Summaries of Sequencing Technologies

Illumina:

To prepare samples for Illumina sequencing, adapters are added to the ends of DNA fragments, which include a sequencing, biding site, indices, and regions complementary to flow cell oligos.

Next, each fragment is amplified on a flow cell that contains several lanes. On each lane is a lawn of two kinds of oligos. One oligo type is complementary to adapters on the DNA fragments, and fragments bind to these oligos on the flow cell lawn. Then, polymerase creates a complement of the fragment, which folds over and binds to the other oligo type. Then polymerase copies this fragment repeatedly, which yields amplification of all the DNA fragments across the lawn. Sequencing is accomplished with specially tagged nucleotides which fluoresce as they bind to template DNA. Since each of the 4 nucleotides are assigned a unique fluorescent color, the Illumina machinery can record the color and intensity of the fluorescence to call bases as they bind to the template DNA. This same process is also done for the reverse strand. For data analysis, forward and reverse sequences are paired and aligned to the reference genome.

Limitations: Because Illumina uses short reads, sequencing can miss a large number of coding exons, including common repeats across the genome.

PacBio “SMRT sequencing”:

PacBio is a long-read sequencing technology used for genomes, transcriptomes, and epigenomes. Long-read sequencing allows more accurate coverage of entire transcripts and regions that may be inaccessible to other technologies.

In the PacBio sequencing process, DNA (or RNA) is isolated and a “SMRTbell Library” of fragments is created by attaching adapters to the ends of the double-stranded DNA, which connect the two strands together in a circular fashion. Then, primer and polymerase are added to the library. These DNA molecules are then fed into the SMRT sequencer, which contains a SMRT cell, which millions of wells. Each DNA molecule lands in one of these wells. As the polymerase adds tagged nucleotides to the template, light is emitted, which is recorded in real time. SMRT sequencing is useful for whole genome sequencing, RNA sequencing, sequencing individuals from genetically variable populations, and epigenetics.

Limitations: Although PacBio uses long-read sequencing, very long transcripts are still likely not fully sequenced, as there are sequencing length limits.

Nanopore

Nanopore is used to sequence DNA, RNA, miRNA and proteins. Unlike the techniques described above, Nanopore sequences fragments regardless of length. This technology uses nanopores, which form holes in membranes. In the Nanopore system a nanopore is placed in a synthetic membrane. An electric potential is applied to the synthetic membrane which creates current that flows through the nanopore. When molecules are passed through the nanopore, it will disrupt this current differently according to the size of the molecule. In the case of

DNA, an enzyme adhered to double-stranded DNA attaches to the nanopore and unzips the strand, feeding single-stranded DNA pulled through the nanopore one base at a time. The characteristic disruption of the electric current caused by each nucleotide as they pass through the nanopore generates the sequence data. If the double-stranded DNA is prepared with a hairpin structure at the end, the nanopore can read both strand in a single read.