Genetic Variation:
What it is and how to summarize it

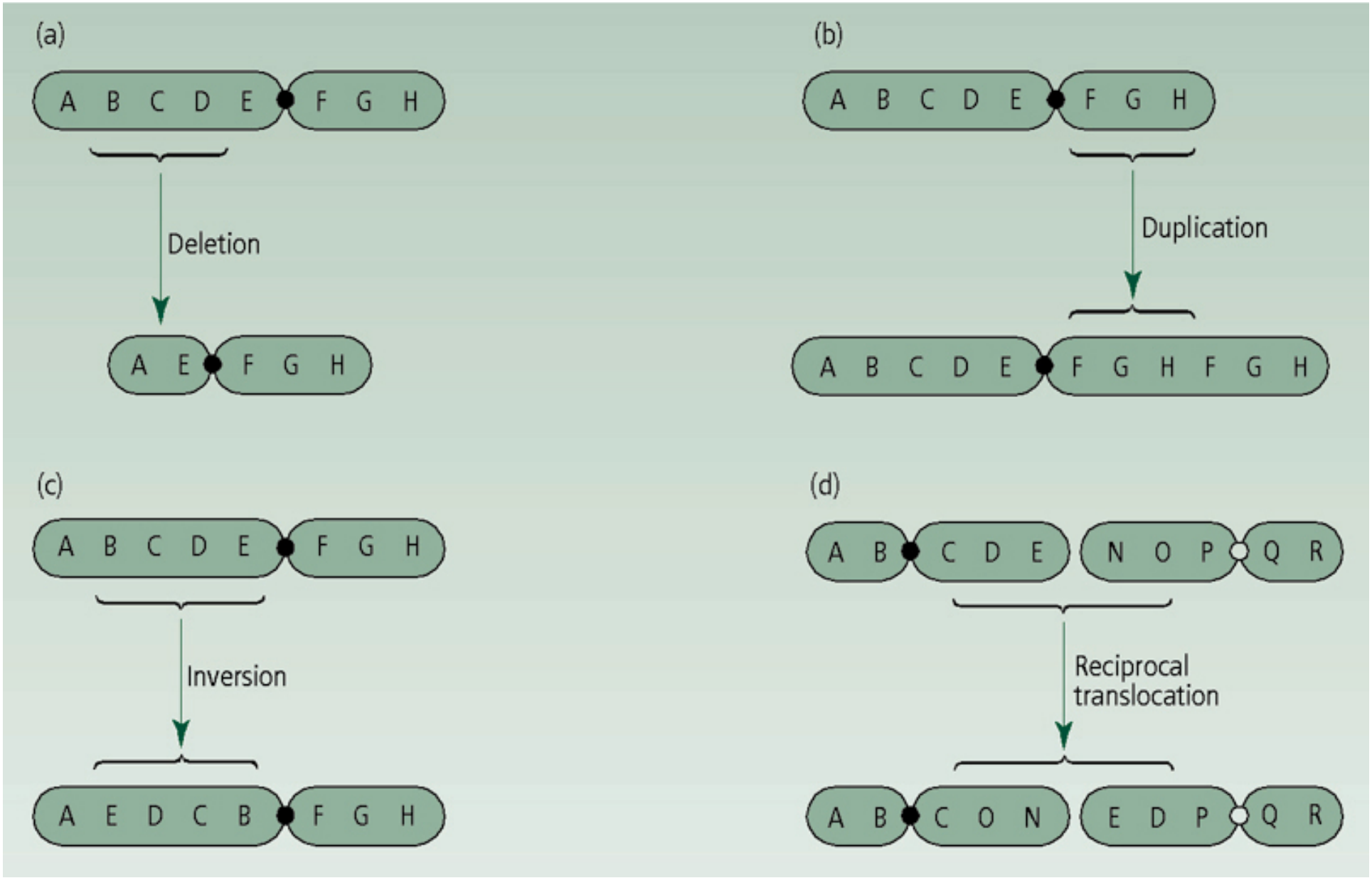
Genetic Variation

- Chromosomal variation
- Protein variation
- DNA variation

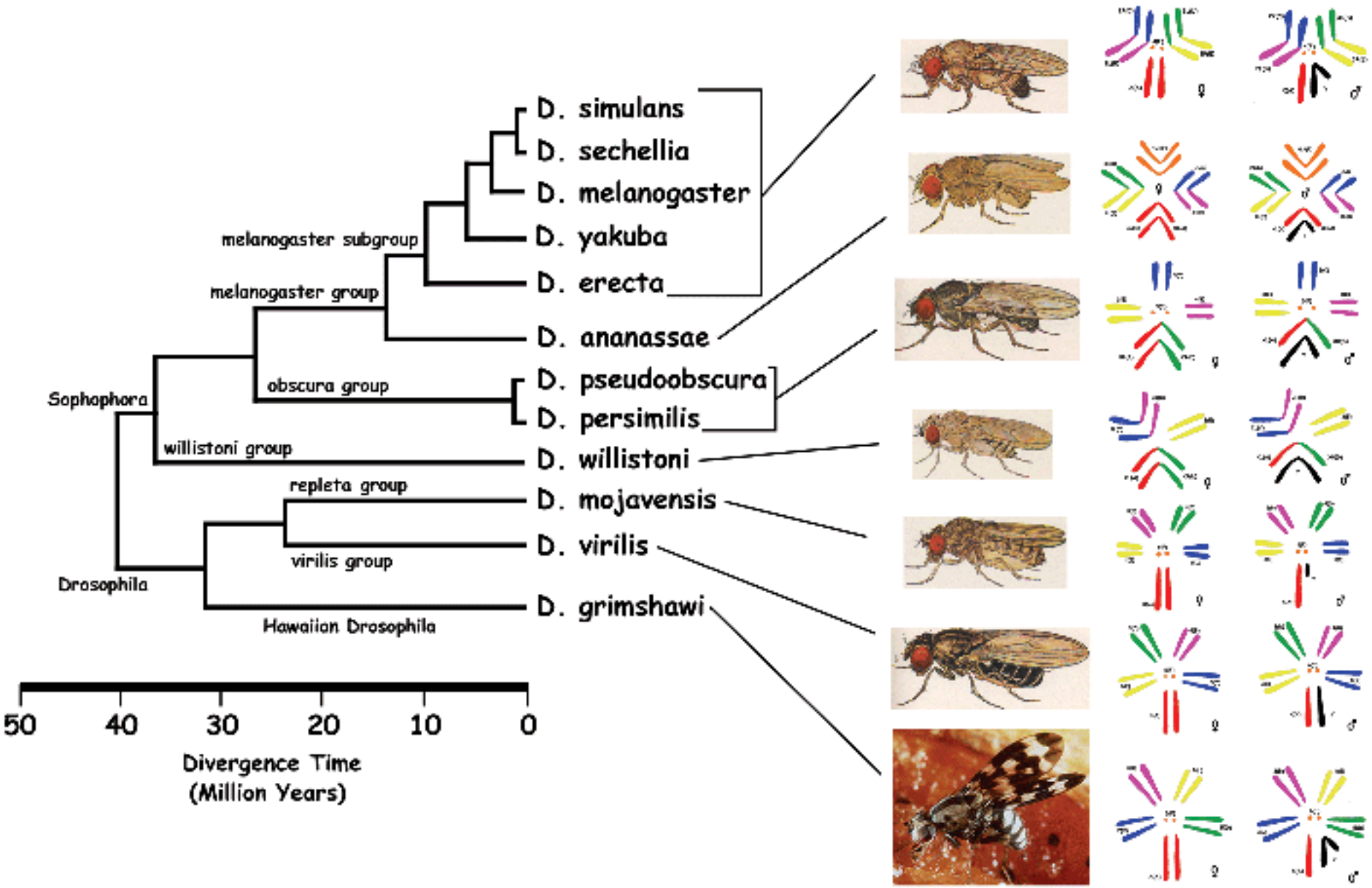
Chromosomal Variation

- Variation in chromosome number, gene number, gene order etc.

Chromosomal rearrangements

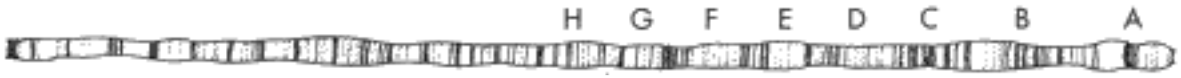


Chromosomal Variation

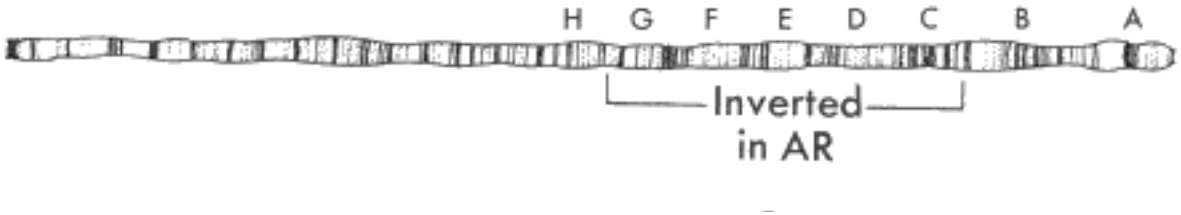


Chromosomal Variation

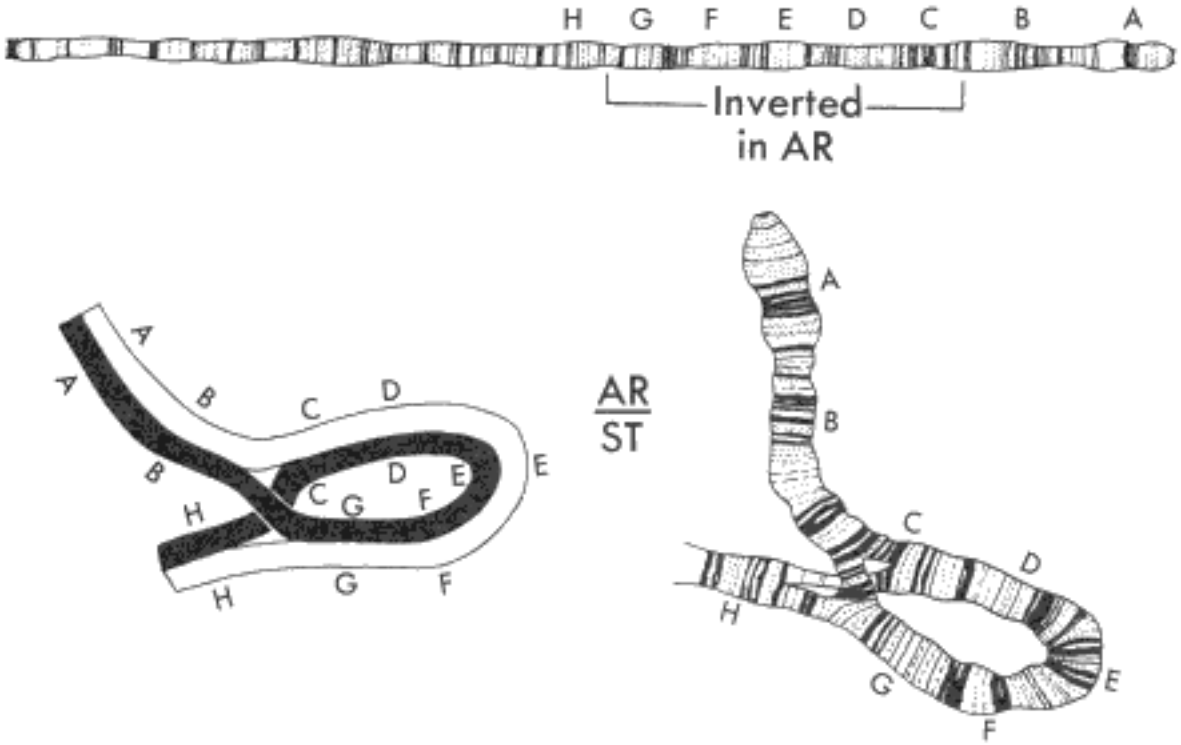
Chromosomal Variation



Chromosomal Variation



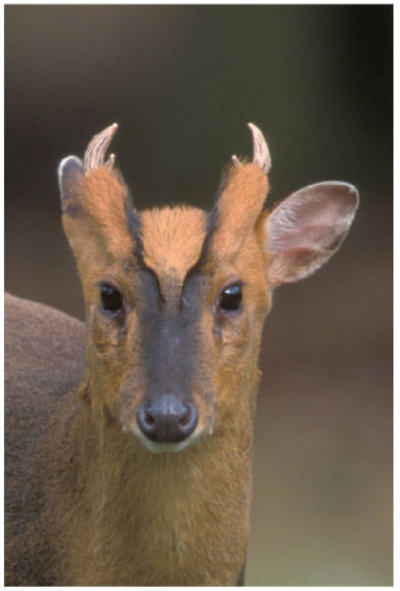
Chromosomal Variation



From Dobzhansky and Sturtevant

Chromosomal number changes

Muntiacus reevesii ($2N = 46$)

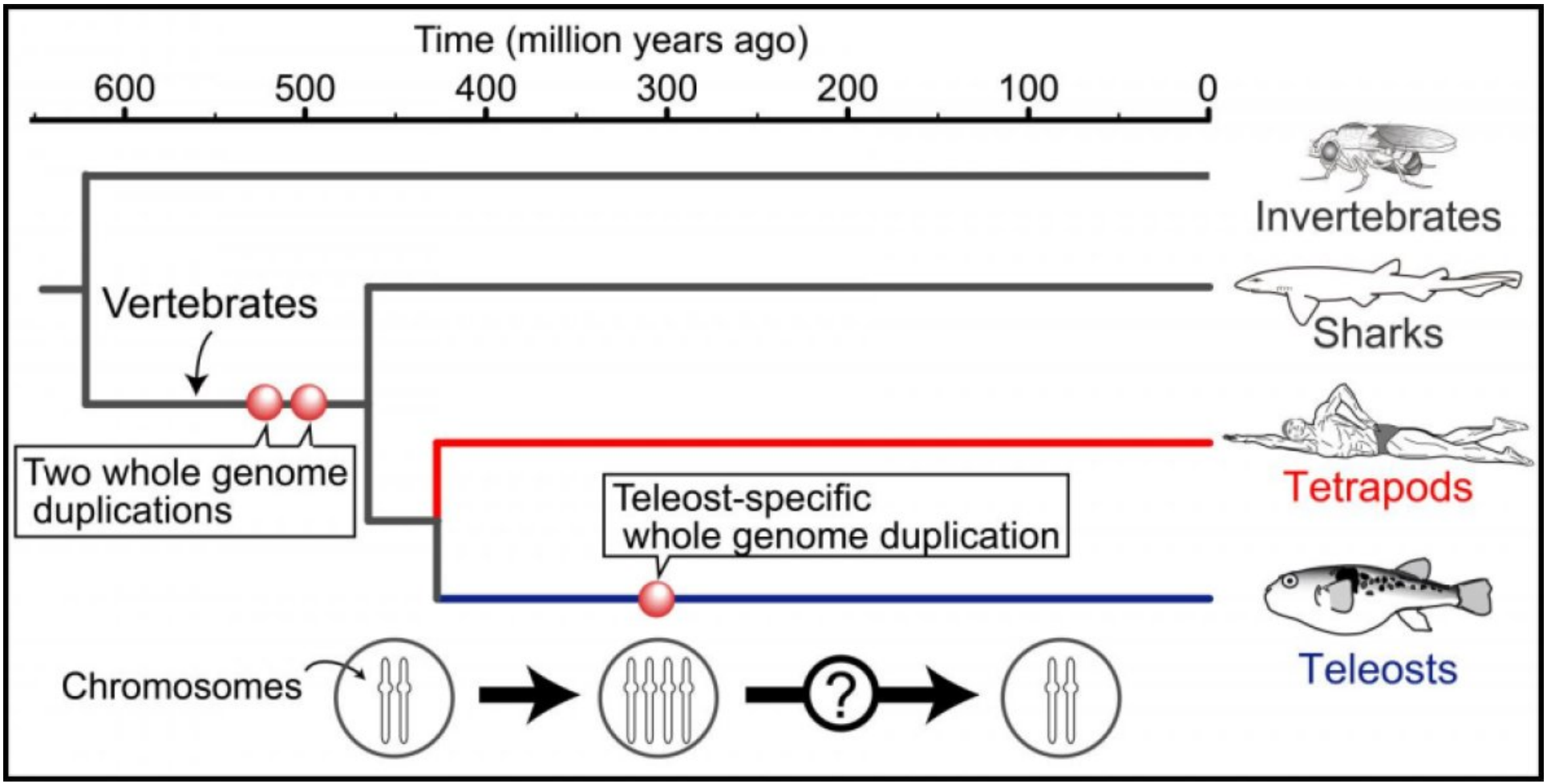


Muntiacus muntiacus ($2N = 8$)

EVOLUTION 2e, Figure 8.26

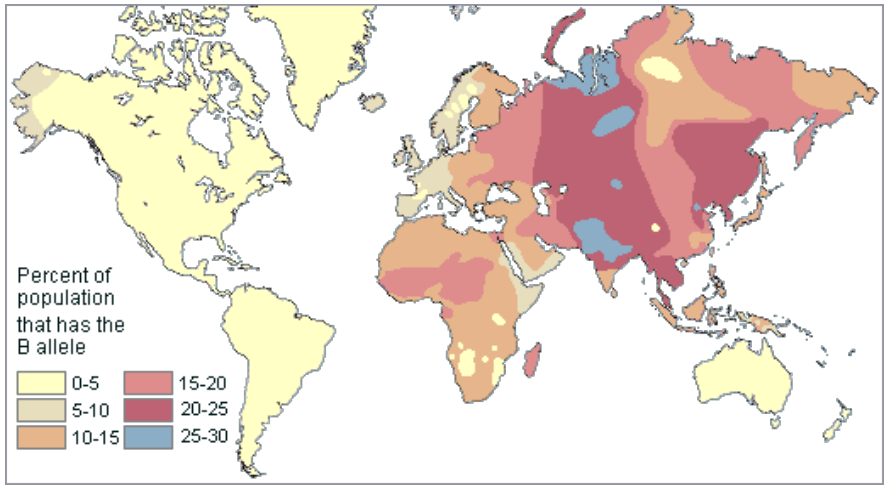
© 2009 Sinauer Associates, Inc.

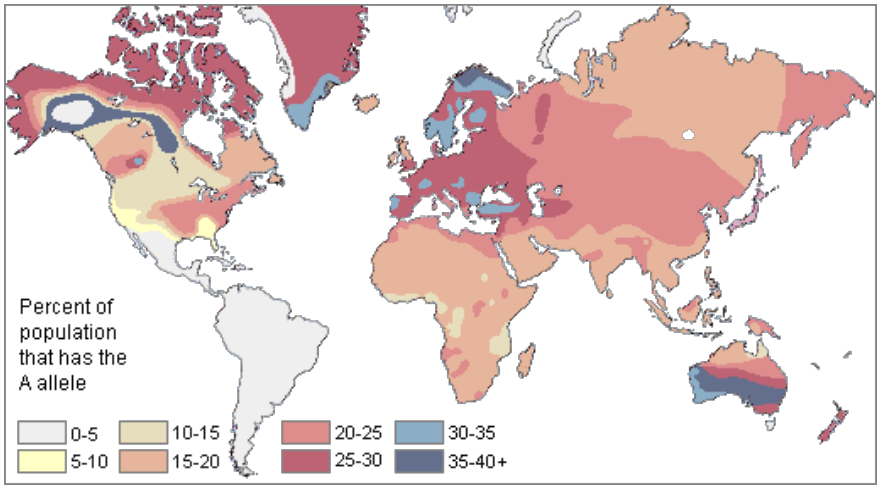
Genome Duplication

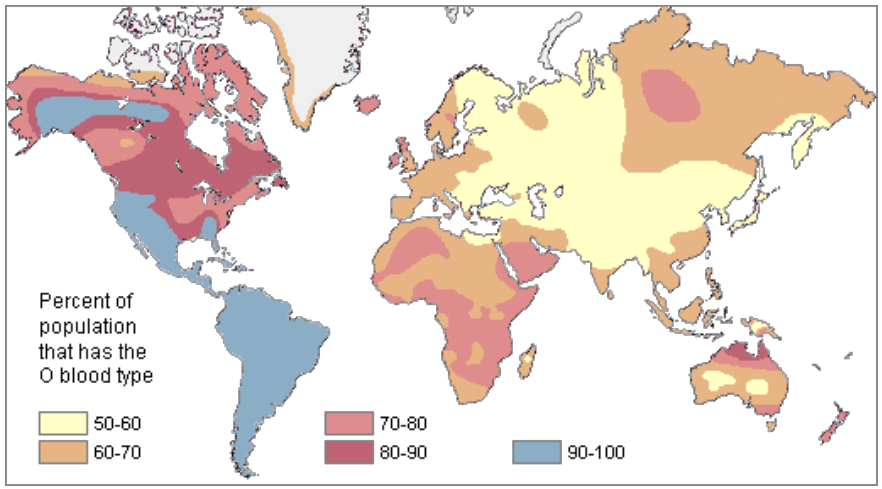


Protein variation

- ABO blood groups







Protein variation

- Allozymes: variant forms of an enzyme encoded by different alleles at the same locus
- Variation revealed using electrophoresis

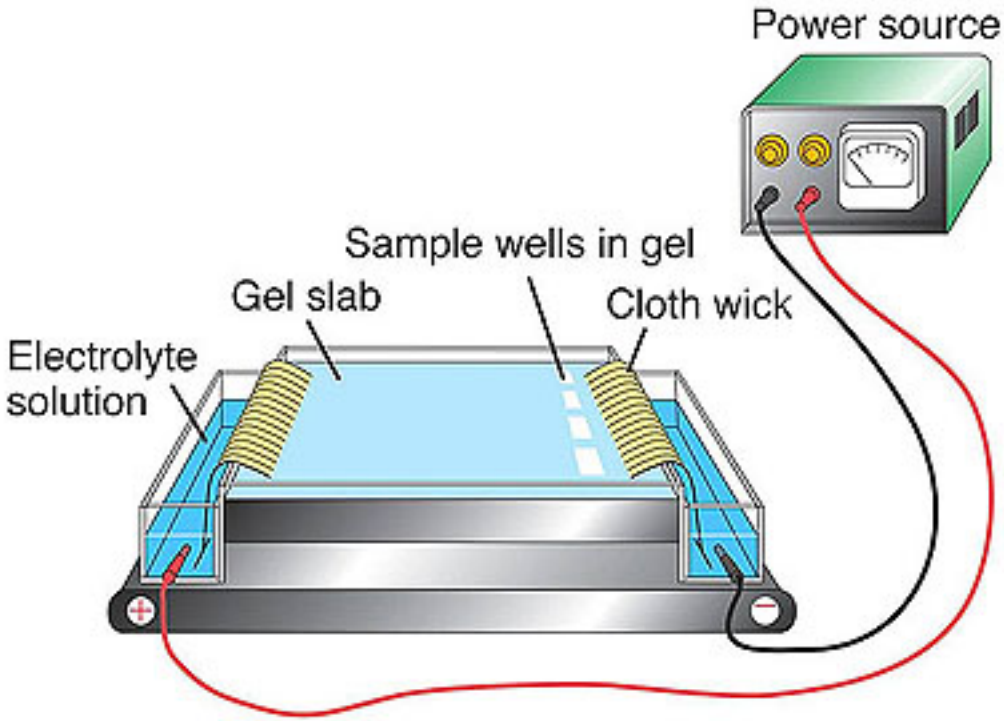
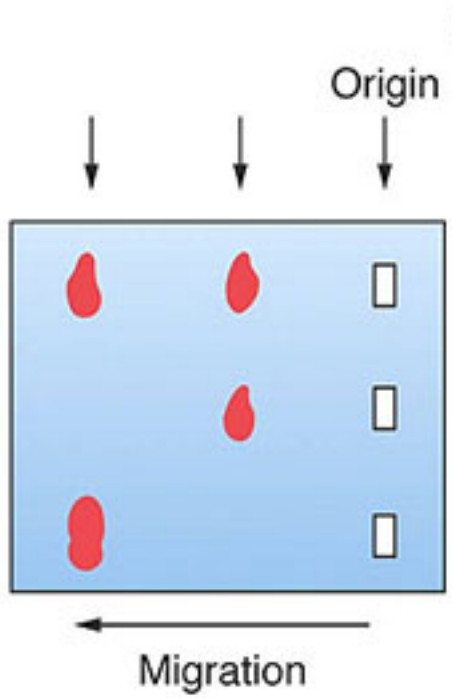


Figure 4: Schematic of devices used in protein electrophoresis

Protein variation

- Allozymes: variant forms of an enzyme encoded by different alleles at the same locus
- Variation revealed using electrophoresis

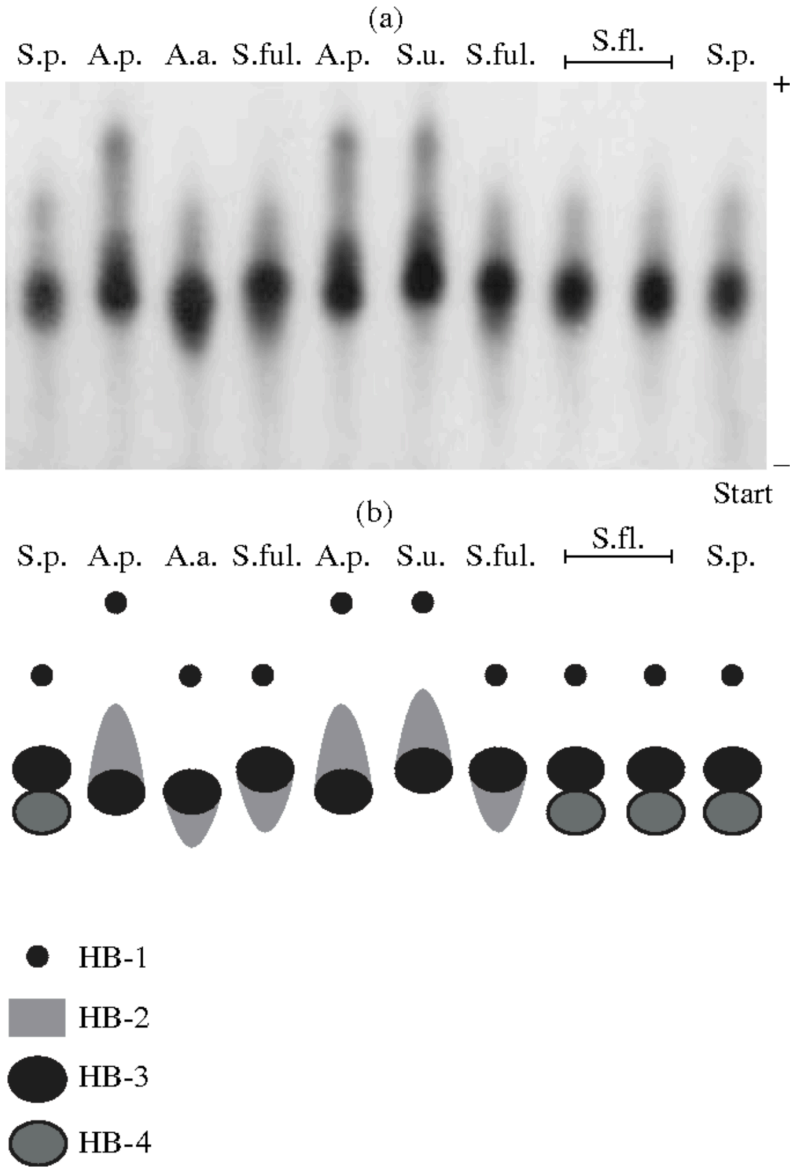


How does it work?

- Nonsynonymous mutations can change enzyme's overall ionic charge
- Leads to differences in electrophoretic mobility

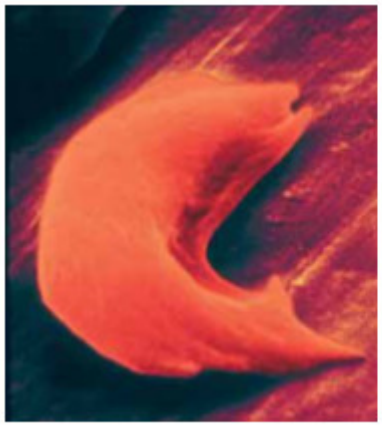
Amino Acid	3-Letter	1-Letter	Side chain polarity	Side chain charge (pH 7.4)
Alanine	Ala	A	nonpolar	neutral
Arginine	Arg	R	polar	positive
Asparagine	Asn	N	polar	neutral
Aspartic acid	Asp	D	polar	negative
Cysteine	Cys	C	nonpolar	neutral
Glutamic acid	Glu	E	polar	negative
Glutamine	Gln	Q	polar	neutral
Glycine	Gly	G	nonpolar	neutral
Histidine	His	H	polar	positive(10%), neutral(90%)
Isoleucine	Ile	I	nonpolar	neutral
Leucine	Leu	L	nonpolar	neutral
Lysine	Lys	K	polar	positive
Methionine	Met	M	nonpolar	neutral
Phenylalanine	Phe	F	nonpolar	neutral
Proline	Pro	P	nonpolar	neutral
Serine	Ser	S	polar	neutral
Threonine	Thr	T	polar	neutral
Tryptophan	Trp	W	nonpolar	neutral
Tyrosine	Tyr	Y	polar	neutral
Valine	Val	V	nonpolar	neutral

Protein variation

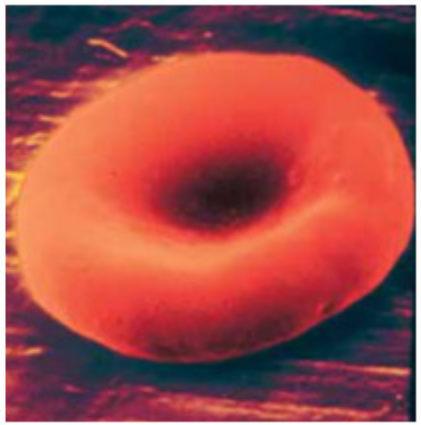


Protein variation

- Amino acid variation
  - Alternative forms of proteins arising from variation in the amino acid sequence
    - Sickle-cell disease (HbS): ONE amino acid change in beta-globin chain of hemoglobin



Sickle-cell phenotype



Normal phenotype

DNA variation

- RFLP: Restriction fragment length polymorphism
  - Created by mutation that changes a restriction site

```
GCCGCATTCTA
CGGCGTAAGAT
```

```
GCCGAATTCTA
CGGCTTAAGAT
```

DNA variation

- RFLP: Restriction fragment length polymorphism
  - Cr n site

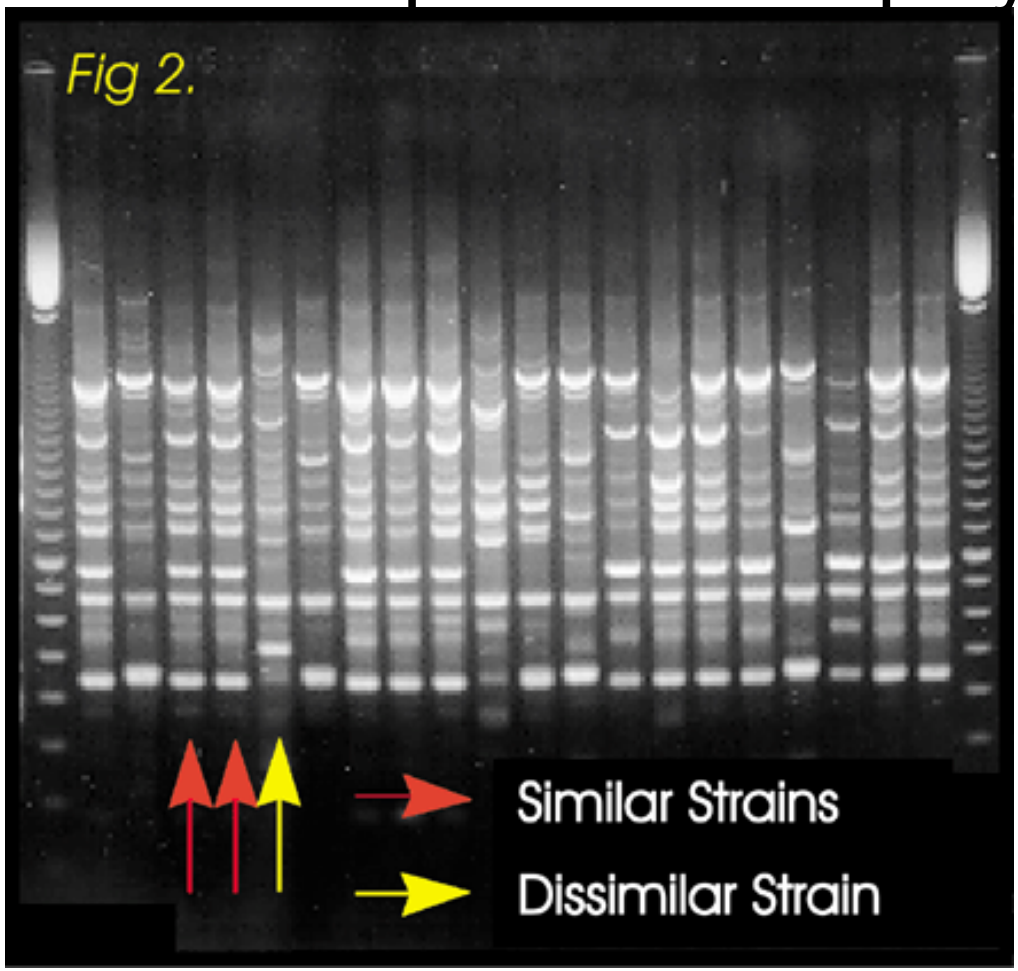
DNA variation

- RFLP
- RAPD: Random amplification of polymorphic DNA
  - Like PCR, but segments are amplified randomly
  - Employs several arbitrary, short primers
  - Need no knowledge of underlying sequence
  - Variation in RAPD profile comes from variation in primer binding sites across individuals

DNA variation

- RFLP
- RAPD: Random amplification of polymorphic DNA

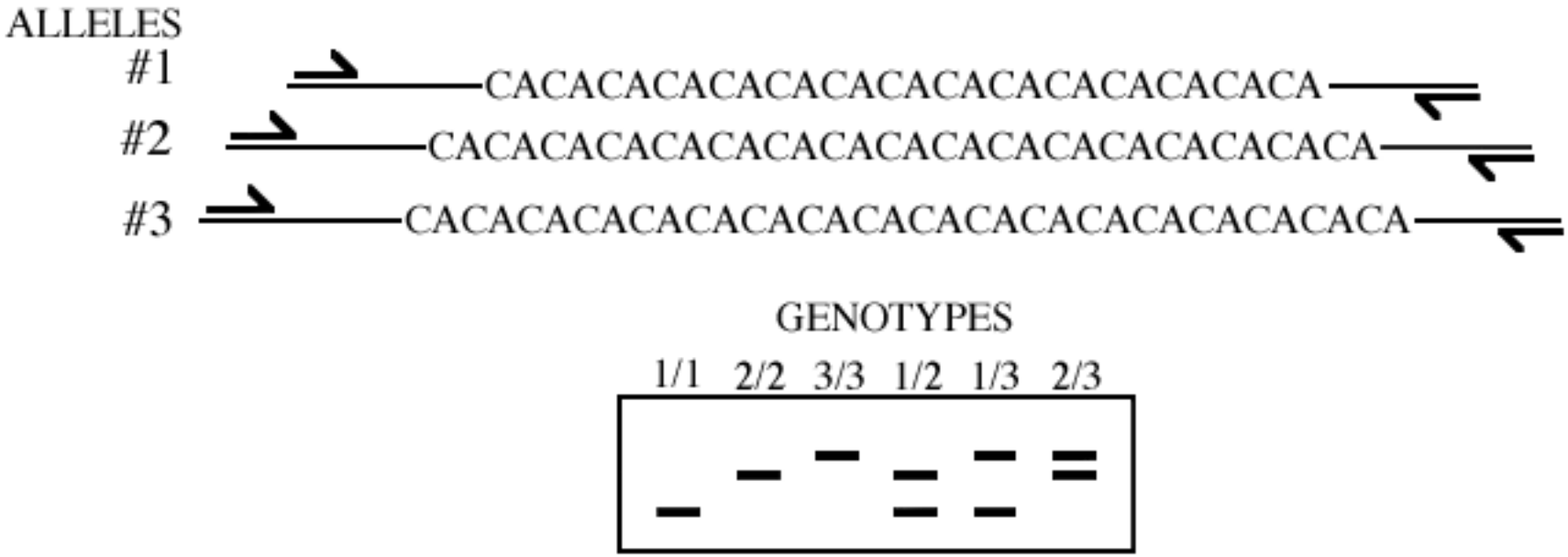
Strains of Lactobacillus from 18 types of Cheddar Cheese



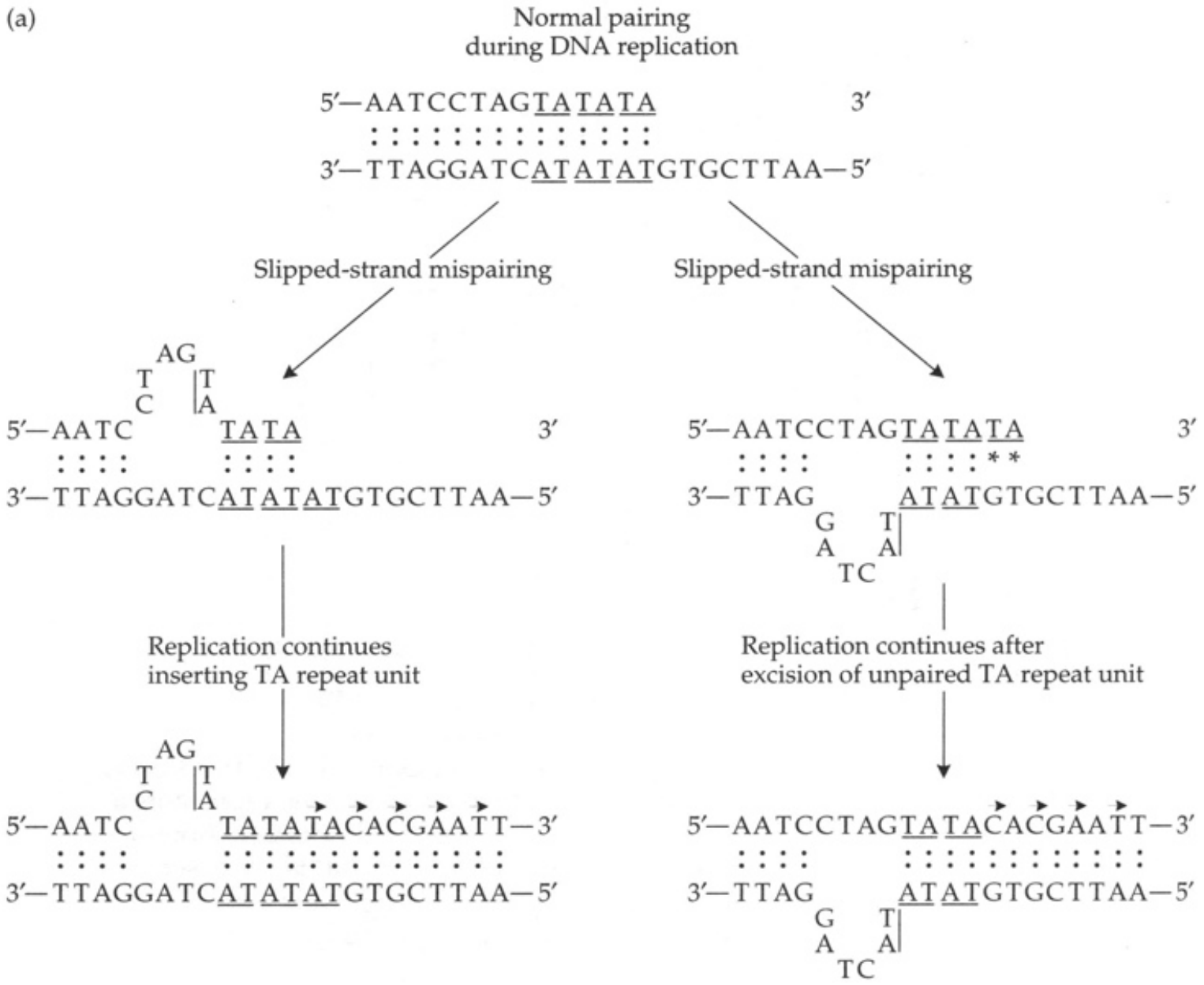
Teagasc 1998

DNA variation

- RFLP
- RAPD
- Microsatellites/minisatellites/VNTRs/SSRs: Tandemly repeated short sequences

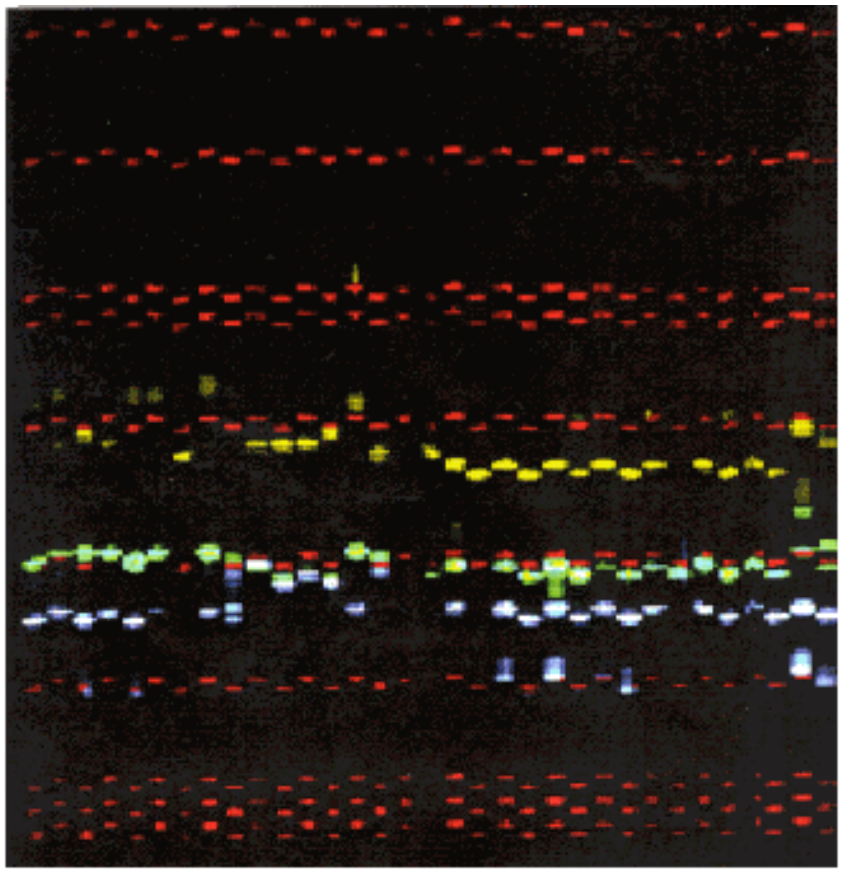


Errors in mispairing during replication and DNA repair



DNA variation

- RFLP
- RAPD
- Microsatellites/minisatellites/VNTRs/SSRs:



Leishmania
(Viannia) isolates

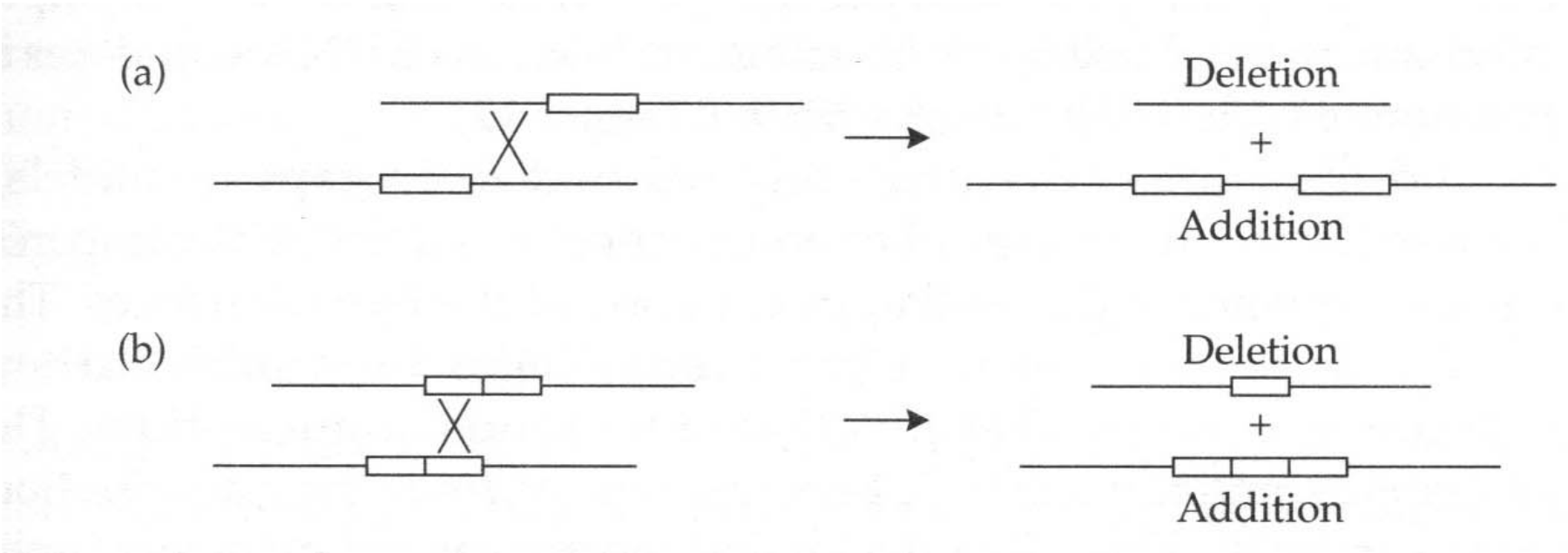
From London School of Hygiene and Tropical Medicine

DNA variation

- RFLP
- RAPD
- Microsatellites/minisatellites/VNTRs/SSRs
- Insertion/Deletion: Gain or loss of DNA segment

From Wei et al. 2011

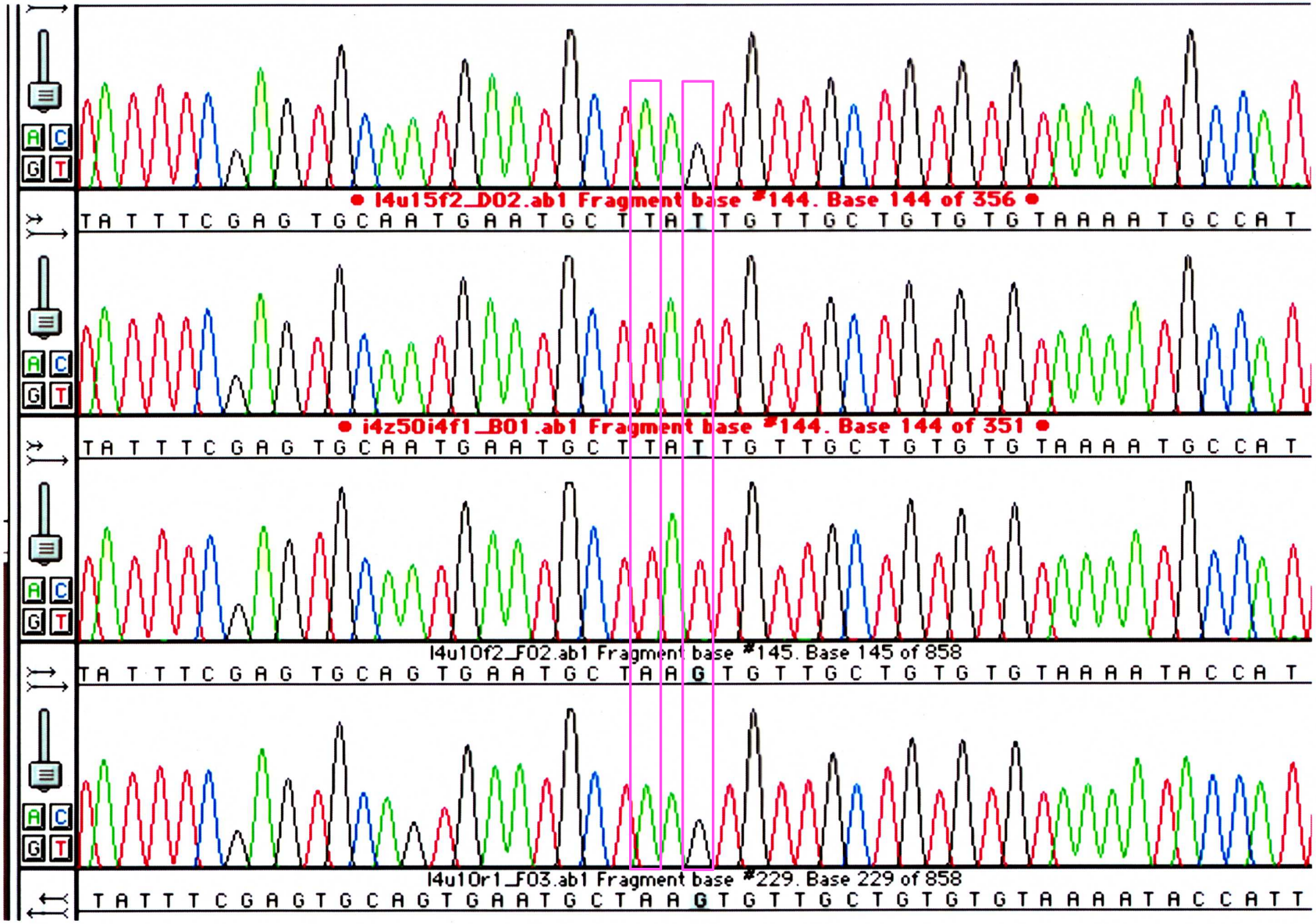
Small Scale Insertion & Deletions (InDels)



DNA variation

- RFLP
- RAPD
- Microsatellites/minisatellites/VNTRs/SSRs
- Insertion/Deletion
- Single Nucleotide Polymorphism: Differences at a single nucleotide

Sequence reads from 4 individuals



Discrete Genetic Variation

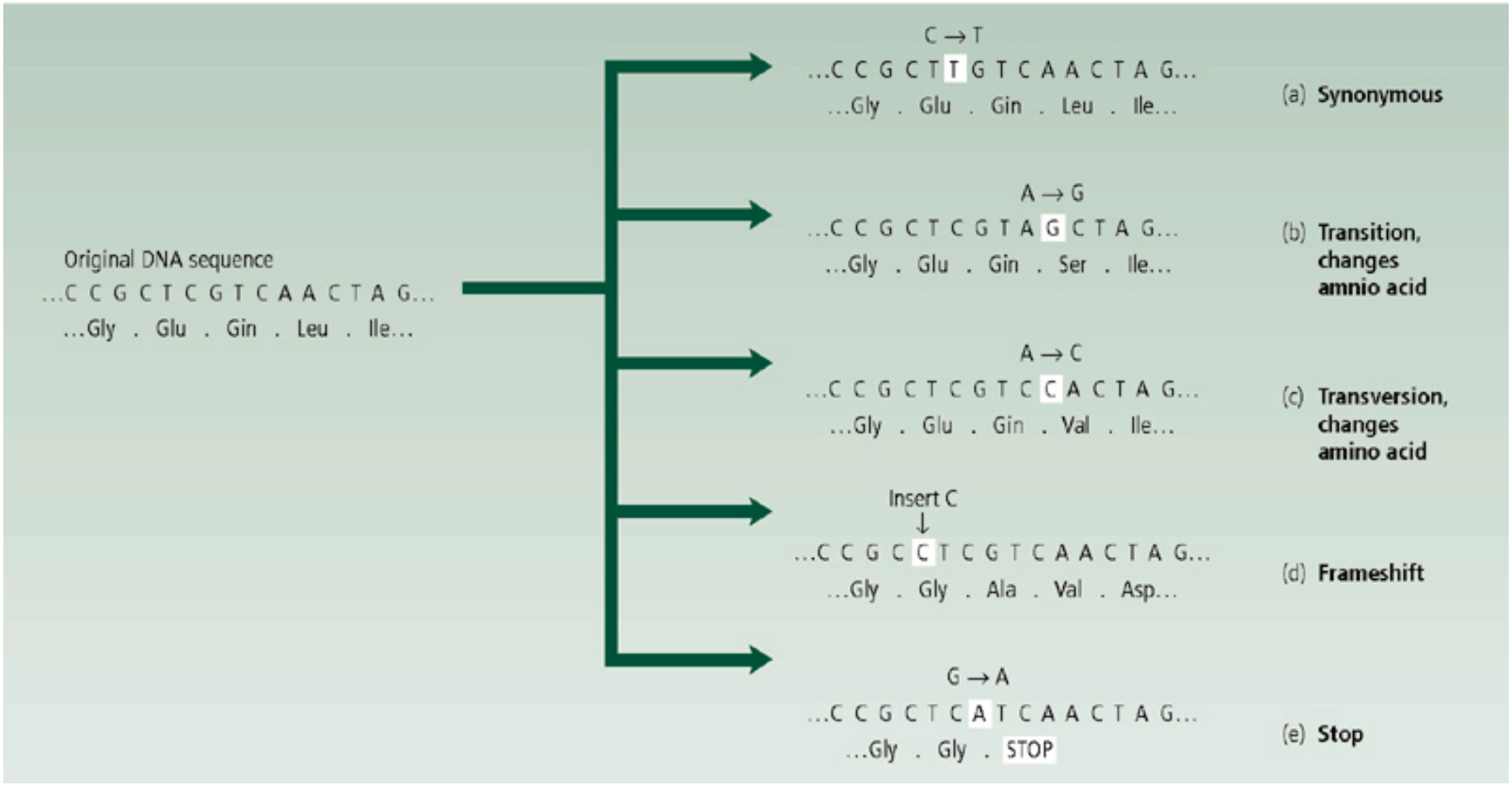
- Chromosomal variation
  - Inversions, chromosomes fusions/fissions
- Protein variation
  - Immunological, allozymes, amino acid variation
- DNA variation
  - RFLP, RAPD, VNTR, Indel, SNP

Where does genetic variation come from?

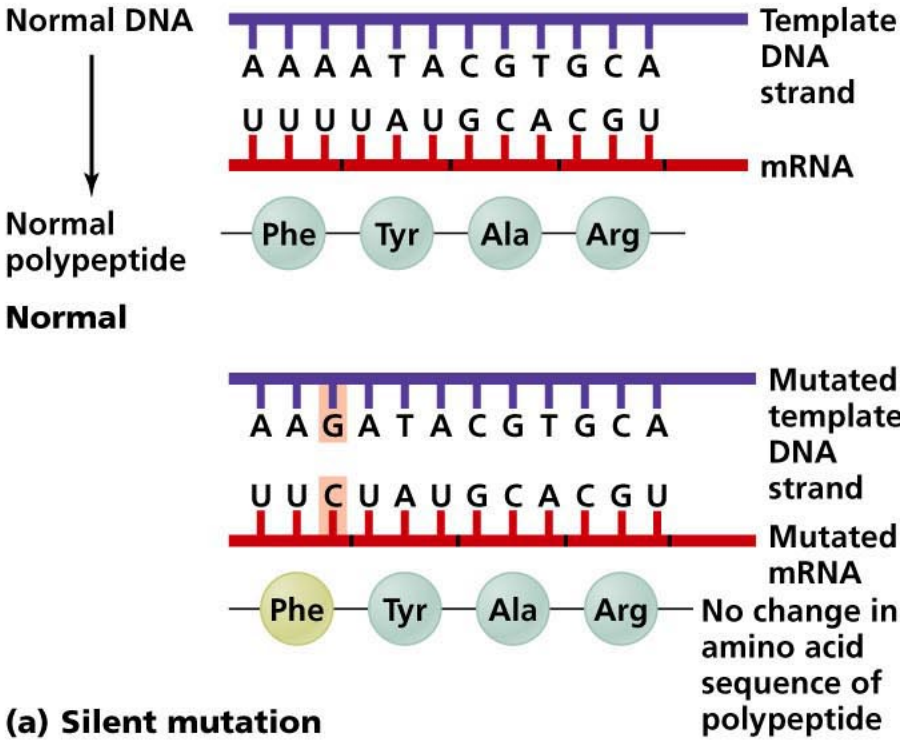
Mutation is the substrate of evolution

- All (genetic) polymorphisms originate with mutation
- Point mutation (one base for another)

Point Mutations



Mutation is the substrate of evolution



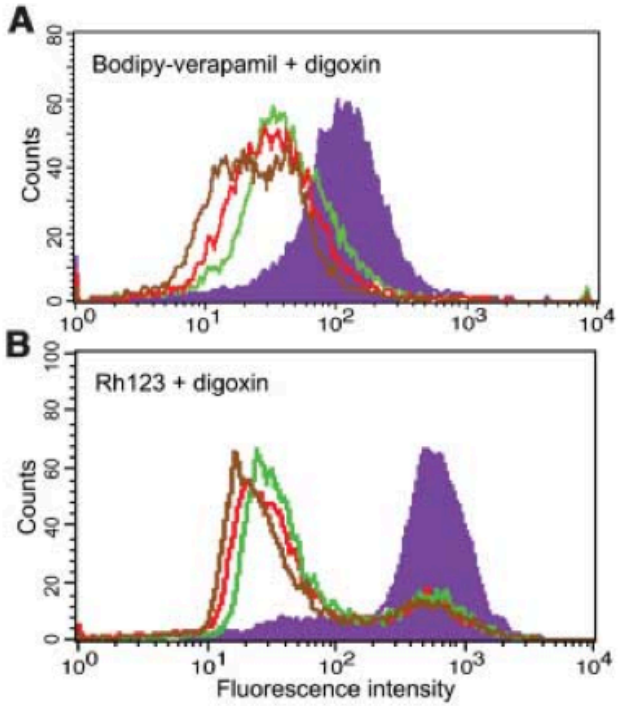
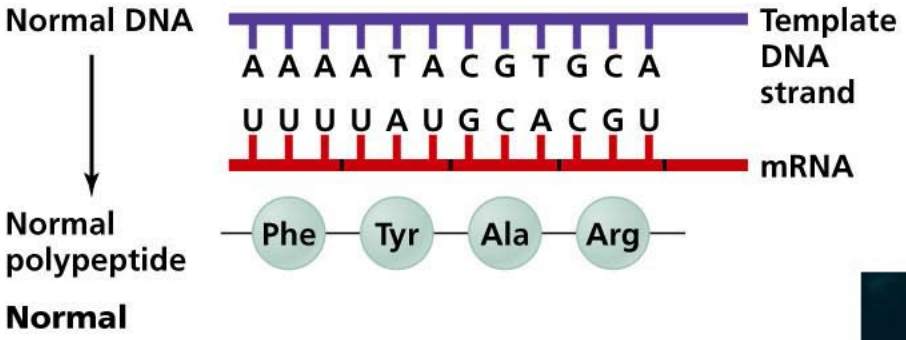
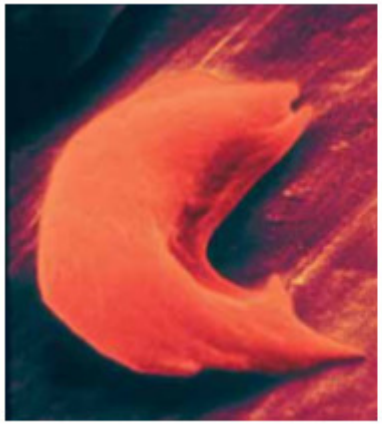


Fig. 4. Drug transport function of wild-type and two MDR1 haplotypes. The drug efflux of vaccinia infected/transfected HeLa cells was determined by FACS analysis (*14*). Cells were transfected with pTM1 (control; purple), MDR1, (wild-type P-gp; green), C1236T-G2677T-C3435T (red), and C1236T-G2677T-C3435A (brown). (**A**) 0.5 μM bodipy-FL-verapamil in the presence of 500 μM digoxin; (**B**) 0.5 μM Rh123 in the presence of 150 μM digoxin.

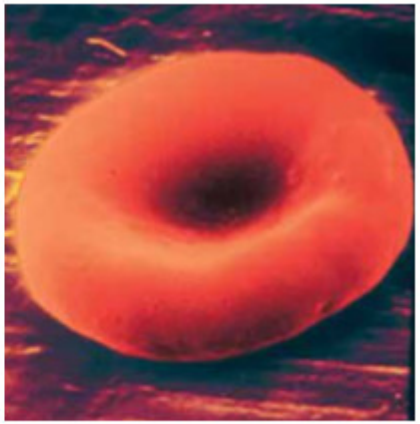
Kimchi-Sarfaty et al. 2007

Mutation is the substrate of evolution

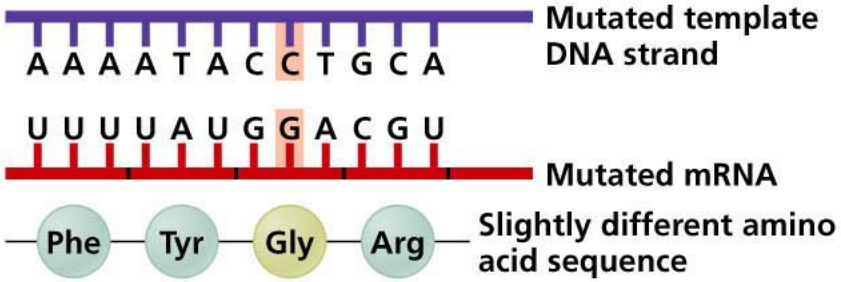




Sickle-cell phenotype

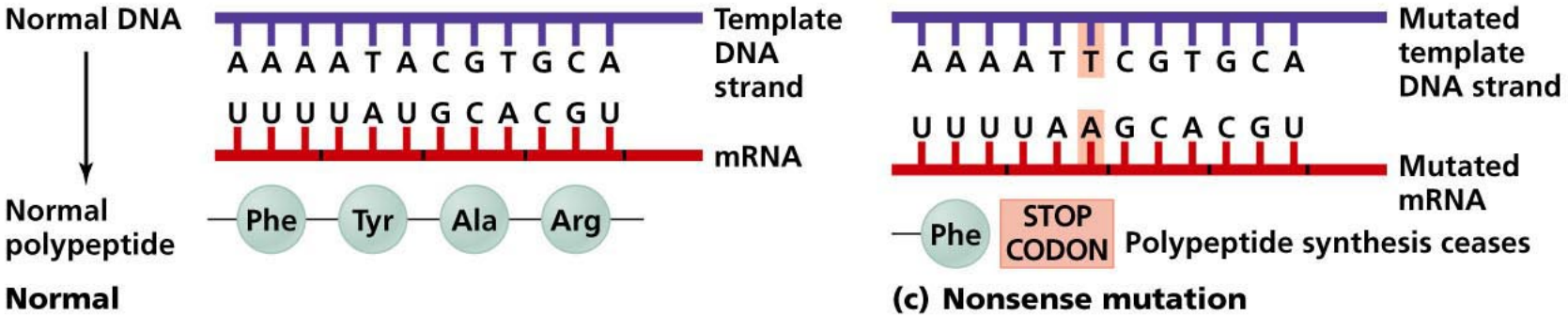


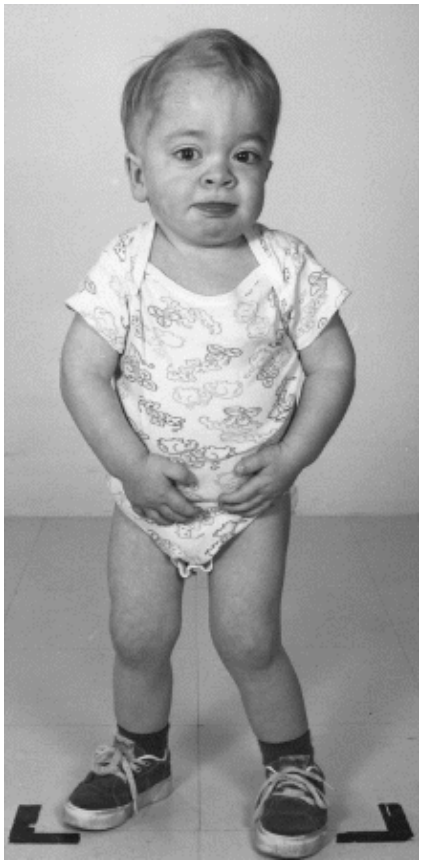
Normal phenotype



(b) Missense mutation

Mutation is the substrate of evolution

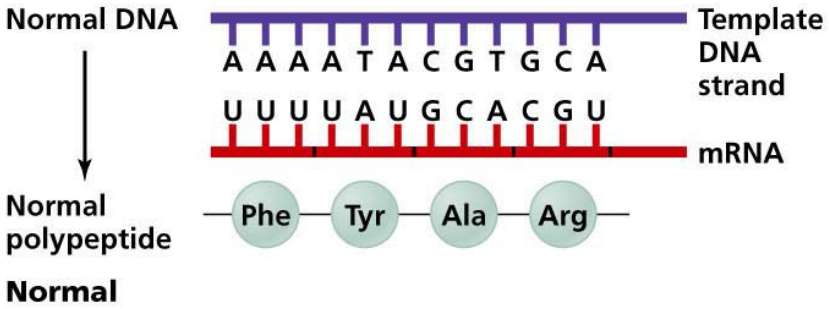


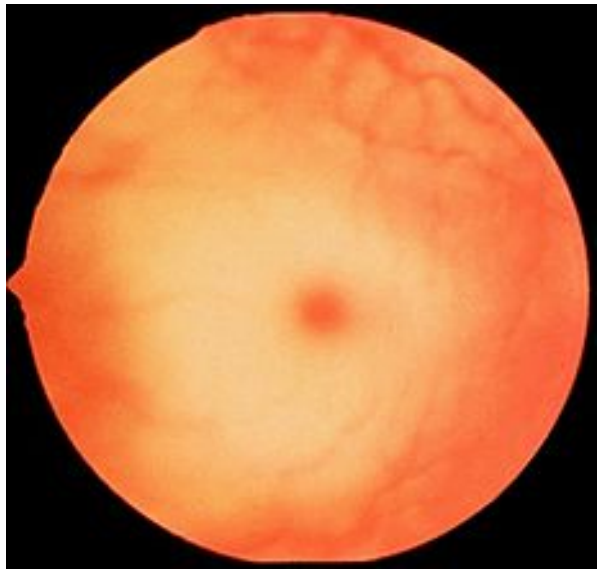


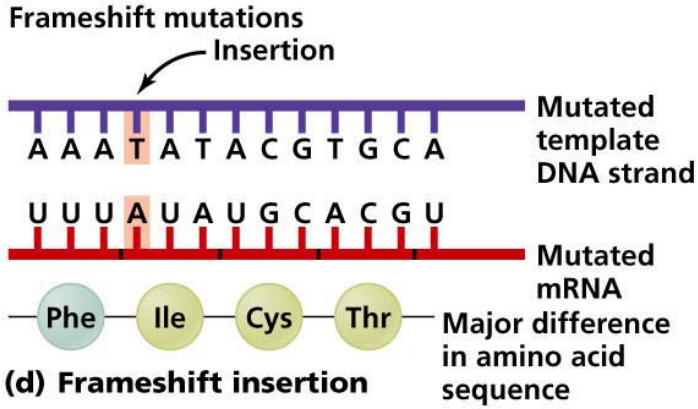
Mutation is the substrate of evolution

- All polymorphisms originate with mutation
- Point mutation (one base for another)
- Insertion (addition of DNA)

Mutation is the substrate of evolution



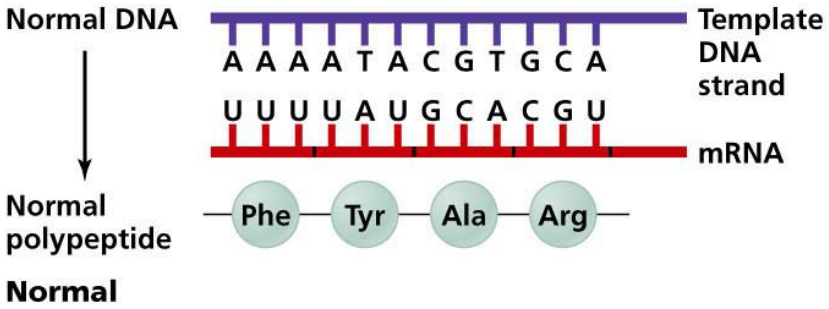


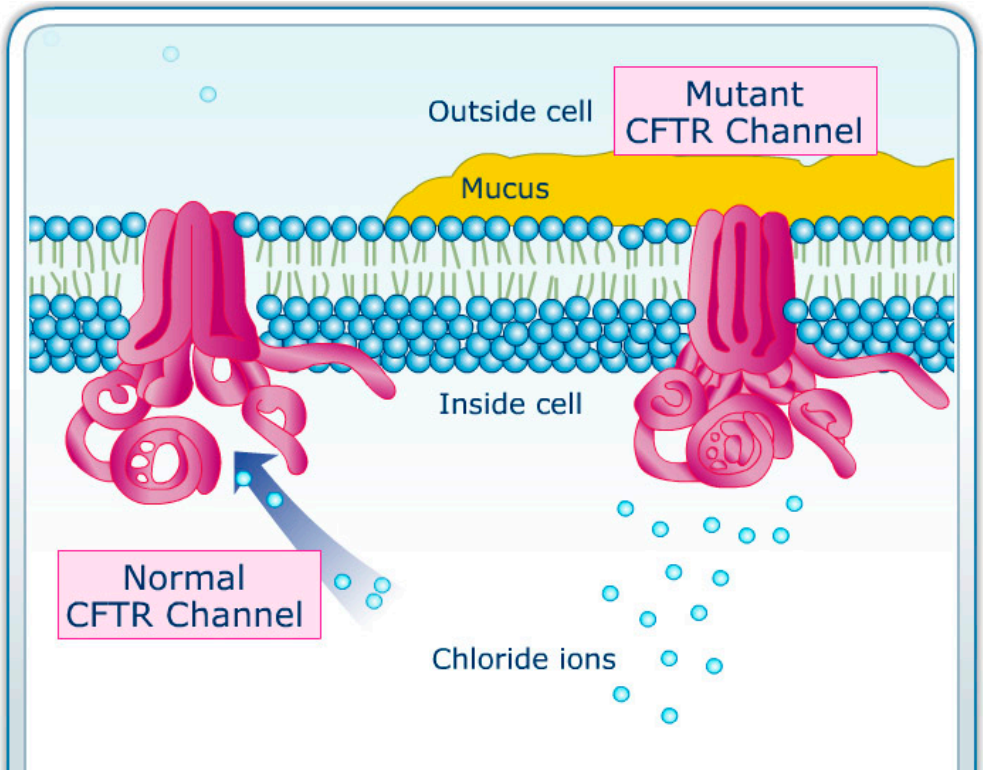


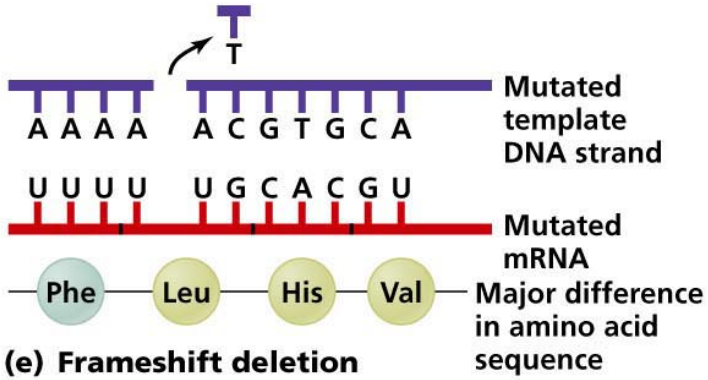
Mutation is the substrate of evolution

- All polymorphisms originate with mutation
- Point mutation (one base for another)
- Insertion (addition of DNA)
- Deletion (loss of DNA)

Mutation is the substrate of evolution



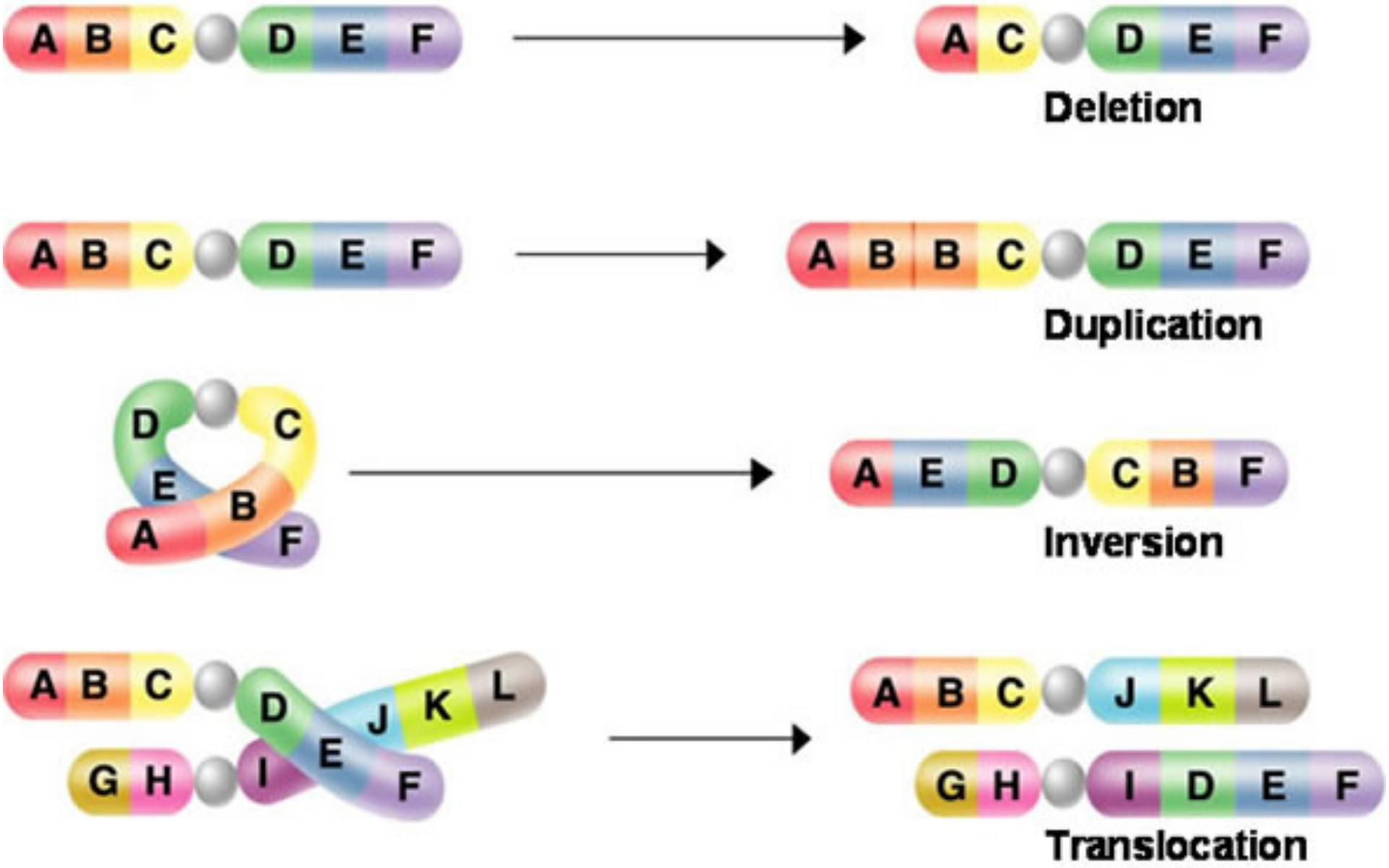




Mutation is the substrate of evolution

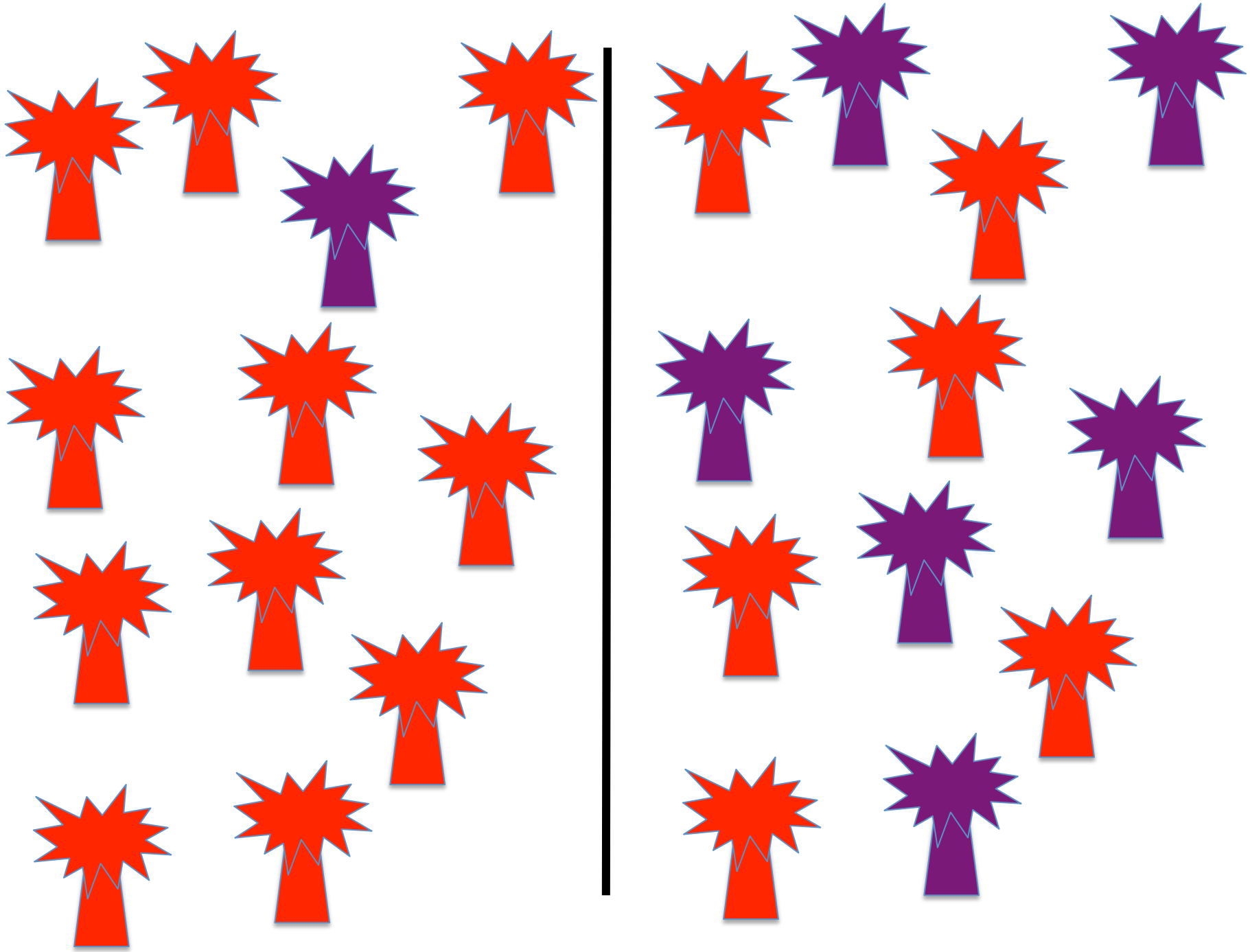
- All polymorphisms originate with mutation
- Point mutation (one base for another)
- Insertion (addition of DNA)
- Deletion (loss of DNA)
- Chromosomal mutations

Mutation is the substrate of evolution



Genetic Variation

- All genetic variation originates with mutation
  - Mutation is the substrate of evolution
- All levels of organization from single base pairs to entire genomes
- Understanding genetic variation has deep implications



Summarizing molecular data

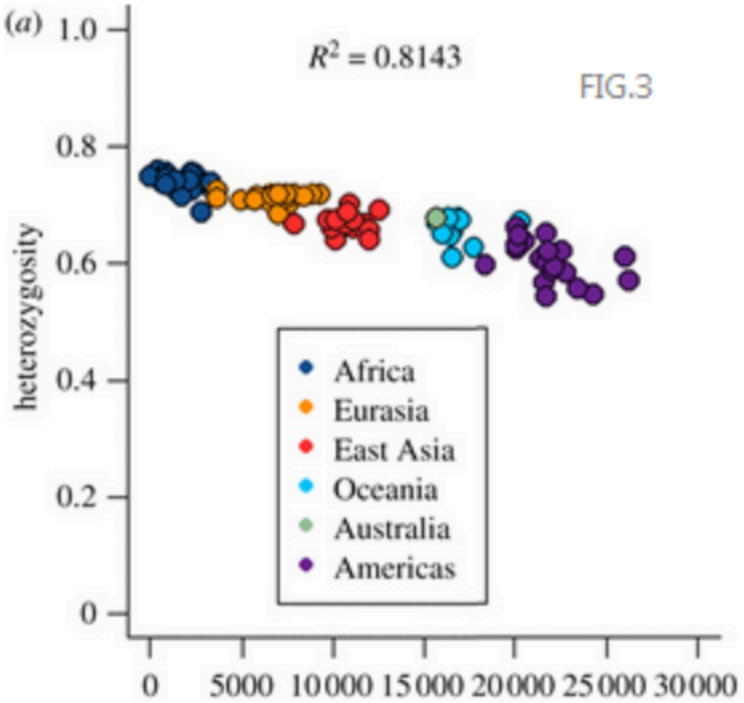
```
#1   CTGCCCACCTTTTGTTTGGGTCTTAGTCCGCAGTGCACTTGTGCCGCCGAGGGGAATGTGGTGCGTTTCCATTGTCCGGATG
#2   ....................C..............................T....................C.......
#3   ..........................................................................................
#4   ..A.................C..............................T....................C.......
#5   ........G...................................A.......................................
#6   ..........................................................................................
#7   ....................C..............................................................C.......
#8   ..............................T.............A.......................................
#9   ..............A...........................................................................
#10  ..............A...........................................................................
#11  ....................C...............C.....................................C.......
#12  ....................C...............C.....................................C.......
```

Key terms

- Gene
- Locus
- Allele
- Genotype
- Haplotype
- Homozygote
- Heterozygote
- Polymorphism

Heterozygosity

- Expected probability that an individual will be heterozygous at a locus
- Expected proportion of heterozygotes in a population at a locus



Microsatellite heterozygosity vs. geographical distance

Hunley et al. 2012

Allelic diversity

- Number of alleles per locus

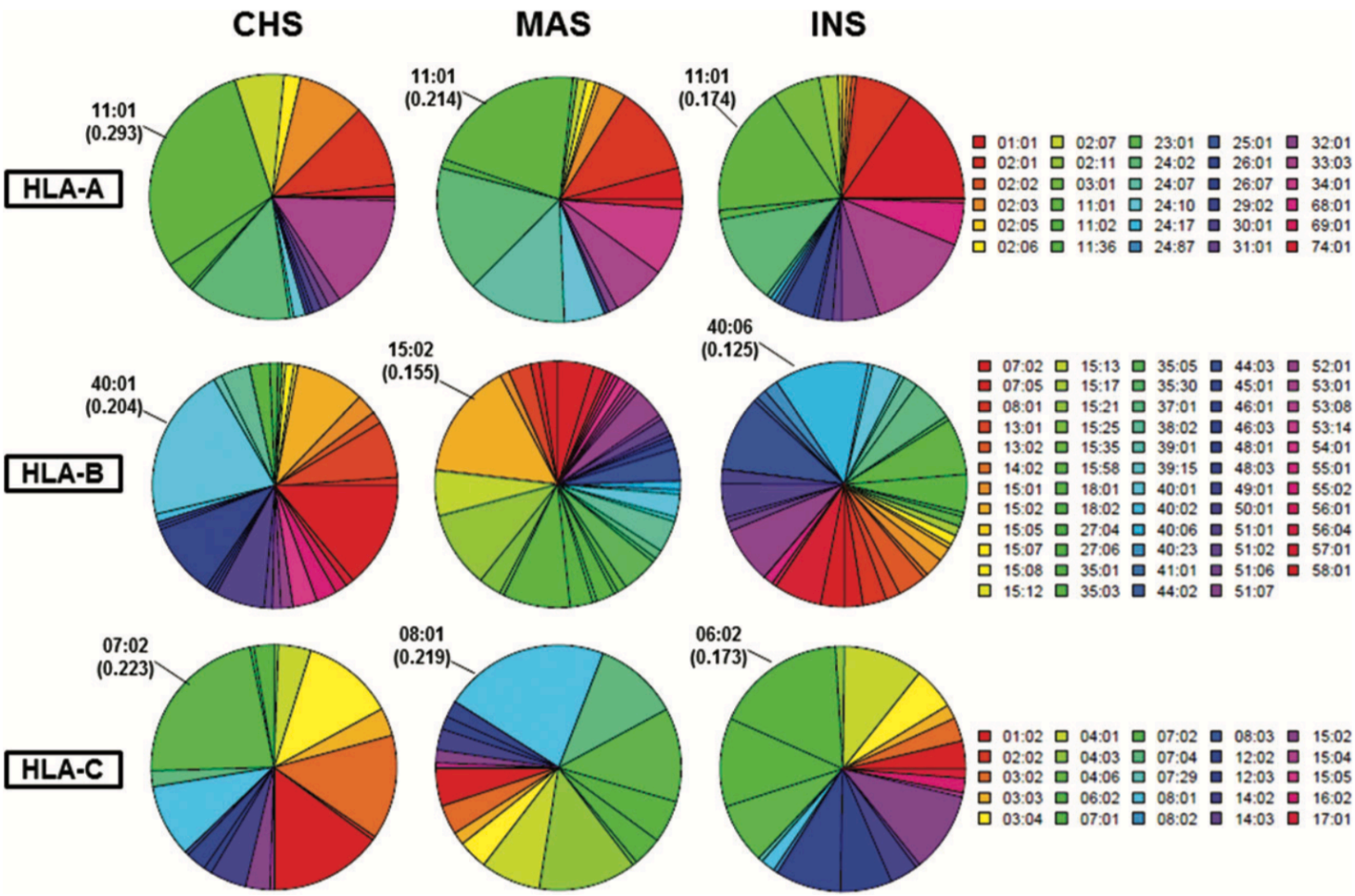
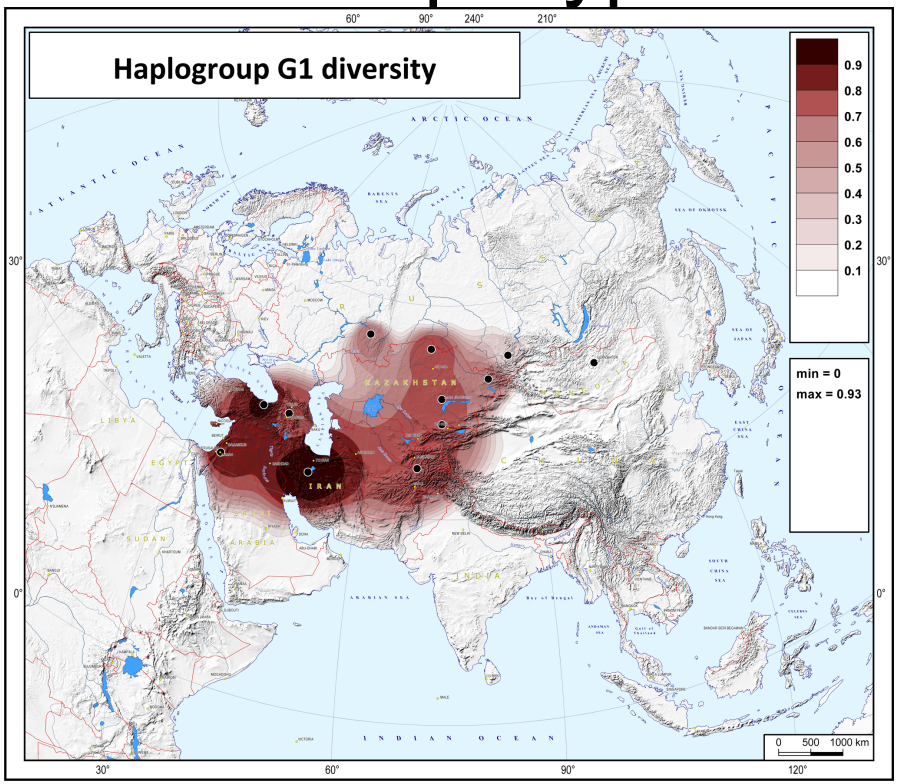


Figure 1. Allelic diversity and distribution in HLA Class I genes. Pie charts illustrating the allelic diversity of the three genes in HLA Class I in the three populations,

Pillai et al. 2014

Haplotype indices

- Haplotype number
- Unique (‘private’) haploytpes
- Haplotype diversity: probability that two randomly chosen haplotypes are different



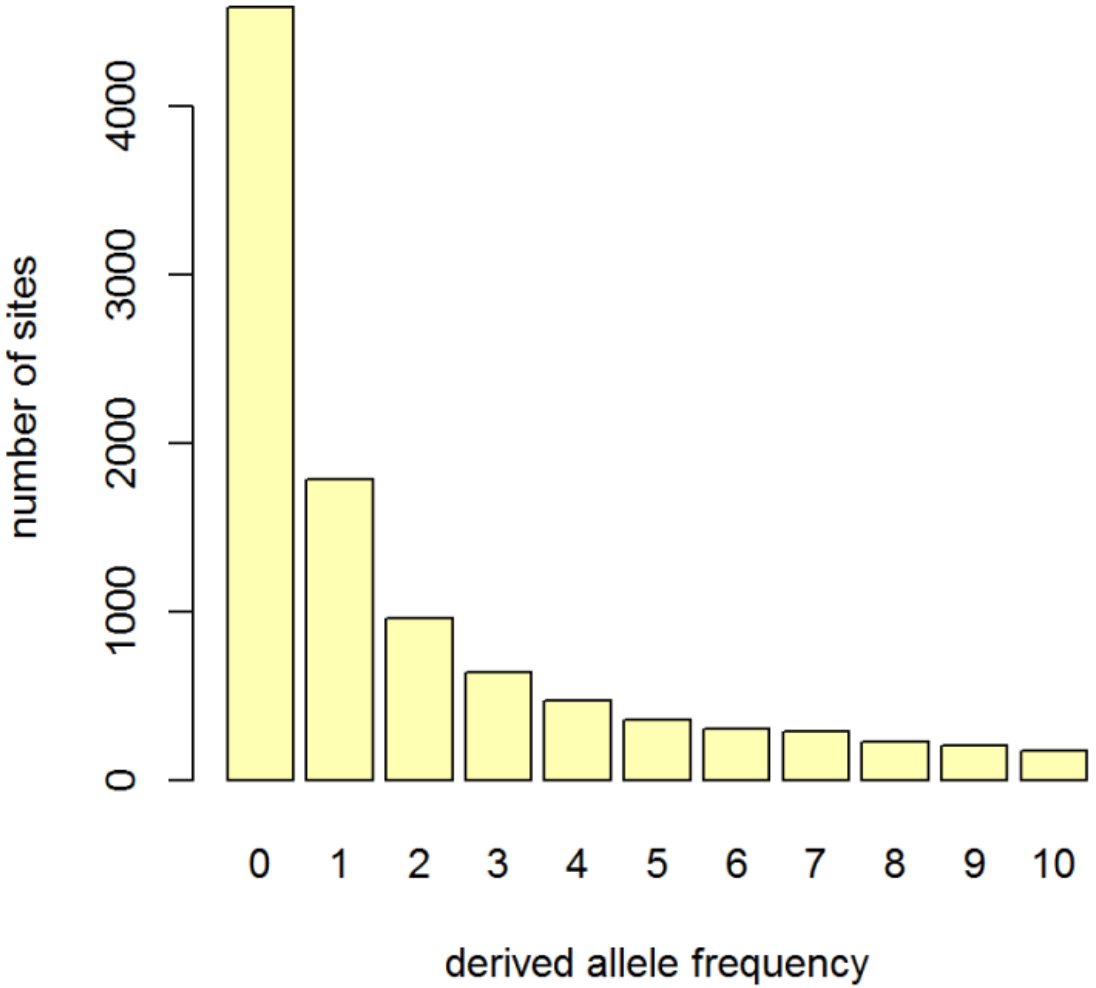
Balanovksy et al. 2015

Nucleotide-level Indices

- Θ_w
  - Watterson's estimator
  - $S/\sum i=1\uparrow n-1 \blacksquare 1/i$
- π
  - Average pairwise difference between alleles

Site Frequency Spectrum

5 diploid individuals
10,000



Genetic Variation: Software

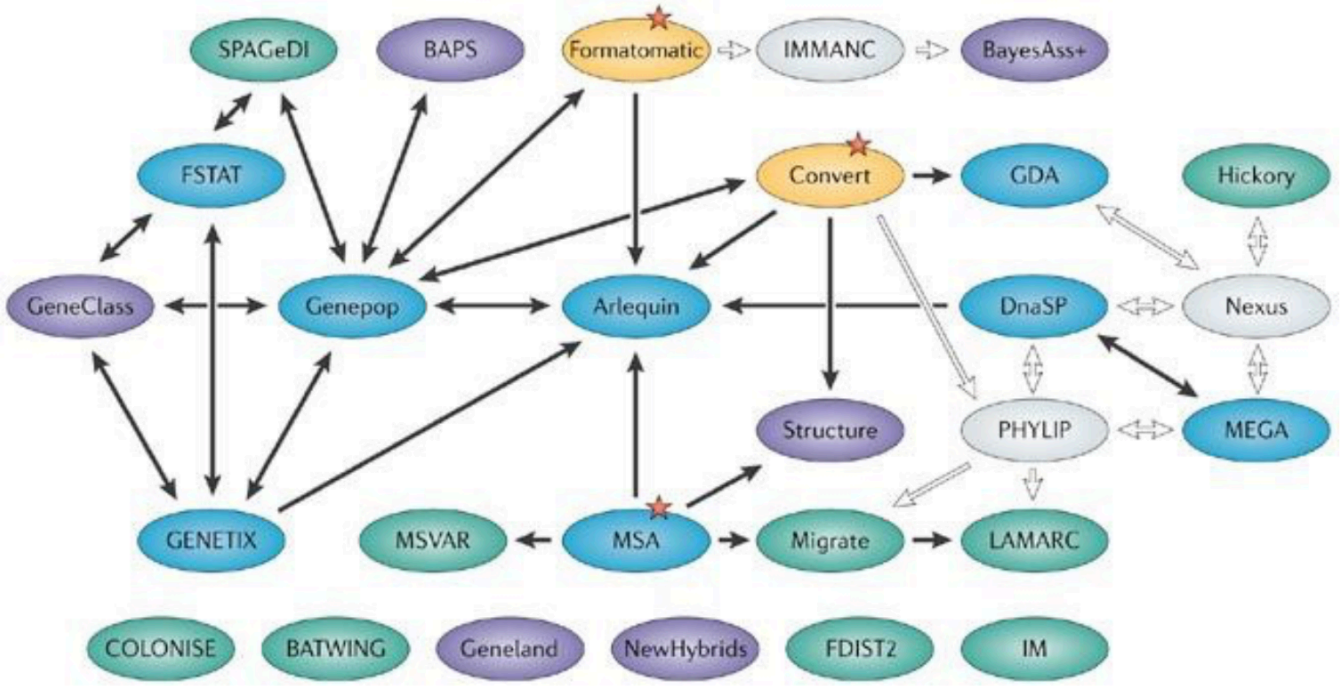
- DnaSP (DNA sequence polymorphism)
- DNA sequence
- Basic population genetic analyses
- Estimate genetic diversity
- Within/Between populations
- LD, recombination, gene flow, gene conversion
- Neutrality tests

Genetic Variation: Software

- Arlequin
- MEGA
- PHYLIP
- …
- Excoffier and Heckel 2006: Computer programs for population genetic data analysis: a survival guide

Genetic Variation: Software

From: Computer programs for population genetics data analysis: a survival guide



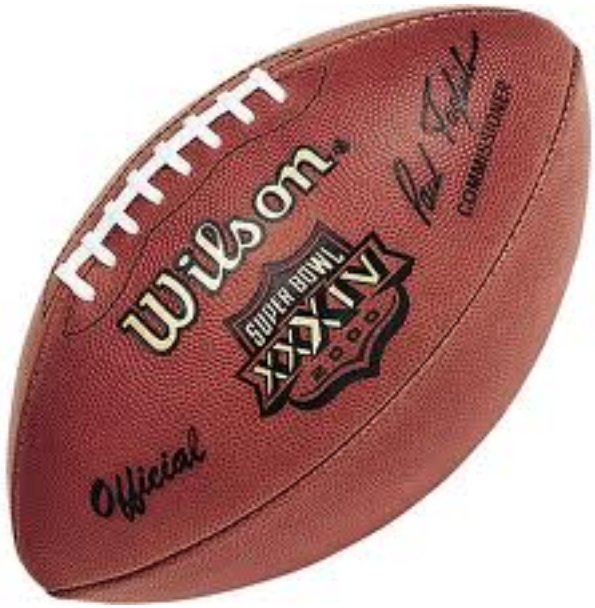
Copyright © 2006 Nature Publishing Group
Nature Reviews | Genetics

Why do we care about genetic variation?

- Genetic variation underlies phenotypic differences among individuals
  - Including disease risk and responses to drugs and environmental factors

Why do we care about genetic variation?

- Genetic variation underlies phenotypic differences among individuals
  - Including disease risk and responses to drugs and environmental factors
- Individual identification



Why do we care about genetic variation?

- Genetic variation underlies phenotypic differences among individuals
  - Including disease risk and responses to drugs and environmental factors
- Individual identification
- Manage resources

Why do we care about genetic variation?

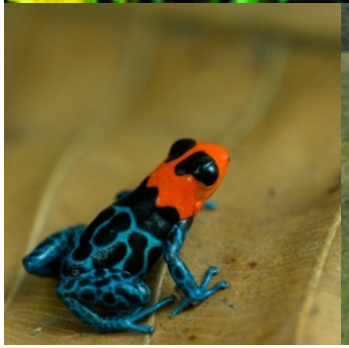
- Genetic variation underlies phenotypic differences among individuals
  - Including disease risk and responses to drugs and environmental factors
- Individual identification
- Manage resources
- Public Health

Why do we care about genetic variation?

- Genetic variation underlies phenotypic differences among individuals
  - Including disease risk and responses to drugs and environmental factors
- Individual identification
- Manage resources
- Public Health
- Improve plant and animal food products
- Understand genetic basis of disease and other complex phenotypes including behaviours
- Insights into evolutionary history, ancestry

Conservation genetics

- Use of genetic theories and techniques to study the loss and restoration of genetic diversity
- Derived from ecological, evolutionary, and quantitative genetics
- Focuses on small populations



Human threats

- Poaching

Human threats

- Poaching
- Habitat loss, degradation, fragmentation

Human threats

- Poaching
- Habitat loss, degradation, fragmentation

Human threats

- Poaching
- Habitat loss, degradation, fragmentation

Human threats

- Poaching
- Habitat loss, degradation, fragmentation
- Invasion

Human threats

- Poaching
- Habitat loss, degradation, fragmentation
- Invasion
- Pollution

Human threats

- Poaching
- Habitat loss, degradation, fragmentation
- Invasion
- Pollution
- Climate change

Why is conservation important?

- Sixth mass extinction
- Aesthetic value
- Bioresources
- Ecosystem services
- Ethical, moral considerations

How can genetics help?

- Minimize inbreeding and loss of genetic variation
- Identify populations of concern
- Resolve population structure
- Resolve taxonomic uncertainty
- Define management units
- Detect hybridization
- Detect and define invasive species
- Estimate population size and sex ratio
- Establish parentage
- Understand population connectivity
- Aid in management
- Enhance reproductive capacity of organisms
- How do genetic factors affect extinction risk?
- How can we best genetically rescue populations?
- What can we use genetics to identify hotspots?
- Applied genetic detective work
- …

Amato et al. 2009

Conservation genetics

- Management strategies
  - Captive Breeding design

Captive breeding

- Establish populations in secure *ex situ* locations
- Educate/engage public
  - Fundraising
- Provide educational opportunities
- Provide animals for reintroduction
- Maintain demographically stable, genetically variable populations with little inbreeding
- 1150 zoos and aquaria worldwide house 1.2 million animals
  - 5-10% used for endangered species

Captive breeding

Table 1. Types of contributions that molecular markers can make to captive breeding programs

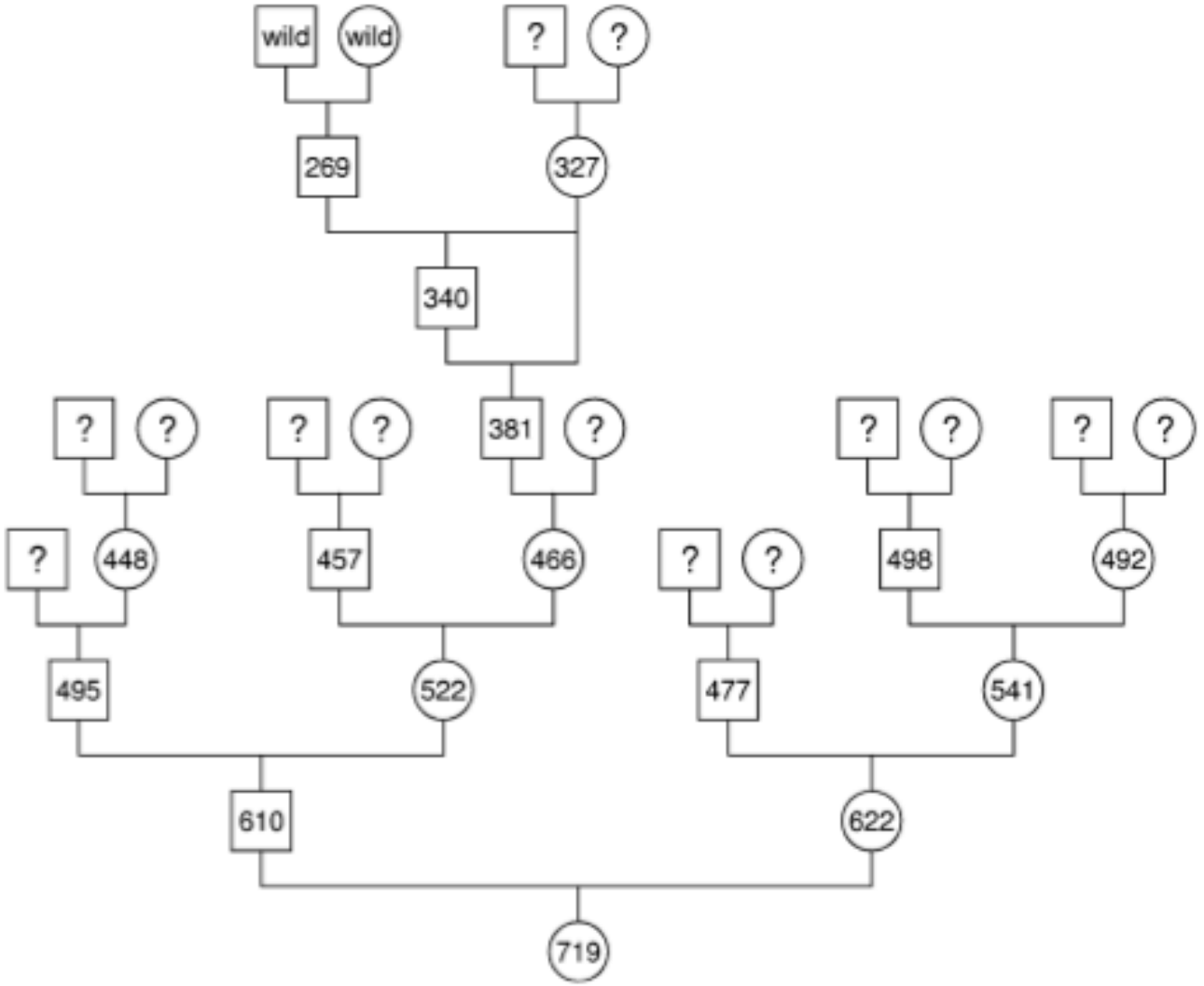
Contribution to breeding program	Species	Citation
Assessment of founder relationships	California condor (*Gymnogyps californianus*)	Geyer et al. (1993)
	Guam rail (*Rallus owstoni*)	Haig et al. (1994)
	Micronesian kingfishers (*Halcyon cinnamomina*)	Haig et al. (1995)
	Bearded vulture (*Gypaetus barbatus*)	Gautschi et al. (2003)
Pedigree reconstruction	Lion-tailed macaques (*Macaca silenus*)	Morin and Ryder (1991)
	Przewalski's horse (*Equus ferus przewalskii*)	Bowling et al. (2003)
Subspecies identification	Lion (*Panthera leo*)	O'Brien et al. (1987)
	Chimpanzee (*Pan troglodytes*)	Ely et al. (2005)
Identification of geographic origin	Galapagos tortoise (*Geochelone nigra*)	Russello et al. (2007)
Quantification of wild genetic diversity captured	Baird's tapir (*Tapirus bairdii*)	Norton and Ashley (2004)
	Iberian wolf (*Canis lupus signatus*)	Ramirez et al. (2006)
Identification of genetically valuable individuals	Whooping crane (*Grus americana*)	Jones et al. (2002)
	St Vincent parrot (*Amazona guildingii*)	Russello and Amato (2004)
Assessment of hybridization	Lesser white-fronted goose (*Anser erythropus*)	Ruokonen et al. (2007)

Ivy et al. 2009

Parma wallaby

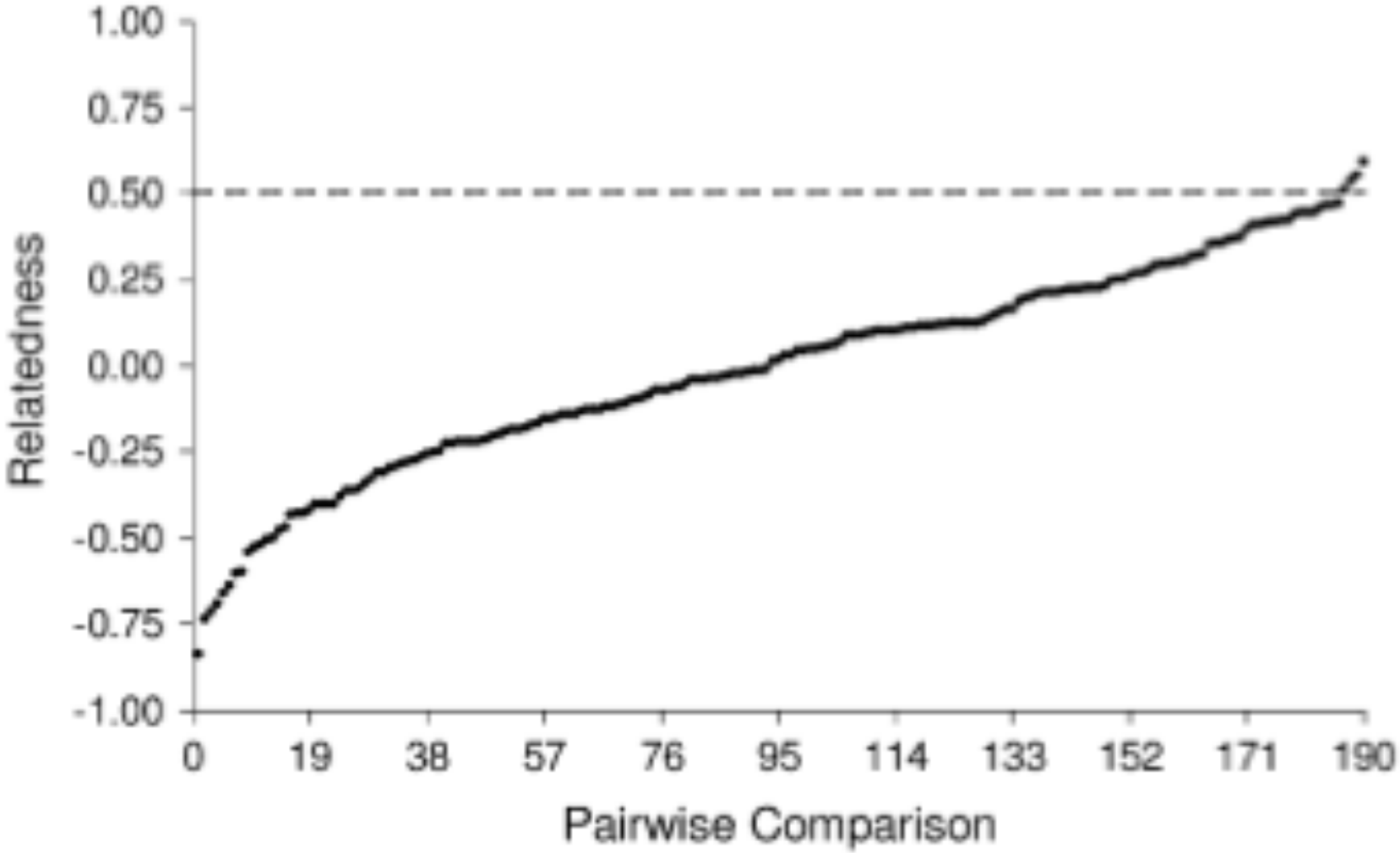
- Marsupial native to Australia
- Threatened by habitat degradation and low census size
- Captive population founded in 1960s
  - Managed by pedigree
  - Breeding recommendations based on kinship
  - 50% of individuals have unknown ancestry

Parma wallaby



Ivy et al. 2009

Parma wallaby



Ivy et al. 2009

Hereditary defects

TABLE 1. Possible, presumed, or proven hereditary defects observed in wild animals bred in captivity

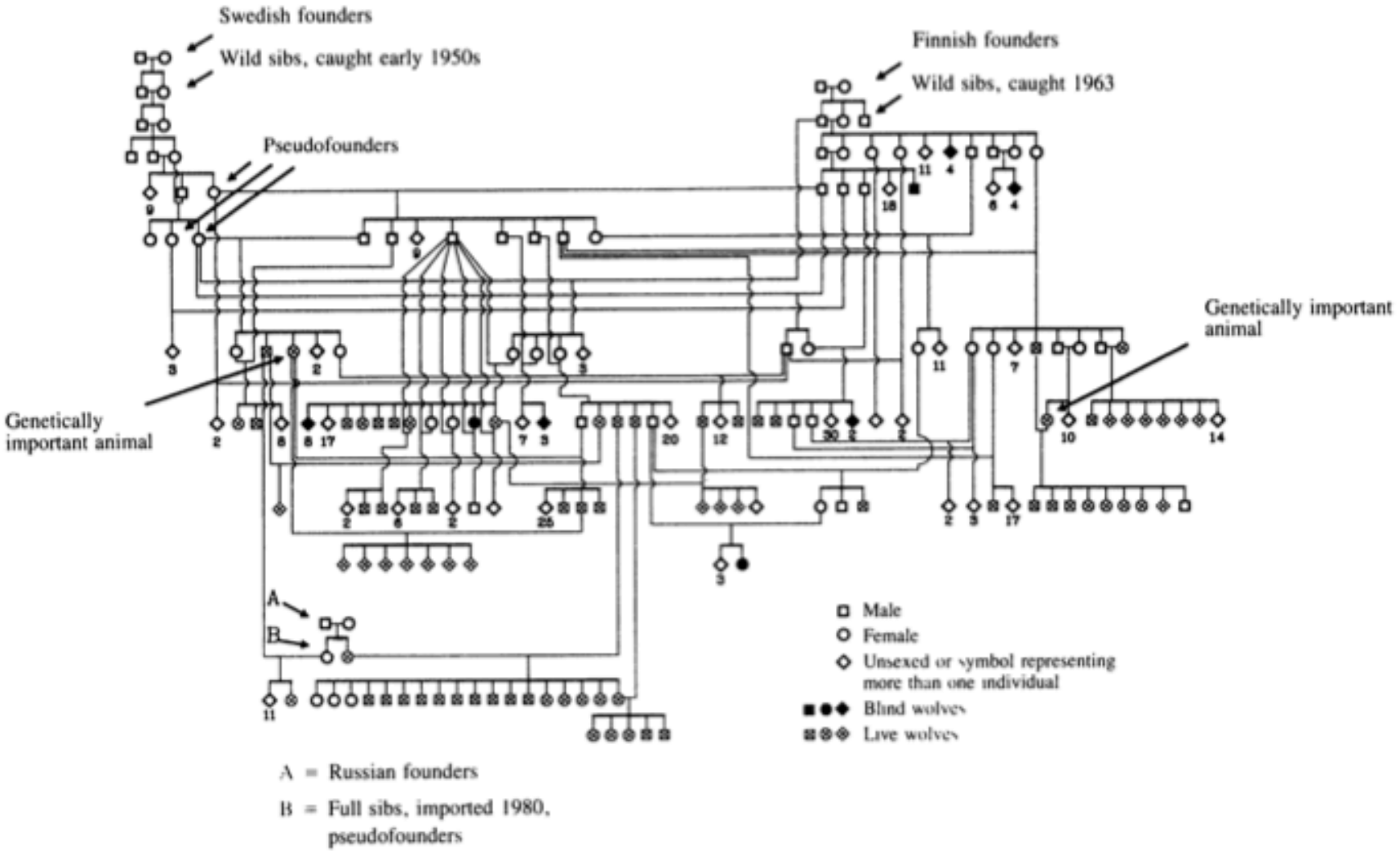
Species/taxonomic group	Defect	Heredity/mode of inheritance	Source
Arctic fox *Alopex lagopus*	Genu flexum (abnormal hind leg)	Hereditary	Krukowski & Klos, 1982
	Gingival hyperplasia	Hereditary/probably autosomal recessive with incomplete penetrance	Dyrendahl & Henricson, 1959
Asian small-clawed otter (*Aonyx cinerea*)	Urolithiasis	Possibly hereditary	Karesh, 1983; Calle, 1985
Axis deer *Axis axis*	Ventricular septal defect	Possibly hereditary	Gendron, 1975
Bison Wood bison *Bison bison athabascae* Plains bison *Bison bison bison*	Arthrogryposis and degenerative joint disease	Possibly hereditary	Schuh & Haigh, 1990
Black-and-white ruffed lemur (*Varecia variegata variegata*)	Pectus excavatum Diaphragm hernia Cleft palate Scoliosis Exencephaly Hydrocephaly	Autosomal recessive Presumed hereditary	Benirschke et al., 1985
Brown bear *Ursus arctos*	Albinism	Single autosomal recessive	Laikre et al., 1996; Andrén et al., 1997; this paper
Carnivores	Various dental diseases	Possibly hereditary	Fagan, 1980
Dear mouse *Peromyscus maniculatus*	Spherocytosis Post juvenile nude, hair deficiency	Autosomal recessive Autosomal recessive	Huestis et al., 1956 Clark, 1934
Dingo *Canis dingo*	Hip joint dysplasia	Presumed hereditary	Christoph et al., 1969
Douc langur *Pygathrix nemaeus*	Pectus excavatum	Presumed hereditary	Sedgwick, 1981

Laikre 1999

Grey wolf

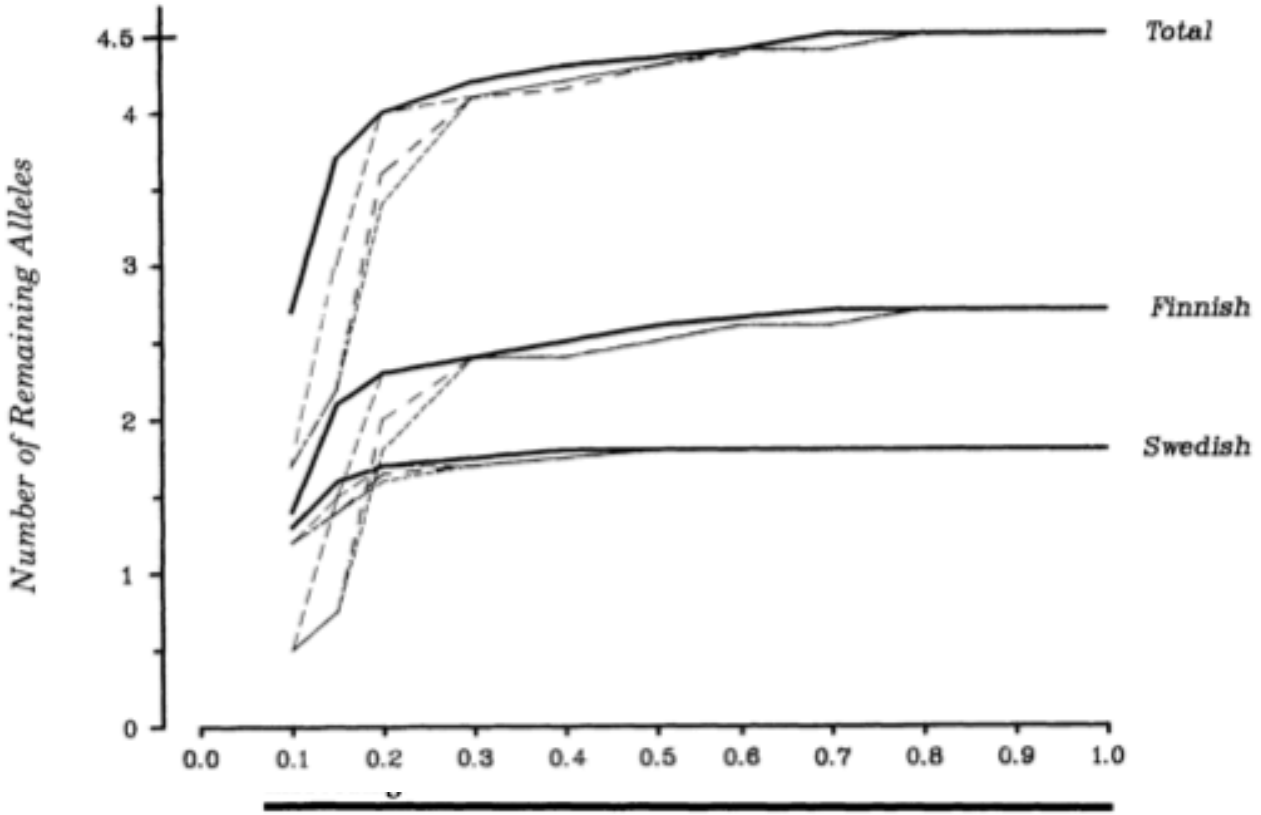
- Largest member of Canidae
- Ancestor of domestic dog
- Once abundant over Eurasia, North Africa, North America
- Current range restricted
- Threatened by habitat loss

Grey wolf



Laikre et al. 199

Grey wolf



Minimum Carrier Probability for Removal	*Number of Animals Remaining*	F_T	F_F	F_S
1.0 (no removals)	55	0.34	0.17	0.17
0.6	49	0.35	0.17	0.18
0.5	43	0.36	0.18	0.18
0.4	40	0.36	0.18	0.18
0.3	36	0.37	0.19	0.18
0.2	29	0.38	0.22	0.16
0.1	2	0.34	0.16	0.18

Laikre et al. 1993

Black footed ferrets

Black footed ferrets

	n	H_O ($\pm$ *SD*)	H_{dem}	A	N_c	$N_{e\text{-dem}}$	$N_{e\text{-gen}}$ (95% *CI*[a])
Agate–Sage Creek	207	0.39 ± 0.12	0.43	2.0	81	75	10.6 (2.2–18.9)
Heck Table	47	0.39 ± 0.16	0.41	2.0	17	14	2.0 (0.0–3.6)
Combined population, 5 loci[b]	254	0.44 ± 0.04	0.43	2.0			

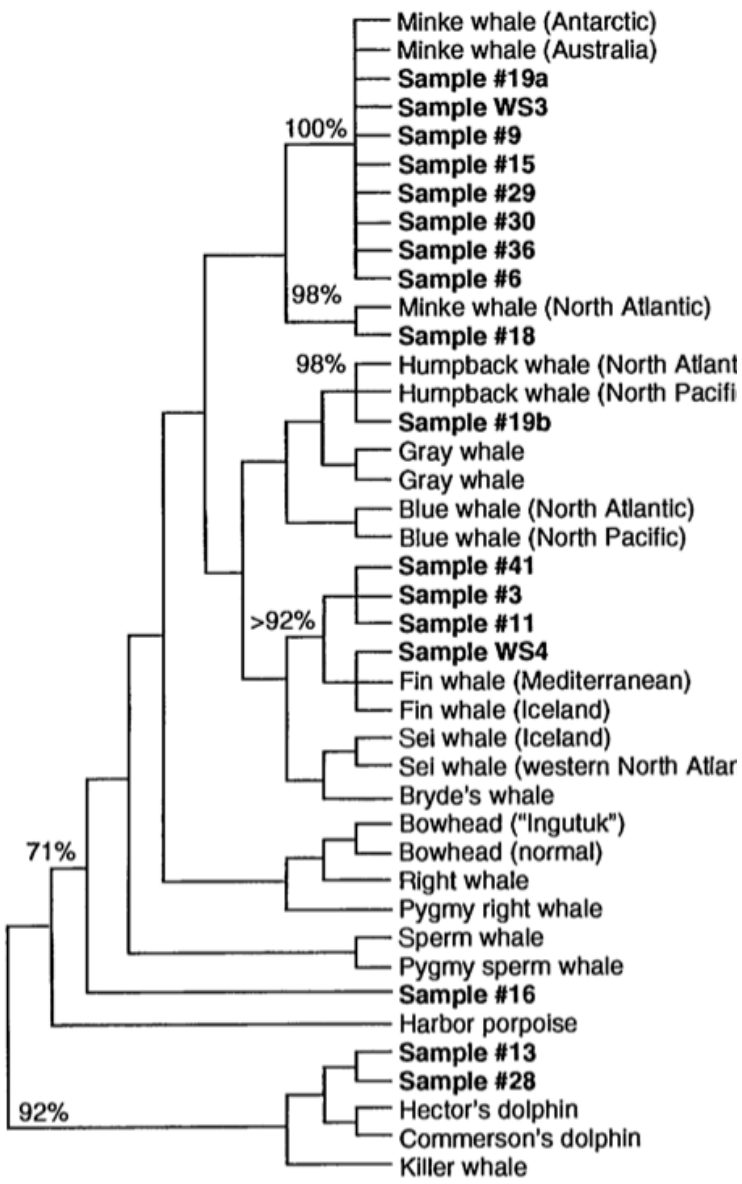
Cain et al. 2011

Conservation genetics

- Management strategies
  - Captive Breeding design
  - Forensics

Whaling

Whaling

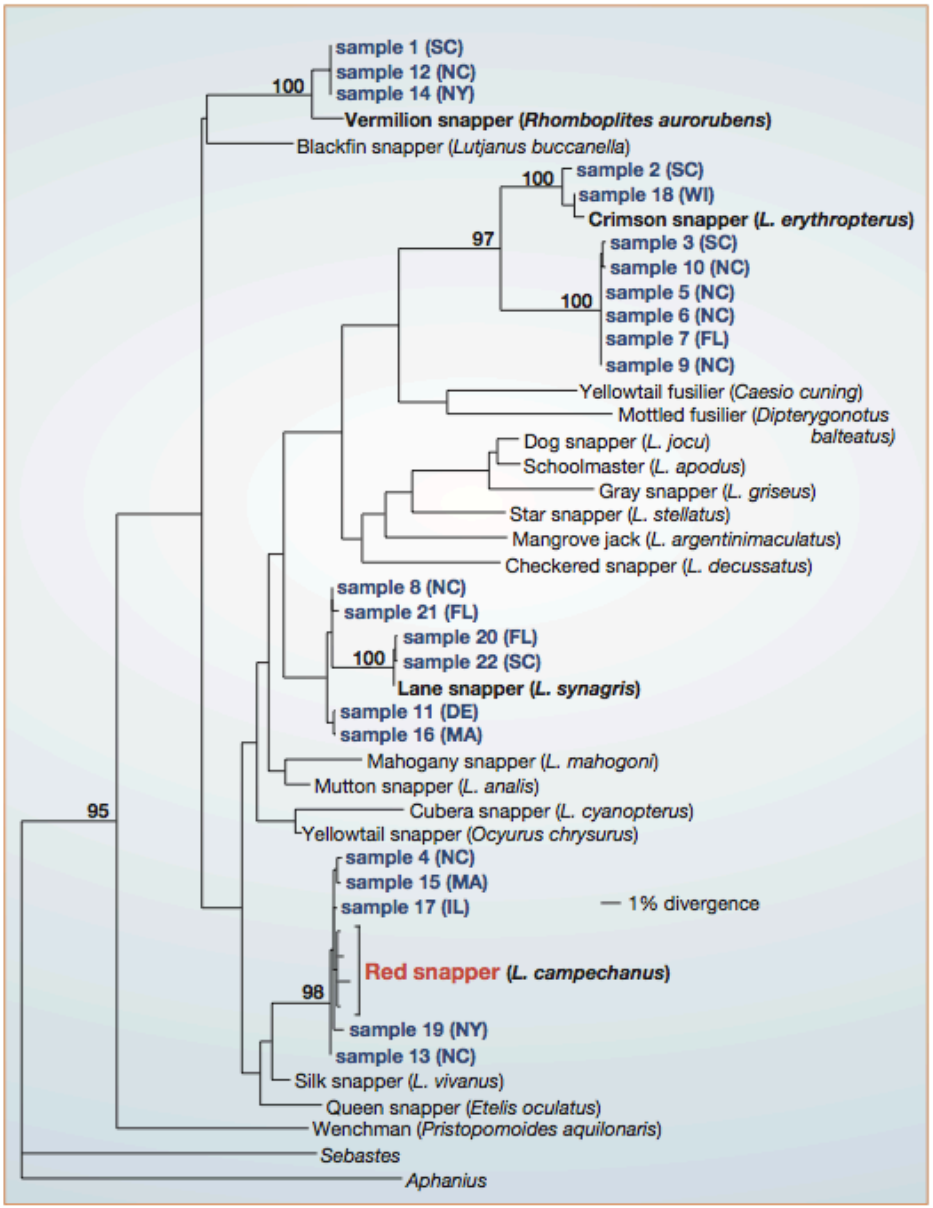


Baker & Palumbi 1993

Red Snapper

- Commercially and recreationally-caught
- Most commonly caught snapper in US
- Restrictions to maintain populations
  - 51% commercial
  - 49% recreational

Red Snapper



Marko et al. 2004

Pacific Red Snapper

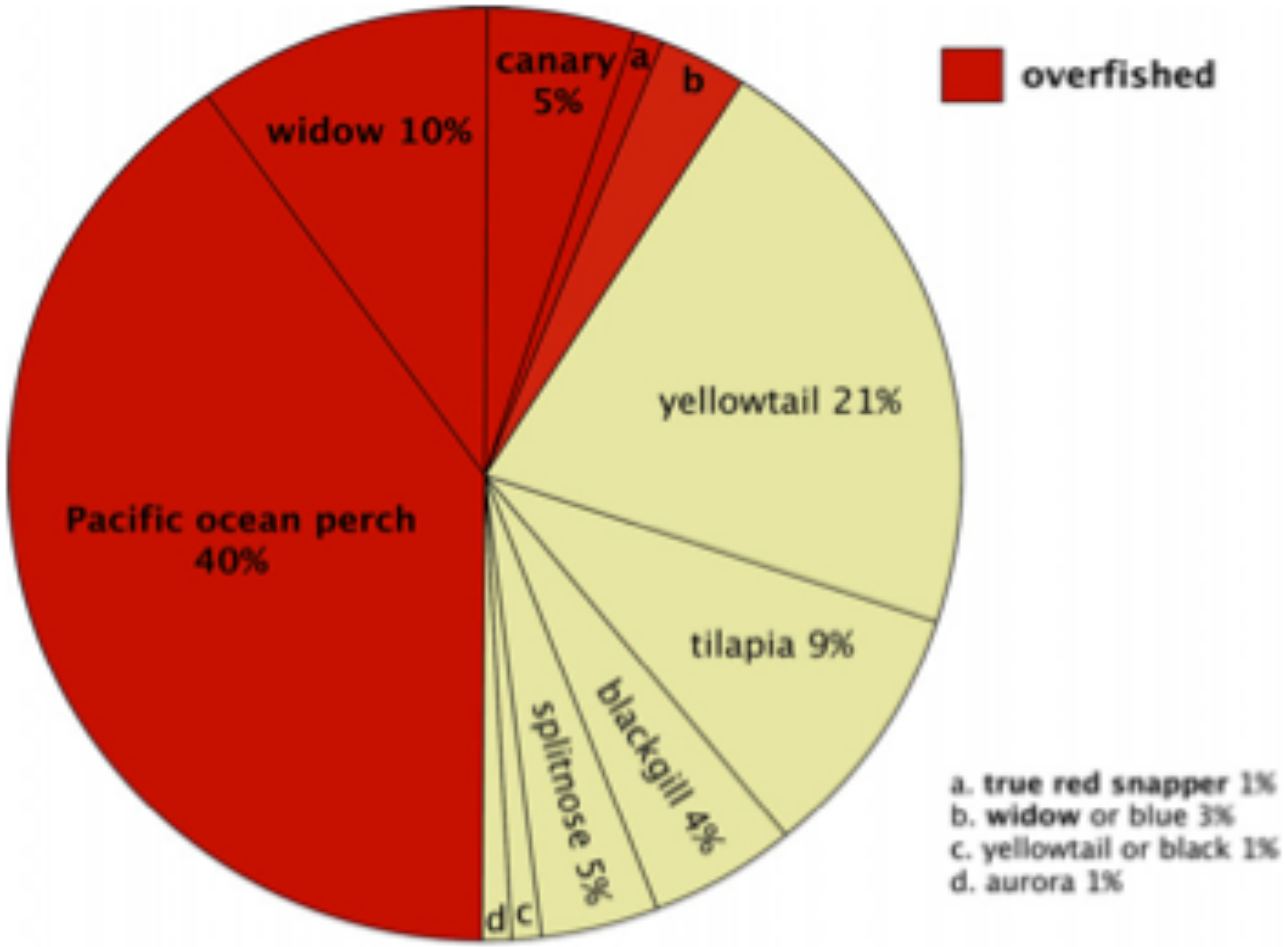


Fig. 2 – Percent of Pacific red snapper samples (*n* = 77) that have been declared overfished within the past three years by the National Marine Fisheries Service.

Logan et al. 2008

Conservation genetics

- Management strategies
  - Captive Breeding design
  - Forensics

Future of conservation genetics

- Open questions:
  - Genetic variation and fitness
  - Mechanism connecting genetic variation and fitness
  - Genotype by environment interaction
  - Intermediate phenotypes
- New technologies
  - DNA sequencing