

## DNA Sequencing



### General Aim

Detection of the sequence of DNA in a sample.

### Method

Next Generation Sequencing (NGS), Whole Genome Sequencing (WGS), Nextera DNA library preparation kit.

### Learning Objectives (ILOs)

- To apply DNA sample purification using exosap IT.
- To apply library preparation steps properly.
- To understand and apply DNA fragmentation, adapter ligation, clean up and amplification of tagged DNA.
- To practice library normalization, denaturing and dilution.
- To perform AMPure beads protocol to purify DNA libraries.

### Theoretical Background/Context

DNA is extracted and purified before sequencing. Some of this DNA is then processed for library preparation. Library is a pool of similarly sized DNA fragments with adapters attached.

During library preparation, DNA first undergoes fragmentation, where DNA is simultaneously fragmented and tagged with adapters using an enzyme called transposome. Transposomes can cut DNA and insert a portion of itself (adapter sequence). After adapters are ligated, DNA is cleaned up from excess transposomes and amplified using PCR. During PCR, additional motifs, such as the sequencing primer binding sites indices and regions that are complementary to the flow cell oligos are also added.

Libraries are then cleaned up using AMPure XP beads, which also provide a size selection feature. This is followed by quantification, using Qubit fluorometric assay, and normalization.

Clustering and sequencing occur in the flow cell. A flow cell is a multilane glass with nano wells coated with two types of oligonucleotides (oligos) on the surface.

#### ► Cluster generation results in clonal amplification of all the fragments through the following steps:

1. Fragments are applied and allowed to hybridize through adapter sequences with complementary oligos.
2. A polymerase creates a complement of the hybrid fragment creating a double stranded hybrid.
3. The double stranded molecule is denatured and the original template is washed away.
4. The strands are clonally amplified through bridge amplification. In this process:
  - a The strand folds over on the adapter region and hybridizes to the second type of oligo on the flow cell.
  - b Polymerases generate the complementary strand forming a double stranded bridge.
  - c This bridge is denatured resulting in two single stranded copies of the molecule that are tethered to the flow cell.
  - d The process is then repeated over and over again simultaneously with millions of clusters resulting in
  - e clonal amplification of all the fragments.

## Theoretical Background/Context (Cont')

5. After bridge amplification, the reverse strands are cleaved and washed off leaving only the forward strands.
6. The 3' ends are blocked to prevent unwanted priming.

### ► Sequencing then begins by synthesis in the following manner:

- a. The extension of the first sequencing primer to produce the first read.
  1. This proprietary process is called sequencing by synthesis. During the synthesis reactions, proprietary modified nucleotides, corresponding to each of the four bases, each with a different fluorescent label, are incorporated in a complementary manner to the template, and are then detected. The number of cycles determines the length of the read. Hundreds of millions of clusters are sequenced in a massively parallel process.
  2. After the completion of the first read, the read product is washed away.
  3. The index one read primer is introduced and hybridized to the template. The read is generated in the same manner as the first read.
  4. After completion of the index one read, the read product is washed off.
  5. The 3' end becomes deprotected.
  6. The template then folds over and binds the second oligo on the flow cell.
  7. Index two is read in the same manner as index one.
  8. Index two read product is washed off at the completion of this step.
  9. Polymerases extend the second flow cell oligo forming a double stranded bridge.
  10. This dsDNA is then linearized and the 3' ends are blocked.
  11. The original forward strand is cleaved off and washed away leaving the reverse strand.
- b. Read two begins with the introduction of the read two sequencing primer. As with read one, the sequencing steps are repeated until the desired read length is achieved. The read two product is washed away.

This entire process generates billions of reads representing all the fragments.

Sequences from pooled sample libraries are separated based on the unique indices introduced during the sample preparation. For each sample, reads with similar stretches of base calls are locally clustered. Forward and reverse reads are paired creating continuous sequences. These continuous sequences are aligned back to the reference genome for variant identification.

## Principle of Work

### ► Steps overview:

- a. Purification of extracted DNA using exosapit.
- b. Library construction (similarly sized dna sequences with known adapter sequences added to the 5' and 3' ends)
  1. Dna tagmentation (fragmentation + adapter ligation).
  2. Clean up of tagmented DNA.
  3. Amplification of tagmented DNA and addition of index adapters using 5 cycle PCR.
  4. Library clean up and size selection (using AMPure beads).
  5. Library normalization and pooling.
  6. Denaturing and Diluting Libraries and PhiX Control for Sequencing.
- c. Cartridges and flow cell handling.
- d. Sequence run.