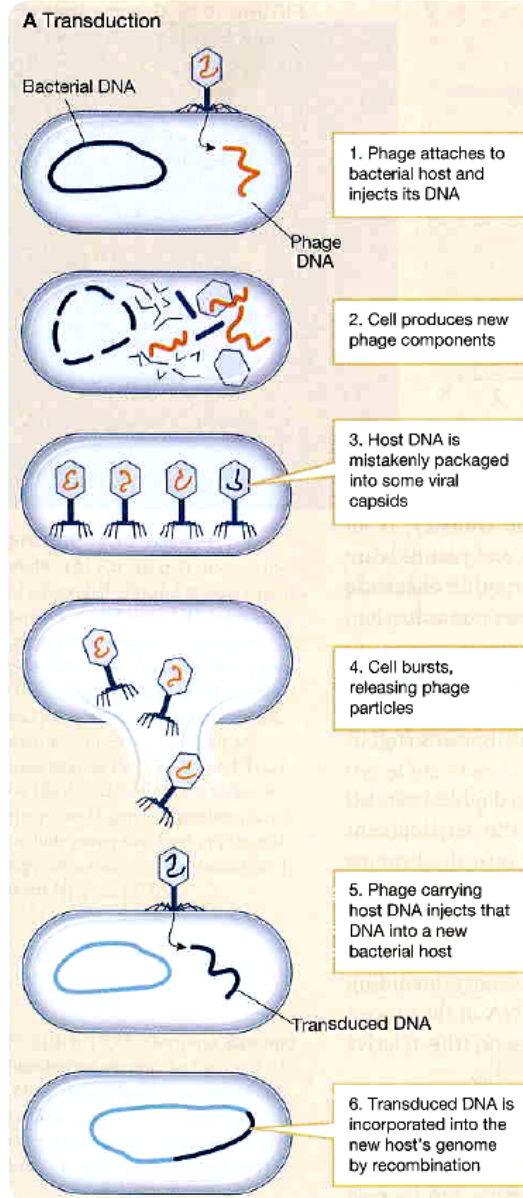
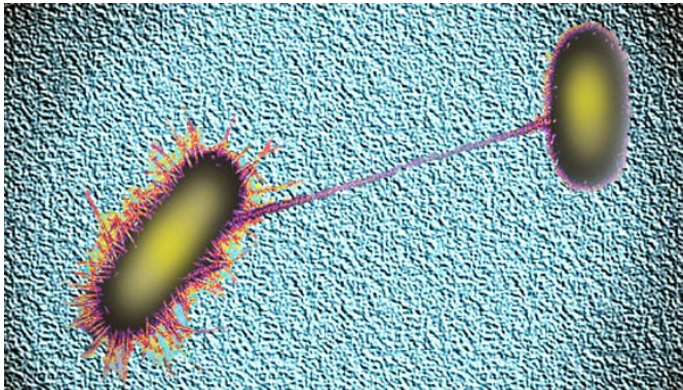
# „sex“ in Prokaryotes:

conjugation

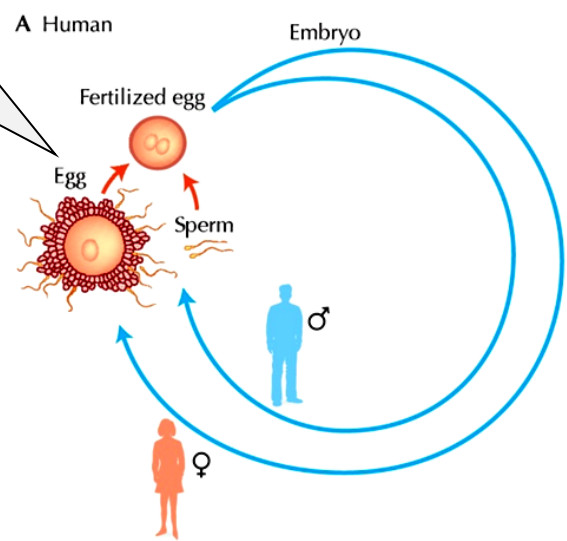
transformation

transduction

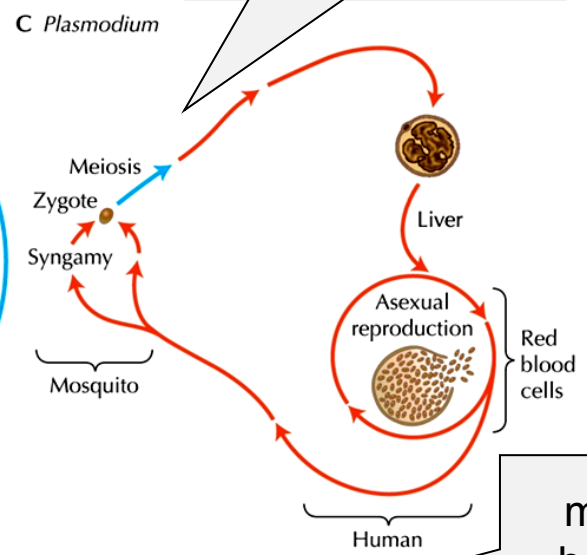
conjugation in *E.coli*:



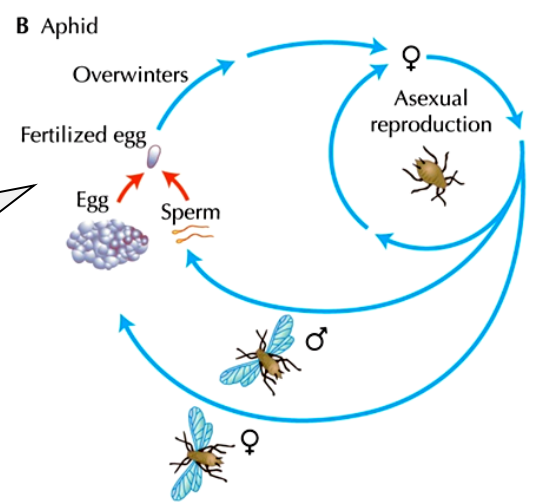
no division in haploid stage



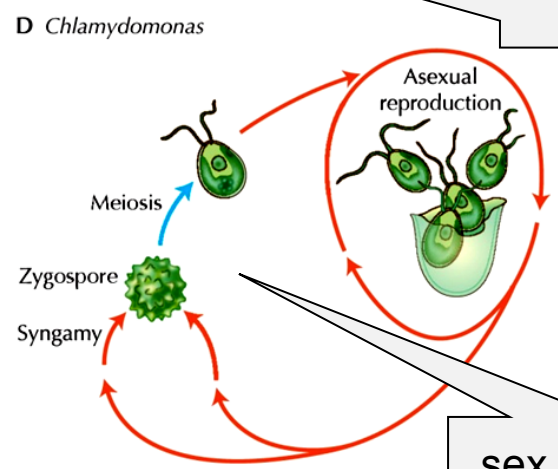
sex at the end of life cycle



facultative sex



most of life in haploid phase

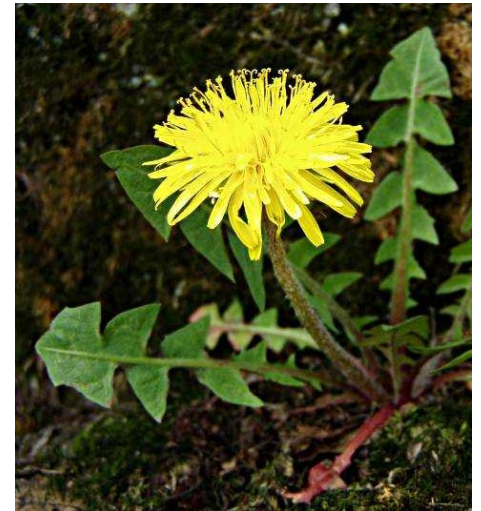
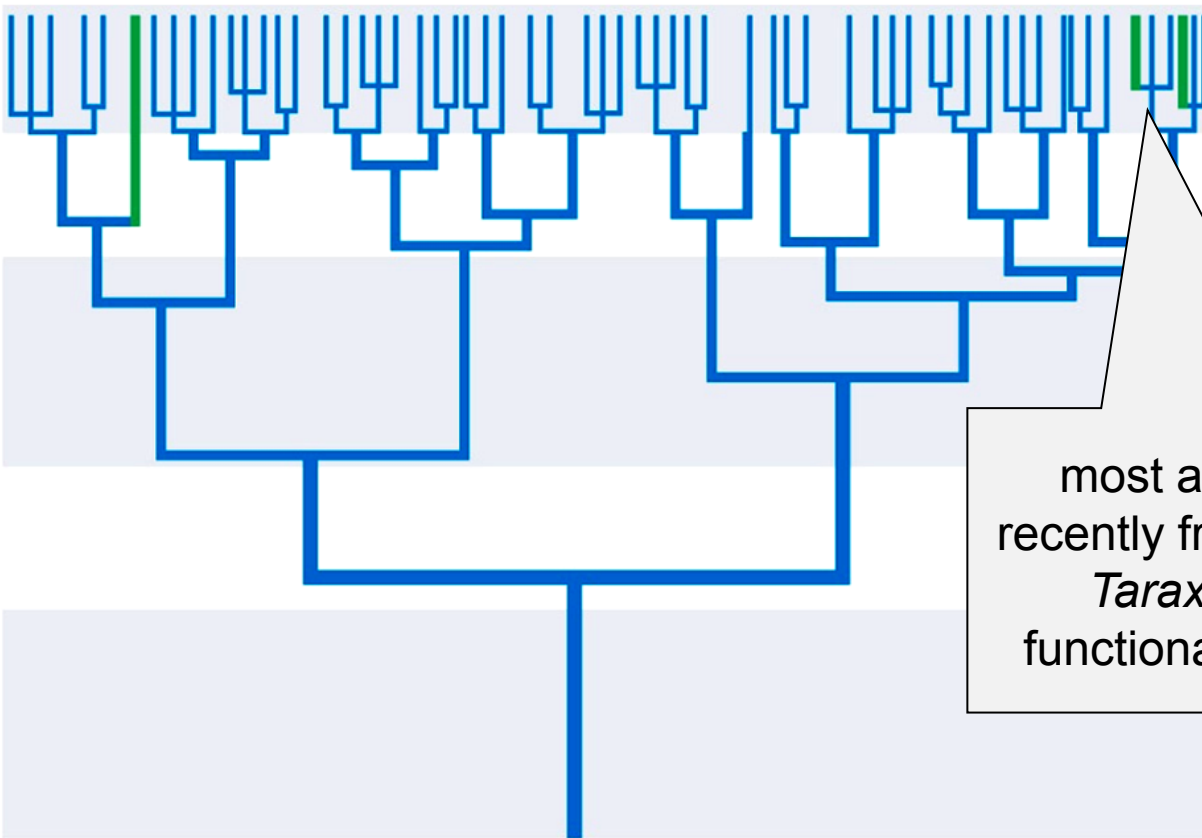


sex triggered by starving (shortage of N<sub>2</sub>)

phylogenetic position of asexual taxa:

mostly recent lineages

taxa scattered



*T. officinale*

most asexual lineages arised recently from sexual; eg. dandelion  
*Taraxacum officinale*: non-functional stamina, yellow colour

exceptions:

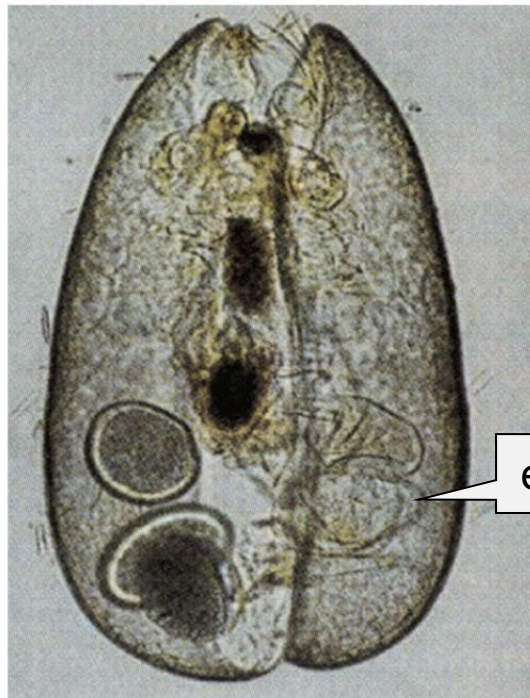
Bdelloidea rotifers:

fossils in amber 35-40 MY  
existency ~100 MY

ostracods:

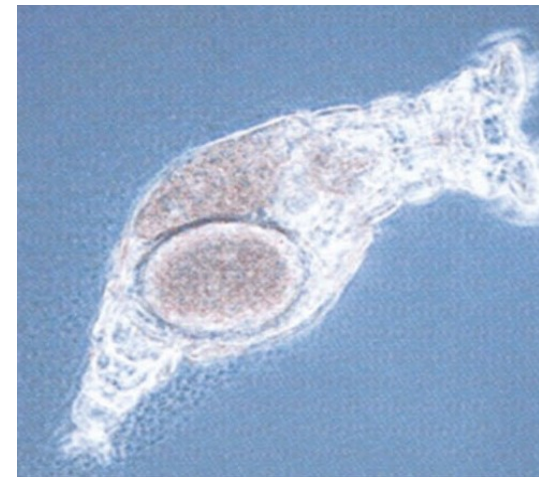
asexual ~100 MY  
× recently males found

*Philodina roseola*



eggs

*Darwinula stevensoni*



*Macrotrachela quadricornifera*

## Disadvantages of sexual reproduction

time and energy necessary for finding a partner (finding itself may be a problem), further effort before copulation

increased risk of predation or parasitism, transmission of venereal diseases

susceptibility to extinction at low  $N_e$

lower capability of colonization

complex meiotic molecular machinery

meiosis: 10-100 h × mitosis: 15 min – 4 h

impact of sexual selection on males → reduction of population fitness

eg. Soay sheep (St. Kilda): males die during the first winter

× females and castrated males several years

# Disadvantages of sexual reproduction:

break-up of advantageous allele combinations by recombination

Eg.:  $A_1$  (dominant) = large claws,  $A_2$  (recessive) = small claws

$B_1$  (dominant) = aggressive,  $B_2$  (recessive) = meek

Gametes produced by  $A_1A_2B_1B_2$  parent  
(large claw, aggressive)

disadvantageous combinations

advantageous combinations

Gametes produced by  $A_2A_2B_2B_2$  parent  
(small claw, meek)

advantageous combinations

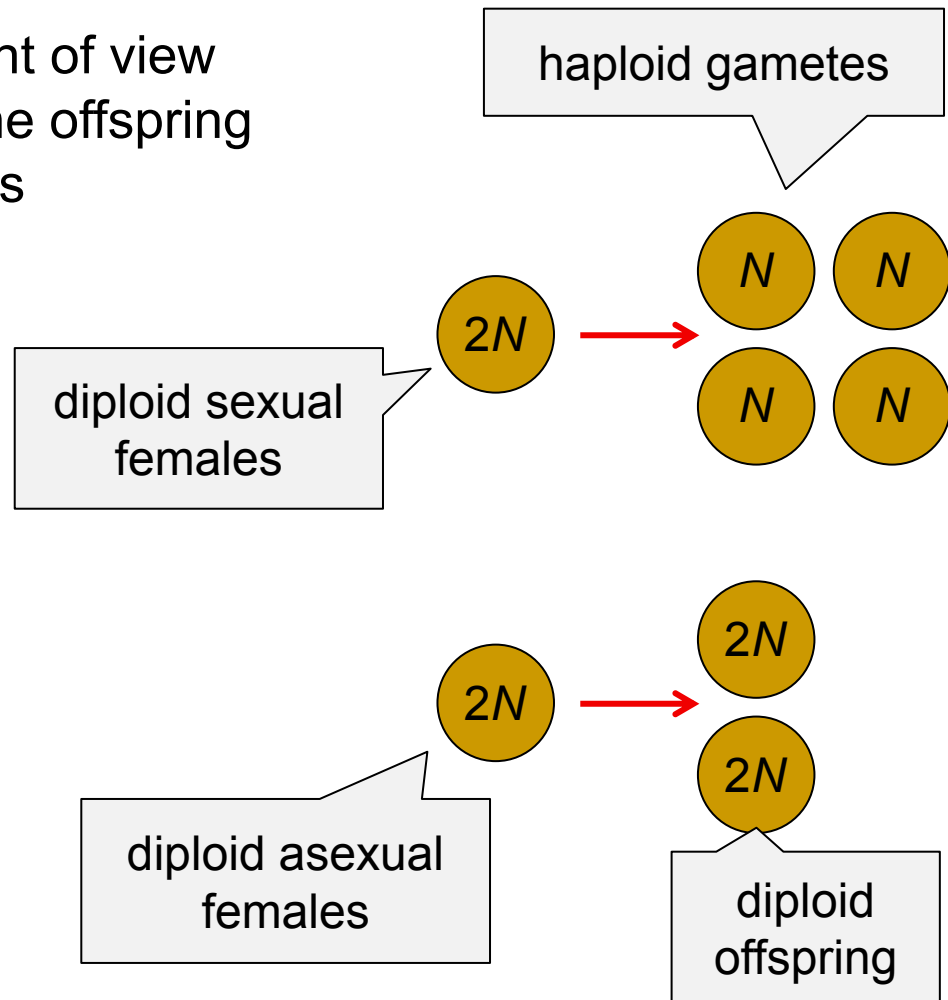
	$A_1B_1$	$A_1B_2$	$A_2B_1$	$A_2B_2$
$A_2B_2$				
$A_2B_2$				
$A_2B_2$				
$A_2B_2$				



# Disadvantages of sexual reproduction:

action of selfish elements (conflict of genes) → reduction of population fitness (B chromosomes, transposons)

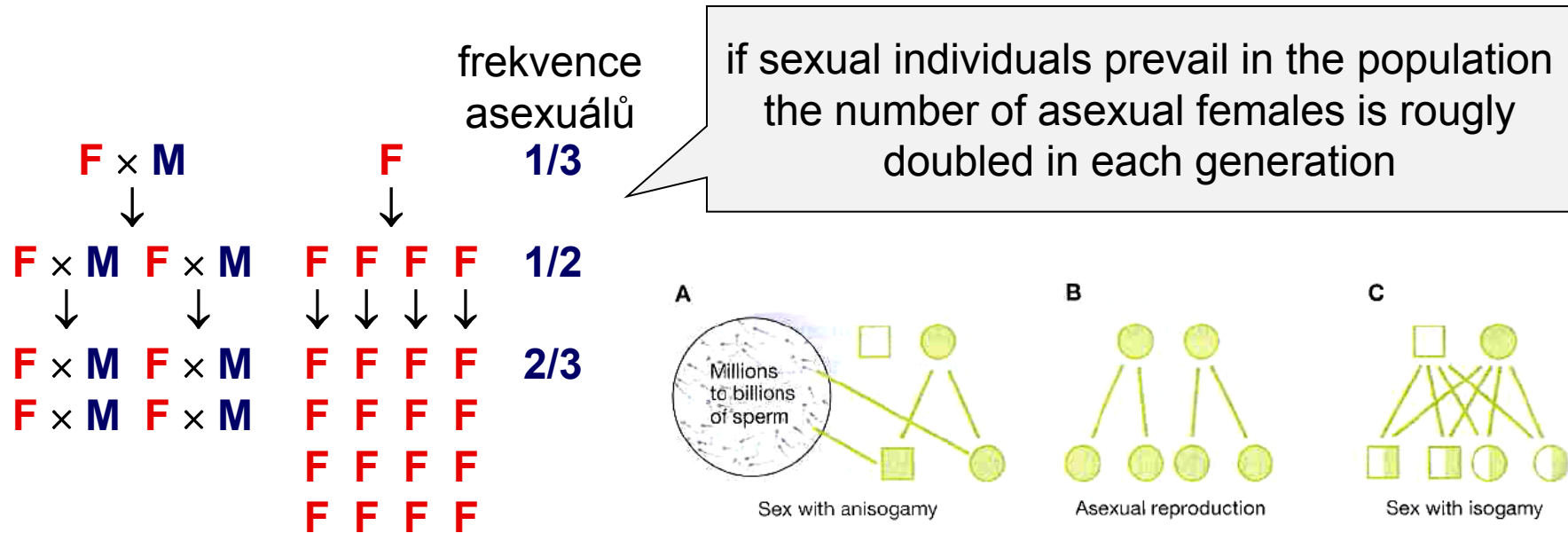
from the sexual female's point of view  
disadvantage, because the offspring  
have only 1/2 of her genes



# J. Maynard Smith: What is the fate of sexual and asexual population?

assumptions: way of reproduction has no effect on

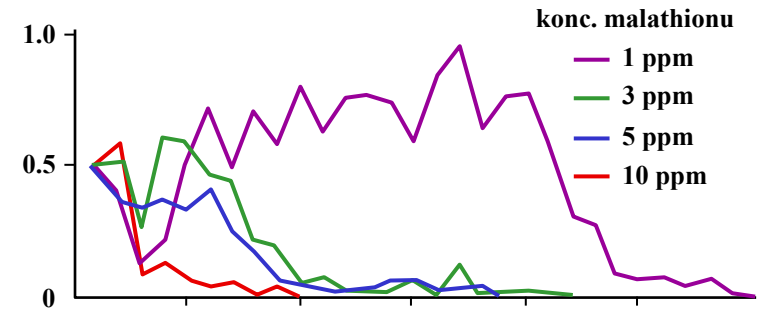
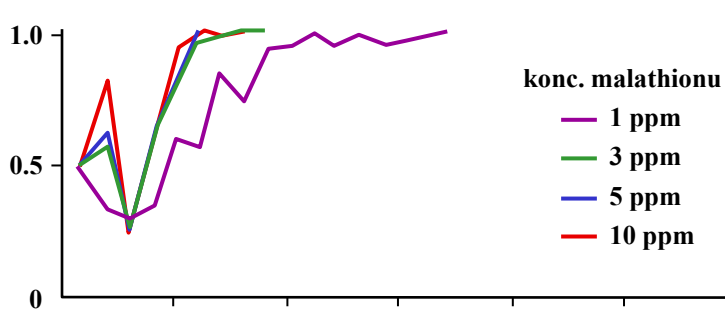
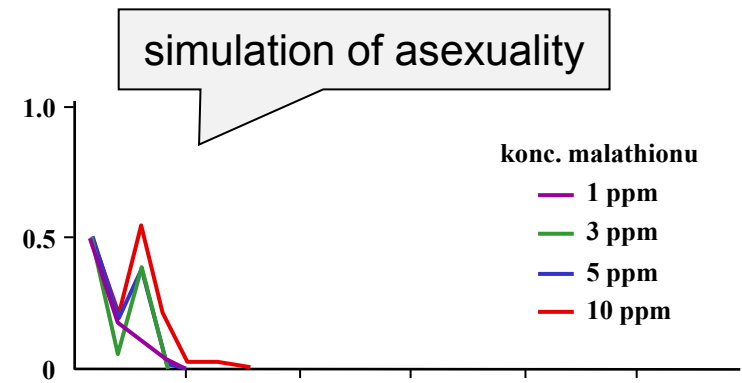
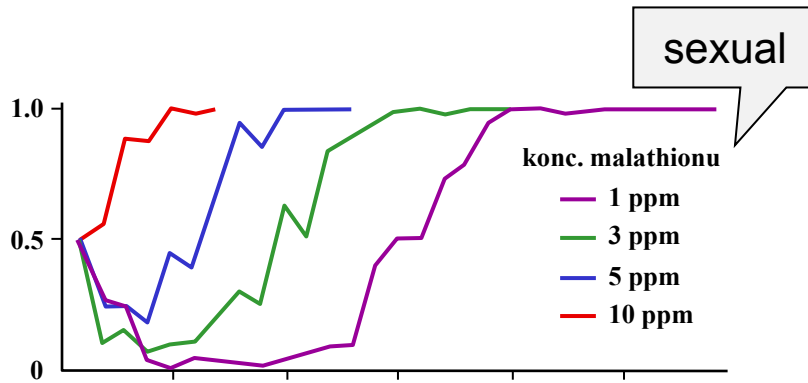
1. number of descendants (eg. when males take care of offspring)
2. probability of offspring survival



⇒ twofold cost of sex, ie. 50% selective disadvantage of sex (not for isogamy! → so rather cost of males)

## ad 2) effect of environment

experiment with *Tribolium castaneum*: competition, insecticide, reproductive advantage of „asexuals“



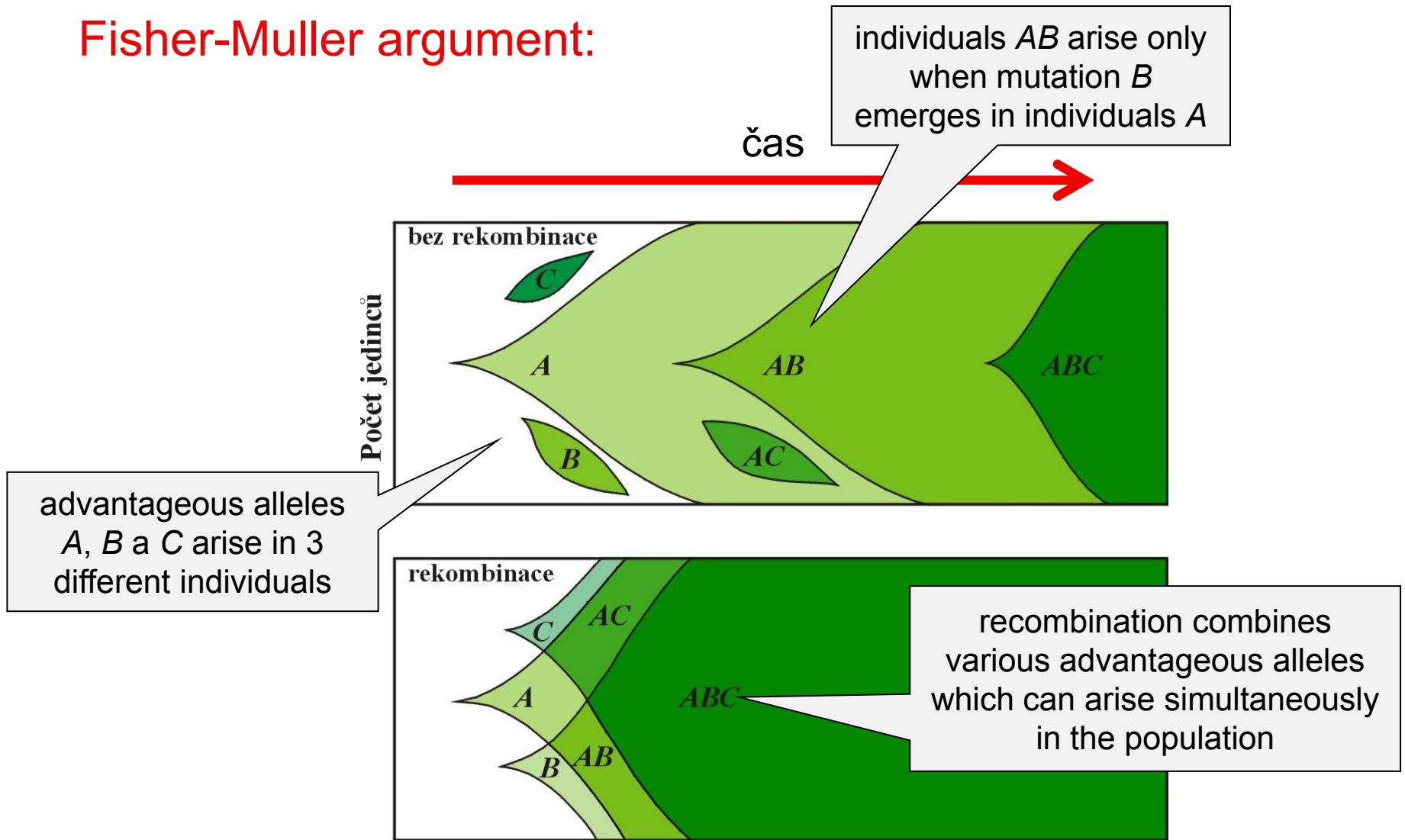
at first prevalence of asexuals, eventually fixation of sexuals

faster at higher insecticide concentrations

offspring of sexuals have higher fitness  $\Rightarrow$  **assumption 2 is not valid**

# Advantages of sexual reproduction

## Fisher-Muller argument:



## Effects of recombination:

1 locus → max. 2 variants of gametes (heterozygote)

2 loci → 4 variants: gametes  $AB/ab$  →  $ab, aB, Ab, AB$

10 loci →  $2^{10} = 1024$  different gametes and  $2^{n-1}(2^n+1) = 524\ 800$  diploid genotypes

for population genetics *the only* consequence of sex is **linkage equilibrium** – when it is reached sex loses sense

every model explaining advantage of sex must include a mechanism which eliminates some gene combinations (LD arises), and explain why genes causing LD are favoured by selection

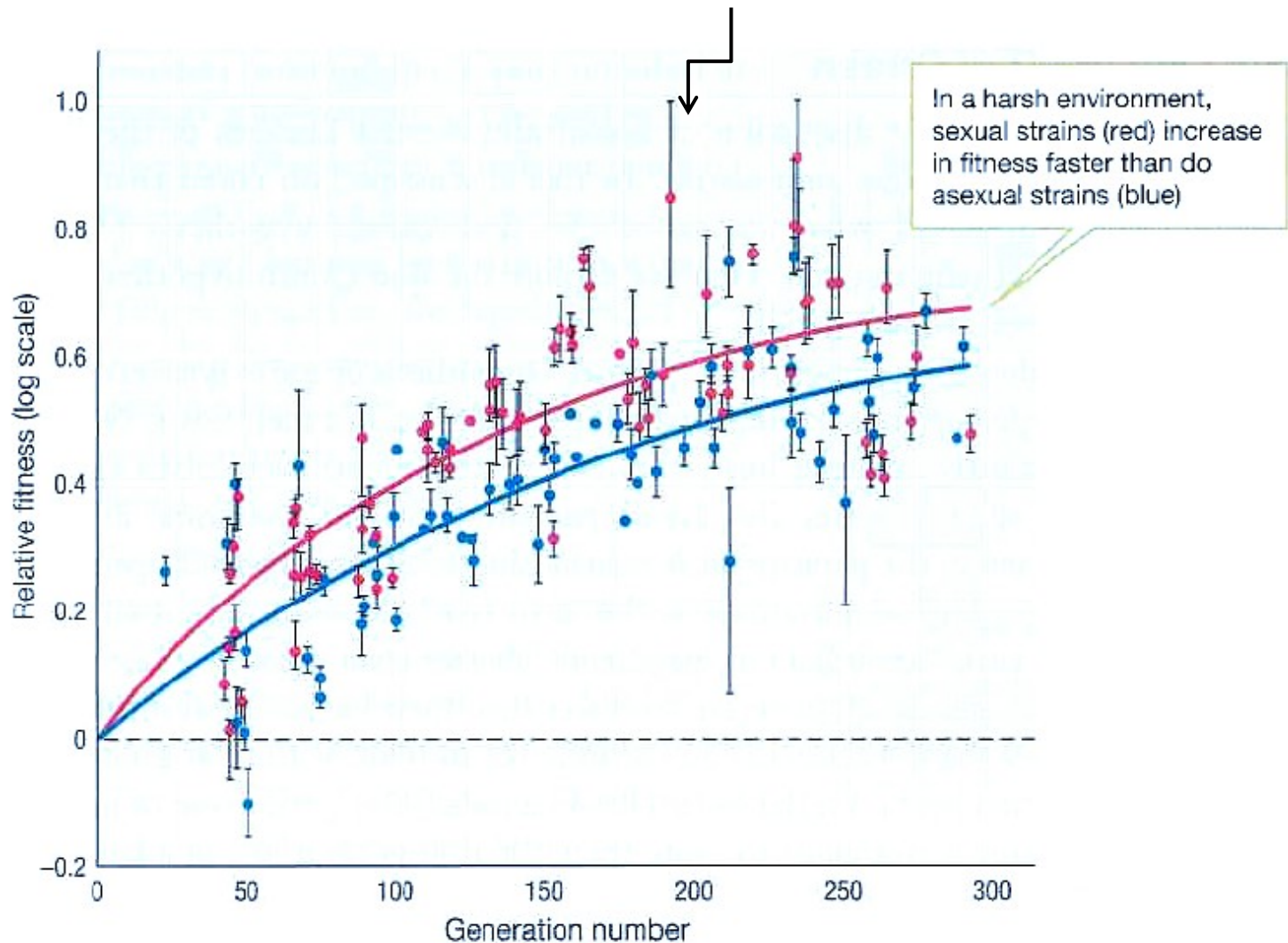
Sexual reproduction increases variation and hence rate of evolution but this advantage mostly in long-term perspective, asexuality in the short-term more advantageous

Eg.: yeast *Saccharomyces cerevisiae*

favourable environment: abundance of glucose, optimal temperature

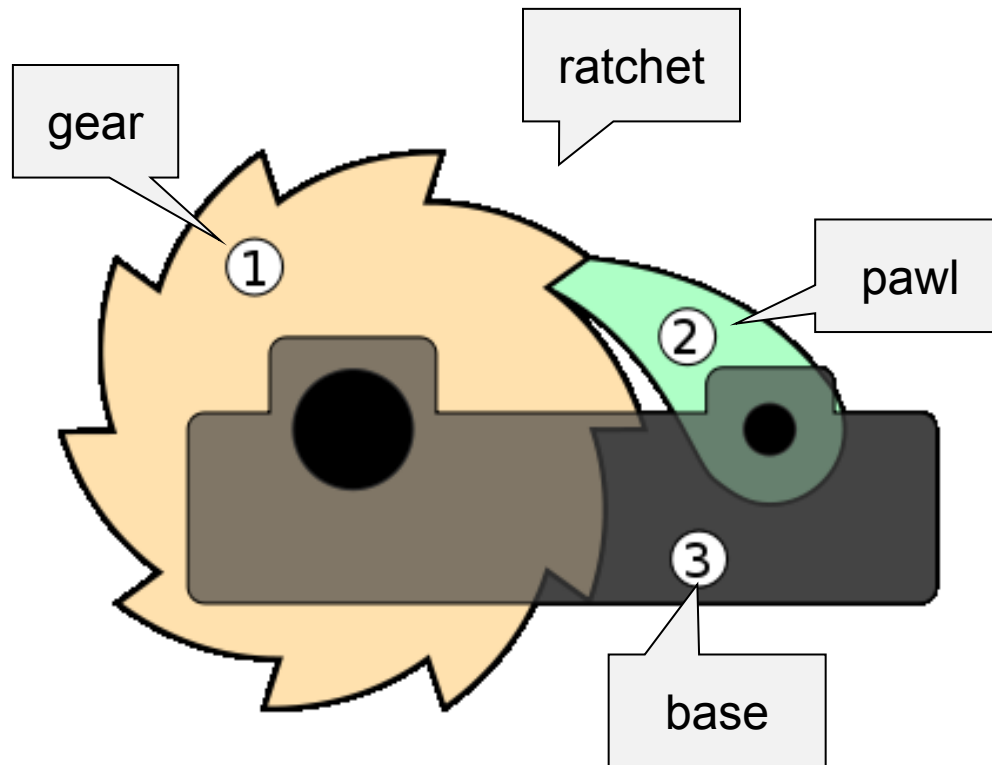
→ no difference

unfavourable environment: shortage of glucose, high temperature



# 1. Elimination of deleterious mutations I. Muller's ratchet:

The only way how to escape from deleterious mutations either  
back mutation, or  
mutation which invalidate effect of the previous mutation



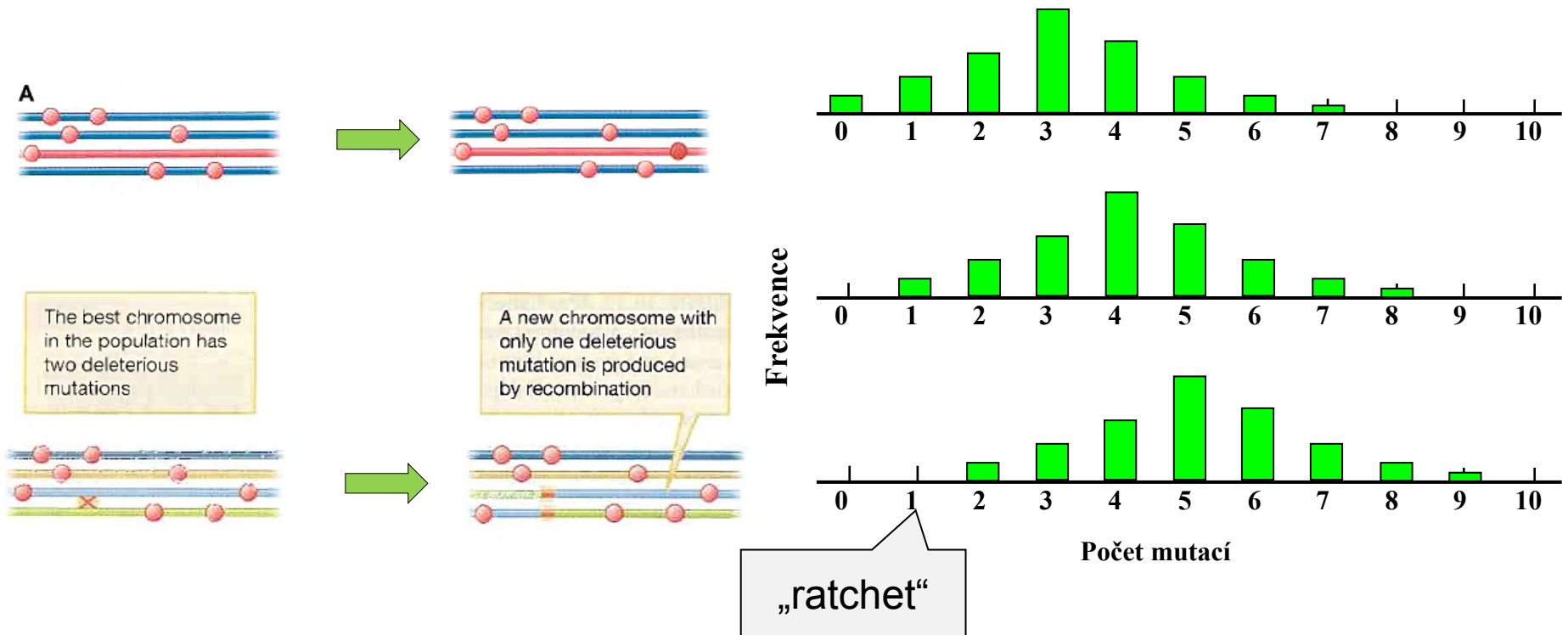
accumulation of deleterious mutations

small population  $\Rightarrow$  role of drift (stochastic process)

with sex chance to avoid „ratchet“

with increase of genotype frequencies without deleterious mutations  
spread of genes responsible for sex

best when mutations are only slightly deleterious





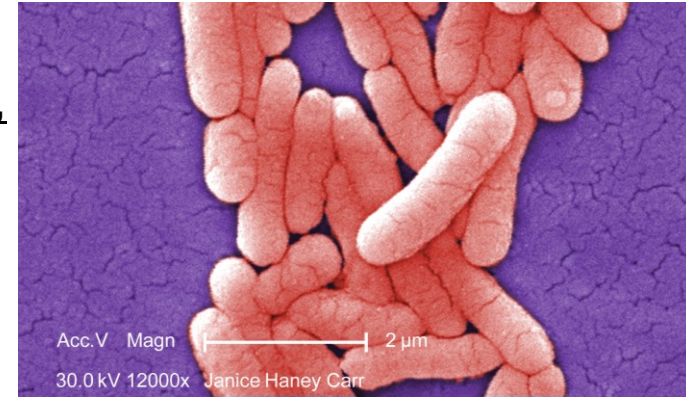
**Andersson and Hughes (1996) - *Salmonella typhimurium***

444 experimental cultures, each from 1 individual → growth overnight

repetition ⇒ repeated drift, total of 1700 generations

comparison with a free-living strain

→ 5 cultures (1%) with significantly reduced fitness, none with higher



**Lambert and Moran (1998) – comparison of fitness of bacteria living within insect cells with free-living species**

9 species of bacteria living only in insect cells

each species had its free-living relative counterpart

thermal stability of rRNA genes

did endosymbionts accumulated deleterious mutations?

→ in all cases rRNA of endosymbionts by 15 - 25% less stable

## 2. Elimination of deleterious mutations II. Kondrashov's model:

Alexey S. Kondrashov (1988)

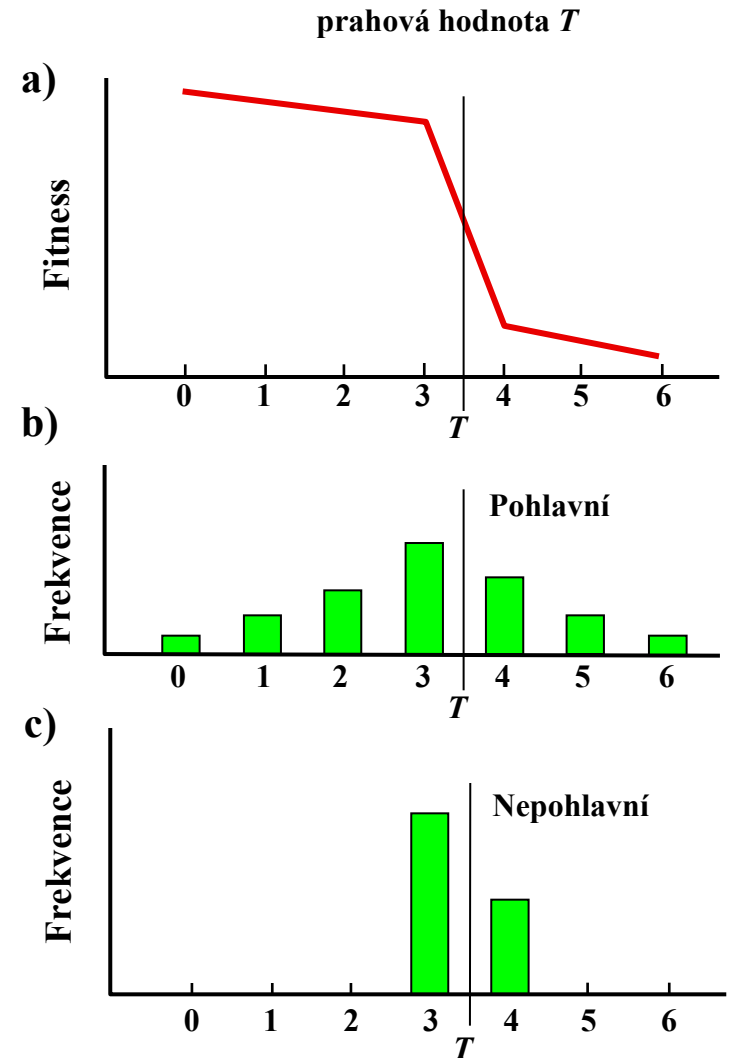
assumption that deleterious mutations act synergically → epistasis

„truncation selection“ (deterministic process)

since in sexuals proportion of deleterious mutations exceeding  $T$  value is higher than in asexuals, elimination of these mutations is faster in the former (recombination combines them)

question if frequencies of deleterious mutations are sufficiently high (at least  $1/\text{generation/genome}$ )

model proven in *E. coli* and *S. cerevisiae*



### 3. Unpredictable environment – lottery model, elm-oyster model

biotope divided into local sites to which descendants randomly „distributed“  
→ only best adapted ones survive, parents cannot know a priori which  
of them will do

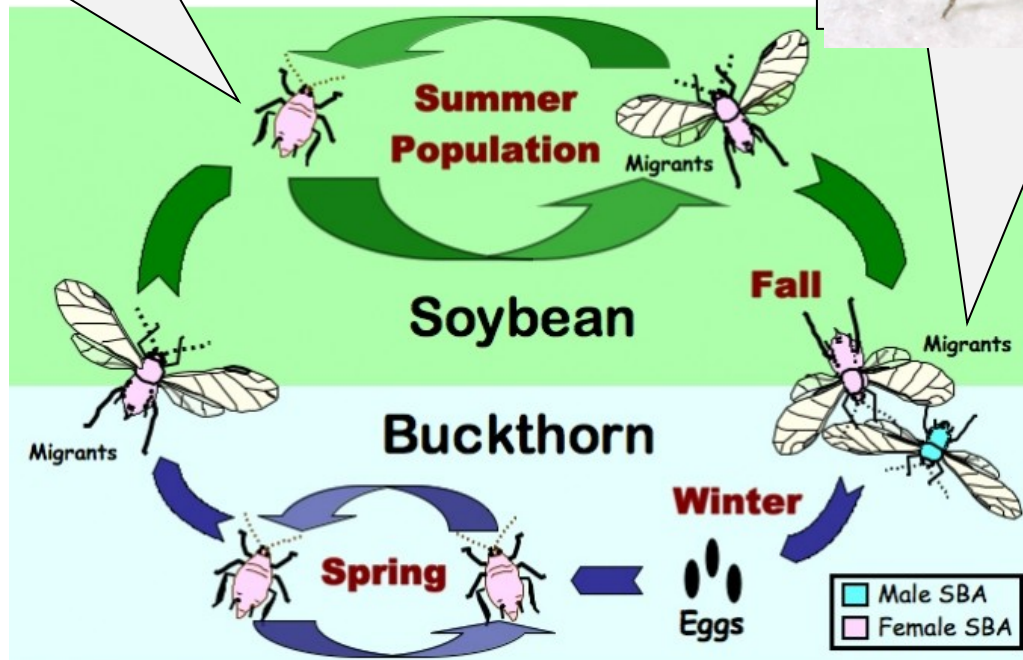
analogy with purchase  
of a lottery ticket



	1	2	3	4
	Temporal variation, predictable	Temporal variation, unpredictable	Spatial variation, predictable over time	Spatial variation, unpredictable over time
1				
2				
3				
4				
5				

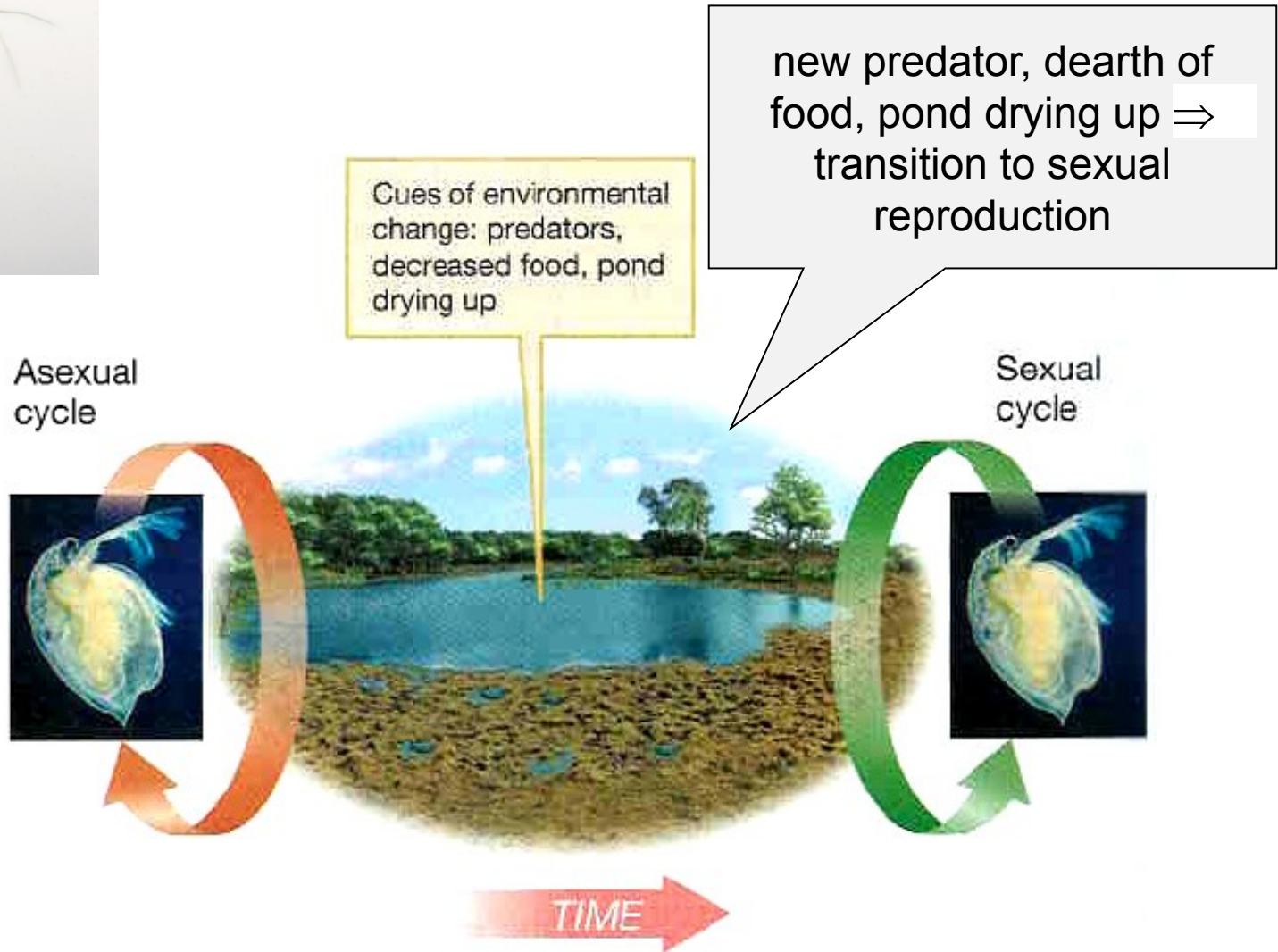


Eg. aphids:





Eg. *Daphnia*:



## 4. Unpredictable environment – elbow room model

assumption that in heterogenous and homogenous biotopes genotypes can differ in usage of limited sources

competition among siblings → more descendants of sexual parents can coexist at the same site because competition of asexual offspring is more intense

**Problem: models 3 and 4 are valid only for organisms with high fertility**

**Fluctuation of environment:**

itself does not maintain sex → fluctuation of epistasis necessary

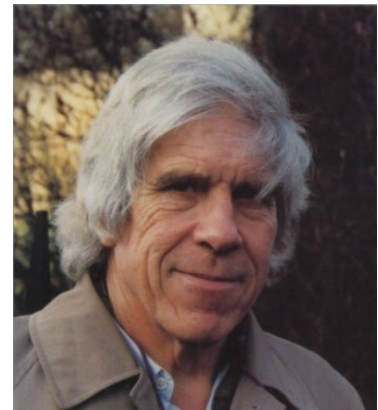
eg. 2 loci: alternatio of association cold-wet and warm-dry ↔ cold-dry and warm-wet

this model can work eg. in parasite-host interaction

## 5. Red Queen hypothesis

William D. Hamilton

based on the Red Queen hypothesis (Leigh Van Valen)



W.D. Hamilton



"The Red Queen has to run faster and faster in order to keep still where she is. That is exactly what you all are doing!"



L. Van Valen



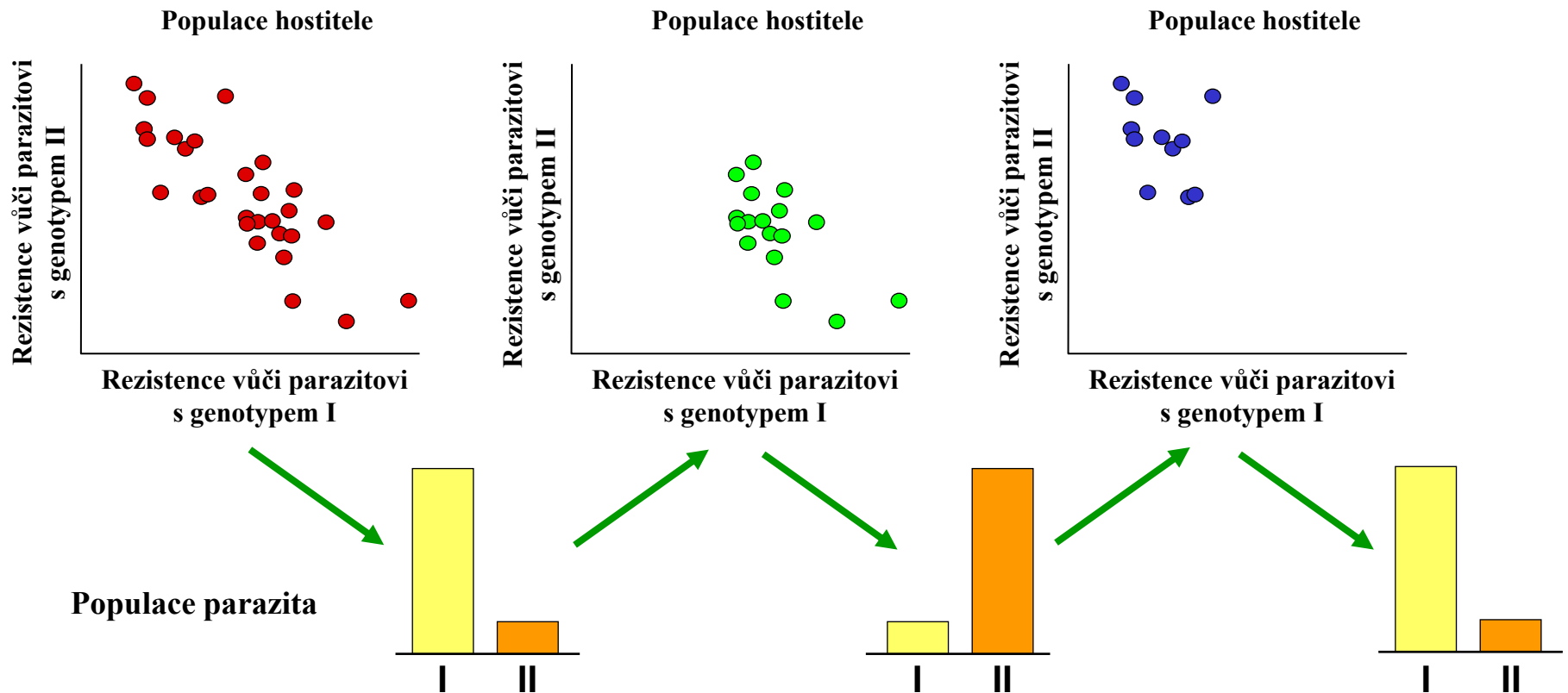
fluctuation of epistasis

fitness and gene frequencies cycles

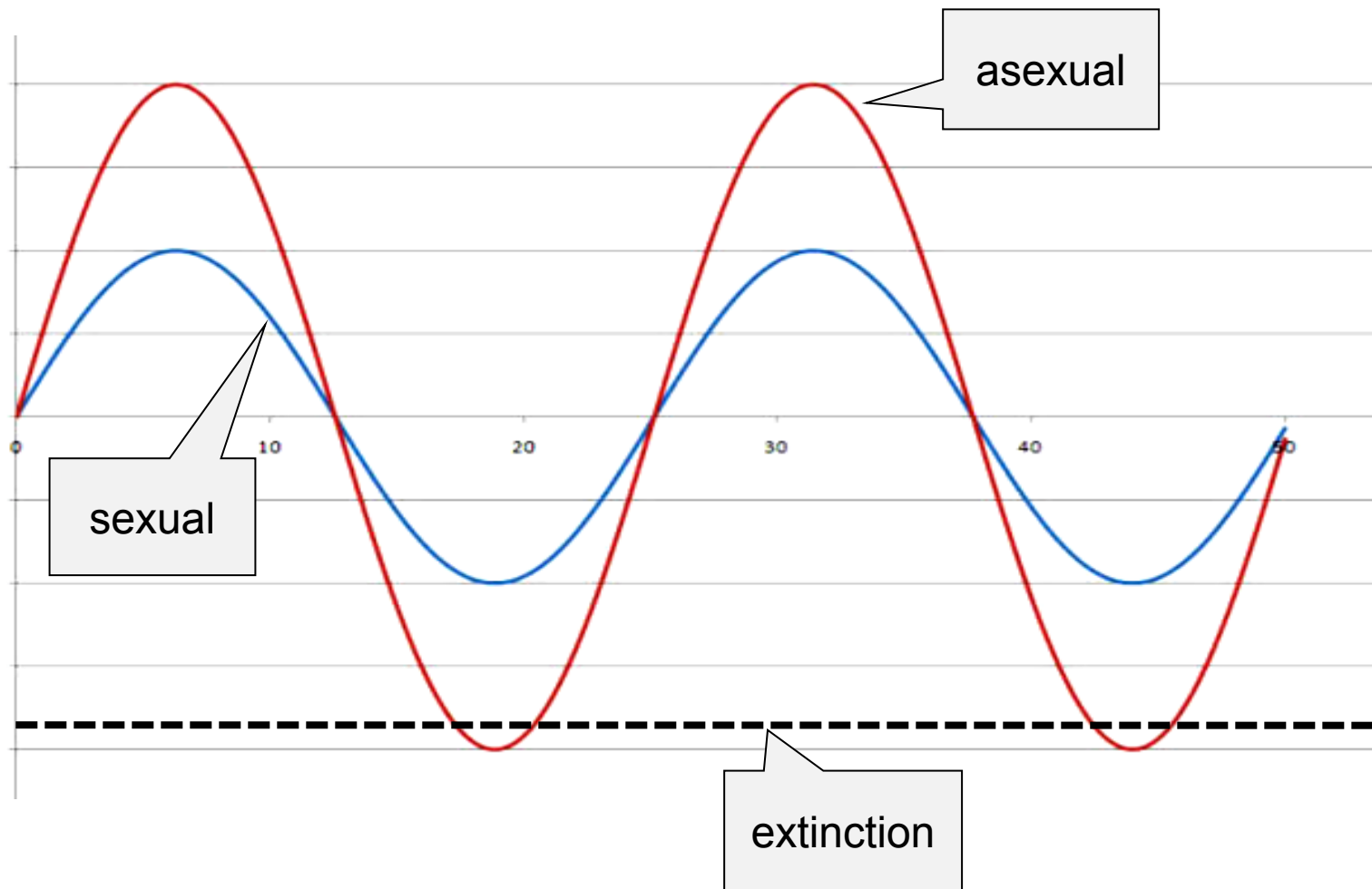
coevolution of parasite and host  $\Rightarrow$  arms races

multilocus „gene-for-gene“ relation

oscillation of gene frequencies higher in asexual individuals







model assumption: in heterogonous organisms (changing of sexual and asexual reproduction) and organisms with facultative sexuality  
**sexual reproduction more frequent in case of increased parasitation**

Curtis Lively (1992): freshwater gastropod *Potamopyrgus antipodarum*

New Zealand lakes and rivers

both sexual and asexual females



Lake Alexandria, South Island, New Zealand



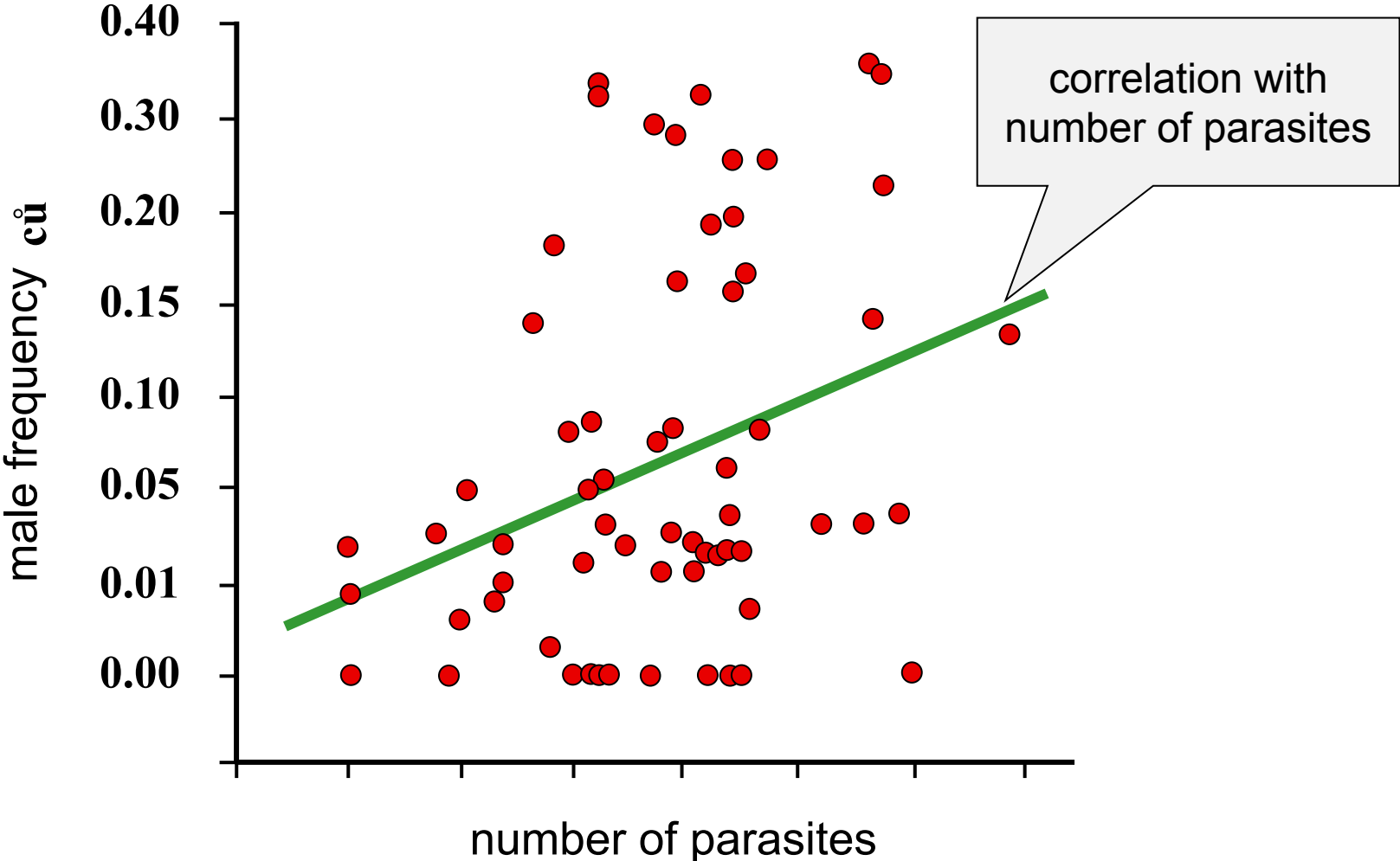
*Potamopyrgus antipodarum*

>12 parasitic trematode species (host castration  $\Rightarrow$  strong selection)

66 lakes

number of males as indicator of sexual reproduction

Lively et al. (1992):



# EVOLUTION OF SEX RATIO

sex ratio often 1:1 → why to waste for males?

R. A. Fisher (1930)

frequency-dependent selection

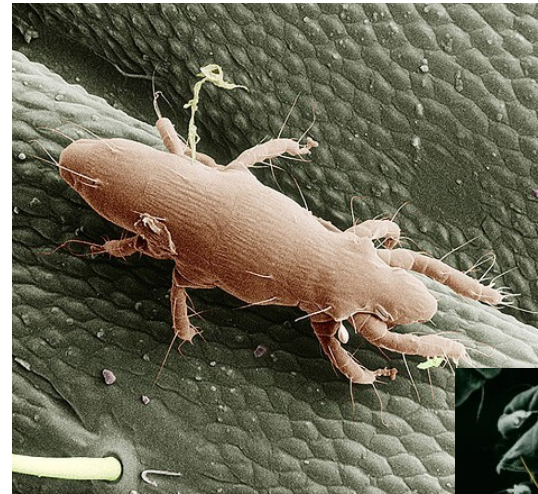
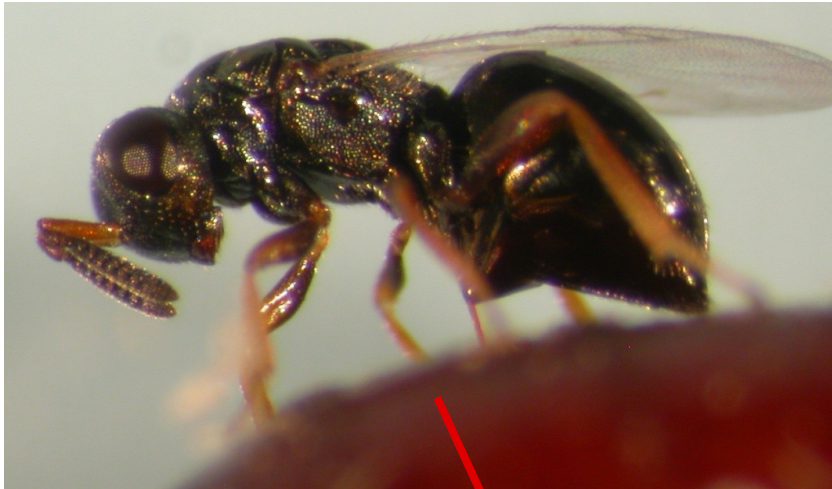
condition for validity of Fisher's argument:

1. random mating
2. same costs of both sexes

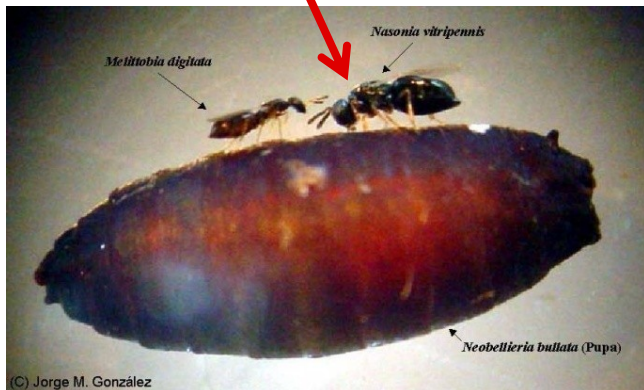
# ad 1) Local mating competition:

mites *Adactylidium*, *Pyemotes ventricosus*, *Acarophenax tribolii*

parasitoid wasps (eg. *Nasonia vitripennis*)



*Pyemotes ventricosus*

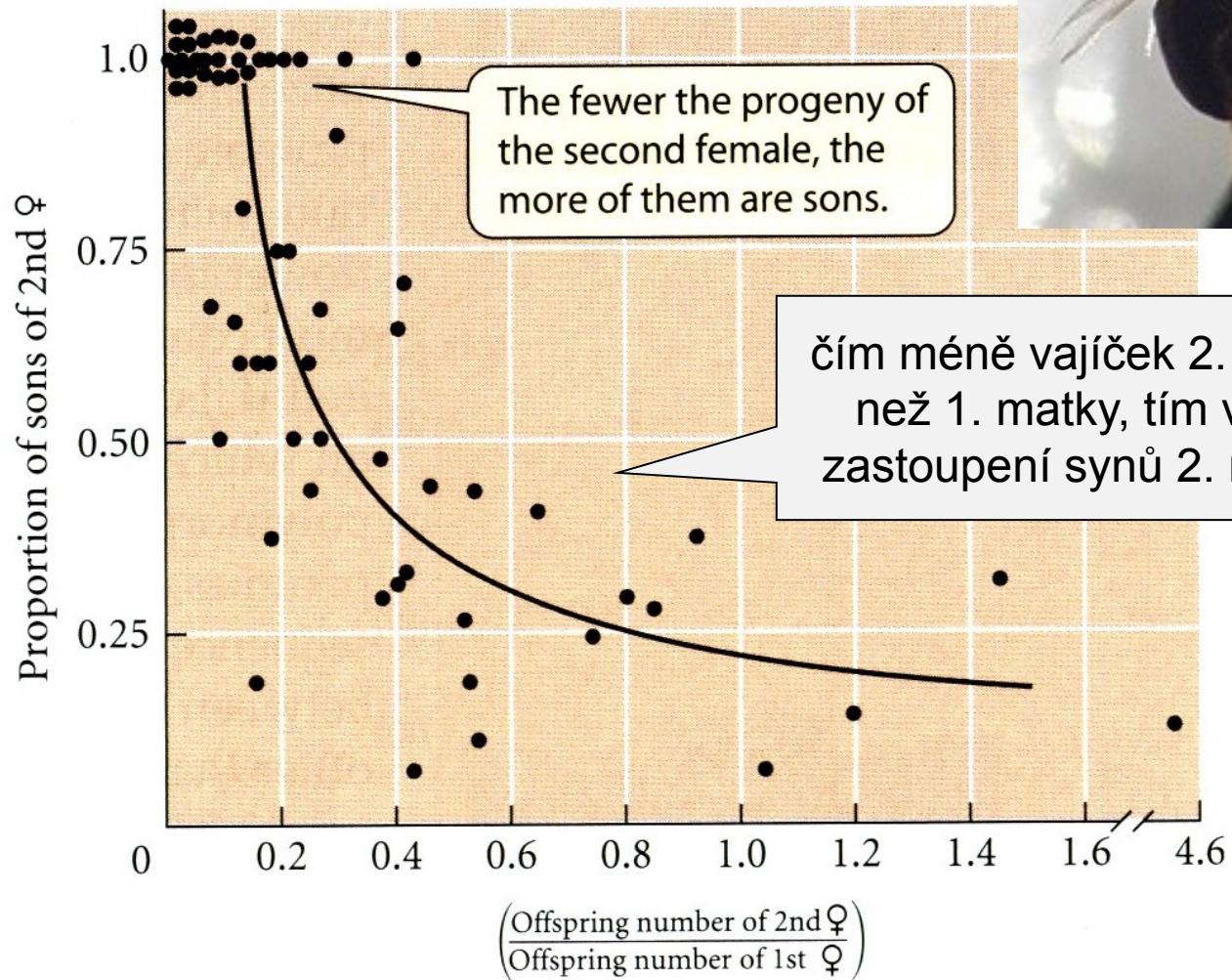


*Nasonia vitripennis*



*Acarophenax tribolii*

theoretical prediction: with increasing number of egg laying females  
percentage of sons increases



## ad 2) Trivers-Willard hypothesis:

Robert L. Trivers, Dan Willard

investment in sex ensuring higher fitness  
in next generation

dominant mother → investment in sons  
and vice versa

sex ratio bias or unequal parental investment

Eg.: deers



R.L. Trivers



D. Willard

