

FASTA vs FASTQ Quick Reference

Understand the two most common sequence file formats used in bioinformatics.

Quick comparison

Feature	FASTA	FASTQ
Purpose	Stores nucleotide or protein sequences	Stores sequencing reads plus quality scores
Typical extension	.fa, .fasta, .fna, .faa	.fastq, .fq, often .gz
Record structure	Header + sequence	4 lines per read
Quality scores	No	Yes
Common use	References, genes, proteins	Raw NGS reads

FASTA example

```
>gene_001
ATGCGTACGTTAGCCTAG
```

FASTQ example

```
@read_001
ATGCGTACGTTAGCCTAG
+
IIIIIIHHHHFFEEEDDD
```

Remember: FASTA answers 'what is the sequence?' FASTQ answers 'what is the sequence and how confident was the sequencer at each base?'