Information Technology Solutions for the Genome Analyzer

Illumina provides robust, preconfigured computing hardware packages to quickly implement an easy-to-use sequence analysis pipeline.

Highlights Of Illumina Computing Solutions For The Genome Analyzer

- **Ready to Use:**
  Eliminates the need for dedicated IT personnel
- **Validated:**
  Specifically designed for the Illumina Genome Analyzer
- **Integrated:**
  End-to-end solutions for meaningful sequence analysis
- **Powerful:**
  Scalable, high-performance computing system

Introduction

The Genome Analyzer can read billions of bases in under a week. To simplify the analysis of the massive quantity of sequence data generated, Illumina provides preconfigured analysis hardware designed for easy setup and use. These modular, scalable informatics solutions accommodate future growth in sequence throughput. These products contribute to an end-to-end sequencing system to rapidly move from raw data acquisition to publishable, biologically meaningful results without the need for dedicated IT personnel.

The IPAR (Integrated Primary Analysis and Reporting) system facilitates real-time image analysis of the primary data output. This allows users to optimize instrument configuration early in the run and vastly reduce the time from sample prep to obtaining useful reads.

After image analysis, the Pipeline software automates base calling and read alignment to a reference sequence. To start sequencing as soon as possible, users can purchase a preconfigured server computer from Illumina with the Pipeline preinstalled.

Both IPAR and the Analysis Server are described in more detail below.

IPAR

IPAR is a hardware and software package that provides real-time quality control and integrated online processing of primary data during sequencing runs. The IPAR system hardware comprises a four-core processing module and 3 TB of usable storage in RAID6 configuration.

Real-Time Quality Metrics

IPAR provides real-time metrics to evaluate performance at each incorporation cycle. Visualization tools help users understand trends in the data during a run (Figure 1) to maximize read length and detect problems in time to act.

Integrated Data Analysis

Preliminary data analysis steps can be performed by IPAR before run completion to reduce the time required for analysis after a run. IPAR also offers the option to reduce the burden of long-term data storage by transferring only digital intensity files of detected clusters.

Automated Data Migration

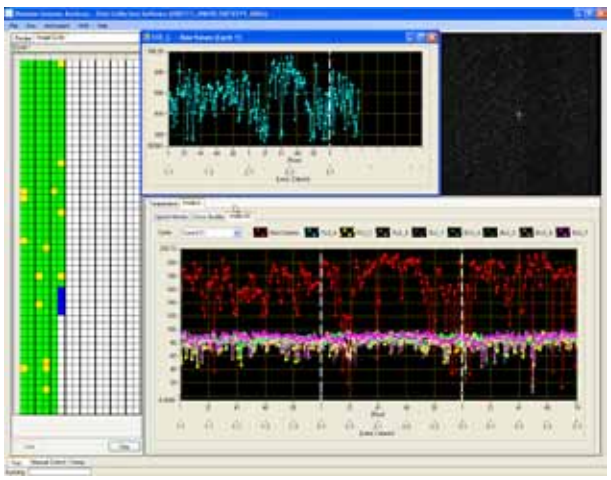
To maximize the efficiency of routine massive data transfer sessions, IPAR includes user-configurable data transfer tools. IPAR can then automatically move data after a sequencing run.

By seamlessly completing a majority of the computationally intensive primary data transformation, IPAR gives researchers usable data for secondary analysis in the least amount of time and with minimal effort.

Pipeline Analysis Server

Pipeline software can be installed on existing IT infrastructure. Since not all users have the IT infrastructure necessary to process the data from the Genome Analyzer, Illumina offers the Pipeline Analysis Server. This is an optional solution for analyzing the images and intensity data generated by the Genome Analyzer.

Figure 1: IPAR Real-Time QC Metrics Display



The IPAR screen displays called-base intensities as a function of cycle. A variety of other QC metrics can be displayed, including tile-by-tile run progress (left panel), tile- and cycle-based measurement statistics, such as cluster intensity and focus quality (top-center, and lower-right panels), and the selected tile image (upper-right).

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA

Table 1: Genome Analyzer Pipeline Server Specifications

HP ProLiant DL580 G5 Rack Server

- **Four quad-core 2.93GHz 64-bit Intel Xeon processors**
- **32 GB RAM**
- **HP Red Hat Enterprise Linux**
- **HP Standard Limited Warranty - 3 Years**

HP StorageWorks 60 Modular Smart Array

- **12 x 750 GB SATA 7,200 rpm drives (9 TB total)**

The Illumina Pipeline Server consists of two modules configured for use with the Genome Analyzer, the HP ProLiant Server for processing, and an HP Modular Smart Array for storage (Table 1).

Validated Hardware System

Users looking for a ready-to-use preconfigured environment for running the pipeline can purchase the Pipeline Server from Illumina. Since the hardware system has been validated by Illumina for use with the Genome Analyzer, researchers can be confident that they will be up and running after installation by Illumina.

Scalable Investment

The Smart Array 20 has the capacity to meet most users' needs for storing 3–4 runs. As the Genome Analyzer evolves, producing more sequence data per run, researchers may desire more storage. The Pipeline Server can be scaled to users' needs by the addition of an unlimited number of SA20 modules.

Preinstalled and Preconfigured

Illumina Sequencing technology has created new opportunities for anyone interested in massively parallel sequencing, from the largest genome core facilities to single principal investigators. This has created a need for even small labs to have access to sophisticated informatics computing power. With a single package, labs can purchase all the computing power needed for primary analysis of Genome Analyzer data, preconfigured and installed. The bundling of hardware, software, and support offers labs the most cost-effective sequencing data analysis infrastructure.

Summary

The Genome Analyzer is supported by a complete portfolio of validated and ready-to-use analysis tools. The informatics hardware consists of the online integrated data analysis tool, IPAR, and the offline data analysis tool, Pipeline Server. All of Illumina's informatics solutions are highly customizable to perform data analysis in a manner that is consistent with virtually any workflow. These easy-to-use and flexible components combine to ensure that the Genome Analyzer is truly a genome center in a box.

Ordering Information

Catalog No.	Product	Description
SY-301-1003	Genome Analyzer Pipeline Server	Includes the HP ProLiant Server for processing and HP Modular Smart Array for storage.

Illumina, Inc. • 9885 Towne Centre Drive, San Diego, CA 92121 USA • 1.800.809.4566 toll-free • 1.858.202.4566 tel • techsupport@illumina.com • illumina.com

FOR RESEARCH USE ONLY

© 2010 Illumina, Inc. All rights reserved.
Illumina, illumina*Dx*, Solexa, Making Sense Out of Life, Oligator, Sentrix, GoldenGate, GoldenGate Indexing, DASL, BeadArray, Array of Arrays, Infinium, BeadXpress, VeraCode, IntelliHyb, iSelect, CSPro, GenomeStudio, Genetic Energy, HiSeq, and HiScan are registered trademarks or trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners. Pub. No. 770-2008-005 Current as of 25 September 2008

illumina®