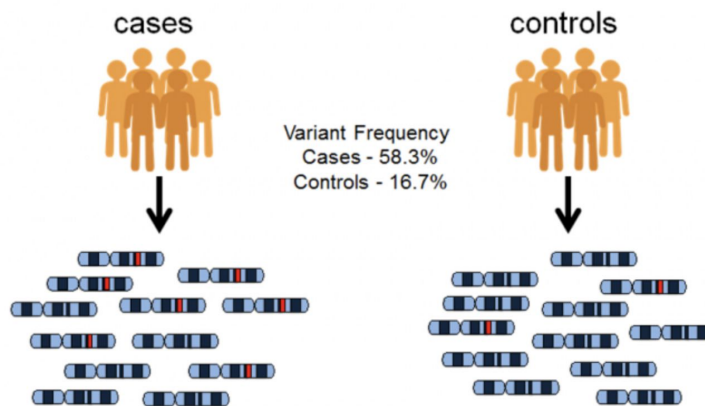


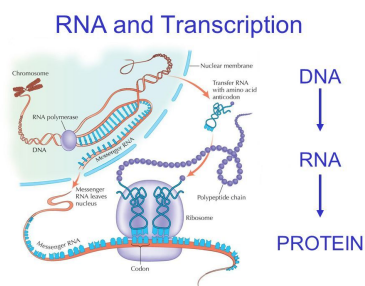
2/26/19: Intro to dataset

Things to remember from high school biology:

- DNA is composed of nucleotides, around 3 billion!
- Humans have variations in their nucleotides (mutations) that make them unique
- Some genetic variants can cause disease
- If a genetic variant is associated with a disease, more people will have the variant in their DNA compared to people who do not have the disease.
 - In the figure below, chromosomes are in blue and black stripes represent variations. The variation that is more frequent in cases compared to controls is highlighted in red, thus the red variant is *associated* with disease (note that we say *is associated with* and not *causes*, we can discuss this more when we meet).



- 1.5% of the genome is protein-coding genes.
- Review translation



- Main point: DNA codes for RNA which makes a protein. Now imagine a mutation in the DNA, the RNA is changed which could also change the protein.

Gene expression:

I'm trying to think back to teaching biology, and I don't think gene expression is covered. Basically, the amount of protein made from a gene can vary between people and between different cell types.

Think about mitosis (cell division), there are proteins that tell your cells when to divide (replicate), so if someone is growing, they probably need more of this protein so that their cells can replicate quickly and they can grow. So, when we compare the gene expression for the gene that makes the “replicate” protein, young children might have more than older adults. I can explain this better in class with a pen and paper.

Gene expression can also be different between tissue types. If there is a gene that helps your body send signals, the gene might be expressed higher in your brain than in your liver. If there is a gene that helps your tissue pulse, the gene might be higher expressed in your heart compared to your brain.

