Sequencing: Illumina Technology

Welcome Page

Narration:

Welcome to the Sequencing: Illumina Technology course. Click Next to continue.

Course Navigation

Narration:

Take a moment to familiarize yourself with the navigation for this course. You can access these instructions at any time by clicking Help in the top right corner of the page.

Course Objectives

Narration:

By the end of this course, you will be able to:

- List the four separate processes of the Illumina Sequencing Workflow.
- Identify where library preparation is performed and the general steps of the preparation.
- Explain the cluster generation process and designate based on instrument if the process is performed on the cBot instrument or on-board the sequencer.
- Describe if either four-channel or two-channel sequencing is utilized based on the sequencer and,
- Understand an overview of the primary and secondary analysis after sequencing is performed.

Illumina Sequencing Workflow

Narration:

The Illumina sequencing workflow can be divided into four separate processes:

- Library Preparation
- Cluster Generation
- Sequencing and,
- Data Analysis

Library Preparation Overview

Narration:

Library preparation is performed on standard laboratory equipment and not on the Illumina sequencing system.

Steps include:
1. Nucleic acid
2. Modify to proper insert size and,
3. Add adapter with sites for flow cell binding and sequencing primer binding

**Library Prep is Critical for Successful Sequencing**

**Narration:**

Regardless of the application or input starting material, all libraries end up looking similar with adapters attached on both ends of the nucleic acid fragments.

For cluster generation: libraries have P5 and P7 binding regions, which interact with oligos of complementary sequence on the surface of the flow cell

For sequencing: Libraries must have sequencing primer binding regions

For multiplexing: libraries must have an index(es).

When library preparation is complete, we are ready to move on to the next stage of the Illumina Sequencing workflow: Cluster generation.

**Cluster Generation**

**Narration:**

Cluster generation occurs on a flow cell.

A flow cell is a thick glass slide with channels or lanes that are specific to each sequencer.

Each channel/lane is randomly coated with a lawn of oligos that are complementary to library adapters.

Cluster generation, depending on the instrument, either is performed on the cBot instrument or on-board the sequencer.

**Cluster Generation and SBS Workflow**

**Narration:**

The “Intro to Sequencing by Synthesis: Industry-leading Data Quality” video explores the following topics:

View the entire video by pressing the play button or view each topic by its link.
**Four- or Two-Channel Sequencing**

_Narration:_

During the imaging step in the Sequencing by Synthesis (SBS) process, either four-channel or two-channel sequencing is used. The type of imaging used depends on the sequencer.

**Four-Channel Sequencing**

_Narration:_

Four-channel sequencing works by using four images to determine which base occurs in each cluster. Each of the four DNA bases emits an intensity of a unique wavelength. Therefore, during a cycle, each cluster appears in only one of the four images. For example, when a strong intensity signal is detected in the wavelength related to the G base, a G is called. When a strong intensity signal is detected in the wavelength related to the T base, a T is called, and so forth. Four-Channel Sequencing requires all four images to build up the DNA sequence.

**Two-Channel Sequencing**

_Narration:_

Rather than using four-images from four channels, two-channel sequencing uses only two images: an image from a red channel and an image from a green channel. In Two Channel Sequencing the intensity emitted by each base is as follows: A emits 50% green and 50% red intensities, C emits 100% red intensity, G is dark and does not emit any intensity, and T emits 100% green intensity.

Therefore if we take a green image and a red image, A is presented in both images represented here as a blend of the red and green image. C is only presented in red. T is only presented in the green, and G is represented here with a blank image. The template is still built up over multiple cycles, so although a cluster starting on a G base would not be detected in the early cycles, it would be detected as it moves though the cycles and changes to A, C, or T.

Let us look at how each base call is made in two-channel sequencing. Intensity from each cluster is plotted onto a scatter plot of red intensity versus green intensity. Each base is called according to the region of the scatter plot, or base population that the cluster falls into.

**Primary and Secondary Analysis Overview**

_Narration:_

The Illumina Data Analysis workflow consists of four processes:

During sequencing, Illumina’s sequencing process uses a technology called sequencing by synthesis, or SBS. During SBS, the instrument takes a color image of each base as it is added to each DNA fragment.
During primary analysis, intensities are extracted from the image data and further translated into base calls. Primary analysis occurs on the sequencing instrument itself or the instrument computer. Resulting base call files are then used for secondary analysis.

Further data analysis occurs after base calling is complete in secondary analysis. Including alignment to a reference genome or variant detection.

Finally, data visualization tools transform sequencing data into various formats to allow you to explore or analyze data from one or many sequencing runs.

**Summary**

*Narration:*

You should now understand the Illumina sequencing workflow including library preparation, cluster generation, sequencing, and data analysis.

Please take a moment to review the topics covered in this course.

- Library preparation consists of modifying nucleic acid to proper size and adding adapters with sites for flow cell binding and sequencing primer binding.
- The goal of cluster generation is to attach DNA libraries into the flow cell and amplify each single DNA library into clonal clusters.
- In SBS technology, just one base is added in each cycle with each base image captured during the cycle.
- Images captured during sequencing are analyzed during the primary data analysis step to generate base call files.
- Sequencing data are aligned during the secondary data analysis step, resulting in universal data output files that are read by downstream software.