

Short-Read vs. Long-Read Sequencing

Next-generation sequencing platforms can be roughly divided into short- (usually <600 bases) and long-read (about 1 kb and above) sequencing according to the size of DNA fragments sequenced. Each technology has advantages and disadvantages that need to be considered prior to their use. This infographic compares the approaches, providing pros and cons, a list of relevant platforms, and expert guidance on which method is ideal for specific applications.

Definition of each method

Short Read

In short-read sequencing, DNA (and cDNA) is broken into fragments, which are typically amplified and then sequenced and finally pieced back together in a continuous sequence. The read lengths in short-read sequencing are generally 600 bases or less.

Long Read

Long-read sequencers generate reads with much longer lengths, typically from 1,000 to 10,000-plus bases at one time and typically require less, if any, fragmentation or amplification of DNA or RNA samples.

Examples of each method

Short Read

- Illumina platform (i.e., iSeq, MiSeq, NextSeq, HiSeq, NovaSeq)
- Thermo Fisher Scientific (Ion Torrent Genexus)
- PacBio (Onso)
- Element Biosciences (AVITI)
- Complete Genomics/MGI (DNBSEQ-G400, DNBSEQ-E25, DNBSEQ-T7, and DNBSEQ-T10x4)
- Ultima Genomics (UG 100)
- Singular Genomics (G4)

Long Read

- Oxford Nanopore platform (i.e., MinION, GridION, PromethION)
- Pacific Biosciences (Sequel, RS, and Revio systems)

Pros and cons of each method

Short Read

- Lower cost per base
- High depth
- High-quality data
- Good for targeting specific regions (targeted resequencing)

Long Read

- Lower start-up cost
- Real-time data streaming
- No sample batching requirements
- Sample prep can be less onerous
- Ability to distinguish large structural variations, regions of high homology or repetition, and splice variations

- Can't sequence long stretches of DNA
- Overlap between DNA fragments can be an issue
- Fixed run time with bulk data delivery can mean long wait times for results
- Sample may need to be batched
- Extensive sample prep and amplification may be required
- Some genome features are difficult to see, including low-complexity regions and repetitive sequences

- Historically has been associated with higher error rates but that is changing
- Data processing can be complex and time consuming for larger genomes
- Assembly may still require polishing/follow-up reads from short-read sequencers

The best of both worlds

It is possible to tap into the advantages of both approaches and combine short-read and long-read data. Sophisticated analysis methods are required, but the benefits justify the effort as the long reads provide genome structure and the short reads are used to correct base identity issues leading to a more complete genome that is more accurate.

Ideal applications

Short Read

- Fragmented or shorter DNA segments
- Low frequency variant calling
- Sequencing FFPE-derived DNA
- ctDNA and cfDNA sequencing
- Target capture sequencing

Long Read

- Sequencing through repetitive regions
- Sequencing full-length RNA transcripts
- Sequencing of native molecules (only with Oxford Nanopore platform), which prevents amplification bias
- Whole genome assembly