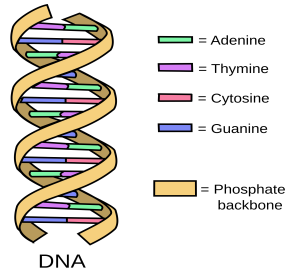# Kollel Genetics

1 Very Abbreviated History of DNA

1952: Hershey-Chase experiment define deoxyribonucleic acid (DNA) as genetic material

1953: DNA double helix structure described by Rosalind Franklin, James Watson, and Francis Crick

like baseball - 4 bases

1977: Sanger sequencing is developed, the foundation of current sequencing methods

1990 - 2003: Human Genome Project estimates about 25K genes

2009: Third generation long-read sequencing is introduced

2 Ancestry: Origin of Your DNA

Your cells contain **3** categories of DNA

1. Somatic: Chromosomal DNA

2. Sex Chromosomes: X & Y Chromosomes

3. Mitochondrial DNA: Only about 16K base pairs encoding 37 genes

There are **3** types of ancestry testing: Mitochondrial (Material), Y-Chromosome (Paternal), Somatic (Regional)
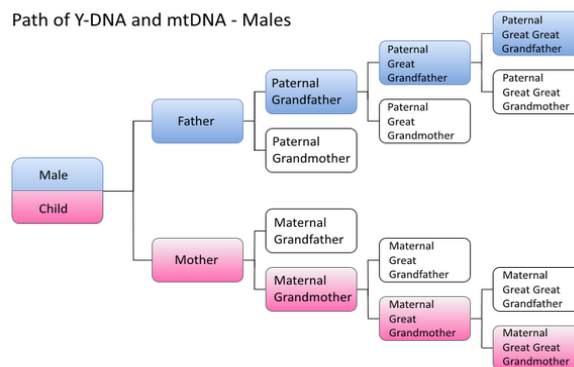

Short Tandem Repeats (STRs) are small, 2-7 base pairs, repeated units in DNA. STRs are typically use for recent genealogical information (i.e. determining extended familial relationships or identifying a suspect in a crime). For example, Cohan Modal Haplotype (CMH) is a pattern of 12 Y-STRs present primarily in the Kohanim.

Hammer, Michael F., et al. "Extended Y chromosome haplotypes resolve multiple and unique lineages of the Jewish priesthood." Human genetics. 2009

ATTCT CTCTCTCT AATTT

STRs are not used in Somatic testing commonly associated with DNA ancestry. Single Nucleotide Polymorphisms (SNPs) are point mutations in DNA strings compared to a genome reference. There are **a lot** of point mutations between groups.
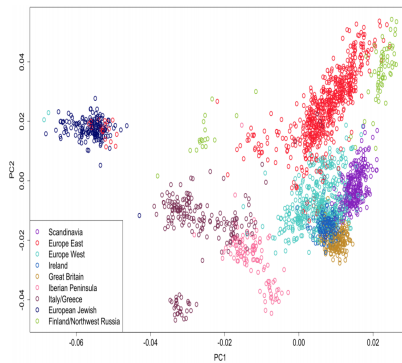
SNP testing is able to identify **3** branches common to Jewish lineages: Ashkinazi (Eastern Europe), Sephardic (North Africa), and Mizrahi (Middle East).
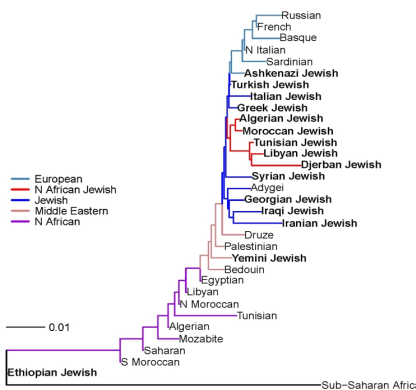
3 Genes Idenifith the Man?

Human identification in criminal and parental testing is done using STRs. 13 core STRs make up the Combined DNA Index System (CODIS). The probability of unrelated people matching at a single STR is 7.5%.

# of STR Matches	Probability
1	0.075
2	0.0056
3	0.00042
4	0.0000024
5	0.000000018

Chances of matching 9 STRs is 1 in 13 billion!. So do unrelated people actually match? Yes, due to the amount of pairwise comparisons! Also, we are looking for any 9 matching STRs, which opens up 715 different combinations. If we compare a database of 65K people our equation is as follows

$$65000 * \frac{64999}{2} * 715 > 1.5 \ trillon$$

$$1.5 \ trillon/13 \ billion = 116$$

Therefore, we would expect **116 unrelated people to share 9 STRs**.

4 You as Defined by Merriam-Webster

You are genetically unique and thus **Genetically Identifiable**. The Genetic Information Nondiscrimination Act (GINA) passed in 2008 to protect from genetic discrimination in health insurance and employment. Sequencing information provided by a close relative can be used to partially identify you. DNA is very stable and can be stored for long periods of time (long after your lifespan).

Open Questions:

Does your genetic information define your identity?

Are STRs enough for genetic identification?

Who should be allowed to see your genetic information?