




Glossary

Term	Translation to English
Human genome	All information about specific human is saved inside DNA. Genome is very long DNA sequence ~2 billions letters length. When we say human DNA we mean specific part of genome. Each (almost) cell inside you contains copy of genome.
Nucleotide or base	A, T, G, C - are chemicals DNA is build of. For example, sequence ATATAGCGATAGC. Nucleotide and base mean the same.
Genome position	If you represent genome as sequence or as array, list, whatever, you can assign position to every nucleotide.
hg19, hg38 and so on	<p>hg states for Human Genome.</p> <p>19, 38 and so on is version, the higher is the newer.</p> <div> <p>Note: hg19 is the same as hg37, just different name.</p> </div> <p>The newest version is more accurate, contains less errors. So, don't be surprised if one states that at position X there is A and second states that G.</p>
Strand	<p>Generally, DNA consists of 2 sequences. Each sequence called strand.</p>  <p>The one that above is top strand, the one that below is bottom strand.</p>
Complementarity	 <p>If you take a closer look, you will notice pattern how two strands are connected. A is connected to T, G is connected to C. They are complementary, and can be connected only that way. Examples,</p> <p>G NEVER go with T or A.</p> <p>T can go with A, A can go with T.</p>
Complementary strand	If you change all bases to its complementary nucleotide, you will get complementary strand.

	Example, we have ATGCATA, complementary strand is TACGTAT
ssDNA	Single Stranded DNA, only one strand of DNA, second is not bound (connected).
dsDNA	Double Stranded DNA, both strands are represented, they are bound (connected).
Reverse sequence	Literally, reverse ATGC goes into CGTA.
Reverse complemented	Reverse and find complementary strand. ATGC goes into GCAT (ATGC → CGTA → GCAT)
Variant	<p>In other words mutation. Variant is described as position of mutation, base before and base after.</p> <div>  Position is called position (yes, just like that). base is before is reference or ref, base after is alternation or alt. </div> <p>Example, 505114A>C. At position 505114 we should have A, but person has C. With this notation you can define many different mutations. See next.</p>
Substitution or change	<p>Basic case, I will simplify a bit. Top is what healthy person has, bottom is what sick person has.</p> <div>  Here top and bottom have nothing in common with top and bottom strands. I used just for clarity which is what. </div> <p>GACGATCGATGCTCAGTCAGTCA GTCGATCGATGCTCAGTCAGTCA Variant is 2A>T</p>
Insertion	<p>GACGATCGATGCTCAGTCAGTCA GAGCGATCGATGCTCAGTCAGTCA Variant is 2A>AG. Meaning that G was inserted after A.</p>
Deletion	<p>GACGATCGATGCTCAGTCAGTCA GAGATCGATGCTCAGTCAGTCA Variant is 2AC>A. Meaning that C is deleted.</p>
Complex variant	<div>  Complex variant is what I just now invented as term </div> <p>However, sometimes you will have shit like this</p>

GACGATCGATGCTCAGTCAGTCA
GTGCCTGGCATGCTCAGTCAGTCA

Variant is 2ACGATCG>TGCCTGGC. Using this notation is pretty easy, other notations can use - for separation or words to describe what happend ins, del and so one.