REF and ALT

The reference allele and alternative allele(s) observed in a sample, set of samples, or a population in general (depending how the VCF was generated). The REF and ALT alleles are the only required elements of a VCF record that tell us whether the variant is a SNP or an indel (or in complex cases, a mixed-type variant). If we look at the following two sites, we see the first is a SNP, the second is an insertion and the third is a deletion:

```
20  10001298    .   T   A   884.77  .   [CLIPPED]   GT:AD:DP:GQ:PL  1/1:0,30:30:89:913,89,0
20  10001436    .   A   AAGGCT  1222.73 .   [CLIPPED]   GT:AD:DP:GQ:PL  1/1:0,28:28:84:1260,84,0
20  10004769    .   TAAAACTATGC T   622.73  .   [CLIPPED]   GT:AD:DP:GQ:PL  0/1:18,17:35:99:660,0,7
```

Note that REF and ALT are always given on the forward strand. For insertions, the ALT allele includes the inserted sequence as well as the base preceding the insertion so you know where the insertion is compared to the reference sequence. For deletions, the ALT allele is the base before the deletion.