Glossary of Genomic Alteration

1. SNP (Single Nucleotide Polymorphism): The most common type of change in DNA (molecules inside cells that carry genetic information). SNPs occur when a single nucleotide (building block of DNA) is replaced with another. These changes may cause disease, and may affect how a person reacts to bacteria, viruses, drugs, and other substances. Also called single nucleotide polymorphism.

2. Haplotype: A haplotype is a set of DNA variations, or polymorphisms, that tend to be inherited together. A haplotype can refer to a combination of alleles or to a set of single nucleotide polymorphisms (SNPs) found on the same chromosome. Information about haplotypes is being collected by the International HapMap Project and is used to investigate the influence of genes on disease.

3. SNV: Single Nucleotide Variation.

4. MNV: Multinucleotide variation.

5. Mutation: Any change in the DNA sequence of a cell. Mutations may be caused by mistakes during cell division, or they may be caused by exposure to DNA-damaging agents in the environment. Mutations can be harmful, beneficial, or have no effect. If they occur in cells that make eggs or sperm, they can be inherited; if mutations occur in other types of cells, they are not inherited. Certain mutations may lead to cancer or other diseases.

6. Substitution: Substitution is a type of mutation where one base pair is replaced by a different base pair. The term also refers to the replacement of one amino acid in a protein with a different amino acid.

7. Point mutation: A point mutation is when a single base pair is altered. Point mutations can have one of three effects. First, the base substitution can be a silent mutation where the altered codon corresponds to the same amino acid. Second, the base substitution can be a missense mutation where the altered codon corresponds to a different amino acid. Or third, the base substitution can be a nonsense mutation where the altered codon corresponds to a stop signal.

8. Silent mutation: A silent mutation alters DNA sequence, but has no apparent detectable effect on a phenotype or a function.

9. Missense (Conservative/Non-conservative) mutation: A missense mutation is when the change of a single base pair causes the substitution of a different amino acid in the resulting protein. This amino acid substitution may have no effect, or it may render the protein nonfunctional.

10. Nonsense mutation: A nonsense mutation is the substitution of a single base pair that leads to the appearance of a stop codon where previously there was a codon specifying an amino acid. The presence of this premature stop codon results in the production of a shortened, and likely nonfunctional, protein.

11. Insertion: Insertion is a type of mutation involving the addition of genetic material. An insertion

mutation can be small, involving a single extra DNA base pair, or large, involving a piece of a chromosome.

12. Deletion: Deletion is a type of mutation involving the loss of genetic material. It can be small, involving a single missing DNA base pair, or large, involving a piece of a chromosome.

13. Indel: insertion and deletion.

14. Frameshift mutation: A frameshift mutation is a type of mutation involving the insertion or deletion of a nucleotide in which the number of deleted base pairs is not divisible by three. "Divisible by three" is important because the cell reads a gene in groups of three bases. Each group of three bases corresponds to one of 20 different amino acids used to build a protein. If a mutation disrupts this reading frame, then the entire DNA sequence following the mutation will be read incorrectly.

15. Somatic/Germline mutation: Like all the cells that constitute the human body, a cancer cell is a direct descendant, through a lineage of mitotic cell divisions, of the fertilized egg from which the cancer patient developed and therefore carries a copy of its diploid genome. However, the DNA sequence of a cancer cell genome, and indeed of most normal cell genomes, has acquired a set of differences from its progenitor fertilized egg. These are collectively termed **somatic mutations** to distinguish them from **germline mutations** that are inherited from parents and transmitted to offspring,

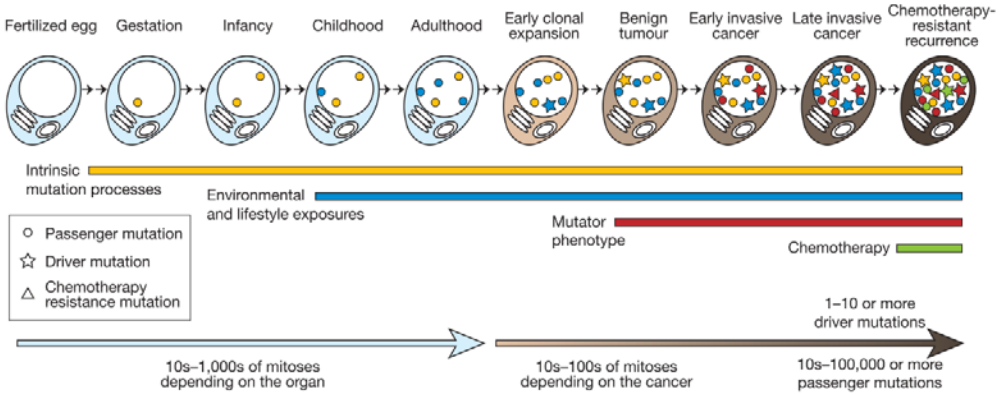

Figure 1 | The lineage of mitotic cell divisions from the fertilized egg to a single cell within a cancer showing the timing of the somatic mutations acquired by the cancer cell and the processes that contribute to them. Mutations may be acquired while the cell lineage is phenotypically normal, reflecting both the intrinsic mutations acquired during normal cell division and the effects of exogenous mutagens. During the development of the cancer other processes, for example DNA repair defects, may contribute to the mutational burden. Passenger mutations do not have any effect on the cancer cell, but driver mutations will cause a clonal expansion. Relapse after chemotherapy can be associated with resistance mutations that often predate the initiation of treatment.

16. Driver/Passager mutation:

All cancers arise as a result of somatically acquired changes in the DNA cancer cells. That does not mean, however, that all the somatic abnormalities present in a cancer genome have been involved in development of the cancer. Indeed, it is likely that some have made no contribution at all. To embody this concept, the terms "driver" and "passenger" mutation have been coined.

A **driver mutation** is causally implicated in oncognesis. It has conferred growth advantage on the cancer cell and has been positively selected in the microenvironment of the tissue in which the cancer arises. A driver mutation need not be required for maintenance of the final cancer (although it often is) but it must have been selected at some point along the lieage of cancer development shown in above figure.

A **passenger mutation** has not been selected, has not conferred clonal growth advantage and has therefore not contributed to cancer development. Passenger mutations are found within cancer genomes because somatic mutations without functional consequences often occur during cell divison. Thus, a cell that acquires a driver mutation will already have biologically insert somatic mutations within its genome. These will be carried along in the clonal expansion that follows and therefore will be present in all cells of the final cancer.

17. Structural variation (SV): Structural variation (SV) is generally defined as a region of DNA approximately 1 kb and larger in size and can include inversions and balanced translocations or genomic imbalances (insertions and deletions), commonly referred to as copy number variants (CNVs). These CNVs often overlap with segmental duplications, regions of DNA >1 kb present more than once in the genome, copies of which are >90% identical . If present at >1% in a population a CNV may be referred to as copy number polymorphism (CNP).

18. Chromosomal rearrangement: In genetics, a chromosomal rearrangement is a type of chromosome abnormality involving a change in the structure of the native chromosome. Such changes may involve several different classes of events, like deletions, duplications, inversions, and translocations.

19. Inversion: A chromosomal rearrangement in which a segment of genetic material is broken away from the chromosome, inverted from end to end, and re-inserted into the chromosome at the same breakage site. Balanced inversions (no net loss or gain of genetic material) are usually not associated with phenotypic abnormalities, although in some cases gene disruptions at the breakpoints can cause adverse phenotypic effects, including some known genetic diseases; unbalanced inversions (loss or gain of chromosome material) nearly always yield an abnormal phenotype.

20. Translocation: Translocation is a type of chromosomal abnormality in which a chromosome breaks and a portion of it reattaches to a different chromosome. Chromosomal translocations can be detected by analyzing karyotypes of the affected cells.

21. Copy number variation (CNV): A copy number variation (CNV) is when the number of copies of a particular gene varies from one individual to the next. Following the completion of the Human Genome Project, it became apparent that the genome experiences gains and losses of genetic material. The extent to which copy number variation contributes to human disease is not yet known. It has long been recognized that some cancers are associated with elevated copy numbers of particular genes.

22. Duplication: Duplication is a type of mutation that involves the production of one or more copies of a gene or region of a chromosome. Gene and chromosome duplications occur in all organisms, though they are especially prominent among plants. Gene duplication is an important mechanism by which evolution occurs.

23. Gene amplification: Gene amplification is an increase in the number of copies of a gene sequence. Cancer cells sometimes produce multiple copies of genes in response to signals from other cells or their environment. The term also can refer to polymerase chain reaction (PCR), a

laboratory technique that is used by scientists to amplify gene sequences in a test tube.

24. Genetic variation: Genetic variation refers to diversity in gene frequencies. Genetic variation can refer to differences between individuals or to differences between populations. Mutation is the ultimate source of genetic variation, but mechanisms such as sexual reproduction and genetic drift contribute to it as well.

25.Tandem Repeat: A tandem repeat is a sequence of two or more DNA base pairs that is repeated in such a way that the repeats lie adjacent to each other on the chromosome. Tandem repeats are generally associated with non-coding DNA. In some instances, the number of times the DNA sequence is repeated is variable. Such variable tandem repeats are used in DNA fingerprinting procedures.

Reference:

- http://www.cancer.gov/publications/dictionaries/cancer-terms
- http://www.genome.gov/glossary/
- https://en.wikipedia.org
- http://ghr.nlm.nih.gov/
- http://www.ncbi.nlm.nih.gov/dbvar/content/overview/
- Michael R. Stratton et al., Nature 2009.