

# Variant Calling

## Reference:

...ATCGA TCGATCGATCG...

## Sample Reads:

TCGATCGAT

ATCGGTCGAT

ATCGATCGAT

ATCGGTCGAT

→ **SNV Detected: A>G**

Allele Frequency: 75% (3/4 reads)

## Confidence Metrics:

- ✓ Coverage depth: 4x
- ✓ Base quality: Q35+
- ✓ Mapping quality: 60

## Variant Calling Process

- Identify differences between sample and reference genome
- Distinguish true variants from sequencing errors
- Calculate confidence scores for each variant
- Filter low-quality and false positive calls

GATK

**Gold Standard**

Genome Analysis Toolkit

FreeBayes

**Bayesian**

Haplotype-based

DeepVariant

**Deep Learning**

Google AI method

## Variant Types

### SNVs/SNPs

Single nucleotide variants - most common (~50M per genome)

### Indels

Small insertions/deletions - 1-50 bp

### Structural Variants

Large deletions, duplications, inversions, translocations (>50 bp)

### Copy Number Variants

Changes in gene copy number