TECHNOLOGY SPOTLIGHT

DNA Sequencing with Solexa® Technology

Generating one billion bases of high quality DNA sequence per run at less than 1% of the cost of capillary-based methods, the Illumina Genome Analyzer is designed to enable researchers to dramatically improve the efficiency and speed of current applications. Now an expanded scale of research that was previously unimaginable with other technology platforms is possible with the Genome Analyzer.

SOLEXA SEQUENCING TECHNOLOGY

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 is employed to create up to 1,000 identical copies of each single 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, Solexa sequencing technology can achieve densities of up to ten million single molecule clusters per square centimeter.

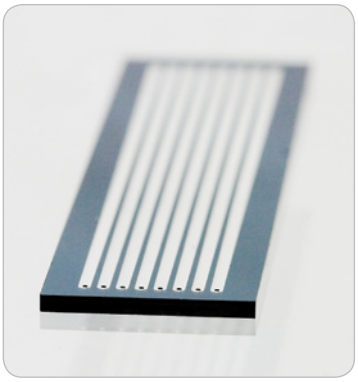
SEQUENCING-BY-SYNTHESIS

Solexa sequencing uses four proprietary fluorescently-labeled modified nucleotides to sequence the millions of clusters present on the flow cell surface (*Figure 2*). These nucleotides, specially designed to possess a reversible termination property, allow each cycle of the sequencing reaction to occur simultaneously 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 one after another (e.g., homopolymers) are addressed like any other sequence and with high accuracy.

ANALYSIS PIPELINE

The Solexa sequencing approach is built around a very large number of short sequence reads. Deep sampling of more than ten-fold even coverage is required to generate a consensus and thus ensure high confidence in determination of genetic differences. Such differences are identified by comparison of sequence reads to a reference. 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.

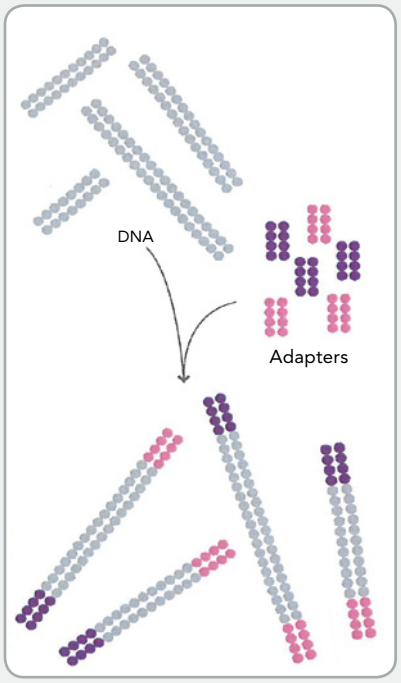
FIGURE 1: ILLUMINA GENOME ANALYZER FLOW CELL



Up to eight samples can be loaded onto the flow cell for simultaneous analysis on the Illumina Genome Analyzer.

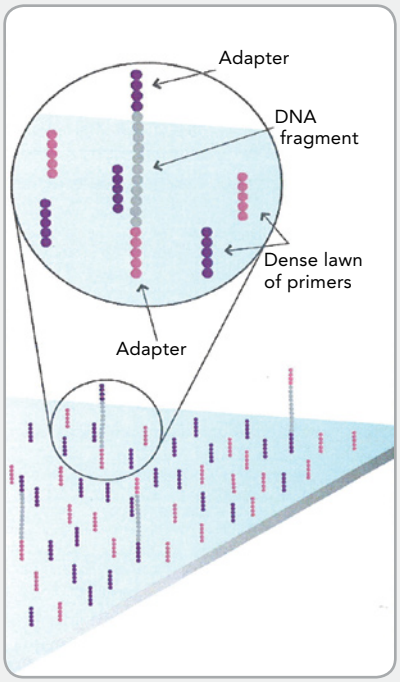
FIGURE 2: SEQUENCING TECHNOLOGY OVERVIEW

1. PREPARE GENOMIC DNA SAMPLE



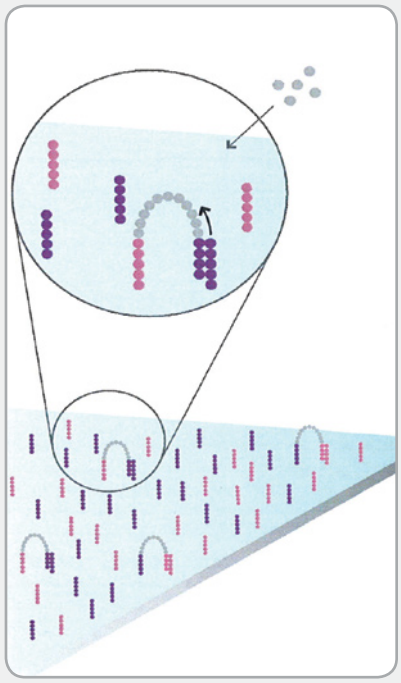
Randomly fragment genomic DNA and ligate adapters to both ends of the fragments.

2. ATTACH DNA TO SURFACE



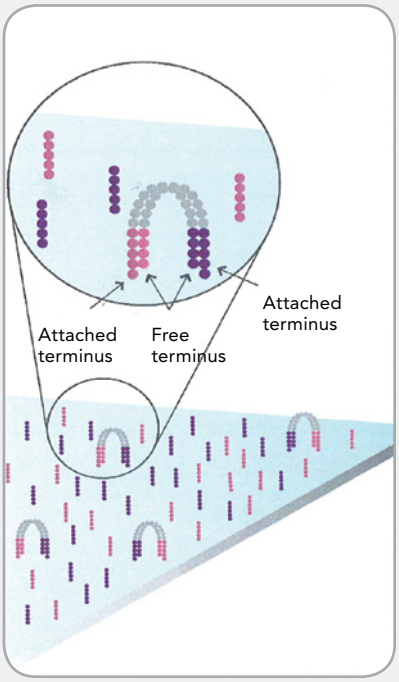
Bind single-stranded fragments randomly to the inside surface of the flow cell channels.

3. BRIDGE AMPLIFICATION



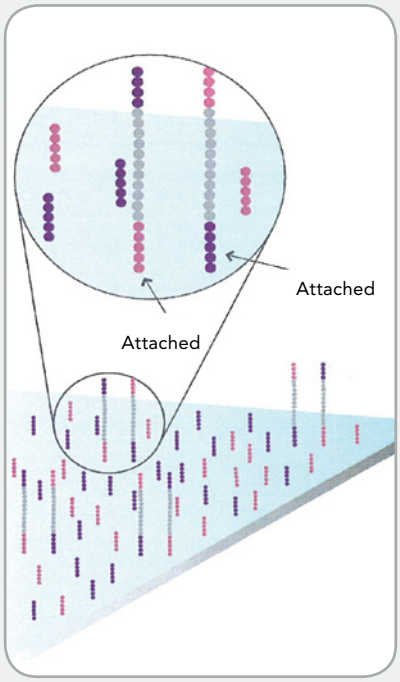
Add unlabeled nucleotides and enzyme to initiate solid-phase bridge amplification.

4. FRAGMENTS BECOME DOUBLE STRANDED



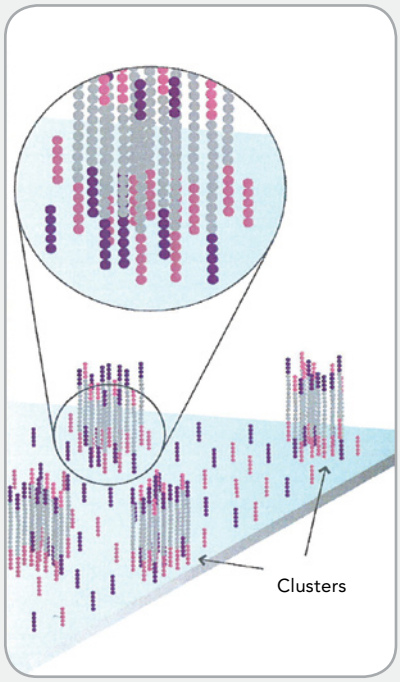
The enzyme incorporates nucleotides to build double-stranded bridges on the solid-phase substrate.

5. DENATURE THE DOUBLE-STRANDED MOLECULES



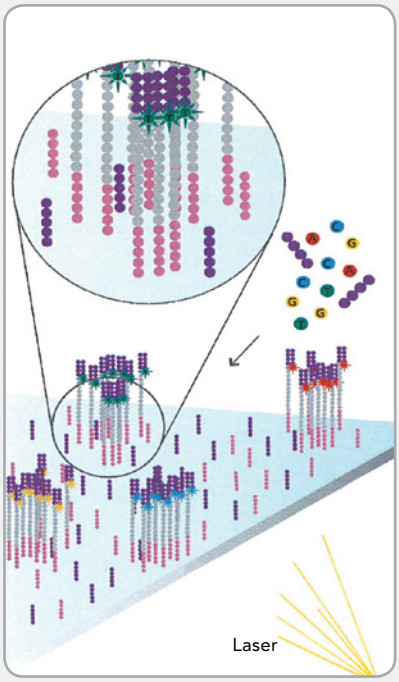
Denaturation leaves single-stranded templates anchored to the substrate.

6. COMPLETE AMPLIFICATION



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

7. DETERMINE FIRST BASE

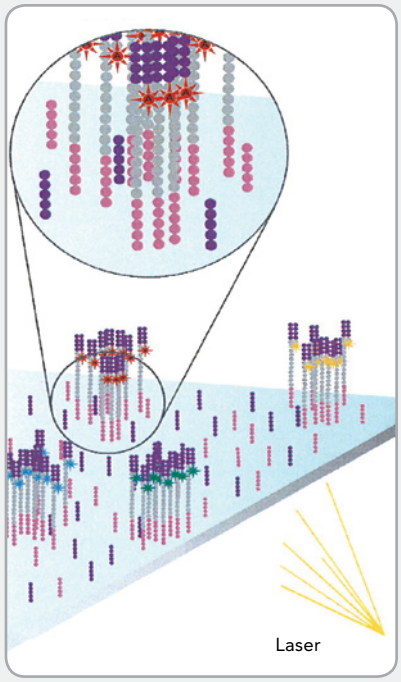


First chemistry cycle: to initiate the first sequencing cycle, add all four labeled reversible terminators, primers and DNA polymerase enzyme to the flow cell.

8. IMAGE FIRST BASE

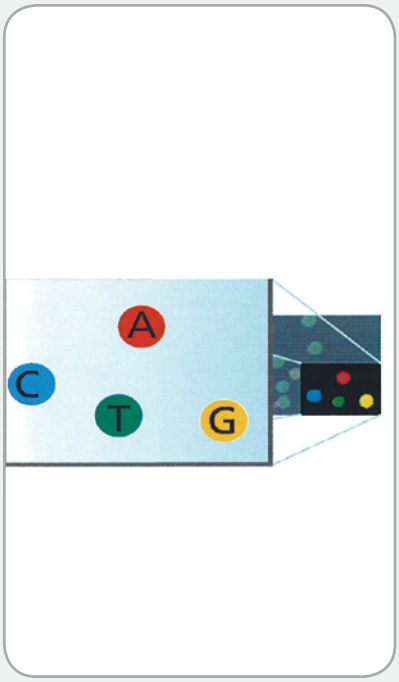
After laser excitation, capture the image of emitted fluorescence from each cluster on the flow cell. Record the identity of the first base for each cluster.

9. DETERMINE SECOND BASE



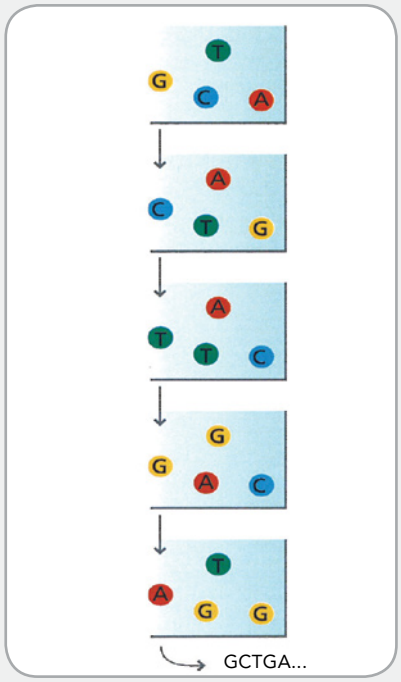
Second chemistry cycle: to initiate the next sequencing cycle, add all four labeled reversible terminators and enzyme to the flow cell.

10. IMAGE SECOND CHEMISTRY CYCLE



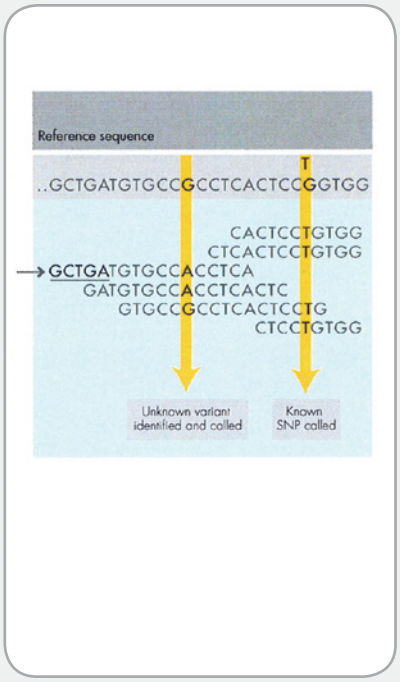
After laser excitation, collect the image data as before. Record the identity of the second base for each cluster.

11. SEQUENCE READS OVER MULTIPLE CHEMISTRY CYCLES



Repeat cycles of sequencing to determine the sequence of bases in a given fragment a single base at time.

12. ALIGN DATA



Align data, compare to a reference, and identify sequence differences.

DATA COLLECTION, PROCESSING, AND ANALYSIS

The Genome Analyzer 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 for easy access to the data at various stages of processing and analysis using simple application program interfaces.

EXPLORE THE POSSIBILITIES

Solexa sequencing technology achieves an unparalleled data density with highly accurate results. With the capability to generate over a billion bases of DNA sequence per run, the Illumina Genome Analyzer provides researchers with the opportunity to sequence mammalian genomes in a matter of 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 up to eight samples per flow cell, the Illumina Genome Analyzer can be tailored to the demands of many applications. Whether project requirements involve a genome, region, or gene, the Illumina Genome Analyzer is the ideal tool for sequencing projects.

ADDITIONAL INFORMATION

Visit our web site or contact us at the address below to learn more about DNA sequencing with Solexa technology or other Illumina products and services.

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

FOR RESEARCH USE ONLY

© 2007 Illumina, Inc.
Illumina, Solexa, Making Sense Out of Life, Oligator, Sentrix, GoldenGate, DASL, BeadArray, Array of Arrays, Infinium, BeadXpress, VeraCode, IntelliHyb, iSelect, and CSPro are registered trademarks or trademarks of Illumina. All other brands and names contained herein are the property of their respective owners.
Pub. No. 770-2007-002 01May07

illumina®