Illumina Sequencing Technology

Massively parallel sequencing technology enables the generation of billions of bases per week at less than 1% of the cost of capillary-based methods. An expansive scale of research unimaginable with other technology platforms is now possible.

INTRODUCTION

Illumina sequencing technology leverages clonal array formation and proprietary reversible terminator technology for rapid and accurate large-scale sequencing. The innovative and flexible sequencing system enables a broad array of applications in genomics, transcriptomics, and epigenomics.

CLUSTER GENERATION

Sequencing templates are immobilized on a proprietary flow cell surface (Figure 1) designed to present the DNA in a manner that facilitates access to enzymes while ensuring high stability of surface-bound template and low non-specific binding of fluorescently labeled nucleotides. Solid-phase amplification creates up to 1,000 identical copies of each single template molecule in close proximity (diameter of one micron or less). Because this process does not involve photolithography, mechanical spotting, or positioning of beads into wells, densities on the order of ten million single-molecule clusters per square centimeter are achieved.

SEQUENCING BY SYNTHESIS

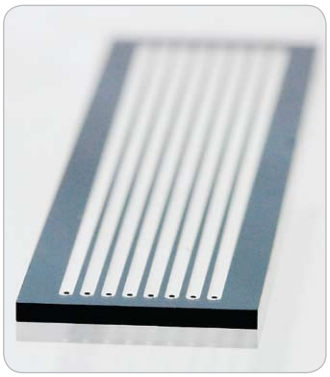
Illumina sequencing uses four proprietary fluorescently-labeled modified nucleotides to sequence the tens of millions of clusters present on the flow cell surface (*Figure 2*). These nucleotides, specially designed with a reversible termination property, allow each cycle of the sequencing reaction to occur simultaneously for all clusters in the presence of all four nucleotides (A, C, T, G). In each cycle, the polymerase is able to select the correct base to incorporate, with the natural competition between all four alternatives leading to higher accuracy than methods where only one nucleotide is present in the reaction mix at a time. Sequences where a particular base is repeated (e.g., homopolymers) are addressed like any other sequence and resolved with high accuracy.

ANALYSIS PIPELINE

The Illumina sequencing approach is built around a massive quantity of sequence reads in parallel. Deep sampling and uniform coverage is used to generate a consensus and ensure high confidence in determination of genetic differences. Deep sampling allows the use of weighted *majority voting* and statistical analysis, similar to conventional methods, to identify homozygotes and heterozygotes and to distinguish sequencing errors.

Each raw read base has an assigned quality score so that the software can apply a weighting factor in calling differences and generating confidence scores.

FIGURE 1: ILLUMINA FLOW CELL



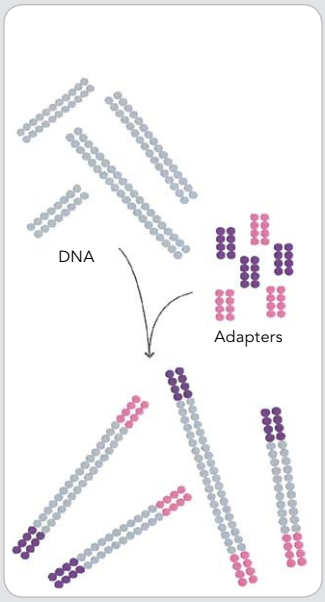
Several samples can be loaded onto the eight-lane flow cell for simultaneous analysis on the Illumina Genome Analyzer.

DATA COLLECTION, PROCESSING, AND ANALYSIS

Illumina data collection software enables users to align sequences to a reference in resequencing applications. Developed in collaboration with leading researchers, this software suite includes the full range of data collection, processing, and analysis modules to streamline collection and analysis of data with minimal user intervention. The open format of the software allows easy access to data at various stages of processing and analysis using simple application program interfaces.

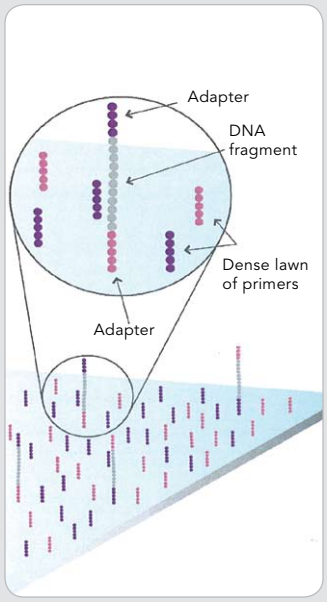
SEQUENCING TECHNOLOGY OVERVIEW

1. PREPARE GENOMIC DNA SAMPLE



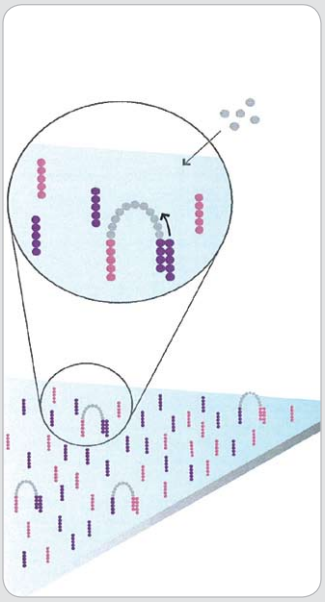
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

2. ATTACH DNA TO SURFACE



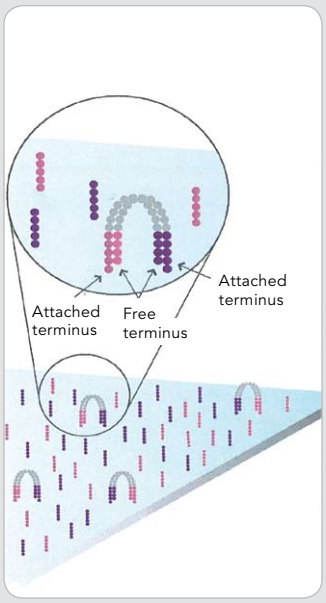
Bind single-stranded fragments randomly to the inside surface of the flow cell channels.

3. BRIDGE AMPLIFICATION



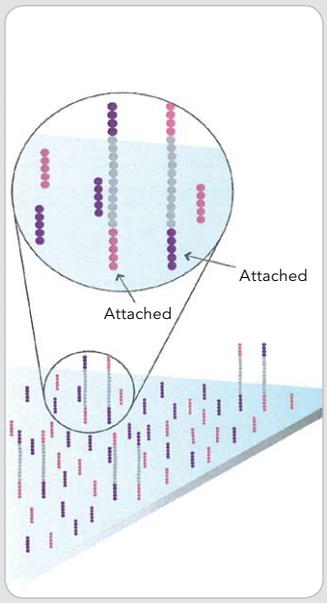
Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.

4. FRAGMENTS BECOME DOUBLE-STRANDED



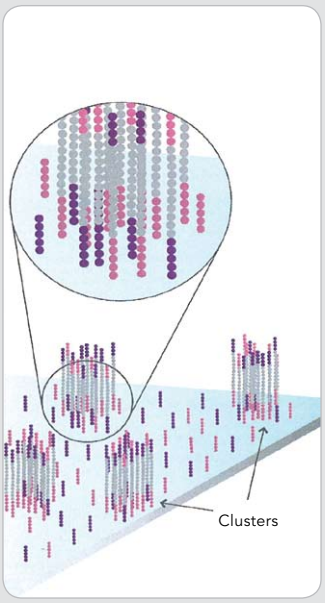
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

5. DENATURE THE DOUBLE-STRANDED MOLECULES



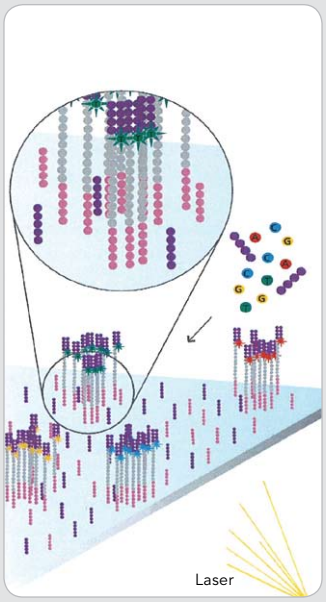
Denaturation leaves single-stranded templates anchored to the substrate.

6. COMPLETE AMPLIFICATION



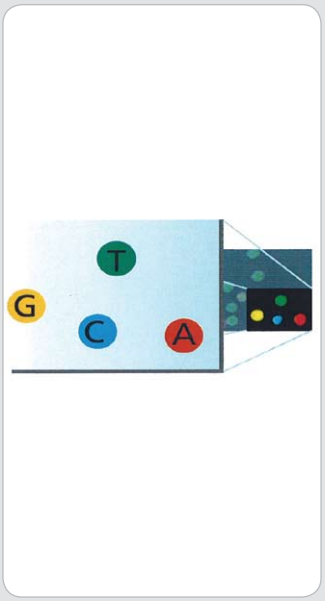
Several million dense clusters of double-stranded DNA are generated in each channel of the flow cell.

7. DETERMINE FIRST BASE



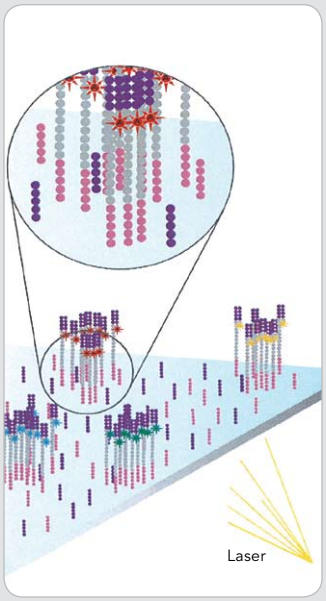
The first sequencing cycle begins by adding four labeled reversible terminators, primers, and DNA polymerase.

8. IMAGE FIRST BASE



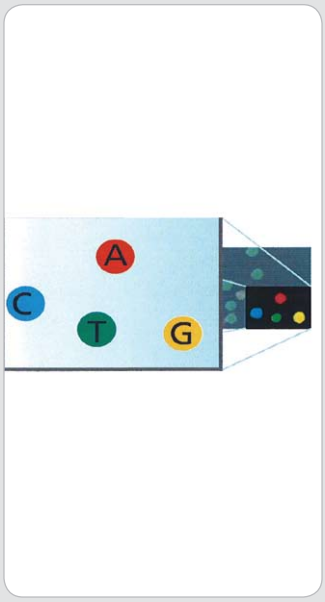
After laser excitation, the emitted fluorescence from each cluster is captured and the first base is identified.

9. DETERMINE SECOND BASE



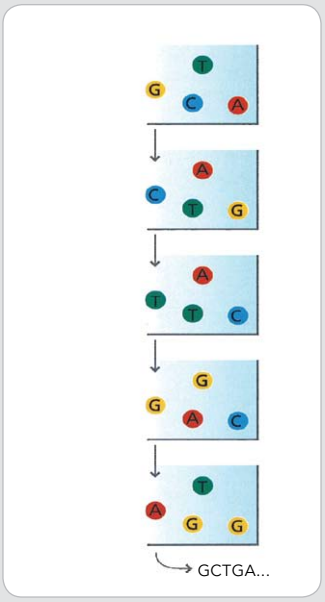
The next cycle repeats the incorporation of four labeled reversible terminators, primers, and DNA polymerase.

10. IMAGE SECOND CHEMISTRY CYCLE



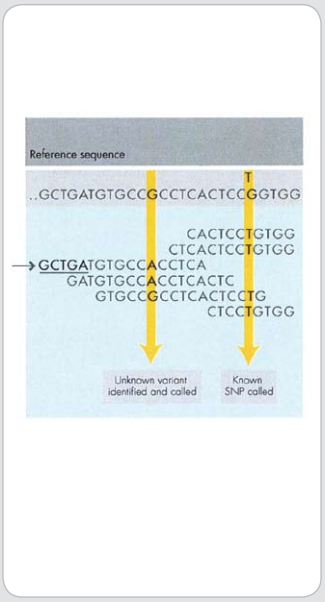
After laser excitation, the image is captured as before, and the identity of the second base is recorded.

11. SEQUENCING OVER MULTIPLE CHEMISTRY CYCLES



The sequencing cycles are repeated to determine the sequence of bases in a fragment, one base at a time.

12. ALIGN DATA



The data are aligned and compared to a reference, and sequencing differences are identified.

EXPLORE THE POSSIBILITIES

Illumina sequencing technology achieves unparalleled data density with highly accurate results. With the ability to generate several gigabases of DNA sequence per run, Illumina sequencing provides researchers with the broadest range of applications and the opportunity to sequence even large mammalian genomes in weeks rather than years.

Leveraging this technology, researchers can potentially resequence genomes for less than 1% of their current costs. With the capacity to accommodate many samples per flow cell, runs can be tailored to the demands of a wide range of applications. Whether project requirements involve a genome, region, or gene, Illumina sequencing is the ideal tool for a wide range of projects.

ADDITIONAL INFORMATION

Visit our website or contact us at the address below to learn more about Illumina sequencing technology and applications.

Illumina, Inc.
Customer Solutions
9885 Towne Centre Drive
San Diego, CA 92121-1975
1.800.809.4566 (toll free)
1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com

FOR RESEARCH USE ONLY

Pub. No. 770-2007-002 Current as of 8 April 2009

illumina®