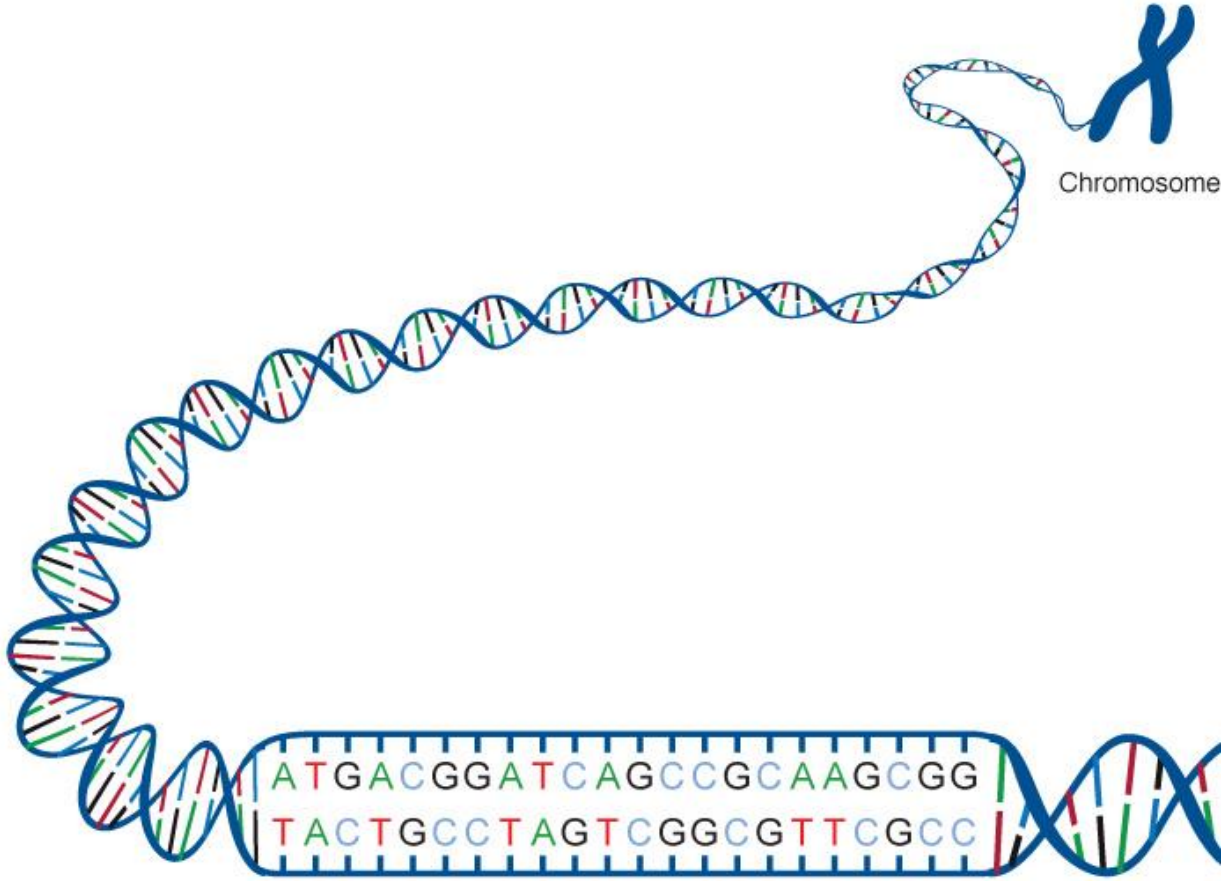
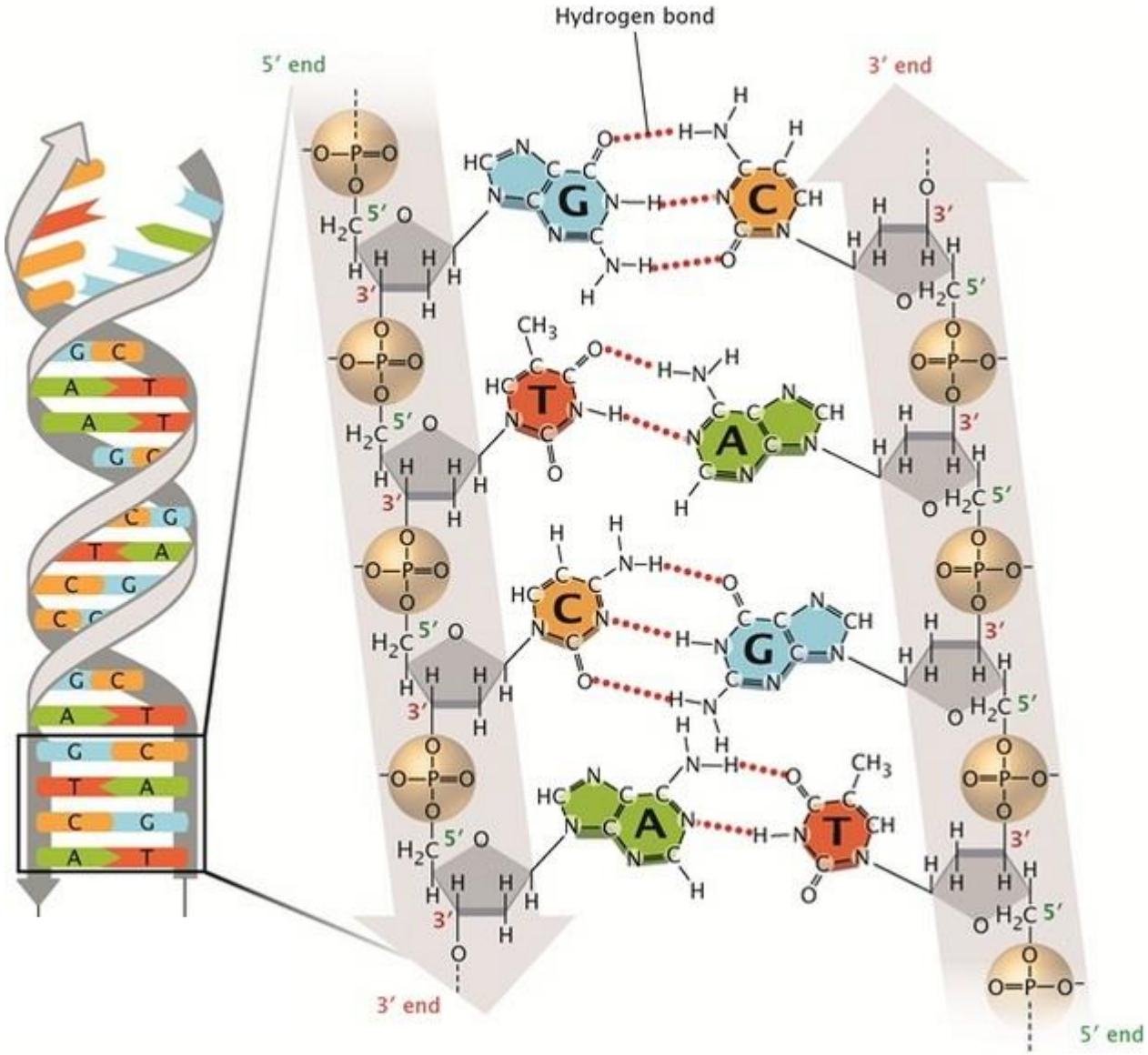
Measuring the Genome

Katrina Grasby

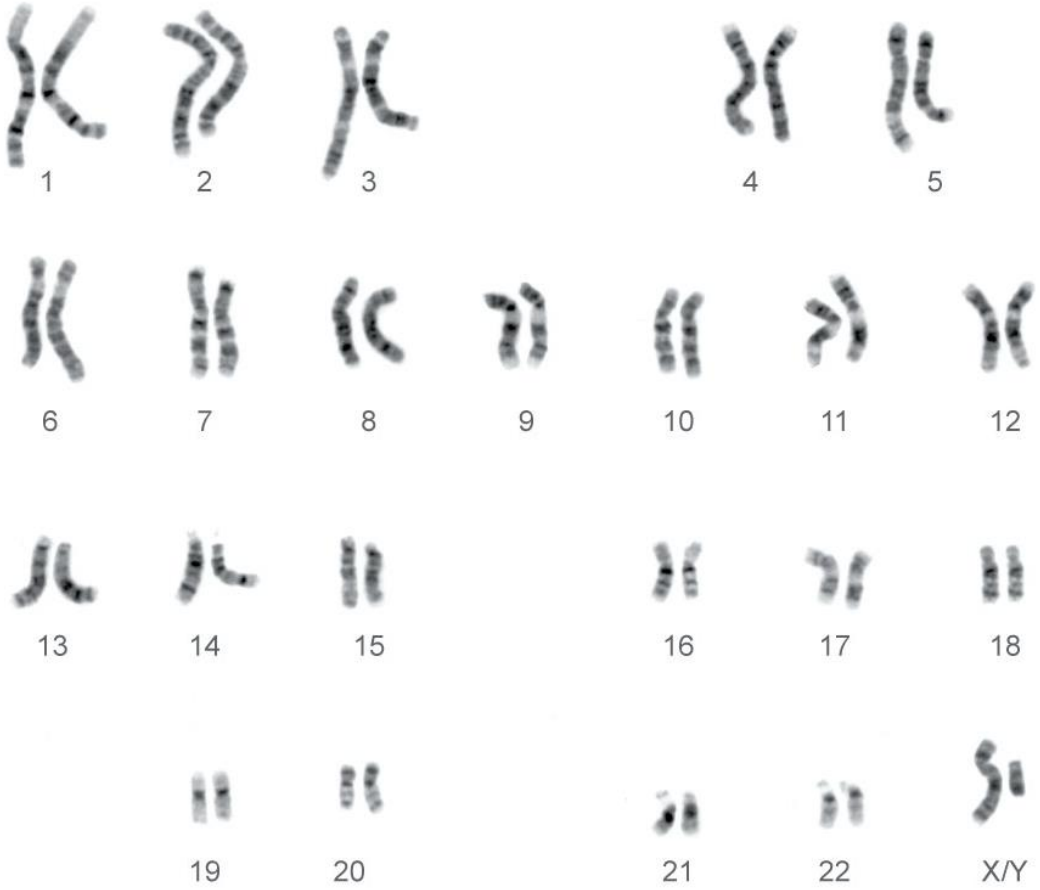


adenine (A), thymine (T), cytosine (C), guanine (G)

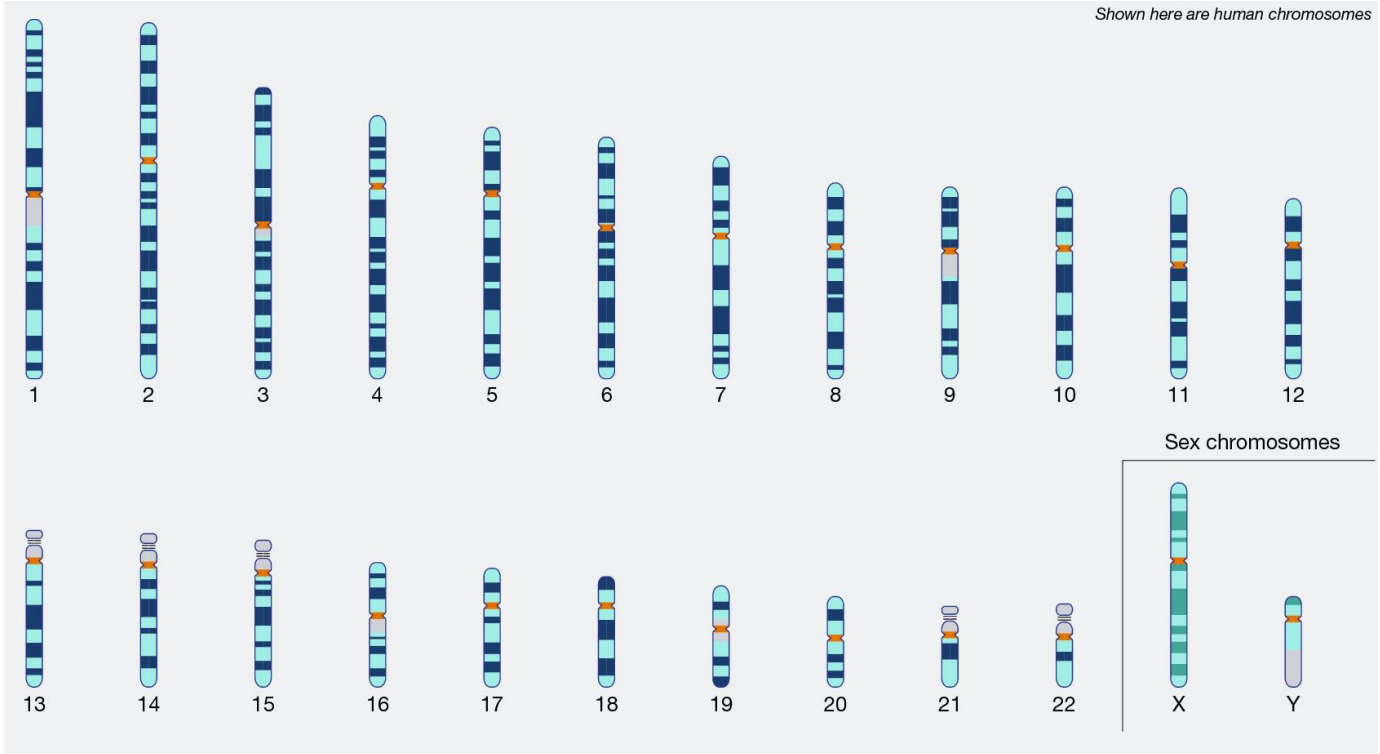
DNA: deoxyribonucleic acid

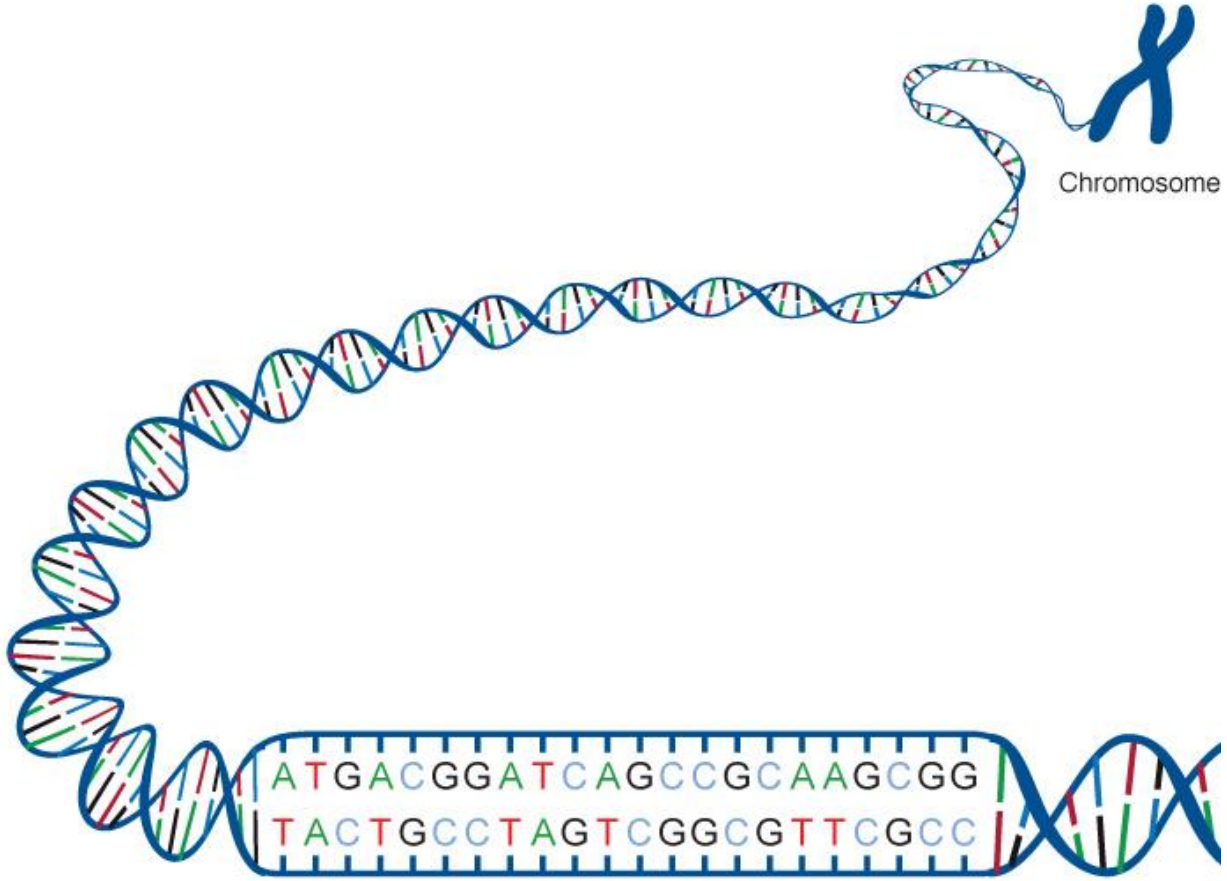


Diploid

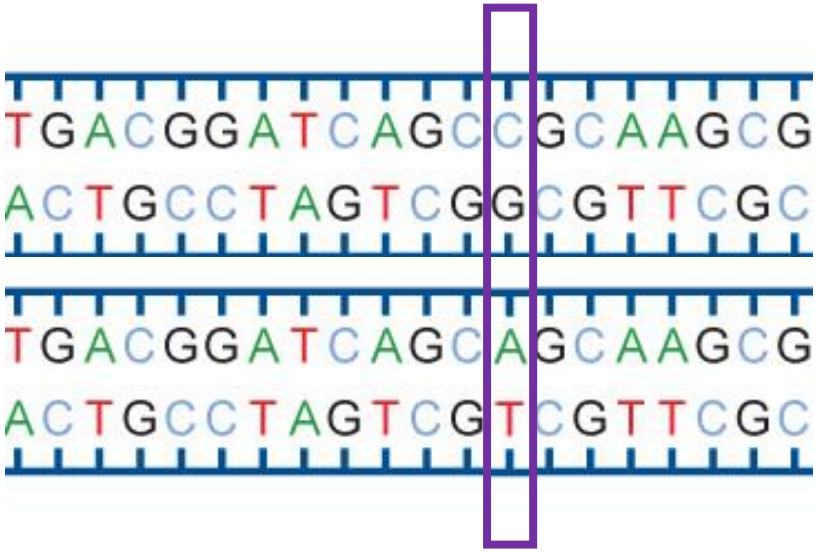


Haploid





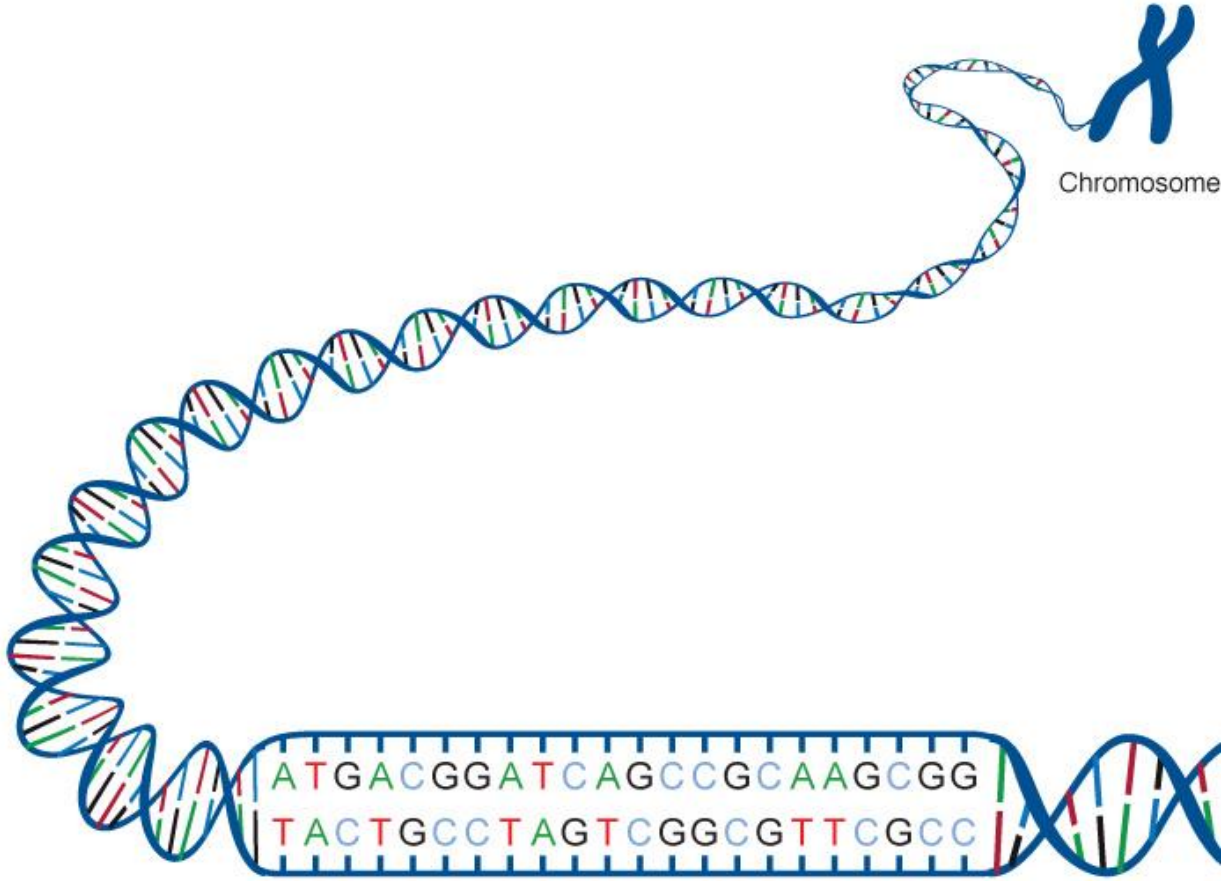
adenine (A), thymine (T), cytosine (C), guanine (G)



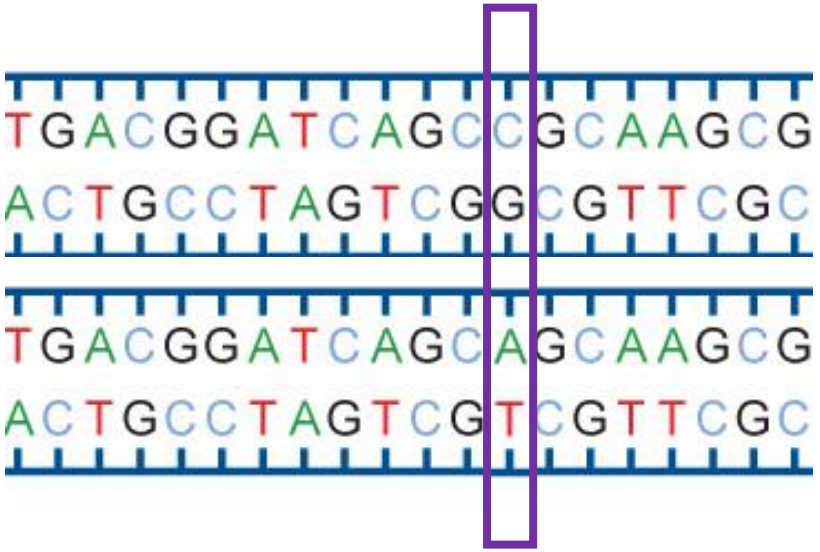
Person 1

Person 2

Genetic variation: differences in the sequence of DNA among individuals.

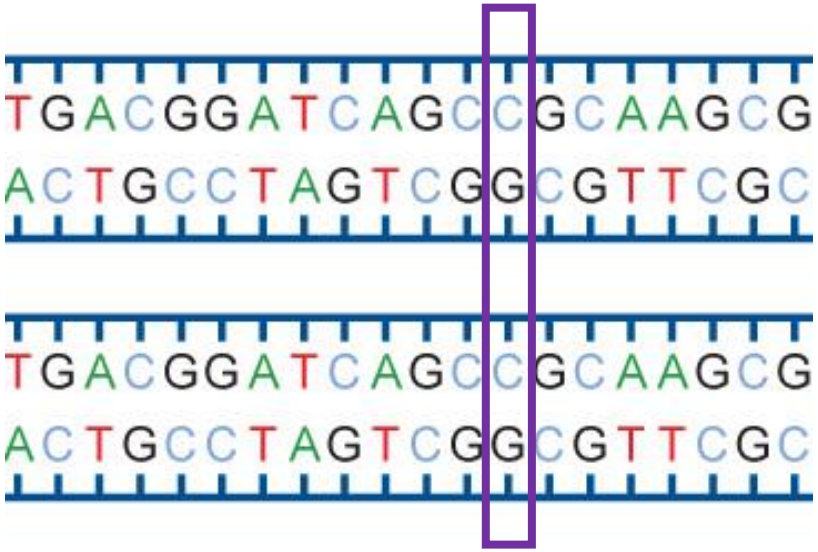


adenine (A), thymine (T), cytosine (C), guanine (G)



Genetic variant
aka locus

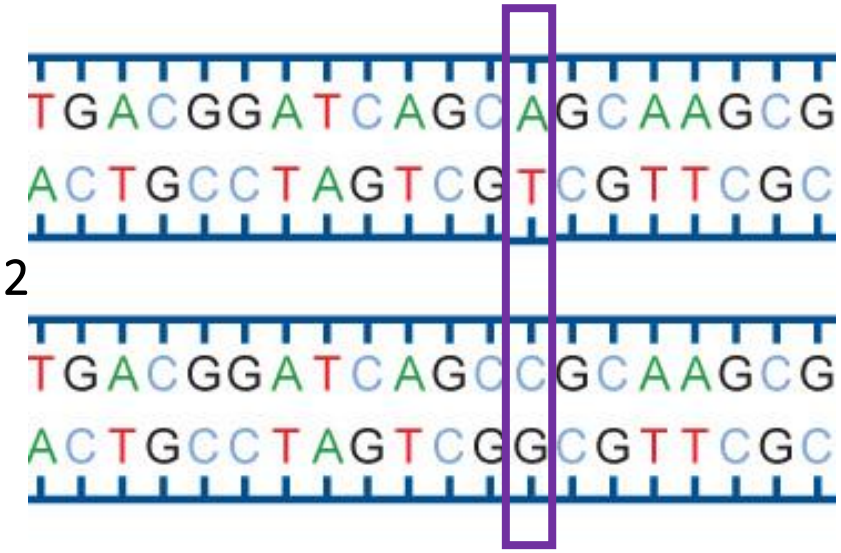
Person 1



Maternal

Paternal

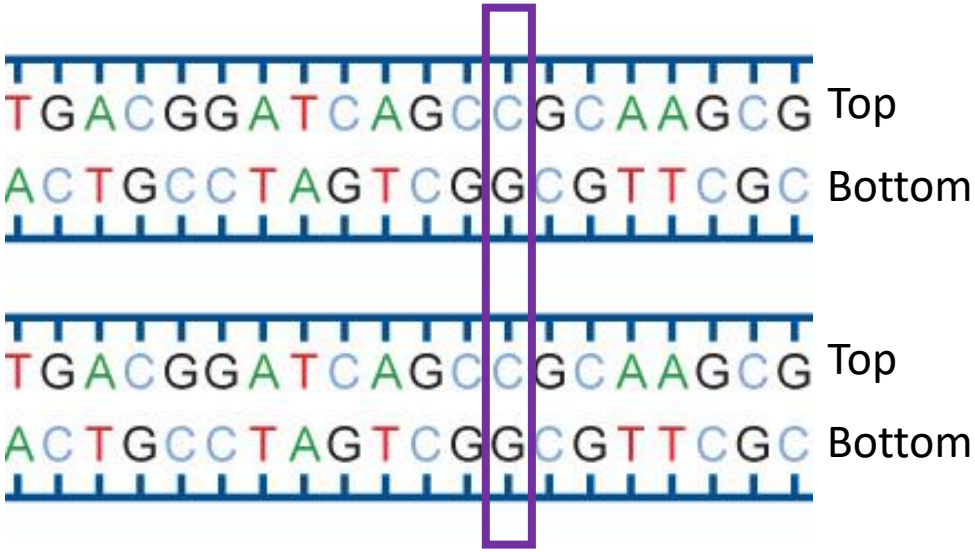
Person 2



Maternal

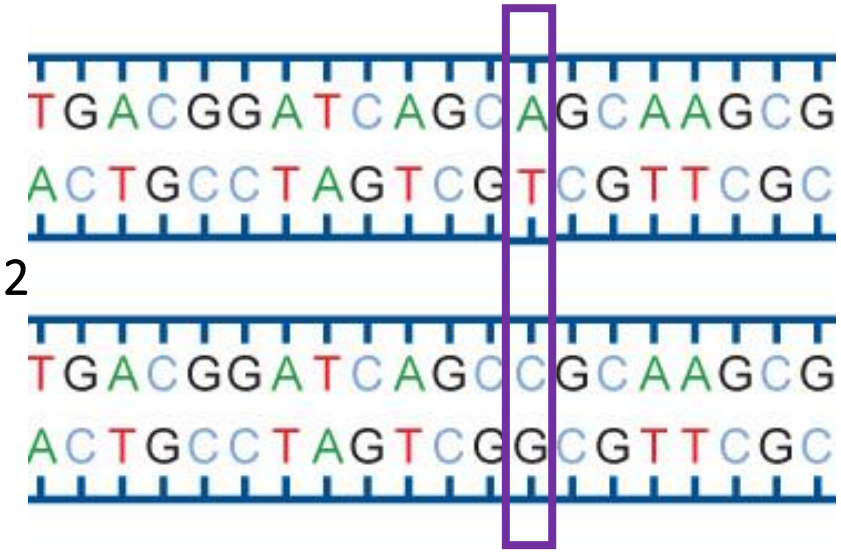
Paternal

Person 1



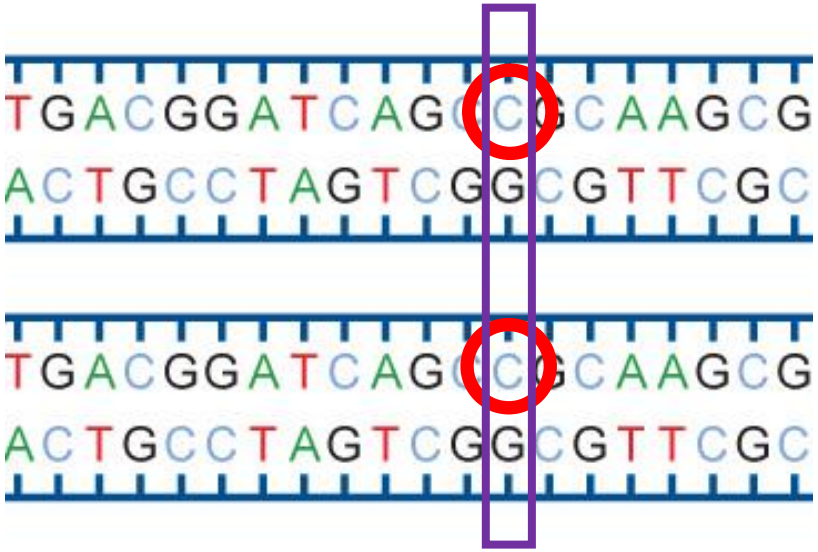
Top
Bottom
Top
Bottom

Person 2

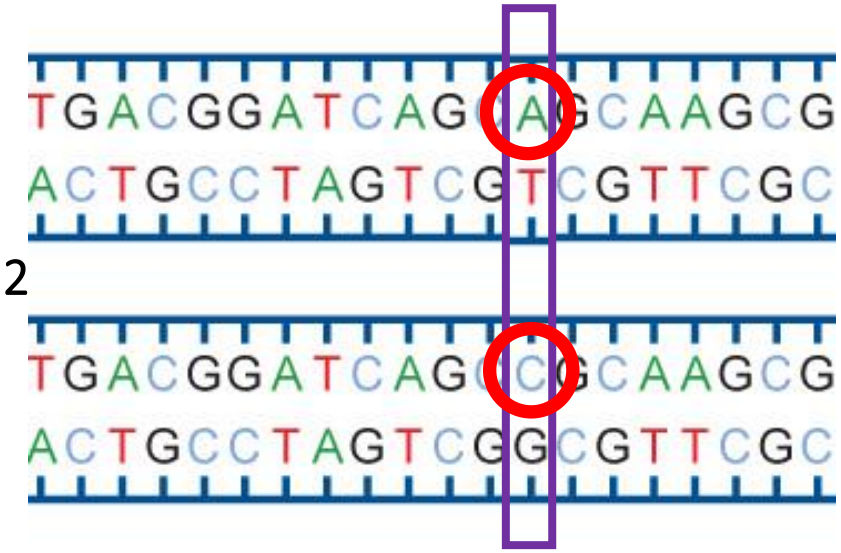


Top
Bottom
Top
Bottom

Person 1



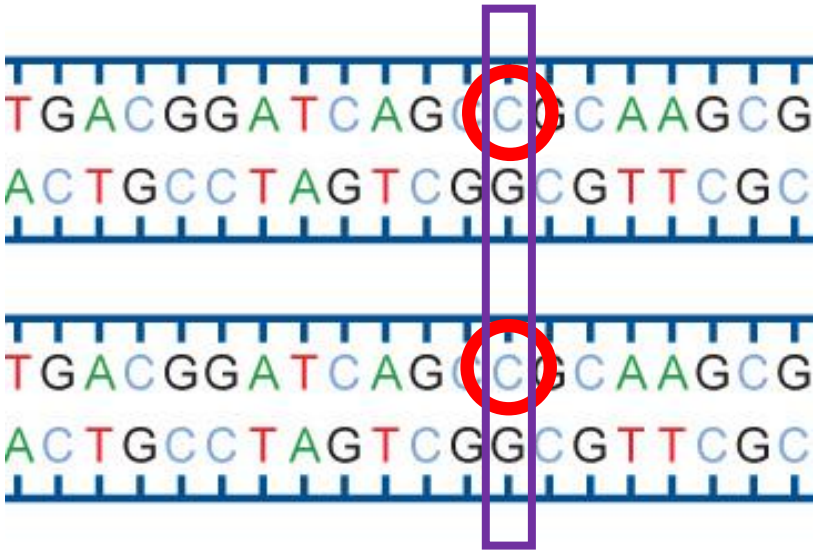
Person 2



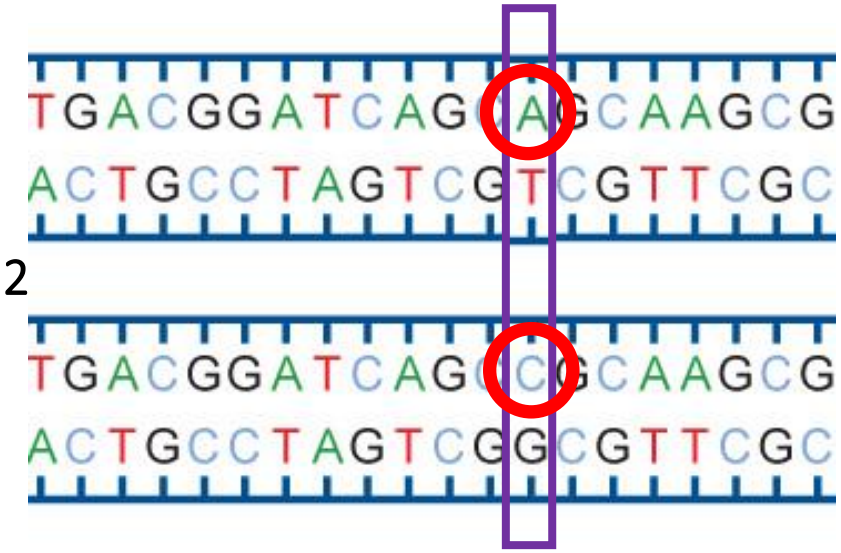
Alleles

A or C at this locus

Person 1



Person 2



Alleles

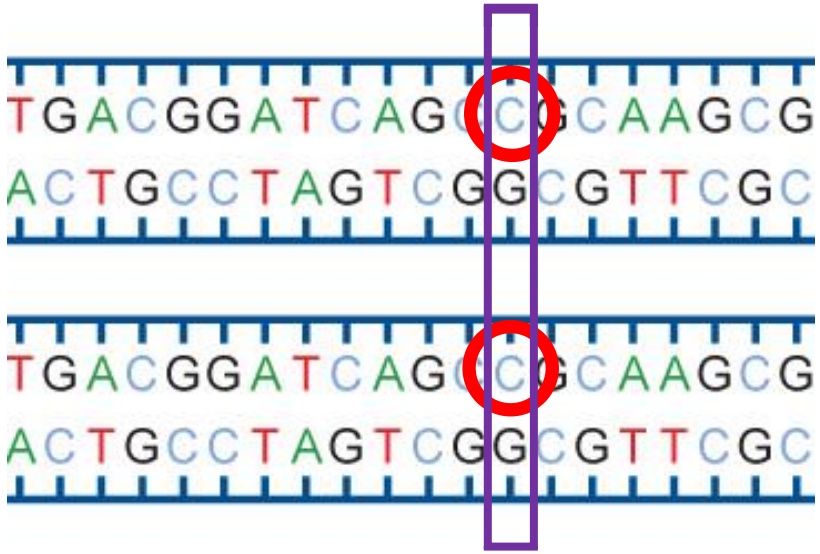
A or C at this locus

Genotypes

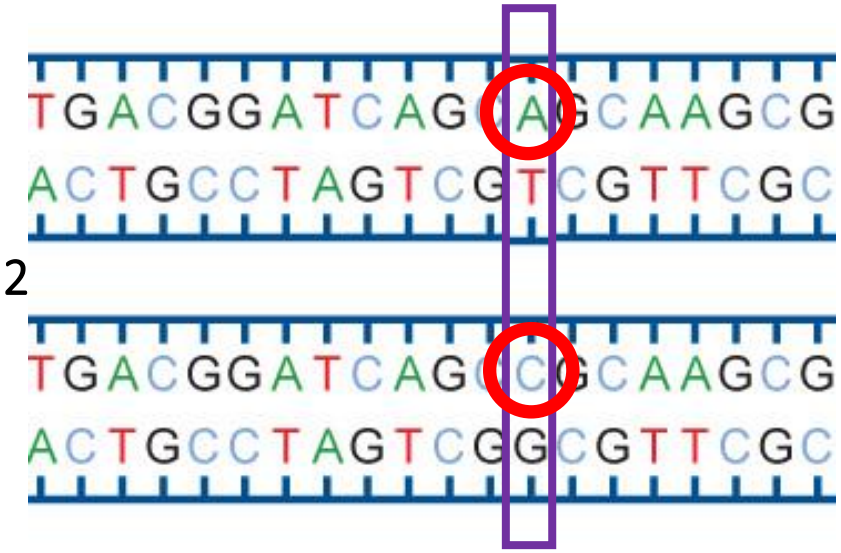
Person 1 = C C

Person 2 = A C

Person 1



Person 2



Alleles

A or C at this locus

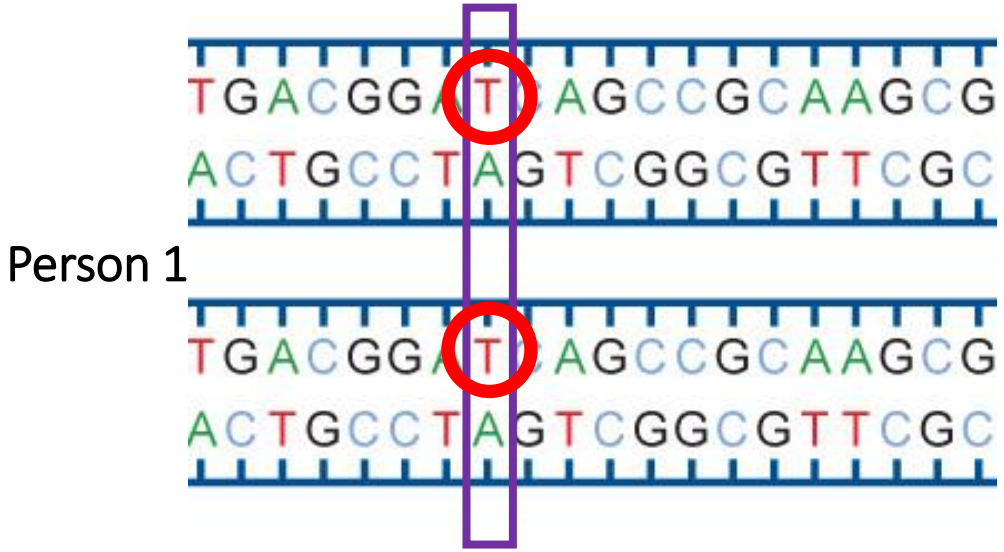
Genotypes

Person 1 = C C

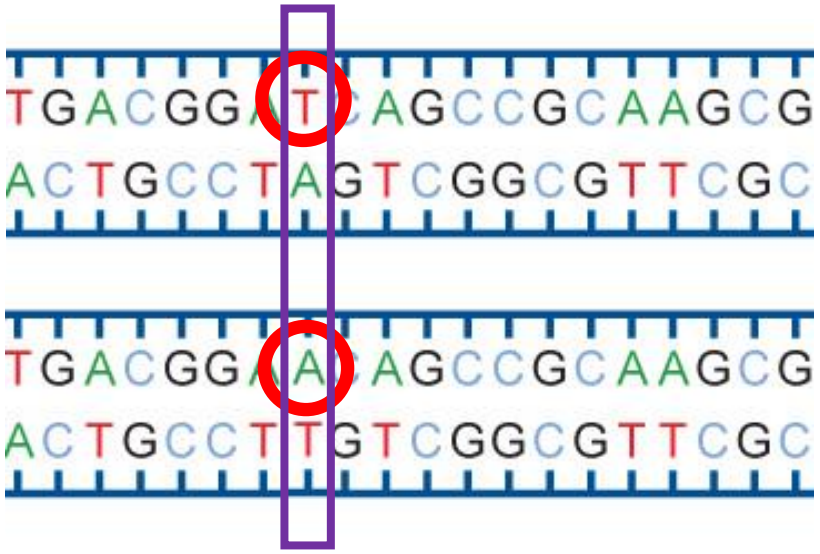
Person 2 = A C

Unambiguous Alleles

Person 1



Person 2



Alleles

A or T at this locus

Genotypes

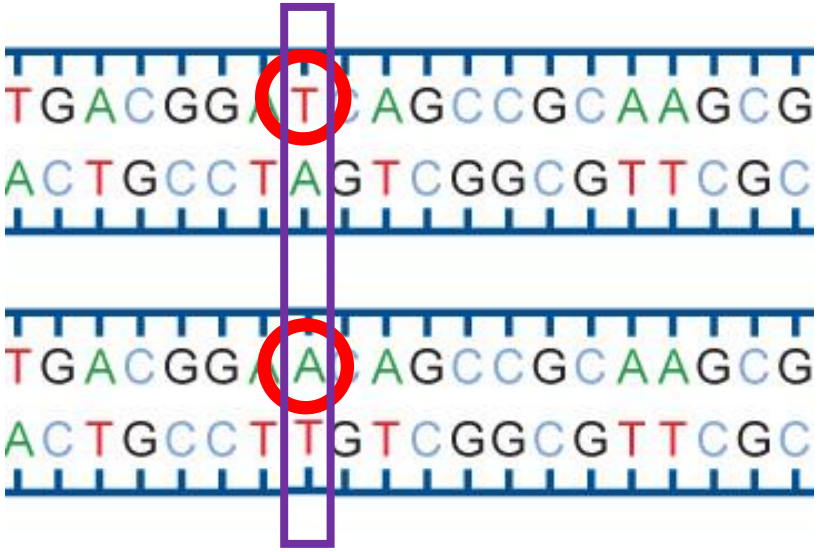
Person 1 = T T

Person 2 = T A

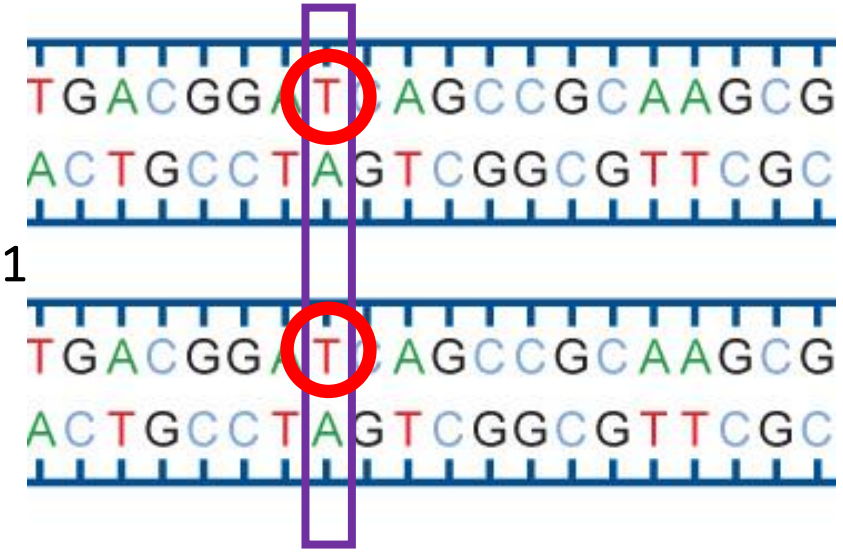
Ambiguous Alleles

Multiallelic

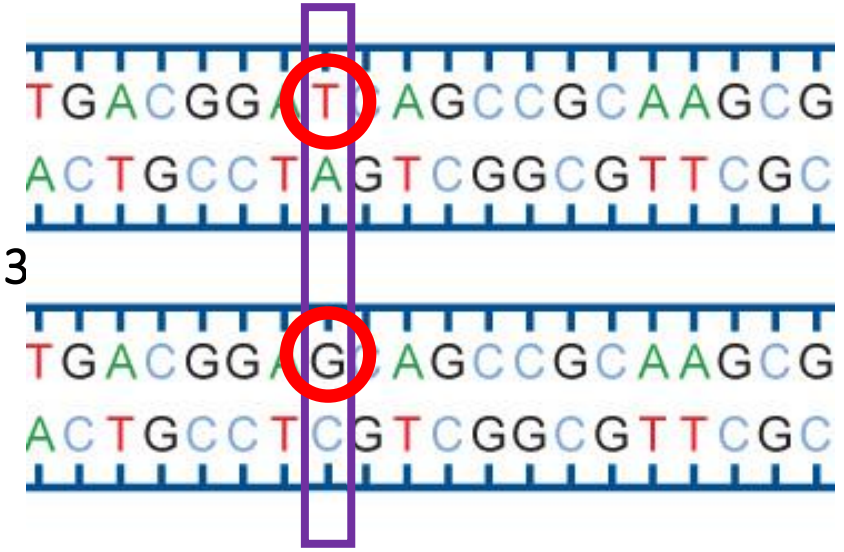
A, T or G at this locus



Person 2



Person 1



Person 3

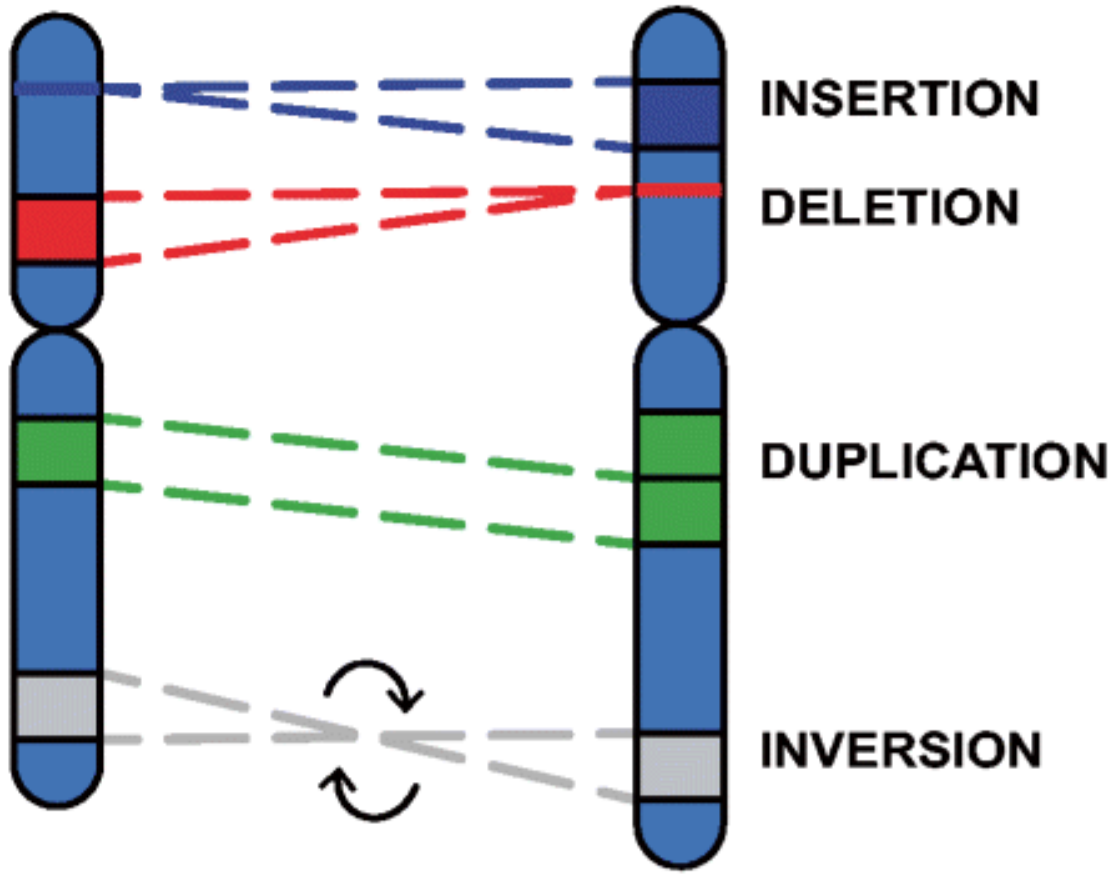
Genotypes:

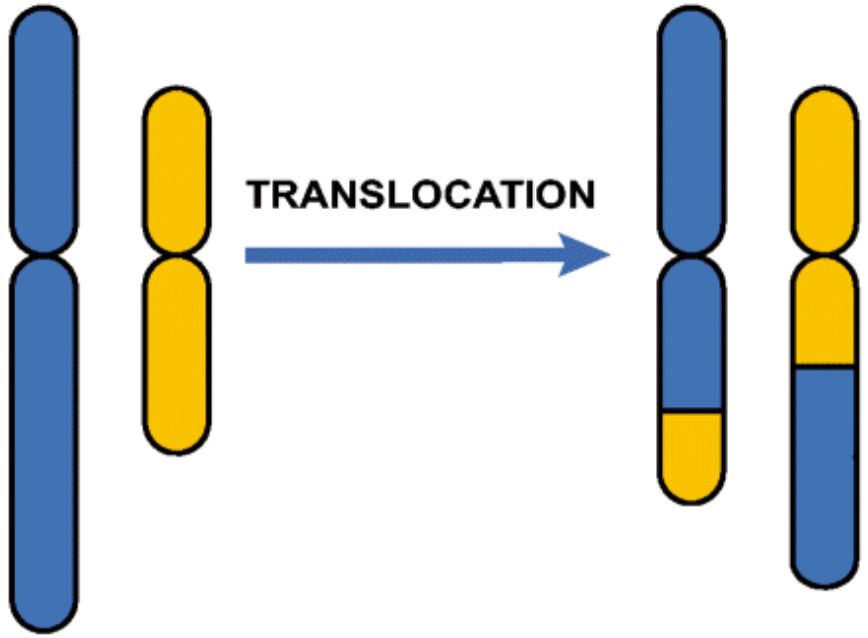
Person 1 = T T

Person 2 = T A

Person 3 = T G

SNPs, INDELs, and Structural Variation





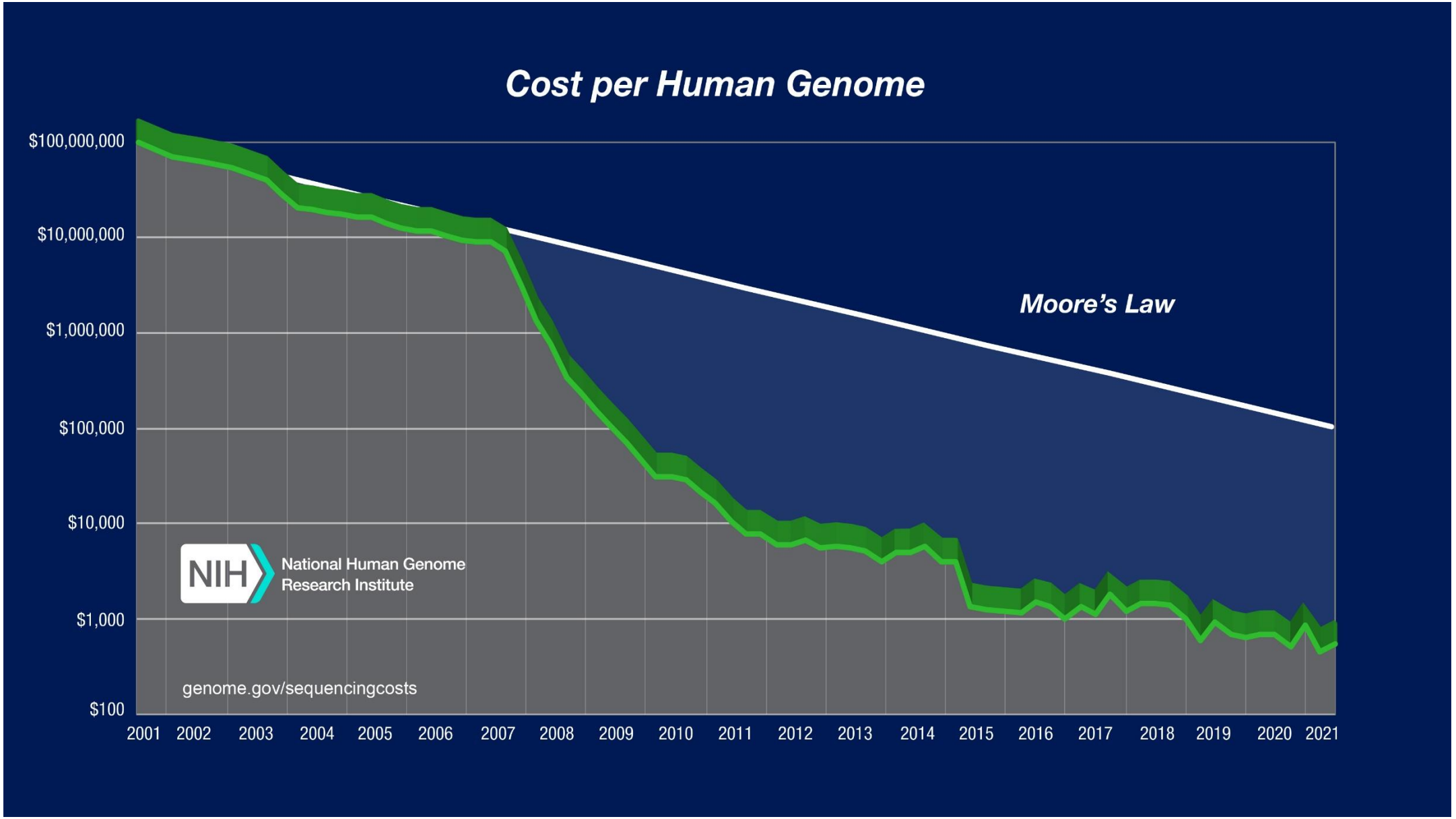
Reference

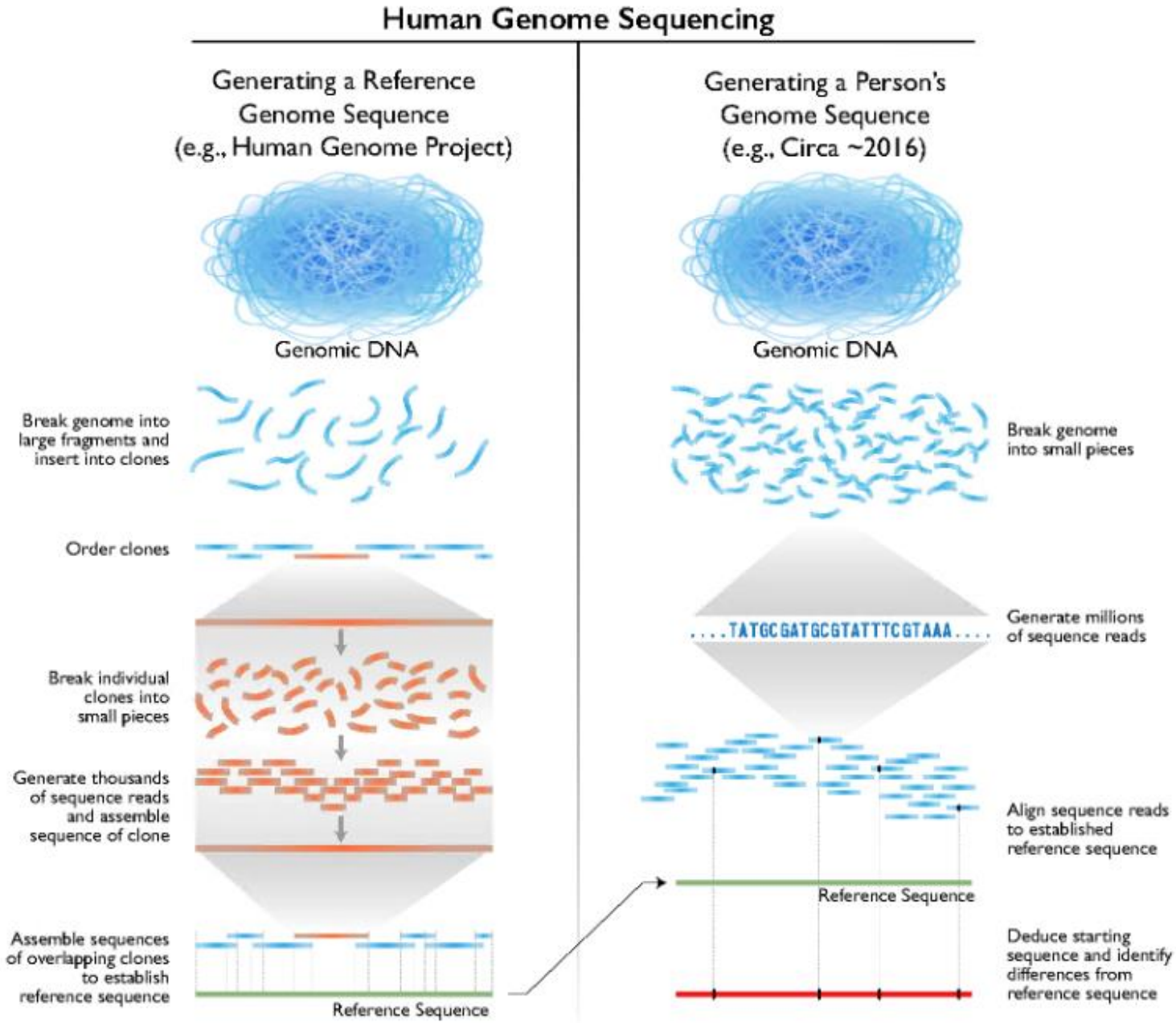
Balachandran & Beck, 2020

Common variant: a genetic variant present in >= 1% of the alleles in the population

Rare variant: a genetic variant present in < 1% of the alleles in the population

Note 1% is arbitrary


Cost per Human Genome
$100,000,000
$10,000,000
$1,000,000
$100,000
$10,000
$1,000
$100
Moore's Law
NIH
National Human Genome
Research Institute
genome.gov/sequencingcosts
2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021



https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost

From DNA to data