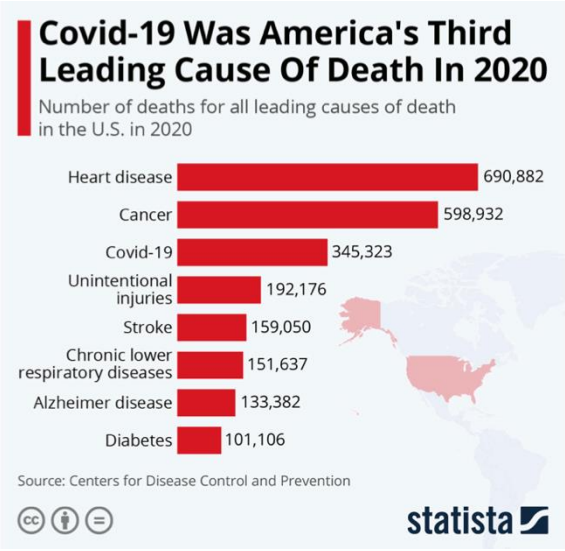
Lecture 16 scribing: cancer genomics

Name: Lau Sin Man

1.Definition: Cancer is a disease in which some of the body's cells grow uncontrollably and spread to other parts of the body.

2.Cancer is heterogeneous, which means mechanism and treatment for different cancers varies. Cancer is the second most severe disease in the US in terms of number of deaths (only lower than heart disease).



3.Genomic VS metabolic?

Cancer is generally believed to be a genomic disease, while some studies focus on metabolic approach.

4.Method to study cancer:

Variant calling (DNA/histone)

Association study/differential peak calling

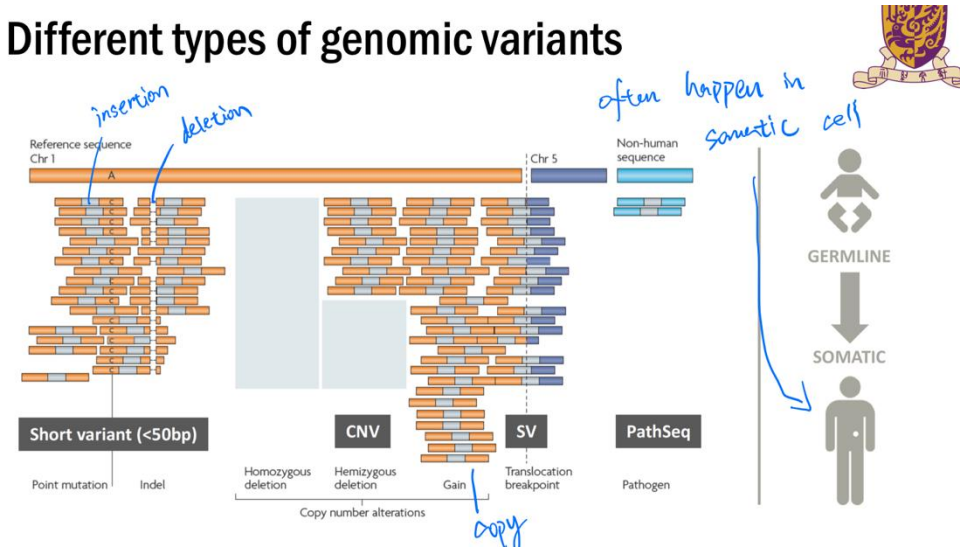
'Gatk' is a useful variant calling tool.

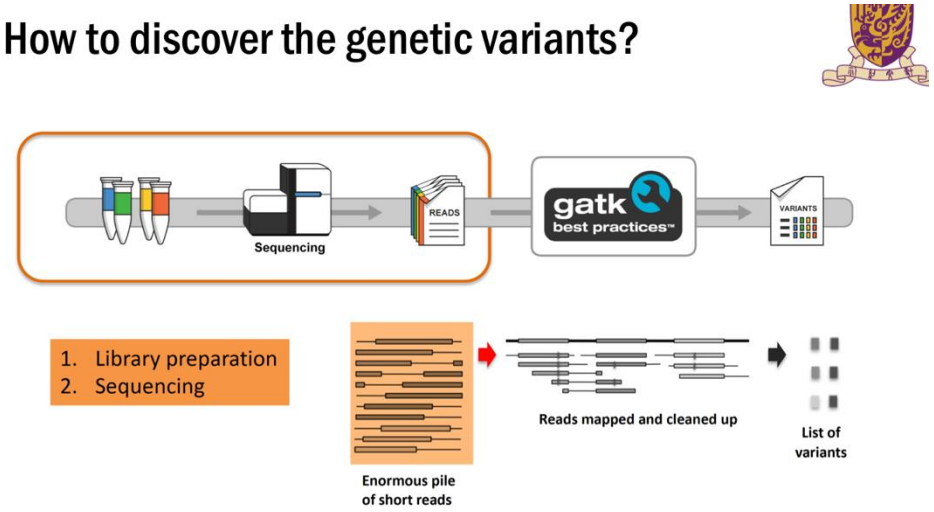
	Morning	Afternoon
Day 1	Introductions to Everything	Case study: End-to-end analysis
Day 2	Germline SNP & Indel calling	Germline variant filtering & evaluation
Day 3	Somatic SNV & Indel calling	Somatic Copy Number Alterations
Day 4	WDL/Cromwell Basics	

5.Significance of genome variance: differences in disease risk and response to treatment

6.How do we find variants?

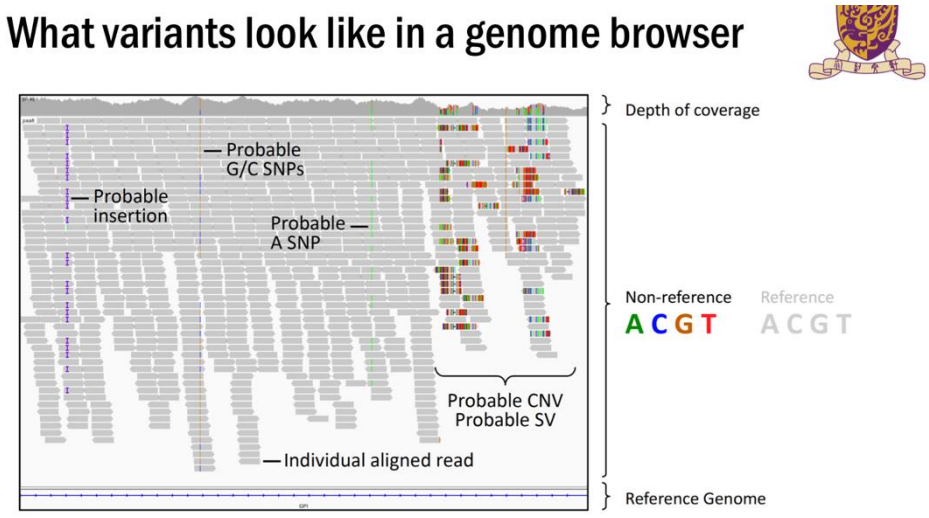
99.5% human DNA is the same. We describe a genome only using the remaining variants. With a reference human genome, we can map raw reads to it and compare differences. Some common variants include mutation, insertion, deletion, copy number, translocation and pathological change(non-human sequence).





7.How can we distinguish error and actual variants?

Depth of coverage serves as a good indicator of reliability of the result. If depth of coverage is great, it is more likely to be an actual variant.



8.Origins of variants?

PCR, sequencing error (raw reads), alignment error, variant calling, genotype (poor annotation).

9.Mendelian disease (single mutation SNPs)

Disease caused by single nucleotide mutation. It is relatively easier to cure such disease than multiple nucleotide mutation diseases.