

**For more information, contact:**

Complete Genomics, Inc.  
Jennifer Turcotte  
Vice President of Marketing  
(650) 943-2846  
[jturcotte@completegenomics.com](mailto:jturcotte@completegenomics.com)

Waggener Edstrom Worldwide Healthcare  
Lisa Osborne  
Account Director  
(202) 261-7806  
[lisao@waggeneredstrom.com](mailto:lisao@waggeneredstrom.com)

## **Complete Genomics Releases Proof-of-Concept Data for Its Sequencing Technology for the First Time**

*Data corroborate the company's promise to deliver high-quality, high-throughput commercial human DNA sequencing at a low cost*

**MOUNTAIN VIEW, Calif. — Feb. 6, 2009** — Complete Genomics, Inc. released proof-of-concept (PoC) data for its human genome sequencing technology for the first time yesterday evening. The analysis results were reviewed by Dr. Clifford Reid, Complete Genomics chairman, president and CEO, during his presentation at the annual Advances in Genome Biology and Technology Meeting held at the Marco Island Marriott Beach Resort, Marco Island, Fla.

Complete Genomics successfully sequenced a Caucasian HapMap sample (Coriell catalog # NA07022; cell line DNA), generating 91x average read coverage of the genome in a matter of days.

This represents the first human genome sequence produced with a third-generation technology and the work was conducted in-house at Complete Genomics' genome center.

Complete Genomics' system delivered unprecedented throughput, producing 254 Gigabases (Gb) of mapped data (reads), which is the most reported for one human genome.

This technology also demonstrated an average run rate of more than 70 billion mapped bases (70 Gb) per run or 8.8 Gb per machine run per day. The entire sequencing process required nine machine runs with a single run taking just eight days. Furthermore, this analysis was conducted on data generated by Complete Genomics' research and development sequencers; the company's production throughput is expected to increase three fold (up to 200 Gb per run) following its commercial launch in June 2009.

"We were able to make high-confidence base calls for 92 percent of the genome," said Dr. Rade Drmanac, chief scientific officer at Complete Genomics. "As expected, the 8

percent that we did not call included long repeats and duplications, which are difficult for all short-read technologies to sequence. We were able to call alleles for both parental chromosomes for 91 percent of the genome. Sequencing this remaining fraction of DNA will require our Long Fragment Read (LFR™) technology addition that is currently being implemented.”

“In a draft assembly, we discovered the expected 3.3 million single-nucleotide polymorphisms (SNPs) and more than 384,000 short (<10b) insertions and deletions. In our analysis, we identified more than 396,000 novel candidate SNPs, which we plan to contribute to the scientific community through dbSNP,” Drmanac added.

Complete Genomics’ sequencing platform employs high-density DNA nanoarrays that are populated with DNA nano-balls™ and uses a non-sequential unchained read technology, called combinatorial probe-anchor ligation or cPAL™, which reduce both reagent consumption and imaging time. These innovations allow genome sequencing at a higher throughput and at a lower cost.

“We are delighted to be demonstrating an initial milestone in advance of our service launch that supports our ability to deliver high-accuracy, high-throughput, low-cost DNA sequencing. Our sequencing service will help researchers to identify the rare genetic variants that play a significant role in drug responses and complex diseases such as cancer,” said Dr. Reid.

“This marks a major achievement for the team at Complete Genomics — they have sequenced a human genome at a high quality and low cost, which surpassed expectations,” said Dr. George M. Church, professor of genetics at Harvard Medical School and director of the Center for Computational Genetics. “My team, having reviewed variation calls from this genome data set, confirmed that it falls in line with what is expected of an individual genome. It is highly concordant with previously published work on this genome and with data from public variation repositories.”

To enable the scientific community to analyze its unique genome sequence data set further, Complete Genomics has sent reads (>350Gb) and base quality measures to the National Center for Biotechnology Information for inclusion in its public database. These data and a technology white paper are also available through Complete Genomics’ Web site at <http://www.completegenomics.com/datarelease/>.

In preparation for its service launch, Complete Genomics is rapidly scaling up its commercial genome center. It plans to sequence 1,000 genomes in the second half of 2009 and 20,000 genomes in 2010. To analyze the enormous amounts of data that will be created, it is also expanding its data center, which will house 5,000 processors and provide five petabytes (5 million gigabytes) of disk storage by the end of 2009, and 60,000 processors and 30 petabytes of disk storage in 2010.

### **About Complete Genomics**

Founded in 2006, Complete Genomics is a California company that has developed a novel approach to sequencing human DNA. Complete Genomics plans to combine its proprietary third-generation DNA sequencing technology and its high-performance computing capabilities to create a commercial human genome sequencing service that will deliver low-cost, high-quality data on an unprecedented scale. The company is currently building the world's largest human genome sequencing center. This development will allow pharmaceutical and biotechnology customers, for the first time, to conduct large-scale human genome studies that will help identify the genetic underpinnings of complex diseases and drug responses. For additional information about the company, please visit <http://www.completegenomics.com>.

# # #