

- AD, ADF, ADR (Integer): Per-sample read depths for each allele; total (AD), on the forward (ADF) and the reverse (ADR) strand.
- DP (Integer): Read depth at this position for this sample.
- EC (Integer): Comma separated list of expected alternate allele counts for each alternate allele in the same order as listed in the ALT field. Typically used in association analyses.
- LEN (Integer): length of the <*> reference block for this sample.
- FT (String): Sample genotype filter indicating if this genotype was “called” (similar in concept to the FILTER field). Again, use PASS to indicate that all filters have been passed, a semicolon-separated list of codes for filters that fail, or ‘.’ to indicate that filters have not been applied. These values should be described in the meta-information in the same way as FILTERS. No whitespace or semicolons permitted.
- GQ (Integer): Conditional genotype quality, encoded as a phred quality $-10\log_{10} p(\text{genotype call is wrong, may be conditioned on the site ~~'s being variant~~ being polymorphic in the population})$.
- GP (Float): Genotype posterior probabilities in the range 0 to 1 using the same ordering as the GL field; one use can be to store imputed genotype probabilities.
- GT (String): Genotype, encoded as allele value preceded by either / or | depending on whether that allele is considered phased. The first phasing indicator may be omitted and is implicitly defined as / if any phasing indicators are / and | otherwise. The allele values are 0 for the reference allele (what is in the REF field), 1 for the first allele listed in ALT, 2 for the second allele list in ALT and so on. For diploid calls examples could be 0/1, 1 | 0, /0/1, or 1/2, etc. Haploid calls, e.g. on Y, male non-pseudoautosomal X, or mitochondria, should be indicated by having only one allele value. A triploid call might look like 0/0/1, and a partially phased triploid call could be |0/1/2 to indicate that the first allele is phased with another variant in the VCF. If a call cannot be made for a sample at a given locus, ‘.’ must be specified for each missing allele in the GT field (for example ‘./.’ for a diploid genotype and ‘.’ for haploid genotype). The meanings of the phasing indicators are as follows (see the PS and PSL fields below for more details on incorporating phasing information into the genotypes):
 - / : allele is unphased
 - | : allele is phased (according to the phase-set indicated in PS or PSL)

For symbolic structural variant alleles, GT=0 indicates the absence of any of the ALT symbolic structural variants defined in the record. Implementer should note that merging a VCF record containing only symbolic structural variant ALT alleles with a record containing other alleles will result a change of the meaning of the GT=0 haplotypes from the record containing only symbolic SVs.

- GL (Float): Genotype likelihoods comprised of comma separated floating point \log_{10} -scaled likelihoods for all possible genotypes given the set of alleles defined in the REF and ALT fields. In presence of the GT field the same ploidy is expected; without GT field, diploidy is assumed.

GENOTYPE ORDERING. In general case of ploidy P and N alternate alleles (0 is the REF and $1 \dots N$ the alternate alleles), the ordering of genotypes for the likelihoods can be expressed by the following pseudocode with as many nested loops as ploidy: †

```

for  $a_P = 0 \dots N$ 
  for  $a_{P-1} = 0 \dots a_P$ 
    ...
    for  $a_1 = 0 \dots a_2$ 
      println  $a_1 a_2 \dots a_P$ 

```

Alternatively, the same can be achieved recursively with the following pseudocode:

```

Ordering( $P, N, \text{suffix}=""$ ):
  for  $a$  in  $0 \dots N$ 
    if ( $P == 1$ ) println str( $a$ ) + suffix
    if ( $P > 1$ ) Ordering( $P-1, a, \text{str}(a) + \text{suffix}$ )

```

Conversely, the index of the value corresponding to the genotype $k_1 \leq k_2 \leq \dots \leq k_P$ is

†Note that we use inclusive for loop boundaries.