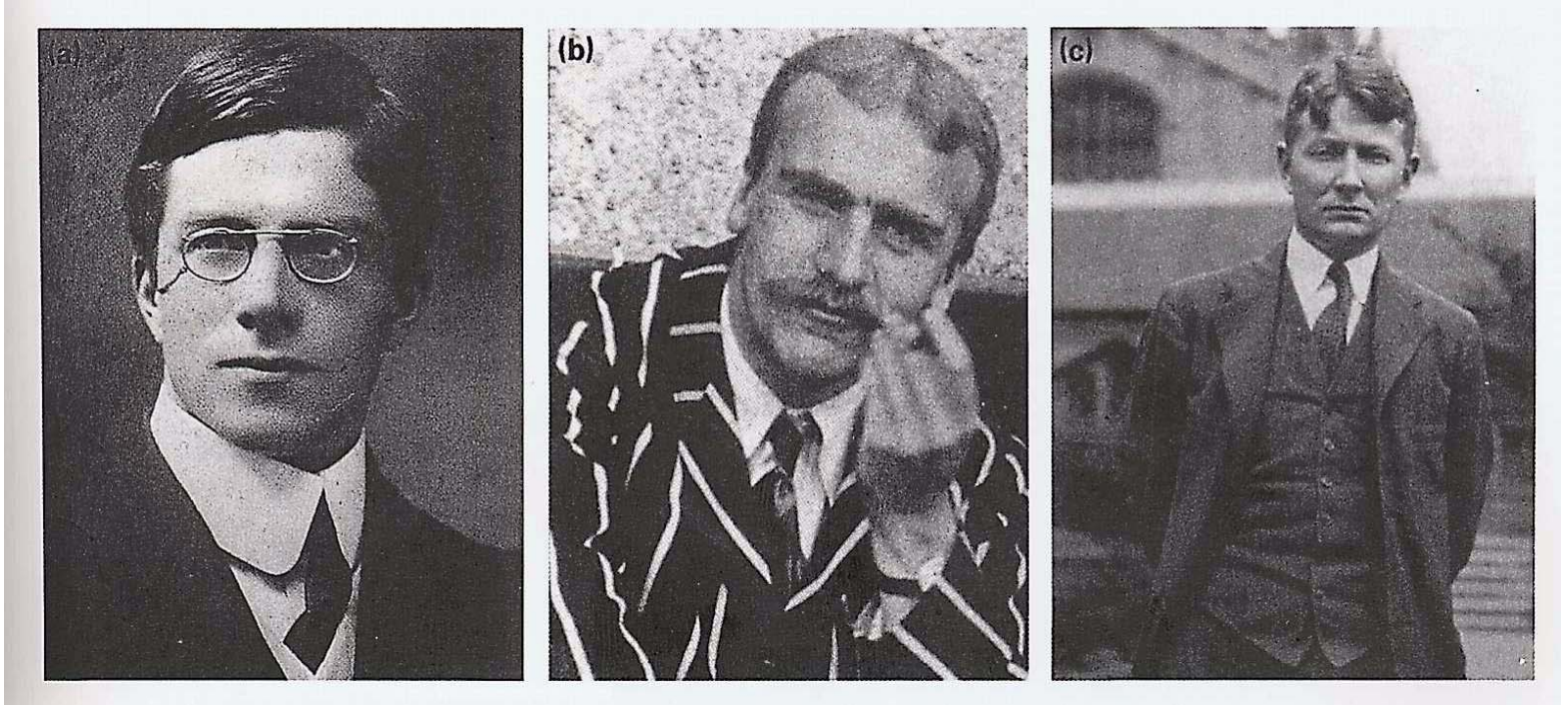
POPULATION GENETICS

POPULATION GENETICS studies the genetic composition of populations and how it changes with time.

It includes the study of forces that induce evolution (the change of the genetic constitution) of populations through time: genetic drift, mutation, gene flow and selection.

R. A. Fisher J. B. S. Haldane S. Wright



“...they had said everything of truly fundamental importance about the theory of genetic change in populations and it is due mainly to man’s infinite capacity to make more and more out of less and less, that the rest of us are not currently among the unemployed.”

R. C. Lewontin (1963)

IMPORTANT CONCEPTS IN POPULATION GENETICS

Genome
Gene
Locus
Allele
Genotype
Homozygote
Heterozygote
Haploid
Diploid
Polyploid

Autosomes
X and Y chromosomes
mtDNA
cpDNA

Linkage disequilibrium
Independency
Recombination
Haplotype

Gene family
Pseudogenes
Exons
Introns
Promotors
Flanking regions

IMPORTANT CONCEPTS IN STATISTICS

Deterministic and stochastic models

Parameters

Monte Carlo Simulations

Sample

Estimates

Randomness

Mean

Variance

Standard deviation

Arithmetic mean

Harmonic mean

Binomial Distribution

Multinomial Distribution

Poisson Distribution

Null hypothesis

Type I and II errors

STRUCTURE OF GENOMES

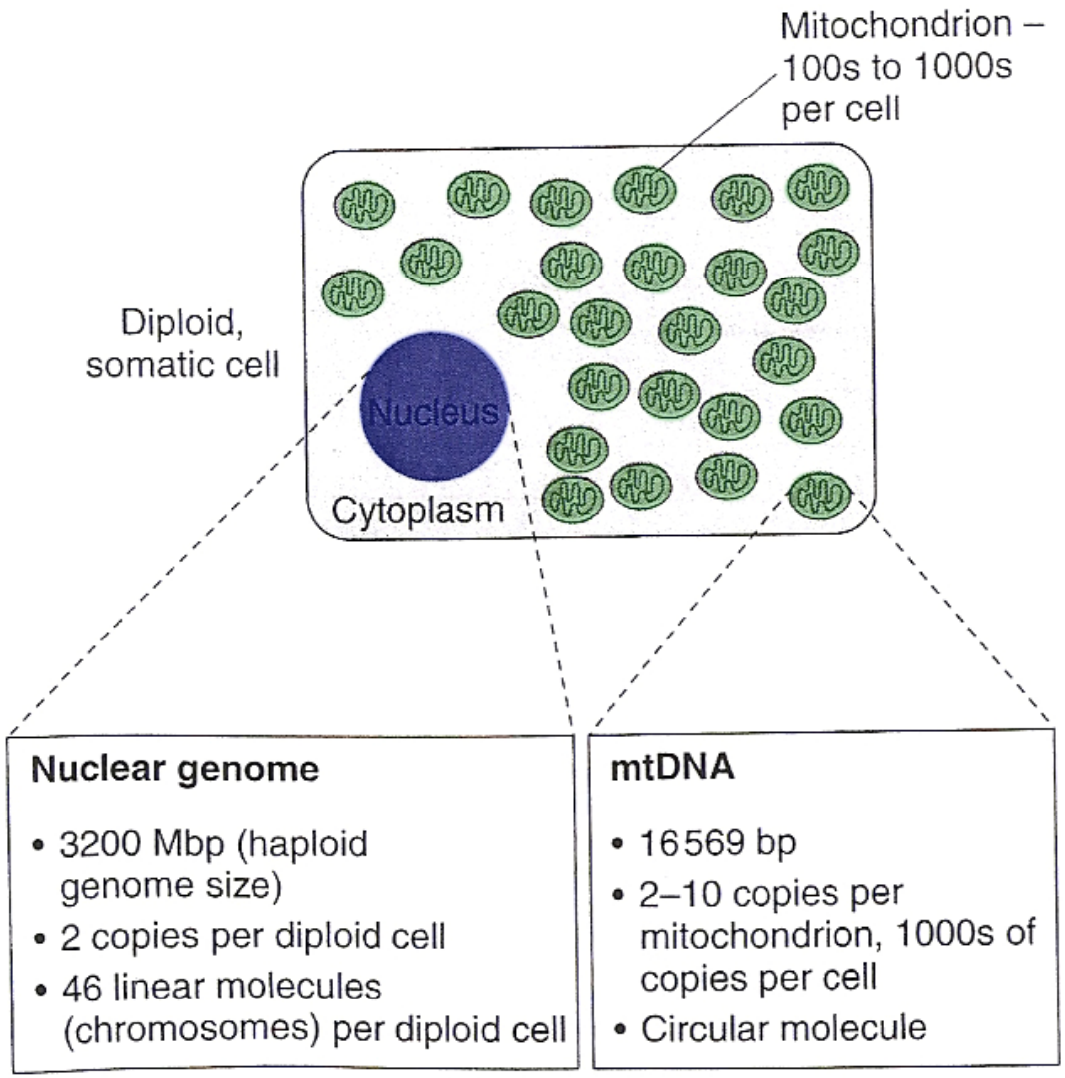
STRUCTURE AND INHERITENCE

The structure and inheritance of genomes described here will be biased towards the most common pattern in mammals.

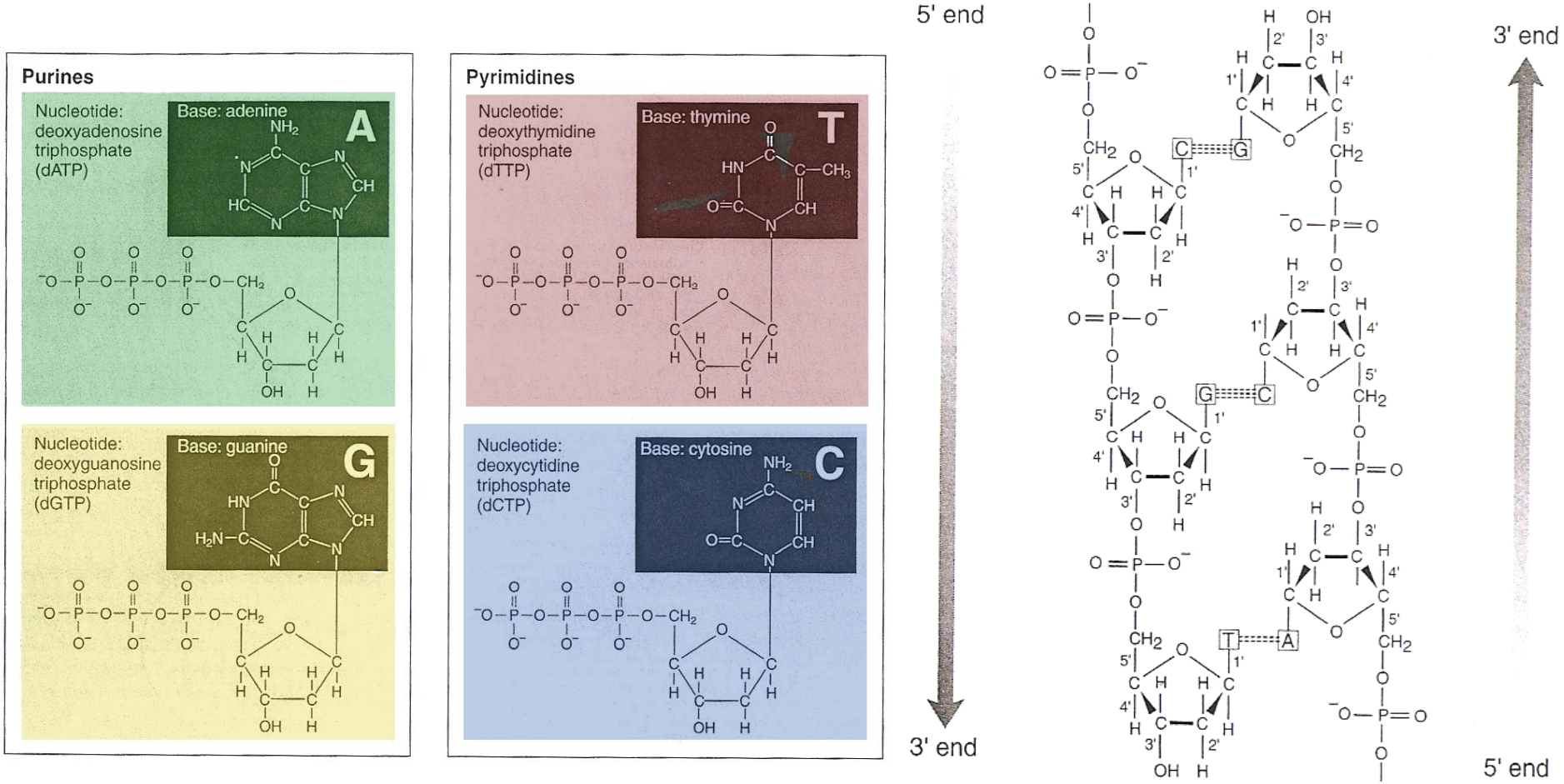
One must be aware that there is variation across organisms and that patterns of diversity also depend on this.

For example:
- Ploidy;
- Sex determination;
- Cytoplasmic compartments of inheritance;
- Etc.

OVERVIEW OF THE GENOME

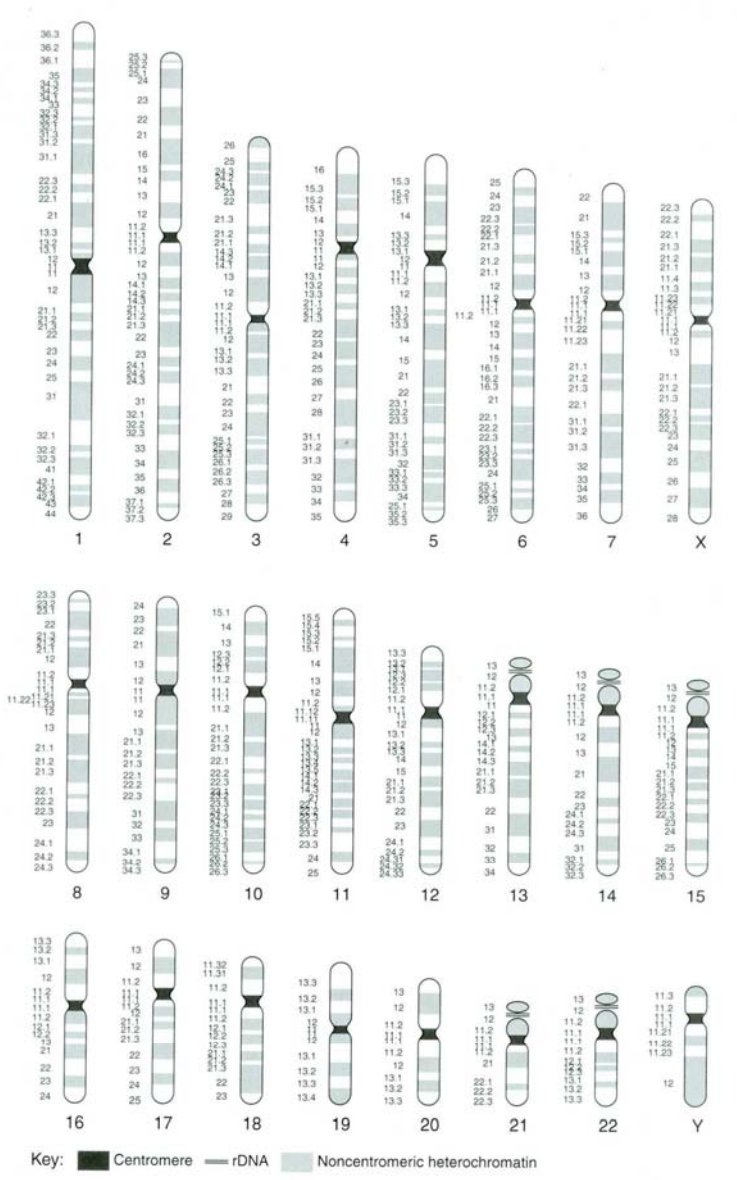


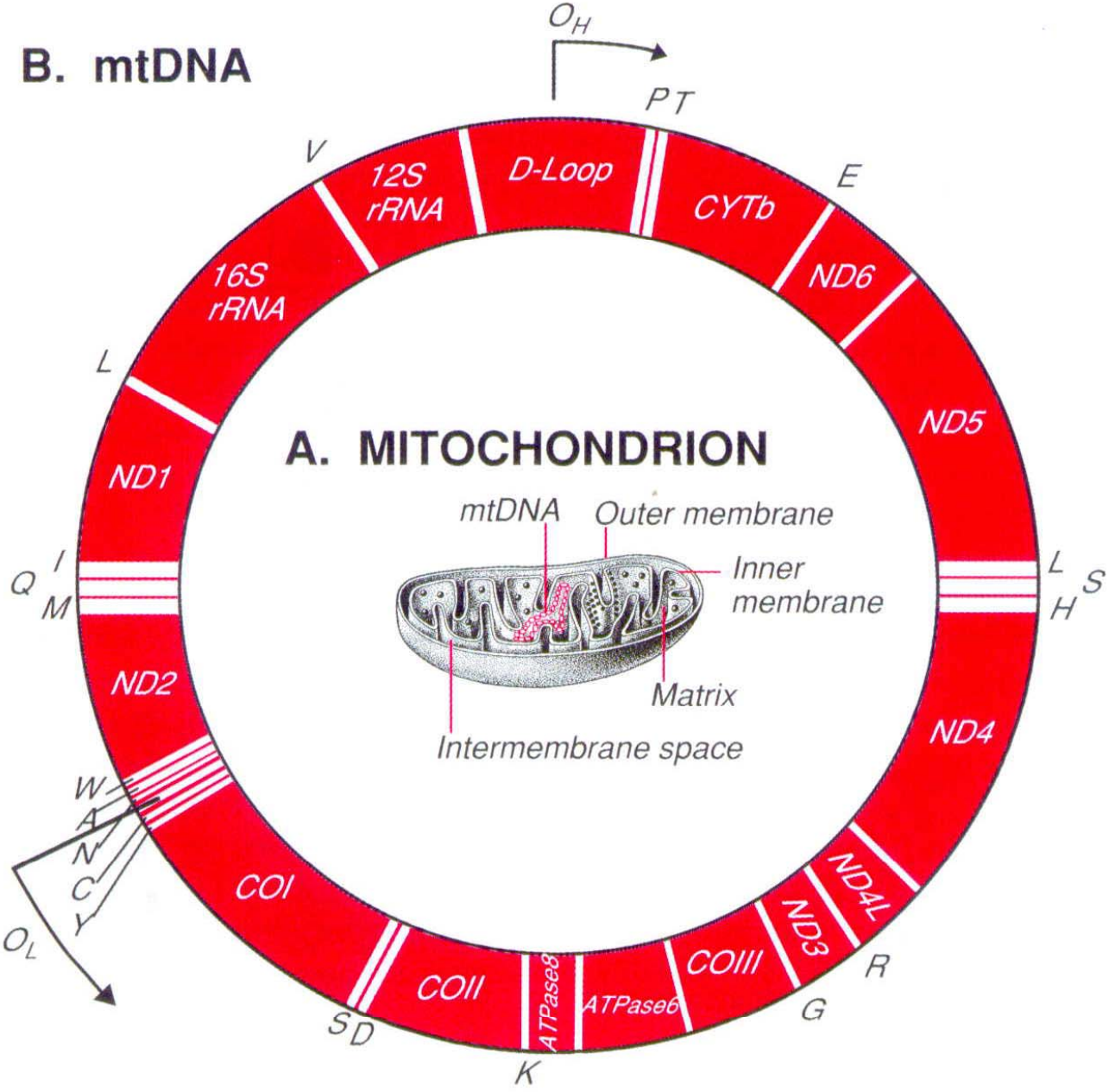
DNA

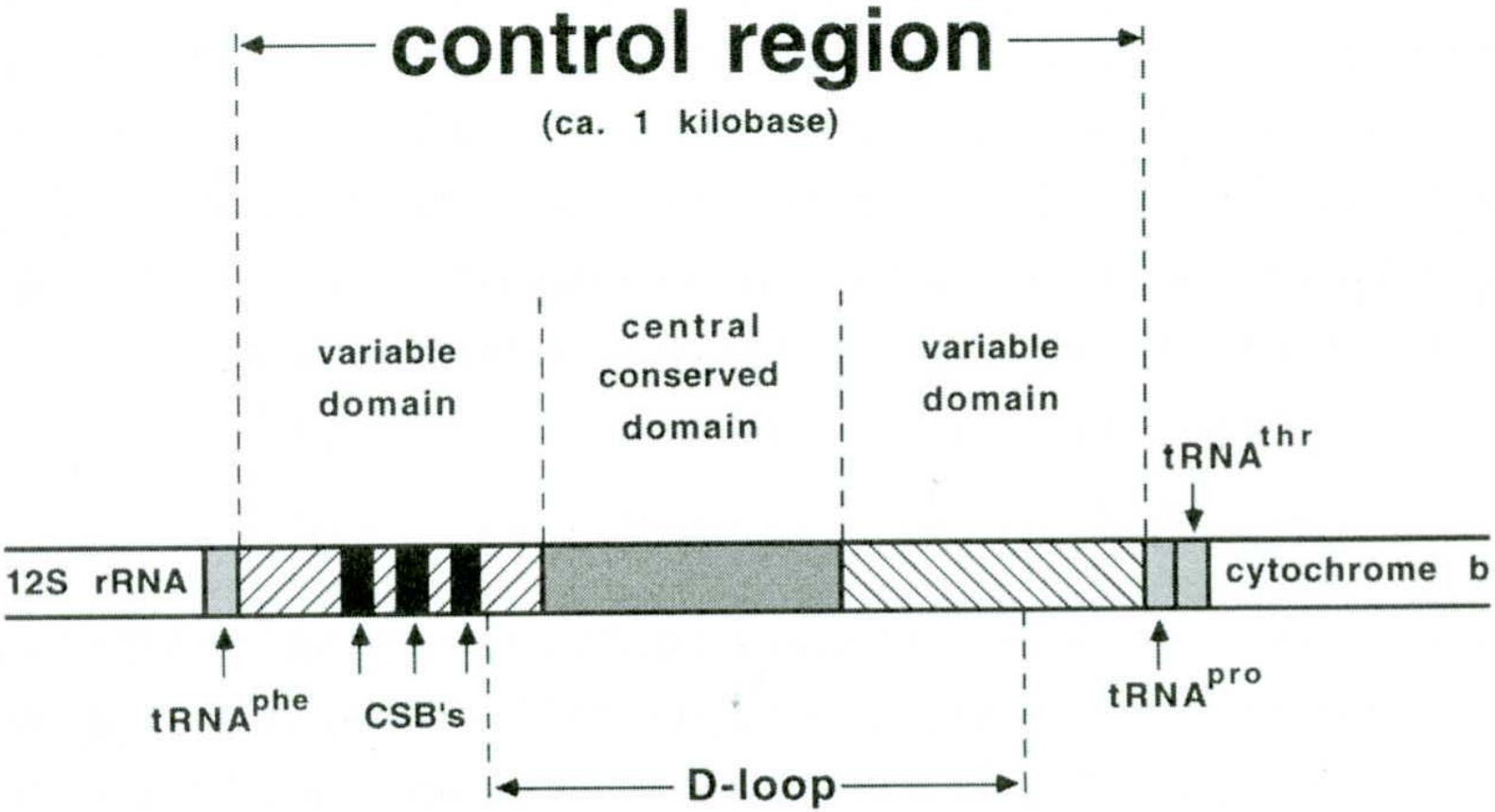


The succession of the nucleotides in a DNA molecule is a DNA SEQUENCE (e.g. AGCTATTTCCAGGA).

CHROMOSOMES

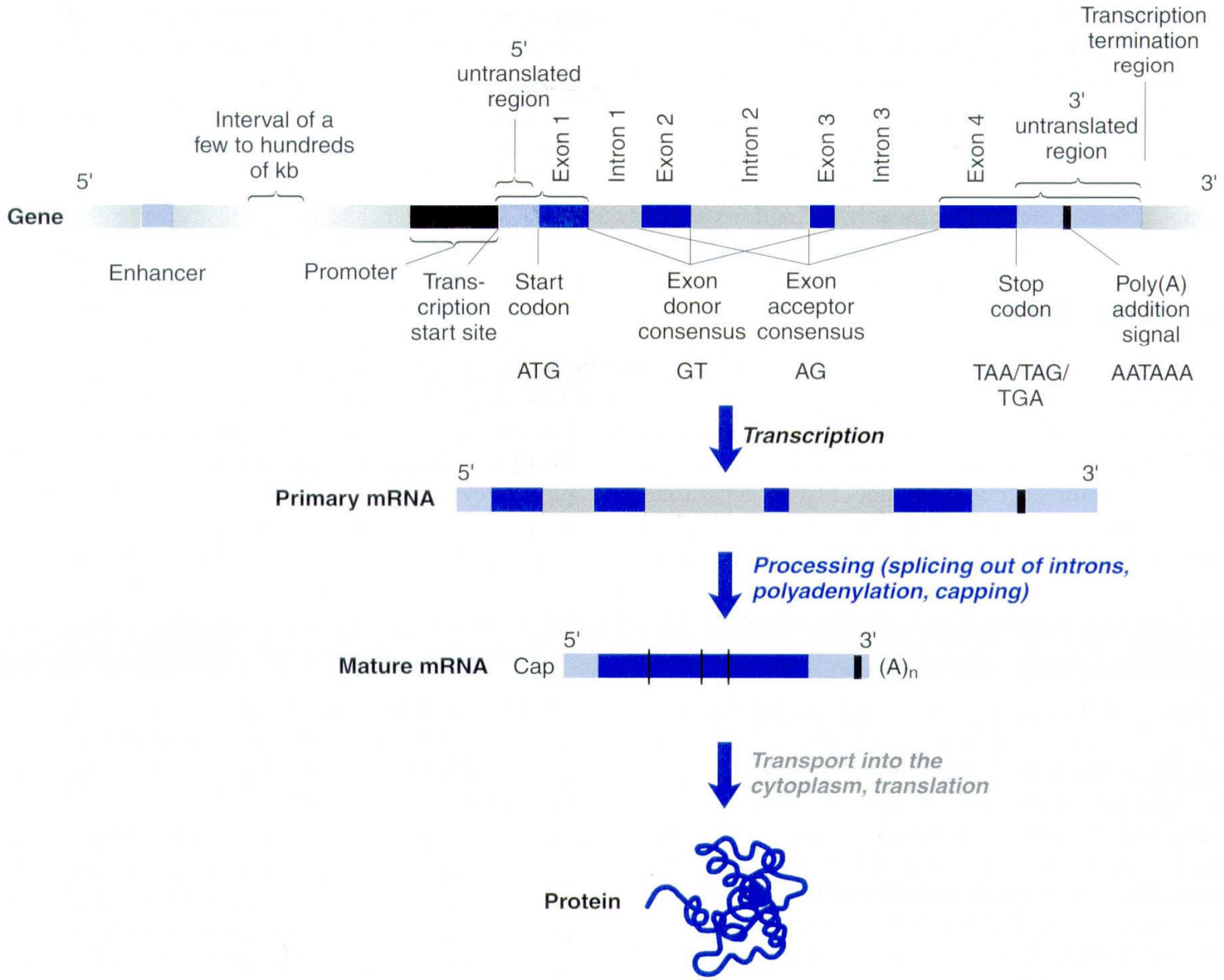


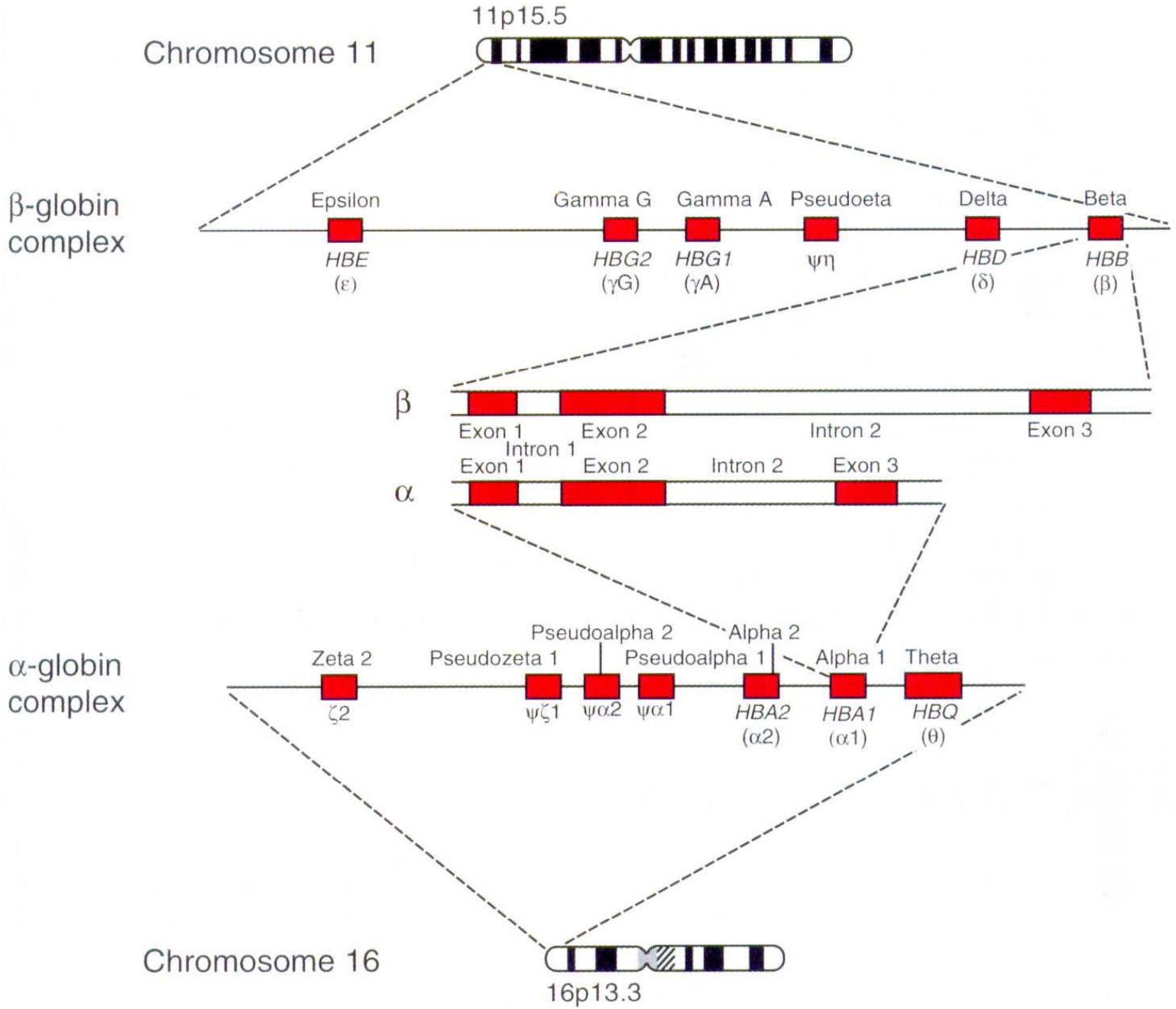

B. mtDNA
O_H
PT
E
CYTb
ND6
ND5
L
H
S
ND4
ND4L
R
ND3
COIII
G
ATPase6
ATPase8
K
D
S
COII
COI
O_L
Y
C
N
A
W
ND2
M
Q
I
ND1
L
16S rRNA
V
12S rRNA
D-Loop
A. MITOCHONDRION
mtDNA
Outer membrane
Inner membrane
Matrix
Intermembrane space


control region
(ca. 1 kilobase)
variable domain
central conserved domain
variable domain
tRNA thr
12S rRNA
cytochrome b
tRNA phe
CSB's
tRNA pro
D-loop

GENES

Most of the DNA present in a genome has no specific function, but some segments contain the instructions for the synthesis of proteins (or sometimes RNA): GENES.


Transcription termination region
5' untranslated region
3' untranslated region
Interval of a few to hundreds of kb
Exon 1
Intron 1
Exon 2
Intron 2
Exon 3
Intron 3
Exon 4
5'
3'
Gene
Enhancer
Promoter
Trans-cription start site
Start codon
ATG
Exon donor consensus
GT
Exon acceptor consensus
AG
Stop codon
TAA/TAG/TGA
Poly(A) addition signal
AATAAA
Transcription
Primary mRNA
Processing (splicing out of introns, polyadenylation, capping)
Mature mRNA
Cap
(A)n
Transport into the cytoplasm, translation
Protein


11p15.5
Chromosome 11
β-globin complex
Epsilon
Gamma G
Gamma A
Pseudoeta
Delta
Beta
HBE (ε)
HBG2 (γG)
HBG1 (γA)
ψη
HBD (δ)
HBB (β)
β
Exon 1
Exon 2
Intron 2
Exon 3
Intron 1
α
Exon 1
Exon 2
Intron 2
Exon 3
α-globin complex
Pseudoalpha 2
Alpha 2
Zeta 2
Pseudozeta 1
Pseudoalpha 1
Alpha 1
Theta
ζ2
ψζ1
ψα2
ψα1
HBA2 (α2)
HBA1 (α1)
HBQ (θ)
Chromosome 16
16p13.3

NONCODING DNA

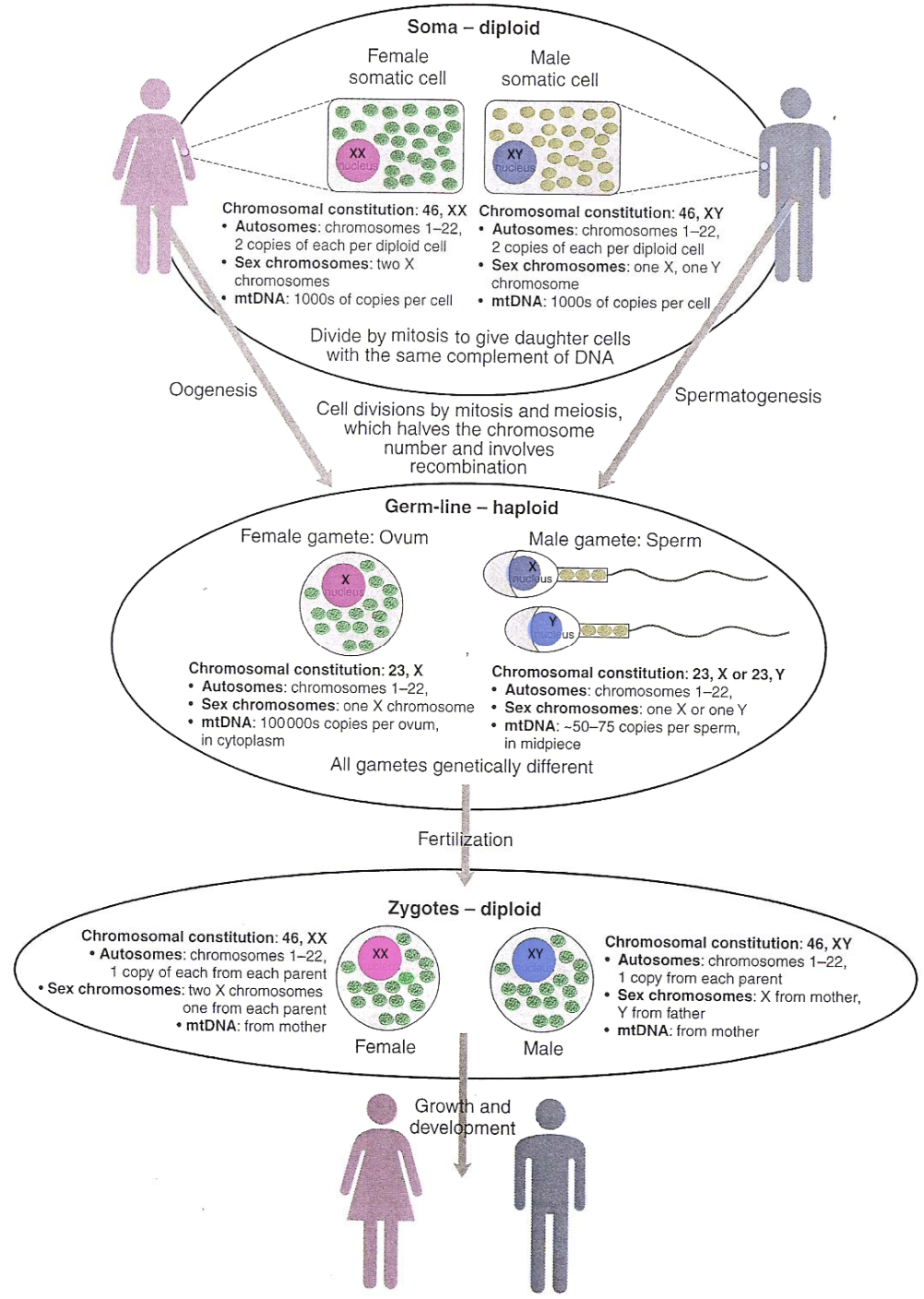
About 98.5% of the DNA does not contain coding sequences:

- Single copy DNA (e.g. most of the introns);

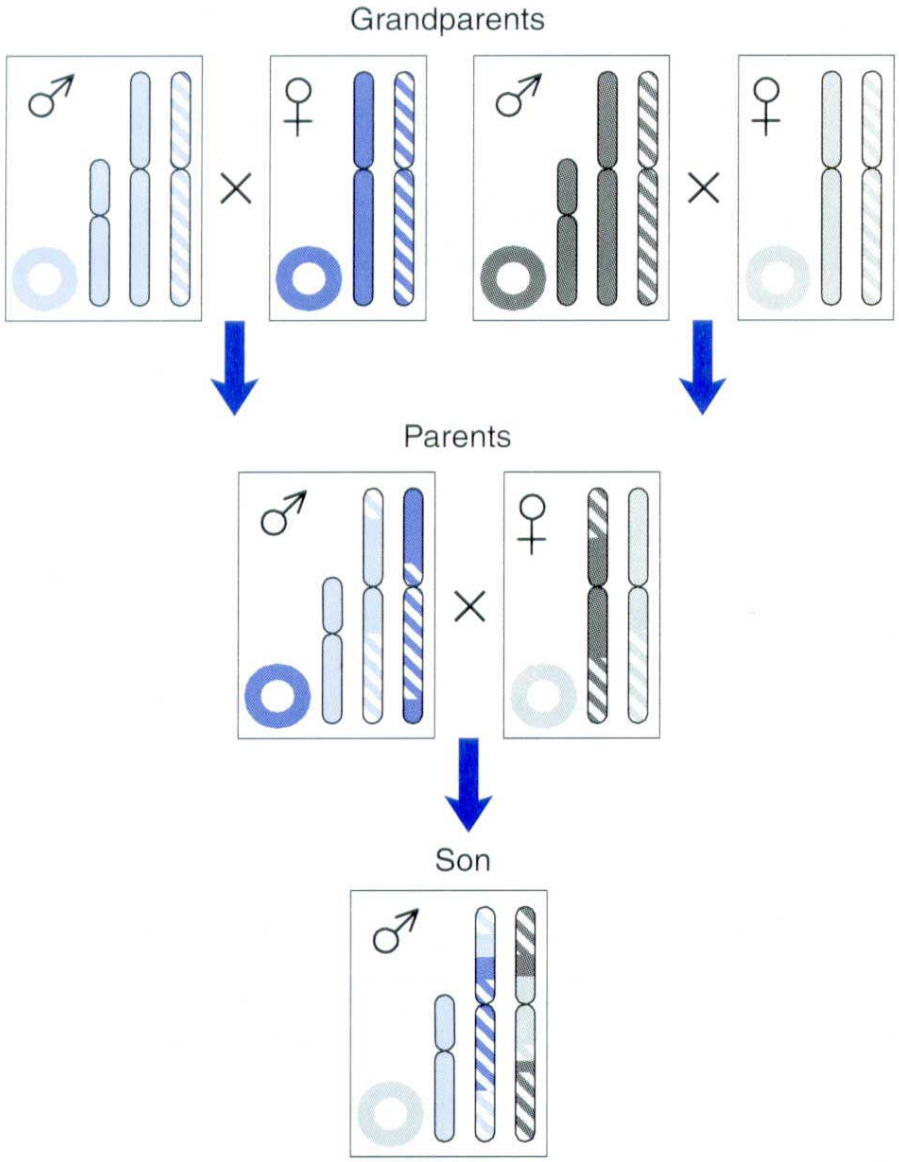
- Repetitive elements (e.g. tandem repeats; interspersed elements).

INHERITANCE OF GENETIC ELEMENTS

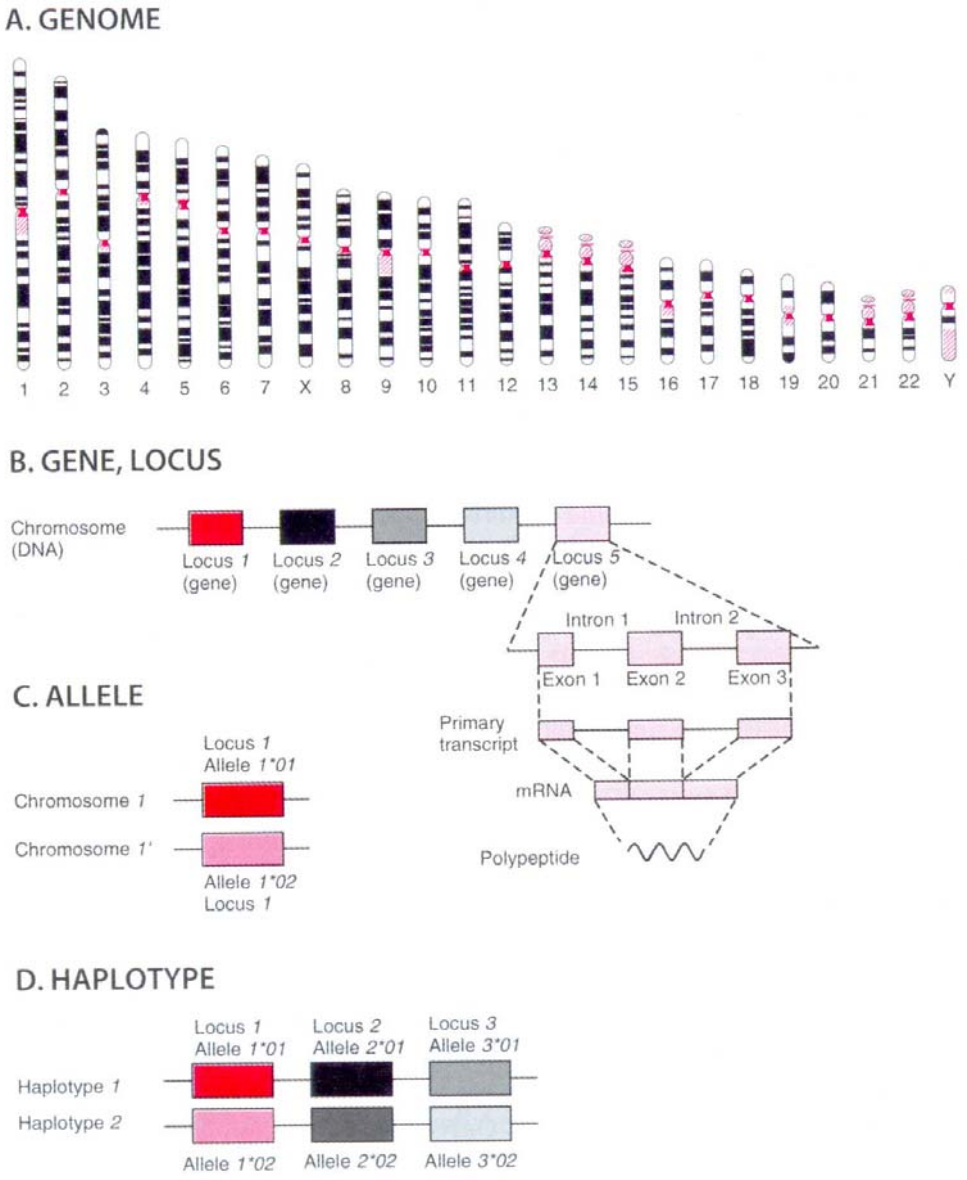
INHERITANCE OF GENOMIC REGIONS



INHERITANCE OF GENOMIC REGIONS



CONCEPTS


A. GENOME
1 2 3 4 5 6 7 X 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 Y
B. GENE, LOCUS
Chromosome (DNA)
Locus 1 (gene)
Locus 2 (gene)
Locus 3 (gene)
Locus 4 (gene)
Locus 5 (gene)
Intron 1
Intron 2
Exon 1
Exon 2
Exon 3
Primary transcript
mRNA
Polypeptide
C. ALLELE
Locus 1
Allele 1*01
Chromosome 1
Chromosome 1'
Allele 1*02
Locus 1
D. HAPLOTYPE
Locus 1
Allele 1*01
Locus 2
Allele 2*01
Locus 3
Allele 3*01
Haplotype 1
Haplotype 2
Allele 1*02
Allele 2*02
Allele 3*02

GENETIC VARIATION

MUTATION

- Is the ultimate source of genetic variation.

- It refers to any change that produces a new allele;

- Only mutations that occur in the germ-line are passed to the next generation;

- Types of mutation:

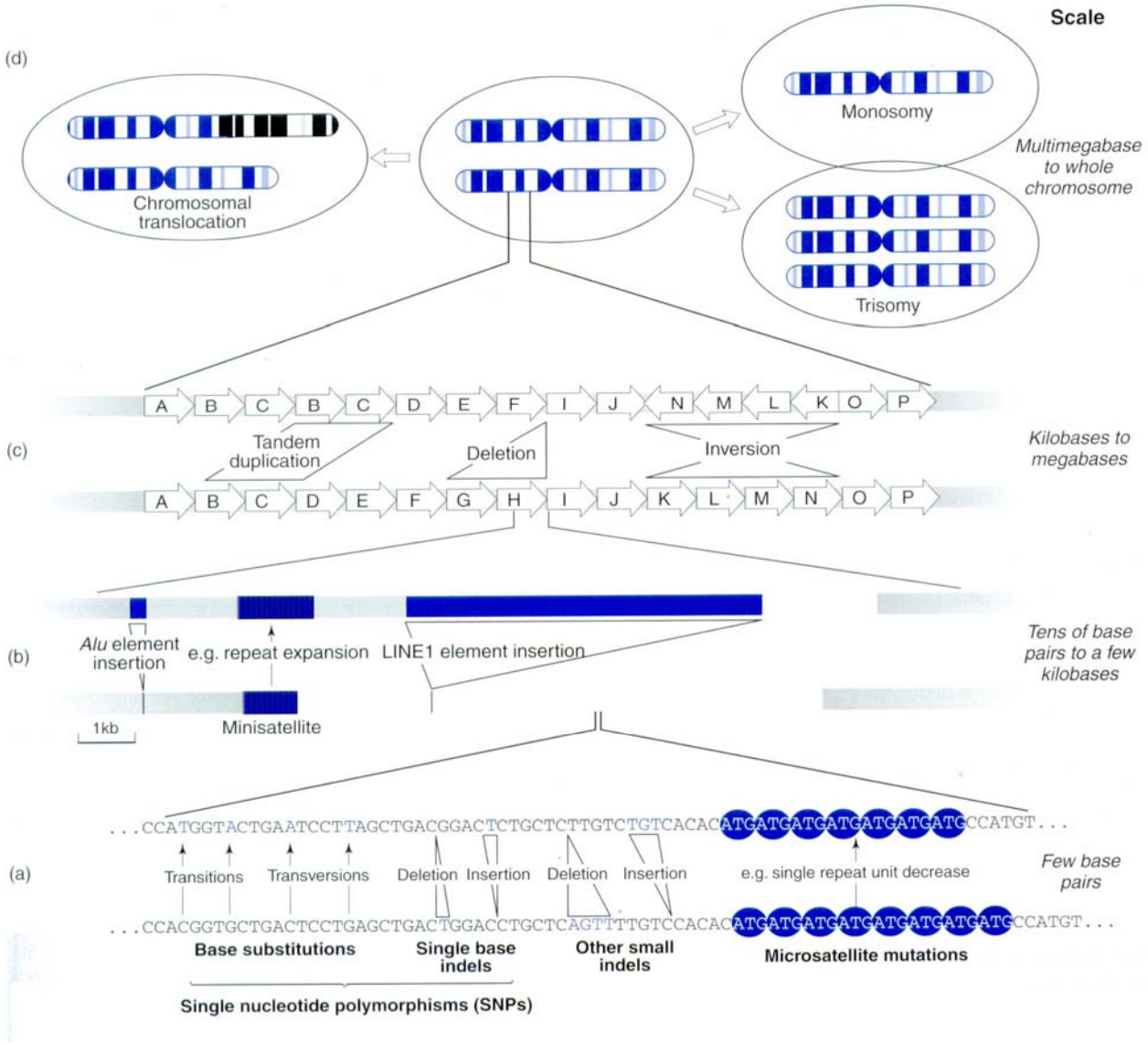
    - Inversions and translocations;

    - Duplications, insertions, deletions;

    - Substitutions.

- It covers a broad spectrum, from the substitution of single nucleotide in the genome to changes in chromosome number.

MUTATION


Scale
(d)
Monosomy
Chromosomal translocation
Trisomy
Multimegabase to whole chromosome
(c)
A B C B C D E F I J N M L K O P
Tandem duplication
Deletion
Inversion
A B C D E F G H I J K L M N O P
Kilobases to megabases
(b)
Alu element insertion
e.g. repeat expansion
LINE1 element insertion
1kb
Minisatellite
Tens of base pairs to a few kilobases
(a)
...CCATGGTACTGAATCCTTAGCTGACGGACTCTGCTCTTGTCTGTCACACATGATGATGATGATGATGATGCCATGT...
Transitions
Transversions
Deletion
Insertion
Deletion
Insertion
e.g. single repeat unit decrease
Few base pairs
...CCACGGTGCTGACTCCTGAGCTGACTGGACCTGCTCAGTTTTGTCCACACATGATGATGATGATGATGATGATGCCATGT...
Base substitutions
Single base indels
Other small indels
Microsatellite mutations
Single nucleotide polymorphisms (SNPs)

BASE SUBSTITUTIONS

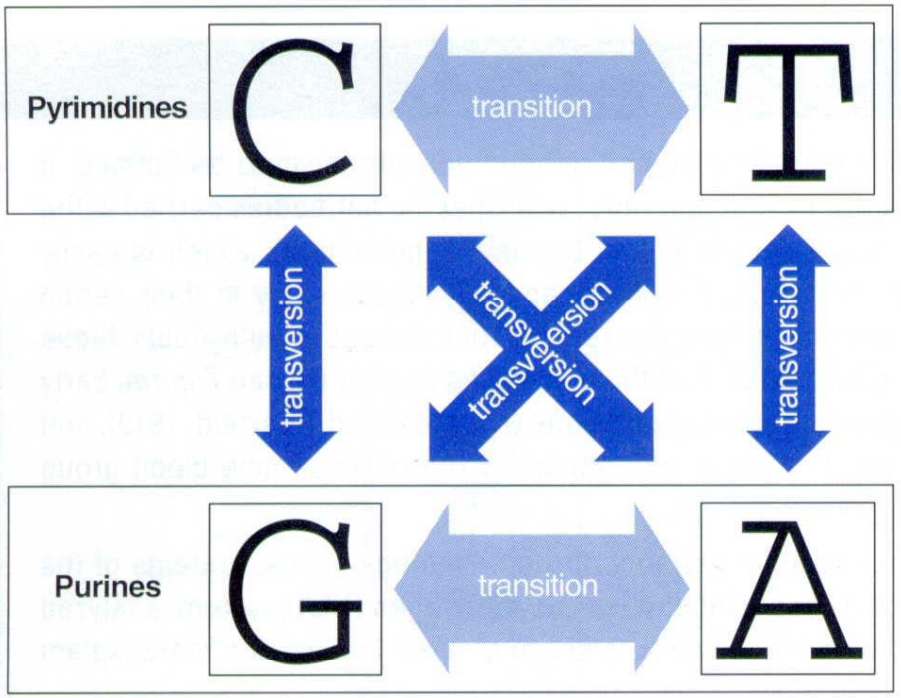
- Is the simplest difference between homologous DNA sequences.

- Are cause by i) misincorporation of nucleotides during replication or ii) chemical or physical mutagenesis.

- The originated polymorphisms are called “Single Nucleotide Polymorphisms” – SNPs (where single base insertions or deletions are also included).

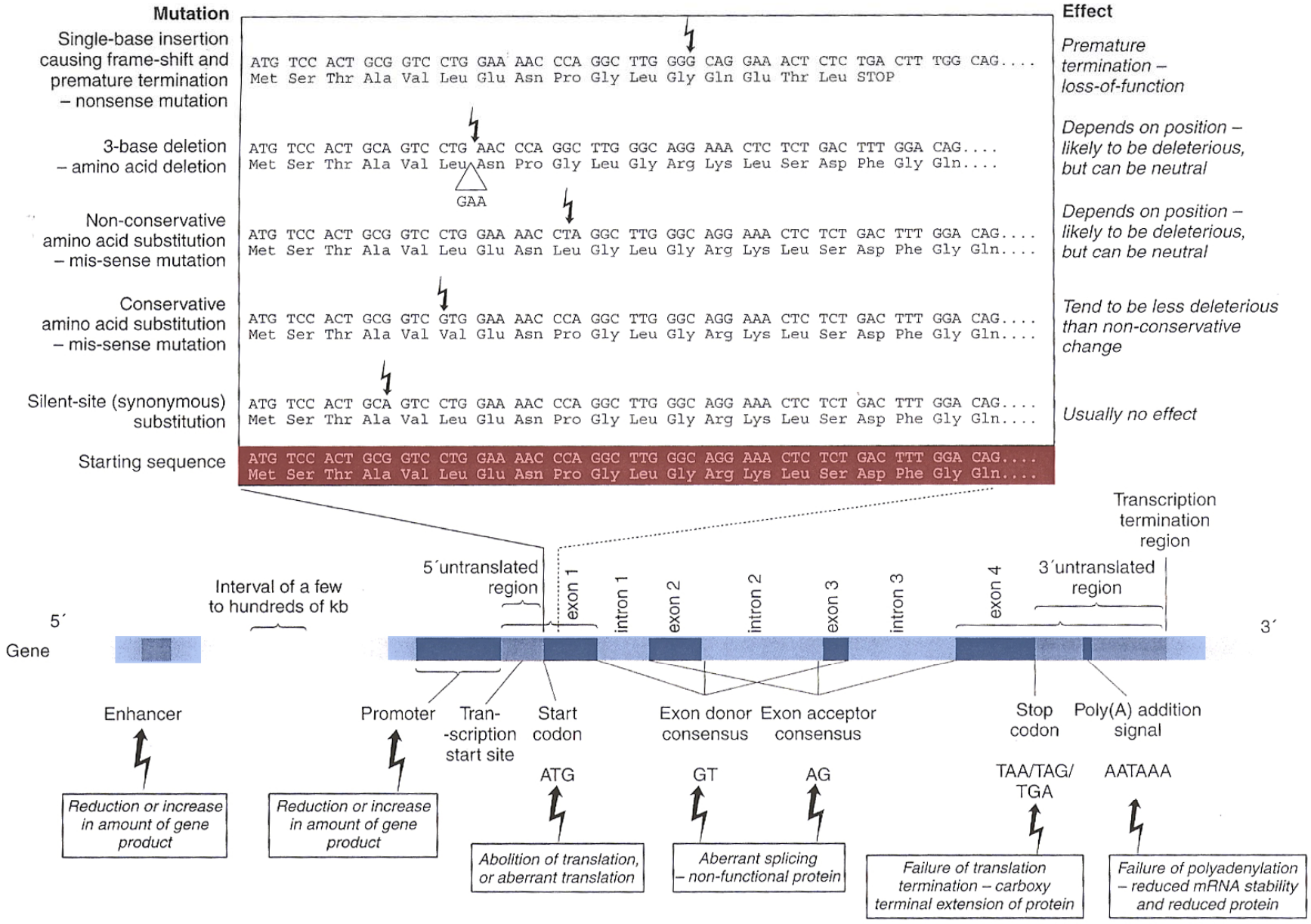
BASE SUBSTITUTIONS

-Transition vs. transversion:

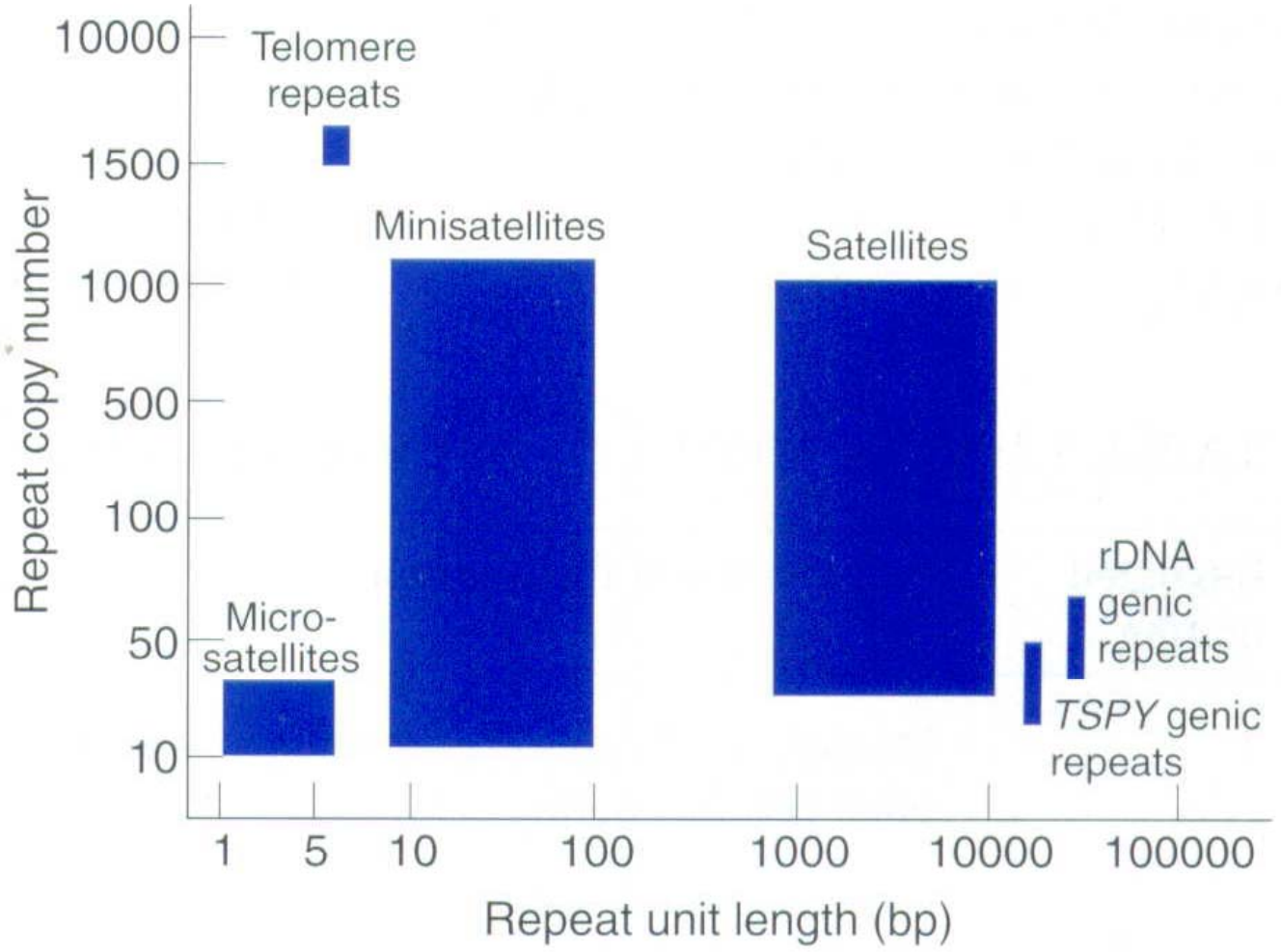


- Synonymous vs. non-synonymous.

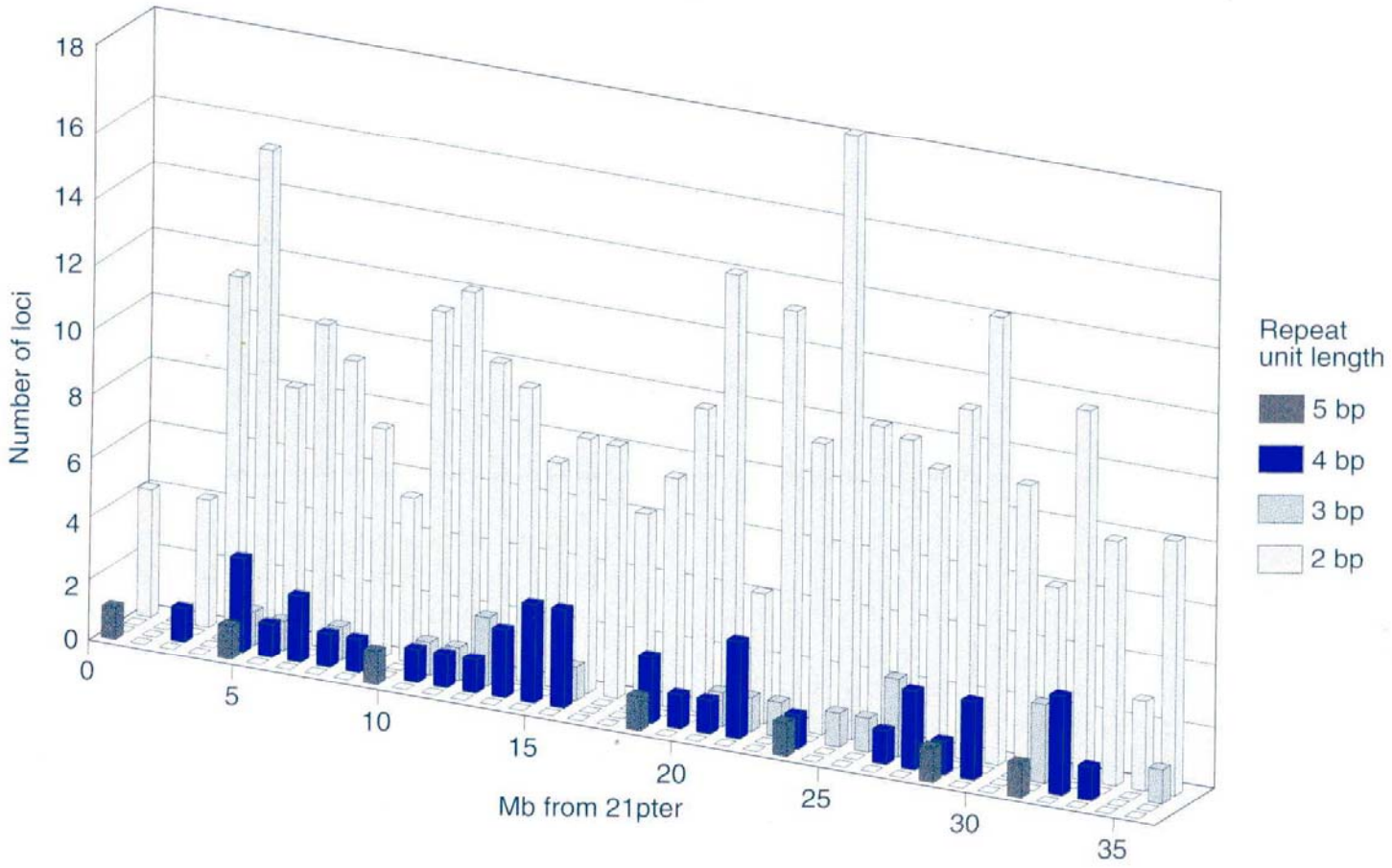
SNPs WITHIN GENES



VARIABLE NUMBER OF TANDEM REPEATS



MICROSATELLITES (STRs)



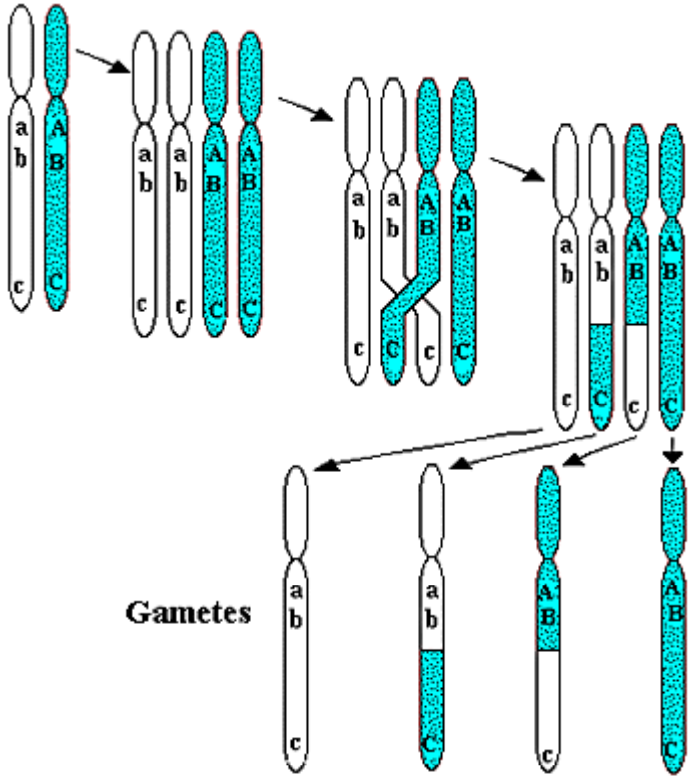
OTHER MUTATIONS

- Transposable element insertions:
  - LINEs
  - SINEs

- Structural mutations:
  - Large inversions, deletions, duplications, lenght variations…

RECOMBINATION

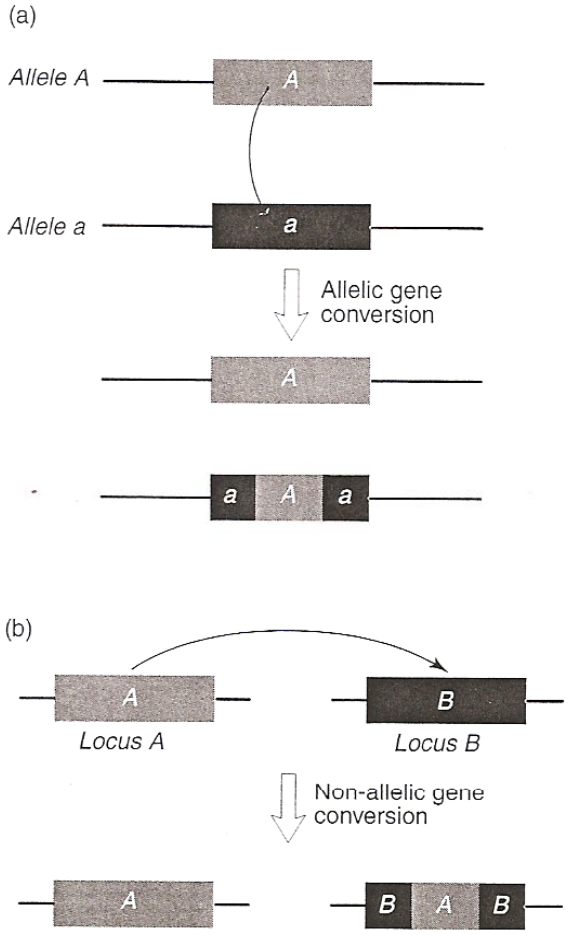
- Reciprocal transfer of genetic information.

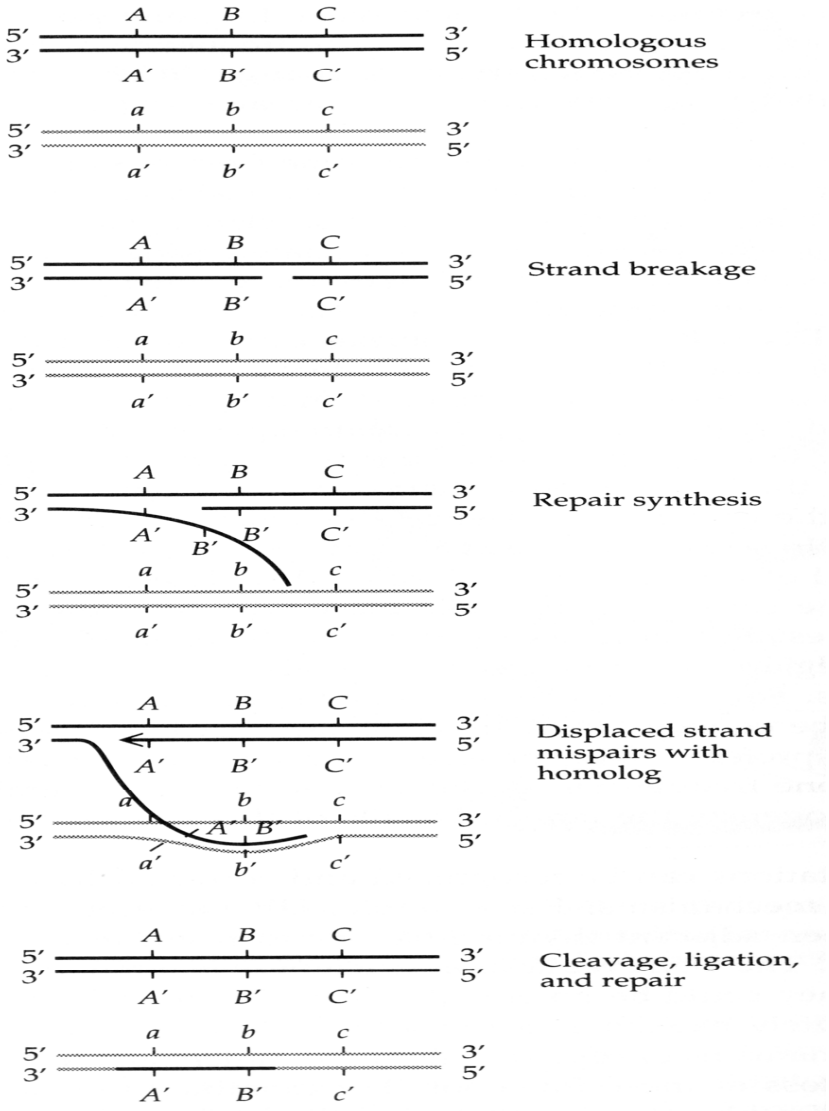


Crossing-over and recombination during meiosis

GENE CONVERSION

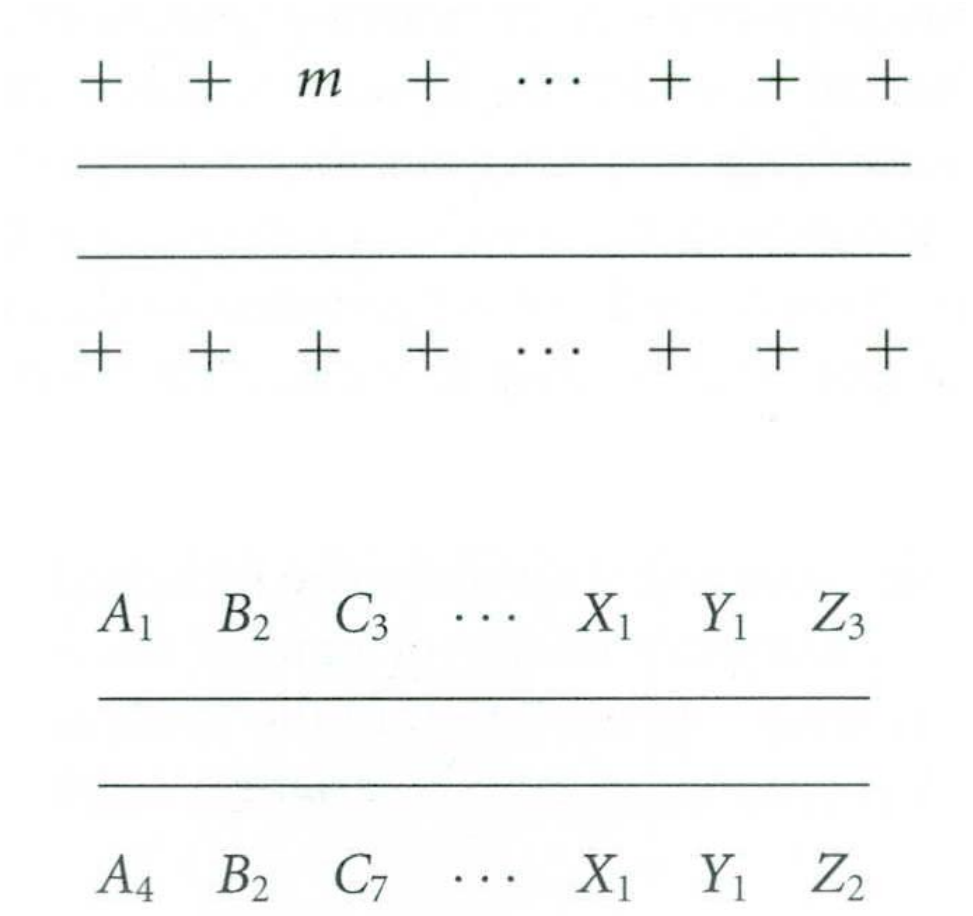
- Nonreciprocal transfer of genetic information.



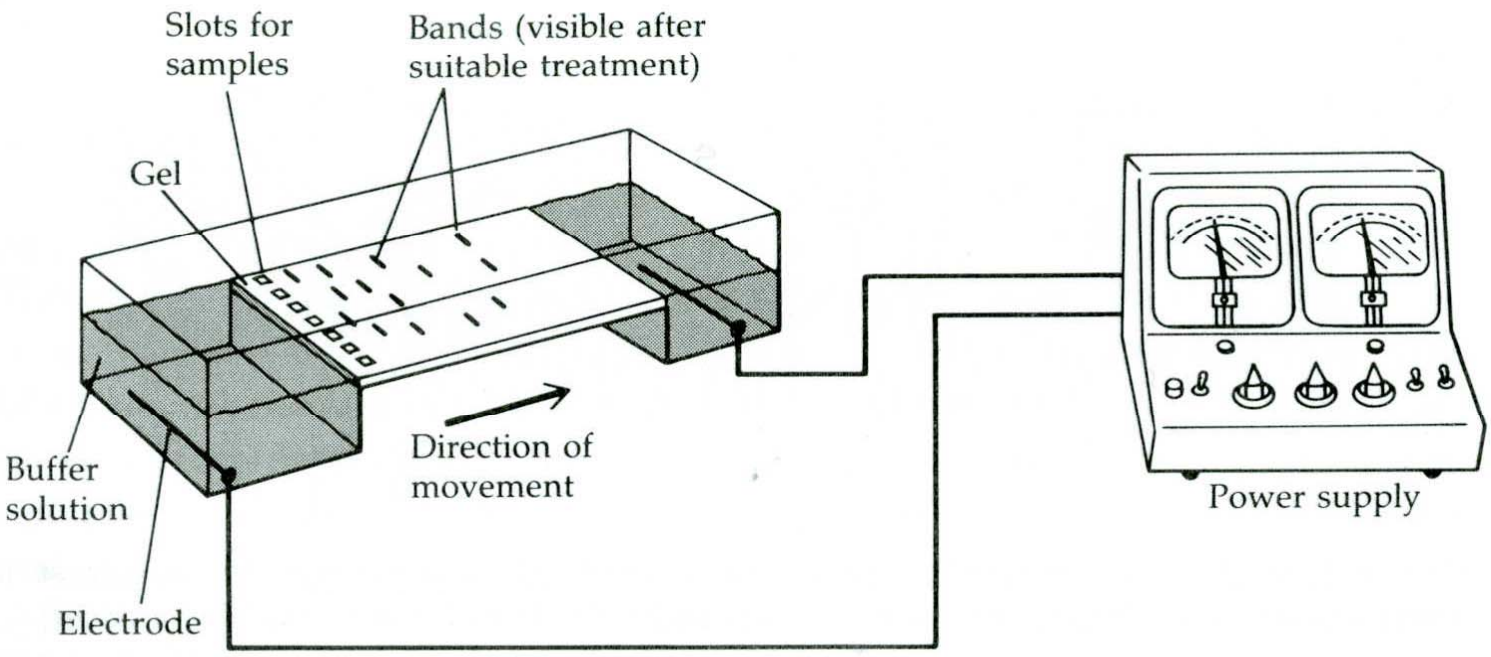


DETECTING GENETIC VARIATION

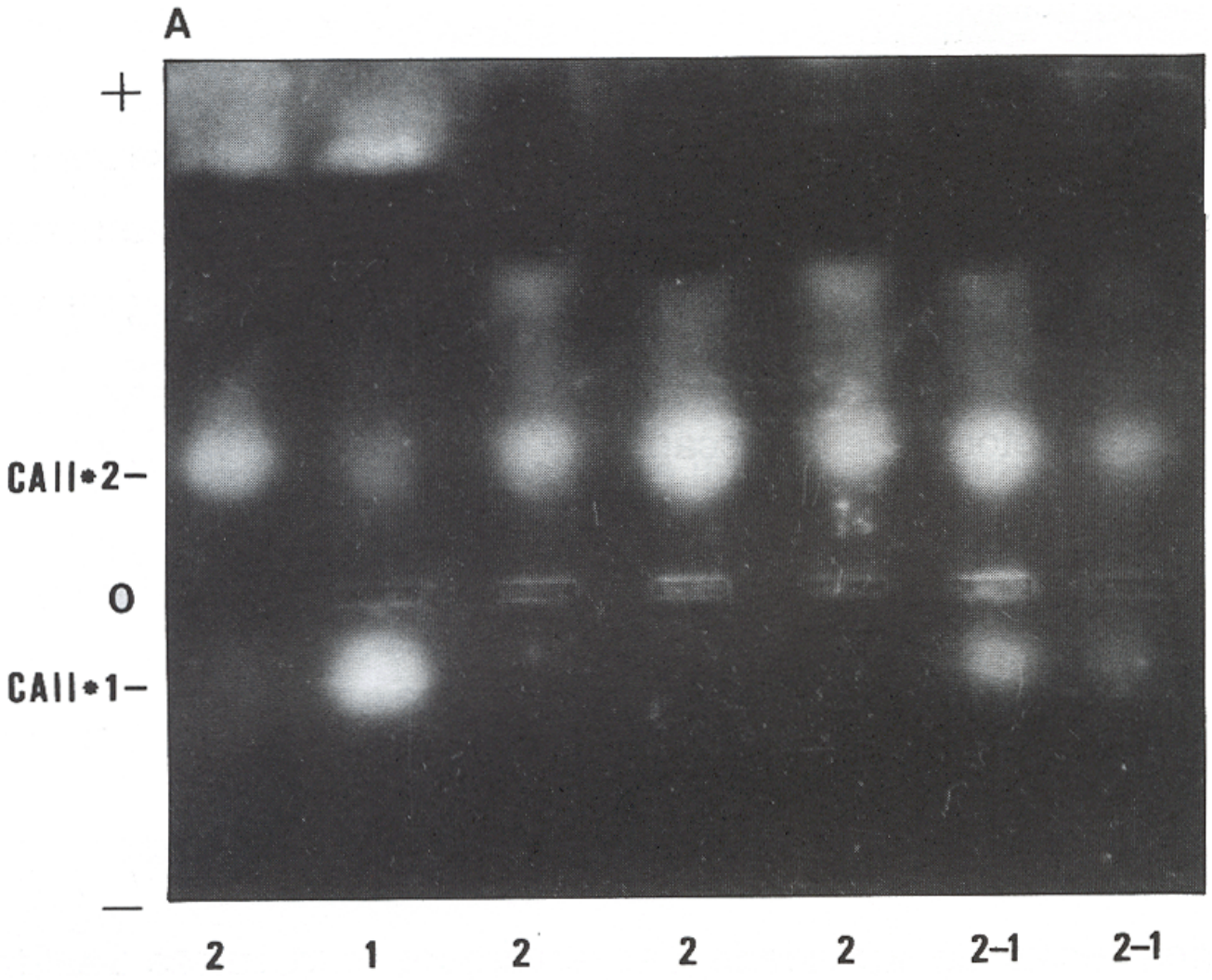
CLASSICAL-BALANCE VIEWS OF GENOME STRUCTURE



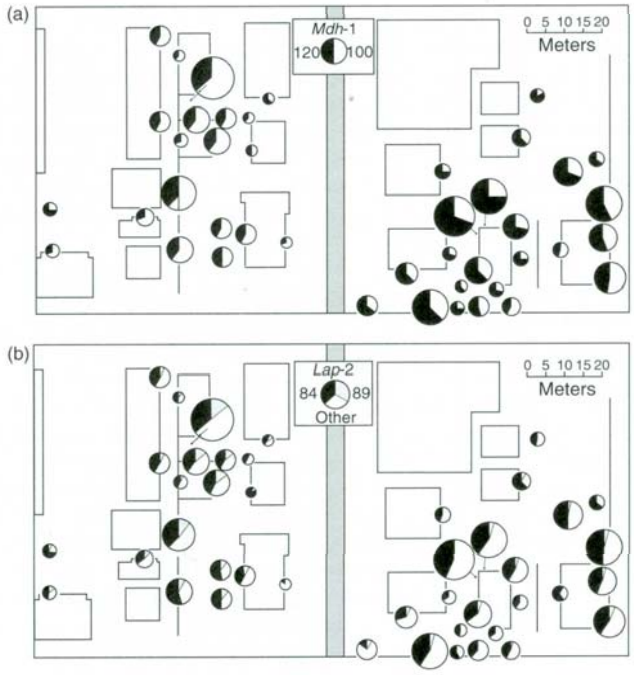
PROTEIN ELECTROPHORESIS

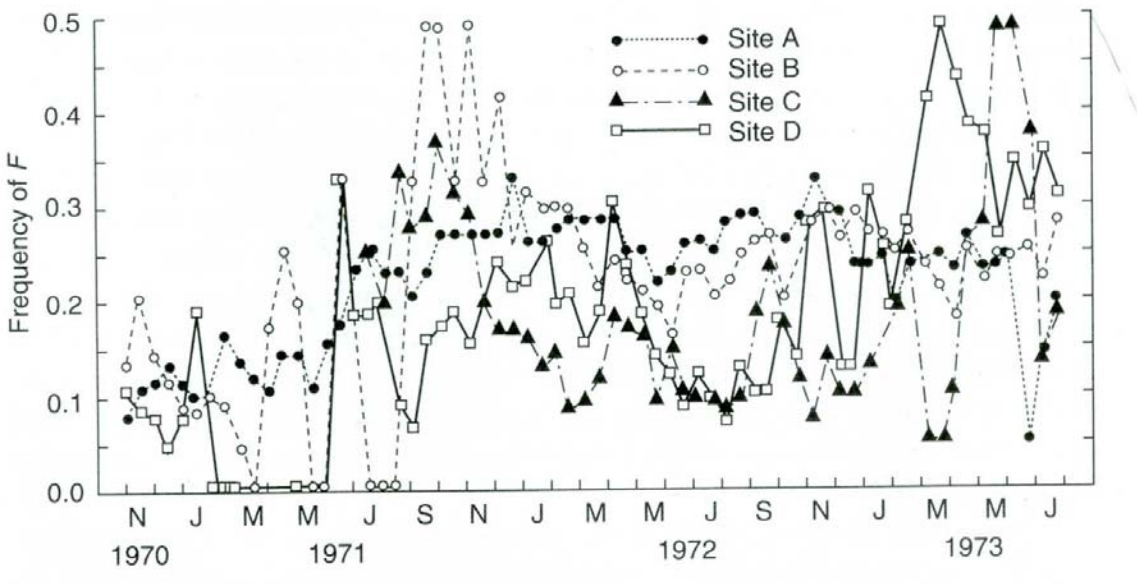


PROTEIN ELECTROPHORESIS

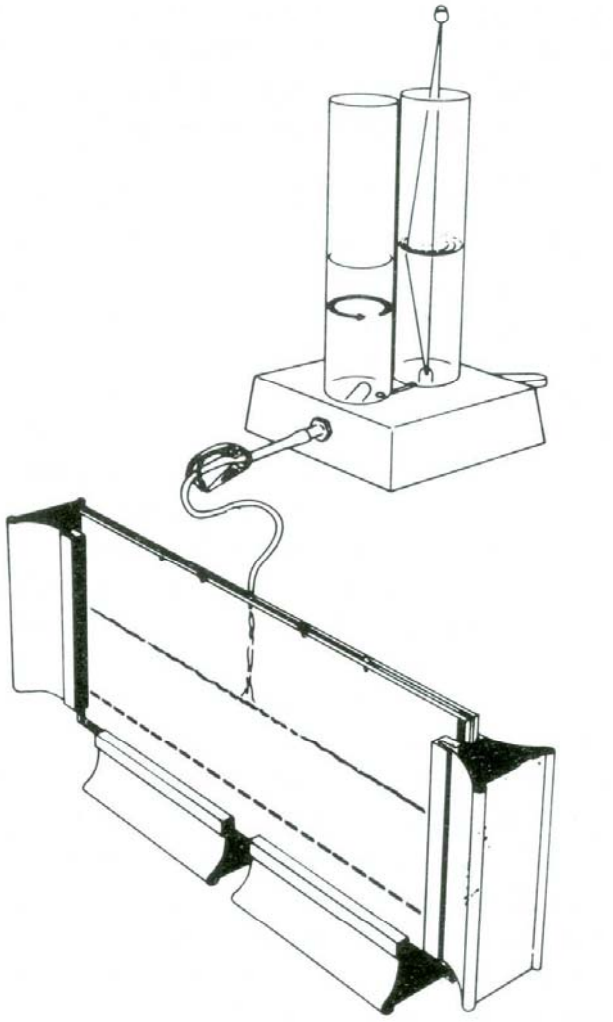


PROTEIN ELECTROPHORESIS

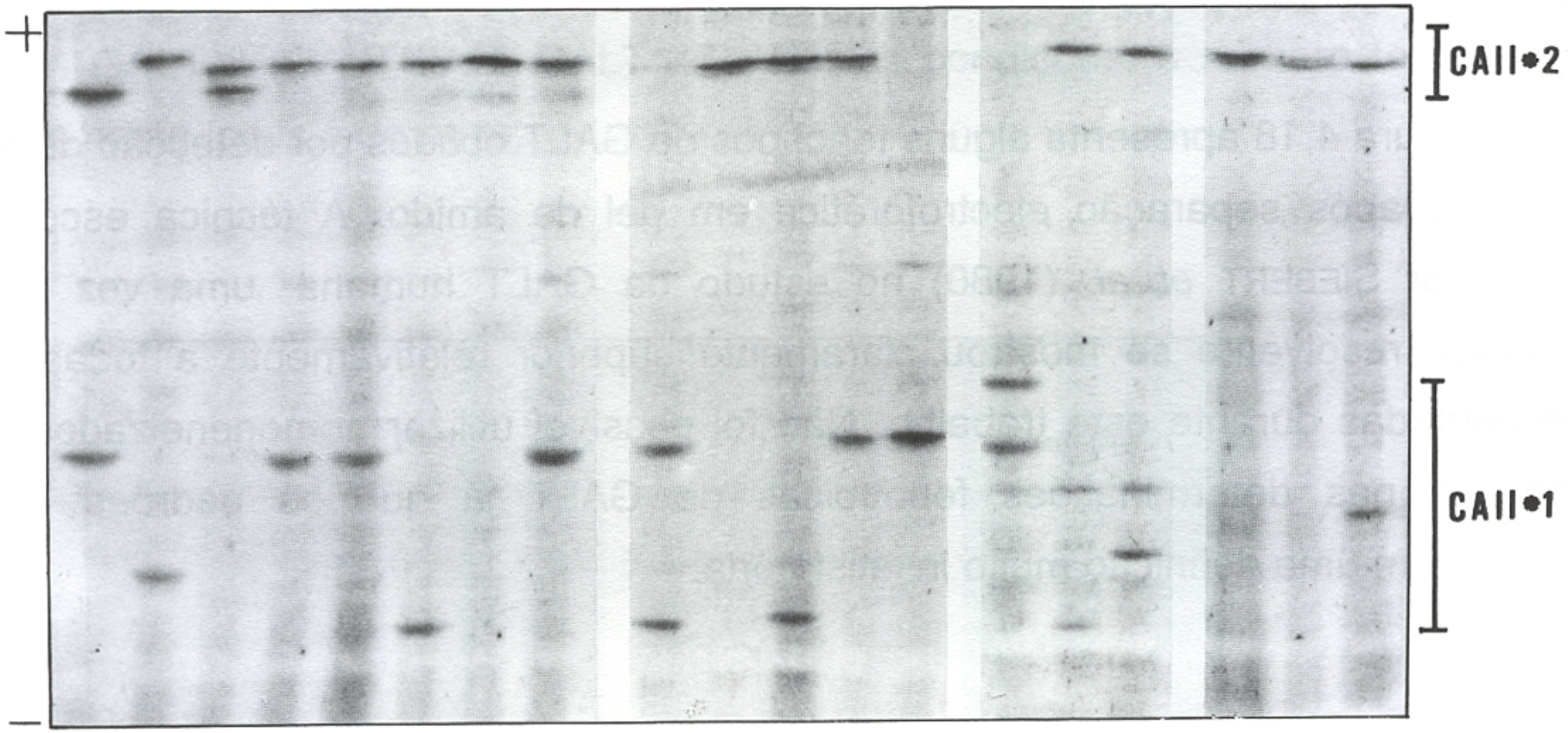




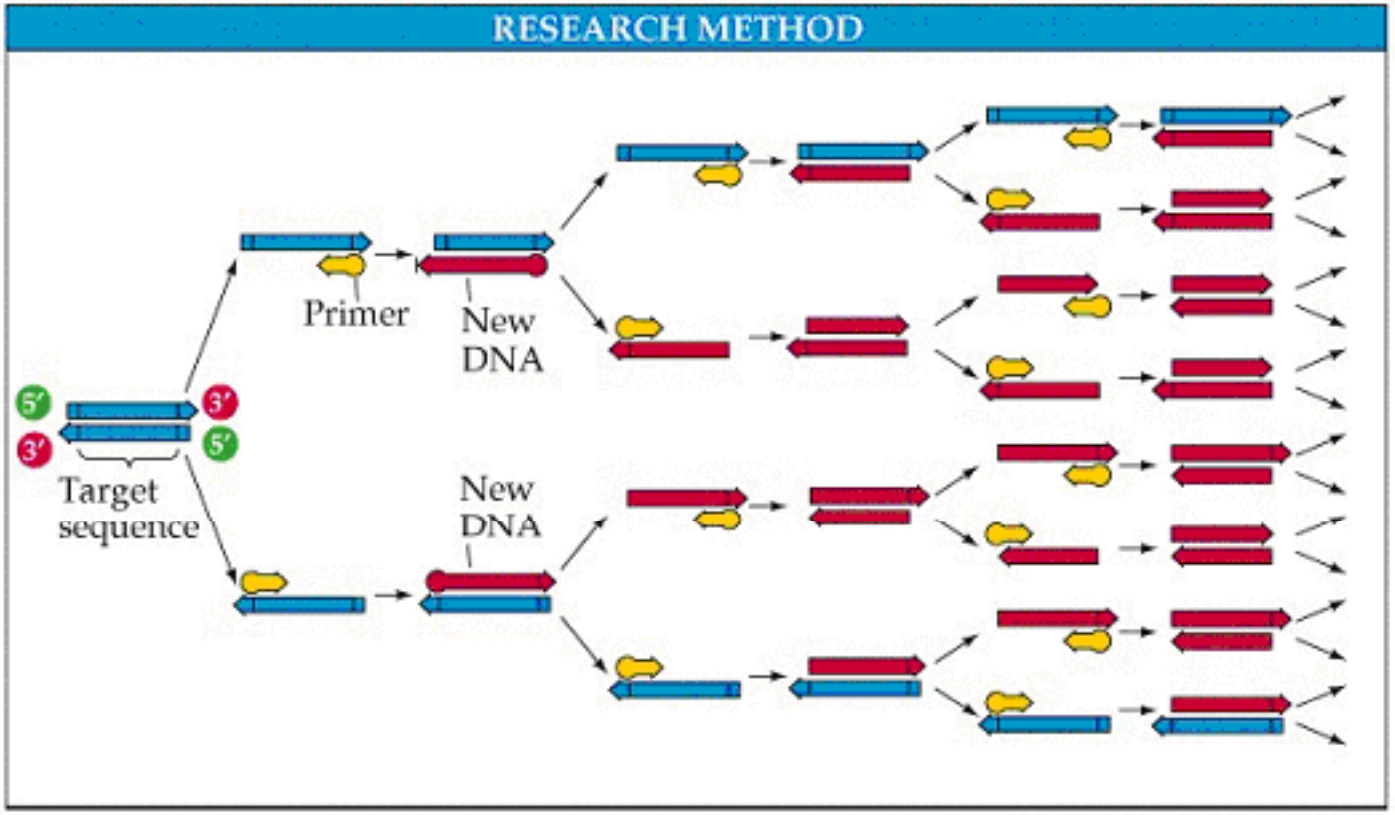
PROTEIN ISOELECTIC FOCUSING



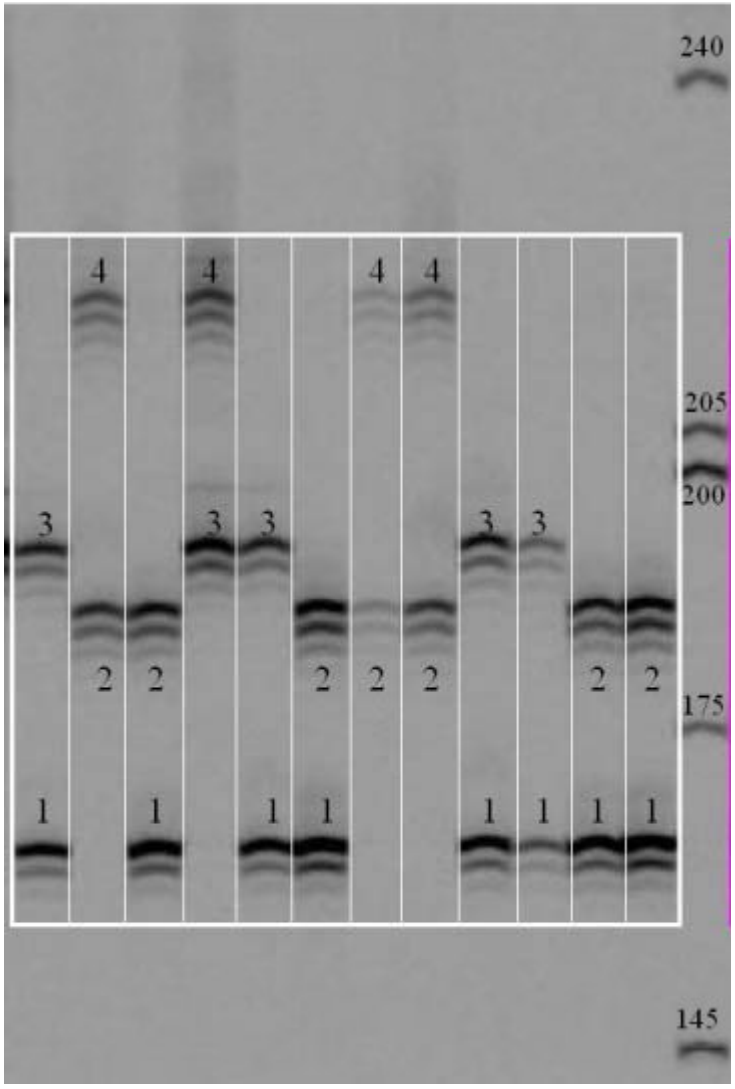
PROTEIN ISOELECTIC FOCUSING

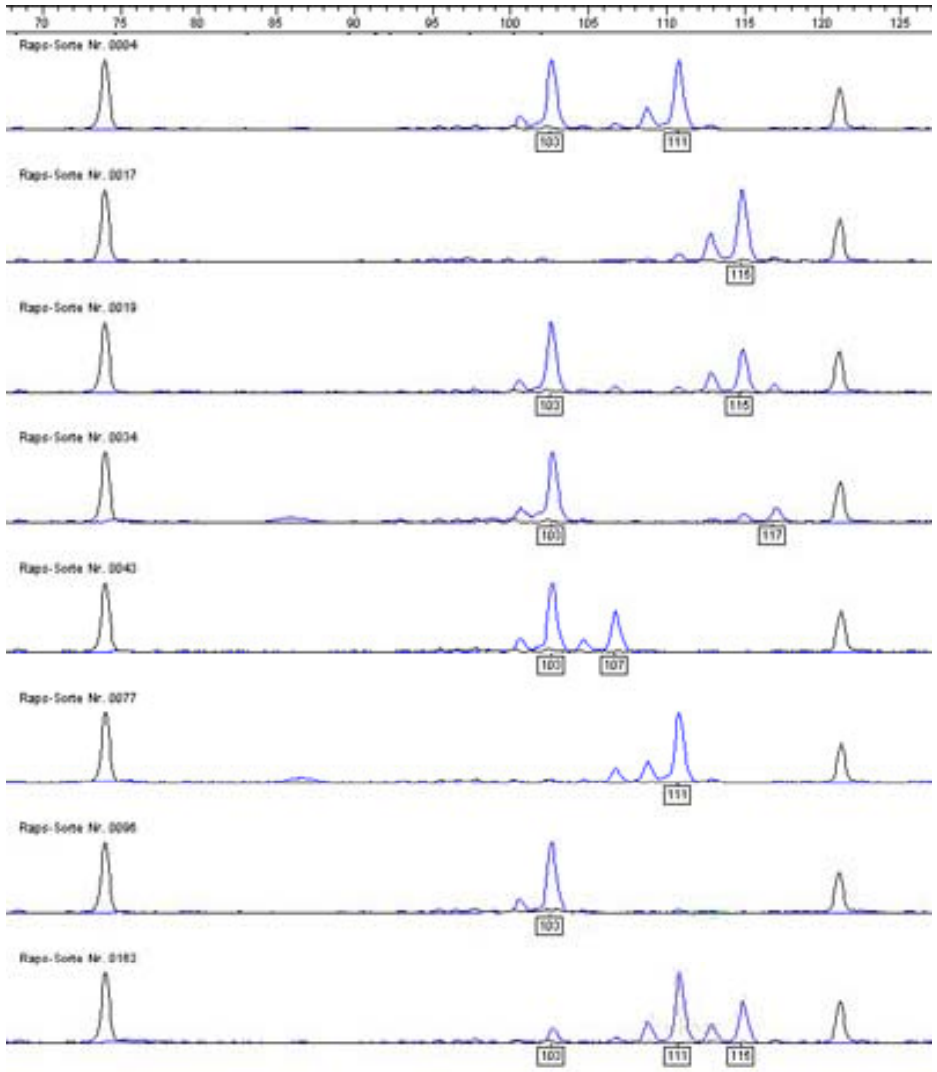


POLIMERASE CHAIN REACTION (PCR)

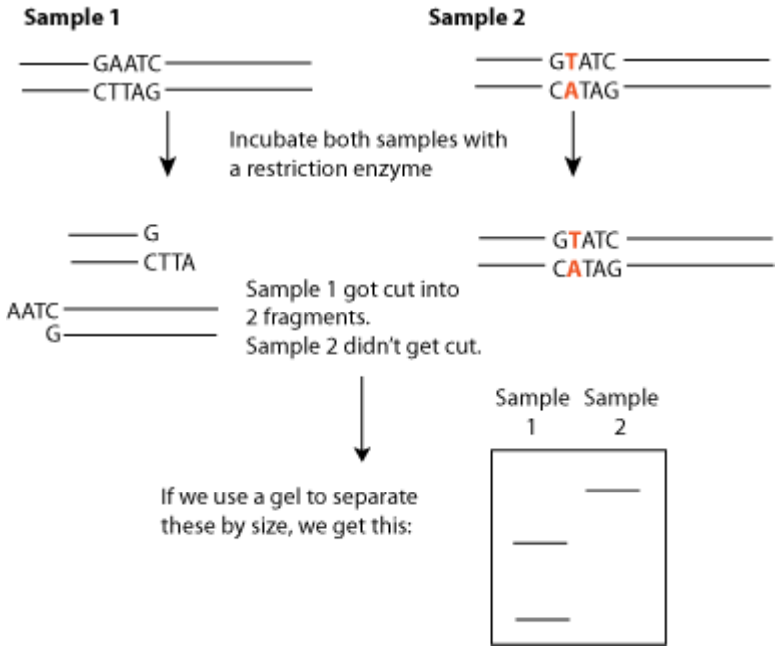


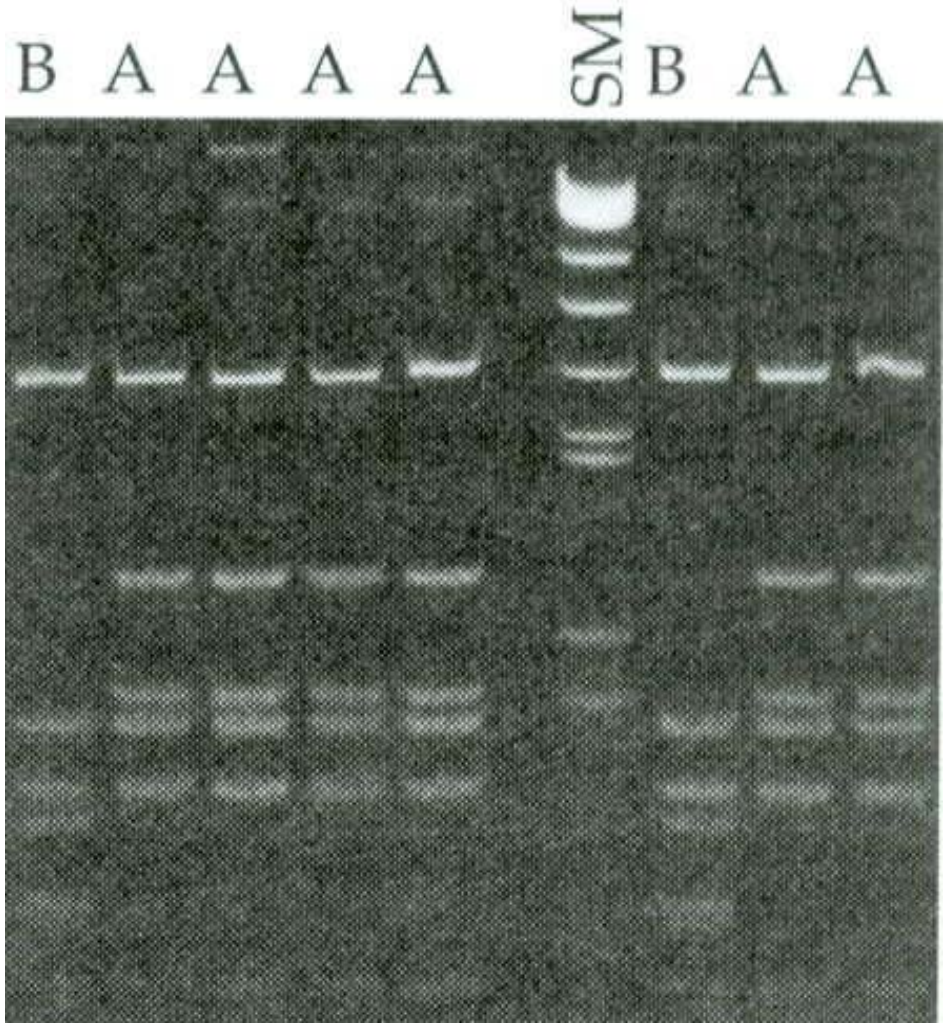
VARIABLE NUMBER OF TANDEM REPEATS



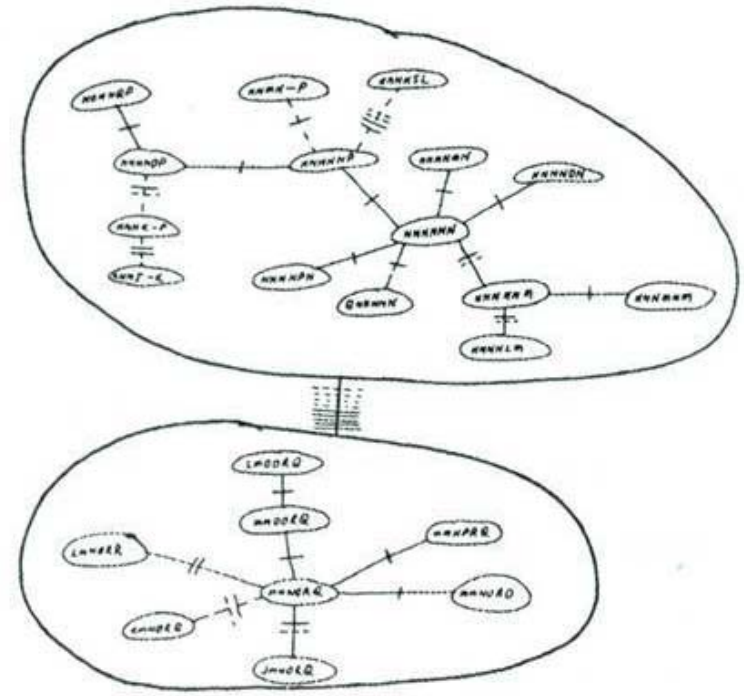


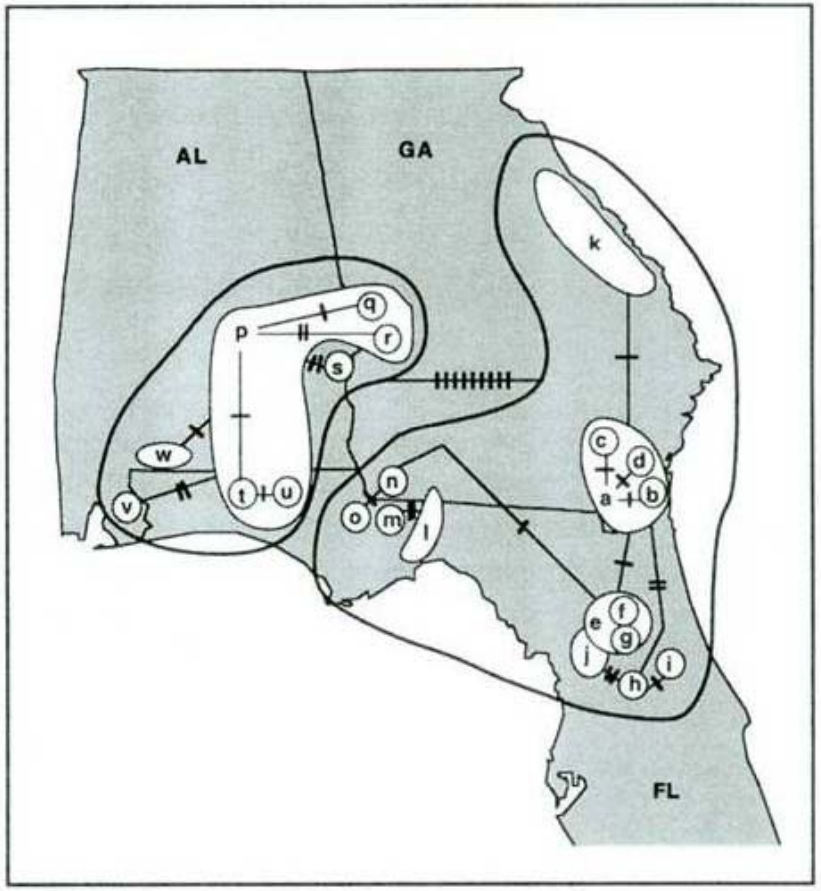
RESTRICTION FRAGMENT LENGHT POLYMORPHISM



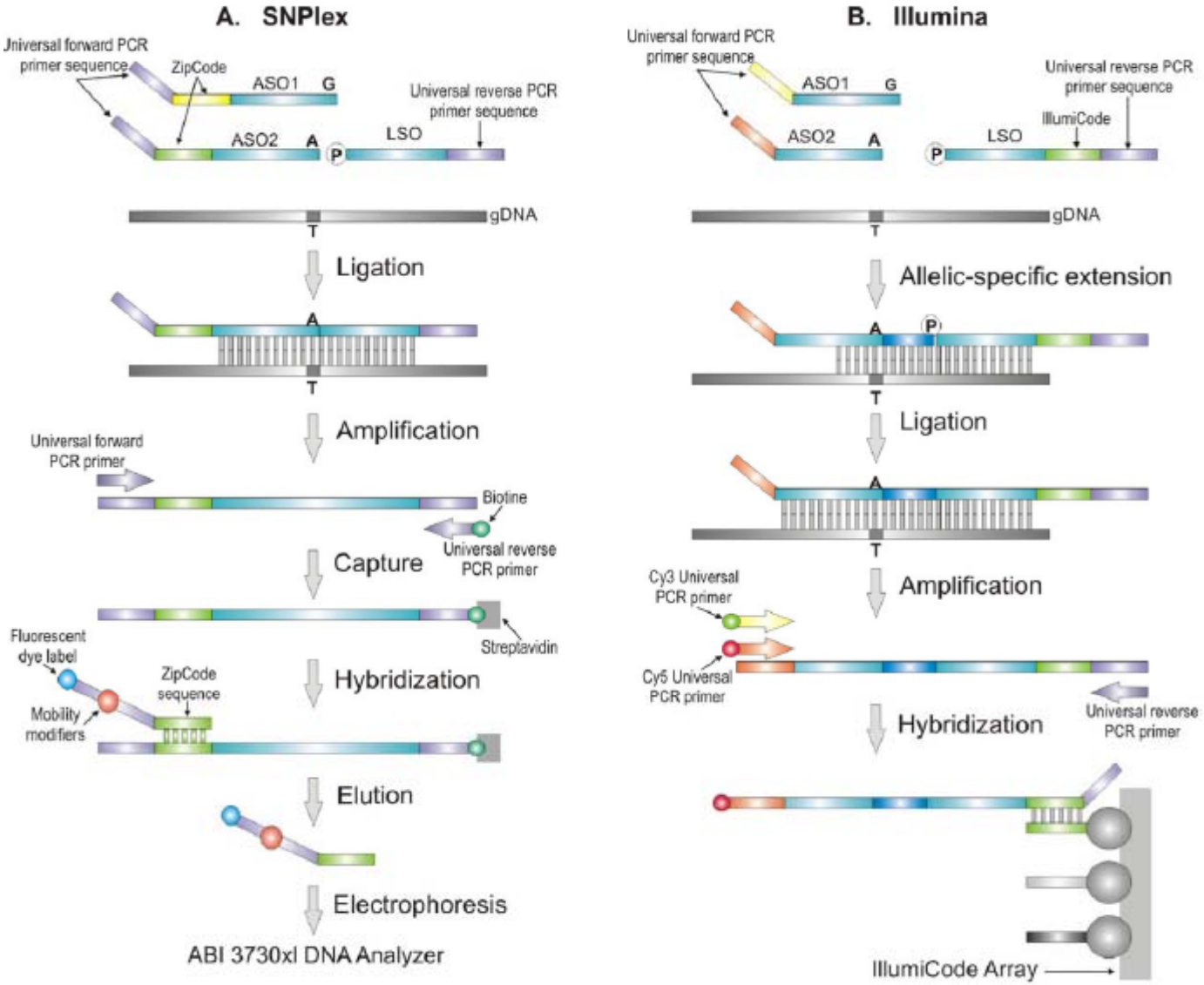


RFLP's

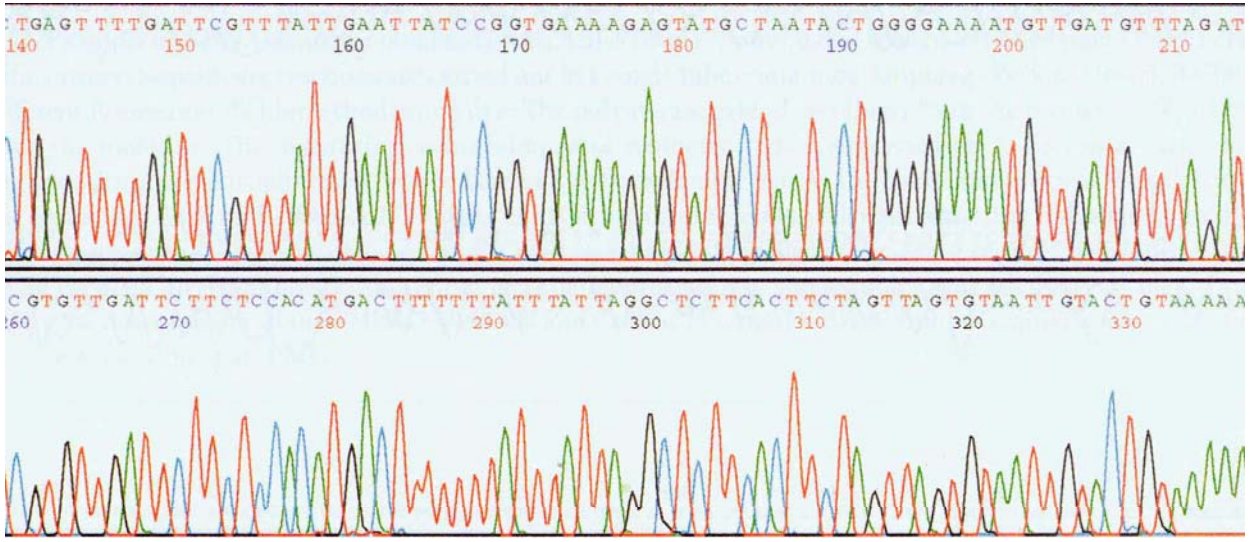




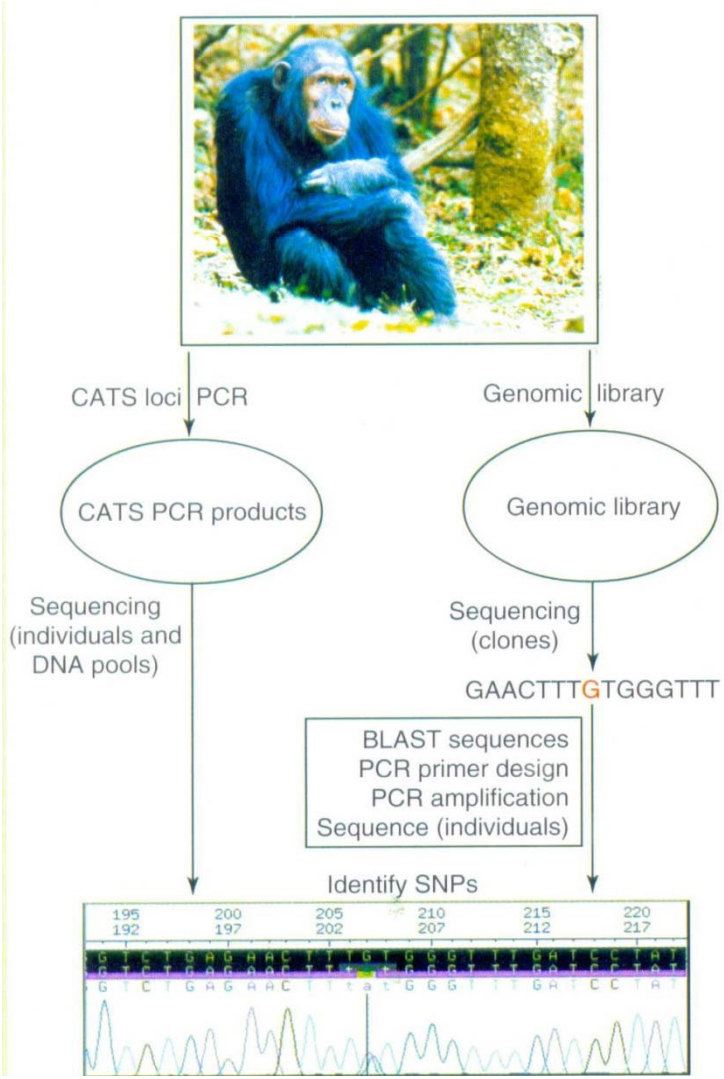
OTHER SNPs GENOTYPING PROCEDURES



SEQUENCING

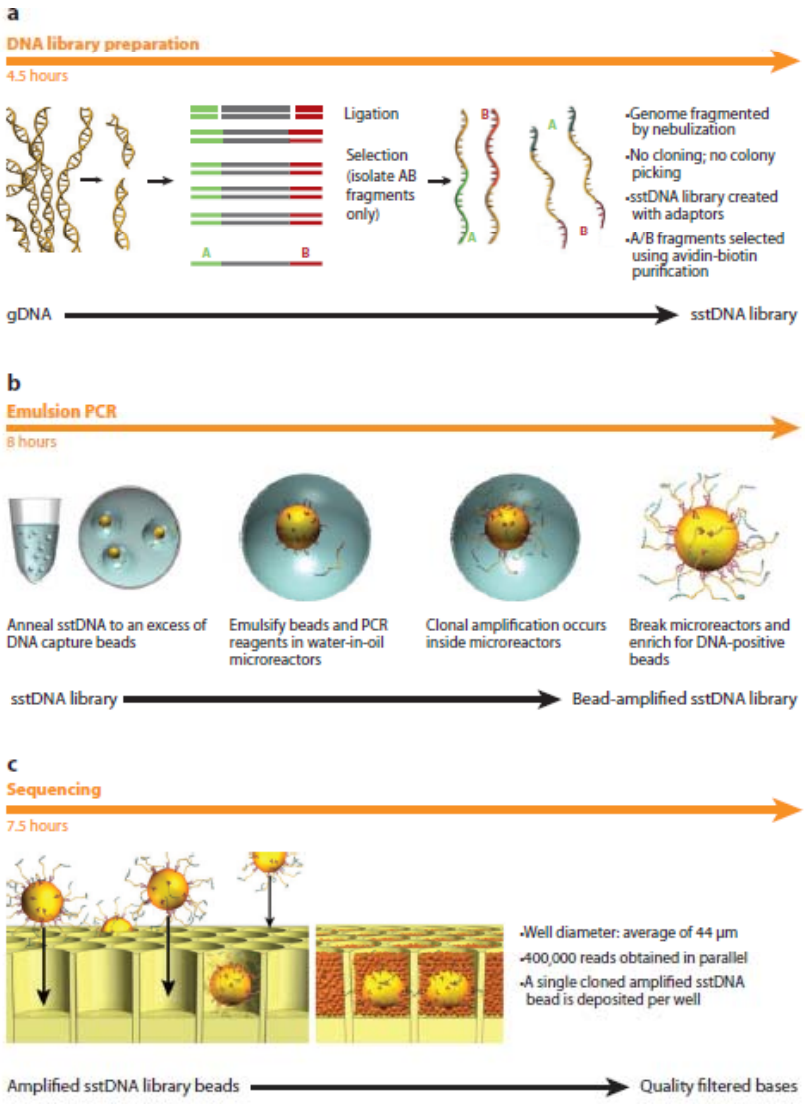

TGAGT TTTGAT TCGTT TATT GAAT TATC CGGT GA AA AGAGTATGC TAATA CT GGGGAAAA AT GT TGAT GT TTAG AT
140 150 160 170 180 190 200 210
C GTGTT GAT TC TTCTCCAC AT GAC TTTTTTTATT TAT TAGGC TC TTCACT TCTA GT TA GTGTAA TT GTAC T GTAA AA A
260 270 280 290 300 310 320 330

SEQUENCING

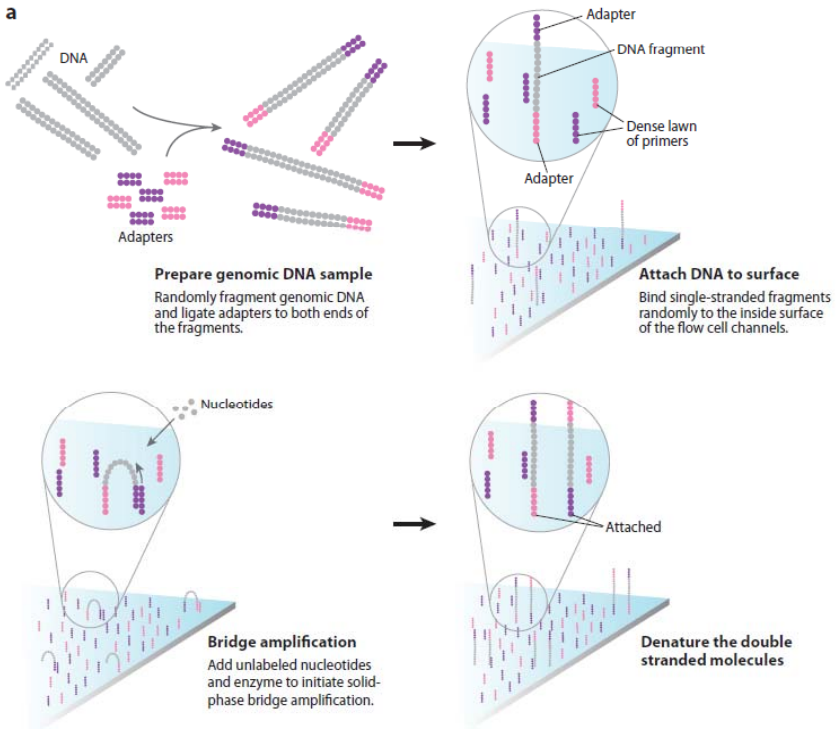

CATS loci PCR
Genomic library
CATS PCR products
Genomic library
Sequencing (individuals and DNA pools)
Sequencing (clones)
GAACTTTGTGGGTTT
BLAST sequences
PCR primer design
PCR amplification
Sequence (individuals)
Identify SNPs
195 200 205 210 215 220
192 197 202 207 212 217

HIGH THROUGHPUT SEQUENCING

454

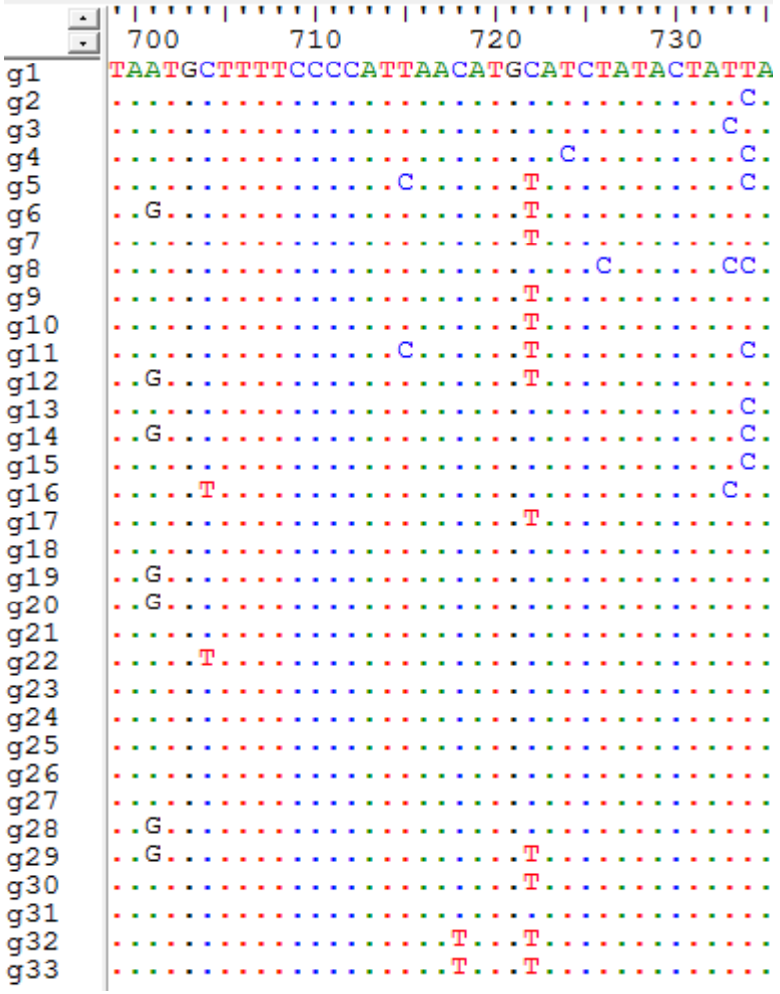


Illumina

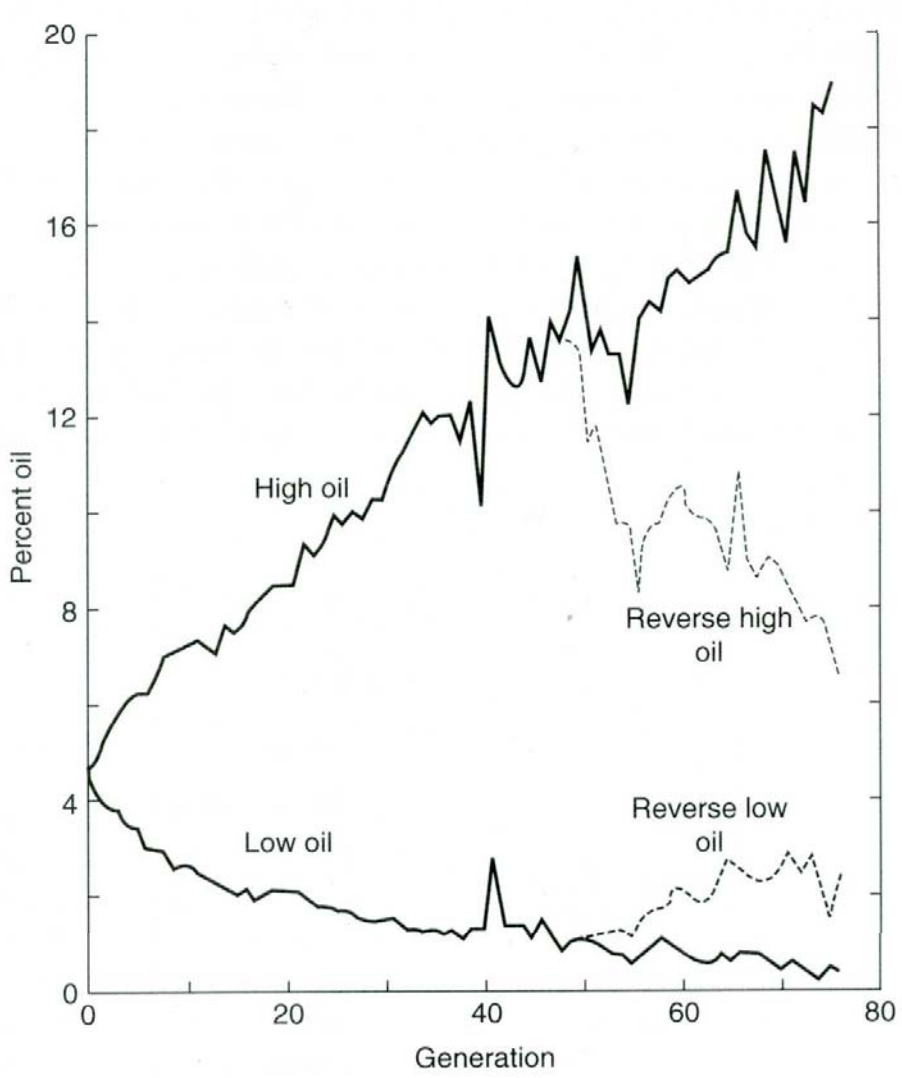


ALLELIC *VS.* SEQUENCE DATA

	CI6	CI39	CL5	CI19	CL136
215	147147	170170	116116	176180	182190
216	147147	170170	116116	172180	178186
217	147147	170170	116116	176188	190190
218	147147	170170	116116	160176	186186
219	147147	170170	116116	172172	190190
220	147147	170170	116116	176180	182190
221	147147	170170	116116	176176	186194
222	147147	170170	116116	176180	182190
502	147147	170170	116116	172176	186190
503	147147	170170	116116	176180	186186
504	147147	170170	116116	180180	182186
505	147147	0	116116	0	182186
506	147147	170170	116116	172176	178186
507	147147	170170	116116	176176	182186
508	147147	0	116116	0	190190
509	147147	170170	116116	176180	182190
510	147147	170170	116116	172180	178182
511	147147	170170	116116	172176	190194
512	147147	170170	116116	176176	186190
513	147147	170170	116116	176180	186190
514	147147	170170	116116	172180	186186
515	147147	170170	116116	172176	186190
516	147147	0	116116	0	186186
517	147147	170170	116116	176176	194194
518	147147	0	116116	0	186190
107	147147	170170	116116	176176	178182
201	147147	170170	116116	176176	178182
202	147147	170170	116116	176184	190190



QUANTITATIVE TRAIT LOCI

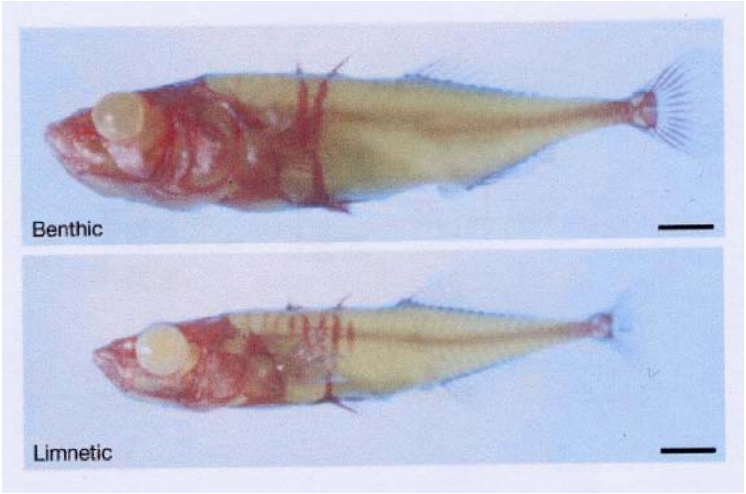


Corn seeds

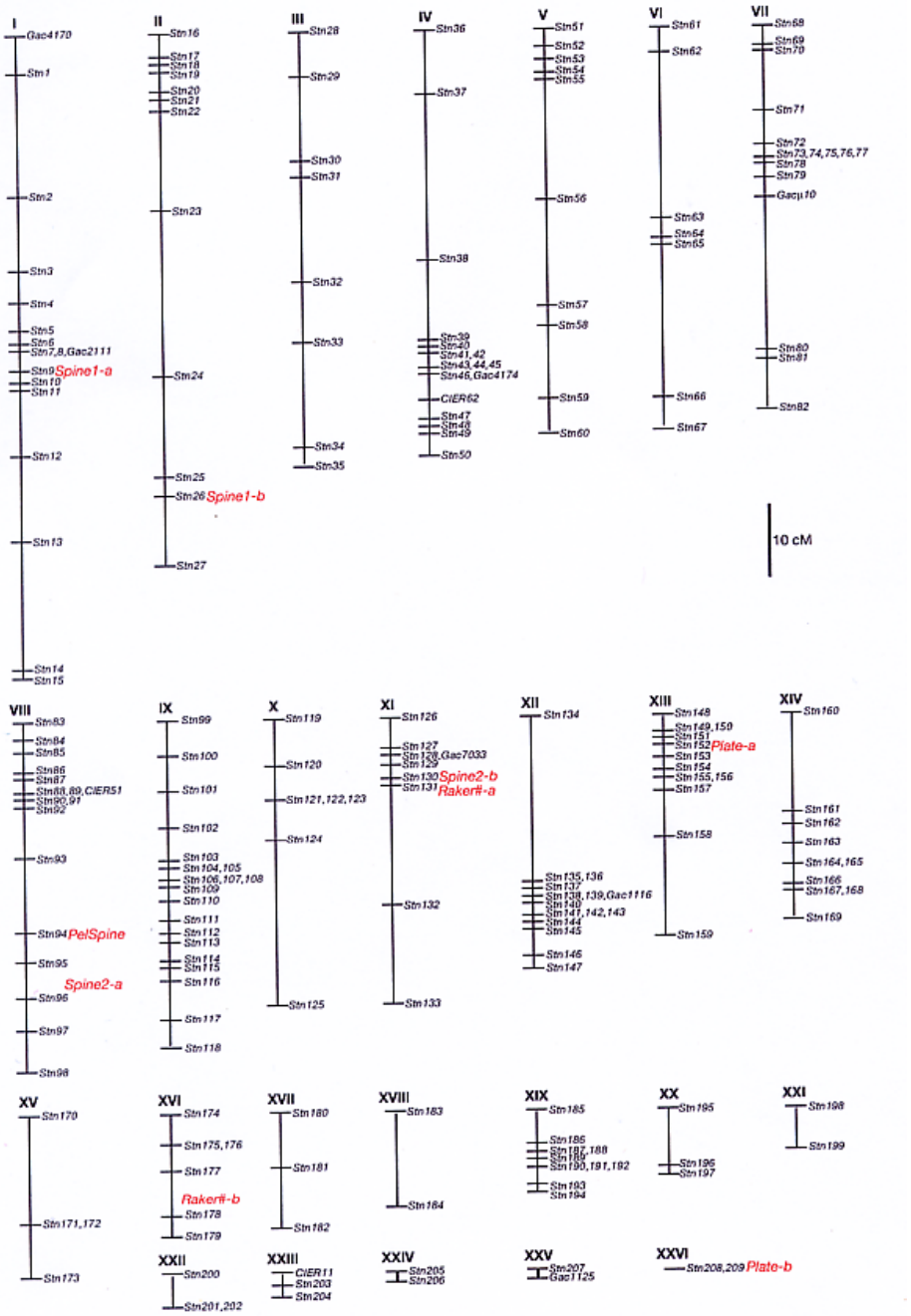
University of Illinois

1896-present)

fied. Members of a species pair are adapted to different niches within a lake, with corresponding changes in feeding morphology and defensive armour occurring in parallel in the different lakes (Fig. 1)[2]. The benthic species feeds on invertebrates near shore and has a great reduction in the amount of body armour, increased body depth, and a decreased number of gill rakers for filtering ingested food. The limnetic species more closely resembles an ancestral marine fish, with more extensive body armour, a longer and more streamlined body, and an increased number of gill rakers. Despite reproductive isolation between the two species in the wild[3–6], it is possible to establish productive matings between the two species under laboratory conditions[2]. The resulting F_1 hybrids are viable and fertile, making it possible to carry out a formal genetic analysis of the number and location of loci responsible for the adaptive morphological differences between these naturally occurring vertebrate species.



Gasterosteus aculeatus



QUANTITATIVE TRAIT LOCI