DNA Sequencing

How a sequence is read, from Sanger to modern methods.

Biology

Sanger sequencing

- Chain-terminating dideoxynucleotides
- Read length ~800–1,000 bases
- Very high accuracy
- One fragment at a time
- Still the standard for single genes

Next-generation

- Millions of fragments in parallel
- Short reads, 100–300 bases typically
- Assembled computationally
- Whole genomes in hours
- Cost per genome has fallen enormously

THE DIFFERENCE THAT MATTERS

Long-read technologies — nanopore, PacBio — read tens of kilobases, resolving repeats that short reads cannot.

The Human Genome Project took 13 years and about $3bn. A genome now costs a few hundred and takes about a day.