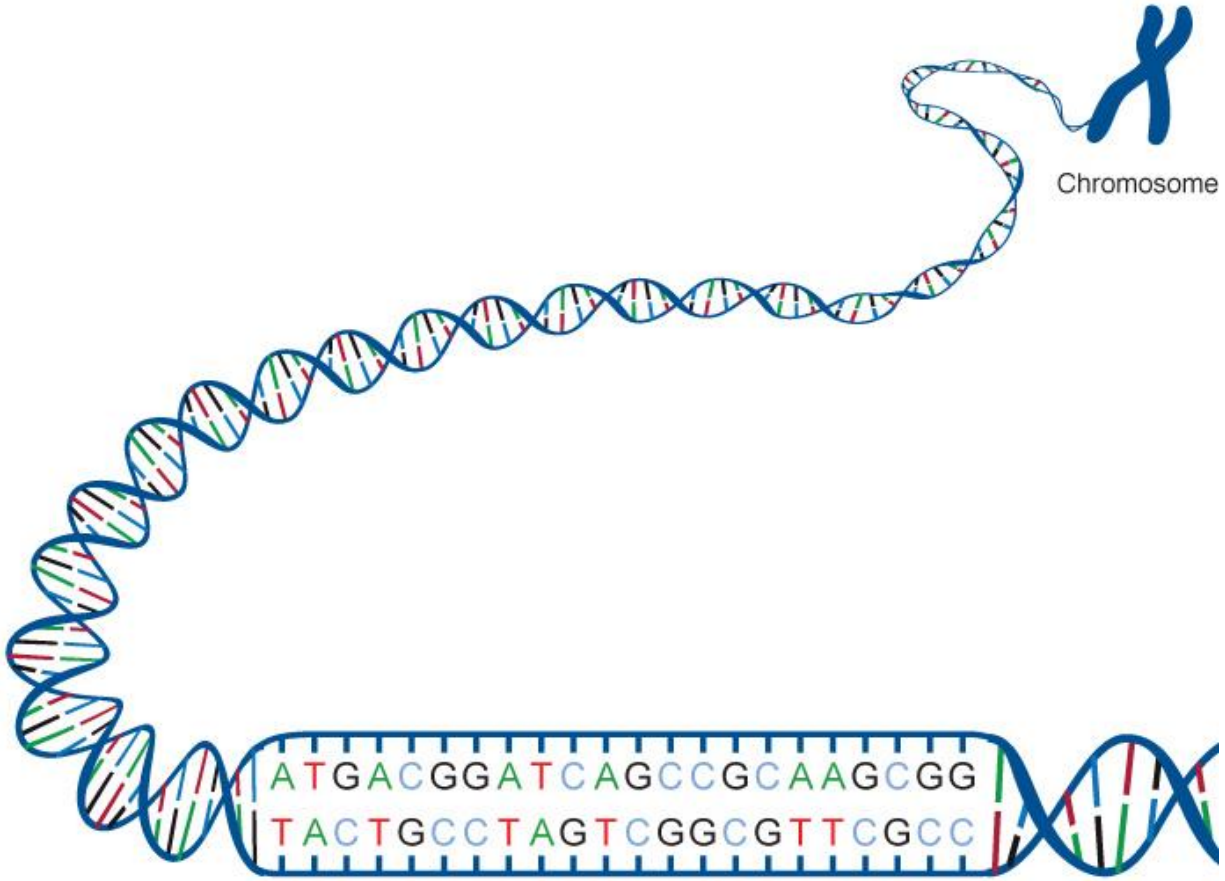
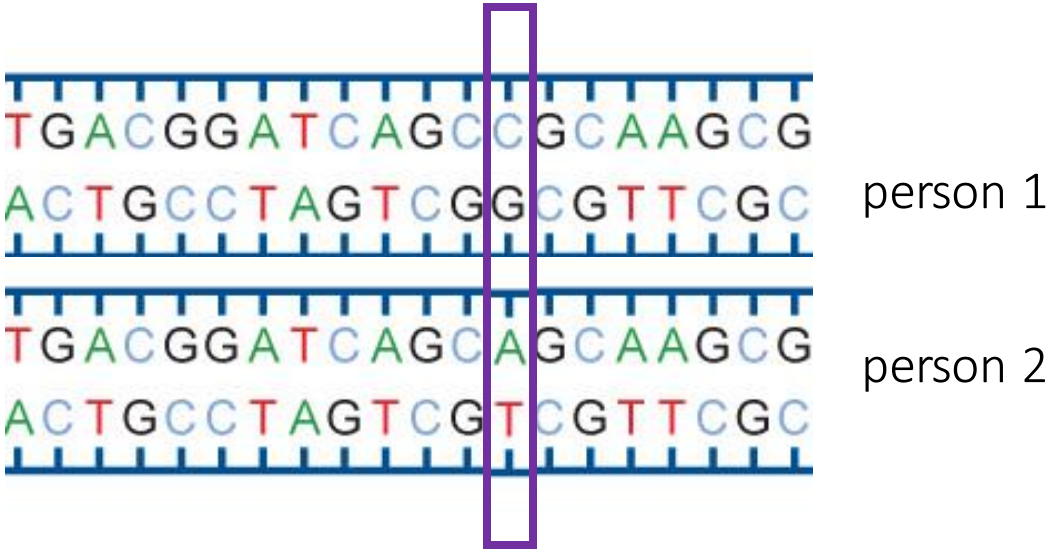
Introduction to common variation (I)

Lucía Colodro Conde and Katrina Grasby



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



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

Mutation: a newly arisen variant

Genetic variant: any specific region of the genome which differs between two genomes.

Allele: version of a variant

Allele frequency: incidence of an allele in a population.

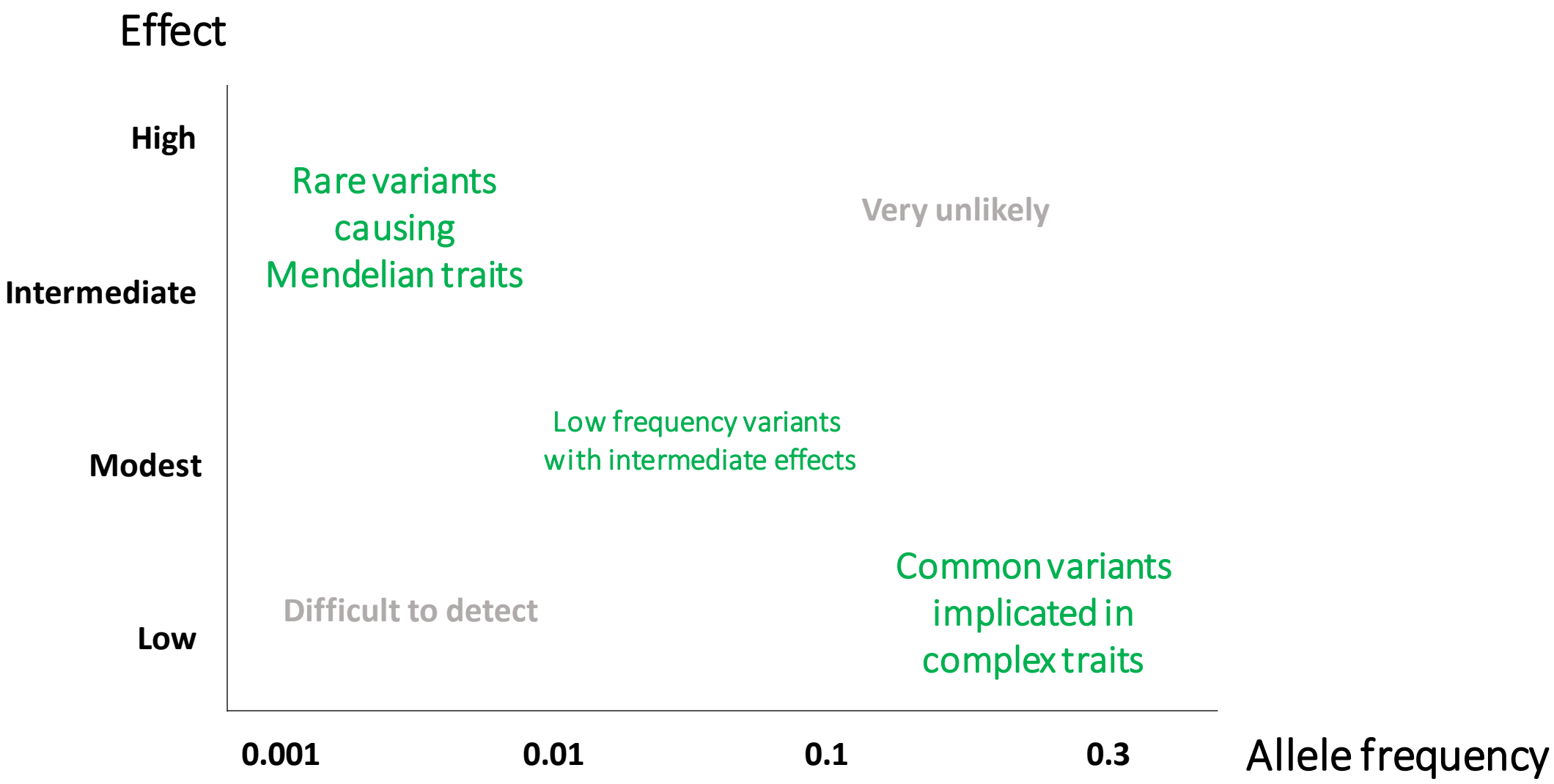
Minor allele frequency (MAF): frequency at which the less common allele occurs in a given population.

Minor allele count (MAC): number of times that allele appears over all individuals

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

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

Note 1% is arbitrary



Based on McCarthy et al (2008) Nature Reviews Genetics 9, 356-369

Examples of genetic variation

GWAS

Sequence variation

Single nucleotide
- substitutions
- insertions | 'indels'
- deletions

Structural variation

2bp to 1,000bp
- VNTRs: microsatellites, minisatellites
- indels
- inversions
- di-, tri-, tetranucleotide repeats

1kb to submicroscopic
- copy number variants
- segmental duplications
- inversions, translocations
- copy number variant regions
- microdeletions, microduplications

Microscopic to subchromosomal
- segmental aneusomy
- chromosomal deletions (losses)
- chromosomal insertions (gains)
- chromosomal inversions
- intrachromosomal translocations
- chromosomal abnormality
- heteromorphisms
- fragile sites

Whole chromosomal to whole genome
- interchromosomal translocations
- ring chromosomes, isochromosomes
- marker chromosomes
- aneuploidy
- aneusomy

Knight JC (2009). Genetics and the general physician: insights, applications and future challenge. *QJM*..

SNP (single nucleotide polymorphism):

variation at a single base pair in a DNA sequence among individuals.

	Chrom.	DNA sequence	Genotype
Person 1	Mat	GTAACTTGGGATCT**A**GACCAATAGAT	**A A**
	Pat	GTAACTTGGGATCT**A**GACCAATAGAT	
Person 2	Mat	GTAACTTGGGATCT**A**GACCAATAGAT	**A C**
	Pat	GTAACTTGGGATCT**C**GACCAATAGAT	
Person 3	Mat	GTAACTTGGGATCT**C**GACCAATAGAT	**C C**
	Pat	GTAACTTGGGATCT**C**GACCAATAGAT	

Slide based on Manuel Ferreira's

TYPES OF SNPs

Non-coding region

They may affect regulation of genes in coding regions

Coding region

Synonymous

They do not change the amino acid sequence

Non-Synonymous

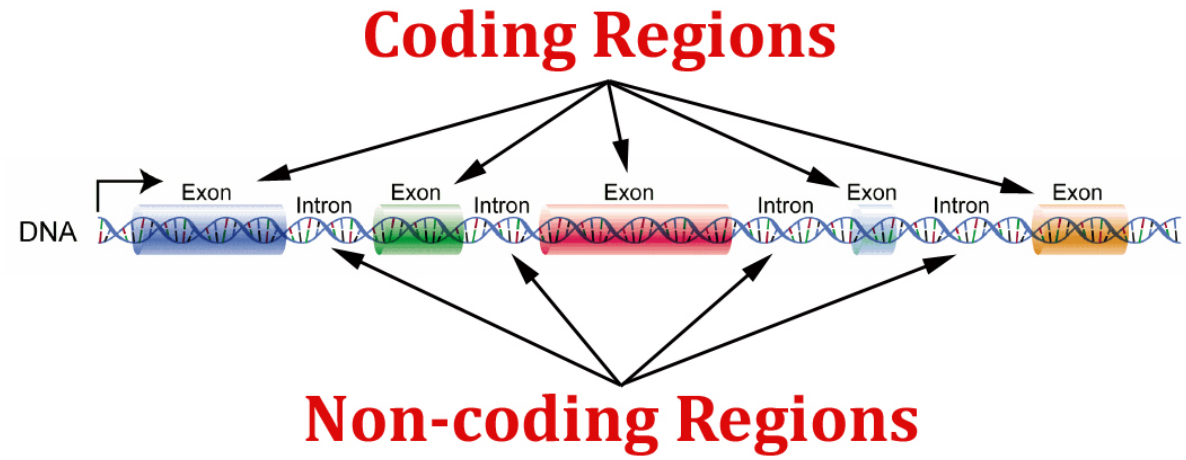
They change the amino acid sequence

Missense

They change one amino acid in a protein (which may have an effect in the protein)

Nonsense

They produce a premature stop codon, that results in an incomplete and usually non-functional protein product



Insertion–deletion variants (indels):

one or more base pairs are present in some genomes but absent in others in relation to the reference

	Chrom.	DNA sequence	Genotype
Person 1	Mat	GTAACTTGGGATCT**GAT**GACCAGATAG	**R D**
	Pat	GTAACTTGGGATCT**---**GACCAGATAG	
Person 2	Mat	GTAACTTGGGATCT**GAT**GACCAGATAG	**R R**
	Pat	GTAACTTGGGATCT**GAT**GACCAGATAG	
Person 3	Mat	GTAACTTGGGATCT**---**GACCAGATAG	**D D**
	Pat	GTAACTTGGGATCT**---**GACCAGATAG	

Introduction to common variation (I)

Lucía Colodro Conde and Katrina Grasby