Molecular population genetics

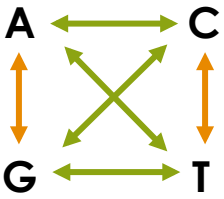
- advent of molecular methods provided direct measures of genetic variation…
- but also resulted in a paradoxical disconnect between genotype and phenotype…
- a connection that is is only now being re-established

Genetic variation

- "classical hypothesis"
  - genetic variation limited and comprised primarily of harmful mutations
- "balance hypothesis"
  - abundant genetic variation is maintained by some form of balancing selection
  - e.g., heterozygote advantage or frequency dependent selection
- the two hypotheses "*sat across the table glowering at each other through most of the 1950's and 1960's*"

Mutation: ultimate source of variation

- point mutations generate new **alleles** (or "haplotypes")
- also insertions, deletions, inversions, duplications (and recombination)



transitions
transversions

Allozymes

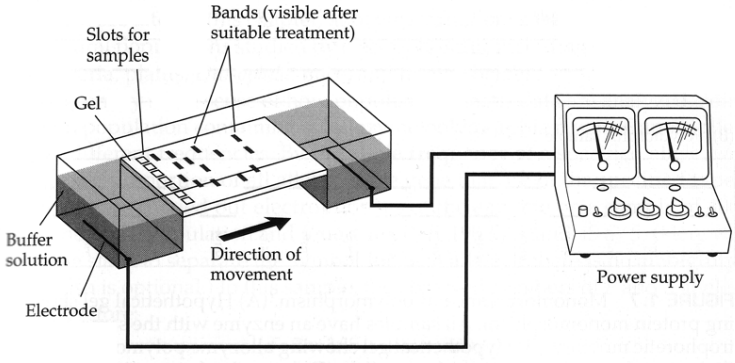
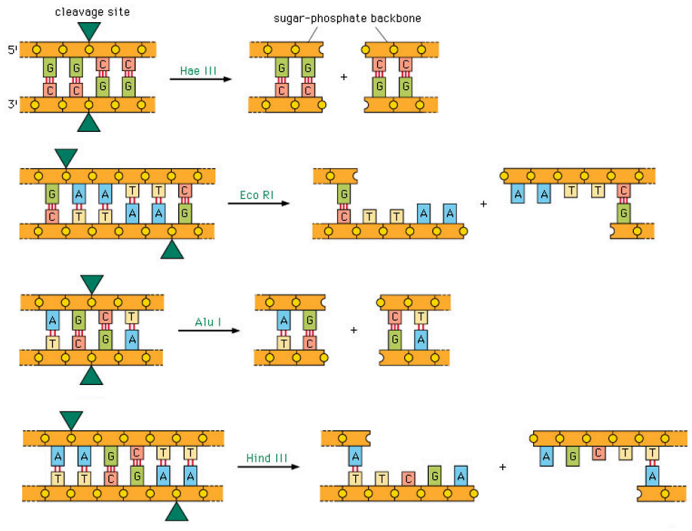


FIGURE 1.6 One type of laboratory apparatus for electrophoresis. The procedure is widely used to separate protein or DNA molecules. In conventional gels, DNA fragments smaller than about 20 kb (1 kb = 1000 nucleotide pairs) migrate approximately in proportion to the logarithm of their molecular weights.

Measuring genetic variation

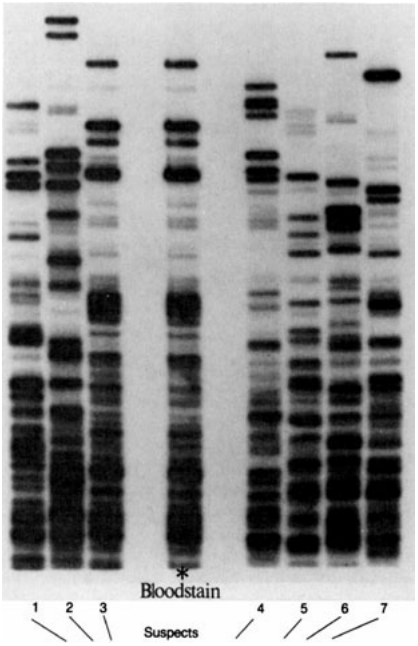
- Allozymes - protein electrophoresis
- RFLPs - restriction fragment length polymorphisms
- mini-satellites (VNTRs), microsatellites (SSRs)
  - often used for paternity analysis
- **DNA sequences**
- SSCP - single-stranded conformational polymorphism
- RAPDs -randomly amplified polymorphic DNA
- AFLPs - amplified fragment length polymorphisms
- **SNPs** - single nucleotide polymorphisms

Restriction enzymes

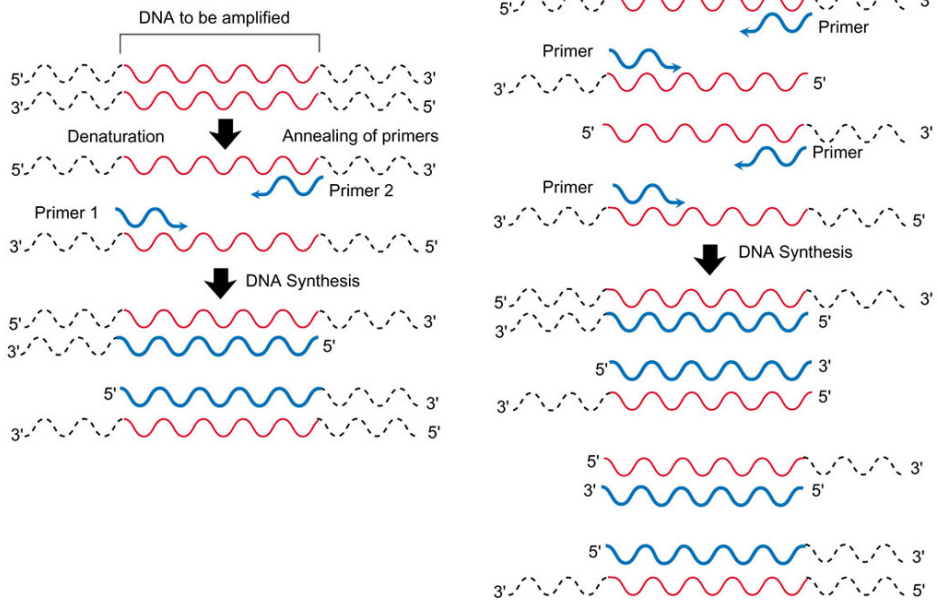


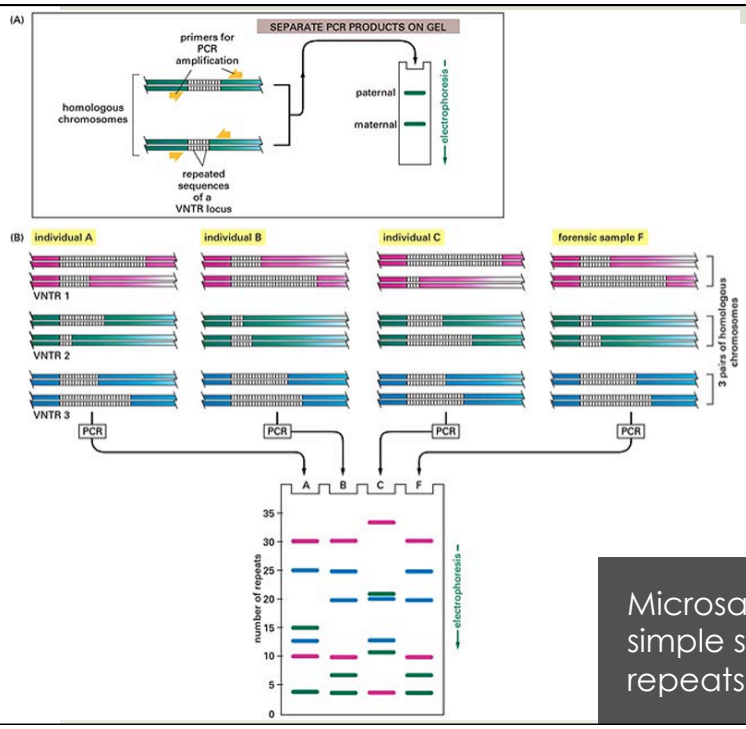
Minisatellite DNA = "multi-locus DNA fingerprinting"

- restriction digested genomic DNA hybridized to a radio-labeled probe
- probe matches highly repeated junk DNA sequence that occurs throughout the genome
- e.g., Jeffries probes 33.15 and 33.6
- why not significant in population genetics?



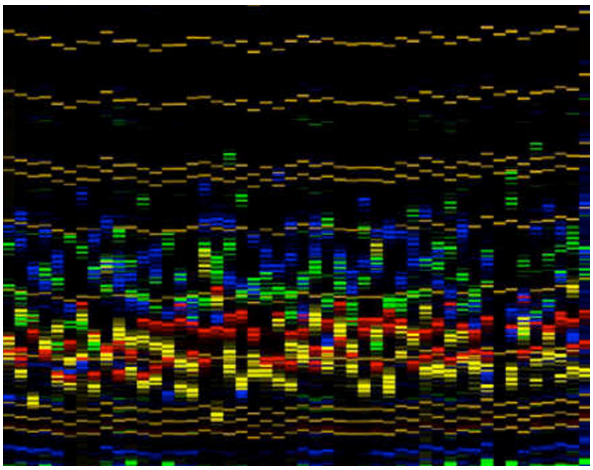
The Polymerase Chain Reaction (PCR)





Microsatellites = simple sequence repeats (SSR)

Microsatellites



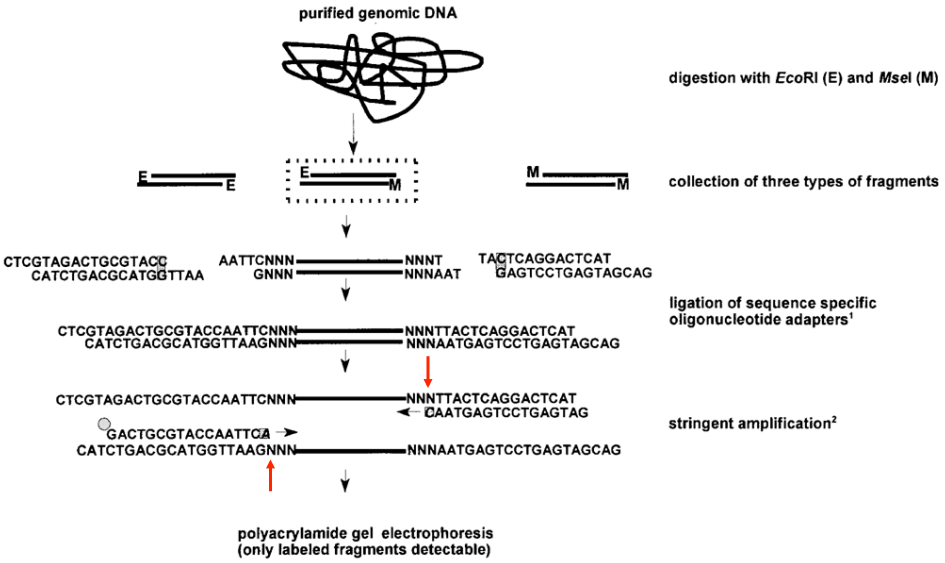
Microsatellites

- Issues in μ-sat data collection
  - null alleles - fail to amplify
  - hidden alleles - differ in sequence but not length
- Issues in μ-sat analysis
  - mutation model - stepwise or not?
  - substantial length "homoplasy"

AFLPs

- amplified fragment length polymorphism
- advantages:
  - fast survey of large number of loci
  - applicable to any organism
- disadvantages
  - generally anonymous loci
  - repeatability across samples?

AFLPs



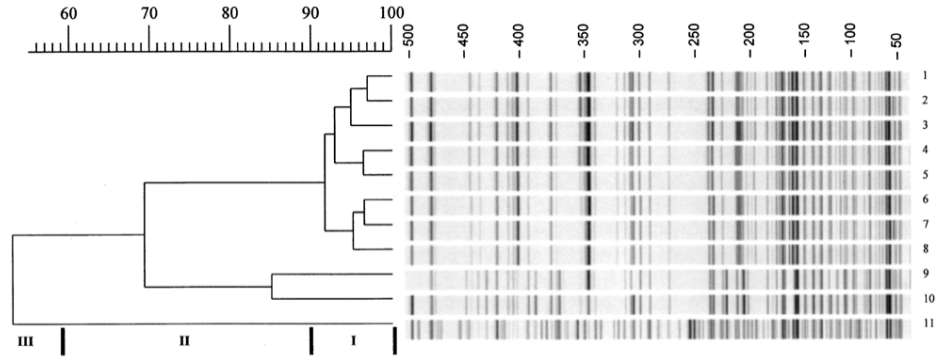
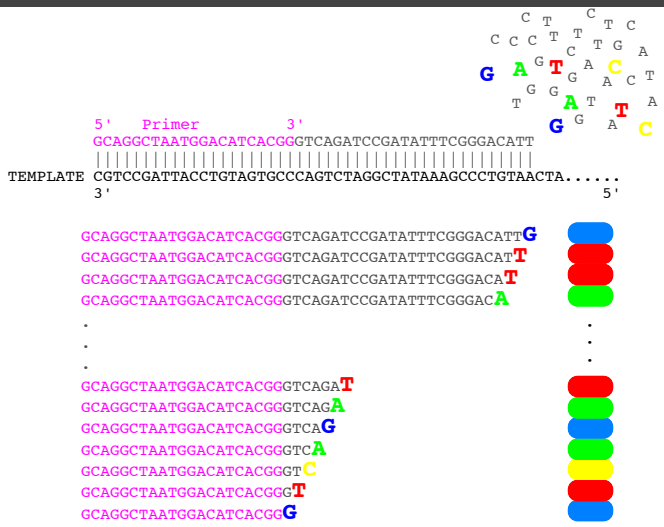
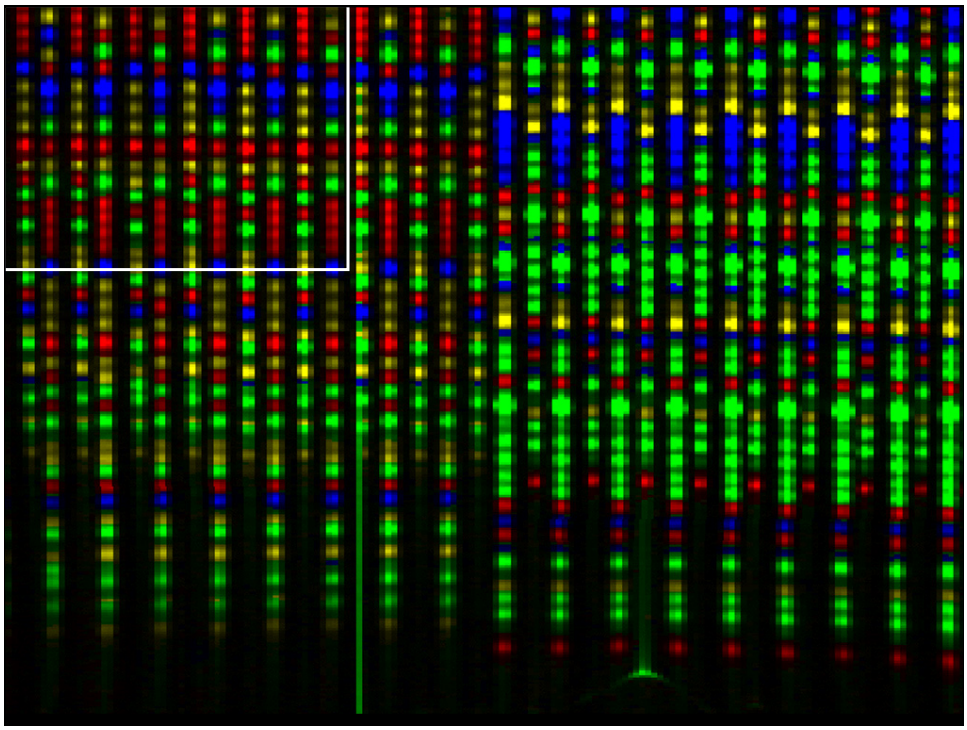
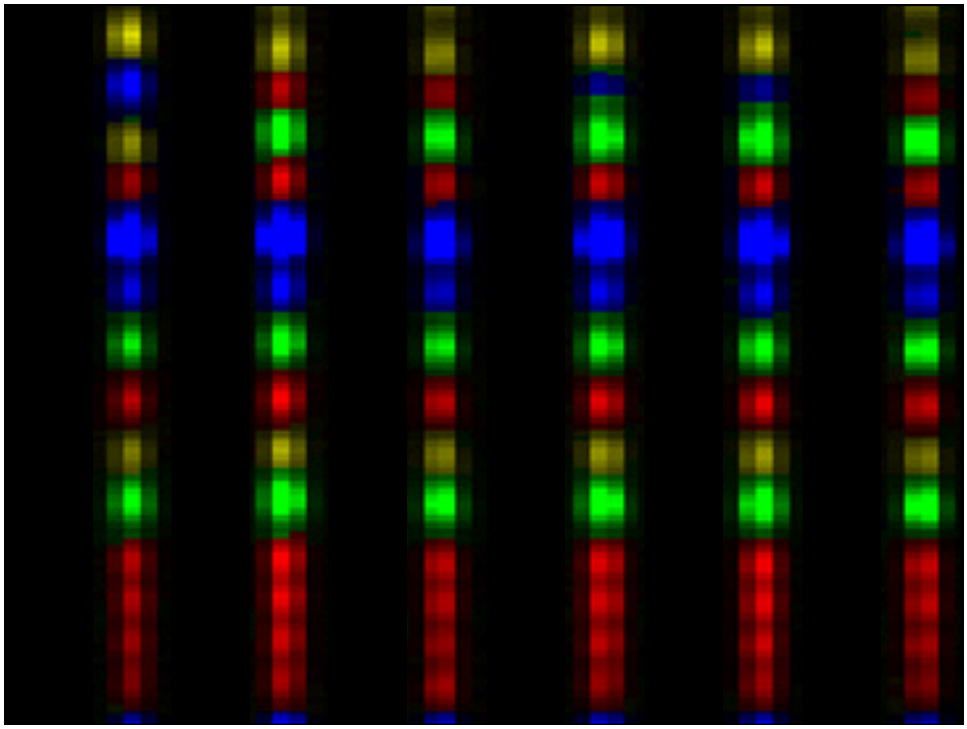


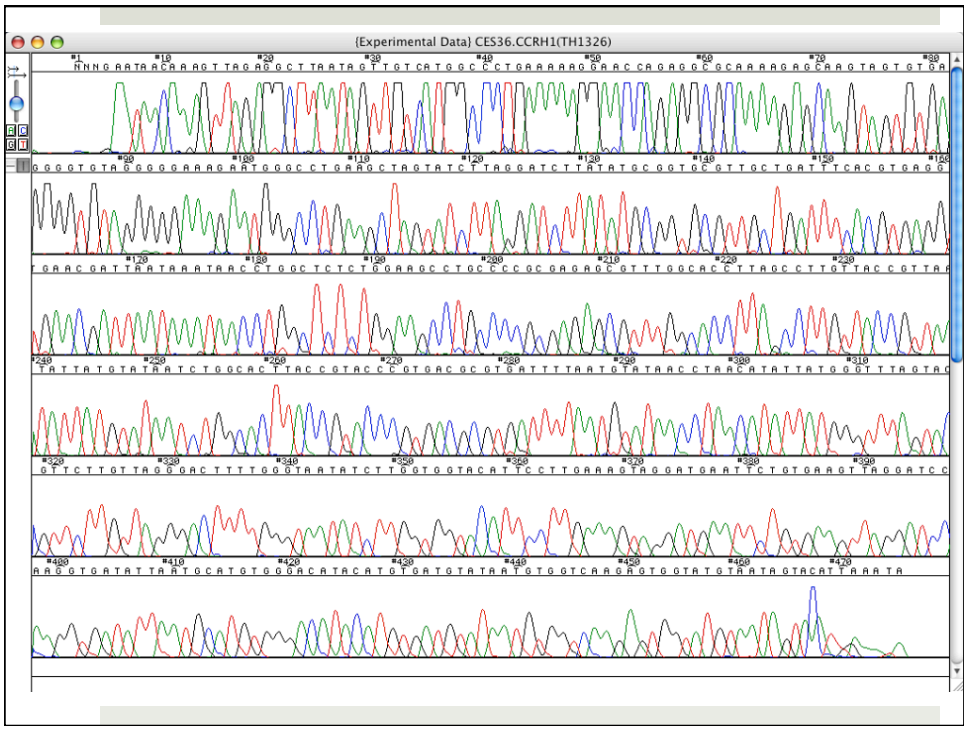
FIG. 2. Example of fluorescently labeled AFLP patterns and dendrogram for 11 different *Klebsiella* isolates. Patterns are the result of amplification of templates generated after restriction and ligation as shown in Fig. 1. The fragments were analyzed on an automated Vistra sequencer (Amersham-Pharmacia Biotech). The dendrogram was constructed with GelCompar (Applied Maths) software by using the Pearson correlation and cluster analysis by the unweighted pair group method using arithmetic averages. Percentages of similarity and molecular sizes (in base pairs) are shown above the dendrogram. Lanes 1 to 8, identical *Klebsiella pneumoniae* isolates; lanes 9 and 10, different *K. pneumoniae* strains; lane 11, a *Klebsiella oxytoca* strain. Within the AFLP patterns from *Klebsiella*, for instance, three windows of similarity may be applicable on the basis of the described experimental conditions: window I, 90 to 100% homology, identical strains; window II, 60 to 90% homology, different strains, same species (e.g., *Klebsiella pneumoniae*); window III, 40 to 60% homology, different species of the same genus; window IV, less than 40% homology, species from different genera.

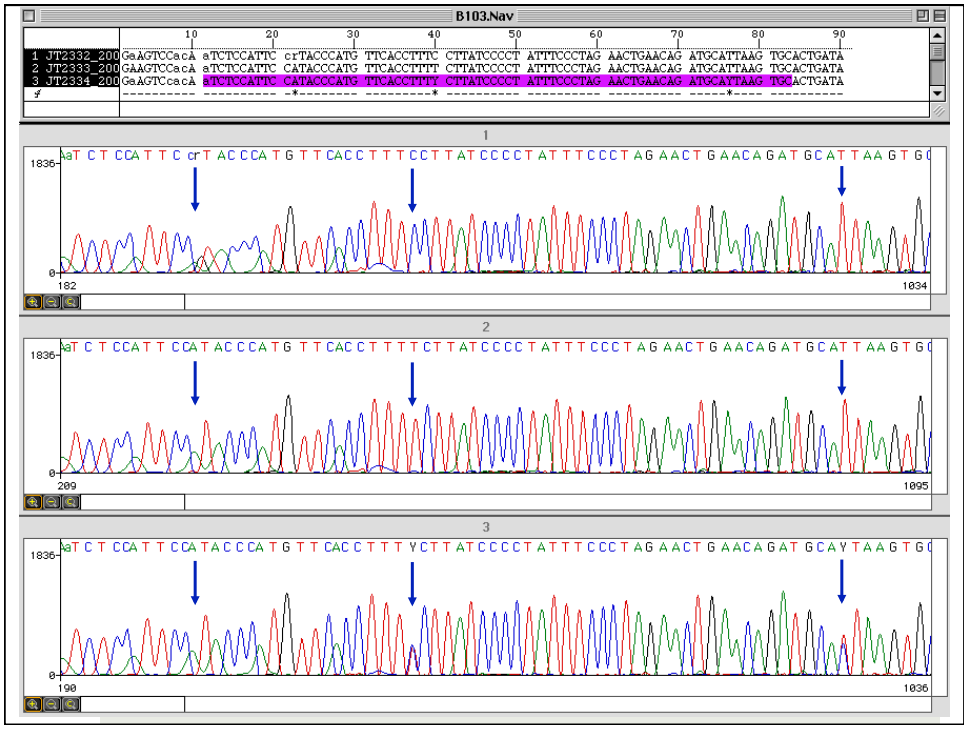
"Sanger" (chain termination) sequencing with fluorescent dye terminators








(Experimental Data) CES36.CCRH1(TH1326)



- DNA sequencing revealed surprisingly high levels of neutral genetic variation

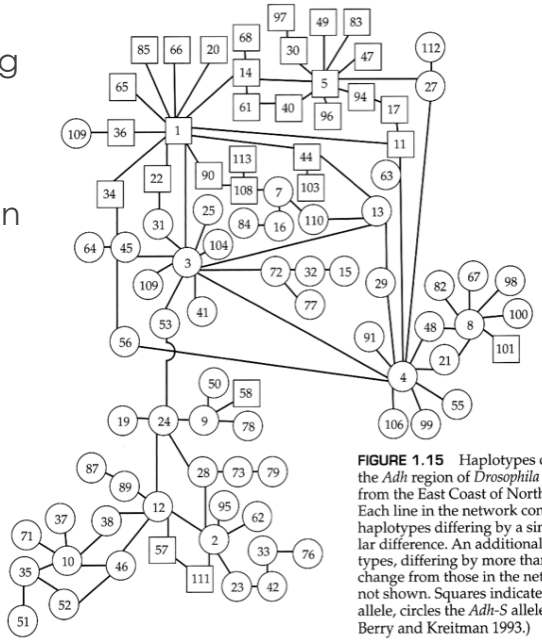


FIGURE 1.15 Haplotypes of alleles in the *Adh* region of *Drosophila melanogaster* from the East Coast of North America. Each line in the network connects two haplotypes differing by a single molecular difference. An additional 20 haplotypes, differing by more than one change from those in the network, are not shown. Squares indicate the *Adh-F* allele, circles the *Adh-S* allele. (From Berry and Kreitman 1993.)