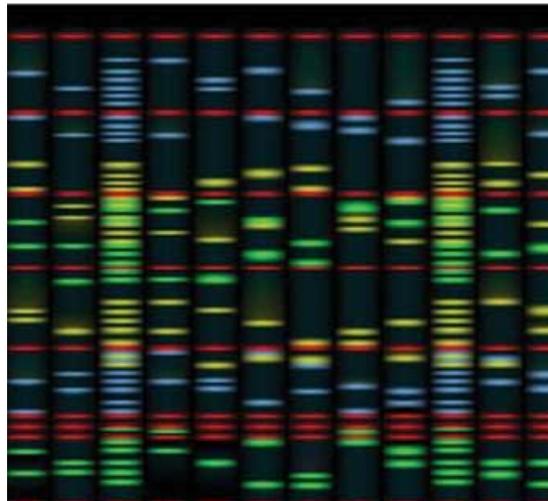
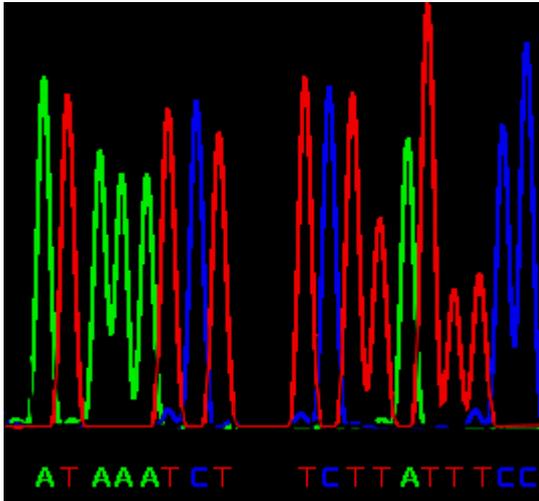


DNA Polymorphism



Polymorphism

- The occurrence in a population of two or more genetically determined forms, each with a frequency greater than 1%.

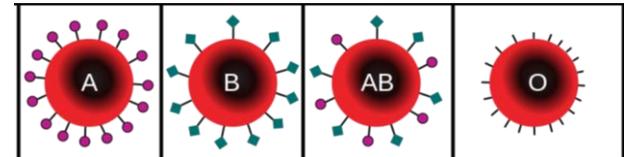
- Polymorphism presents at different level:

-Phenotype polymorphism



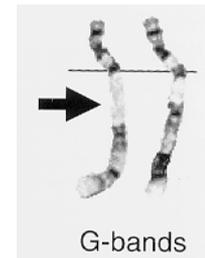
-Protein polymorphism

e.g. ABO blood type or immunoglobulin



-Chromosomal polymorphism

e.g. minor variant in chromosomal structure



-DNA sequence polymorphism



DNA Polymorphism

- **Definition** : more than one normal allele at a genomic locus in the population, with the rarest allele having frequency of $>1\%$.
- **Characteristics:**
 - The frequency of the rarest allele is more than 1%
 - Inherited in Mendelian pattern in the families
 - No functional consequences

Major types of DNA Polymorphism Markers

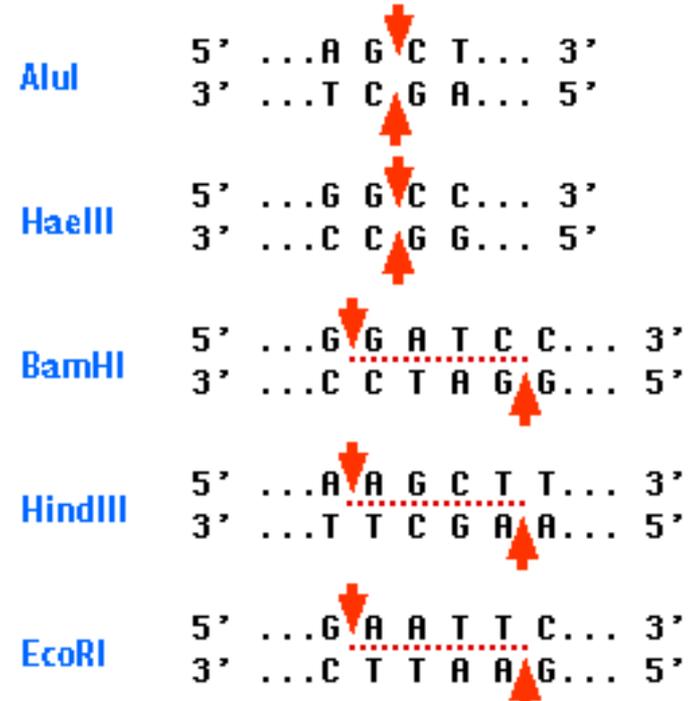
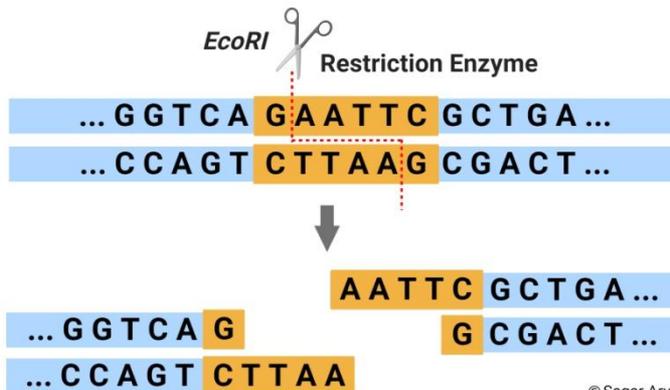
- Restriction fragment length polymorphism (RFLP)
- Randomly amplified polymorphic DNA (RAPD)
- Short tandem repeat polymorphism (STR)
 - Mini-satellite repeat polymorphism
 - Micro-satellite repeat polymorphism
- Single nucleotide polymorphism(SNP)

Restriction Fragment Length Polymorphism

- **RFLP: Polymorphism that alter the length of restriction fragments**
- **Result from changes that introduce or delete a restriction enzyme site**
- **Two alleles**
- **Genotyping by Southern or PCR-RFLP**

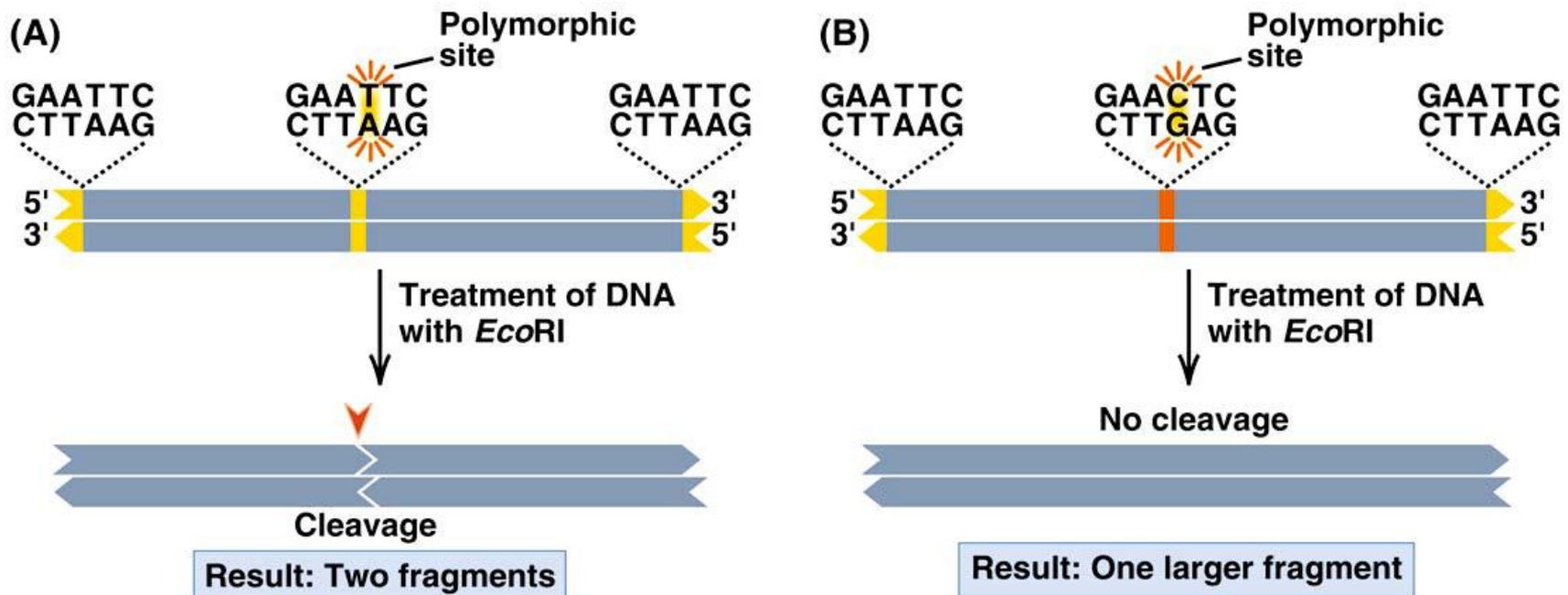
Restriction Endonucleases

- Restriction endonucleases are bacterial enzymes that recognize specific 4, 5, 6, or 8 base pair (bp) nucleotide sequences and cut DNA wherever these sequences are encountered, so that changes in the DNA sequence due to *indels*, *base substitutions*, or *rearrangements* involving the restriction sites can result in the gain, loss, or relocation of a restriction site

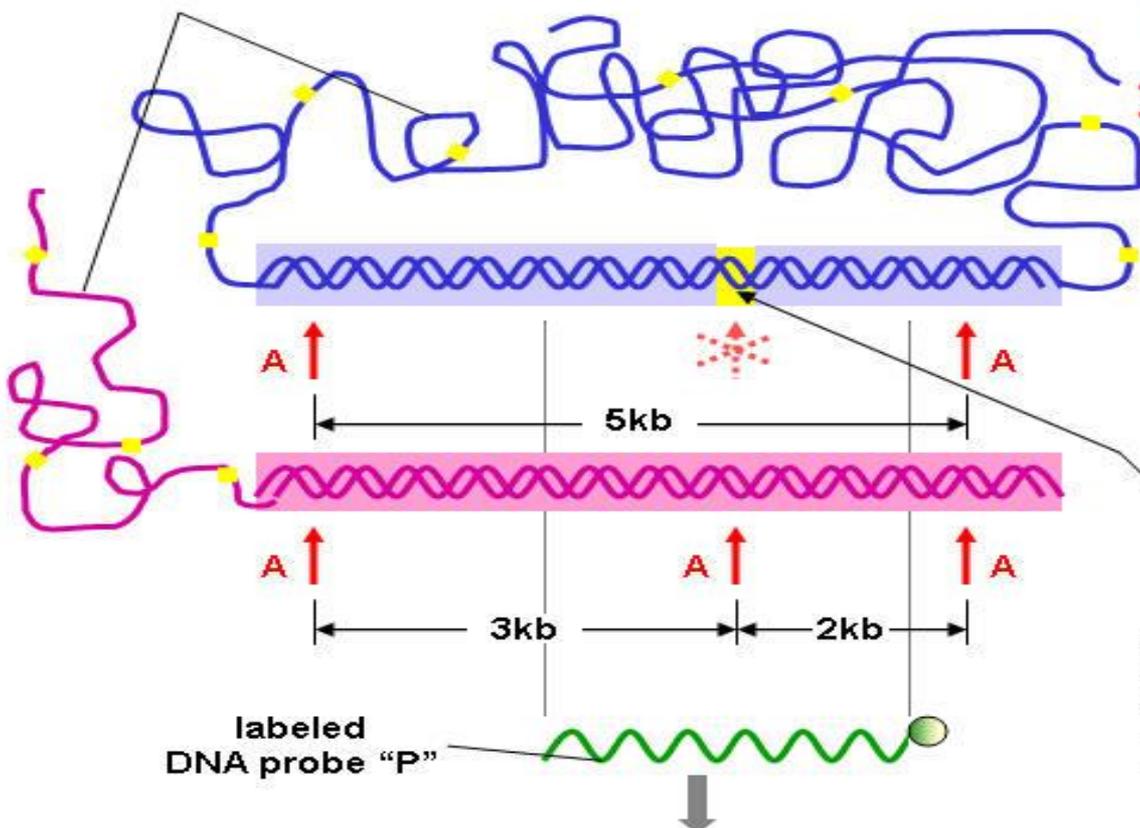


Restriction Fragment Length Polymorphism (RFLP)

RFLP is a difference in homologous DNA sequences that can be detected by the presence of fragments of different lengths after digestion of the DNA samples in question with specific restriction endonucleases.



Isolated genomic DNA



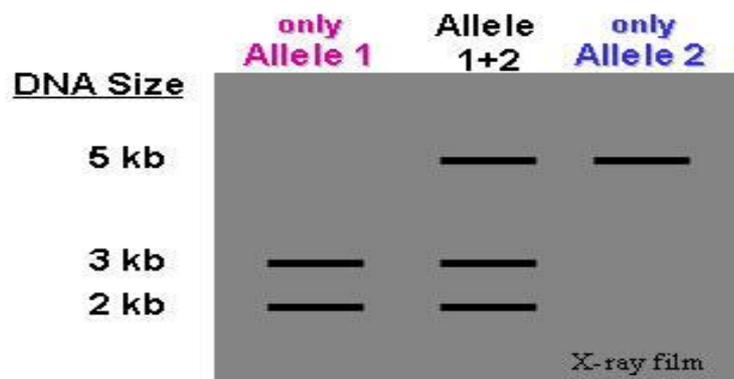
↑ = Cut sites of restriction enzyme A
 ☆ = Loss of cut sites of restriction enzyme A

Allele 1

Allele 2

Loss or addition of restriction site due to:
 - Mutation/SNP
 - Transposon insert (LINE, SINE/Alu)

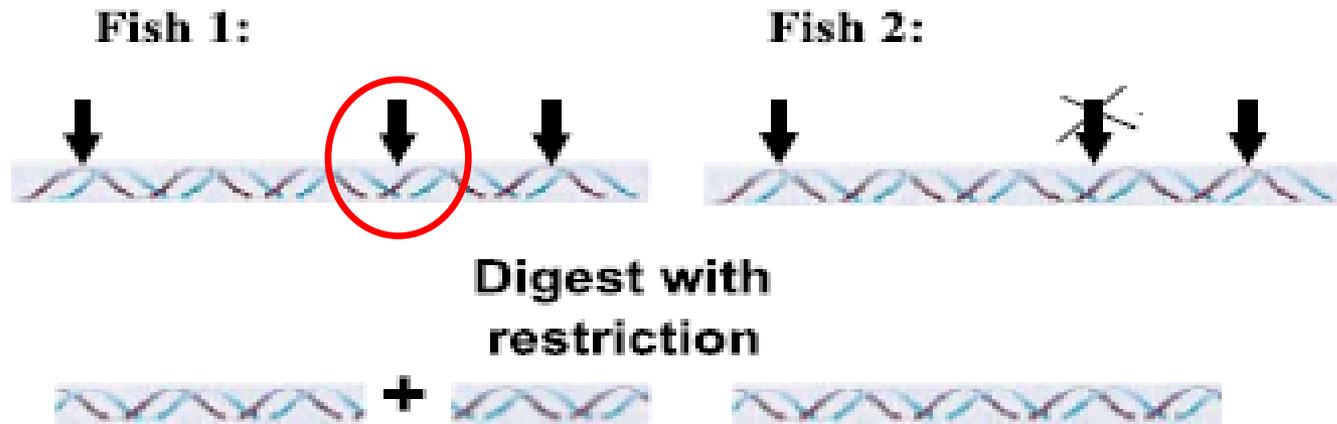
labeled DNA probe "P"



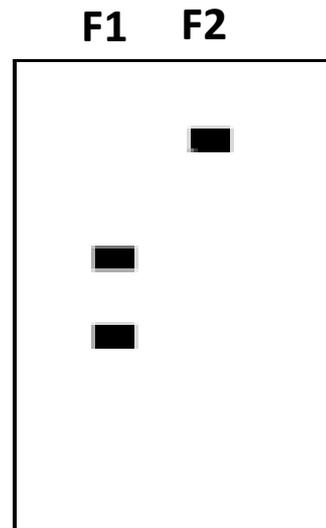
Genotypes: Homozygous, Heterozygous, Homozygous

DNA fragments after:
 1. Restriction digest
 2. Gel electrophoresis &
 3. Southern blotting (using DNA probe "P")

A. Base substitutions at the restriction sites



Gel electrophoresis and Southern blot

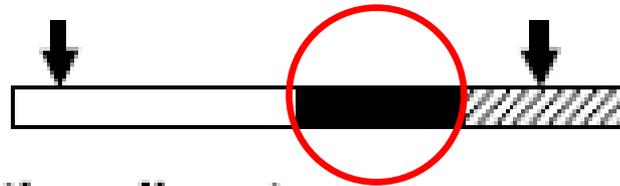


B. Insertions or deletions

Fish 1:



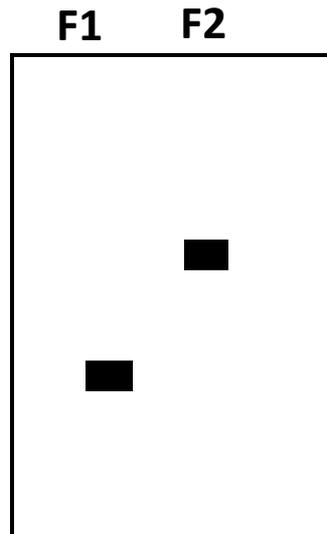
Fish 2:



Restriction digest



Gel electrophoresis and Southern blot



Randomly Amplified Polymorphic DNA (RAPD)

RAPD is a random amplification of anonymous loci by PCR.

The technique is based on the PCR amplification of discrete regions of genome with single short oligonucleotide primers (8-10 bp length) of arbitrary sequence.

Performed at relatively low annealing temperatures (often 36–40°C).

Each product (presumably) representing a different locus.

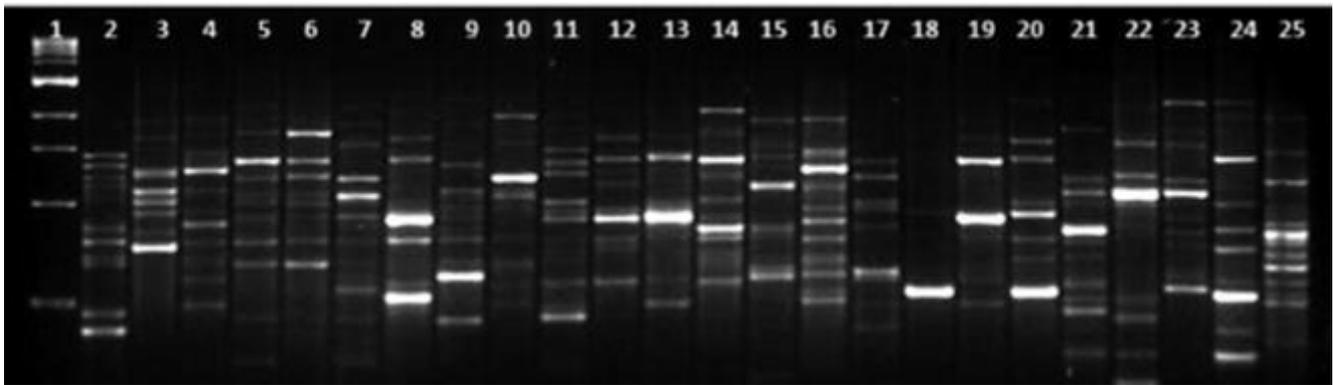
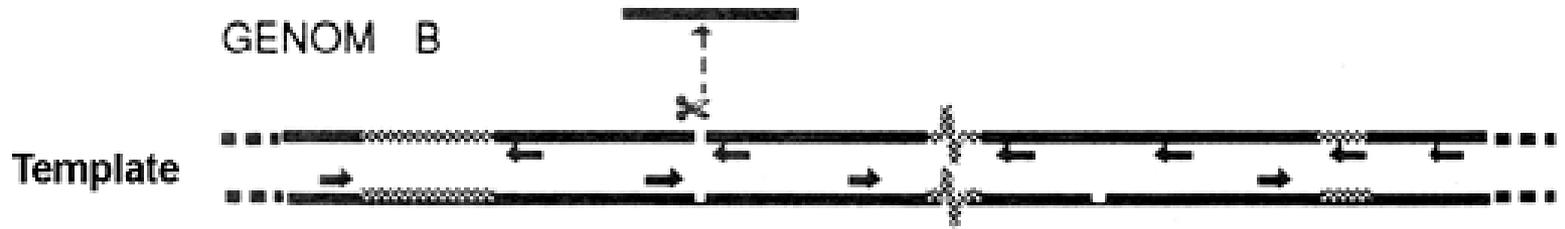
Genetic variation and divergence within and between the taxa of interest are **assessed by the presence or absence of each product**, which is dictated by changes in the DNA sequence at each locus.

RAPD polymorphisms can occur **due to base substitutions at the primer binding sites or to indels in the regions between the sites.**

GENOM A



GENOM B



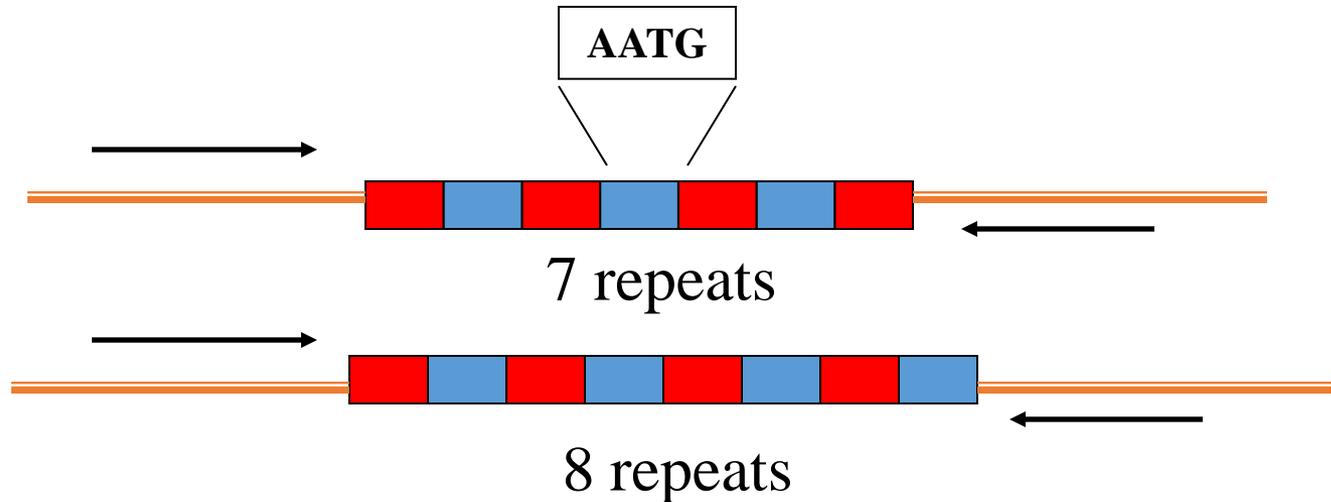
Advantage

- Simple
- Rapid and Cheap
- High polymorphism
- Only a small amount of DNA is required
- No need for molecular hybridization
- No prior knowledge of the genetic make-up of the organism

Disadvantage

- Poor reproducibility-resulting pattern of bands is very sensitive to variations in reaction conditions, DNA quality, and the PCR temperature profile
- Dominant marker-homozygous and heterozygous states cannot be differentiated
- Less sensitive

Short Tandem Repeats (STRs)/microsatellite



the repeat region is variable between samples while the flanking regions where PCR primers bind are constant

Homozygote = both alleles are the same length

Heterozygote = alleles differ and can be resolved from one another

- **Also called as STR, SSR, VNTR**
- **Tandemly repeated DNA sequences with the repeat/size of 1 – 6 bases repeated several times**
- **Highly polymorphic; can be analysed with the help of PCR**
- **Individual alleles at a locus differ in number of tandem repeats of unit sequence owing to gain or loss of one or more repeats and they can be differentiated by electrophoresis according to their size**
- **Powerful DNA markers for quantifying genetic variations within & between populations of a species**

Short Tandem Repeats (STR)

➤ Variable number of tandem repeats

➤ Multiple alleles

CACACACACACACA.... (CA)⁷

....CACACACACACACACACA.... (CA)¹⁰

....CACACACACACACACACACA.... (CA)¹²

....CACACACACACACACACACACA.... (CA)¹⁴

Short Tandem Repeats Polymorphism (STR)

100bpCACACACACA.....100bp

Allele 1 (CA)⁵ 210bp

Allele 2 (CA)⁶ 212bp

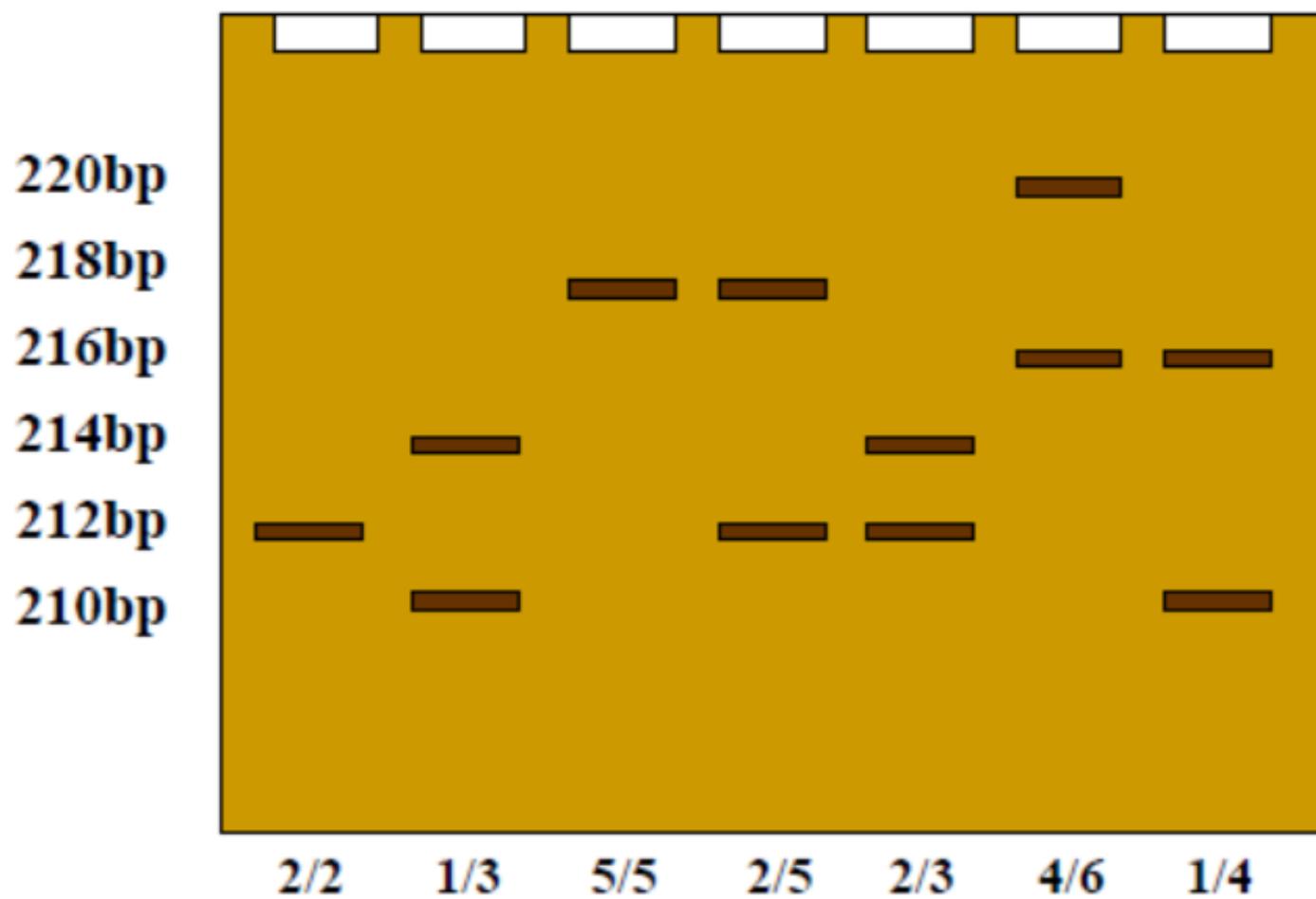
Allele 3 (CA)⁷ 214bp

Allele 4 (CA)⁸ 216bp

Allele 5 (CA)⁹ 218bp

Allele 6 (CA)¹⁰ 220bp

Result:



Microsatellite

- Microsatellites (1 to 10 nucleotides) and minisatellites (>10 nucleotides) are subcategories of short tandem repeats (STRs) that make up genomic repetitive regions.
- Repeat polymorphisms usually result from the addition or deletion of the entire repeat units or motifs. Therefore, different individuals exhibit variations as differences in repeat numbers.

Microsatellites – Types

Based on repeat pattern

1. Perfect – 
2. Imperfect – 
3. Compound – 
4. Complex – 

Based on number of base pairs

- 1) Mono (e.g. CCCCCCCC or AAAAAA)
- 2) Di (e.g. CACACACACA)
- 3) Tri (e.g. CCA CCA CCA CCA)
- 4) Tetra (e.g. GATA GATA GATA GATA GATA GATA GATA)

Minisatellites: - (9 – 65 base pairs repeated from 2 to several hundred times)

———— CGCCATTGTAGCCAATCCGGGTGCGATTGCAT CGCCATTGT
AGCCAATCCGGGTGCGATTGCAT CGCCATTGTAGCCAATCCGGG
TGCATTGCAT CGCCATTGTAGCCAATCCGGGTGCGATTGCAT
CGCCATTGTAGCCAATCCGGGTGCGATTGCAT ———

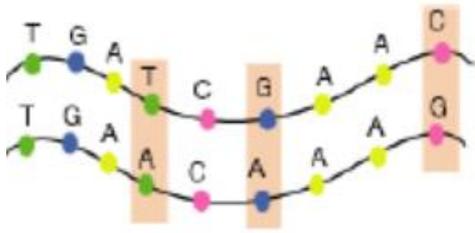
Microsatellites – Properties

- ✓ Co-dominant
- ✓ Polymorphic loci with allele number as high as 14 – 15 per locus
- ✓ Mostly reported from non-coding region, hence can be independent of selection
- ✓ Flanking region is highly conserved in related species
- ✓ Can be obtained from small amounts of tissues [STR analysis can be done on less than one billionth of a gram (a nanogram) of DNA (as in a single flake of dandruff)]
- ✓ PAGE separation; silver staining/automated genotyping
- ✓ Abundant in the eukaryote genome ($\sim 10^3$ to 10^5 loci dispersed at 7 to 10^{100} kilobase pair (kb) intervals)

➤ The **microsatellite**, or short sequence repeat (SSR), is a powerful genetic marker, useful in many areas of fish genetics and breeding.

➤ Polymorphic microsatellite loci have been frequently applied to the **analysis of genetic diversity, population genetic structure, and genomic mapping.**

➤ These co-dominant markers have also been applied to the **classification and systematics, parentage identification, germplasm conservation, and breeding programme of food fish.**



SNP

(single nucleotide polymorphism)



❖ Definition:

- ❖ the variation of only one base pair on one polymorphic locus
- *the most simple form*

Single nucleotide polymorphism (SNP) describes polymorphisms caused by point mutations that give rise to different alleles containing alternative bases at a given nucleotide position within a locus.

They are the most abundant polymorphism in any organism, adaptable to automation, and reveal hidden polymorphism not detected with other markers and methods.

SNP markers are inherited as co-dominant markers.

SNPs

Variations in DNA at a single base that are found in at least in 1% of the population

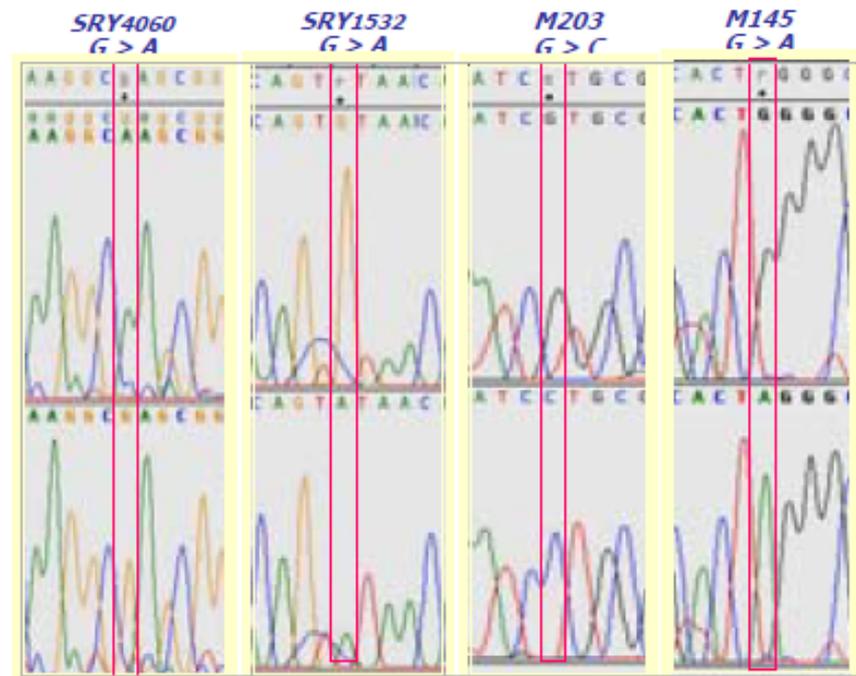
- Stable inheritance
- ~ 1/ 0.5 kb throughout the genome
- Bi-allelic

Expected SNPs in human genome = 2.8 million
Functionally interesting SNPs = 50,000 to 250,000

..great potential in Biomedical research, Disease diagnosis..

Can be detected by:

- Sequencing



Ideal DNA markers

- ❑ Be polymorphic and evenly distributed throughout the genome
- ❑ Provide adequate resolution of genetic differences
- ❑ Generate multiple, independent and reliable markers
- ❑ Repeatable between assays, people, machines and labs
- ❑ Low errors
- ❑ Be simple, quick and inexpensive
- ❑ Need small amounts of tissue and DNA samples
- ❑ Link to distinct phenotypes
- ❑ Require no prior information about the genome of an organism

No molecular marker presents all the listed advantages

Choice of marker systems

One of the questions at the beginning of any genome research is what type of marker is most suitable given the project at hand and the species of interest. There is no simple answer to this question, and much depends on the specific objectives of the study. However, with a good understanding of the DNA marker technologies, appropriate decisions can be reached.

Applications of DNA markers in aquaculture genetics

Tasks	Recommended marker system	Other useful marker types
Species identification	RAPD/mtDNA (Barcoding)	AFLP, microsatellites, isozymes
Strain identification	AFLP, microsatellites	RAPD
Hybrid identification	RAPD	AFLP, microsatellites, mitochondria ^a
Paternity determination	Microsatellites	
Genetic resource/diversity analysis	AFLP, microsatellites	RAPD
Genetic mapping	Type I markers, Microsatellites SNP	AFLP, RFLP
Comparative mapping	Type I markers	ESTs, conserved microsatellites

^a Use of mitochondrial markers should also allow determination of maternity.