VCF Format

VCF (Variant Call Format)

```
##fileformat=VCFv4.2
##reference=GRCh38
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT SAMPLE1
chr1 10177 . A AC 50 PASS DP=32;AF=0.5 GT:DP:GQ 0/1:32:50
chr1 10352 rs123 T A 100 PASS DP=45;AF=1.0 GT:DP:GQ 1/1:45:99
```

VCF Columns

CHROM
Chromosome

POS
Position

REF
Reference

ALT
Alternate

QUAL
Quality

INFO
Annotations

Genotype (GT)

0/0 = homozygous reference

0/1 = heterozygous

1/1 = homozygous alternate

Key INFO Fields

DP = Total depth

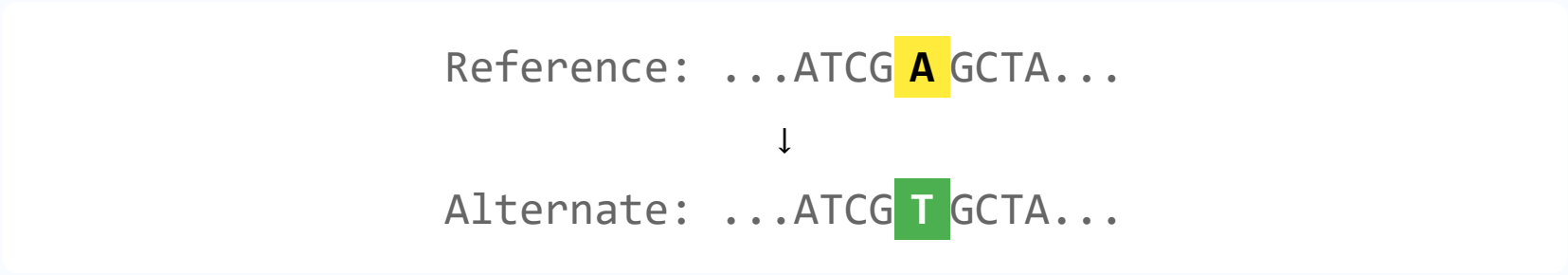
AF = Allele frequency

AC = Allele count

AN = Total alleles

Variant Types with Examples

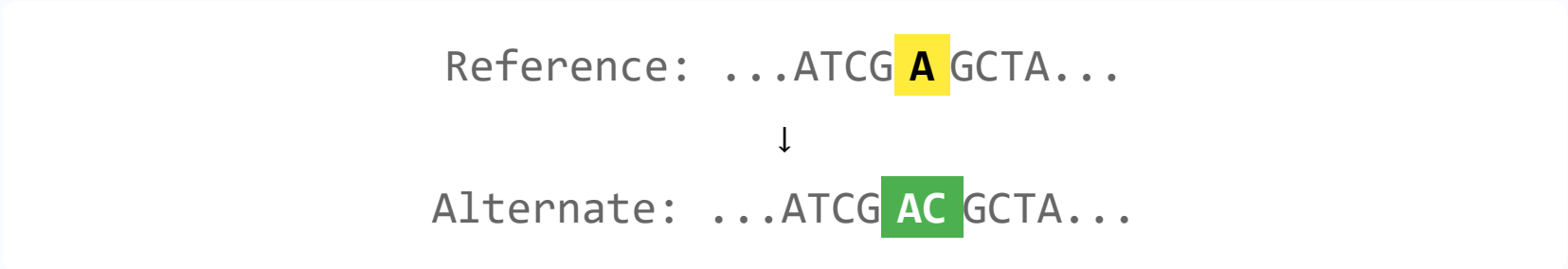
1. SNP (Single Nucleotide Polymorphism)




```
chr1 10352 rs123 A T 100 PASS
```

The most common type of variant, where a single nucleotide is substituted with another. In this example, A at position 10352 is changed to T.

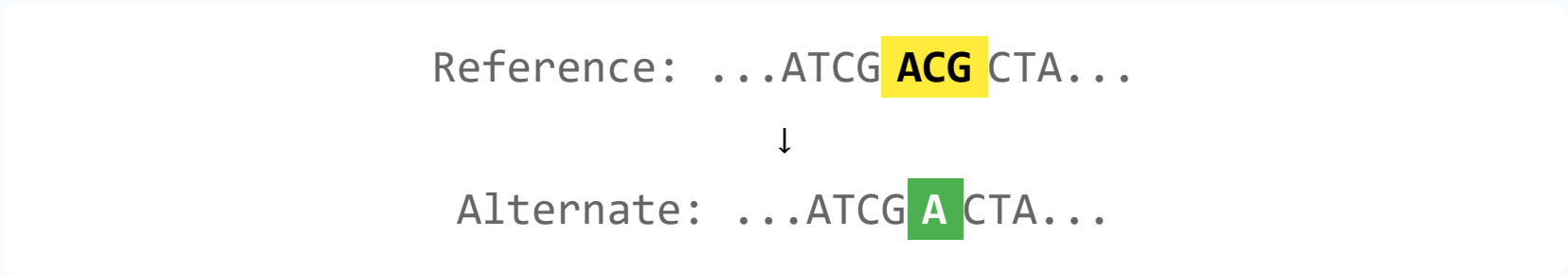
2. Insertion




```
chr1 10177 . A AC 50 PASS
```

Nucleotides are added to the reference sequence. REF is the base at the insertion position (A), and ALT includes REF + the inserted base (AC).

3. Deletion




```
chr1 20000 . ACG A 80 PASS
```

The sequence is removed. REF includes the deleted bases (ACG), and ALT shows only the remaining base (A).

Genotype Interpretation

0/0

Homozygous Reference

Allele 1: A
Allele 2: A

Both alleles match the reference sequence

0/1

Heterozygous

Allele 1: A
Allele 2: T

One reference allele, one alternate allele

1/1

Homozygous Alternate

Allele 1: T
Allele 2: T

Both alleles are alternate variants

Quality Metrics Explained

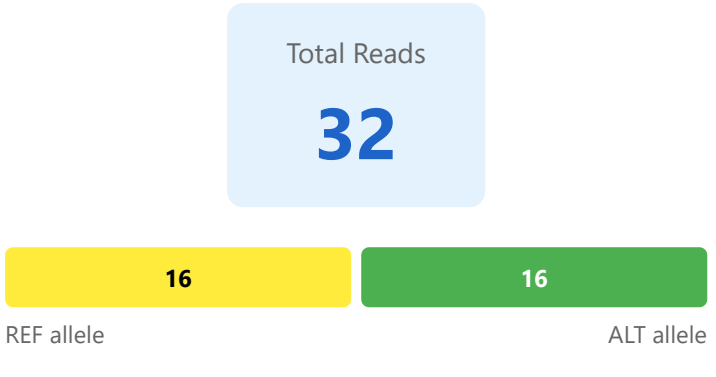
QUAL (Quality Score)

Phred-scaled quality score

```
QUAL = 10 → 90% accuracy
QUAL = 20 → 99% accuracy
QUAL = 30 → 99.9% accuracy
QUAL = 40 → 99.99% accuracy
```

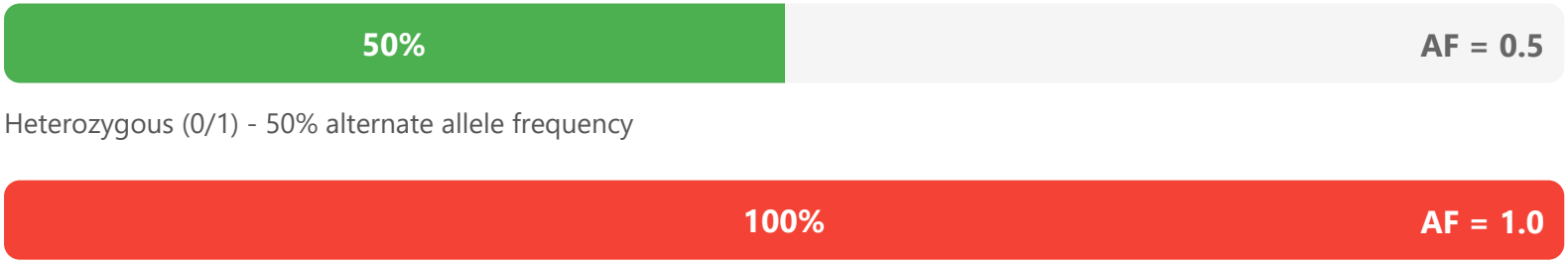
Higher values indicate greater confidence in the variant call. Generally, values of 30 or above are considered reliable variants.

DP (Depth)



The total number of reads covering this position. Higher depth increases confidence in variant calling.

AF (Allele Frequency)



Heterozygous (0/1) - 50% alternate allele frequency

Homozygous alternate (1/1) - 100% alternate allele frequency

Complete Example Analysis

```
chr1 10352 rs123 T A 100 PASS DP=45;AF=1.0 GT:DP:GQ 1/1:45:99
```

Location Information

- Chromosome: chr1
- Position: 10,352
- dbSNP ID: rs123

Variant Information

- Reference base: T
- Alternate base: A
- Variant type: SNP

✓ Quality Information

- Quality Score: 100 (very high)
- Filter: PASS
- Depth: 45 reads

Sample Information

- Genotype: 1/1 (homozygous)
- Allele Frequency: 100%
- Genotype Quality: 99

Interpretation

This sample shows a homozygous variant at position 10,352 on chromosome 1, where both alleles have changed from T to A. Confirmed by 45 reads, with high quality score (100) and genotype quality (99), indicating a highly reliable variant call.