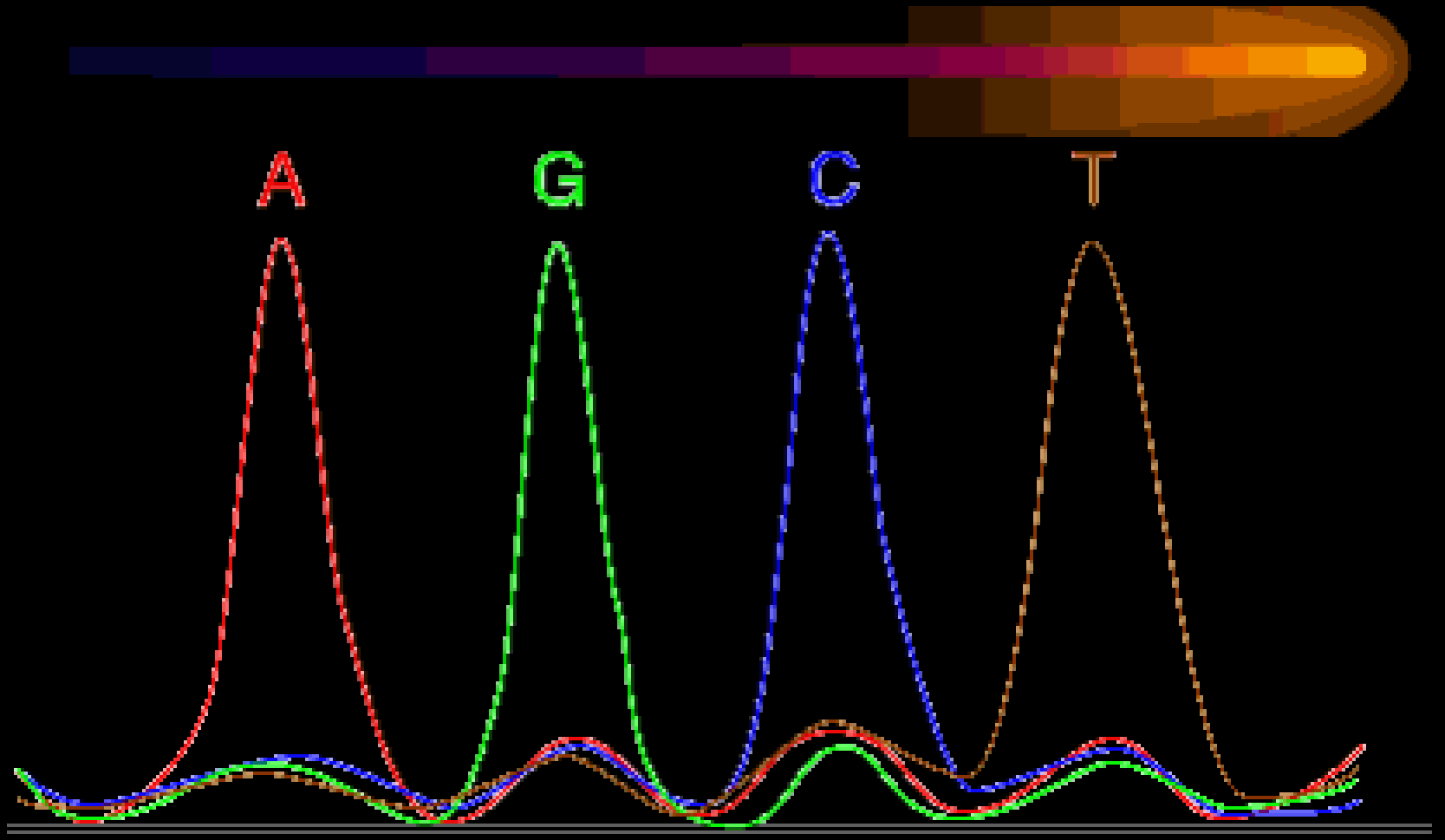
BENG 183

Trey Ideker

Genotyping



To be covered in one 1.5 hr lecture

Genetic variation: Some basic definitions

- **Allele**
  Alternative form of a genetic locus inherited separately from each parent
- **Polymorphism**
  A locus with two or more alleles where each allele occurs at a frequency > 1% in the population
- **Mutation**
  Any heritable change in DNA
- **Genotyping**
  Determining the specific set of alleles present in an individual's genome

Types of markers for genotyping

- VNTRs
- STRs
- RFLPs
- AFLPs
- SNPs
- STSs

PHYSICAL

- Markers linked to a PHENOTYPIC trait

Interspersed repeats

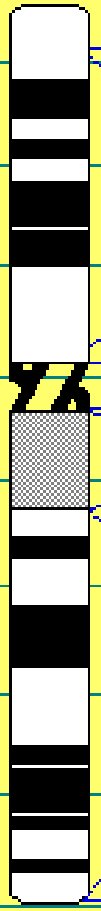
- Tandem
  - Satellites, minisatellites (VNTRs) microsatellites (STRs)

- Interspersed
  - Retrotransposons (class I)
  - SINE (Alu- 10% of human genome)
  - Transposons (class II)
  - LINE (10% of human genome)

Resources: Repbase, RepeatMasker

Chromosomal organization of repeated sequences

chr 16



subtelomeric repeats

interspersed and tandem repeats — retroposons, minisatelites, microsatellites

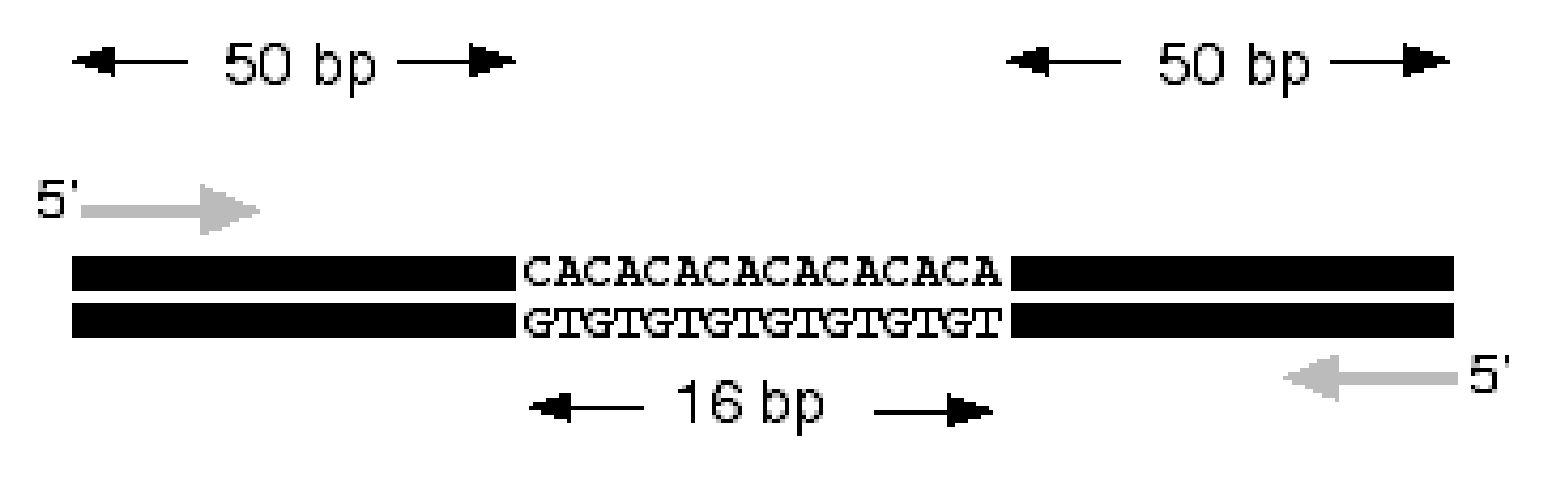
tandem repeat — alpha satellite (340 bp dimer; 1-2 Mb)

tandem repeat — satellite II (5 bp monomer; 2-7 Mb)

interspersed and tandem repeats — retroposons, minisatelites, microsatellites

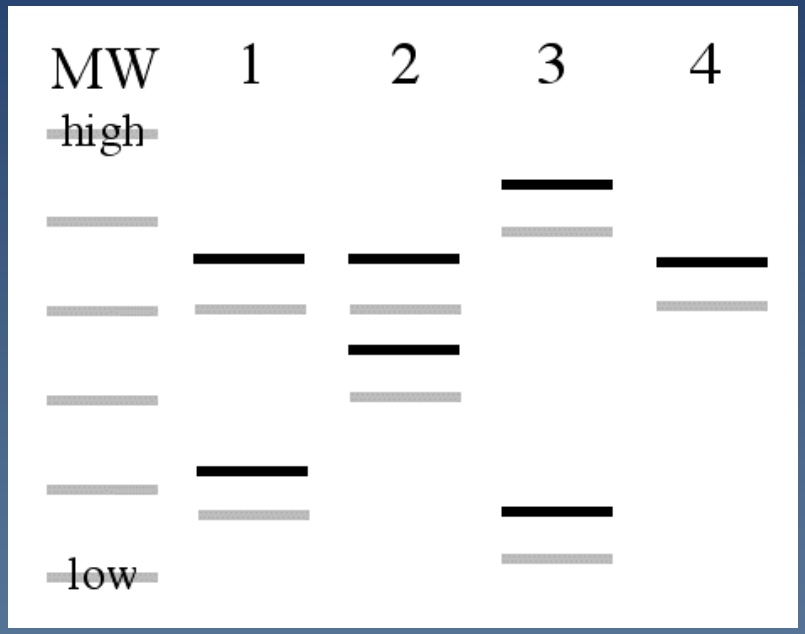
tandem repeat — telomer $(TTAGGG)_n$ 5-12 kb

Microsatellite detection (STRs)



Two PCR primers (forward and reverse gray arrows) are designed to flank the microsatellite region. If there were zero repeats, the PCR product would be 100 bp in length. Therefore, by determining the size of each PCR product (in this case 116 bp), you can calculate how many CA repeats are present in each microsatellite (8 CA repeats in this example).

Microsatellite detection (2)



Gel electrophoresis

Capillary instruments also used

Fluorescence detection

Microsatellite Questions

- 1) Why do you think microsatellites with repeating units of two nucleotides are usually located in non-coding DNA?
- 2) Why might a stutter band differ from a major band by two nucleotides instead of one?
- 3) In the previous figure, one animal only has a single band. How can this be?
- 4) Compare the two forms of data in the previous fig. Are the data identical or not?

High-throughput PCR for (μsatellite) genotyping

The 96-well x 4 tetrad

→ → →

Types of markers for genotyping

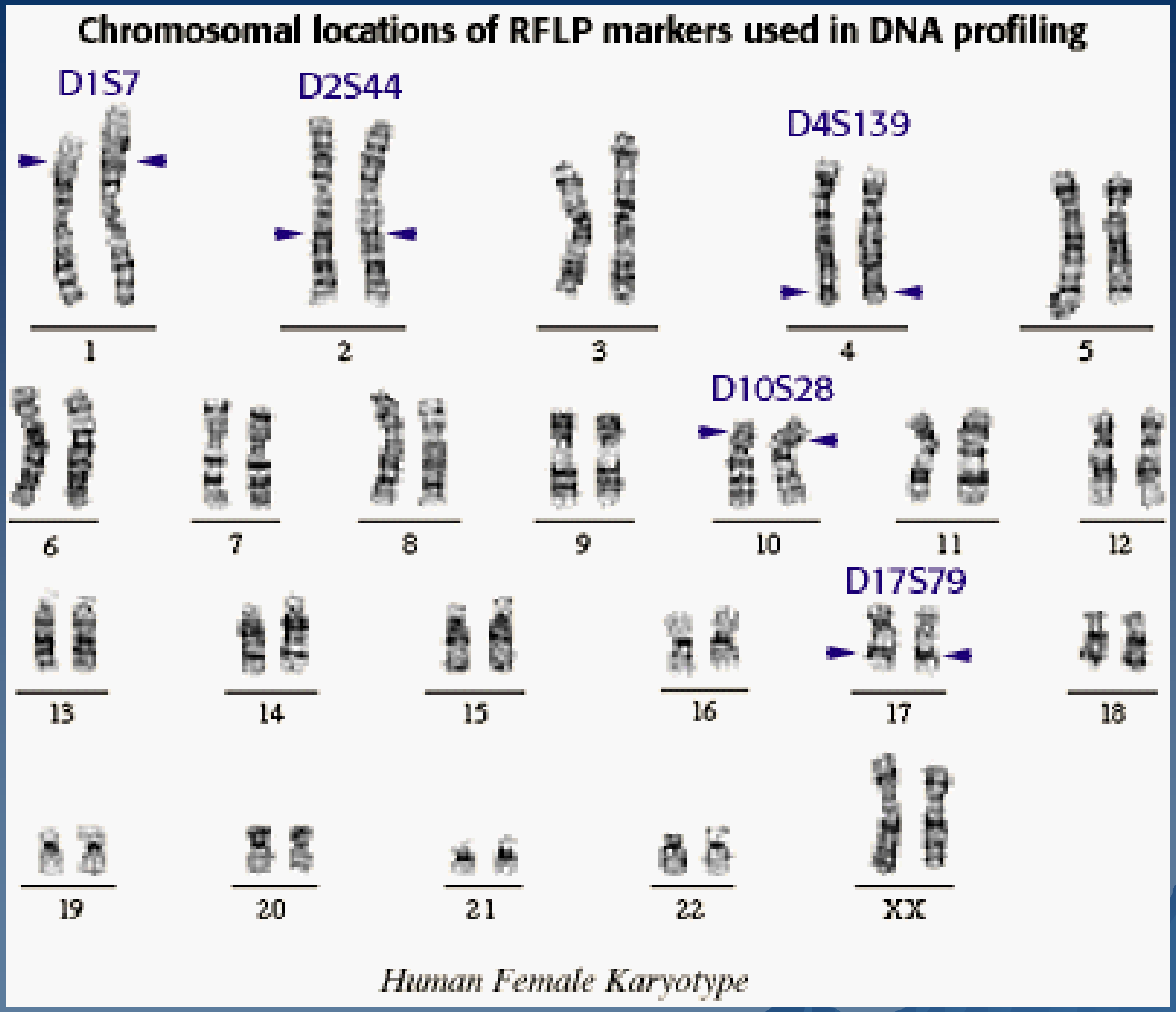
- VNTRs
- STRs
- RFLPs
- AFLPs
- SNPs
- STSs

PHYSICAL

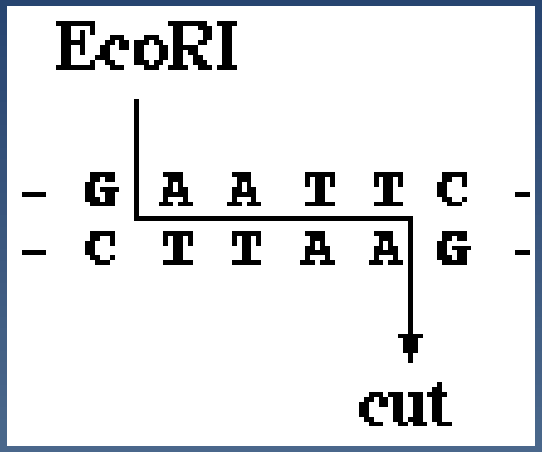
- Markers linked to a PHENOTYPIC trait

RFLPs

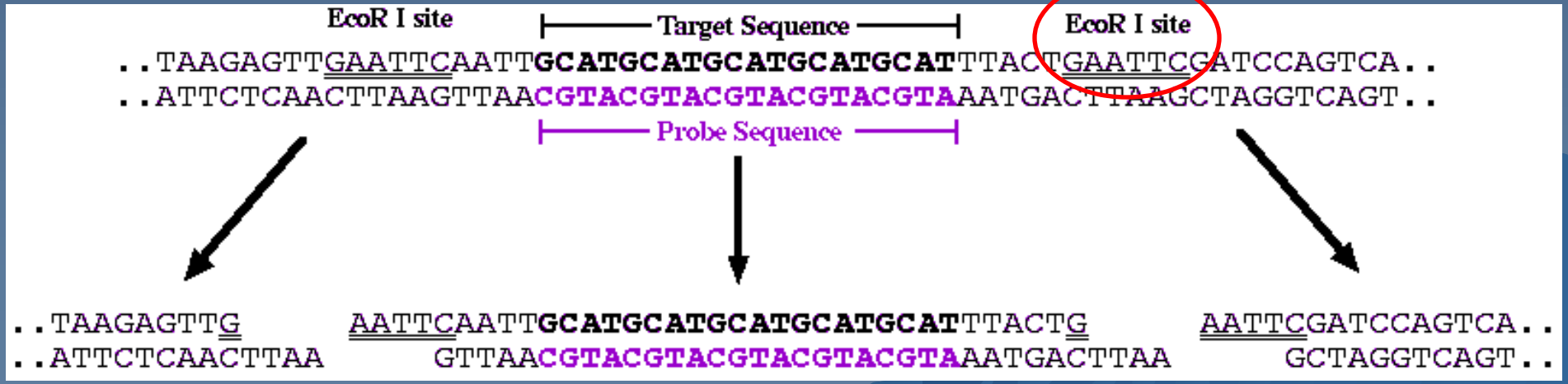
- Restriction Fragment Length Polymorphism
- Is a sequence of DNA that has a restriction site on each end with a "**target**" sequence in between.
- An investigator uses a labeled probe to identify this target sequence in a Southern blot
- The presence/absence of a restriction site will correlate with the size of the target

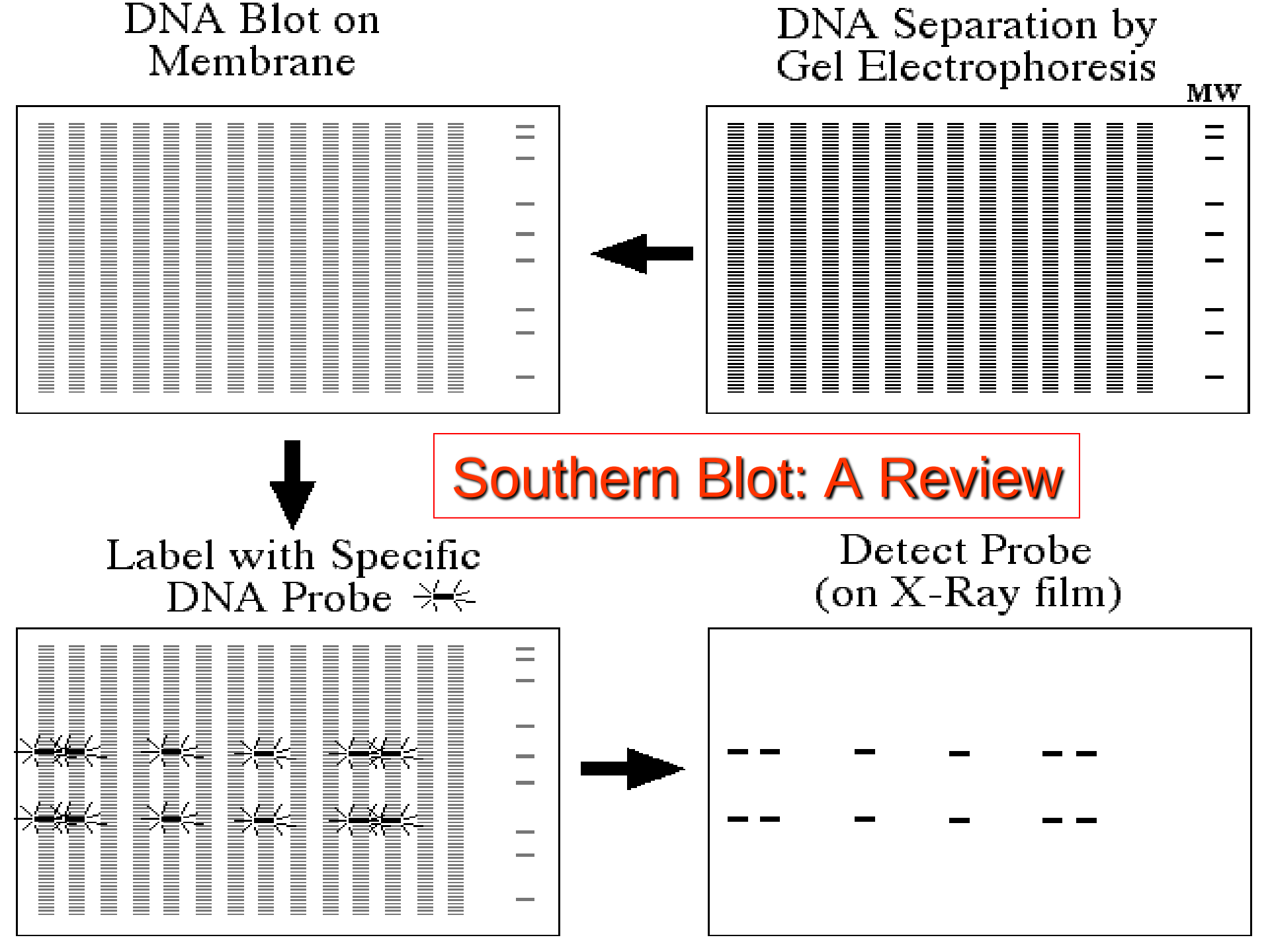

Chromosomal locations of RFLP markers used in DNA profiling
D1S7
D2S44
D4S139
D10S28
D17S79
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
XX
Human Female Karyotype

RFLPs (2)



Present/Absent



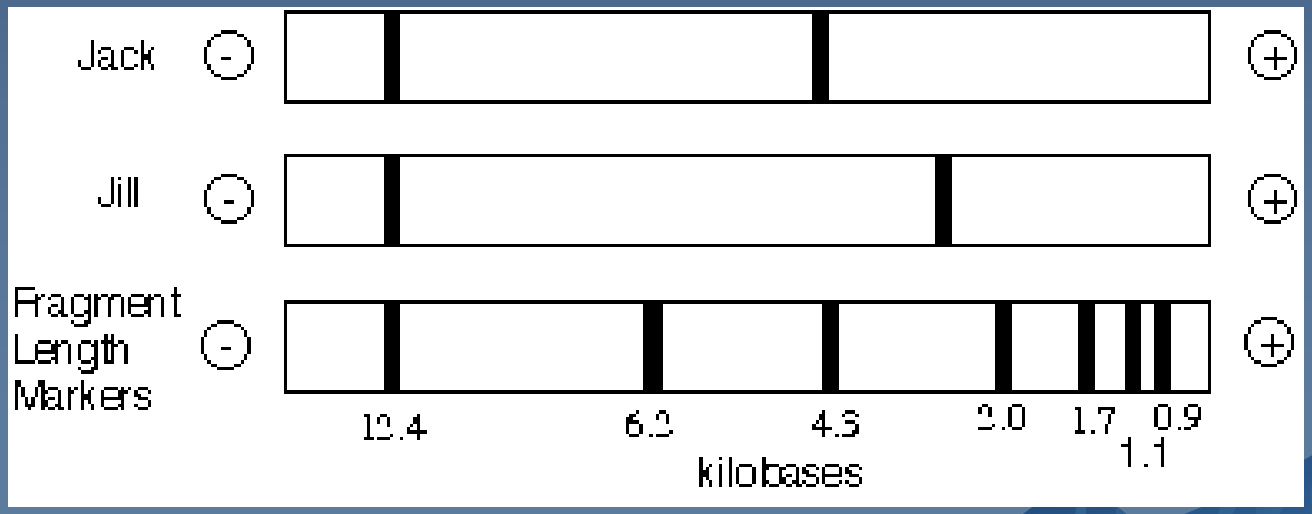

DNA Blot on Membrane
DNA Separation by Gel Electrophoresis
MW
Southern Blot: A Review
Label with Specific DNA Probe
Detect Probe (on X-Ray film)

RFLPs (3)

```
IDENTICAL RFLPs
Jack 1: -GAATTC---(8.2 kb)---GCATGCATGCATGCATGCAT---(4.2 kb)---GAATTC-
Jill 1: -GAATTC---(8.2 kb)---GCATGCATGCATGCATGCAT---(4.2 kb)---GAATTC-

DIFFERENT RFLPs
Jack 2: -GAATTC--(1.8 kb)-GAATTT--(1.2 kb)--GCATGCATGCATGCATGCAT--(1.3 kb)-GAATTC-
Jill 2: -GAATTC--(1.8 kb)-GAATTC--(1.2 kb)--GCATGCATGCATGCATGCAT--(1.3 kb)-GAATTC-
```

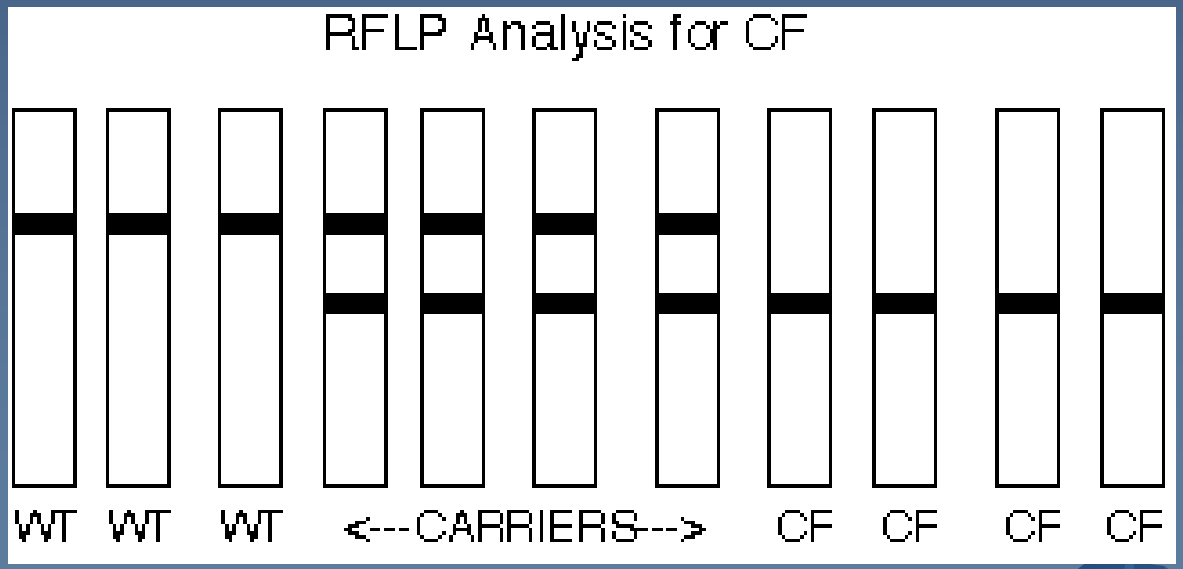
A Southern Blot

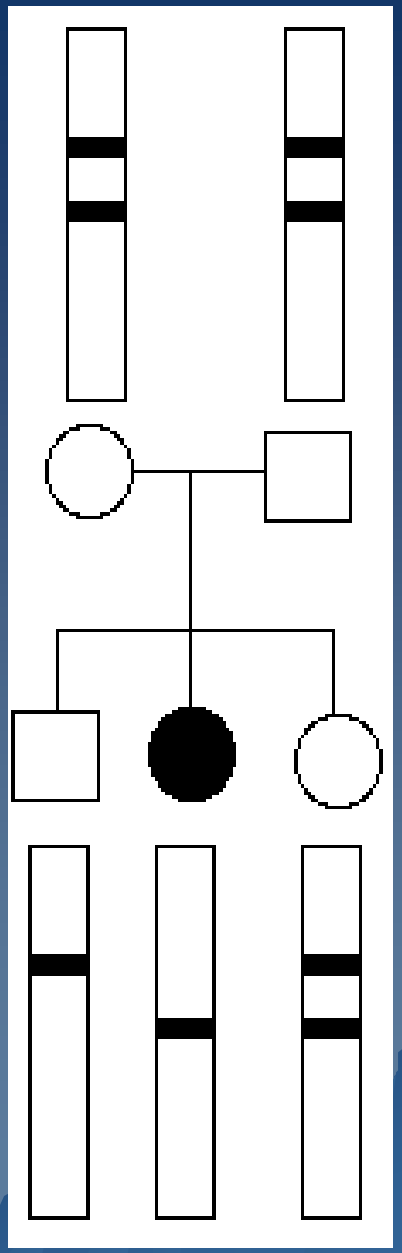


An example: Cystic Fibrosis (CF)

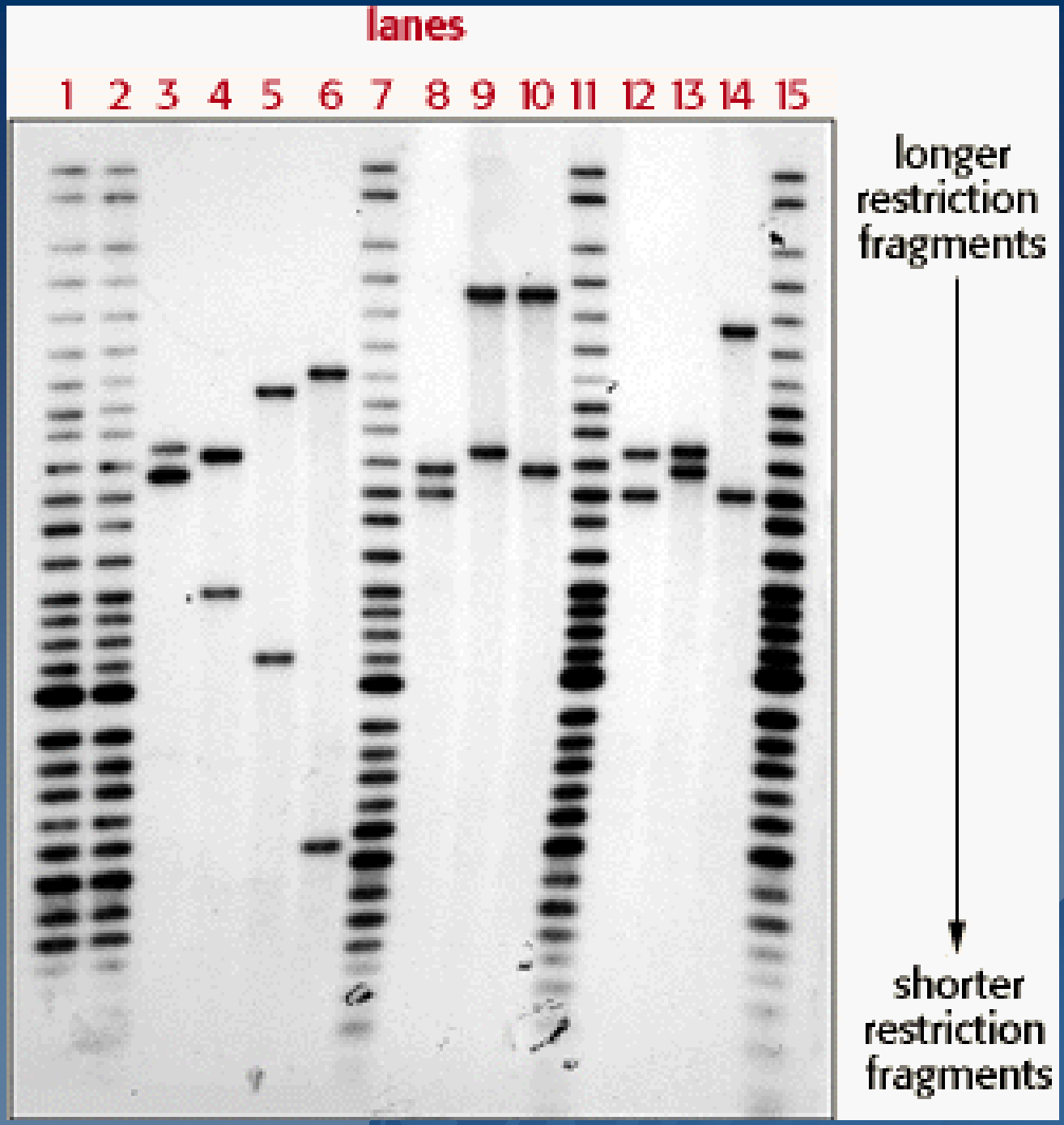
The gene for CF has been mapped, and RFLPs are known.

If a couple comes to a genetic counselor, often an RFLP analysis is performed on the couple's DNA.





An example autorad



SNPs

GCATGCA**a**GCATGCAT

GCATGCA**c**GCATGCAT

GCATGCA**a**GCATGCAT

GCATGCA**a**GCATGCAT

GCATGCA**a**GCATGCAT

This is either…

- A bona-fide SNP
- A sequencing error
- A paralogous gene

Consequences and Types of SNPs

- Changes to protein structure and function
  - Sickle cell anemia is classic example
- Silent mutations that may or may not have a phenotype
  - Regulatory elements
  - Splicing signals (in introns OR exons!)
    e.g., some SNPs in *BRCA1*
- Mitochondrial SNPs
  - Associated with defective oxidative phosphorylation; affect tissues requiring massive energy production such as cardiac and skeletal muscle

SNP Detection

- Resequencing of genomic DNA from many individuals in a population
- EST (cDNA) sequencing
- Resequencing by hybridization
- Bioinformatic mining and comparison of sequence data in databases

SNP Typing

- TaqMan assay – PCR based
- Non-PCR based methods, e.g. invasive cleavage assay
  (covered in reading material, not lecture!)

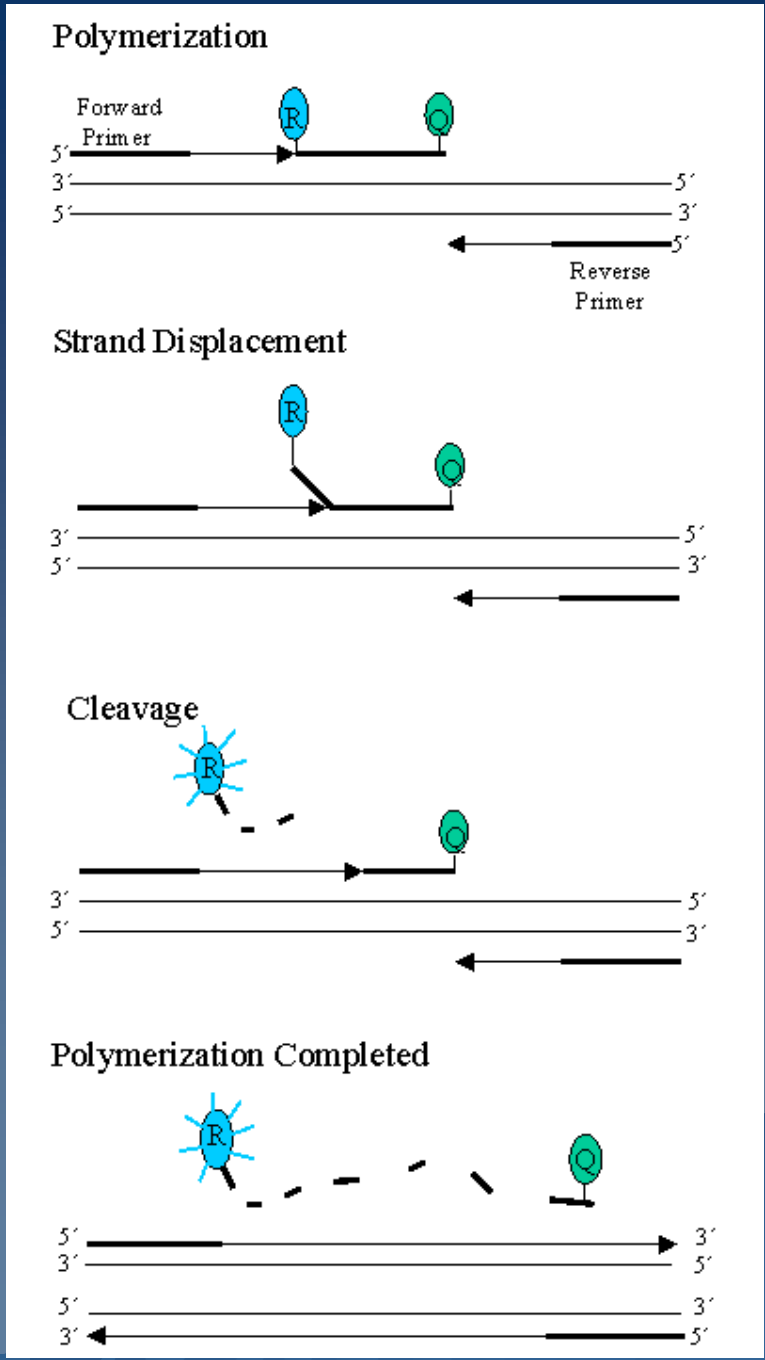
The TaqMan Method

A ~100 bp region flanking the SNP is amplified in the presence of two probes, each specific for one or the other allele.

The probes have a fluor reporter at the 5’ end and a quencher at the 3’ end.

The reporter only fluoresces when free in solution because of the quencher.

The presence of two probes of different colors allows simultaneous detection of both alleles



The TaqMan Instrument