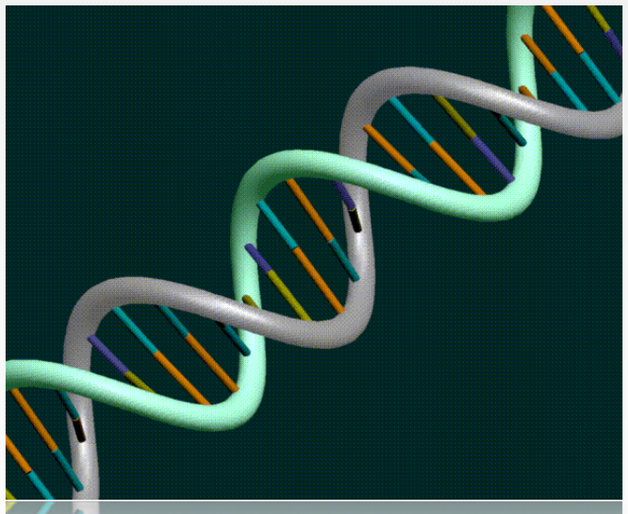
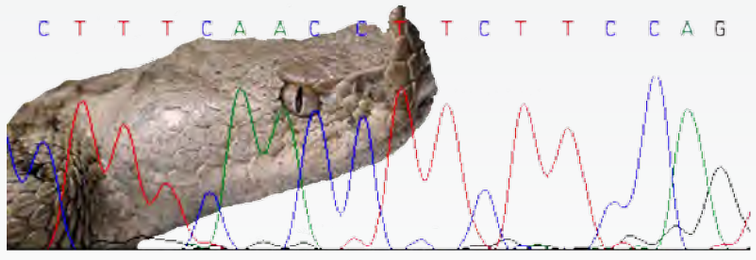
Conservation genetics

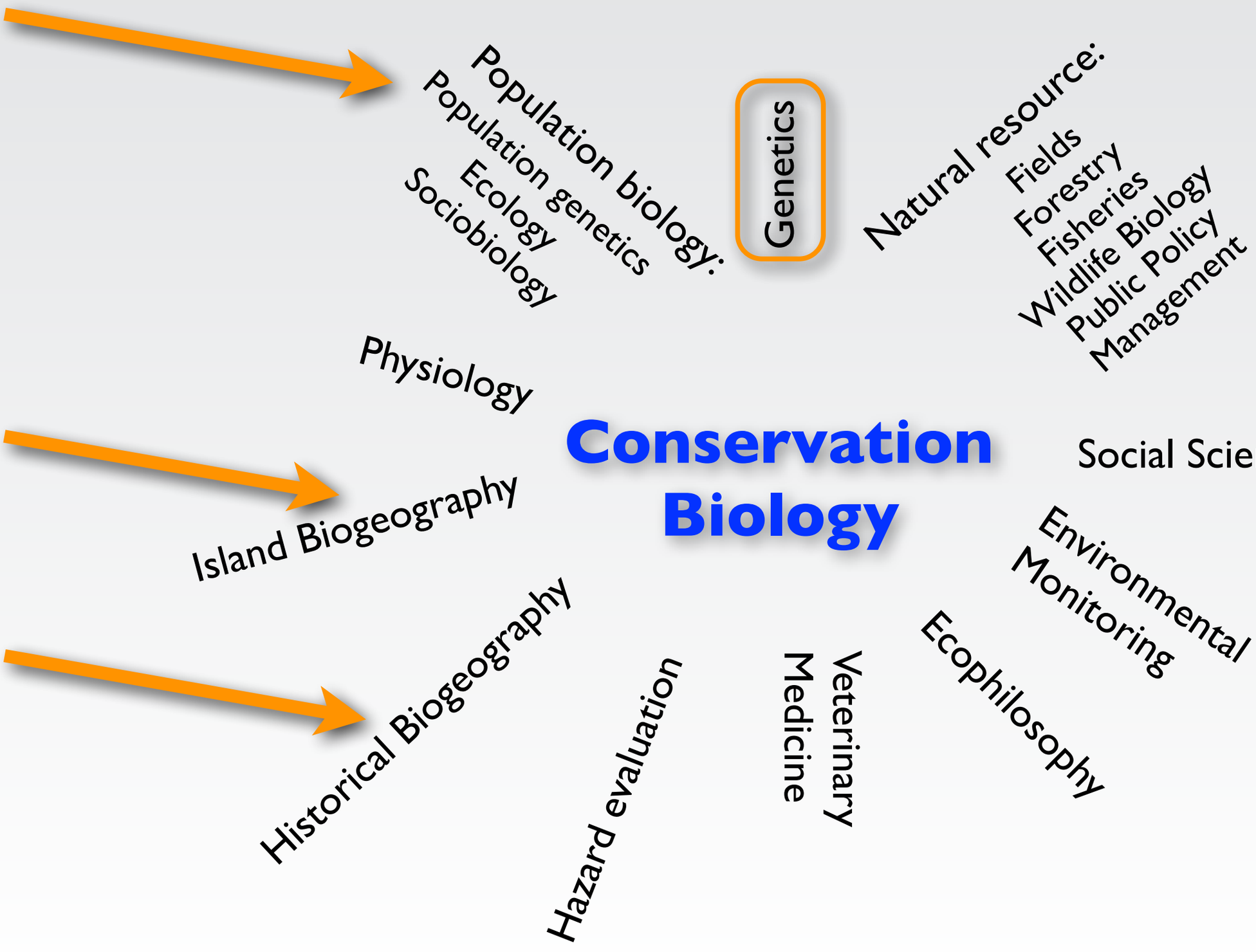




Sylvain Ursenbacher
NLU, room 36

UNI BASEL

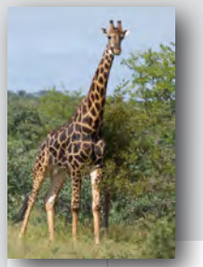
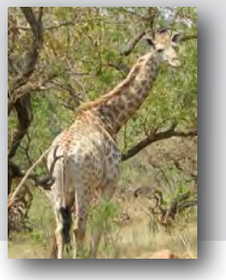
Introduction: *Conservation biology*

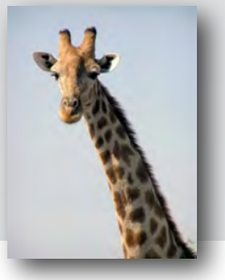


adapted from Soulé, 1985

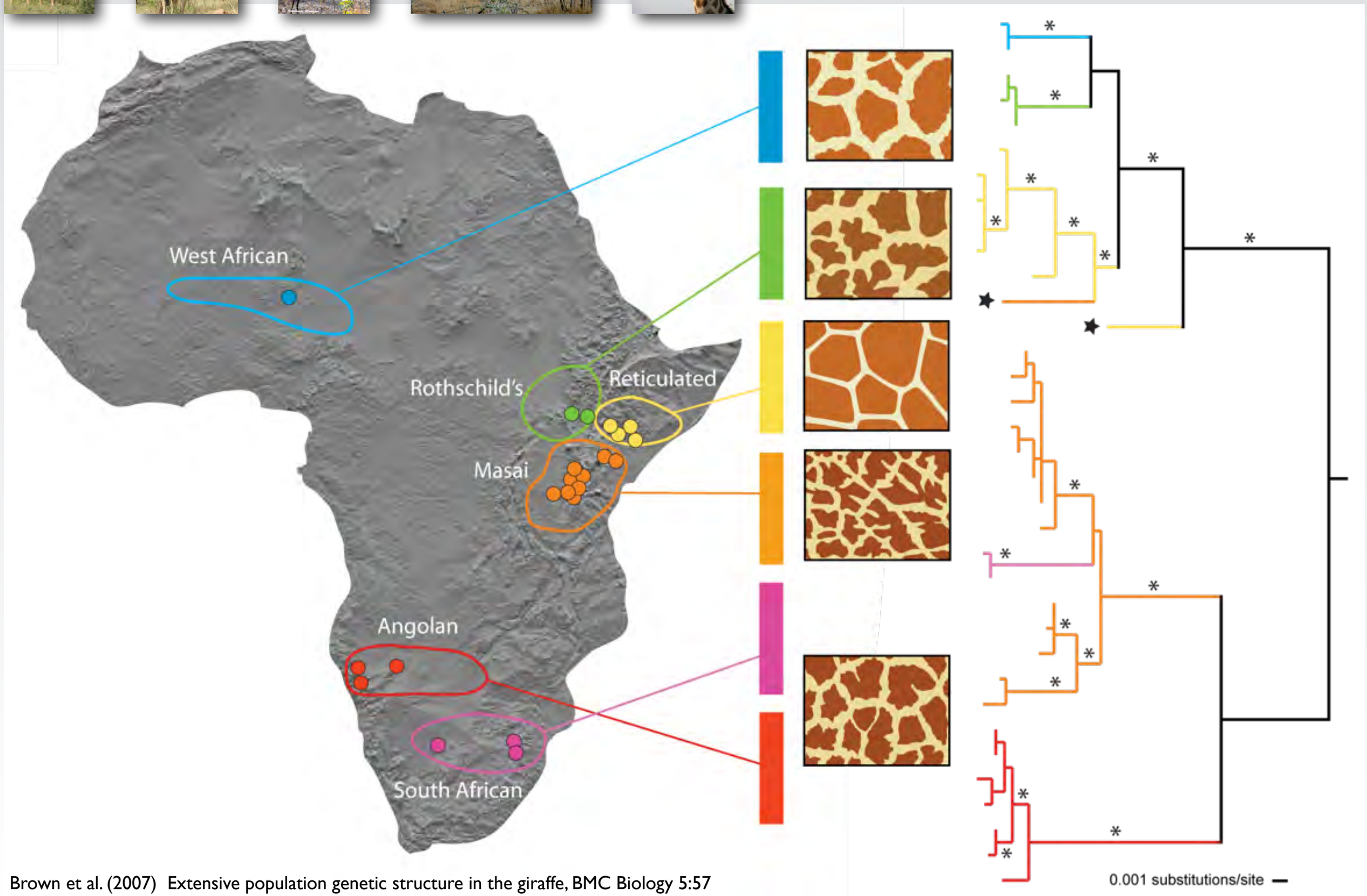
Introduction: *Conservation genetics*

- how genetic analyses can help threatened species: some examples...

Giraffa camelopardalis



Brown et al. (2007) Extensive population genetic structure in the giraffe, BMC Biology 5:57

Vipera berus

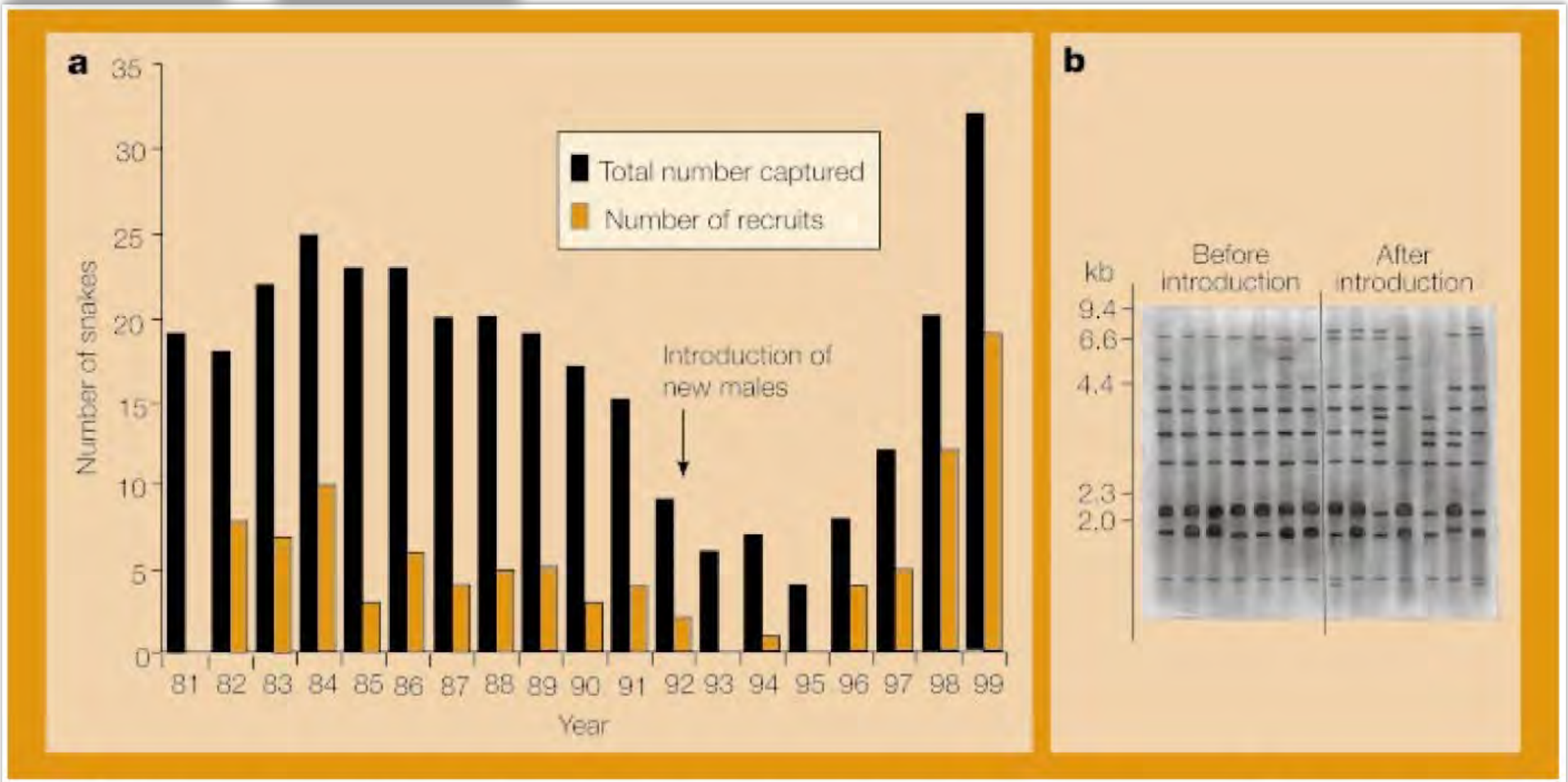


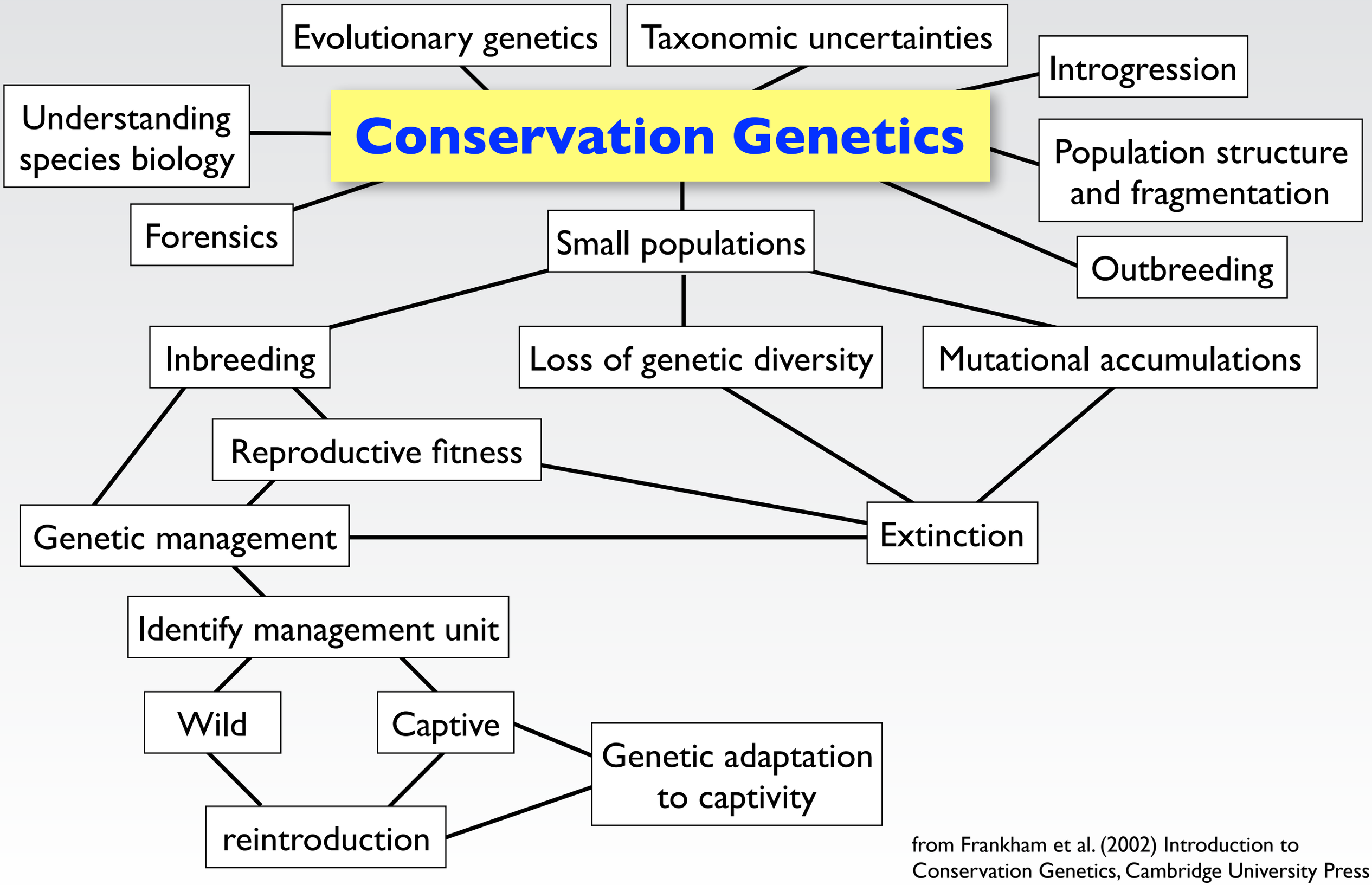
Figure 1 Introducing new males increases the genetic diversity and enables the adder population to recover. **a,** Total number and number of recruited male adders captured in Smygehuk from 1981 to 1999. **b,** Southern-blot analysis of major histocompatability complex (MHC) class I genes in seven males sampled before the introduction of new males (left) and in seven recruited males sampled in 1999 (right).

Madsen et al. (1999) Restoration of an inbred adder population, Nature 402, 34-35

Introduction: *Conservation genetics*

- how genetic analyses can help threatened species: some examples...
  - measure inbreeding / outbreeding depression
  - loss of genetic diversity
  - fragmentation of population and reduction of gene flow
  - genetic drift
  - define management unit
  - understand aspects of species biology important for their conservation

Introduction: *Conservation genetics*



from Frankham et al. (2002) Introduction to Conservation Genetics, Cambridge University Press

Introduction: *why genetic diversity is important in populations...*

- genetic diversity reflects evolutionary potential
  - genetic diversity required to evolve or to adapt to new environment or environmental modifications.
  - ↗ genetic difference between individual ⇒ ↗ fitness of the most adapted

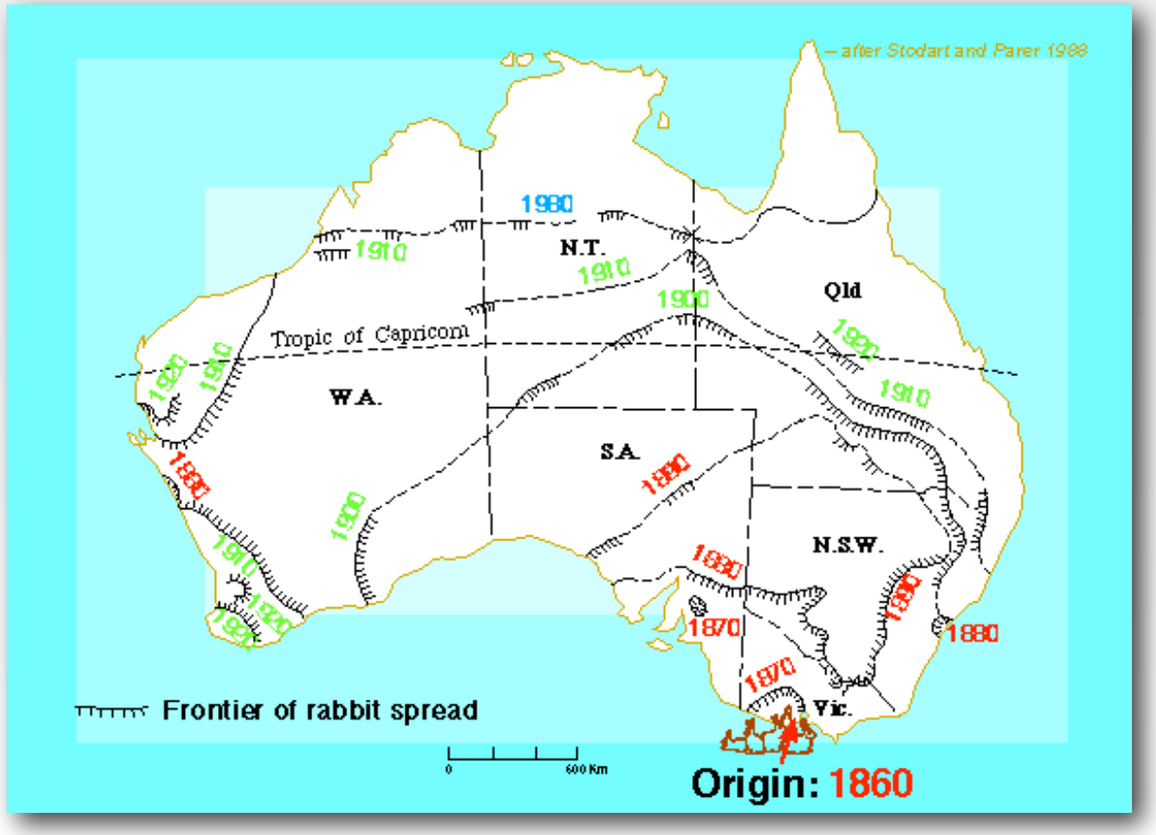
Introduction: *why genetic diversity is important in populations...*

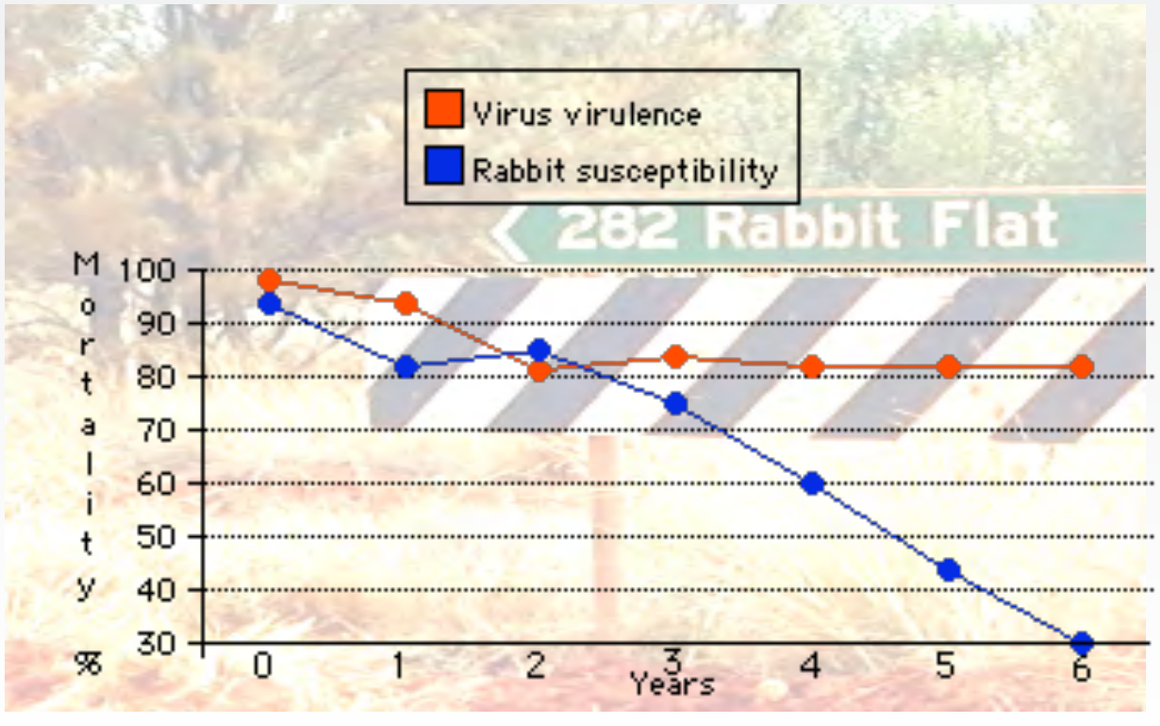
- genetic diversity reflects evolutionary potential
  - example 1 - habitat selection: peppered moth (*Biston betularia*) in UK
    - dark and light forms
    - night: active / day: resting on trees
    - ➡ camouflage critical for survival
    - light form: camouflaged on lichen-covered tree trunks
    - Industrialisation: kill lichen by sulphur pollution
    - ➡ light form: visible / dark form: camouflaged

Grant (1999) Fine tuning the peppered moth paradigm, Evolution 53, 980-984
Kettlewell (1973) The Evolution of Melanism, Clarendon Press, Oxford, UK
Majerus (1998) Melanism: Evolution in Action. Oxford University Press
Kettlewell (1958) A survey of the frequencies of *Biston betularia* (L.) (Lep.) and its melanic forms in Great Britain, Heredity 12, 551-572
but see also: Rudge (2006) Myths about moths: a study in contrasts, Endeavour 30, 19-23

Introduction: *why genetic diversity is important in populations...*

- genetic diversity reflects evolutionary potential
  - example 2 - disease resistance: resistance to myxoma virus in Australian rabbits
    - introduction of rabbits in Australia: 1860
    - control measure: introduction of myxoma in 50'
    - ➡ high mortality rate first years
    - high selection for resistance





Introduction: *why genetic diversity is important in populations...*

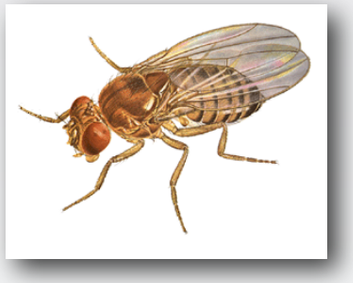
- genetic diversity reflects evolutionary potential
  - genetic diversity required to evolve or to adapt to new environment or environmental modification.
  - ↗ genetic difference between individual ⇒ ↗ fitness of the most adapted
- loss of genetic diversity often associated with inbreeding, reduction of reproductive fitness and extinction risk

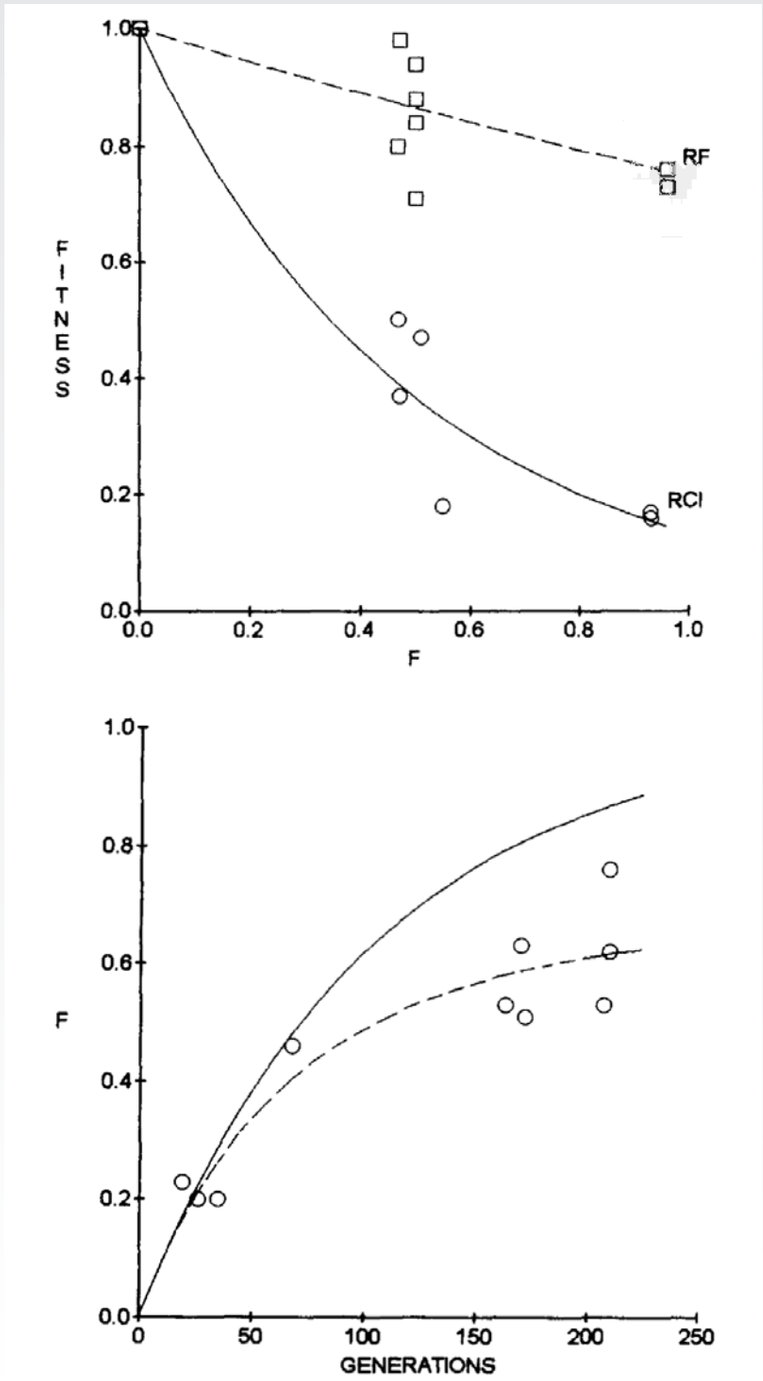
Introduction: *why genetic diversity is important in populations...*

- loss of genetic diversity often associated with inbreeding, reduction of reproductive fitness and extinction risk

  - example 3 - captive fruit fly/housefly populations

    - 60 captive fruit fly (*Drosophila melanogaster*) populations, maintained during 210 generations
      mean population size: 67 individuals
    - ➡ 15/60 populations extinct after 210 generations
    - 6 captive housefly (*Musca domestica*) populations, maintained during 64 generations
      population size: 50 individuals
    - ➡ 5/6 populations extinct after 64 generations





Latter & Mulley (1995) Genetic Adaptation to Captivity and Inbreeding Depression in Small Laboratory Populations of *Drosophila melanogaster*, Genetics 139, 255-266
Reed & Bryant (2000) Experimental tests of minimum viable population size, Animal Conservation 3, 7-14

Introduction: *why genetic diversity is important in populations...*

- loss of genetic diversity often associated with inbreeding, reduction of reproductive fitness and extinction risk
  - example 4 - large metapopulation (Finland) of the Glanville fritillary butterfly (*Melitaea cinxia*)
    - 42 butterfly populations genotyped in 1995
    - survival and extinction recorded in 1996
    - ➡ 36 populations survived
    - extinction rate high for populations with lower heterozygosity even corrected for demographic and environmental variables (pop. size, area, ...)

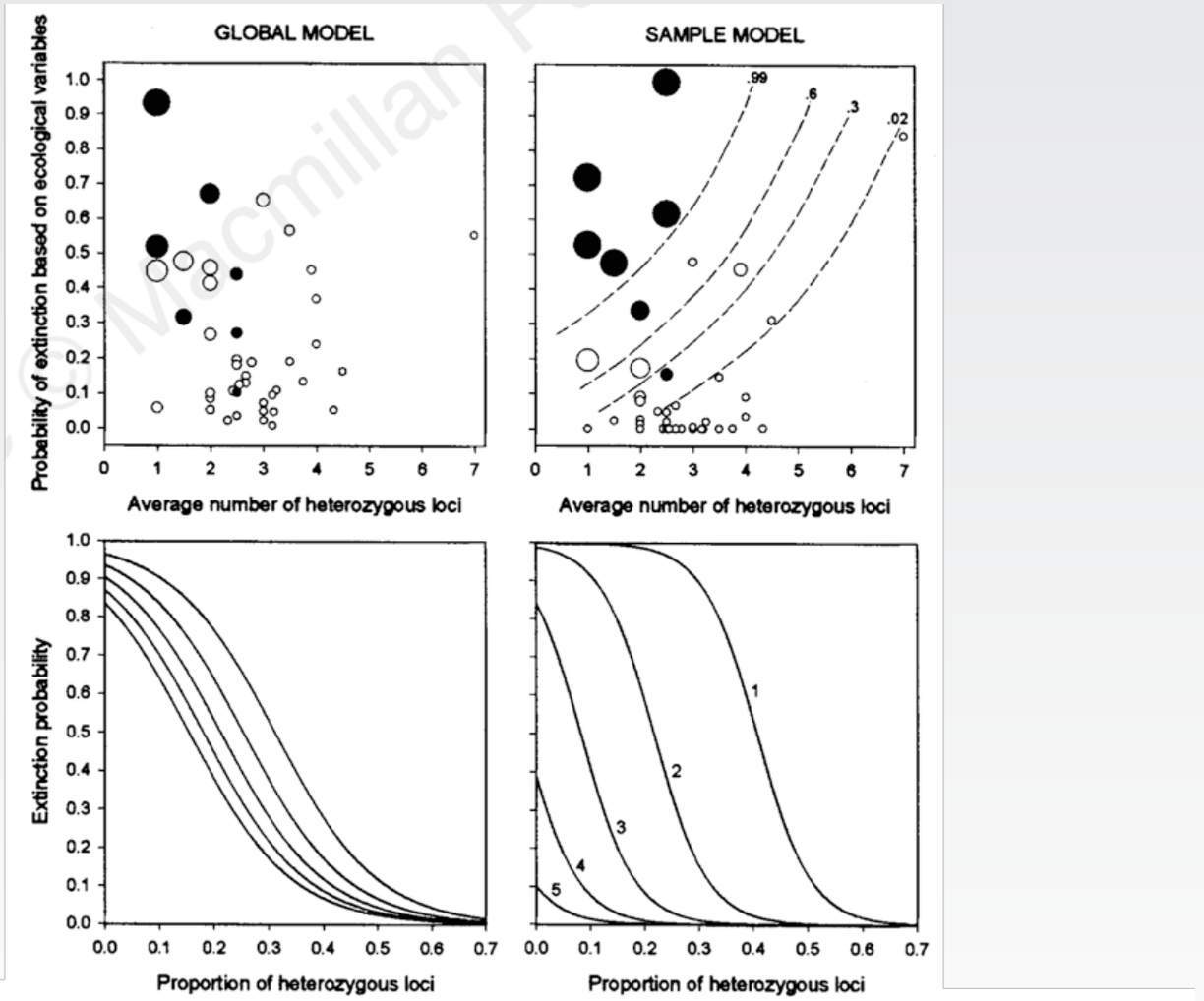
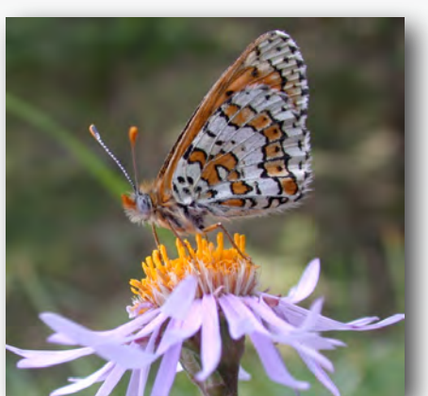
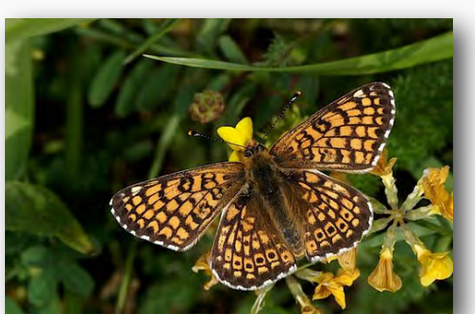


Figure 2 For both global and sample models (Table 1), the upper panels show: (1) the observed average number of heterozygous loci in extinct (black) and surviving (white) populations; (2) the probability of extinction predicted by the models without heterozygosity compared with the observed heterozygosity; (3) the probability of extinction predicted by the full model, including heterozygosity (proportional to circle size). For the sample model, we have drawn appropriate isoclines for the extinction risk predicted by the model, including ecological factors and heterozygosity. These figures illustrate that both the ecological factors and heterozygosity influence the extinction risk (for statistical analysis, see Table 1). Lower panels show the relationship between the risk of local extinction and heterozygosity predicted by the global and sample models (Table 1). Model predictions are shown for local population sizes of 1–5 larval groups, fixed at the lower quartile value of change in regional density (N_{trend}) and the lower quartile value of meadow area in the global model; and fixed at the lower quartile value of regional density (N_{neigh}) and median flower abundance in the sample model.





Saccheri et al. (1998) Inbreeding and extinction in a butterfly metapopulation, Nature 392, 491-494

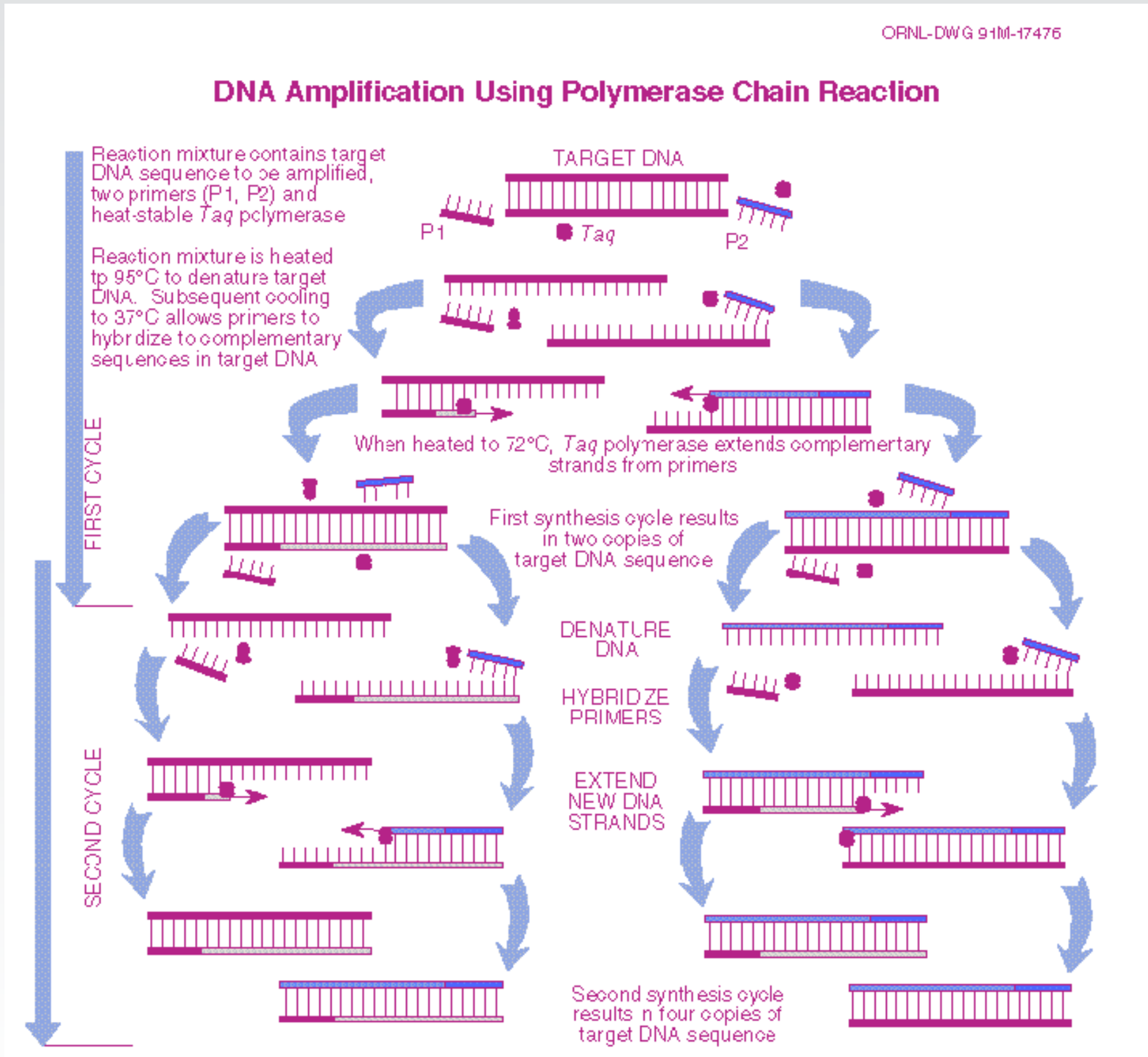
Genetic tools: *DNA sampling*

- invasive methods (dead animals)
  - entire animal/plants (e.g. insects)
  - internal tissue: liver, heart, ...
- non-invasive methods
  - blood sample
  - part of the body: hairs, feathers, scales, sloughed skin, ... leafs, flowers, ...
  - buccal swab
  - faeces
  - ...

Genetic tools: *DNA extraction*

- first: lysis of the tissue/sample using proteinase
- numerous protocols
  - standard Phenol/Chloroform (Sambrock et al. 1989)
    ⊕ low cost / ⊖ high toxicity
  - CTAB: more adapted to plants (or amphibians)
  - CHELEX:
    ⊕ quick / ⊖ not for long storage
  - columns: several companies, e. g. Qiagen, Promega, Sigma,...
    ⊖ expensive / ⊕ high purity DNA

Genetic tools: *DNA amplification (PCR)*



Source: *DNA Science*, see Fig. 13.

Measuring genetic diversity

- different markers (regions)
  - under selection or not
  - lineage: maternal, paternal or both
  - easy/difficult to develop, use or analyse
  - cheap/expensive

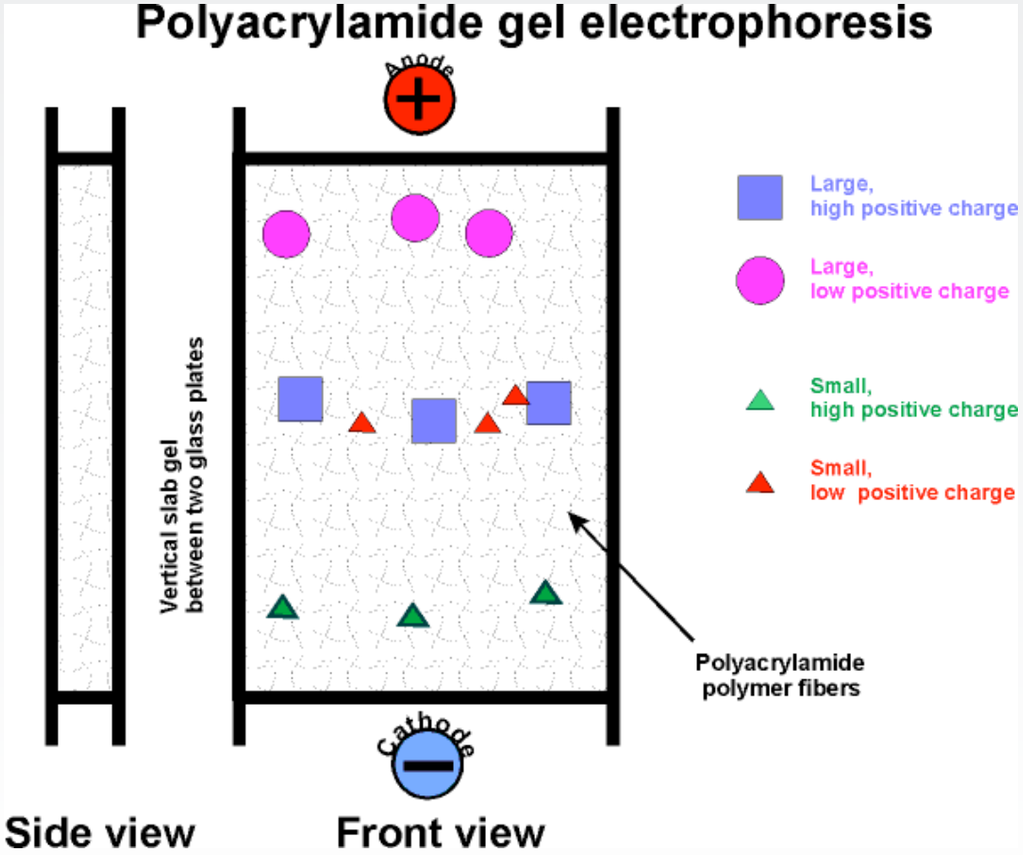
  - Proteins / Allozymes
  - sequencing
  - Restriction Fragment Length Polymorphism (RFLP)
    Amplified Fragment Length Polymorphism (AFLP)
    Random Amplified Polymorphic DNA (RAPD)
    DNA fingerprints (minisatellites)
  - microsatellites
    Single Nucleotide Polymorphism (SNP)
    Single Strand Conformational Polymorphism (SSCP)

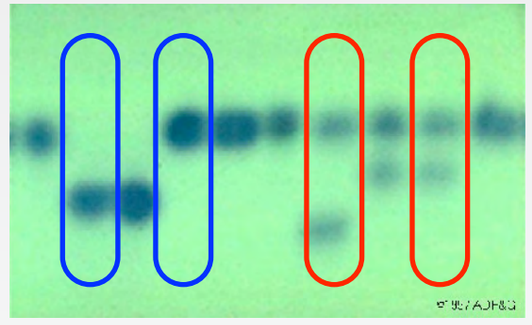
Genetic markers: *Proteins / Allozymes*

- separate proteins according to their electric charge and molecular weights

DNA coding for a protein	...ATG CTT GAC GTT ...	...ATG CTT G**G**C GTT ...
mRNA	...AUG CUU GAC GUU ...	...AUG CUU G**G**C GUU ...
amino acid composition	... -met - leu - asp - val - ...	... - met - leu - **gly** - val - ...

- electrophoresis



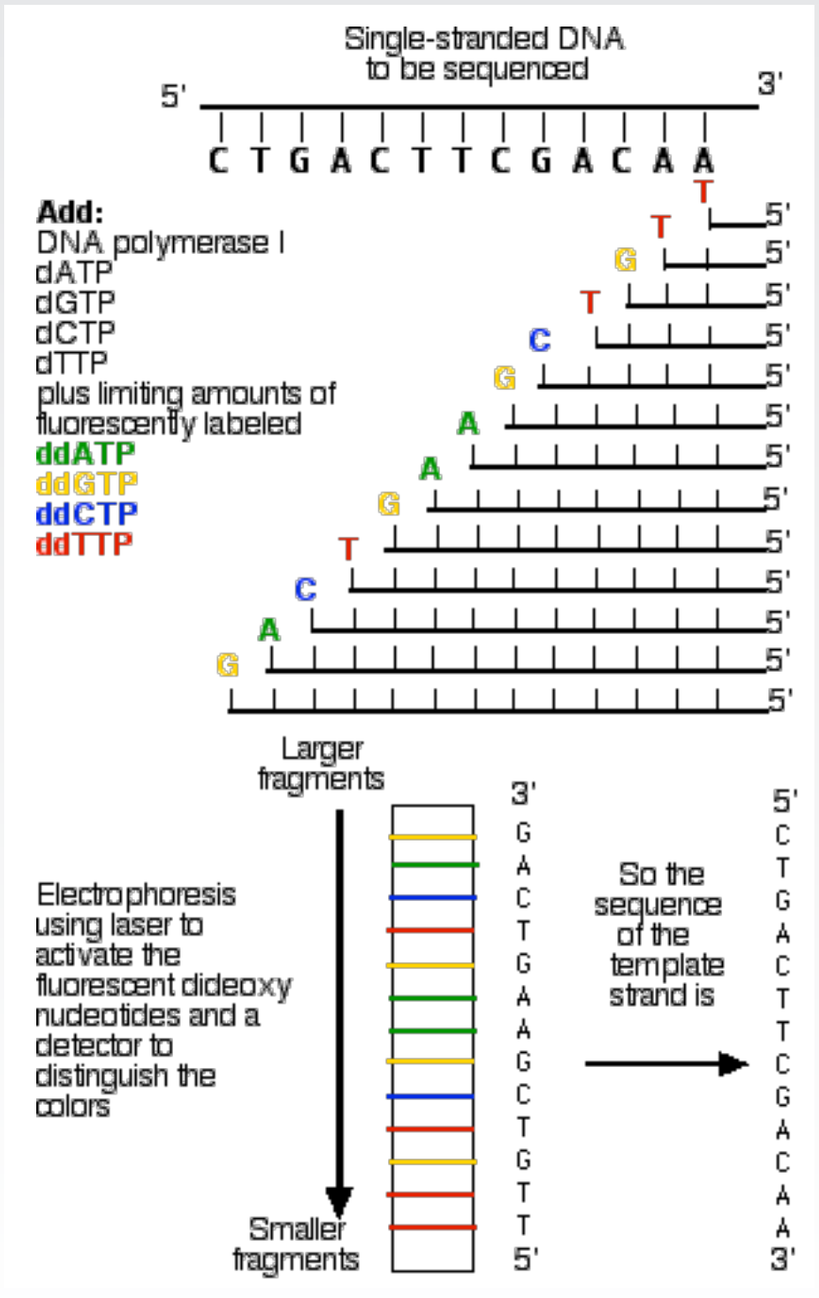


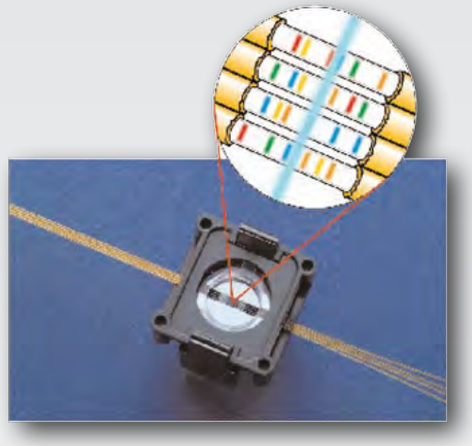
⊖ only 30% of DNA base changes result in charge changes: underestimation of genetic diversity

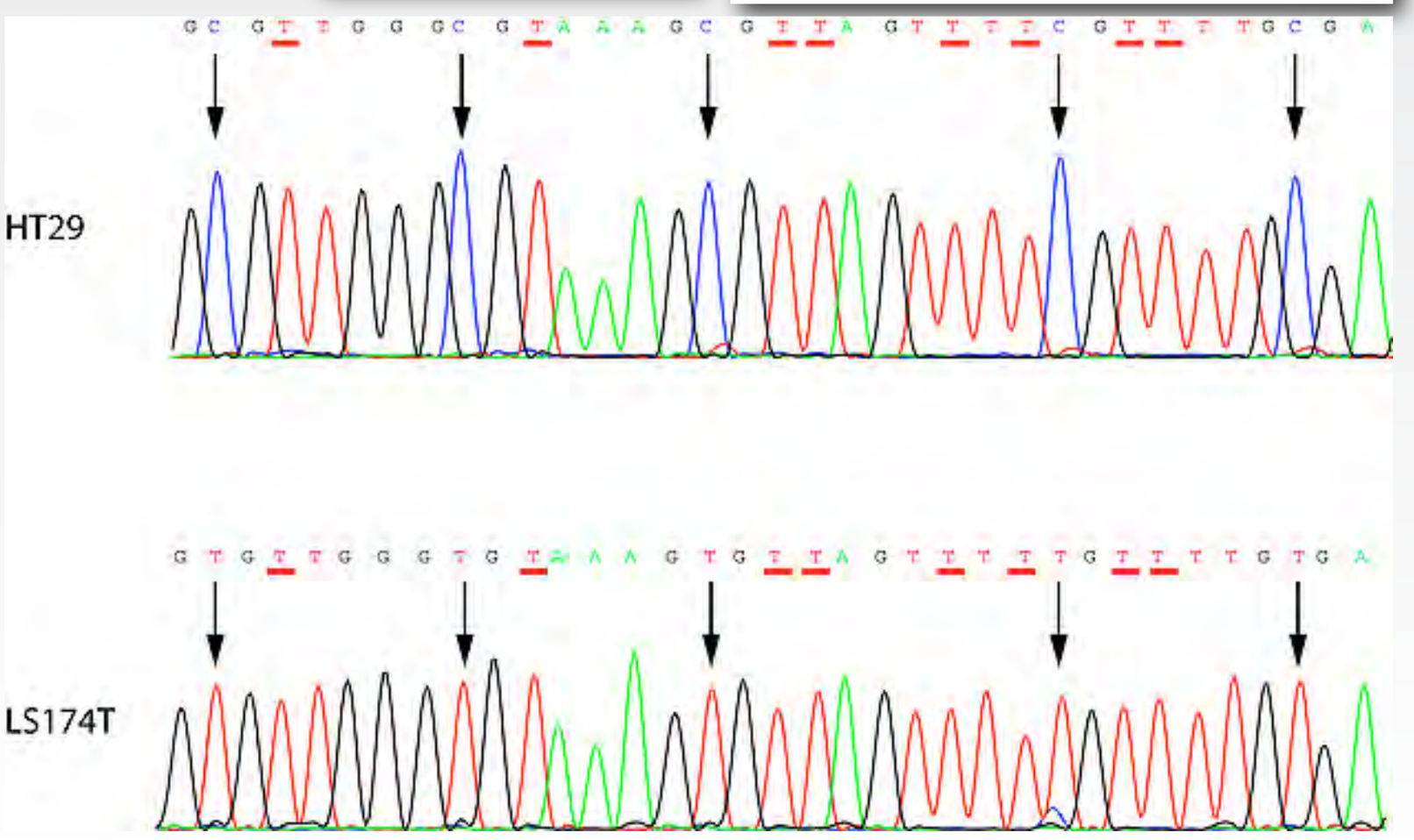
⊖ need high quantity of material (blood, kidney, liver) not really useful for endangered species

Genetic markers: *Sequencing*

- "reading" DNA sequences





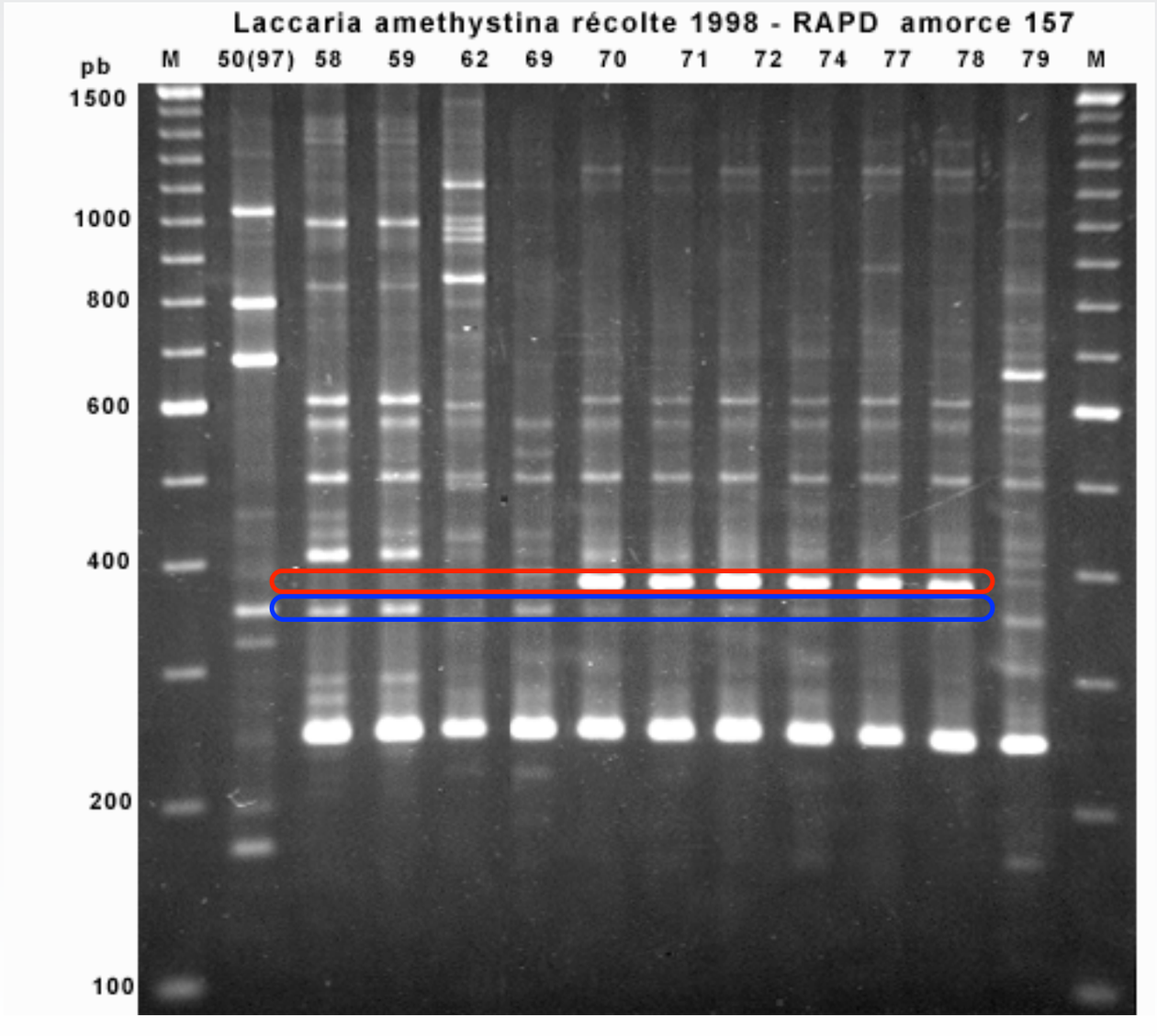


Genetic markers: *Sequencing*

- "reading" DNA sequences
  - ⊖ high cost
  - ⊖ problems with heterozygosity
  - ⊖ primers sequences must be known

Genetic markers: *Random Amplified Polymorphic DNA (RAPD)*

- PCR reaction using random primers (10-20 bp), producing several fragments with different length
- electrophoresis to see the different fragments



⊖ repeatability of the results not always good...

⊖ dominant markers

Genetic markers: *dominance / co-dominance principle*

AA	Aa	aa
---P___P----	---P___P----	---P----------
---P___P----	---P----------	---P----------
DNA fragment on a gel		
___	___	no band

A: dominant / a: recessive
P: primer similar to the DNA seq.
_____ PCR product

- dominance: when heterozygotes are not distinguishable from homozygotes
  - AA with PCR product, aa without PCR product, Aa with PCR product
- co-dominance: when heterozygotes are distinguishable from homozygotes
  - AA with low mobility, aa with high mobility, Aa with a medium mobility

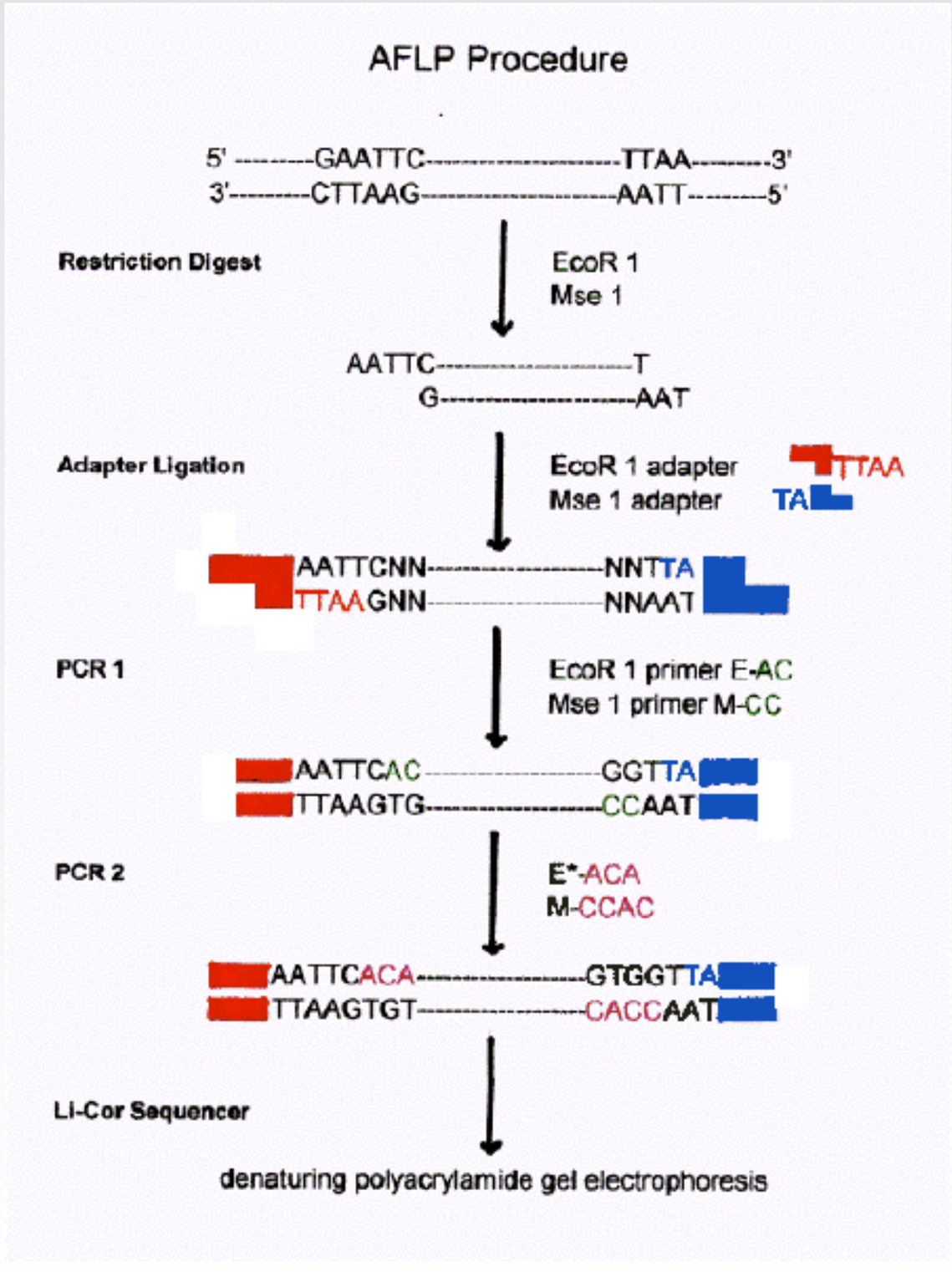
→ difficulties in the analyses

Genetic markers: *Amplified Fragment Length Polymorphism (AFLP)*

- method close to RAPD
- DNA cut with a restriction enzyme, and short DNA fragments of known sequence are attached to the cut ends

  ⊕ more accurate than RFLP
  no repeatable problems

  ⊖ dominant markers



Genetic markers: *microsatellites*

- also named STR (short tandem repeats) or
  SSR (simple sequence repeats)

- tandem repeats of a short DNA segment (1-5 bp)

  maternal origin ATATATATATATATATATAT $(AT)_9$

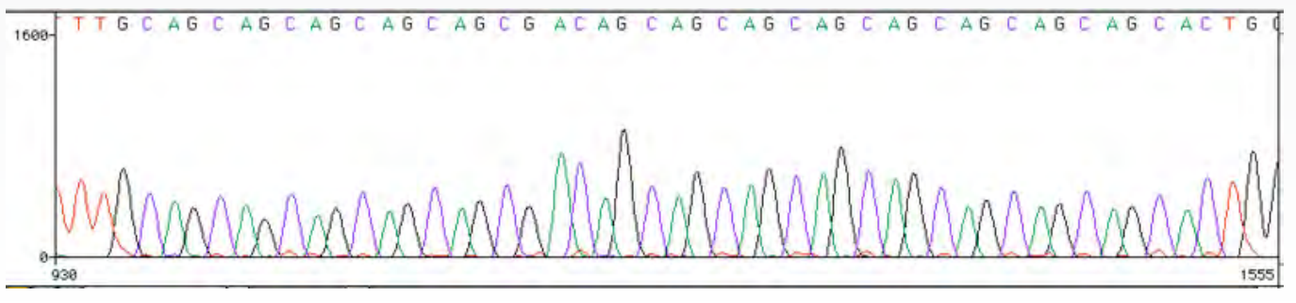
  paternal origin ATATATATATATATATATATATAT $(AT)_{11}$

- between two conserved regions flanking the microsatellites

  stable ATATATATATATATATATAT stable

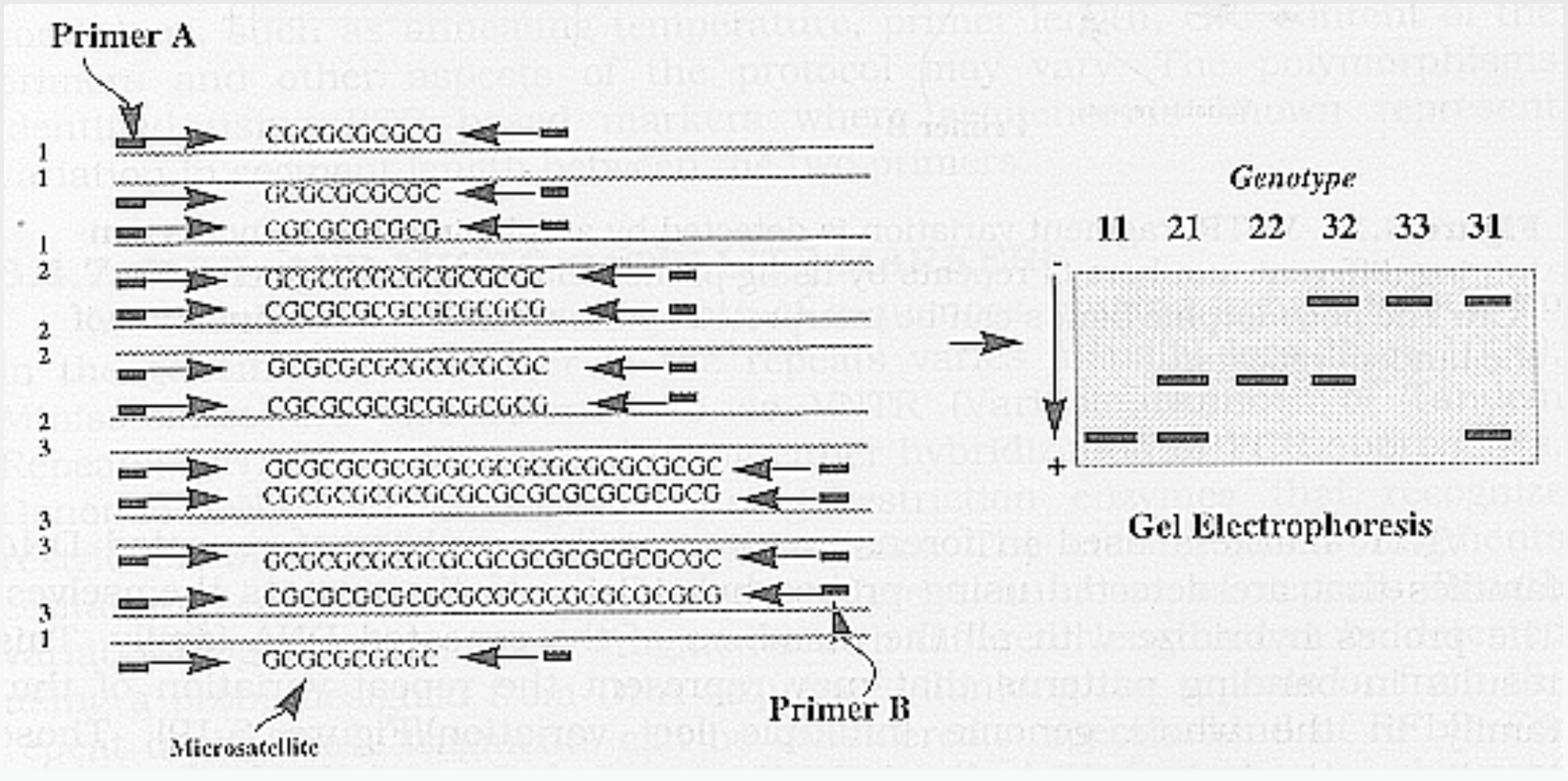
  stable ATATATATATATATATATATATAT stable

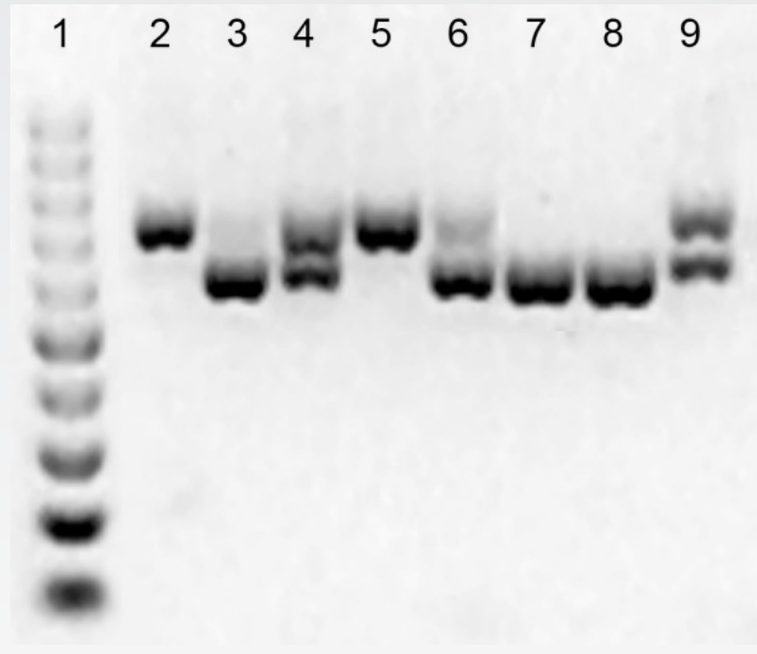
- reason of the polymorphism: polymerase "slippage" or "stuttering"

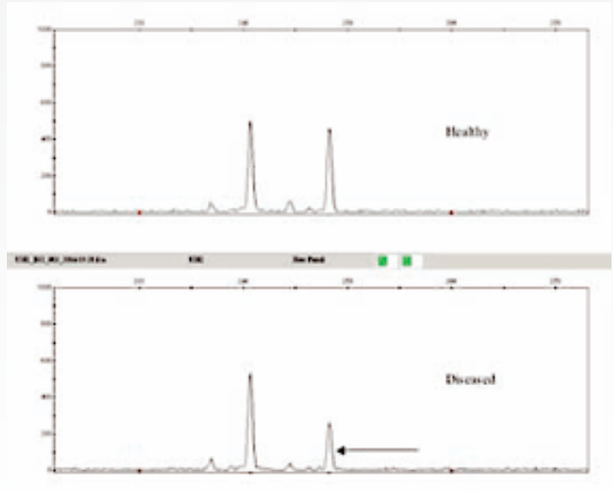


Genetic markers: *microsatellites*

- must found the conserved regions flanking the microsatellites
- separation using electrophoresis (agarose gel or sequencer)



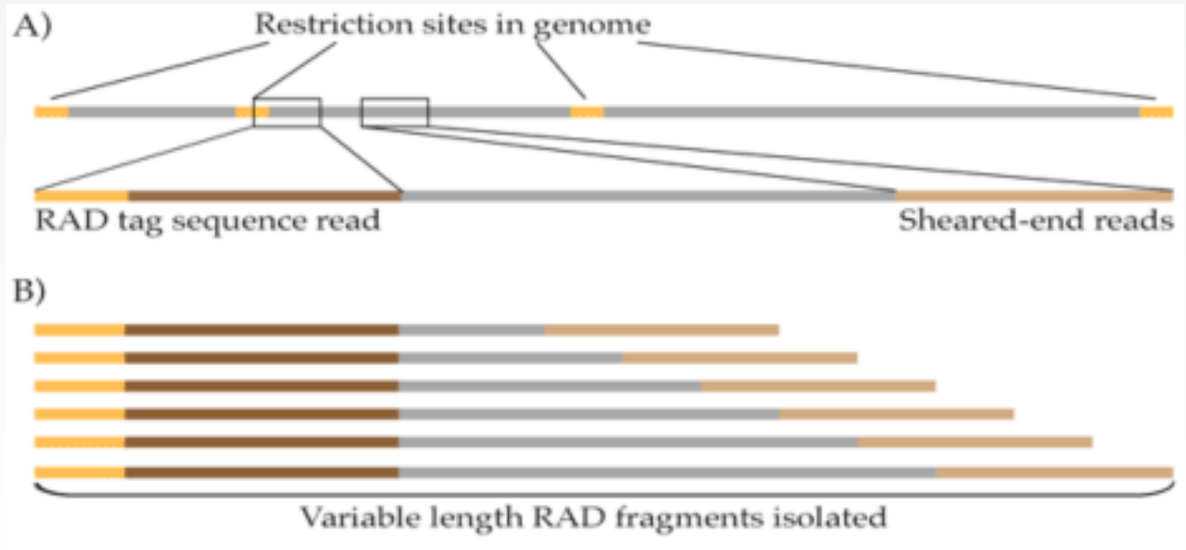




⊖ difficult to identify the conserved regions flanking the microsatellites

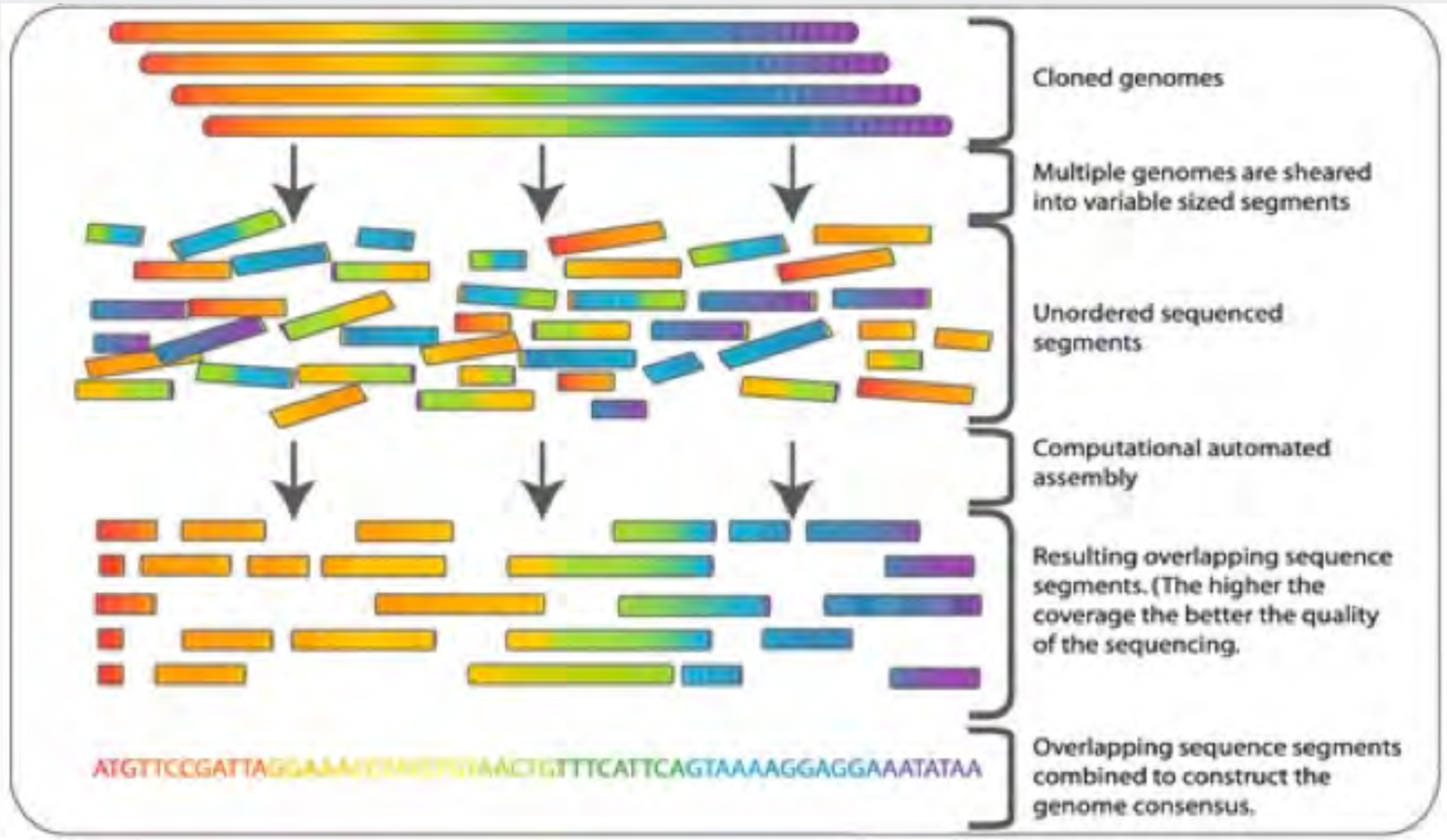
Genetic markers: *Next Generation Sequencing (NGS)*

- several methodologies and machines
  from "old" 454 Roche, Ion Torrent to currently used Illumina or PacBio RS II
- used for short or long amplification, numerous copies
  - whole genome
  - SNP (Single Nucleotid polymorphism)
    punctual mutation in a gene, present in >1% of the population
  - RAD sequencing / RAD markers (Restriction-site associated DNA)
    sequencing DNA just after a known restriction site



Next-generation sequencing - *sequencing*

- sequencing part or the entire genome (nDNA or mtDNA)
- appropriate for phylogeny or phylogeography
- limited number of analysed individuals



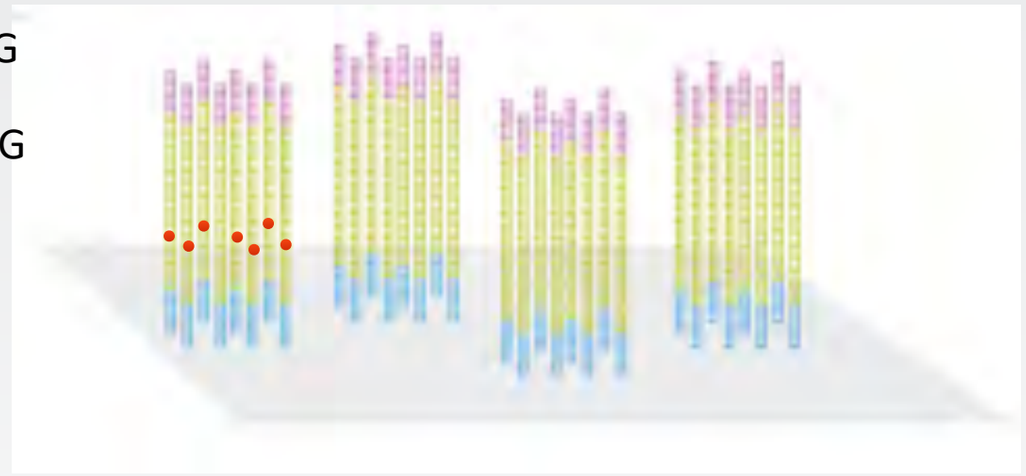
Next-generation sequencing - *RAD markers*

- Restriction Site Associated DNA markers (RAD markers)
- same approach as for RAPD or AFLP, but with the sequences (not the length)

GCGTTGTCTTGCTAGTATCGTCGTAATTCTCTTCCACCAGCAGTCG

GCGTTGTCTTGGTAGTATCGTCGTAATTCTCTTCCACCAGCAGTCG

SNP: Single Nucleotide Polymorphism



- typically hundreds of SNPs can be identified and used as markers

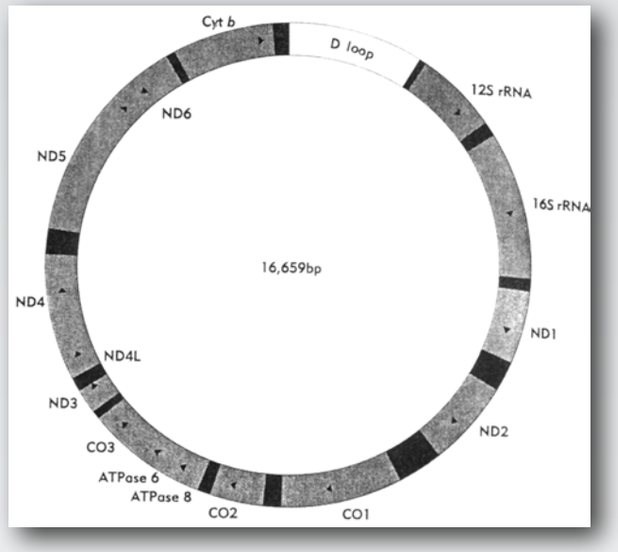
Next-generation sequencing - *RAD markers*

- Advantages
  - no need for time-consuming locus development
  - adequate for fingerprinting at individual level
  - high number of loci
  - technique rather sophisticated but the process can be automated
  - no high DNA quality needed
  - at the population level: ≥ 4 individuals only
- Disadvantages
  - bioinformatics
  - repeatability!
  - 2000 CHF, up to 5000 per case

Genetic markers - *summary*

	Basis	Polymorphism	Level of polymorphism	Dominance / co-dominance	selection	development	cost	non-invasive sampling
Allozymes	amino-acid polymorphism	change in amino-acid	low	co-dominant	under	none	low	no
sequencing	sequencing of PCR product of a defined gene/region	nucleotide polymorphism, inserts, deletion	low/high	co-dominant	no or under	none	high	yes
RFLP	Randomly fragmented DNA	length of the fragments	medium	co-dominant	no (rarely under)	limited	moderate	no
RAPD	Random amplified DNA fragments	amplifiable or not amplifiable fragment	medium	dominant	no (rarely under)	limited	low/ moderate	yes
AFLP	Random amplified DNA fragments	amplifiable or not amplifiable fragment	medium	dominant	no (rarely under)	limited	moderate / high	yes
microsatellites	PCR amplification of a unique loci, harbouring simple sequences repeats	variation in the number of repeats	high	co-dominant	no	long time, high cost	moderate	yes
NGS-RAD	Random amplified DNA fragments	single nucleotide polymorphism	medium/high	co-dominant	no or under	none	high / very high	yes

Mitochondrial markers



- numerous copies in a cell
- only maternal lineage / no (limited) heterozygosity
- animals: about 15-17k bp
  - well known: sequencing from defined primers
  - most interesting regions:
    - Control region (highly variable non-coding region: intra population → species)
    - cytochrome b (subspecies → genus)
    - NADH dehydrogenase 1-6 (subspecies → genus)
    - CO1 (species → order)
    - 12S / 16S (species → order)
- plants: 200k bp to >2400k bp (chloroplastic DNA)
  - sequencing of some parts
  - presence of microsatellites in the chloroplastic DNA

Mitochondrial analyses

- only maternal lineage / no (limited) heterozygosity
- limited mutation rate: 1-10% / million of years
- methods used: SEQUENCING
- reconstruction of lineage, relationship between genus, species: PHYLOGENY
- relationship within a species, with implication of the geography e.g. PHYLOGEOGRAPHY: geographical distribution of genealogical lineages

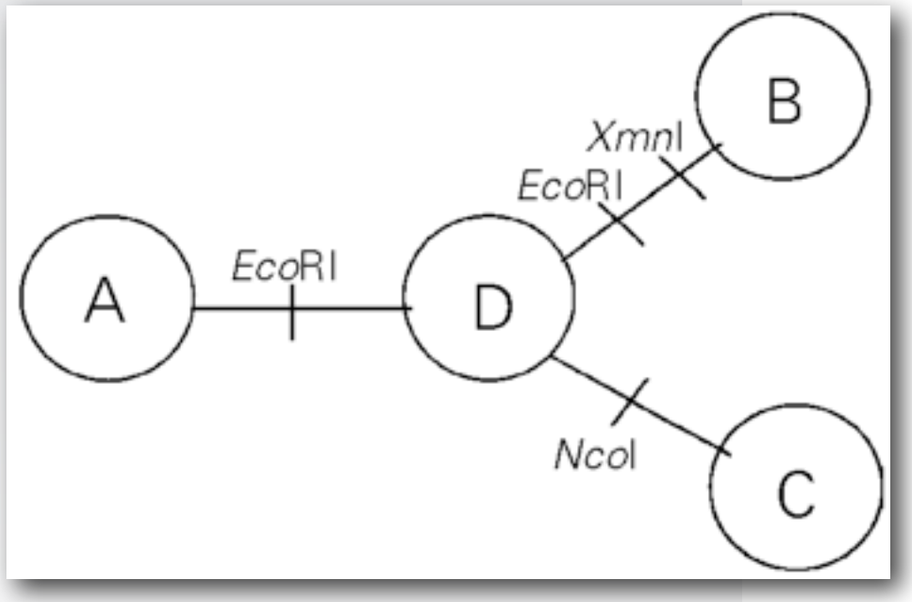
Mitochondrial analyses: *phylogenetic trees*

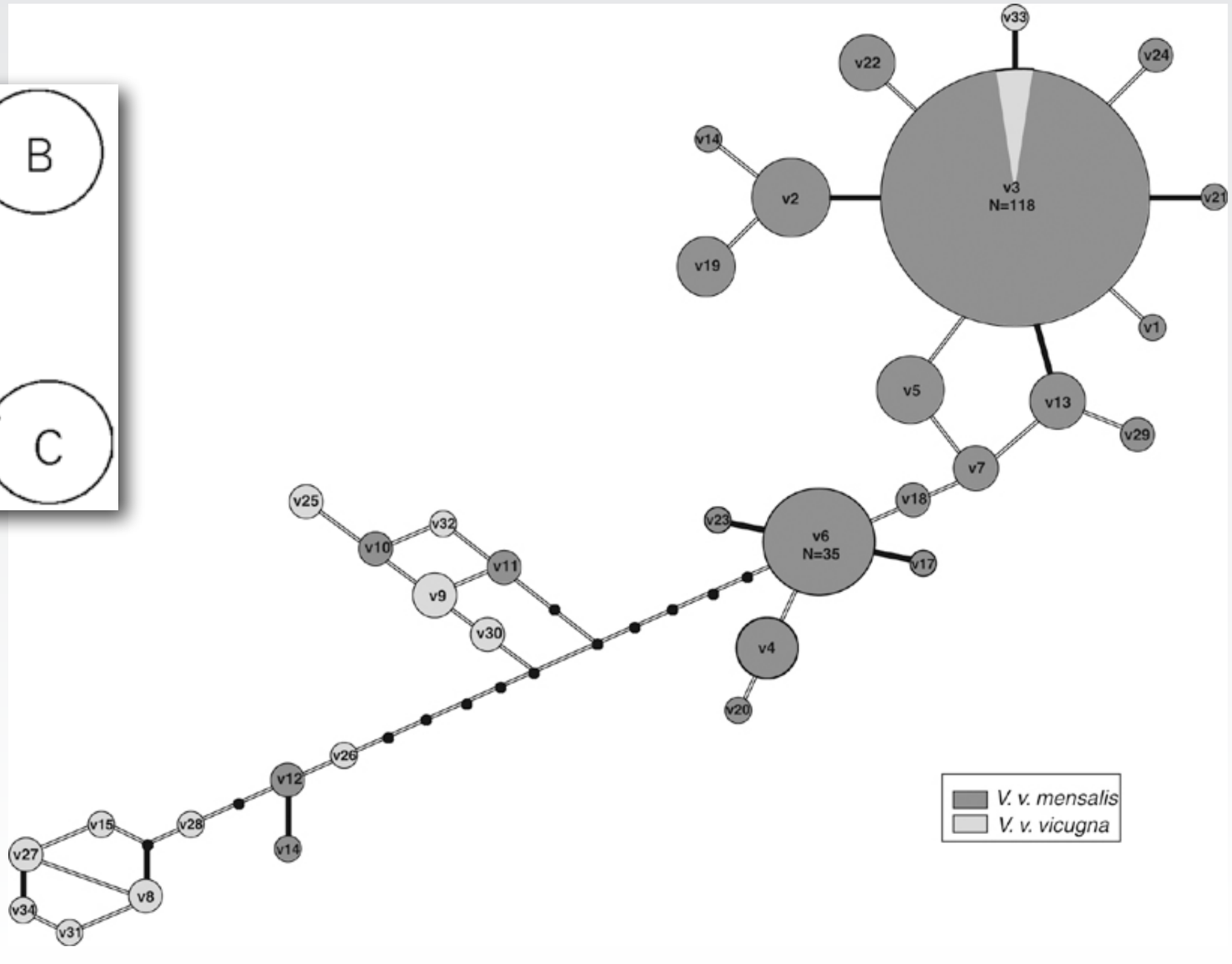
- regrouping most similar haplotypes

- different methods:
  - maximum likelihood: the tree with the highest probability
  - maximum parsimony: the less number of steps (mutations)
  - genetic distance (Neighbour joining): regrouping most similar haplotypes
  - Bayesian method: posterior probability, after simulating and keeping the most probable trees

Mitochondrial analyses: *network*

- re-create all steps (mutation) between all haplotypes with a minimum steps





Nuclear markers

- paternal and maternal lineages: 2 copies ⇒ heterozygosity
- mutation rate:
  very low (e. g. coding region) to very high (e. g. microsatellites)
- use for
  - pedigree reconstruction (maternal-paternal lineages)
  - level of inbreeding
  - population differentiation
  - migration estimation
  - differentiated behaviour (migration, ...) between sexes
  - ...

Nuclear markers: *some definitions*

- **Locus**: a segment of DNA, e.g. a microsatellites, a region coding for a protein, ...
- **Alleles**: different forms of the same locus, e.g. different length of a microsatellite, different amino-acidic chain in a protein, ...
- **Heterozygote**: an individual with two different allele at a locus e.g. alleles A_1A_2 for the locus A
- **Average heterozygosity**: mean of the heterozygosity at all loci
- **Allelic diversity**: average number of alleles per locus

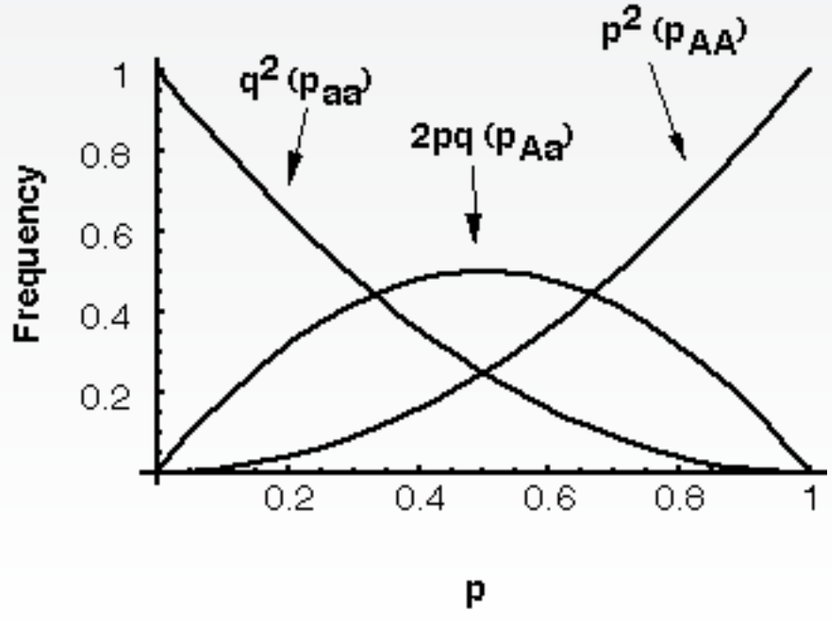
Nuclear markers

- markers used
  - microsatellites
    - when microsatellites already developed
    - no limitation by cost
    - more for animals (sometimes difficult to find in plants)
    - neutral markers
  - AFLP
    - when no microsatellites already exist and cannot been developed (time, cost)
    - plants
    - dominance is not a problem
  - RFLP
    - when no microsatellites already exist and cannot been developed (time, cost)
    - limit the cost
    - plants
    - dominance is not a problem
  - RAPD, enzymes, sequencing, SSCP, fingerprints, ...
    - particular cases

Nuclear marker analyses: *Hardy-Weinberg (HW) equilibrium*

- in large population, with random mating and no mutation, migration or selection
- allele and genotypes frequencies in equilibrium
- e.g. locus with alleles A_1 and A_2, relative frequency of p and q, where p+q=1
  - proportion of A_1A_1: (♀p - ♂p) p*p = p^2
  - proportion of A_2A_2: (♀q - ♂q) q*q = q^2
  - proportion of A_1A2: (♀p - ♂q AND ♀q - ♂p) 2* p * q = 2pq
- at the HW equilibrium: $p^2 + 2pq + q^2 = 1$

	A(p)	a (q)
A(p)	AA(p^2)	Aa (pq)
a (q)	Aa (pq)	aa (q^2)



Nuclear marker analyses: *genetic diversity characteristics*

- expected heterozygosity (gene diversity): H_e
  - for p and q allele frequency: $H_e=2pq$
  - for more alleles: $H_e = 1-\sum p_i^2$ for all alleles frequencies
- observed heterozygosity: H_o
  - proportion of heterozygotes at a locus
- allelic richness: A (or A_R)
  - average number of alleles per locus

Nuclear marker analyses: *genetic diversity characteristics*

- example 1

	AA	AB	BB	total
number	27	23	5	55
genotype frequency	0.49	0.42	0.09	1

estimation of alleles frequency:

$p = [(2*27)+(1*23)] / [2*55] = 0.70$ OR $p = [(2*0.49)+(1*0.42)] / 2$

$q = [(2*5)+(1*23)] / [2*55] = 0.30$ OR $p = [(2*0.09)+(1*0.42)] / 2$

$p + q = 0.70 + 0.30 = 1$

- expected heterozygosity: H_e

$H_e = 1-\sum p_i^2 = 1 - [0.70^2 + 0.30^2] = 1-[0.49+0.09] = 1-0.58 = 0.42$

- observed heterozygosity: H_o

no heterozygotes / total number = 23 / 55 = 0.42

- allelic richness: A (or A_R)

average number of alleles per locus = 2

- Population at the HW-equilibrium (compare H_e and H_o)

$2pq = 2x0.70x0.30 = 0.42 \approx Ho$

$p^2 + 2pq + q^2 = 0.70^2 + 2*0.70*0.30 + 0.30^2 = 0.49 + 0.42 + 0.09 = 1$

Nuclear marker analyses: *genetic diversity characteristics*

- example 2

	91/91	91/95	91/97	95/95	95/97	97/97	total
number	10	24	6	23	9	8	80
genotype frequency	0.125	0.3	0.075	0.2875	0.1125	0.1	1

- estimation of alleles frequency:

  p = [(2*10)+(1*24)+(1*6)] / [2*80] = 0.312

  q = [(2*23)+ (1*24)+(1*9)] / [2*80] = 0.494

  r = [(2*8)+ (1*6)+(1*9)] / [2*80] = 0.194

  p + q +r = 0.312 + 0.494 + 0.194 = 1

- expected heterozygosity: H_e

  $He = 1-\sum pi^2 = 1 - [0.312^2 + 0.494^2 + 0.194^2] = 1-0.38 = 0.62$

- observed heterozygosity: H_o

  no heterozygotes / total number = 24 + 6 +9 / 80 = 0.49

- allelic richness: A (or A_R)

  average number of alleles per locus = 3

- Population not at the HW-equilibrium (compare H_e and H_o)

  $2pq + 2pr + 2qr = 2*0.312*0.494 +2*0.312*0.194 + 2*0.494*0.194 = 0.62 \neq H_o$

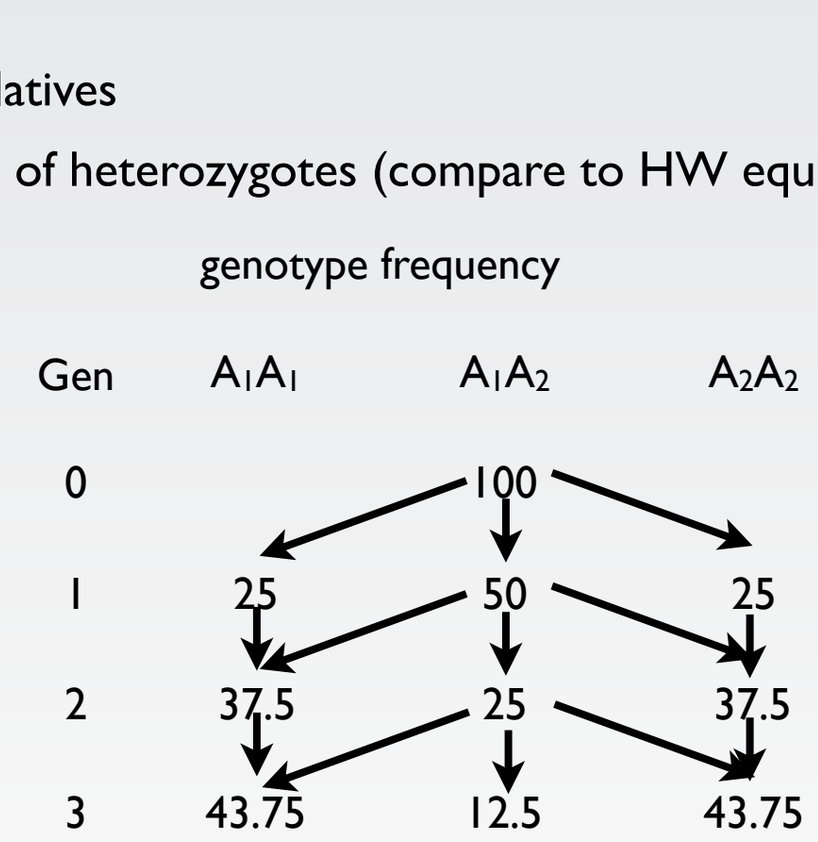
Nuclear marker analyses: *Deviations from Hardy-Weinberg (HW) equilibrium*

- causes
  - inbreeding
  - assortative and disassortative mating
  - fragmented populations

Nuclear marker analyses: *Deviations from Hardy-Weinberg (HW) equilibrium*

- causes
  - inbreeding
    - definition: mating with relatives
    - with inbreeding: decrease of heterozygotes (compare to HW equilibrium)

      e.g.: selfing



  - assortative and disassortative mating
  - fragmented populations

Nuclear marker analyses: *Deviations from Hardy-Weinberg (HW) equilibrium*

- causes
  - inbreeding
  - assortative and disassortative mating
    - preferential selection of mate with similar (assortative) or different (disassortative) genotype

      e.g.: human female selection:

      disassortative odour preferences in human (Wedekind et al., 1995; Wedekind & Furi 1997; Thornhill et al. 2003) ➡ disassortative

      MHC-disassortative mating observed between partners (Ober et al., 1997)

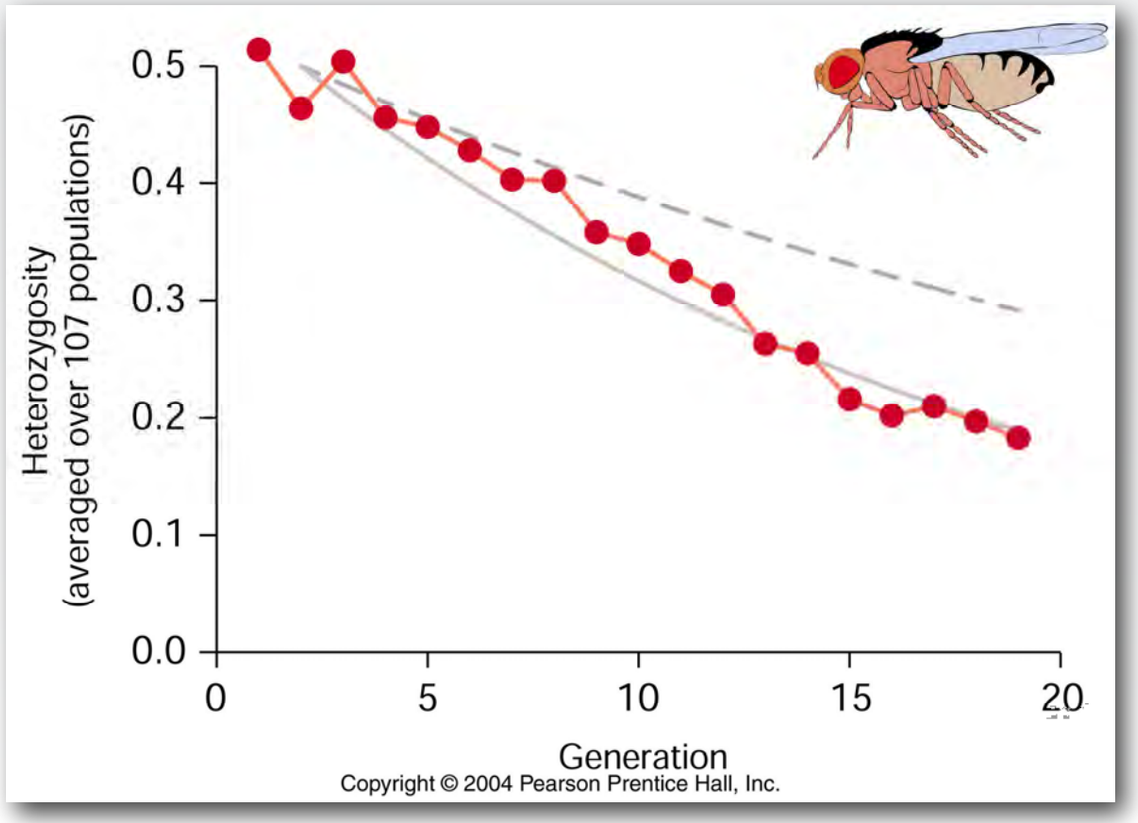
      BUT: MHC-similar facial preferences
  - fragmented populations

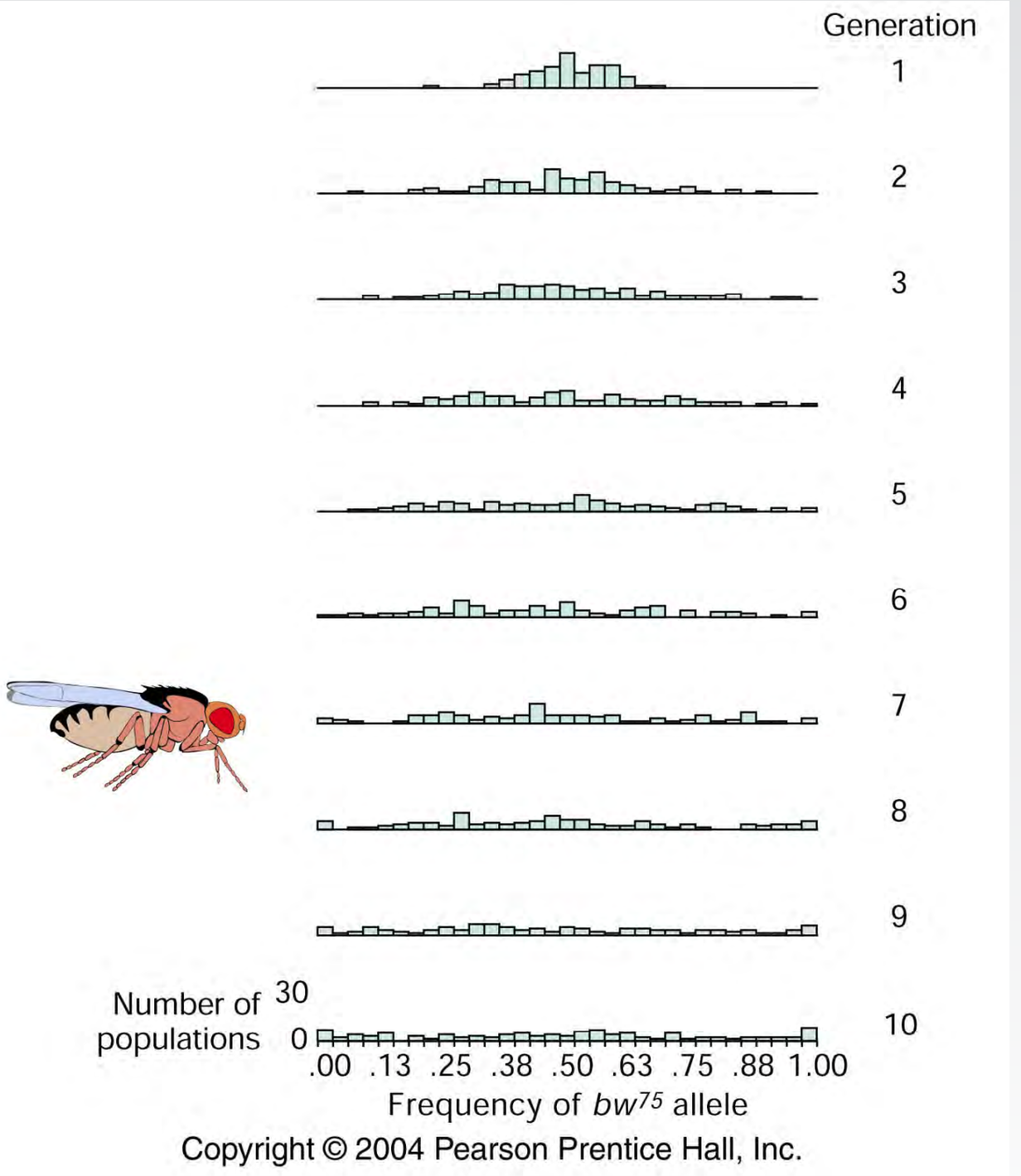
MHC (Major histocompatibility Complex): is a large genomic region or gene family found in most vertebrates. It plays an important role in the immune system, autoimmunity, and reproductive success.

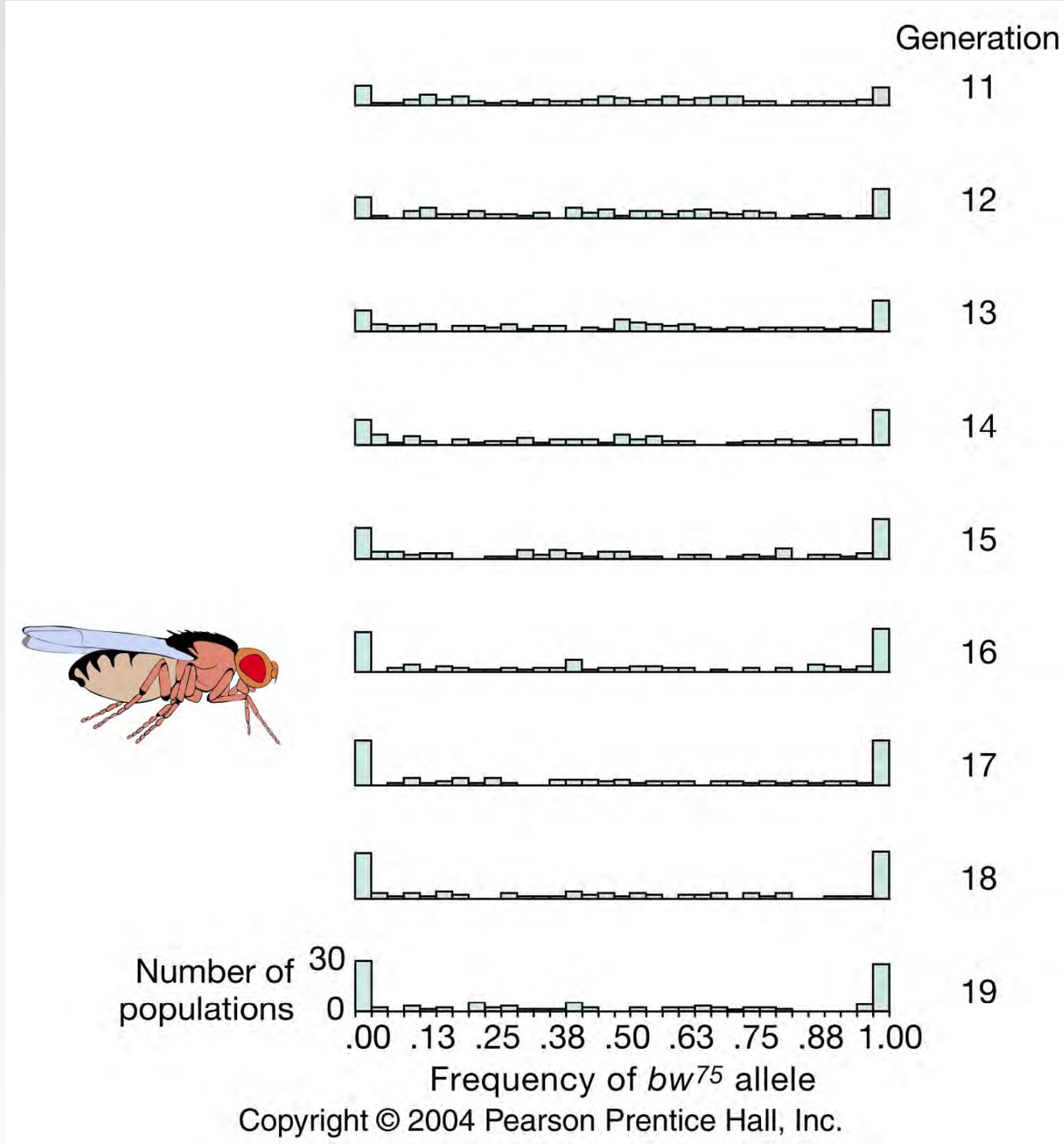
Nuclear marker analyses: *Deviations from Hardy-Weinberg (HW) equilibrium*

- causes
  - inbreeding
  - assortative and disassortative mating
  - fragmented populations
    - small isolated population fragments will differentiate at random due to genetic drift

e. g. Buri 1956: evolution of heterozygosity in bw^{75} allele over 19 generations in 105 replicate populations maintained with 16 parents per generations



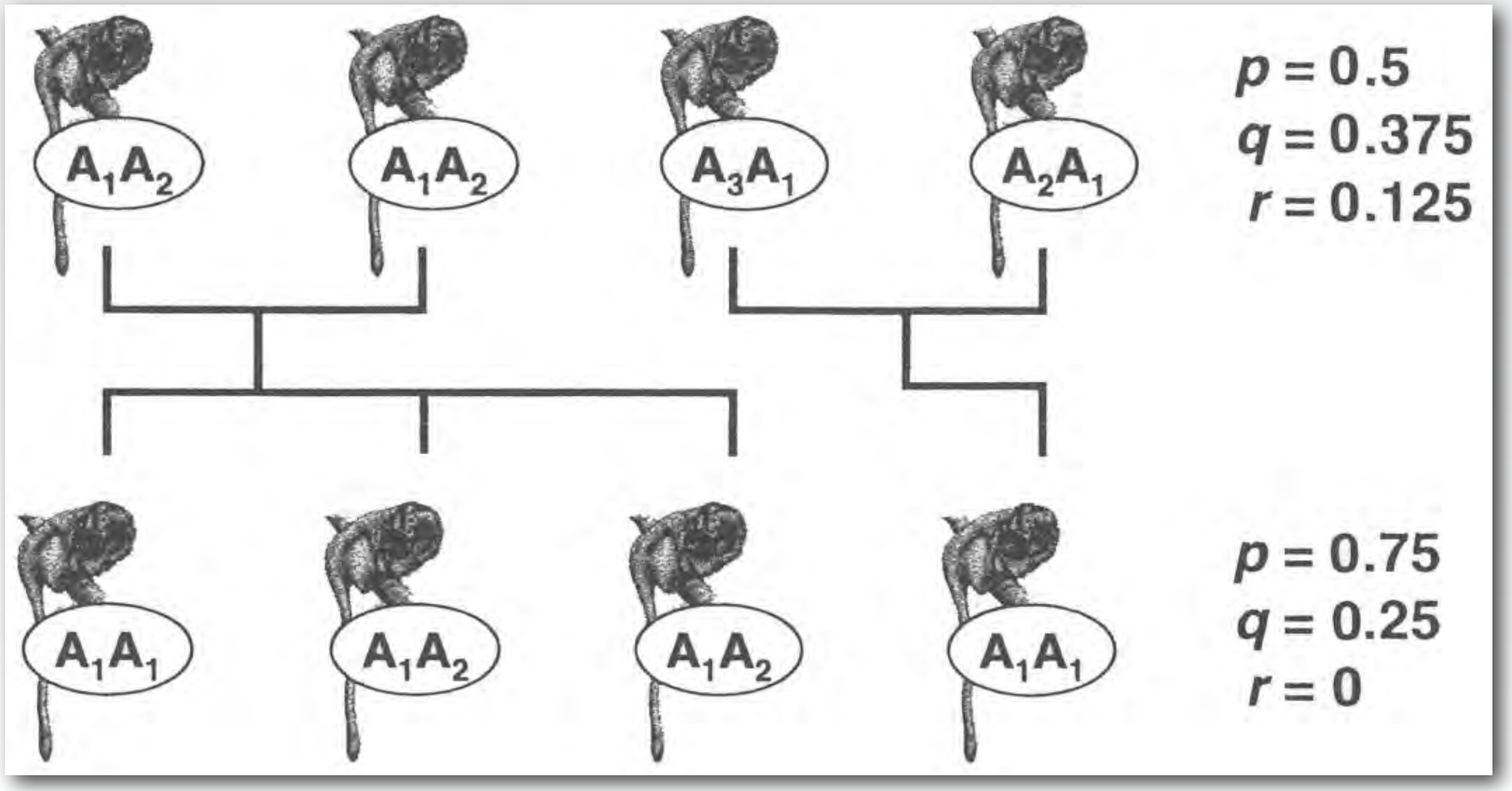




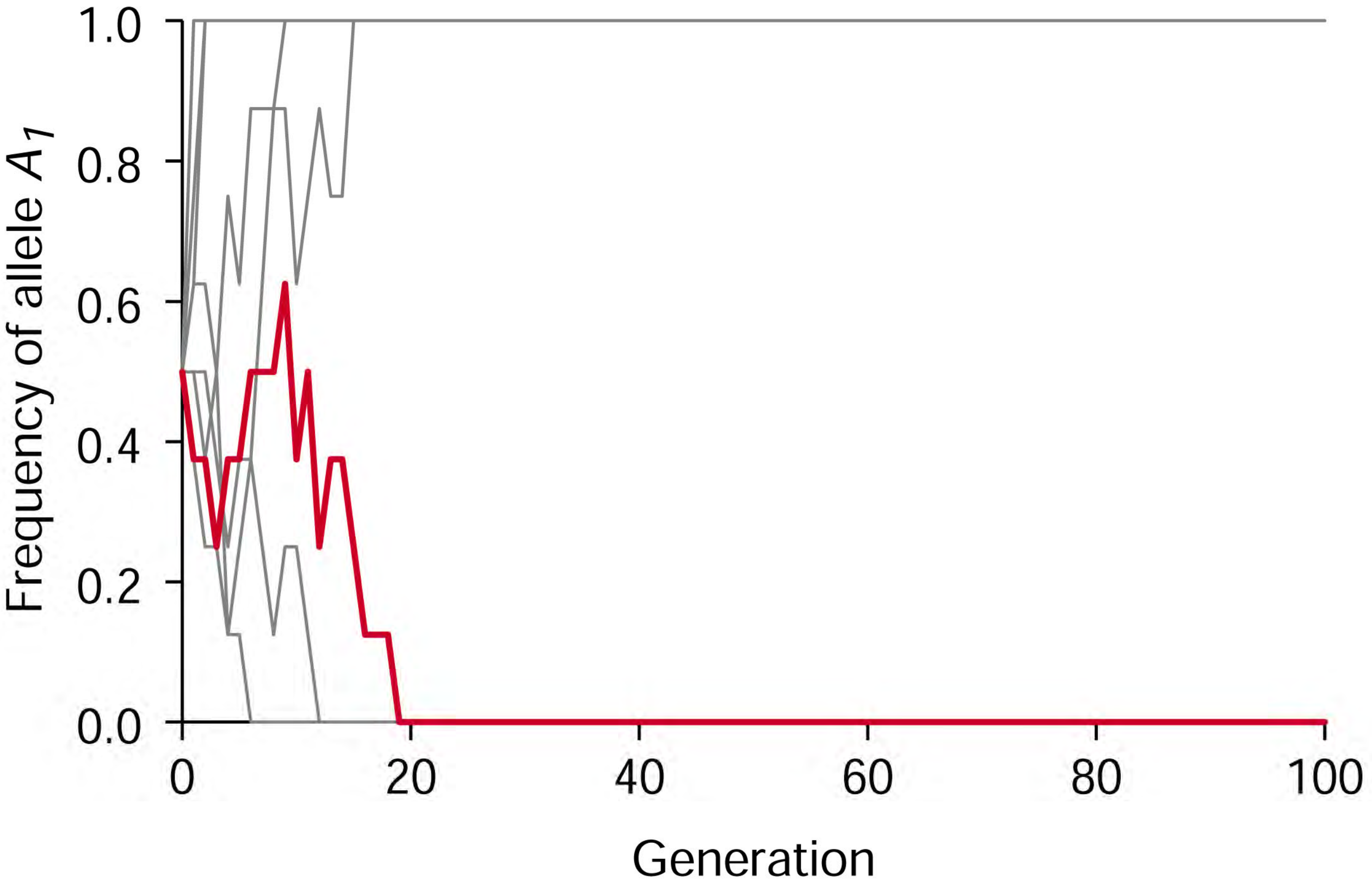
Buri, 1956: frequency distribution of the bw^{75} allele over 19 generations in 105 replicate populations maintained with 16 parents per generations

Small population problems: *impact of the population size on the genetic diversity*

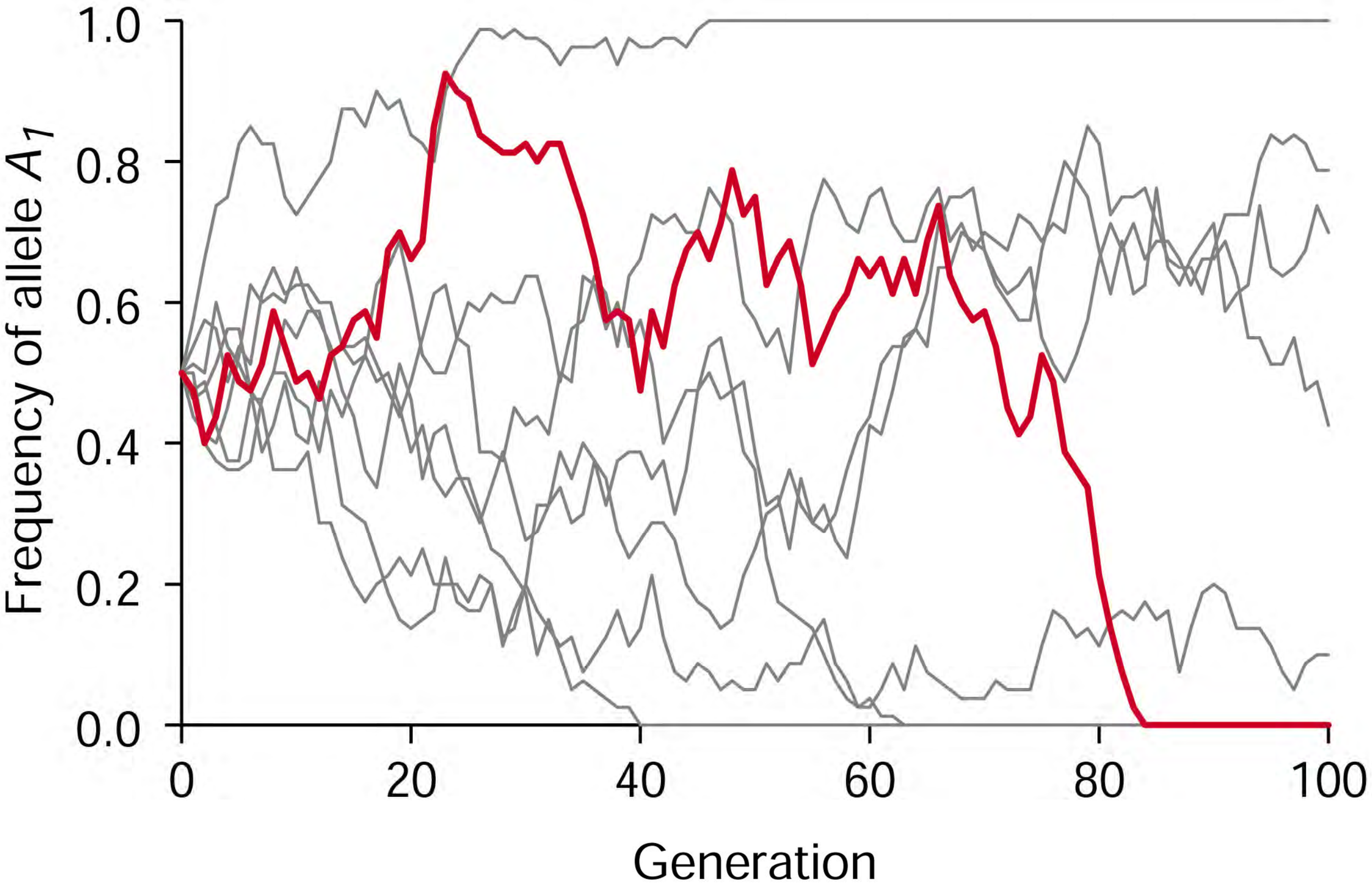
- stochasticity
  - just by chance, some alleles (especially the rare ones) may not be passed to the next generation and are consequently lost.
    ➡ frequency of alleles change over generation



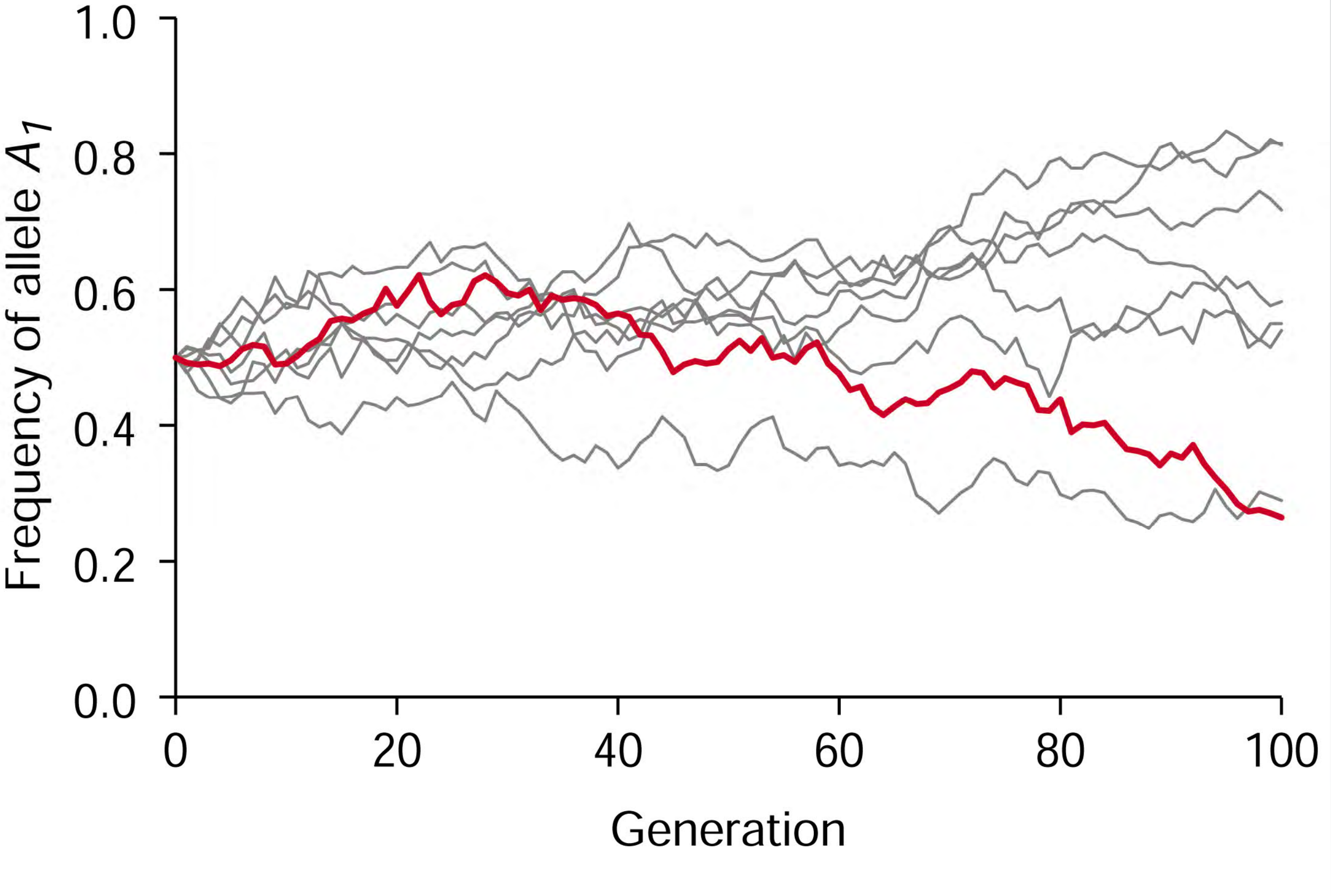
(a) Population size = 4



(b) Population size = 40

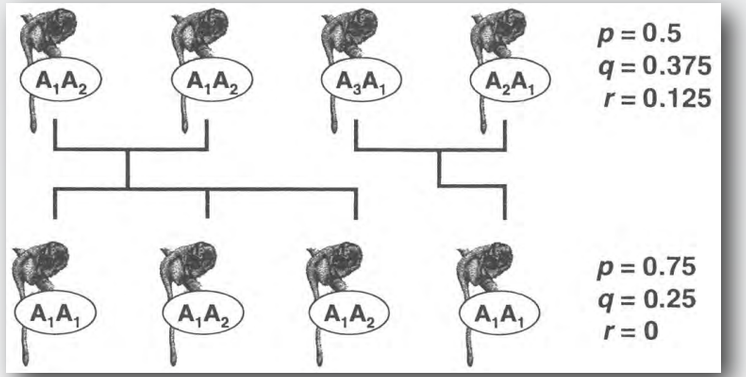


(c) Population size = 400



Small population problems: *impact of the population size on the genetic diversity*

- stochasticity
  - just by chance, some alleles (especially the rare ones) may not be passed to the next generation and are consequently lost.
    ➡ frequency of alleles change over generation



  - genetic drift: allele frequency change over generation, with a general reduction of the global genetic diversity

    consequences:
    - random changes in allele frequencies from one generation to the next one
    - loss of genetic diversity and fixation of alleles within populations
    - diversification among replicate population from the same original sources (e. g. fragmented populations)

Small population problems: *lost of genetic diversity*

- reasons of the lost of genetic diversity in small populations
  - genetic drift
  - inbreeding reducing heterozygosity
  - selection reducing genetic diversity by favouring one allele at the expense of other ➡ fixation

- impact:
  - reduce the ability to evolve in response to environmental changes

    e.g.: peppered moth in UK / resistance to myxoma virus in Australian rabbits

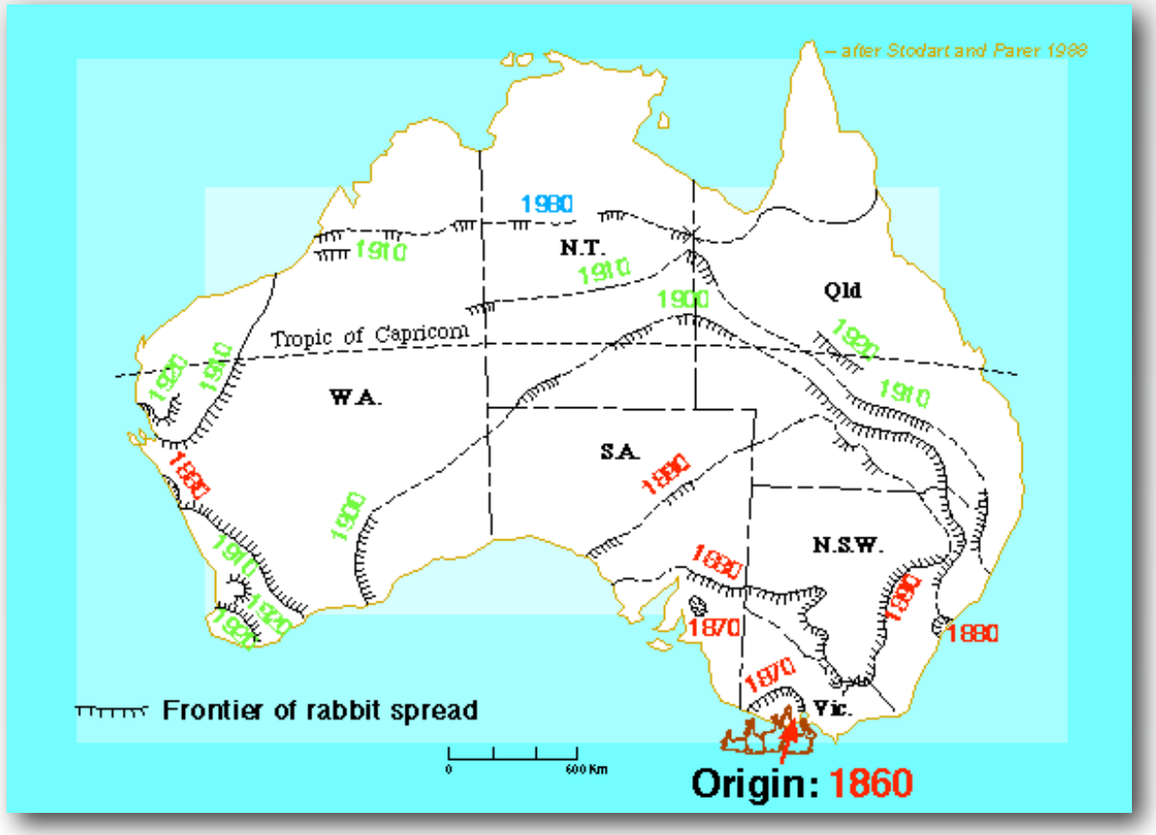
Introduction: *why genetic diversity is important in populations...*

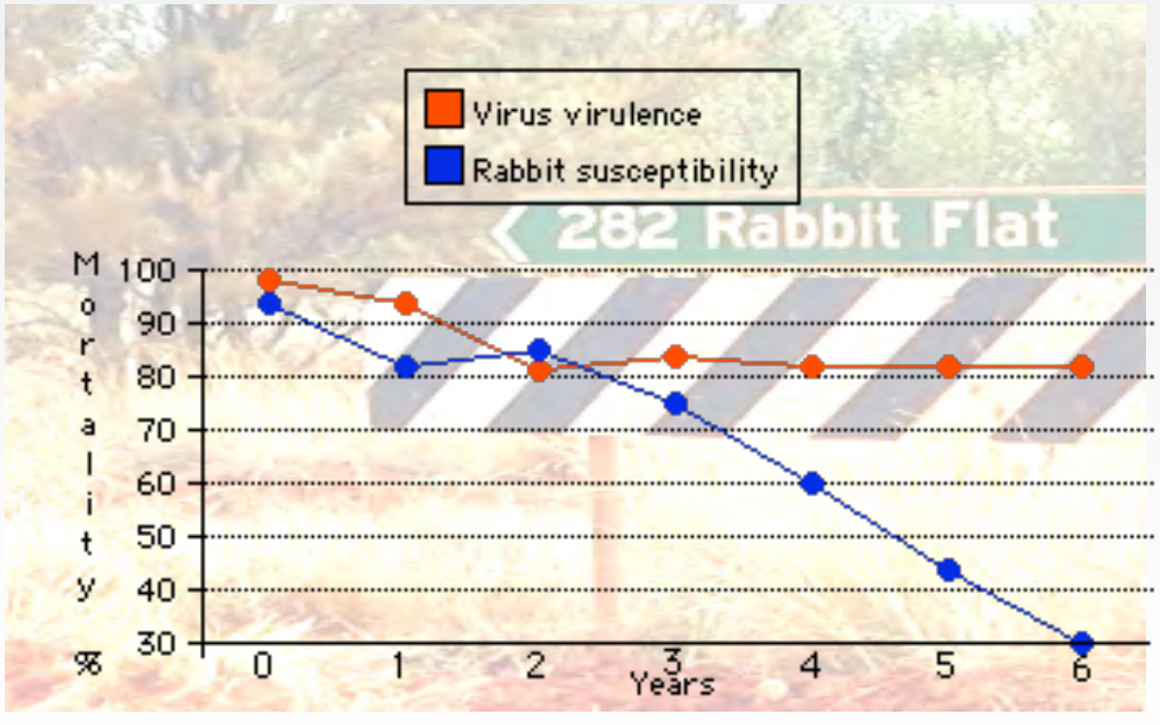
- genetic diversity reflect evolutionary potential
  - example 1 - habitat selection: peppered moth in UK
    - dark and light forms
    - night: active / day: resting on trees
    - ➡ camouflage critical for survival
    - light form: camouflaged on lichen-covered tree trunks
    - Industrialisation: kill lichen by sulphur pollution
    - ➡ light form: visible / dark form: camouflaged

Grant (1999) Fine tuning the peppered moth paradigm, Evolution 53, 980-984
Kettlewell (1973) The Evolution of Melanism, Clarendon Press, Oxford, UK
Majerus (1998) Melanism: Evolution in Action. Oxford University Press
Kettlewell (1958) A survey of the frequencies of *Biston betularia* (L.) (Lep.) and its melanic forms in Great Britain, Heredity 12, 551-572
but see also: Rudge (2006) Myths about moths: a study in contrasts, Endeavour 30, 19-23

Introduction: *why genetic diversity is important in populations...*

- genetic diversity reflect evolutionary potential
  - example 2 - disease resistance: resistance to myxoma virus in Australian rabbits
    - introduction of rabbits in Australia: 1860
    - control measure: introduction of myxoma in 50'
    - ➡ high mortality rate first years
    - high selection for resistance





Small population problems: *lost of genetic diversity*

- reasons of the lost of genetic diversity in small populations
  - genetic drift
  - inbreeding reducing heterozygosity
  - selection reducing genetic diversity by favouring one allele at the expense of another ➡ fixation

- impact:
  - reduce the ability to evolve in response to environmental changes
  - reduce the fitness

    e.g.: *Gentiana pneumonanthe* (Oostermeijer et al, 1995)
    see DH Reed, R Frankham (2003) for a review

Analysis of the relationship between allozyme heterozygosity and fitness in the rare Gentiana pneumonanthe L.
Oostermeijer et al. (1995) **J. Evol. Biol.** 8: 739-759 (1995)

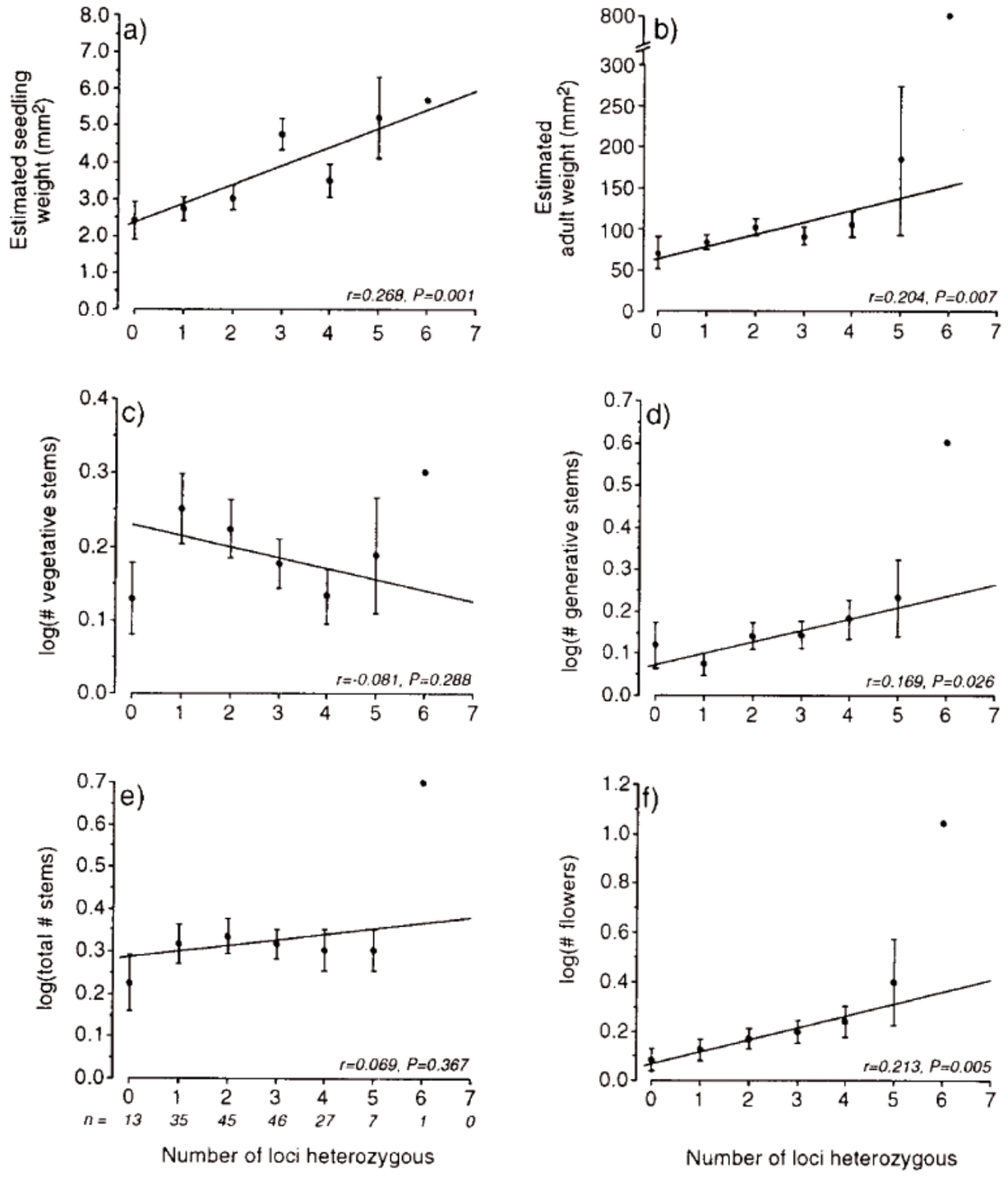


Fig. 1. Relationship between the number of heterozygous loci per individual (out of seven assayed polymorphic loci) and six components of individual fitness, (a) seedling weight, (b) adult weight, (c) number of vegetative stems, (d) number of generative stems, (e) total number of stems, and (f) number of flowers. Note that parameters (c) to (f) have been ln-transformed. Regression lines are based on values of individual plants and not on the class means shown in these graphs with their standard errors. In the right hand corner of each graph, the correlation coefficient (r) and its probability (P) is given. Below graph (e) the number of individuals per heterozygosity class (n) is shown. Only one individual was heterozygous for 6 of the seven loci (hence this class has no standard error), and none were heterozygous for all seven.

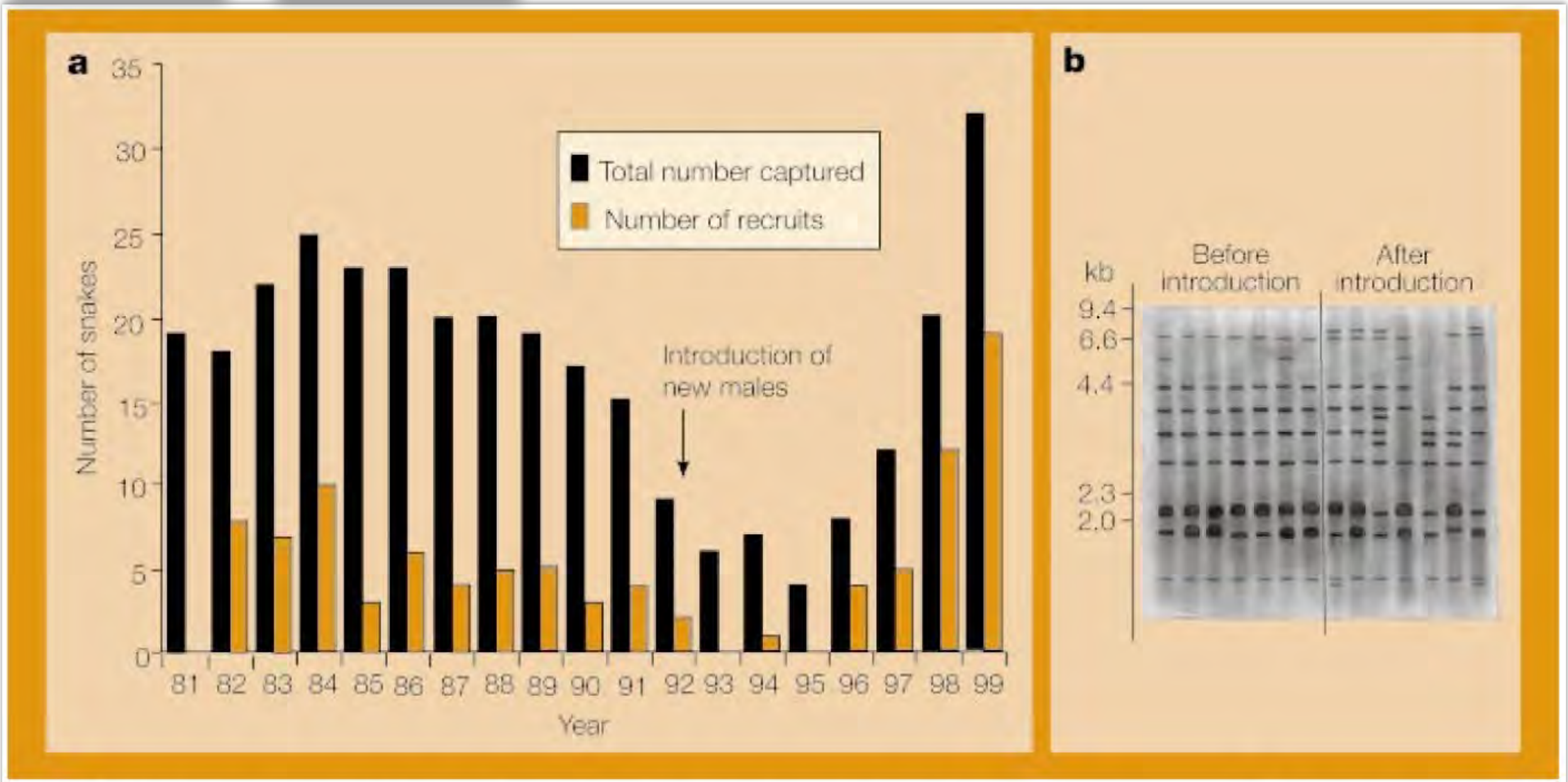
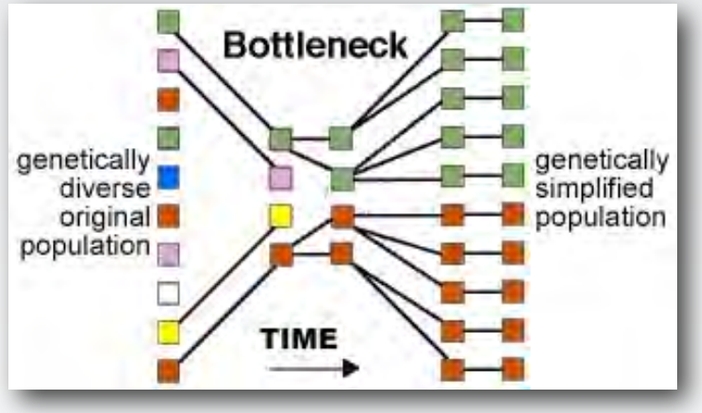


Figure 1 Introducing new males increases the genetic diversity and enables the adder population to recover. **a,** Total number and number of recruited male adders captured in Smygehuk from 1981 to 1999. **b,** Southern-blot analysis of major histocompatability complex (MHC) class I genes in seven males sampled before the introduction of new males (left) and in seven recruited males sampled in 1999 (right).

Madsen et al. (1999) Restoration of an inbred adder population, Nature 402, 34-35

Small population problems: *bottleneck*

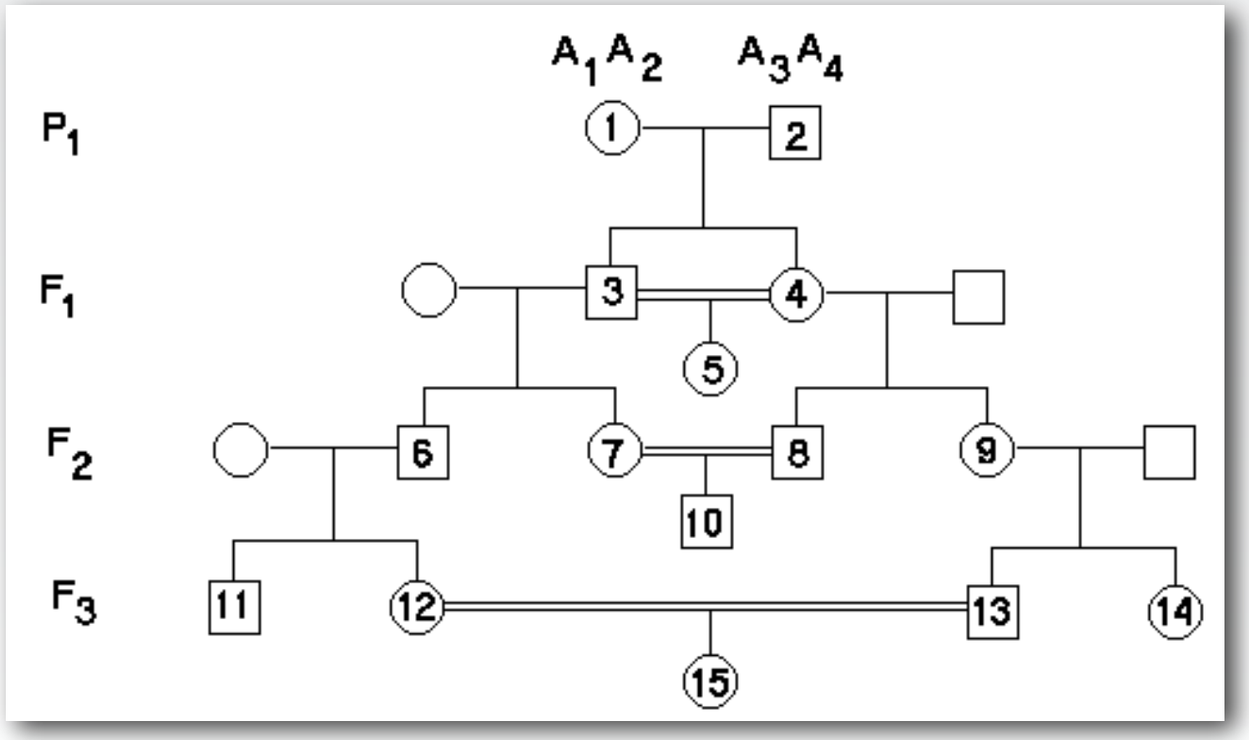
- bottleneck: large reduction of N_e in a period of time
  - consequence: lost of genetic diversity, especially rare alleles
  - impact depends on the population size during the bottleneck and the duration of it (nb generation)
  - e.g.: northern elephant seal (*Mirounga angustirostris*)
    - large reduction of the population size due to hunting
    - 20-30 survived in Isla Guadalupe (probably only a single harem)
    - mtDNA:
      - before 1892: ≥4 haplotypes (only 5 samples)
      - after 1892: only 2 haplotypes (>150 samples)
    - 20 allozymes:
      - no diversity in the northern elephant seal
      - normal level for the southern elephant seal (*Mirounga leonina*)



Small population problems: *Inbreeding estimations*

- inbreeding: mating of individuals related by ancestry
  measured as the probability that two alleles at a locus are identical by descent (*F*). Recent copies of the same allele are referred to as identical by descent, or autozygote

- also named as pedigree inbreeding



Relationship	***Description***	***Example***	***F of offspring***
Parent / Offspring	mother or father, to son or daughter	2 & 4	1/2
Full sibs	offspring of same parents	3 & 4	1/4
Half sibs	offspring with one parent in common	not shown	1/8
1st cousins	offspring of full sibs	7 & 8	1/16
2nd cousins	offspring of 1st cousins	12 & 13	1/64

Small population problems: *Theory of inbreeding in small populations*

in an hermaphroditic species

N = nb individuals

2N ancestral alleles (diploïd species)

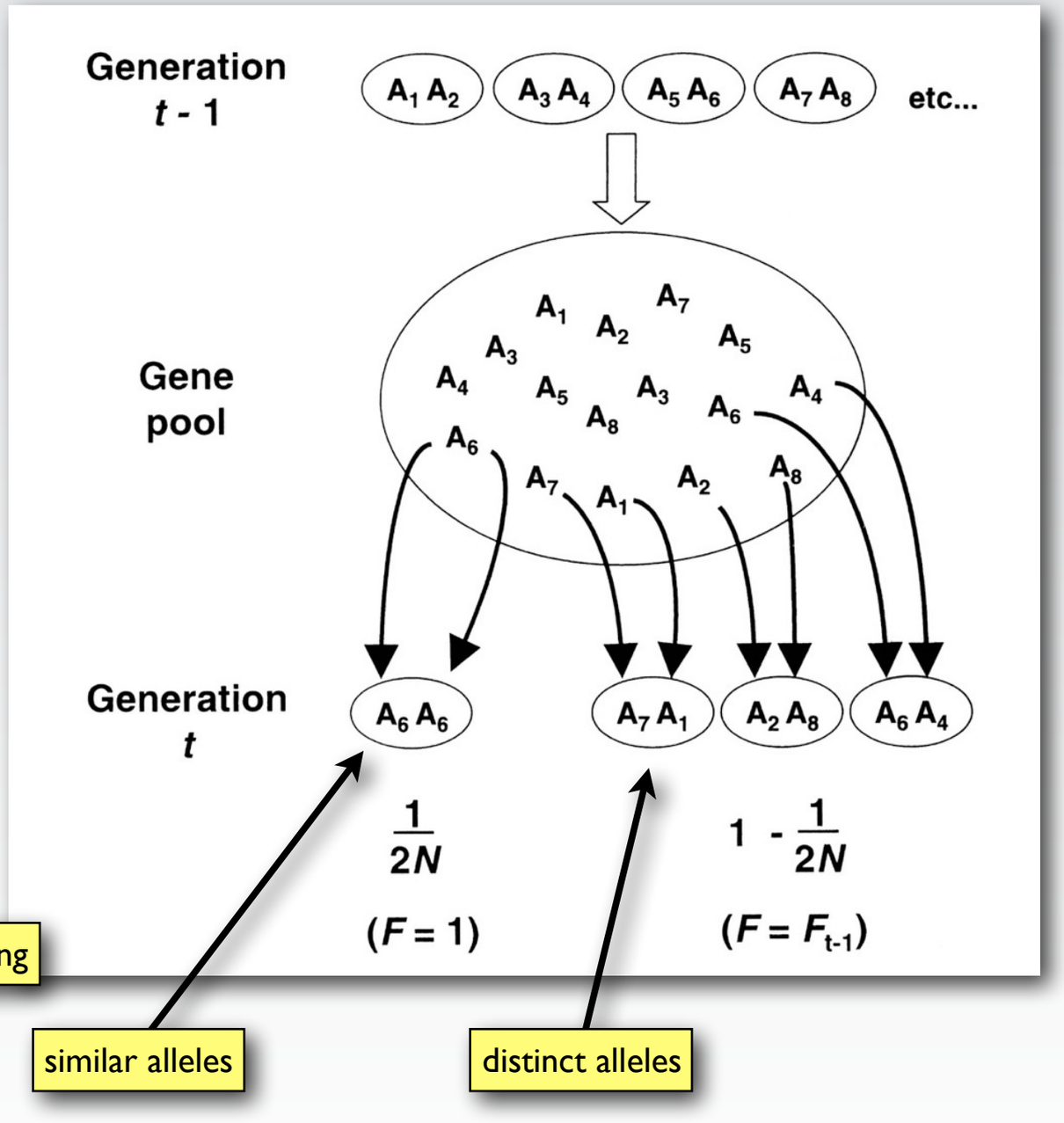
each individual at t: randomly sampling with replacement of two alleles

e.g. A6 first sampled:
prob. that the second is A6 for 1 individual:
= 1/2N

probability of sampling distinct alleles:
= 1 - 1/2N

➡ probability of creating a zygote with both alleles identical by descendent (F_t):
$F_t = 1/2N + [1-1/2N]F_{t-1}$ ← previous inbreeding

➡ increase of inbreeding per generation:
$\Delta F = 1/(2N)$



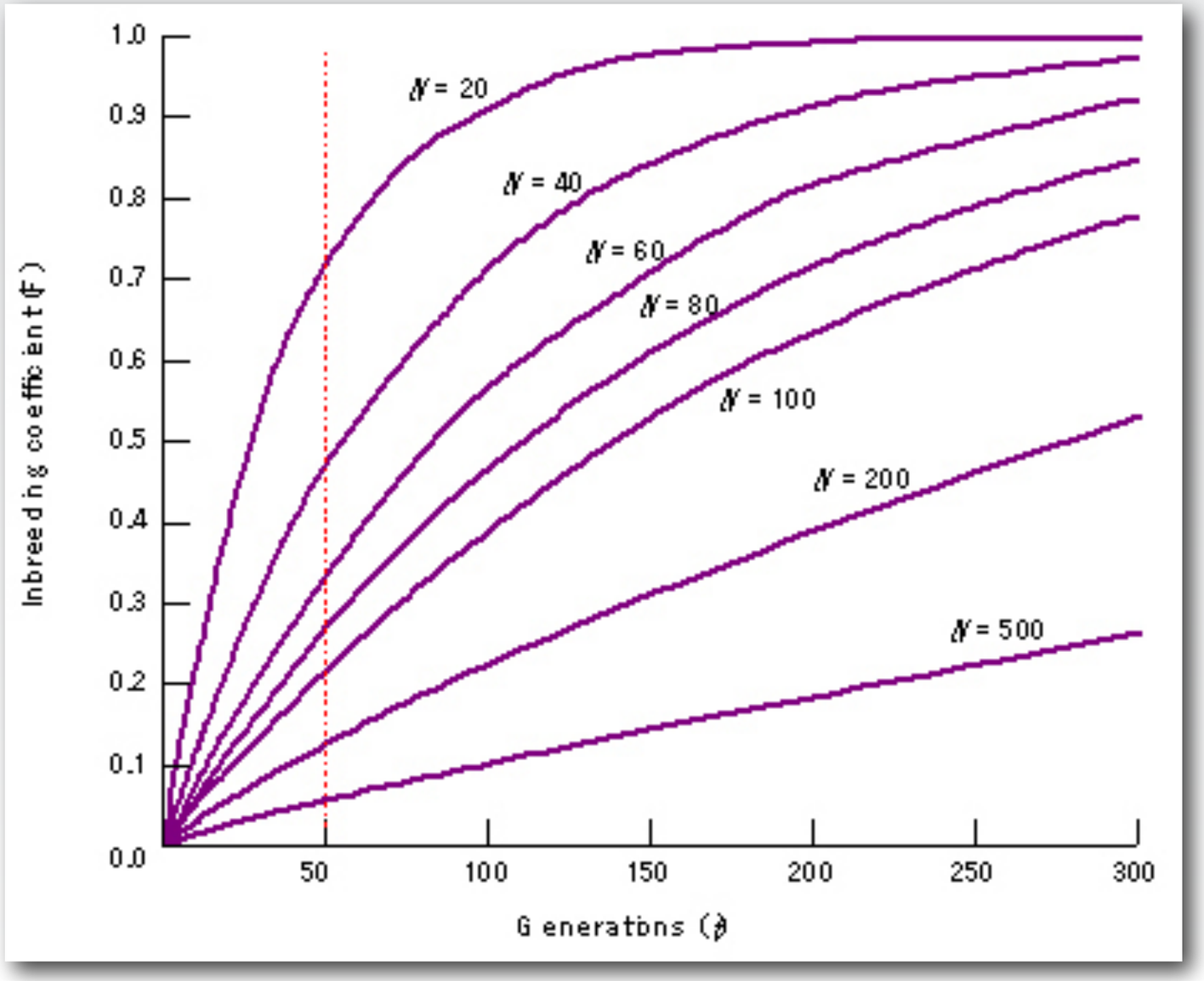
Small population problems: *Theory of inbreeding in small populations*

➡ probability of creating a zygote with both alleles identical by descendent (F_t):

$F_t = 1/2N + [1-1/2N]F_{t-1}$

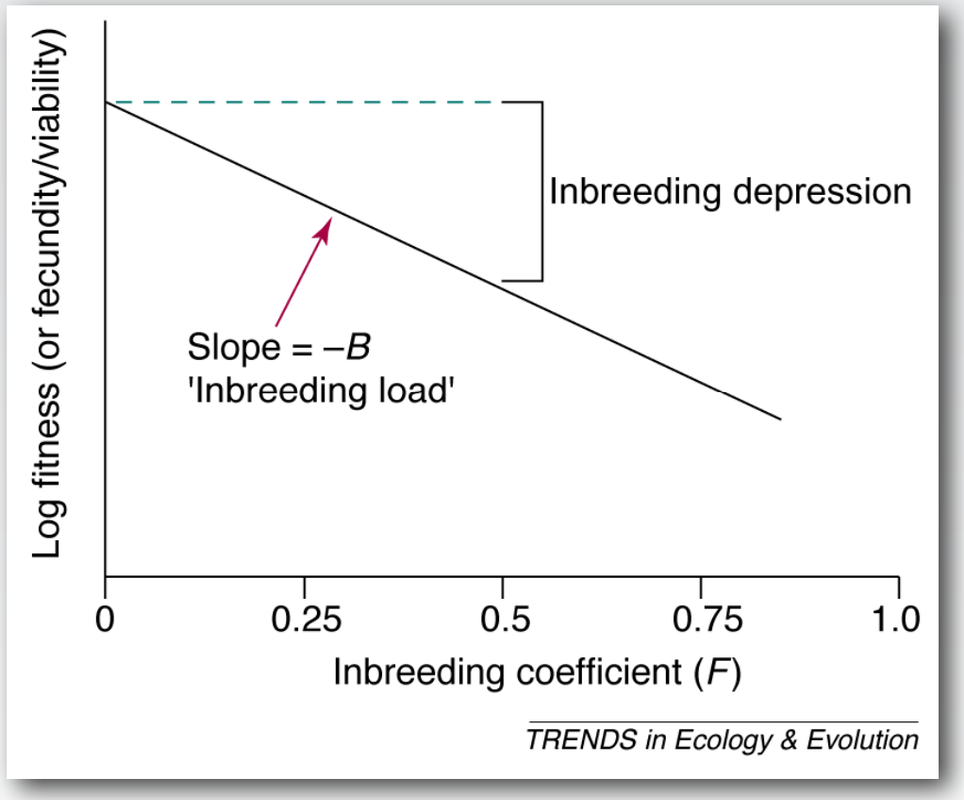
➡ increase of inbreeding per generation:

$\Delta F = 1/(2N)$



Small population problems: *Inbreeding depression*

- population size reduction increase inbreeding rate in closed populations ➡ inbreeding results in a decline of the global fitness, named as inbreeding depression



- purging
  - elimination due to a strong negative selection on rare deleterious recessive alleles
    purging highly effective for alleles with large effects (e. g. lethal)

Small population problems: *Inbreeding depression*

Charpentier et al. (2006), Life history correlates of inbreeding depression in mandrills (*Mandrillus sphinx*), **Molecular Ecology** 15:21-28

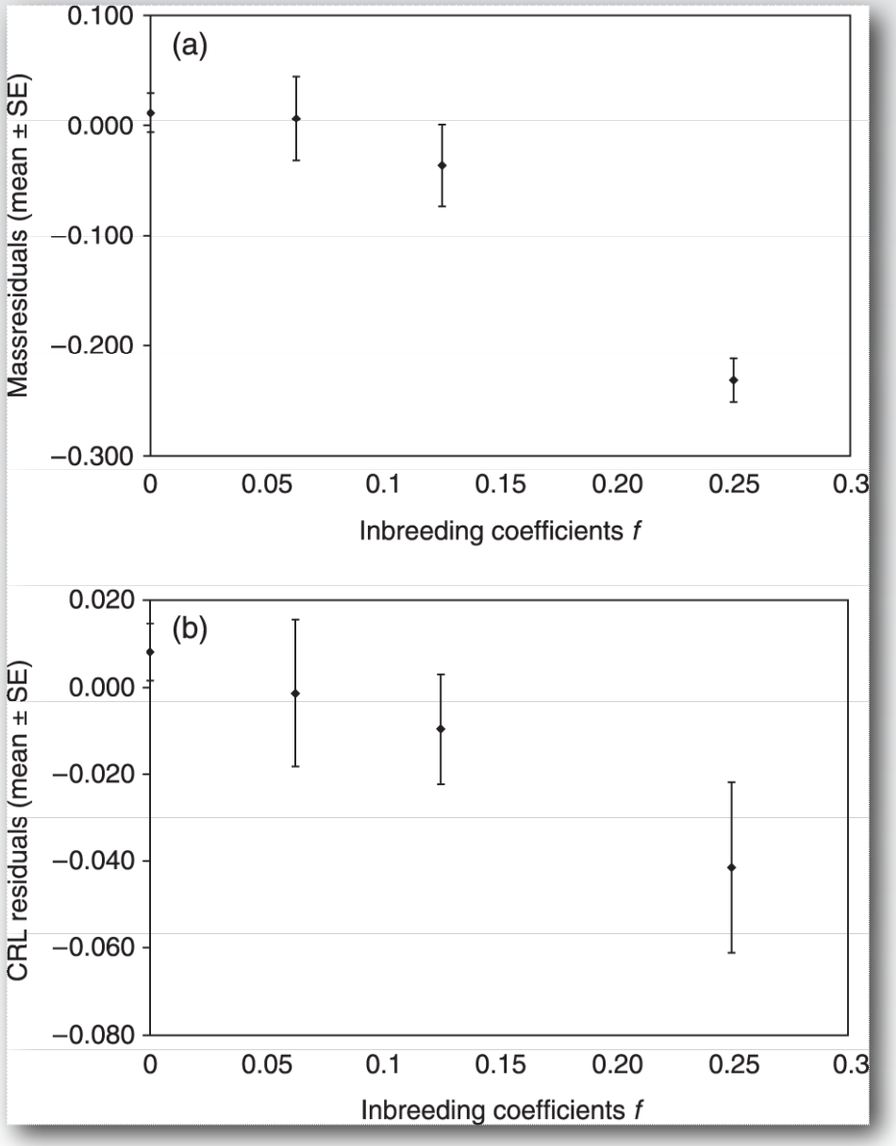


Fig. 1 Relationship between inbreeding coefficients and growth in females. Figures show mean ± SE for each inbreeding value. (a) Mass-for-age; (b) Crown-rump length -for-age (= embryos length)

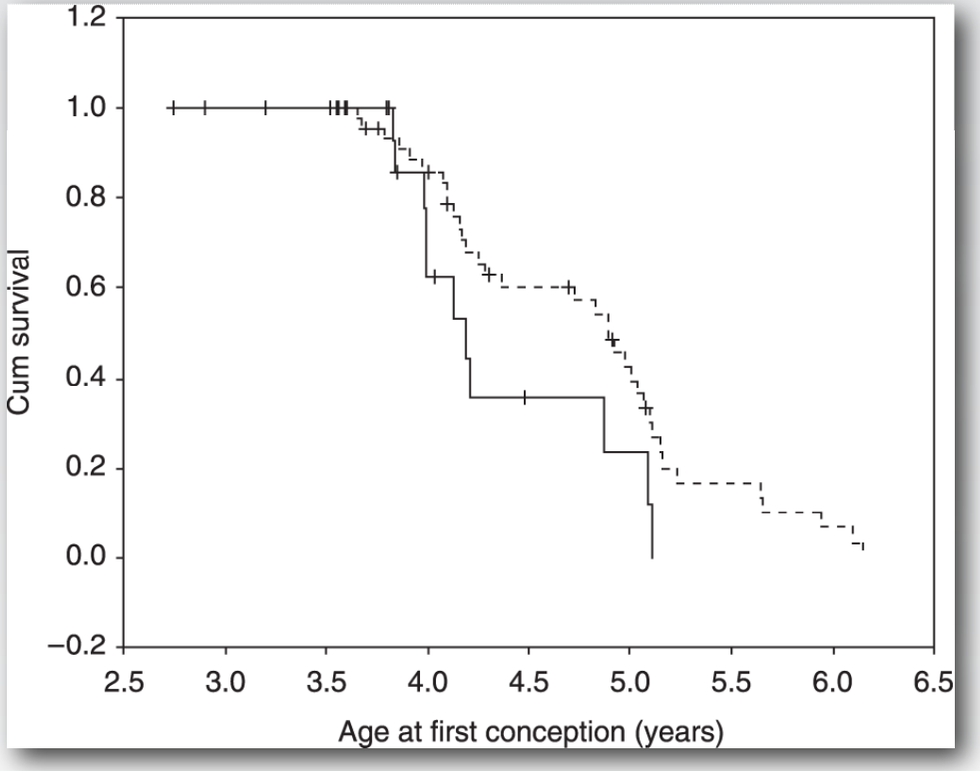


Fig. 2 Cumulative survival curve showing age at first reproduction in inbred (solid line) and noninbred (dashed line) female mandrills. Crosses indicate censored cases.

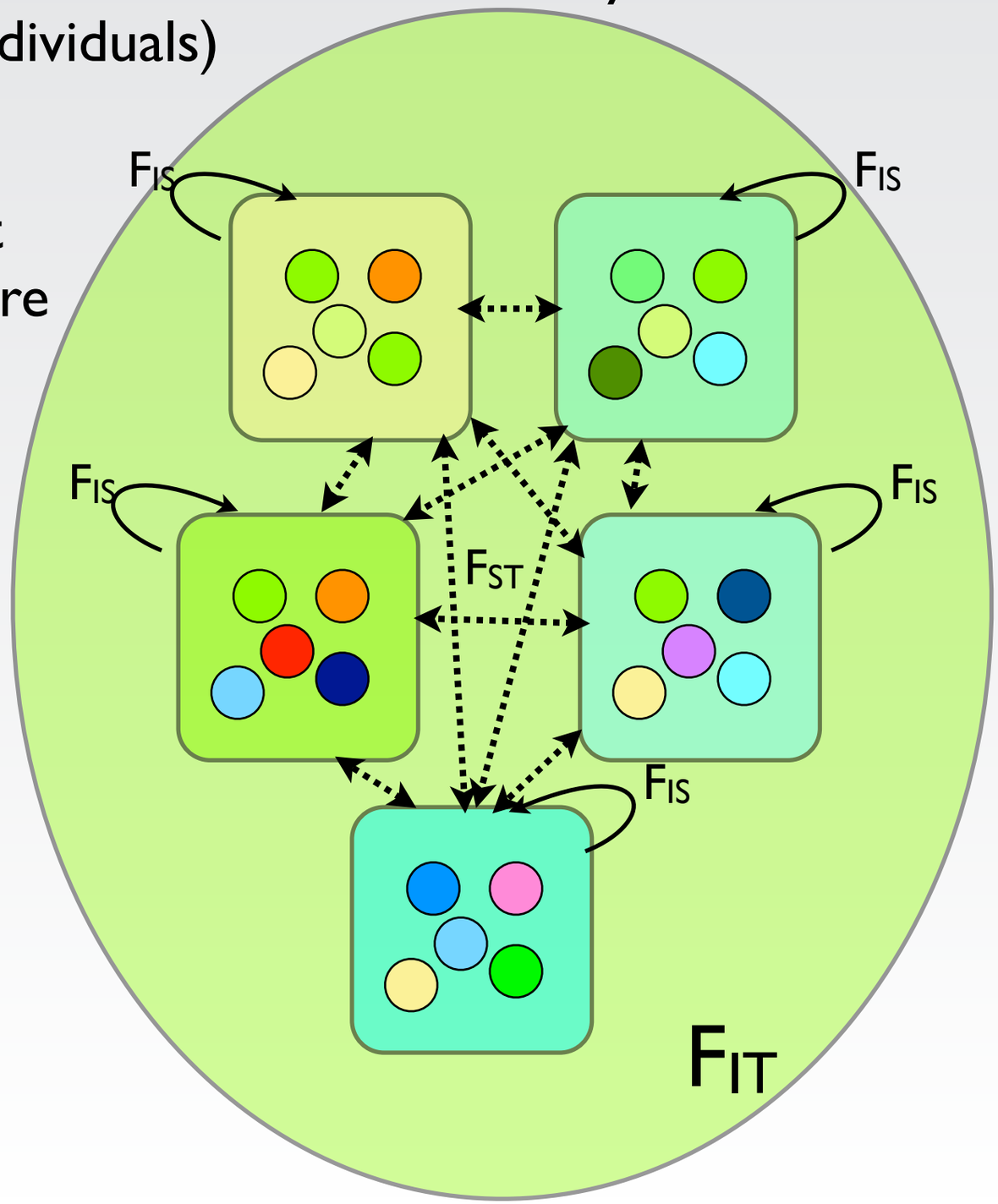
Population differentiation

- high fragmentation of habitats
  - instead of one continuous habitat (panmixia) ➡ separated populations without or with limited migration between them
- genetic differentiation between populations
  - due to genetic drift, stochasticity, selection, etc...
- measuring population fragmentation: *F*-statistics (Wright, 1969)

panmictic population is one where all individuals are potential partners

Population differentiation: *F-statistics*

- F_{IS}: probability that two alleles in an individual are identical by descent (≈ *F* averaged across all individuals) intra-population
- F_{ST}: fixation index - probability that two alleles from two populations are identical by descent between population structure between populations
- F_{IT}: general genetic structure
- $F_{IT} = F_{IS} + F_{ST} - (F_{IS})(F_{ST})$



Population differentiation: *F-statistics*

- $F_{IT} = F_{IS} + F_{ST} - (F_{IS})(F_{ST})$

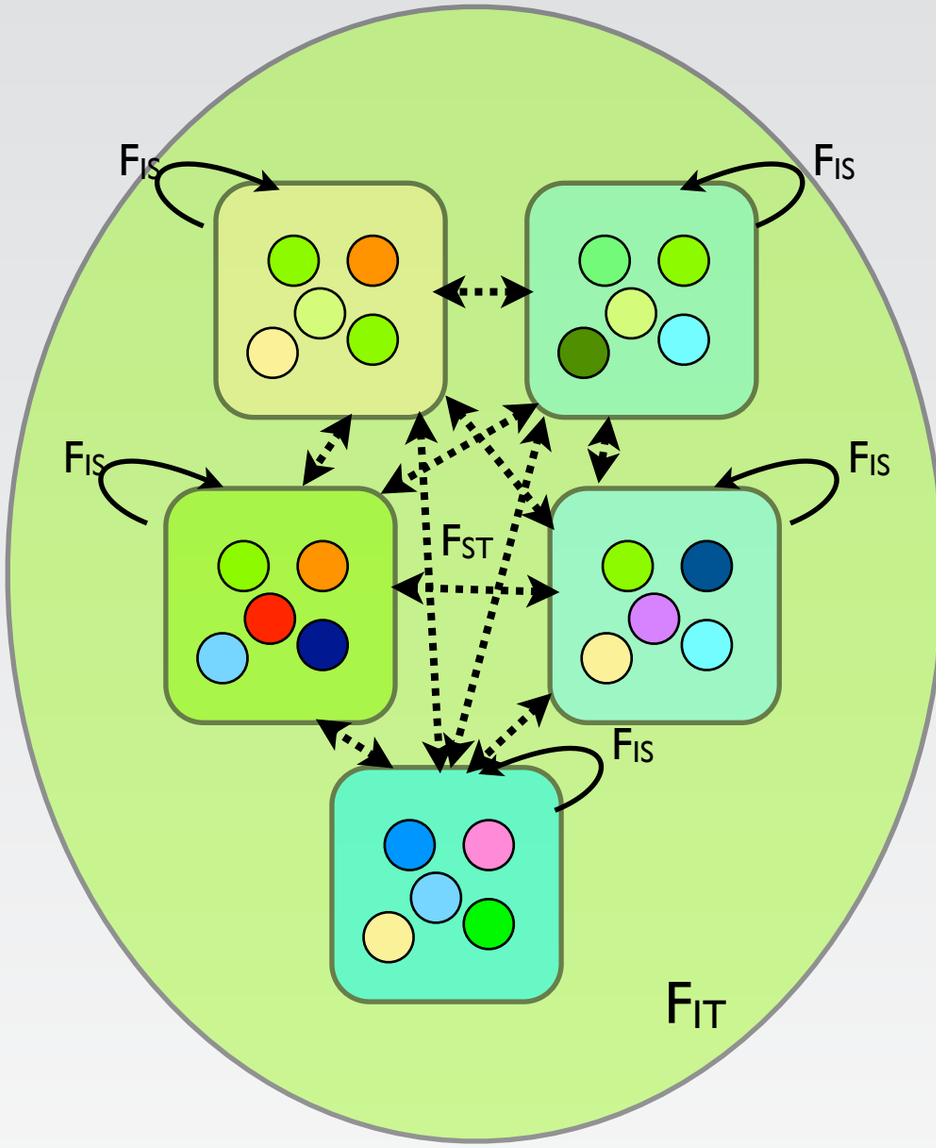
  or $F_{ST} = (F_{IT}-F_{IS})/(1-F_{IS})$

- but inbreeding and heterozygosity related:
  $F = 1-(H_o/H_e)$

  $F_{IS} = 1- (H_I/H_S)$

  $F_{ST} = 1-(H_S/H_T)$

  $F_{IT} = 1-(H_I/H_T)$



H_I = observed heterozygosity averaged across all population fragments

H_S = expected heterozygosity averaged across all population fragments

H_T = expected heterozygosity for the total population ($=H_e$)

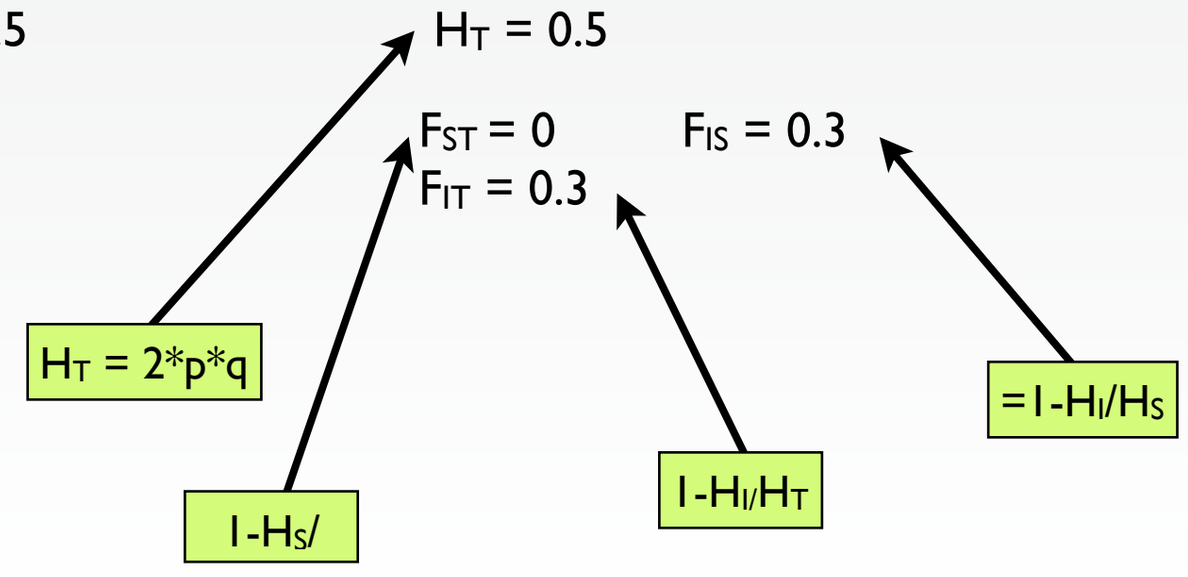
Population differentiation: *F-statistics*

- example 1:

size Pop 1 = size Pop 2

	Genotypes						
Population	A_1A_1	A_1A_2	A_2A_2	Allele frequency	H_o	H_e =2pq	F =1-(H_O/H_E)
1	0.25	0.5	0.25	A_1: p=0.5 A_2: q=0.5	0.5	0.5	0
2	0.4	0.2	0.4	A_1: p=0.5 A_2: q=0.5	0.2	0.5	0.6
mean:		H_I = 0.35				H_S = 0.5	

combined: A_1: p=0.5
A_2: q=0.5



Population differentiation: *F-statistics*

- example 2:

	Genotypes						
Population	A_1A_1	A_1A_2	A_2A_2	Allele frequency	H_o	H_e =2pq	F =1-(H_O/H_E)
1	0.25	0.5	0.25	A_1: p=0.5 A_2: q=0.5	0.5	0.5	0
2	0.14	0.12	0.74	A_1: p=0.2 A_2: q=0.8	0.12	0.32	0.625
mean:						H_I = 0.31	H_S = 0.41
combined:				A_1: p=0.35 A_2: q=0.65		H_T = 0.455	

$p = 2*A_1A_1 + A_1A_2$

F_{ST} = 0.099 F_{IS} = 0.244
F_{IT} = 0.319

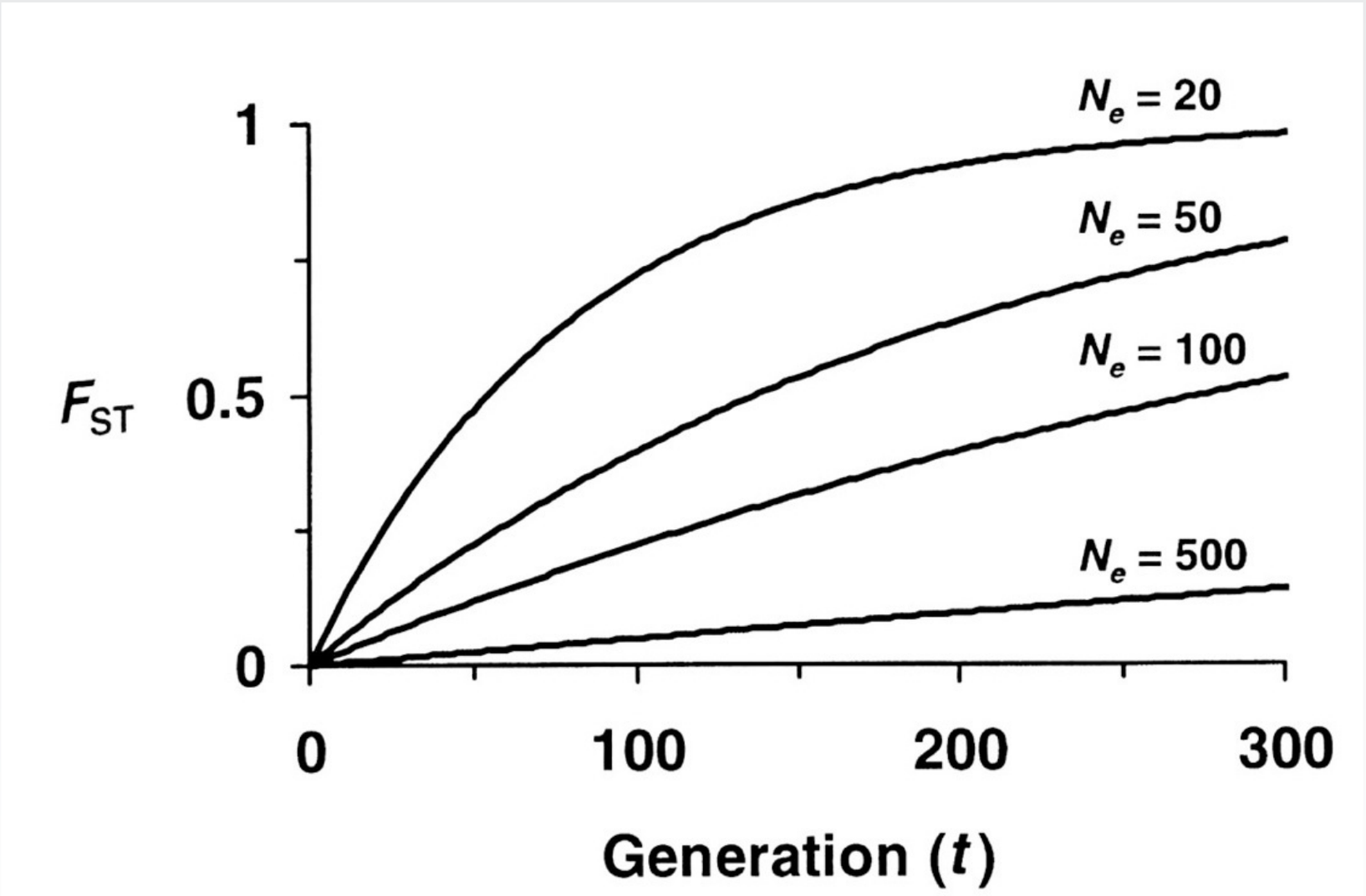
$1-H_S/H_T$

$1-H_I/H_T$

$=1-H_I/H_S$

Population differentiation: *evolution over time*

- when populations are isolated (no gene-flow):
  increase of the genetic differentiation between populations (F_{ST})



Population differentiation: *gene flow*

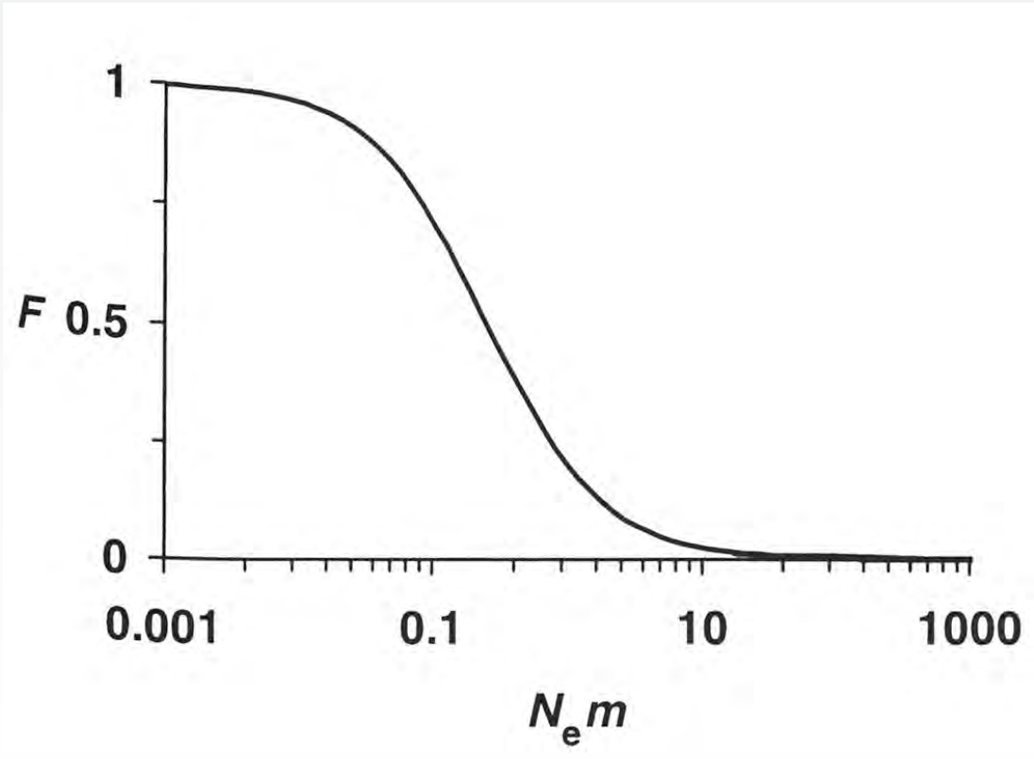
- gene flow reduce the isolation
- gene flow must be sufficient to avoid genetic differentiation
- measuring gene flow: very difficult on the field
  rough estimation using the function:

  $F_{ST} = 1/(4N_em+1)$

N_e= effective population size

m = migration rate

N_em = number of migrant per generation



Relationship between inbreeding, heterozygosity, genetic diversity and population size

- numerous relationships between these parameters
- theory (for random mating populations)
  - relationship between inbreeding and heterozygosity
    $F = 1-(H_o/H_e)$
  - relationship between increase of inbreeding per generation and population size
    $\Delta F = 1/(2N)$
  - loss of genetic diversity ≈ inbreeding coefficient
- in practice (rarely completely random mating in all pop.)
  - large plant populations doing selfing: high inbreeding coefficient, low heterozygosity but high overall genetic diversity (alleles randomly distributed in the population but not within the individuals)
- relationship between inbreeding and loss of genetic diversity more complex in species with regularly high level of inbreeding

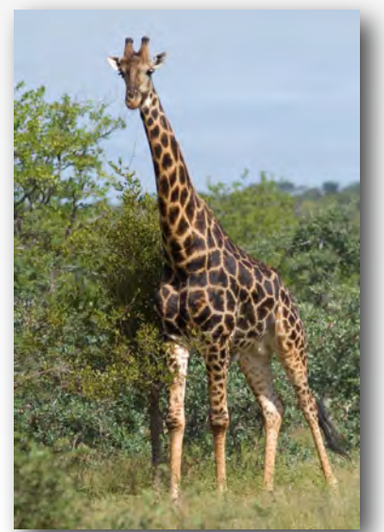
Management of wild populations: *genetic can be used for...*

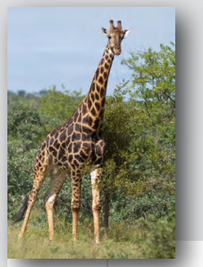
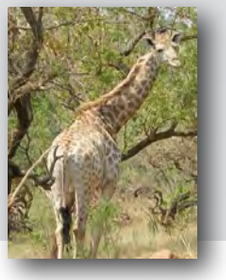
- resolving taxonomic uncertainties, defining management units
- genetically viable populations
- habitat fragmentation
- management of wild populations
- management of captive populations
- reintroduction
- forensic

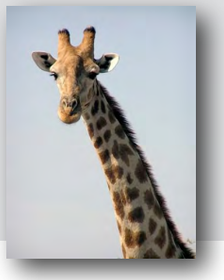
Management of wild populations: *resolving taxonomic uncertainties, defining management units*

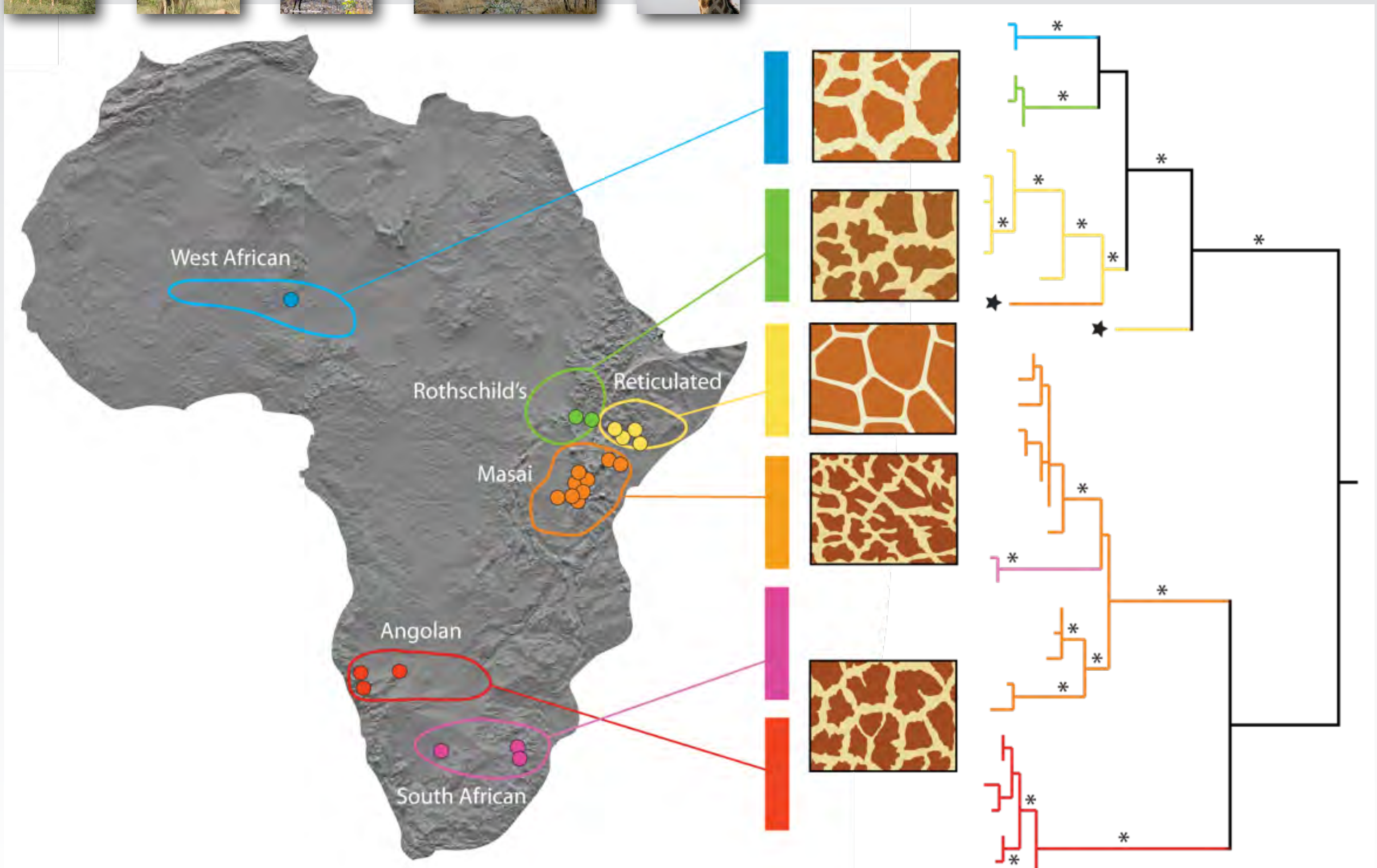
- large scale genetic structure

  e.g. : Giraffe (*Giraffa camelopardalis*)

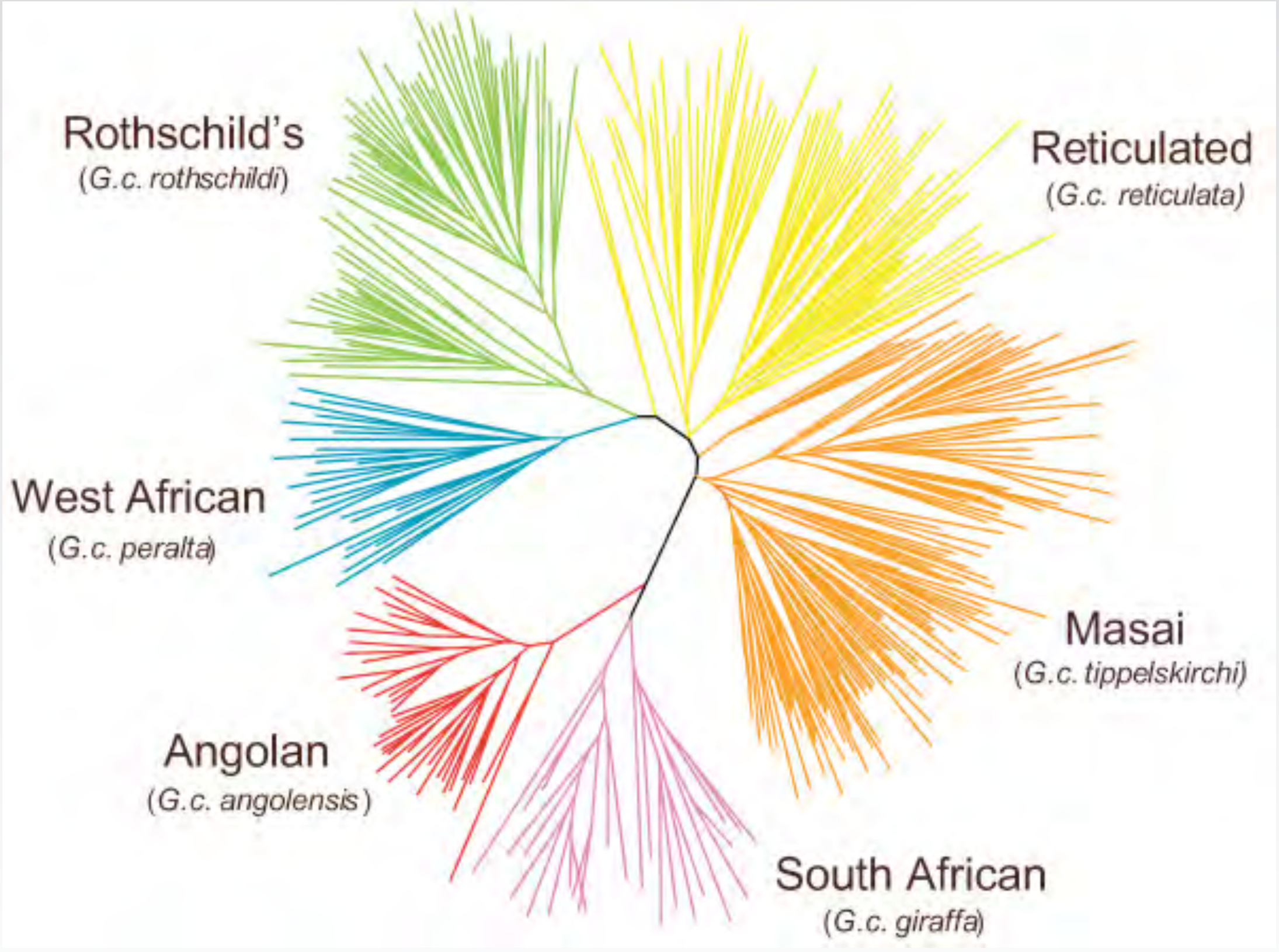




Brown et al. (2007) Extensive population genetic structure in the giraffe, BMC Biology 5:57

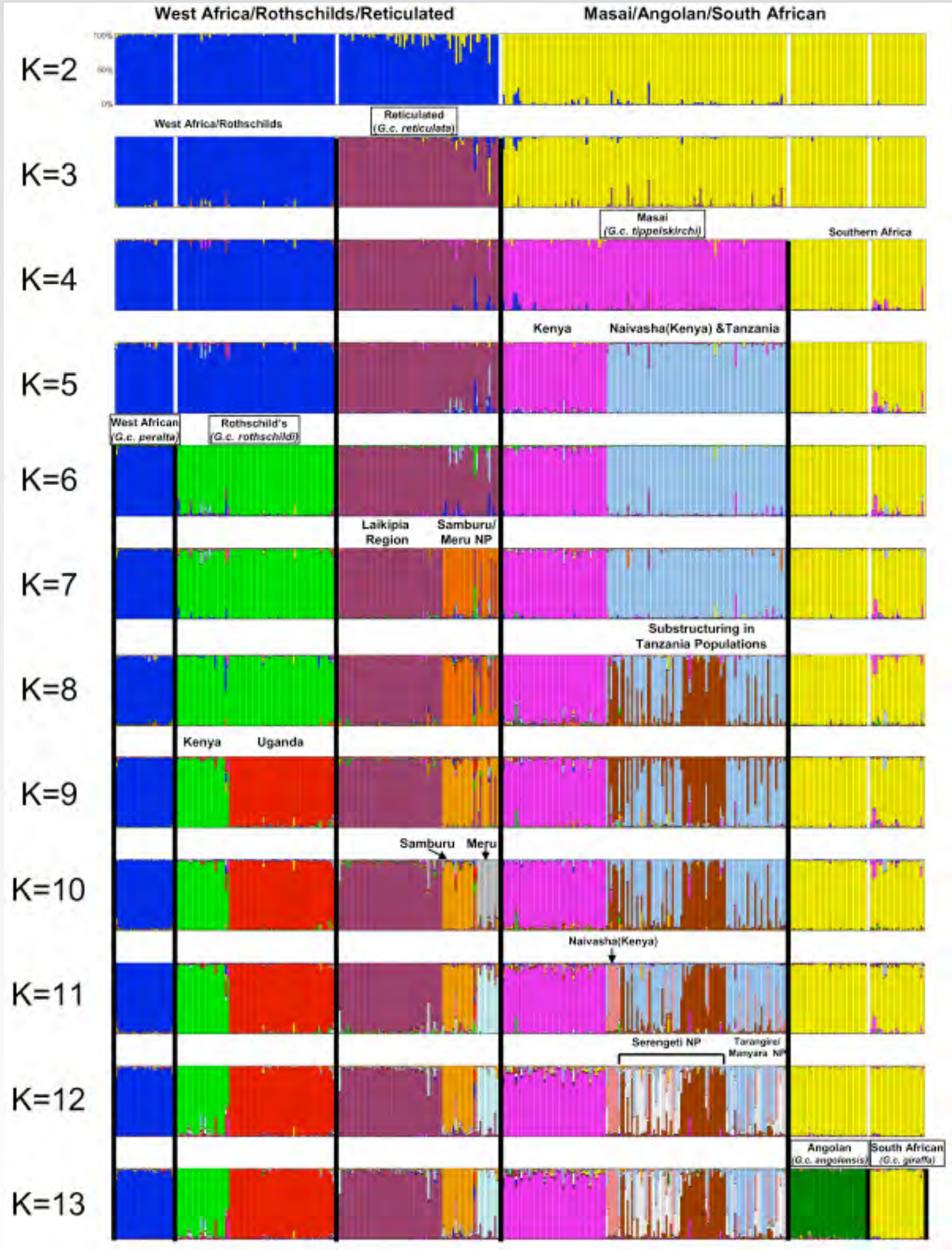


Genetic subdivision in the giraffe based on microsatellites alleles. Neighbor-joining network of allele-sharing distances (D_s) based on 14 microsatellite loci typed in 381 giraffes. Colors are coded as in Figure 1A.

Brown et al. BMC Biology 2007 5:57

Genetic subdivision among giraffe groups and populations based on Bayesian cluster analysis [23] of 14 microsatellite loci from 381 individuals. Shown are the proportions of individual multilocus genotypes attributable to clusters (K) indicated by different colors. Sample group designations and sampling locations are denoted. We varied K from 2–16 and at least six groups corresponding to currently defined subspecies and 11 geographic clusters are resolved as indicated.

Brown et al. BMC Biology 2007 5:57

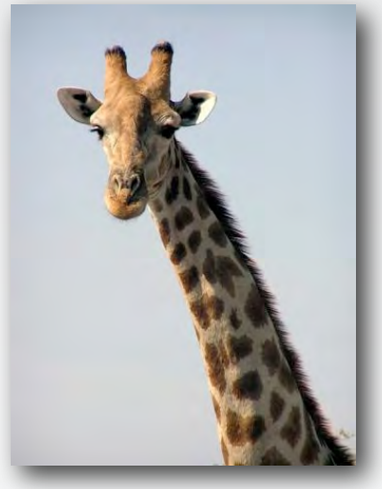


Management of wild populations: *resolving taxonomic uncertainties, defining management units*

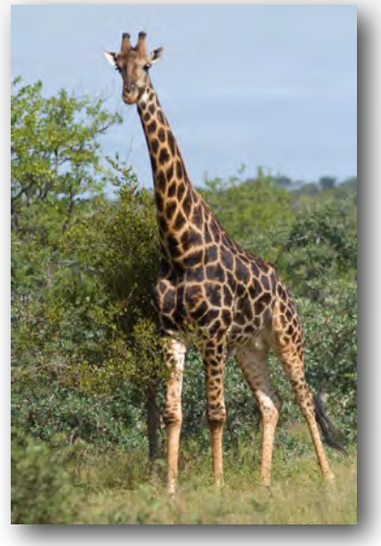
- large scale genetic structure

  e.g. : Giraffe (*Giraffa camelopardalis*)

  - previously considered as one unique Evolutionary Significant Unit
  - after this work: 6 genetically distinct lineages, with limited interbreeding between them.
    6 ESU
  - possible presence of different species (parapatric subspecies, e.g. among Masai, reticulated and Rothschild's giraffes)



  - previous conservation status: Low Risk for the UICN Red List
  - some genetically isolates groups: highly endangered



Management of wild populations: *genetically viable populations*

- how large do populations need to be to ensure their genetic "health"?
  - avoid inbreeding depression
  - ability to evolve in response to environmental changes
  - avoid accumulation of deleterious mutations

- based on different simulations: variable results...

goal	N_e	Recovery time (in generation)
retain reproductive fitness	50	
retain evolutionary potential	500	100-1000
	5000	
	570-1250	
retain single locus genetic diversity	100'000 - 1'000'000	100'000-10'000'000
avoid accumulating deleterious mutations	1000	
	100	
	12	

Franklin (1980), Soulé (1980), Lande & Barrowclough (1987), Lande (1995), Franklin & Frankham (1998), Lynch et al. (1995), Charlesworth et al (1993)

Management of wild populations: *examples*

- adder (*Vipera berus*)
  - small venomous snake
  - limited dispersal

  - Madsen and collaborators (Sweden)

Table 1. Effective population sizes of reproducing adults in an isolated population of adders, *Vipera berus*, at Smygehuk in southern Sweden

Year	No. of adult females	No. of reproductive females	No. of adult males	No. of males that mated	Total no. of adults	No. of reproductive adults	Effective population size[a]
1984	13	9	25	13	38	22	21·27
1985	17	1	23	2	40	3	2·67
1986	11	5	23	13	34	18	14·44
1987	22	14	20	16	42	30	29·87
1988	17	6	20	12	37	18	16·00
1989	22	15	19	18	41	33	32·73
1990	17	4	17	10	34	14	11·43
Mean values	17·0	7·71	21·0	12·0	38·0	19·71	Arithmetic = 18·34 Harmonic = 9·9

[a]Effective population size is calculated from Wright's (1940) equation to compensate for skewed sex ratios: $4N_mN_f/(N_m + N_f)$ where N_m = number of males, N_f = number of females.

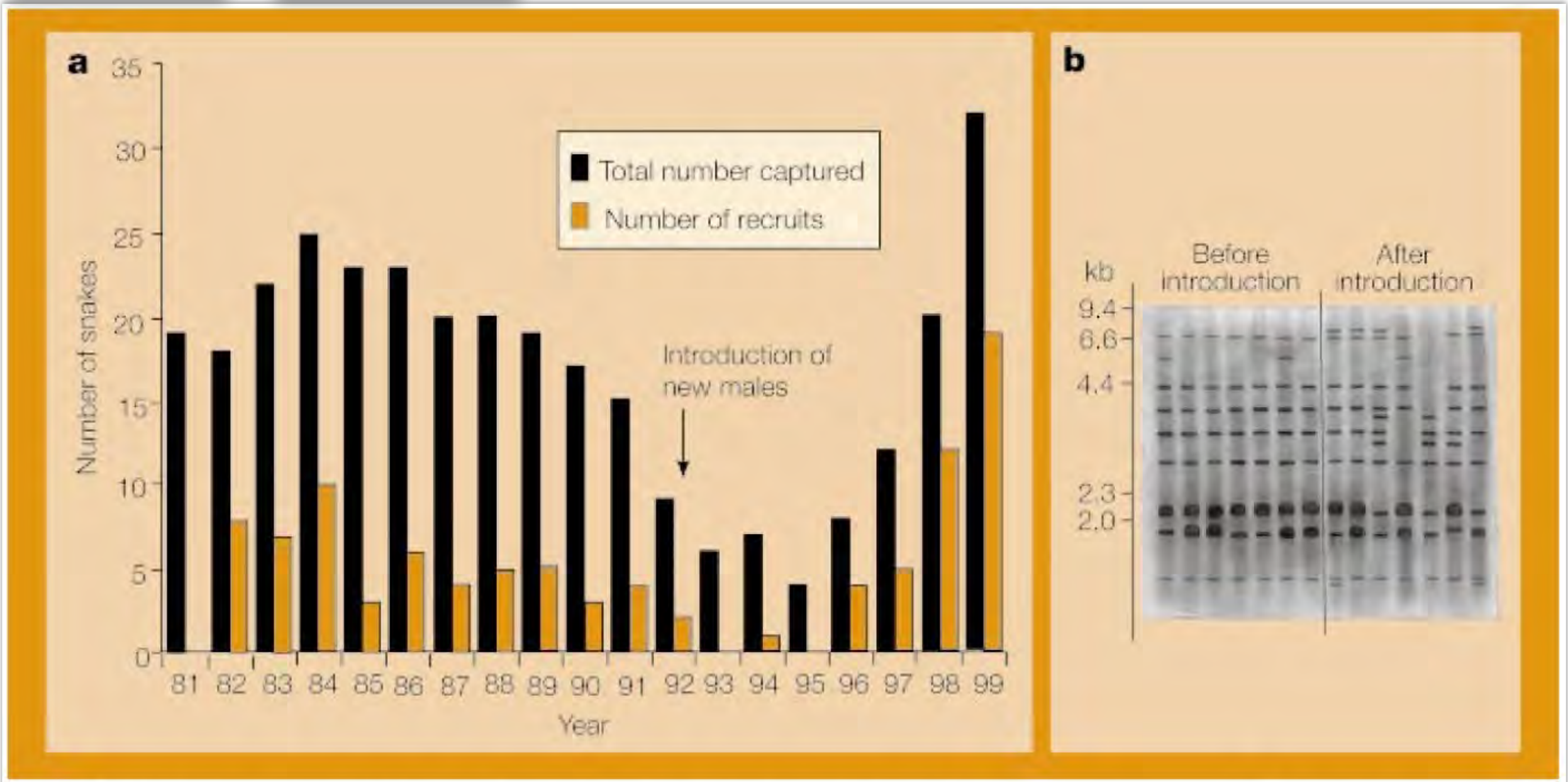


Figure 1 Introducing new males increases the genetic diversity and enables the adder population to recover. **a,** Total number and number of recruited male adders captured in Smygehuk from 1981 to 1999. **b,** Southern-blot analysis of major histocompatability complex (MHC) class I genes in seven males sampled before the introduction of new males (left) and in seven recruited males sampled in 1999 (right).

Madsen et al. (1999) Restoration of an inbred adder population, Nature 402, 34-35

Management of wild populations: *examples*

	Smygehuk	Genarp	Captivity
population	isolated (Ne<20)	not close	
body condition	similar (p=0.34)		
Pesticides	no significantly differences		
genetic variation (2 enzymes and with DNA fingerprinting)	significant lower genetic diversity (p>0.001) higher band-sharing (similar bands) in Smygehuk (p<0.001))		
neonate mortality rate	0.31	0.09	0.05
	significantly higher in Smygehuk (p=0.0001)		
offspring mass	similar (p=0.73)		
brood size	significantly smaller for Smygehuk		

Madsen et al., 1996

- 3 females from Smygehuk cross with males from Genarp: neonate mortality rate = 0

Management of wild populations: *examples*

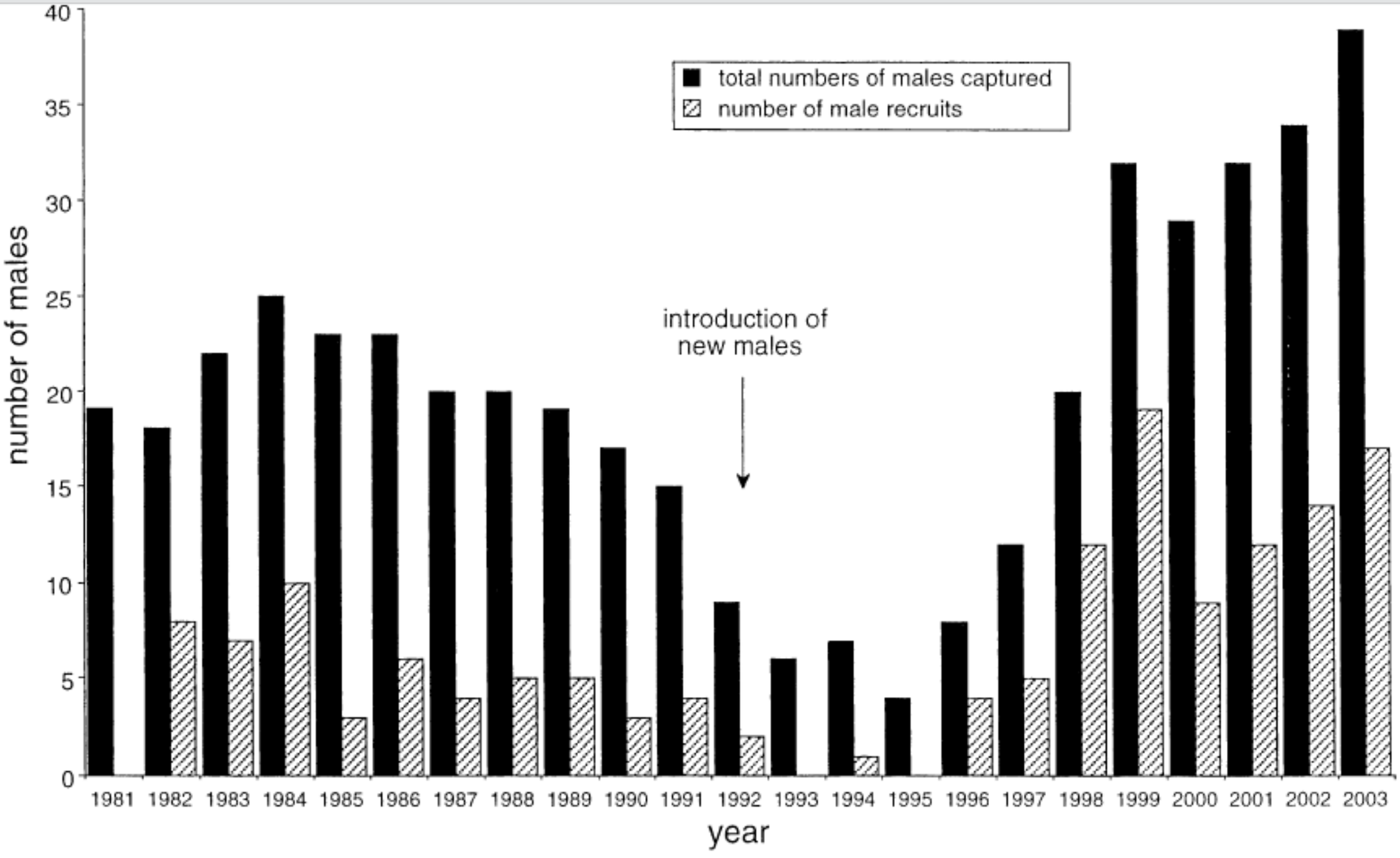


Fig. 1. Total number and number of recruiting males captured in Smygehuk from 1981 to 2003. The introduced males are not included in the figure.

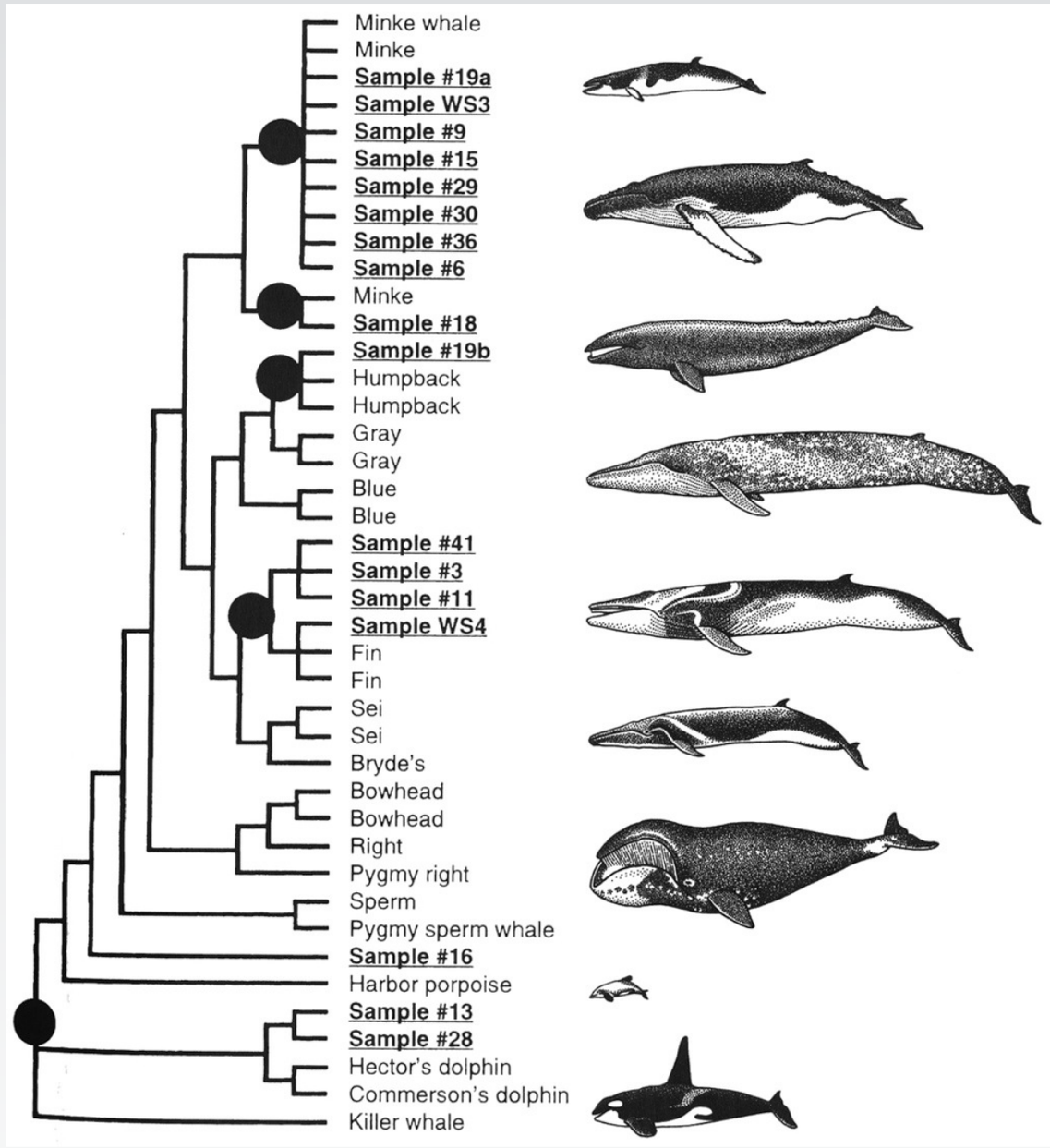
Madsen et al., (2004) Novel genes continue to enhance population growth in adders (*Vipera berus*) , Biological Conservation

Forensic

- can help to determine illegal hunting
- estimation of bottlenecks
- estimation of effective size
- parentage analyses
- determination of sex
- hybridisation
- ...

Forensic: *illegal hunting*

- e. g.: meat of whales
  - Baker & Palumbi (1996)
    - using portable PCR laboratory, installed in a hotel room
    - amplification of the control region (mtDNA), sequencing in their lab.
    - 16 samples taken
    - 9 = mink whale (legal "scientific" whaling)
    - 1 = humpback whale, 4 = fin whales
      3 = dolphins
  - Dizon et al (2000)
    - 954 samples
    - 773 = whales, 9% coming from protected species
    - rest: dolphins, porpoises, sheep and horses






Forensic: *hybridisation*

- e. g. hybridisation between introduced and native species
  - domestic cat and wild cats (Randi, 2008)

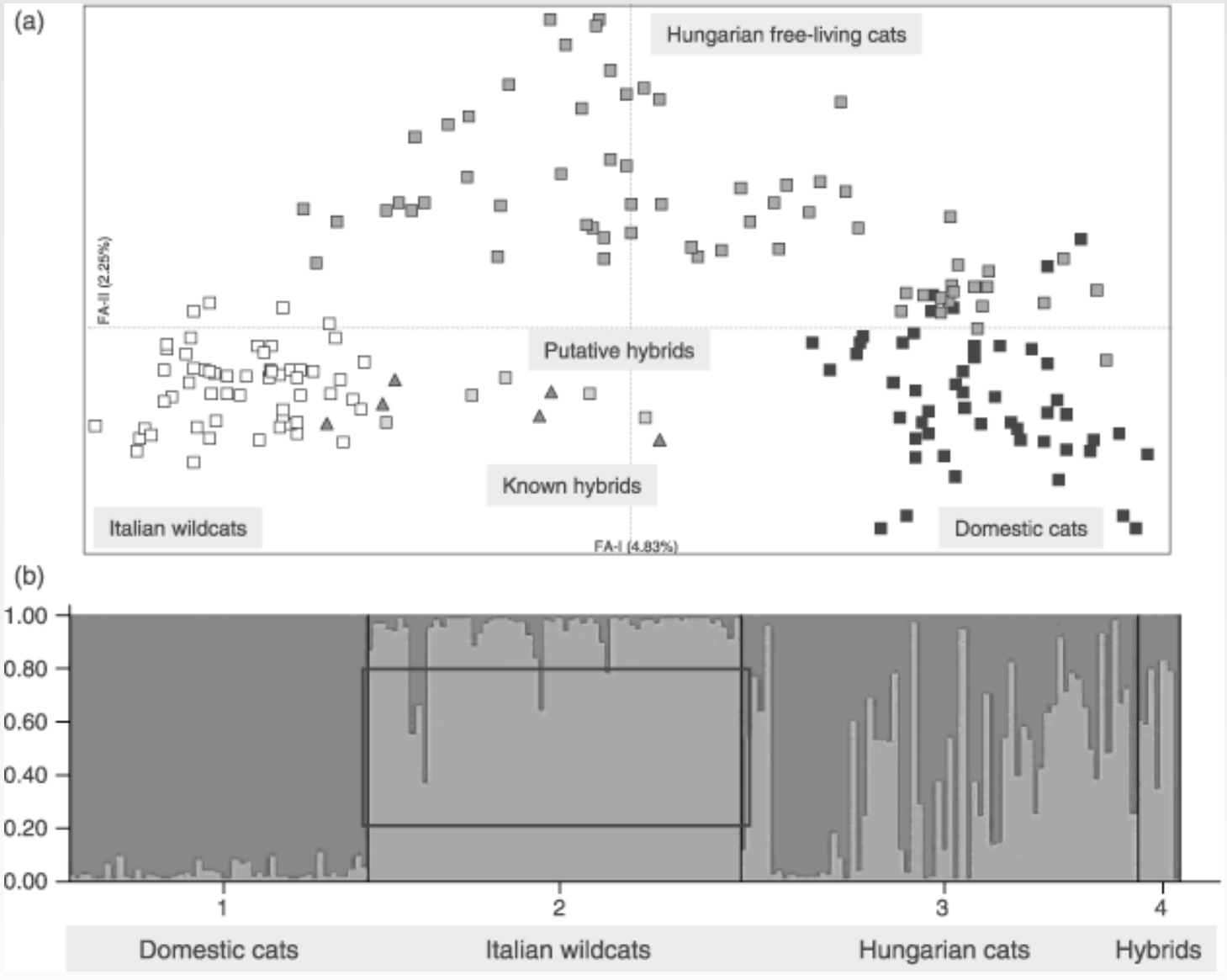
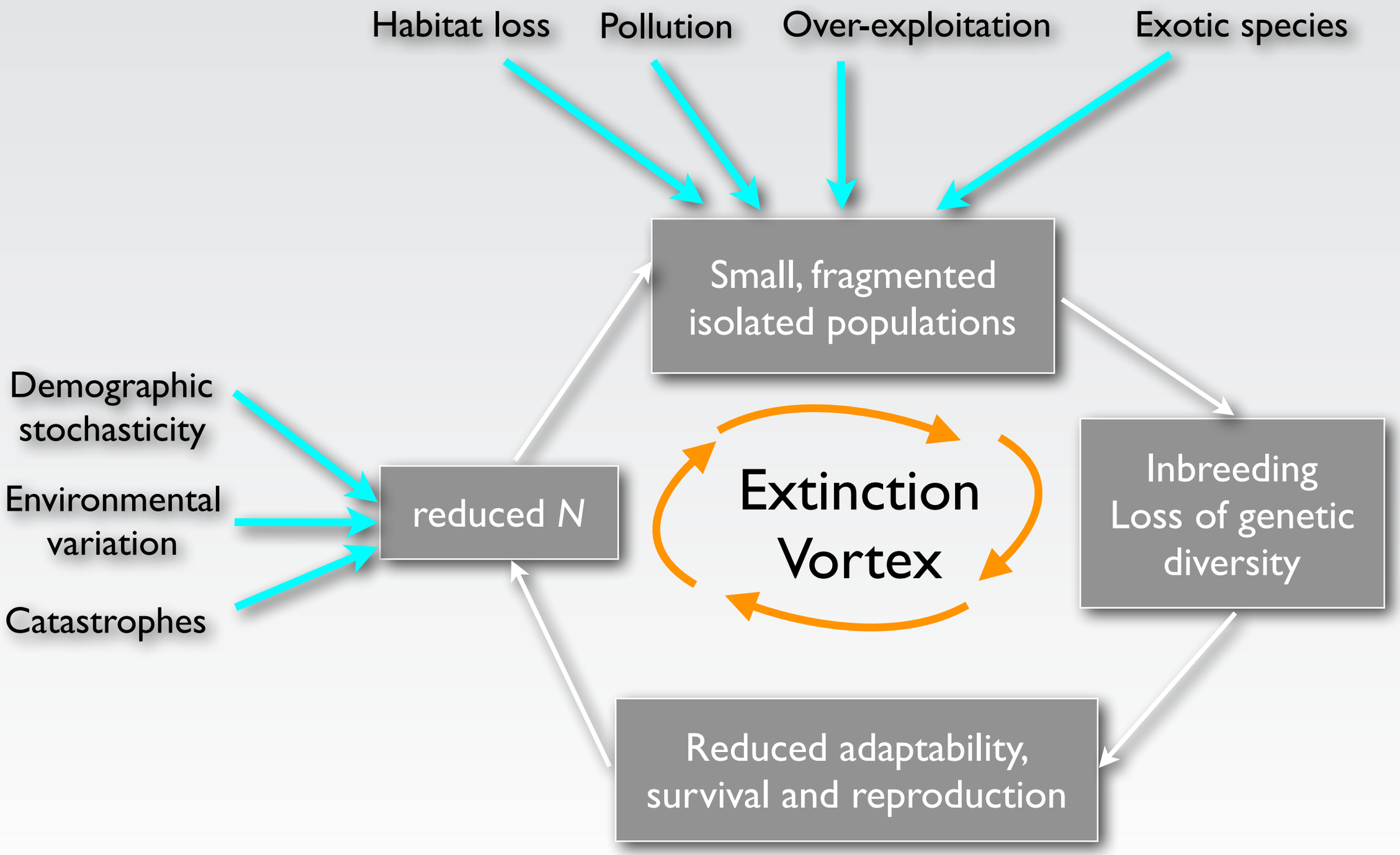


Fig. 2 (a) Factorial correspondence analysis showing relationships among the multilocus genotypes of individual Italian wildcats, freeranging Hungarian cats, domestic cats and putative or known hybrids. FA-I and FA-II are the first and second principal factors of variability. (b) structure analyses (performed assuming k = 2 distinct genetic clusters) of multilocus microsatellite genotypes of domestic cats, Italian wildcats, Hungarian free-ranging cats and captive-reproduced hybrids. Admixed genotypes, at threshold qi = 0.80, are evidenced.

General conclusions: *Extinction vortex*



Frankham et al. (2002) Introduction to Conservation Genetics, Cambridge University Press

General conclusions

- reduction of genetic diversity inevitable in small closed populations
- to maintain genetic diversity
  - populations must be big enough
    - Ne > 50 to avoid inbreeding depression
    - Ne = 500-5000 to retain evolutionary potential
  - gene flow must be maintain between populations
    - with about 1 migrant between pop. per generation, genetic drift is limited
  - manage to reduce genetic drift in fragmented populations and the lost of genetic diversity within them
- why?
  - genetic factors generally contribute to extinction risk
  - inbreeding has deleterious effect on reproduction and survival
  - loss of genetic diversity reduce the ability to adapt in response to environmental changes
- resulting mainly from fragmentation and population size reduction

General conclusions

- genetics: interesting tools to answer several questions
  - global genetic structure
    - species
    - ESU
  - population history
    - phylogeny / phylogeography
    - bottleneck detection
  - population structure
    - population differentiation / migration
    - inbreeding, inbreeding depression
  - forensic
    - e.g. illegal hunting
  - better understanding of the biology of the species
    - pedigree, paternity assignment
    - ....

General conclusions

- genetics: interesting tools to answer several questions

BUT

- must be combined with other approaches
  - biology of the studied species
  - history
  - Population viability analysis (PVAs)
    - as well as combining with e. g. genetic impacts of inbreeding
  - ...

for conservation, genetic is just a tool

additional information

- books
  - Frankham, Ballou & Briscoe (2002) Introduction to Conservation Genetics, Cambridge University Press
  - Allendorf & Luikart (2007) Conservation and the Genetics of Populations, Blackwell Publishing
- articles
  - inbreeding: Keller & Waller (2002) Inbreeding effects in wild populations, **TRENDS in Ecology & Evolution** 17: 230-241
  - analyses softwares: Excoffier & Heckel (2006) Computer programs for population genetics data analysis: a survival guide, **Nature Reviews Genetics** 7:745-758
- technical and analyses
  - DNA manipulation (PCR, sequencing, etc.): http://www.dnai.org/b/index.html
  - softwares: e. g. http://www.biology.lsu.edu/general/software.html
    http://evolution.genetics.washington.edu/phylip/software.html

Vorlesung mit Übungen: Conservation Genetics (2 KP)

Herbst Semester

1. Introduction
2. Non-invasive sampling and techniques
3. Next-generation sequencing and its applications to conservation biology
4. Mitochondrial DNA: taxonomic aspects in conservation
5. Metabarcoding & Environmental DNA
6. Hardy-Weinberg Equilibrium: detection of and reasons for HWE bias
7. Inbreeding and inbreeding depression
8. Outbreeding depression
9. Genetic drift and genetic of small populations
10. Bottleneck
11. Population fragmentation
12. Sex-biased dispersal
13. Landscape genetics
14. Captive populations and reintroduction
15. Approximate Bayesian Computation
16. Analyses with some datasets

+ exercises, presentation of journal articles, ..
applied examples
"playing" with some datasets, use of most important softwares, ...

Exercise: *Hardy-Weinberg equilibrium*

- 1)

	100/100	100/75	75/75	total
	1	22	7	30
observed frequency				1

prop 100:
prop 75:

Test of HW equilibrium: via X^2 test

	100/100	100/75	75/75	total
observed	1	22	7	30
expected following Hardy-Weinberg				30
chi-square				

d.f. = p =

Exercise: *Hardy-Weinberg equilibrium*

- 2)

	A1A1	A1A2	A1A3	A2A2	A2A3	A3A3	Total
	4	8	15	6	21	28	82
observed frequency							1

prop A1:
prop A2:
prop A3:

Test of HW equilibrium: via X^2 test

	A1A1	A1A2	A1A3	A2A2	A2A3	A3A3	Total
observed	4	8	15	6	21	28	82
expected freq. if Hardy-Weinberg Equilibrium							1
expected							82
chi-square							

d.f. = p =