Types/Sources of Genetic Data

Mendelian Genetics

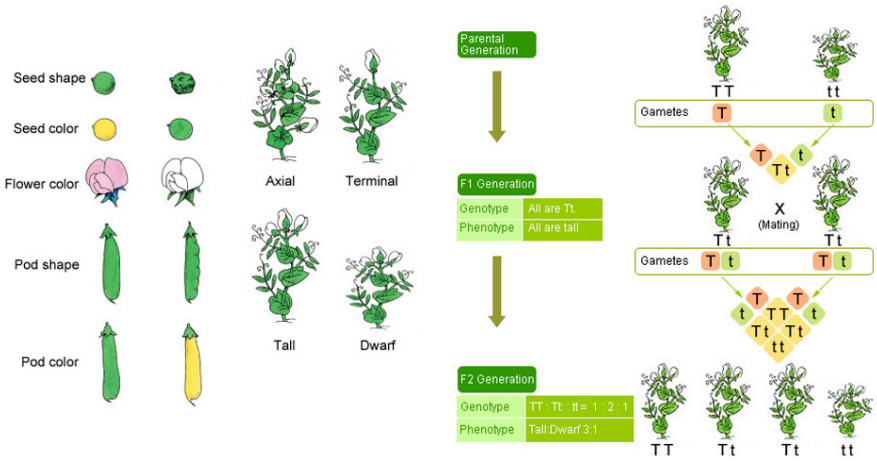
Before Mendel…

- Problem of "blending inheritance"
  - Darwin: *"I have lately been inclined to speculate very crudely & indistinctly, that propagation by true fertilisation, will turn out to be a sort of mixture & not true fusion, of two distinct individuals, or rather of innumerable individuals, as each parent has its parents & ancestors."*
- Jean-Baptiste Lamarck
  - inheritance of acquired characteristics
- Galton vs. Mendel
  - continuous phenotypic variation versus discrete traits

Gregor Mendel (1822-1884)

- 1866 paper detailed results of breeding experiments on garden peas
- observed classic ratios of discrete phenotypes in F2 generation
- results too good to be true?
  - perhaps "filtered" by Mendel
  - what's the chance of all seven traits being independent - i.e., on separate chromosomes?
- essentially ignored until early 1900's

Traits studied by Mendel

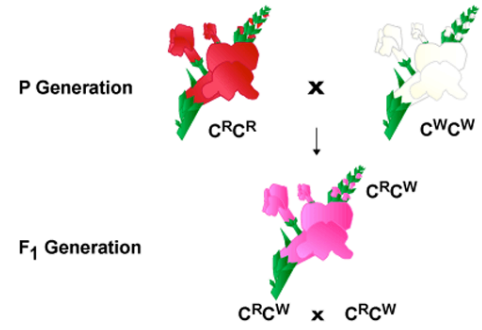


Genetic Dominance

- For a simple Mendelian trait determined by two alleles at one locus, the "dominant" trait/allele is "expressed" in heterozygotes
  - individuals homozygous for the dominant allele and heterozygous individuals have the same phenotype
- Dominant traits (are/do) not necessarily:
  - more frequent (common) in the population
  - produce bigger, stronger, faster, or more beautiful phenotypes
  - produce higher fitness
- Dominance may be incomplete (or partial)
  - codominance, over-dominance, under-dominance

"Mendelian" trait with incomplete dominance

- Snapdragon (*Antirrhinum majus*): cross of red and white parents yields pink "F1" individuals due to incomplete dominance
- In the "F2" generations, there is a 1:2:1 ratio of phenotypes



Human Mendelian traits?

- Tongue-rolling?
  - Martin NG 1975 No evidence for a genetic basis of tongue rolling or hand clasping. *J. Heredity* 66: 179-80.
- albinism, polydactyly
  - yes, but mutations in a number of different genes can cause these phenotypes
- eye color, hair color, freckles
  - one or a few genes of large effect, but also many modifying genes and environmental effects

Molecular population genetics

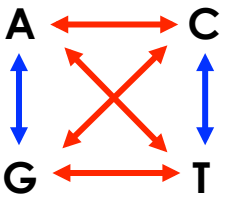
- advent of molecular methods provided direct measures of genetic variation…
- but also resulted in a paradoxical disconnect between genotype and phenotype…
- a connection that is is only now being re-established

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 is the ultimate source of genetic variation

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



transitions
transversions

Allozymes

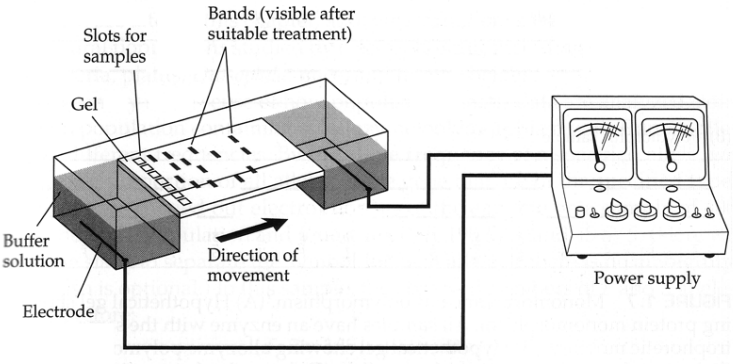
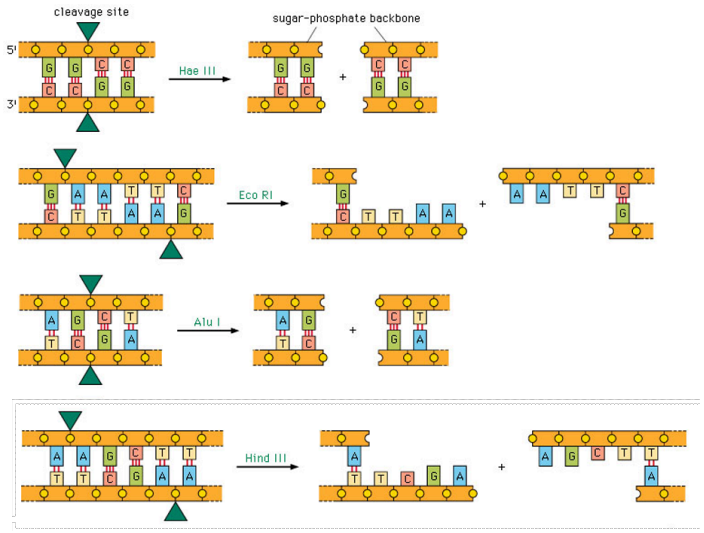


FIGURE 1.6 One type of laboratory apparatus for electrophoresis. The procedure is widely used to separate protein or DNA molecules. In conventional gels, DNA fragments smaller than about 20 kb (1 kb = 1000 nucleotide pairs) migrate approximately in proportion to the logarithm of their molecular weights.

Measuring genetic variation

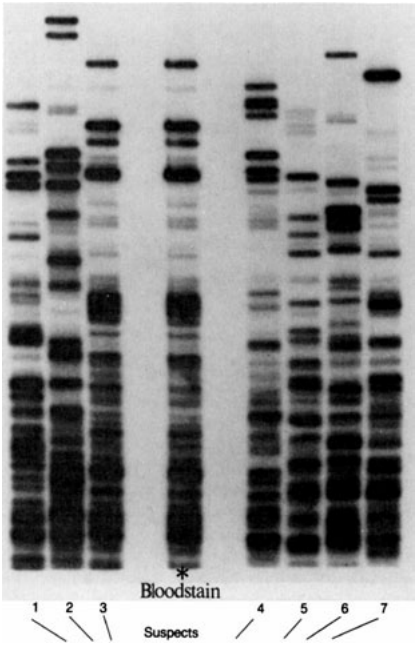
- Allozymes - protein electrophoresis
- RFLPs - restriction fragment length polymorphisms
- mini-satellites (VNTRs), microsatellites (SSRs)
  - often used for paternity analysis
- DNA sequences (esp. mtDNA: late 1980's-2000's)
- SSCP - single-stranded conformational polymorphism
- RAPDs - randomly amplified polymorphic DNA
- AFLPs - amplified fragment length polymorphisms
- **SNPs** - single nucleotide polymorphisms

Restriction enzymes

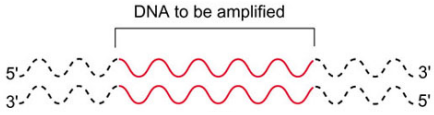


Minisatellite DNA = "multi-locus DNA

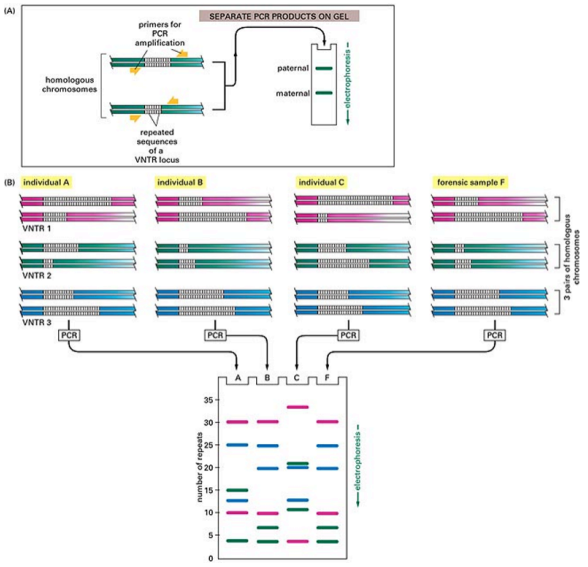
- restriction digested genomic DNA hybridized to a radio-labeled probe
- probe matches highly repeated junk DNA sequence that occurs throughout the genome
- e.g., Jeffries probes 33.15 and 33.6
- why not significant in population genetics?



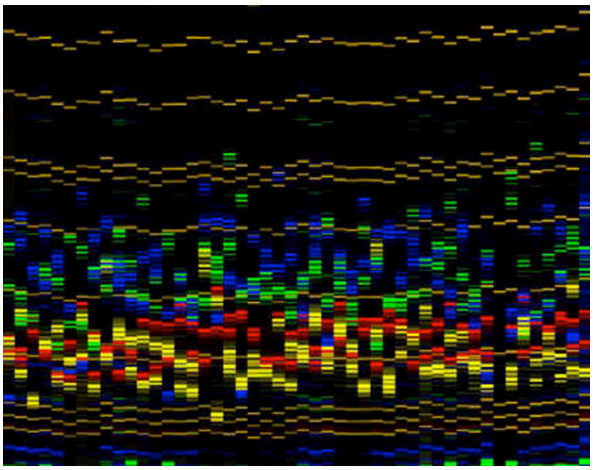
The Polymerase Chain Reaction (PCR)



Microsatellites



Microsatellites



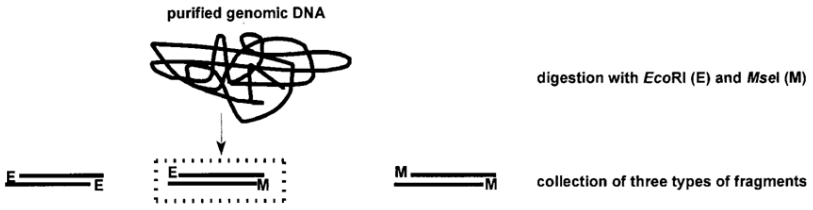
Microsatellites

- 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



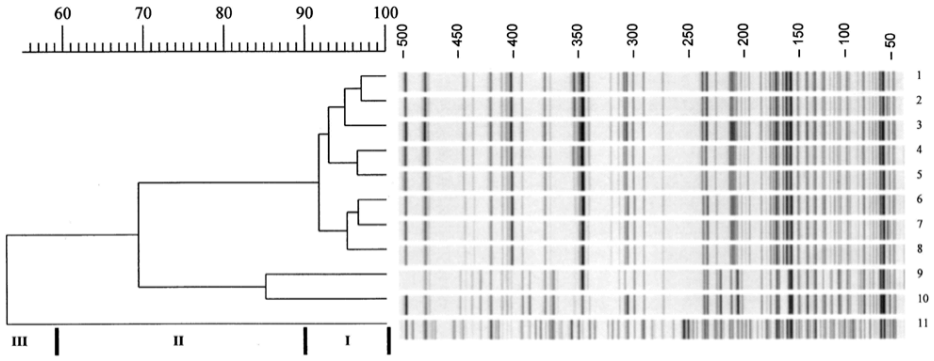
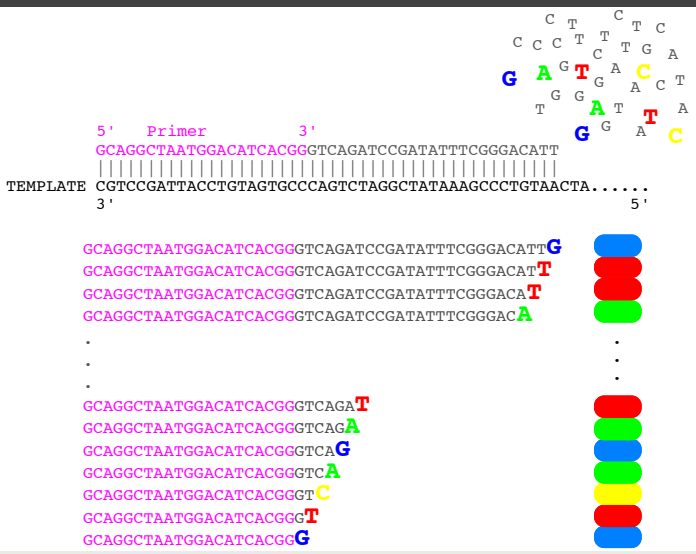
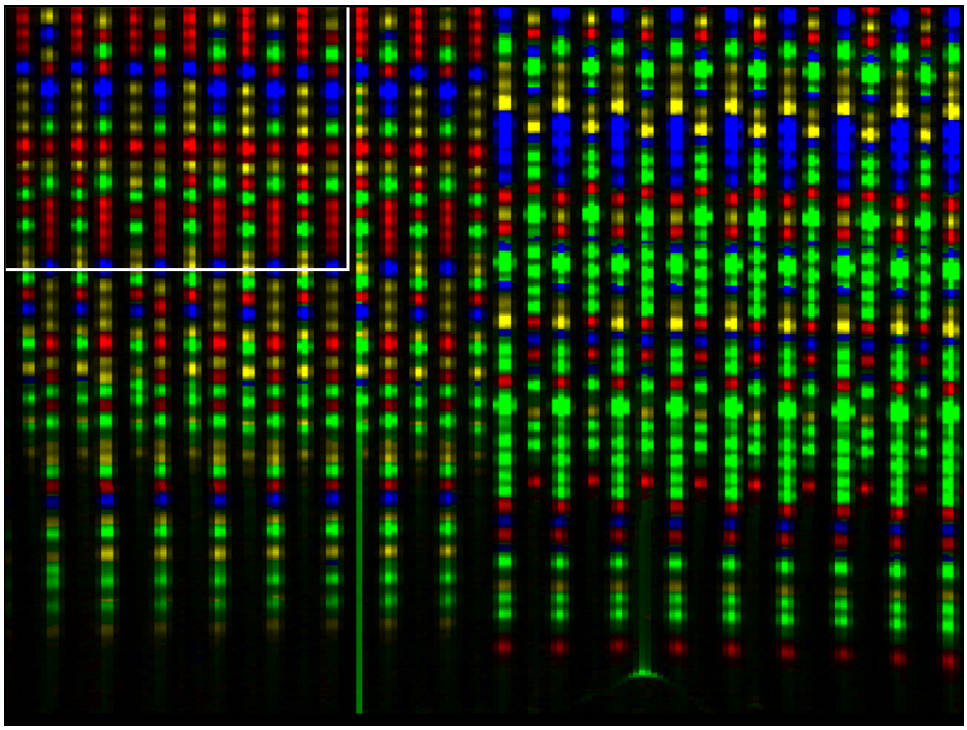
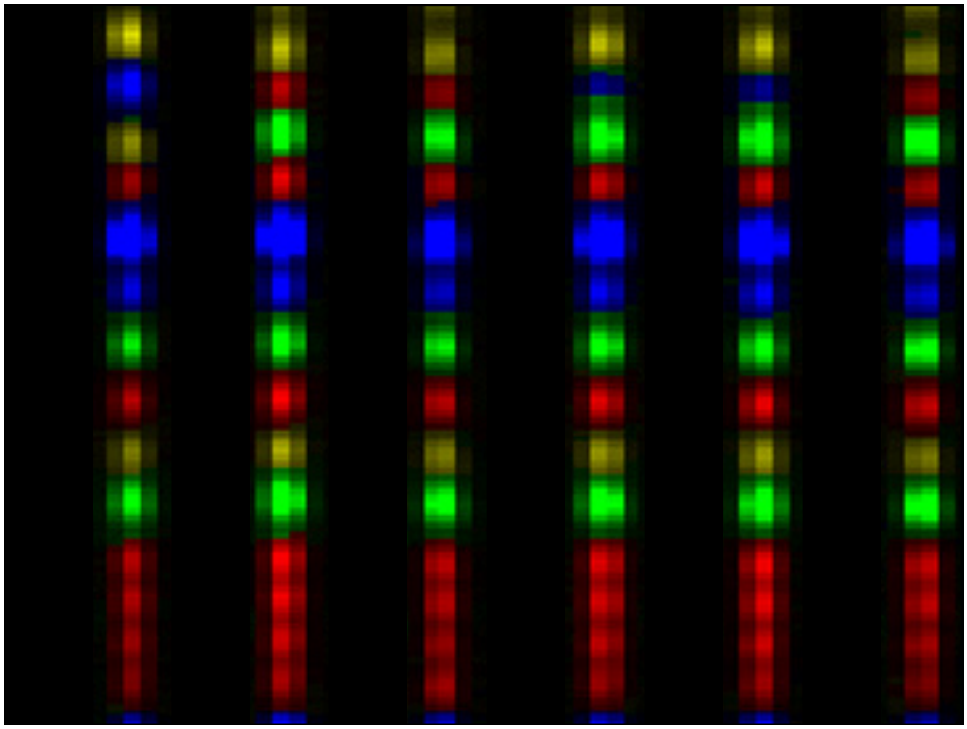


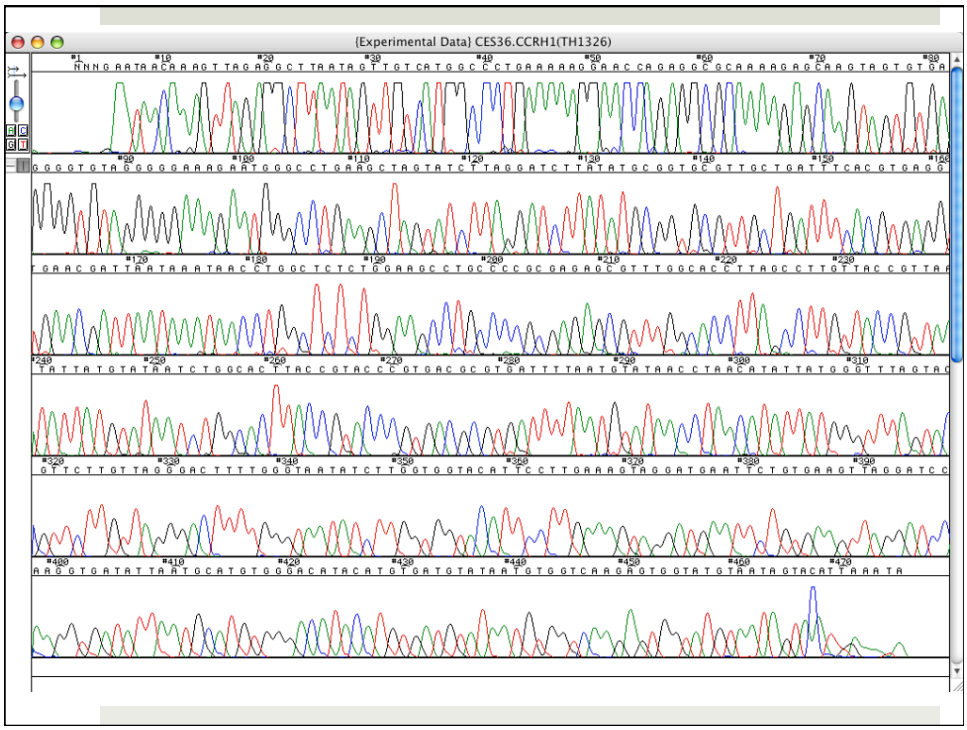
FIG. 2. Example of fluorescently labeled AFLP patterns and dendrogram for 11 different *Klebsiella* isolates. Patterns are the result of amplification of templates generated after restriction and ligation as shown in Fig. 1. The fragments were analyzed on an automated Vistra sequencer (Amersham-Pharmacia 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.

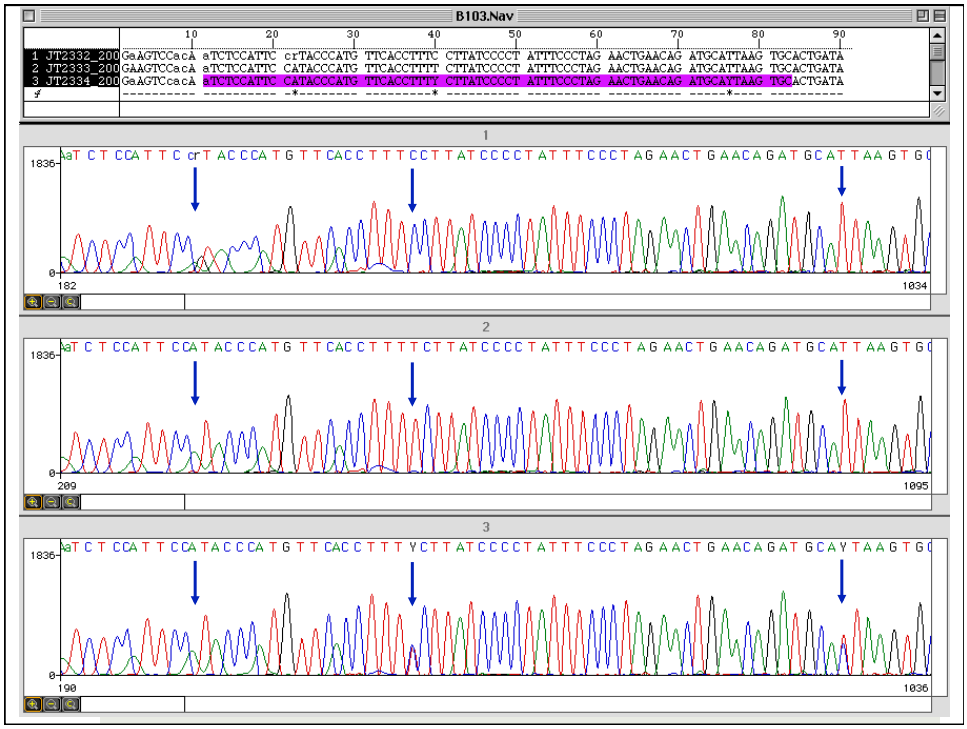
"Sanger" (chain termination) sequencing with fluorescent dye terminators







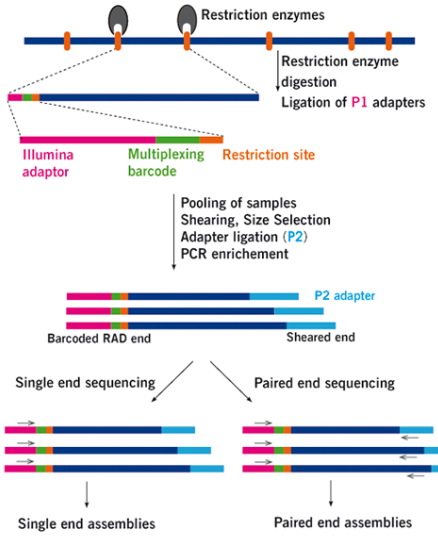

(Experimental Data) CES36.CCRH1(TH1326)

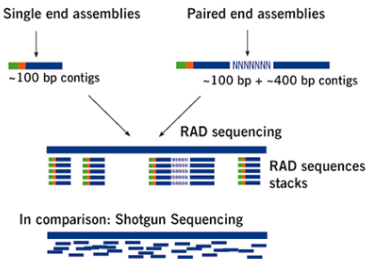


"Next-Gen" Methods

- Restriction Site Associated DNA Sequencing (RAD-seq)
  - Double-digest RAD-seq (ddRAD-seq)
  - Type IIB RAD-seq (2b-RAD-seq)
- Genotype by Sequencing (GBS)
- Complexity Reduction of Polymorphic Sequences (CRoPS)
- Sequence-Based Genotyping (SBG)
- Multiplexed Shotgun Genotyping (MSG)

e.g., RAD-seq





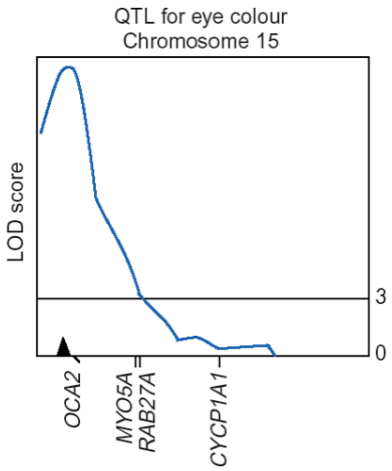
DNA is digested with a specific set of restriction enzymes and two adaptors ligated to the fragments afterwards.

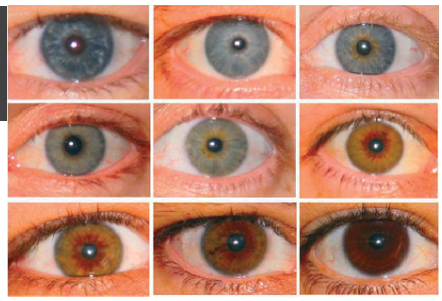
Sequencing the fragments from either one side or both sides results in RAD sequences stacks.

The stacks are the perfect starting point to identify and analyse genetic markers.

Genetic Basis for Human Mendelian traits?

Eye color?



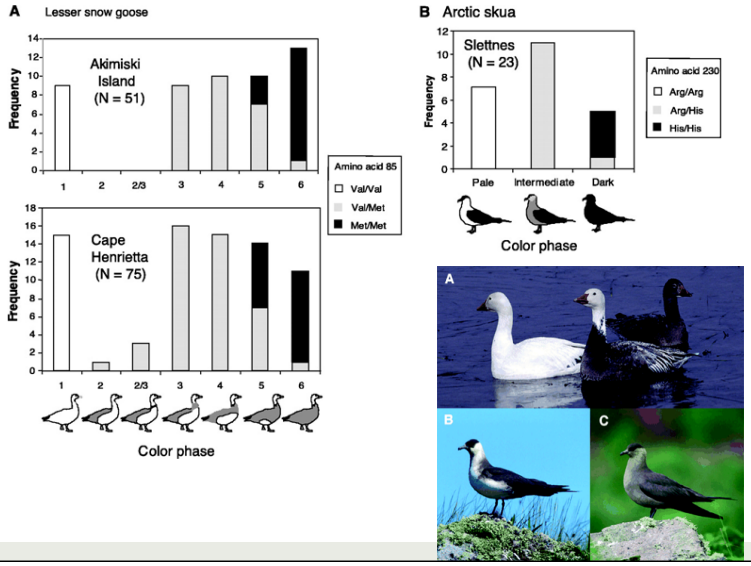


- the OCA2 gene is a principal determinant of blue versus brown eyes in Europeans, but many other genes are also correlated with eye color variation
- Sturm & Frudakis 2004 *Trends in Genetics* 20: 327-332.

Are there any truly discrete traits (controlled by a single-locus)?

- in most cases in which genes of large effect are important, the phenotype is also influenced by other genes and environmental effects
- MC1R - melanocortin 1 receptor
  - a.k.a.: melanocyte-stimulating hormone receptor
  - member of the G-protein-coupled receptor family
  - functions at the surface of specialized pigment producing cells (melanocytes)

MC1R in birds



Pocket mice

- Nachman et al. 2003 *PNAS* 100: 5268-5273.
- light and dark coloration produced by alternative alleles of MC1R at Pinacate but not at Armendaris
- independent evolution of similar phenotypes due to changes in different genes

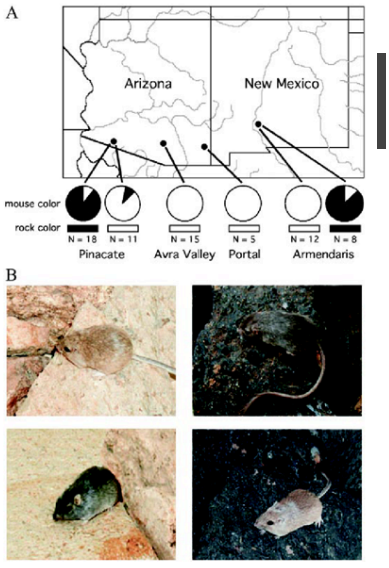


Fig. 1. (*A*) Collecting localities, substrate color, and mouse color. Sample sizes at each site are given. Pie charts indicate the proportion of light and dark mice at each site. Rectangles indicate the substrate color at each site. Mice from Pinacate and Armendaris were sampled on dark lava and also on light rock adjacent to the lava, whereas mice from Avra Valley and Portal were sampled only on light rock. (*B*) Light and dark *C. intermedius* from the Pinacate locality on light and dark rocks.

Neanderthal MC1R

- Lalueza-Fox *et al*. 2007 *Science* 318: 1453-1455.
- point mutation in the Neanderthal MC1R gene suggests inactive variant that may have resulted in red hair!

"Asian flush"

"Asian flush"

- ethanol broken down to acetaldehyde by ADH (alcohol dehydrogenase), then to acetic acid by ALDH2 (aldehyde dehydrogenase)
- "defective" ALDH2*2 allele is relatively common in Asian populations
- reduced enzyme function due to a single amino acid substitution results in buildup of toxic acetaldehyde in the bloodstream
- incomplete dominance: stronger effect in homozygotes than heterozygotes

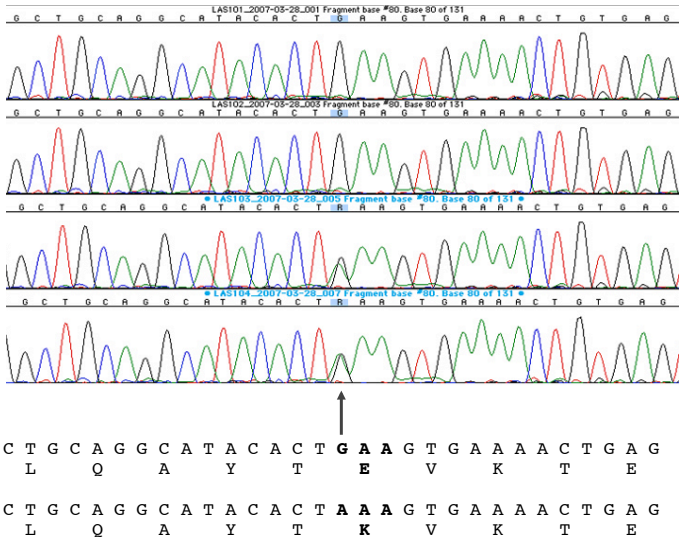
Functional variant (SNP) in ALDH2 gene

Subject1

Subject2

Mother

Grandma



E = glutamic acid, K = lysine