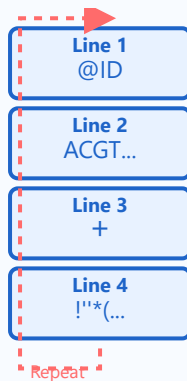


FASTQ Format

FASTQ File Structure



@SEQ_ID (Sequence identifier)

GATTGGGGTTCAAAGCAGTATCGATCAAAATAGTAAATCCATTGTTCAACTCACAGTTT

+ (Separator)

!'"*(((((***+))%%%+)(%%%) . 1***-+* ' '))**55CCF>>>>>CCCCCCC65

Line 1: @Identifier

Unique read ID with instrument and run information

Line 2: Sequence

Raw nucleotide sequence (A, T, C, G, N)

Line 3: +

Separator (sometimes repeats identifier)

Line 4: Quality Scores

Phred quality scores (ASCII encoded)

Phred Score: $Q = -10 \times \log_{10}(P)$ | Q30 = 99.9% accuracy, Q40 = 99.99%

