SISG 2023 - Module 2

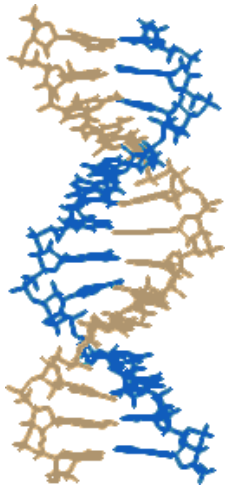
Introduction to Genetics and Genomics

What is a gene?

Block #1 – Monday, July 10

Joe Lachance and Greg Gibson

joseph.lachance@biology.gatech.edu



Date	Time (PDT)	Reading	Topic	Instructor
Monday, July 10	8:30am – 8:45am		Introductions	
	8:45am – 10:00am	Portin and Wilkins (Genetics, 2017)	What is a Gene ?	JL
	10:00am – 10:30am		Break	
	10:30am – 11:45am	Visscher et al. (Nat Rev Genet, 2008)	Heritability	GG
	11:45am – noon		Q&A Discussions	
Monday, July 10	1:30pm – 3:00pm	Goddard et al. (Proc Roy Soc B, 2016)	Quantitative Genetics	GG
	3:00pm – 3:30pm		Break	
	3:30pm – 4:45pm	Deichmann (Dev Biol, 2016)	Molecular Biology	JL
	4:45pm – 5:00pm		Q&A Discussions	
Tuesday, July 11	8:30am – 10:00am	Abdellaoui et al. (Am J Hum Genet, 2023)	Genome-Wide Association Studies	GG
	10:00am – 10:30am		Break	
	10:30am – 11:45am	Christmas et al. (Science, 2023)	Molecular Evolution	JL
	11:45am – noon		Q&A Discussions	
Tuesday, July 11	1:30pm – 3:00pm	Wang et al. (Nat Rev Genet, 2009)	Gene Expression Profiling	GG
	3:00pm – 3:30pm		Break	
	3:30pm – 4:45pm	Prohaska et al. (Cell, 2023)	Population Genetics	JL
	4:45pm – 5:00pm		Q&A Discussions	
Wednesday, July 12	8:30am – 10:00am	Ashley (Nat Rev Genet, 2016)	Genomic Medicine	GG
	10:00am – 10:30am		Break	
	10:30am – 11:45am	Bergström et al. (Nature, 2021)	Genetic Ancestry	JL
	11:45am – noon		Q&A Discussions	

What is a gene?

- How would you define a gene?

- Which matters more, structure or function?

Terminology

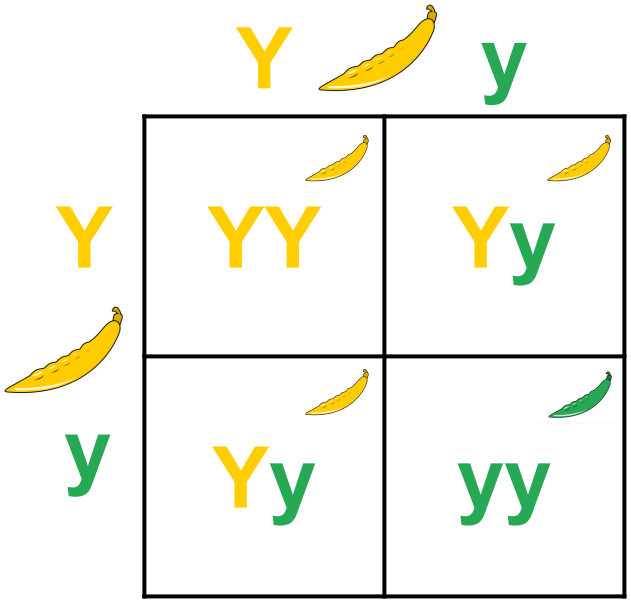
- **Genes**: DNA sequences that encode a functional protein or RNA molecule
- **Allele:** one of two or more alternative forms of a gene
- **Genotype:** the genetic makeup of an individual
- **Phenotype:** the observable characteristics and traits of an organism
- **Genome**: the complete set of genetic material in a cell or organism
- **Haplotype**: closely linked DNA sequencies on the same chromosome that are co-inherited

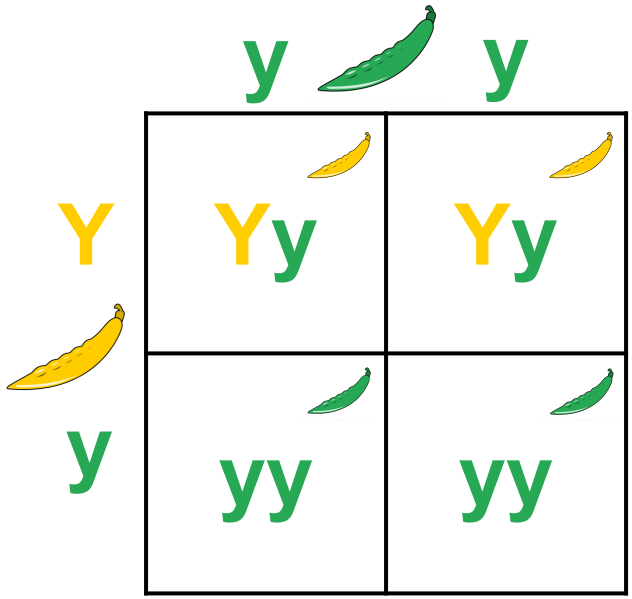
Mendel's laws of inheritance

- Law of segregation (1st law)
  - Parental pairs of alleles separate during gamete formation
- Law of independent assortment (2nd law)
  - Pairs of alleles for different traits segregate independently
- Law of dominance (3rd law)
  - Heterozygotes manifest the trait associated with the dominant allele
- *These rules are often broken!*

Image source: Wikipedia

Mendelian ratios

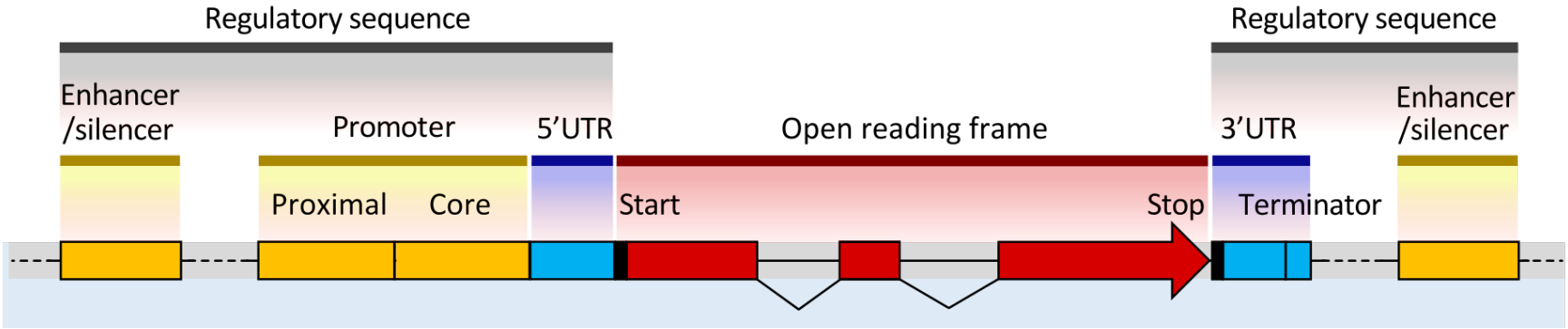




- Punnett squares can be used to predict the products of breeding
- Genetics is not always this simple!

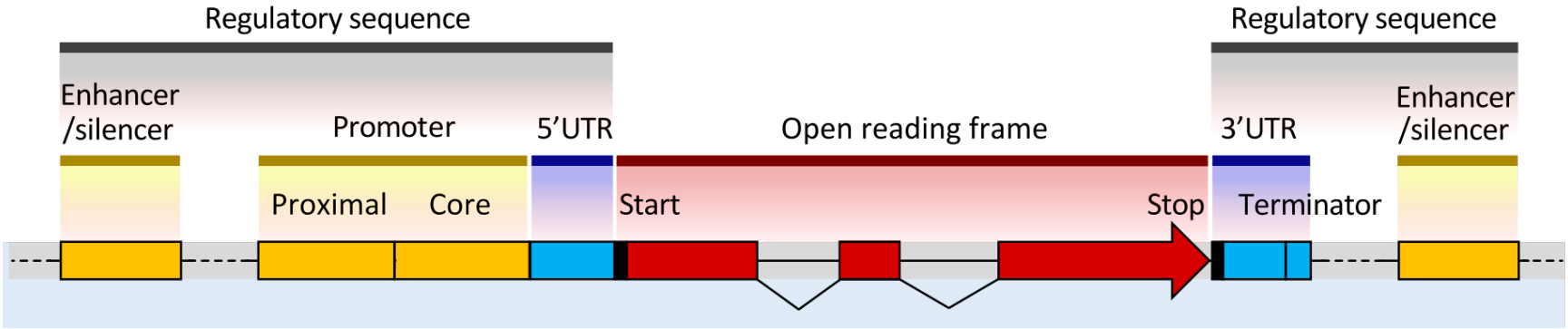
	Prokaryotes	Eukaryotes
Internal structures	No organelles	Organelles
DNA	Circular No introns DNA in cytoplasm	Linear Introns DNA in nucleus
Genome size	Tend to be < 5Mb	10Mb-100,000Mb
Chromatin	No histones	Histones
Ploidy	Haploid	Usually diploid
Reproduction	Asexual (binary fission)	Asexual (mitosis) and sexual (meiosis)

The structure of (protein coding) genes



- Exons: nucleotide sequence not removed by splicing (coding DNA)
- Introns: nucleotide sequence removed by splicing (noncoding DNA)
- Cis-regulatory elements
  - Enhancers: increase the likelihood of transcription when bound to activators
  - Silencers: decrease likelihood of transcription when bound to repressors
  - Promoters: region of DNA where transcription is initiated
- UTRs: untranslated regions

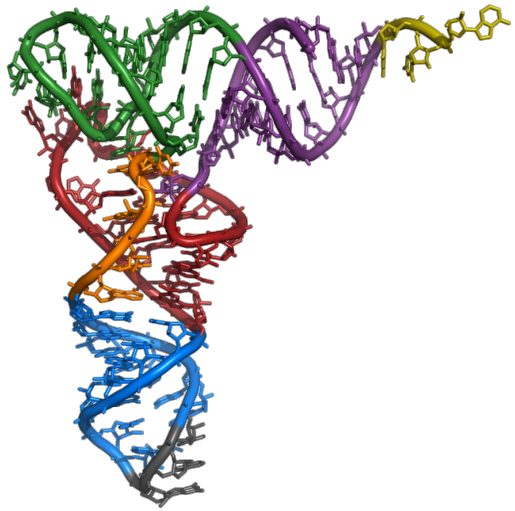
The structure of (protein coding) genes



- Which parts would you consider to be part of a gene?

RNA genes

- DNA sequences which encode function non-coding RNA are called RNA genes
- Transfer RNAs (tRNAs)
- Ribosomal RNAs (rRNAs)



- Different types of small RNAs (e.g., microRNAs, siRNAs)

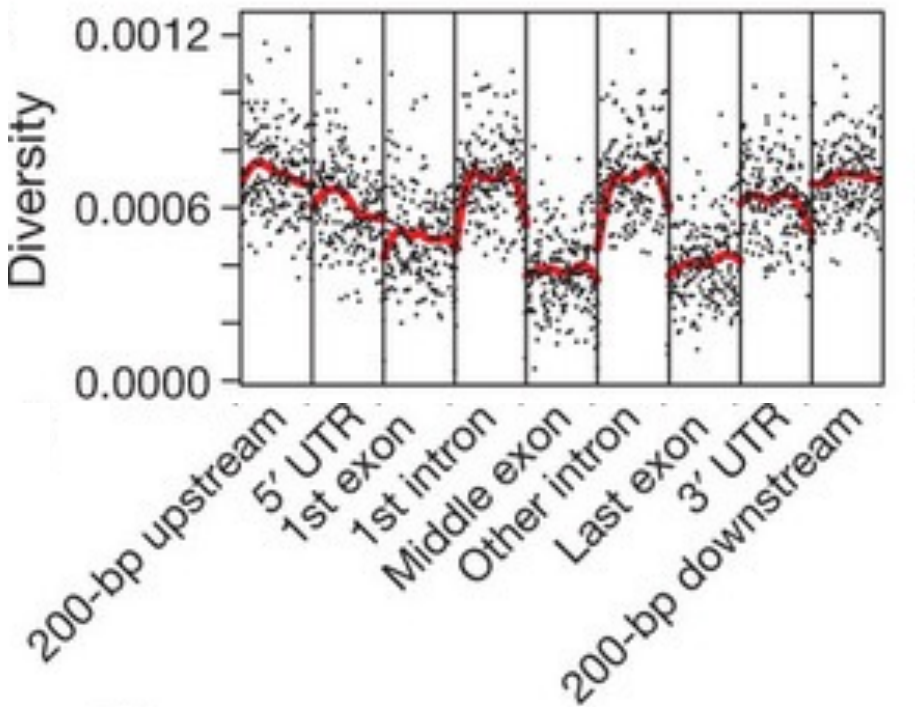
The locus of evolution

- What sort of genetic changes underlie morphological adaptations?

- Hopi Hoekstra and Jerry Coyne
  - Protein-coding DNA matters

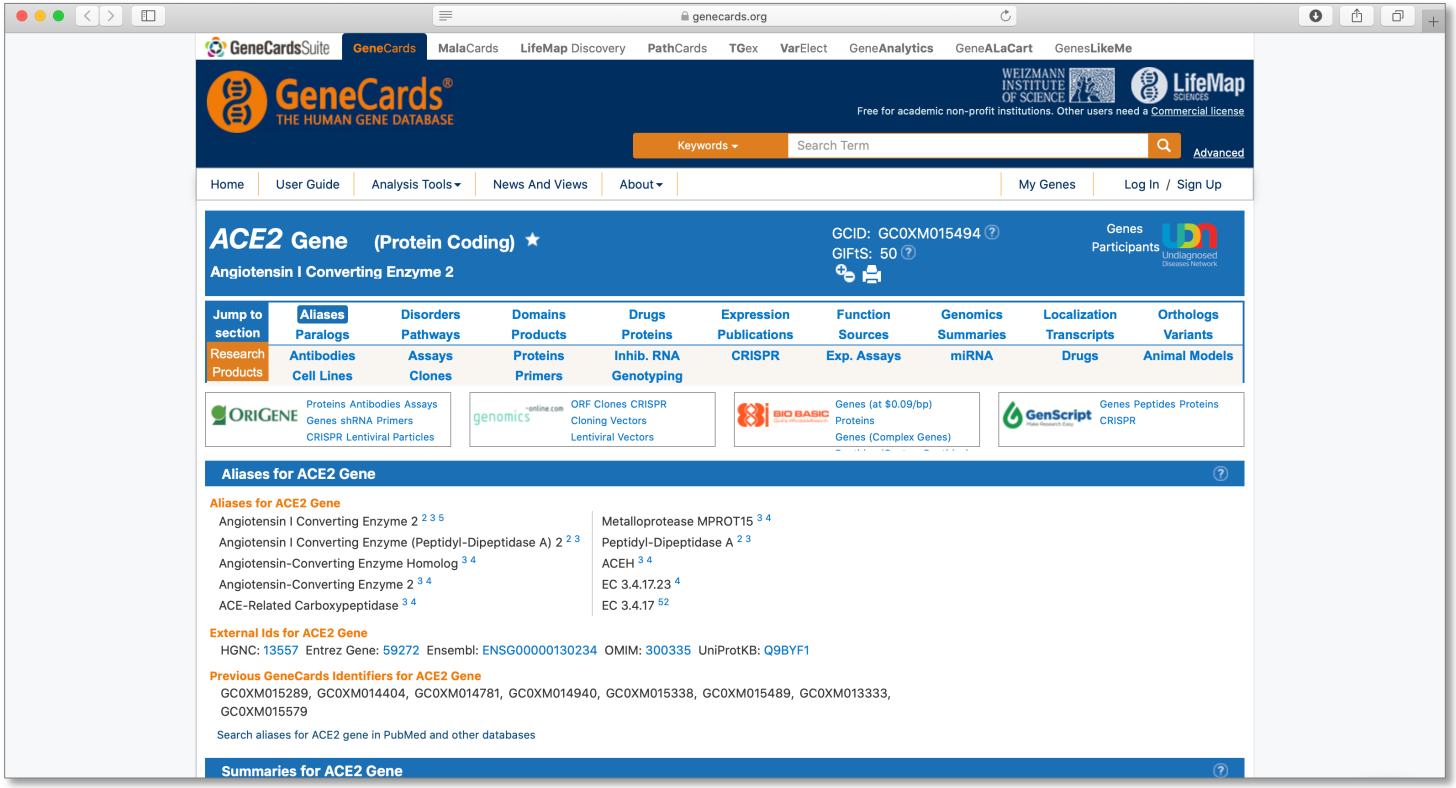
- Sean Carroll
  - It's all about regulatory DNA

Polymorphism near genes



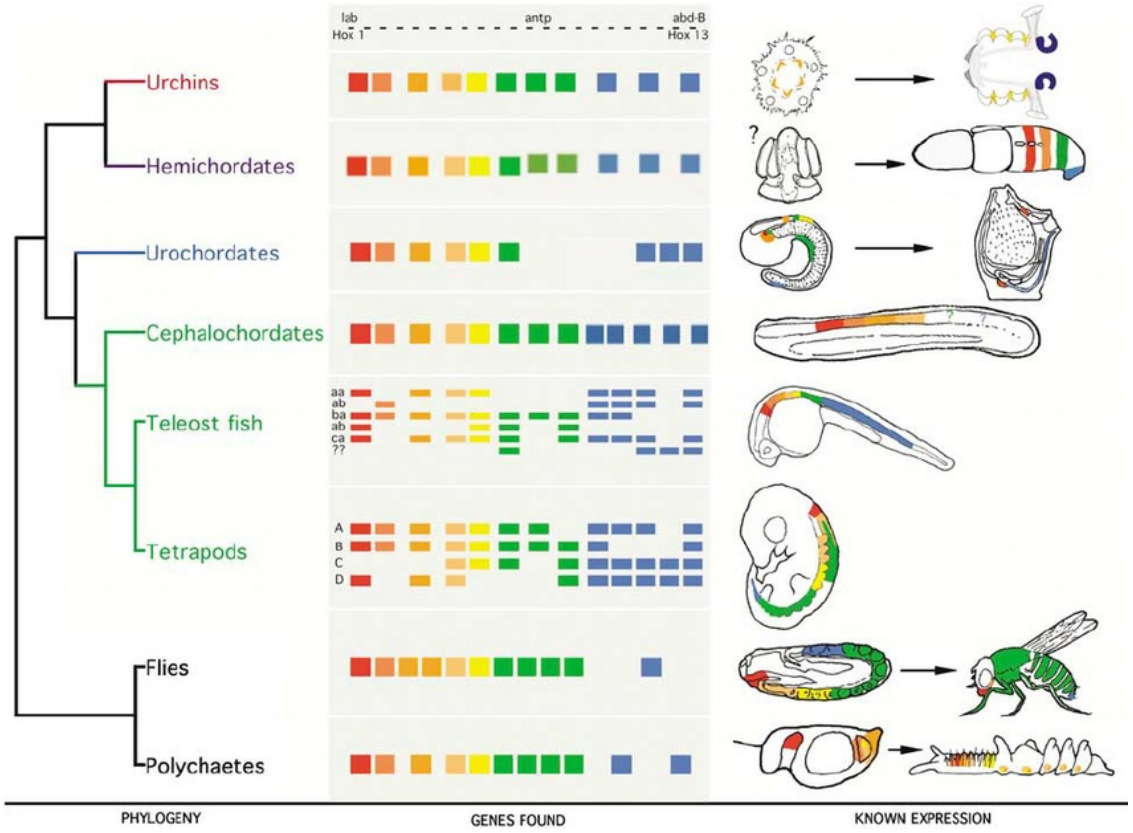
- Exons contain less genetic variation than non-coding DNA

GeneCards



- How to find more information about a particular gene?
- https://www.genecards.org

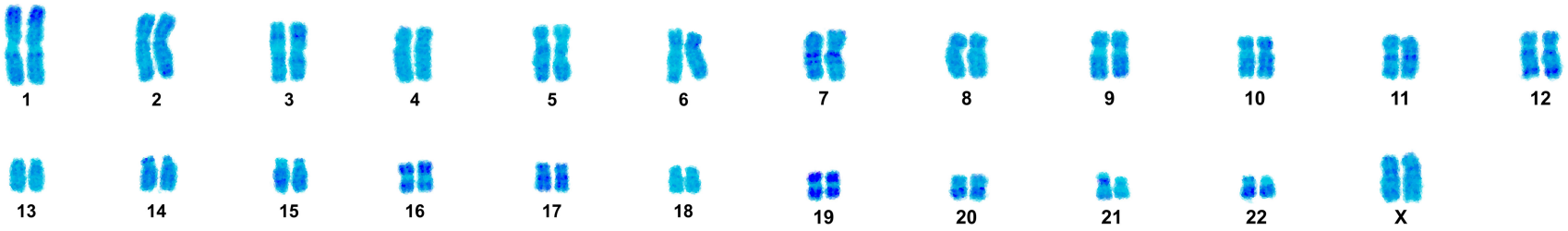
Different species share many of the same genes



- Homologous genes are two or more genes that descend from a common ancestral DNA sequences

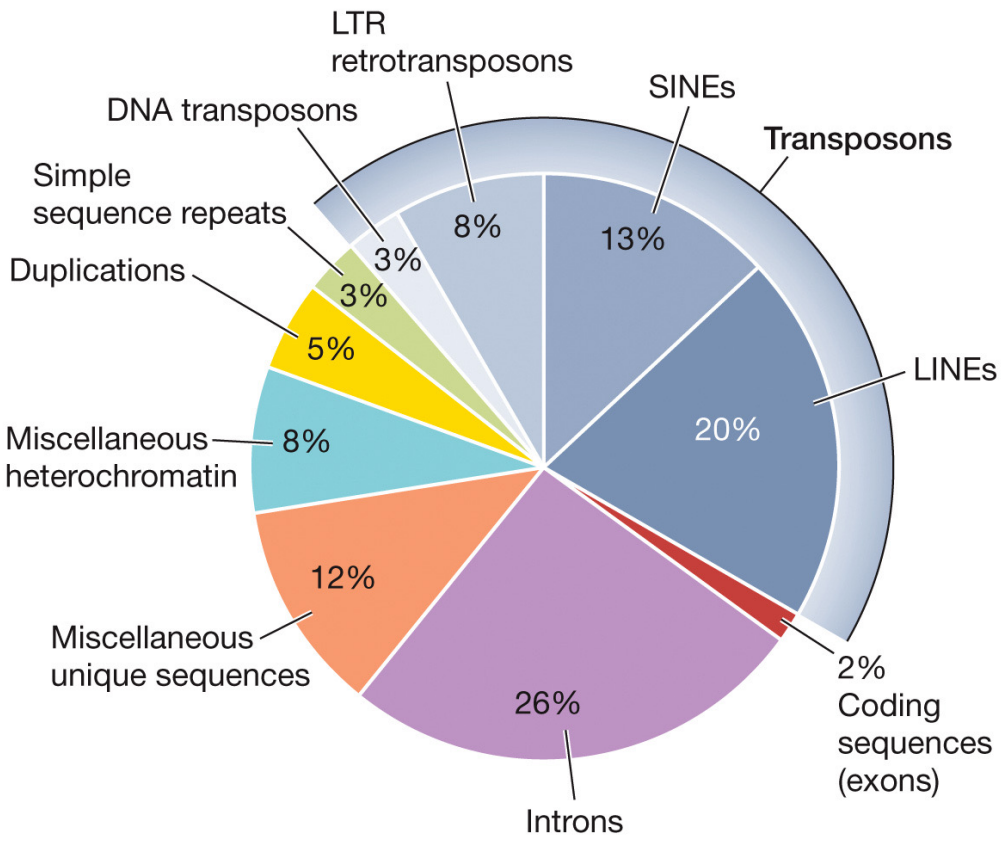
Figure from Swalla (*Heredity*, 2006)

Genes do not exist in isolation



- Genes are found on chromosomes
- DNA sequences that are close together on the same chromosome tend to be inherited together (**linkage**)
- Genes can interact with other genes (**epistasis**)

Genomic fractions



EVOLUTION 4e, Figure 14.14

- Most of the human genome is non-coding
- Transposons (selfish DNA) make up a large % of the human genome

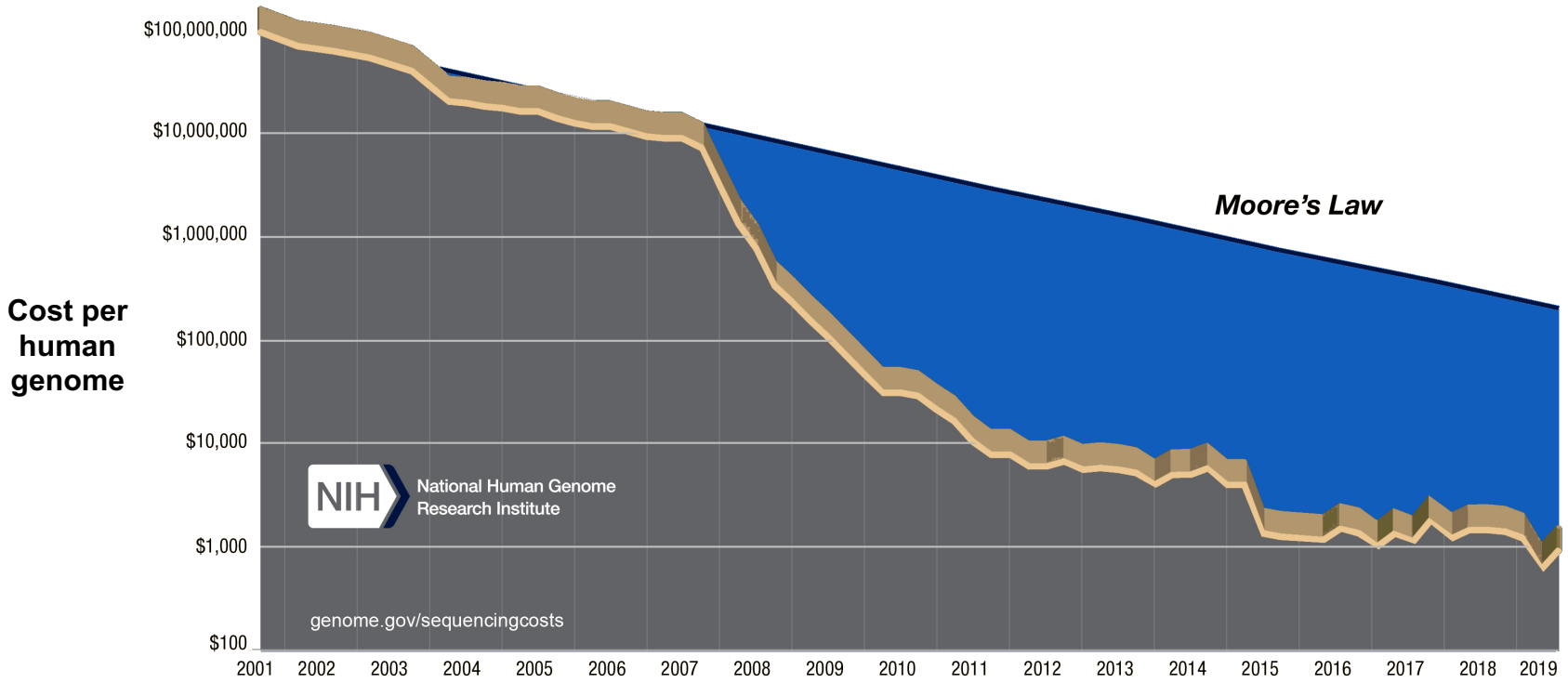
Whole genome sequencing (WGS)

- WGS is sometimes called next-generation sequencing
- Depth of coverage: average number of reads per base pair in a genome (low coverage = 5-10X, high coverage: >30X)
- One error per 100,000 base pairs (high coverage)
- Relatively expensive, but getting cheaper

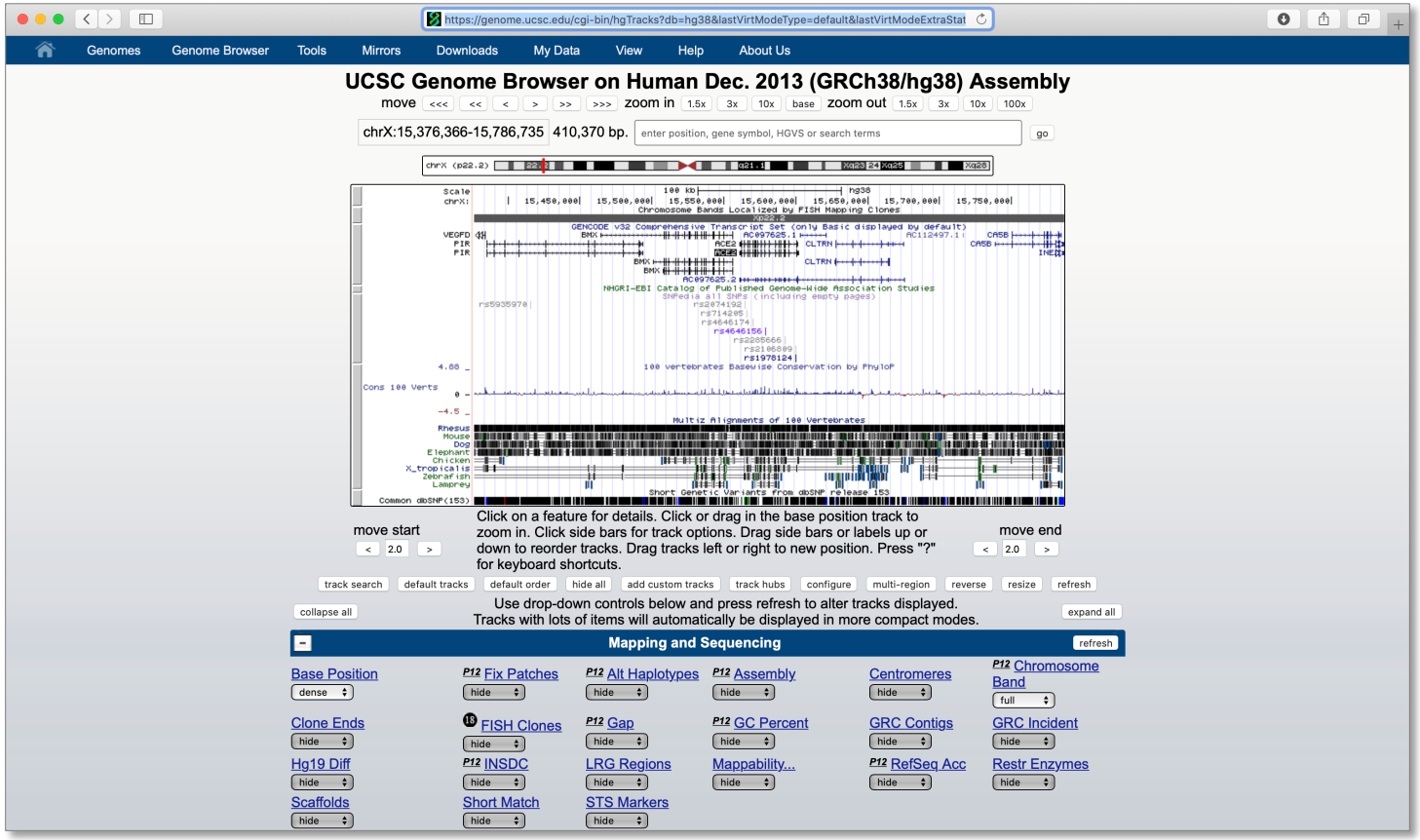
NovaSeq 6000

illumina

Declining sequencing costs

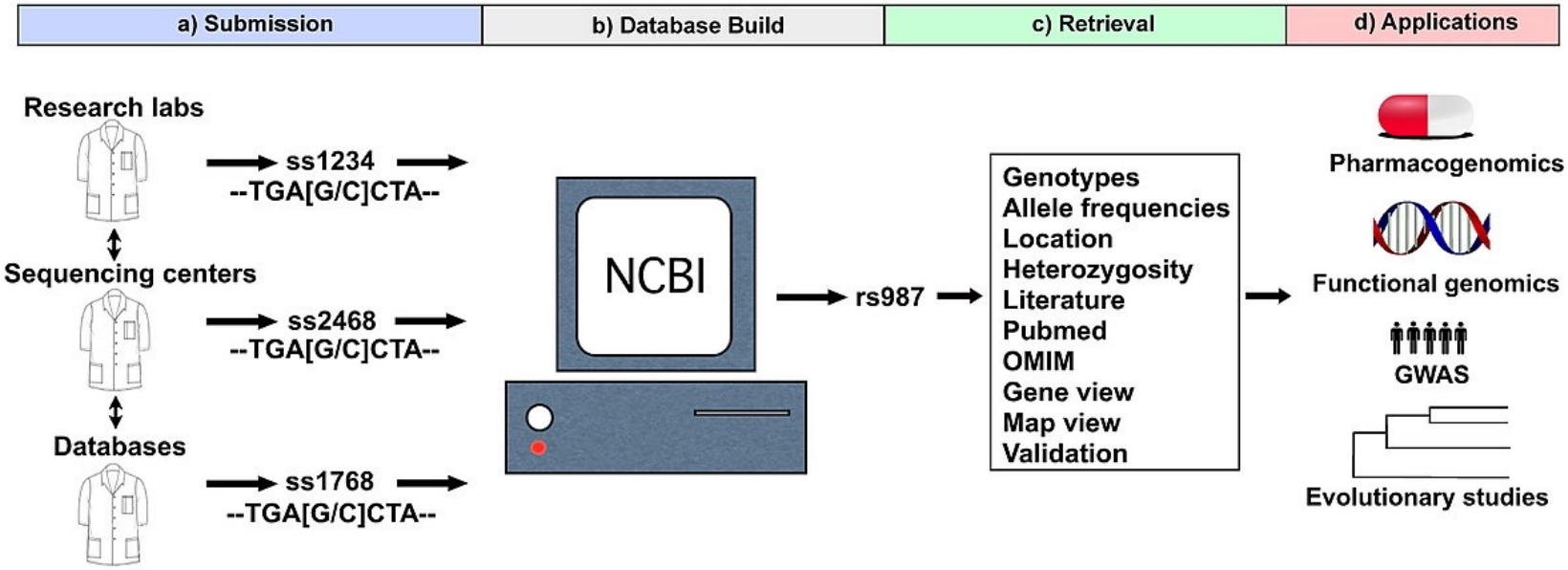


UCSC Genome Browser



- An online resource for exploring the human genome
- https://genome.ucsc.edu

dbSNP



- A comprehensive resource of known polymorphisms
- rs numbers refer to specific genetic variants
- https://www.ncbi.nlm.nih.gov/snp/

Environmental context matters

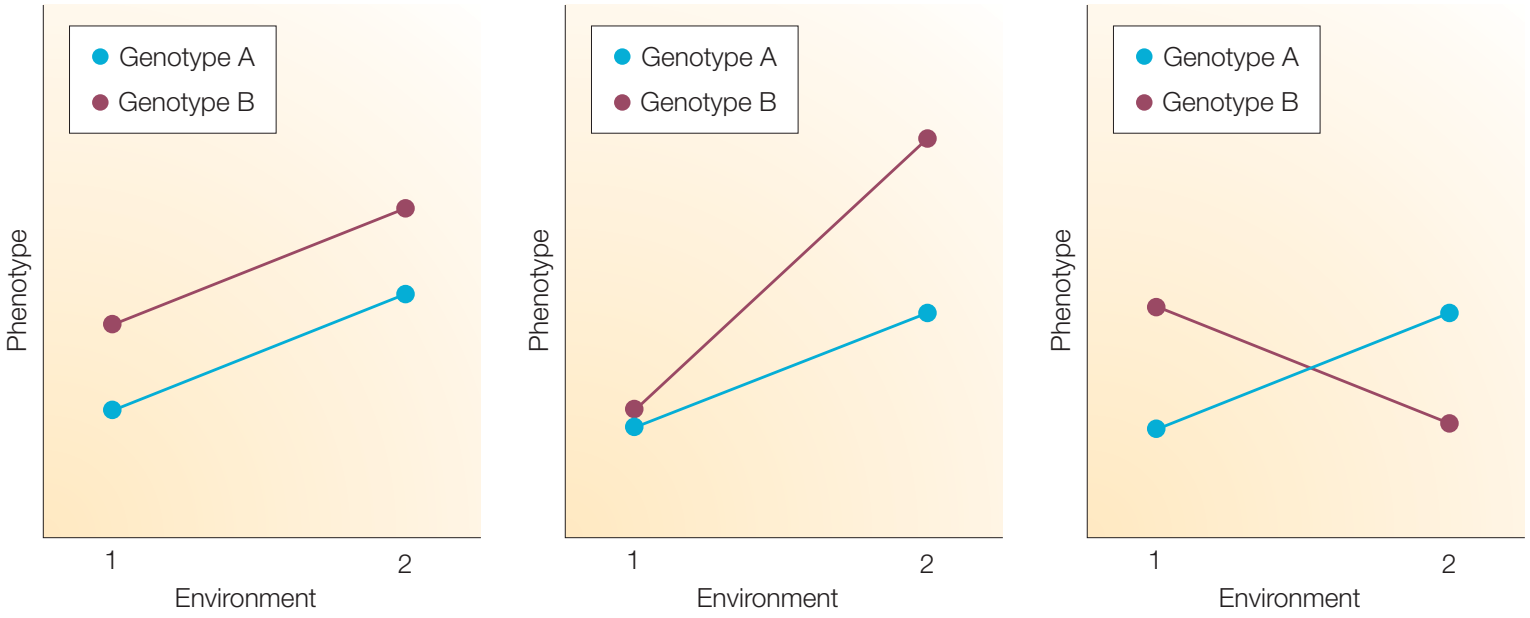
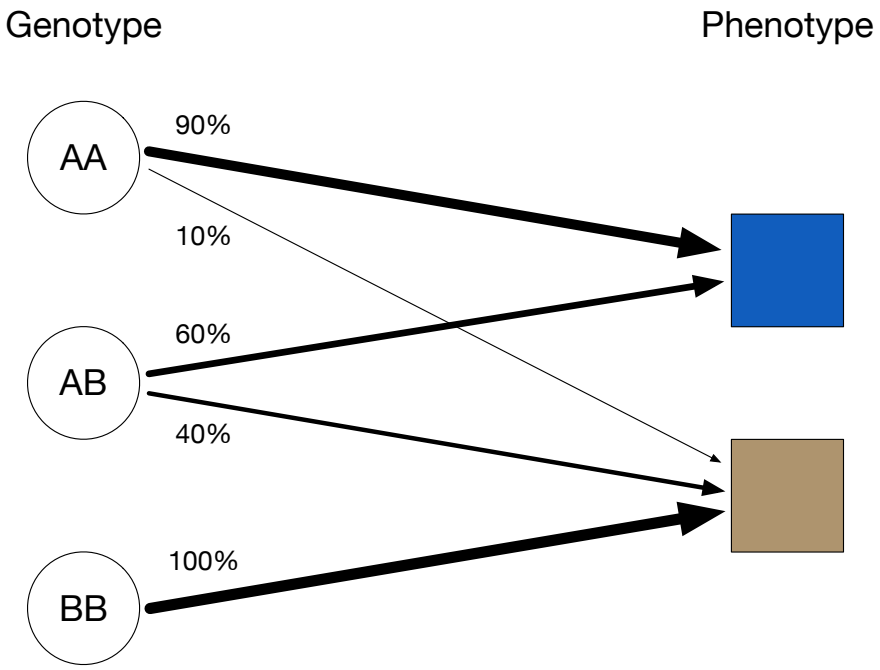


Image from Anholt and Mackay (*Nature Reviews Genetics*, 2004)

Incomplete penetrance



- Genotype-phenotype maps are not always one-to-one (e.g., some alleles increase your chances of getting hypertension)
- **Penetrance** refers to the proportion of individuals with a given genotype that show the expected phenotype

Pleiotropy

- Be careful to avoid terminology like "cancer gene" or "height gene"
- This is because genes often contribute to multiple phenotypes (i.e., they are **pleiotropic**)
- Example: A mutation in the *Frizzle* gene results in feathers that curve outward, fewer eggs laid, and high body temperatures

Cis and trans in genetics

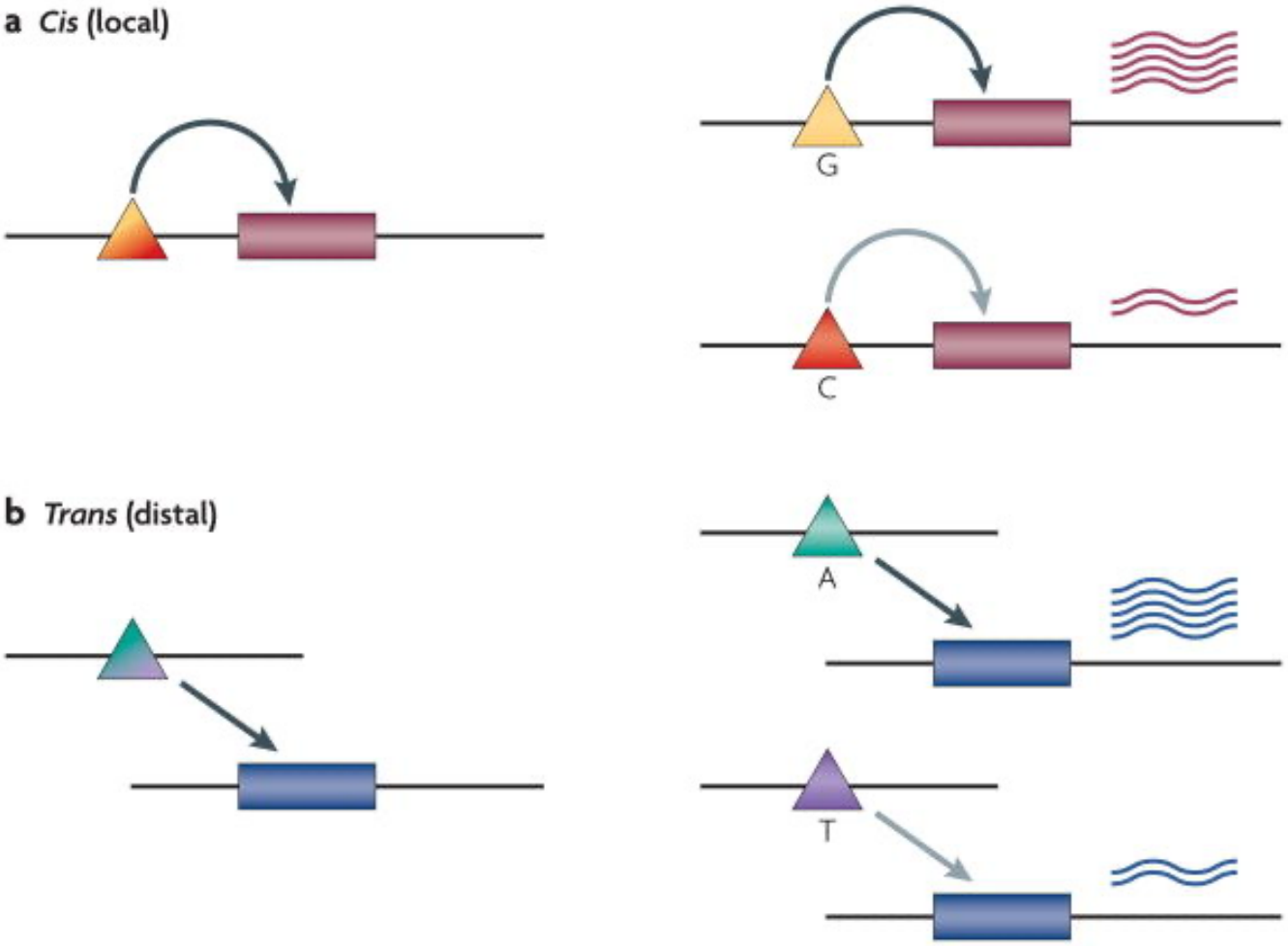
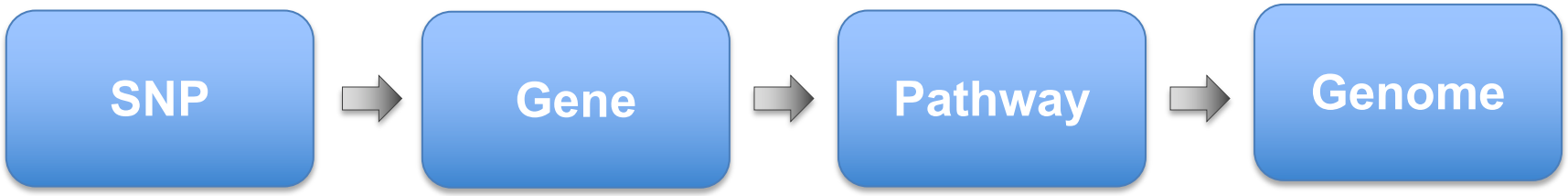


Figure from Cheung and Spielman Nat Rev Genet, 2009)

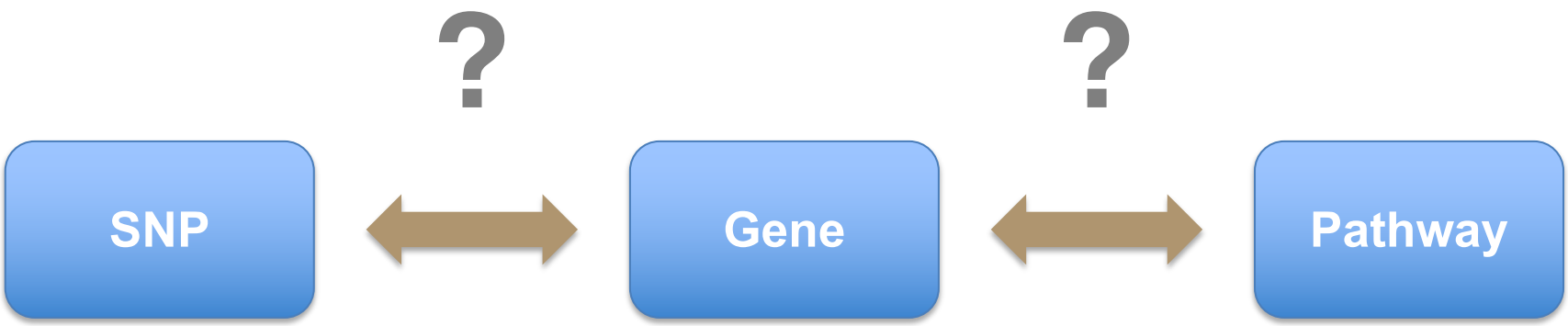
Units of analysis in genetics



- Genetic data be analyzed on **population** as well as **individual** scales

- Sometimes we are more focused on **traits**...

The mapping problem



- SNPs can affect more than one gene (or none)
- Gene to pathway mappings are hindered by ascertainment bias
- Mappings need not be one-to-one