Genetics: From Genes to Genomes
Chapter 10 Analyzing Genomic Variation

10.1 Variation Among Genomes (343 - 347)

- There is no human wild-type

DNA Polymorphisms - sequence differences

Only a small minority of DNA variation is responsible for phenotypic differences

- less than 2% of the human genome consists of codons within genes
- even when they occur, many mutations of codons are silent (don't change amino acid)
- if a particular mutation is not silent and is deleterious, natural selection could get rid of it

Anonymous DNA Polymorphisms - affect neither nature or any protein in the body

DNA Marker - even an anonymous polymorphism can act as a signpost in the genome

Categories of Genetic Variants

Single Nucleotide Polymorphisms (SNPs) - alternate letter of alphabet

Derived Allele - changed nucleotide

Ancestral Allele - present long before human species took form

- many humans share unlinked SNP (showing common ancestry)

Deletion-Insertion Polymorphisms (DIPs)

- include problems in replication or recombination
- act as frameshift mutations unless number deleted or inserted is multiple of 3

Simple Sequence Repeats (SSRs)

- arise from rare, random events that produce a short repeated sequence
- can expand by *slipped mispairing* or *stuttering*

Copy Number Variants

- unequal crossing over
- genetic recombination takes place between mispaired repeating units
- gametes produced that have more or fewer copies of the repeating unit

10.2 Genotyping a Known Disease-Causing Mutation (347-353)

- must be able to isolate the disease-causing gene and analyze the purified DNA

PCR - (polymerase chain reaction) - amplifies a target region of DNA

- two oligonucleotides, PCR primers, define the ends of the target region
- each oligonucleotide is complementary is complementary to a strand
- denatured, primers hybridize, copying occurs, denatured again, etc.
- amount of DNA is doubled each time
- DNA polymerase used is from bacterium that grow in hot springs

PCR Products

- genotyped by either sequencing or sizing

Prenatal Genetic Diagnosis - genotyping fetal cells using…

amniocentesis - amniotic fluid surrounding fetus extracted using needle

preimplantation embryo diagnosis - stimulate maturity of ten eggs, cells are extracted, put through PCR, and sequenced or sized

10.3 Forensic DNA Fingerprinting Examines Multiple SSR Loci (353-356)

- SSR Loci are highly polymorphic
- the chance that two people will have the same alleles at all 13 SSR locations is $(0.1)^{13}$
- one chance in 10 trillion

DNA Fingerprint - genotype of 13 unlinked, polymorphic SSR loci

- 13 pairs of PCR primers labeled with fluorescent dyes
- FBI maintains CODIS (Combined DNA Index System) to share and compare DNA profiles
- Siblings and parents share 50% of SSR loci, allowing partial familial matches

Nucleic Acid Hybridization - the ability of complementary strands of DNA / RNA to come together

- a double-stranded hybrid is significantly less stable if one of the base pairs doesn't match

Allele Specific Oligonucleotide - 20-40 base oligonucleotide that will hybridize to only one allele

DNA Microarray - ASO attached to a solid support (like a silicon chip)

Probe - DNA fragmented and fluoresced

- fluorescence of shows number of copies of allele

10.4 Positional Cloning (357-363)

Disease Gene - genes whose mutant alleles cause a disease phenotype

Positional Cloning - strategy to obtain information about location of disease gene by finding polymorphic loci to which the mutation is genetically linked

- look for mutations that consistently appear in patients
- track one locus by phenotype (un/affected) and the other by direct genotyping
- can use DNA microarrays to follow millions of anonymous loci
- DNA microarrays are so densely packed with polymorphic loci that they must be linked
- multipoint information about loci in respect to each other in microarrays

  Lod Score - (log of the odds) whether or not data is significant
  - score of 3 means it is 1000 times more likely linked that not (10^3)
  - Lod = log(P(linked)/P(unlinked))
- Once you narrow down location of the gene, best strategy is to use PCR and sequence all of the candidate genes

Allelic Heterogeneity - caused by a variety of different mutations in the same gene

Compound Heterozygote - one copy of chromosome has one mutation and the other, another

- medicine could treat only one mutation

Locus Heterogeneity - caused by different genes

10.5 The Era of Whole-Genome Sequencing

Whole-exome sequencing - enrich genomic fragments that correspond to exons and then sequence, which constitutes less than 2% of whole genome

- high throughput methods
  - straightforward extensions of Sanger sequencing
  - DNA molecules being synthesized are anchored in place
  - control base addition temporally and read each one before the next is sequenced
  - sensitivity of detection is high enough to make PCR unnecessary

- variants that have been documented in databanks as common are probably not diseases
- sequence, compare to databases, filter out missense, filter frequent mutations