Complete Genomics, Inc.

Initiating Coverage With an Outperform Rating

Unique model enables cost reduction by leveraging scale efficiencies. Complete Genomics offers a full-service model based on its proprietary sequencing technology. Its key differentiator is its sole focus on whole-human-genome sequencing. Ultimately, this has enabled Complete Genomics to build instruments and workflows that are relatively inflexible but highly economical in terms of reagent usage and instrument efficiency, driving down ASP while providing accuracy levels in line with or better than competing instruments.

Validated approach and large orders from major research institutes lend credibility to the model. With peer-reviewed data now published from early adopters and a building customer list and backlog, CGI's platform has been increasingly validated and endorsed by the research community—further confirmed by our recent channel checks. As of early March 2011, the company had more than 45 customers, had sequenced 1,000 human genomes (roughly one-half of the total whole human genomes sequenced in the world last year), and had a backlog of just under 2,000 genomes, suggesting its model is gaining traction.

We expect demand for outsourced sequencing to accelerate. DNA sequencing has become a critical tool in understanding, diagnosing, and treating disease. Although sequencing costs have declined exponentially, the up-front investment and production cost outlays (particularly for whole human genomes) remain significant. We believe the expense and IT horsepower needed to analyze DNA sequence data will drive demand for the outsourced model. As Complete Genomics continues to reduce its pricing, we see three primary markets it can penetrate: research, drug discovery/clinical trials, and clinical diagnostics. We expect CGI to sequence a cumulative 69,000 genomes through 2013, representing 22% market penetration.

Stock thoughts/valuation. We view Complete Genomics as a compelling way for small-cap investors to participate in the high-growth sequencing market and a diversification mechanism for investors concerned about researchers' ability to secure funding for large-box purchases. Our DCF analysis (based on a weighted average cost of capital of 14.8% and a terminal growth rate of 2%) suggests an intrinsic value of $18.45 (a 58% premium to where the stock is trading on April 14).

Investment risks. These include 1) inability to drive down ASP, which could hamper adoption; 2) an increasingly competitive service provider landscape; and 3) significant execution risk given its early stage of commercialization.

Amanda Murphy, CFA
312.364.8951
amurphy@williamblair.com

Sylvia Chao
312.364.8654
schao@williamblair.com

Healthcare | **Life Sciences**

April 18, 2011
Basic Report (11-044)

Stock Rating:	**Outperform**
Company Profile:	**Aggressive Growth**
Symbol:	GNOM (NASDAQ)
Price (4/14):	$11.70 (52-Wk.: $8-$12)
Market Value (mil.):	$303
Fiscal Year End:	December
Dividend Yield:	None

Estimates	**2010A**	**2011E**	**2012E**
EPS FY	($13.60)	($1.97)	($0.96)
Sales (mil.)	$9.4	$34.2	$89.8
Valuation			
EV/Sales	26.3x	7.2x	2.8x

Trading Data	
Shares Outstanding (mil.)	26.0
Float (mil.)	0.9
Average Daily Volume (thous.)	127

Financial Data	
Debt/Total Capital	12%
Book Value Per Share	$16.24
Enterprise Value (mil.)	$247
EBITDA (mil.)	($39.2)
Enterprise Value/EBITDA	NM
Return on Equity	NM

William Blair & Company, L.L.C.
222 West Adams Street
Chicago, Illinois 60606
312.236.1600
www.williamblair.com

Complete Genomics offers whole-human-genome sequencing services using its proprietary sequencing platform. Incorporated in 2005, CGI began commercial operations in 2010 and raised more than $50 million in an IPO in late 2010.

Please consult pages 44-45 of this report for all disclosures.

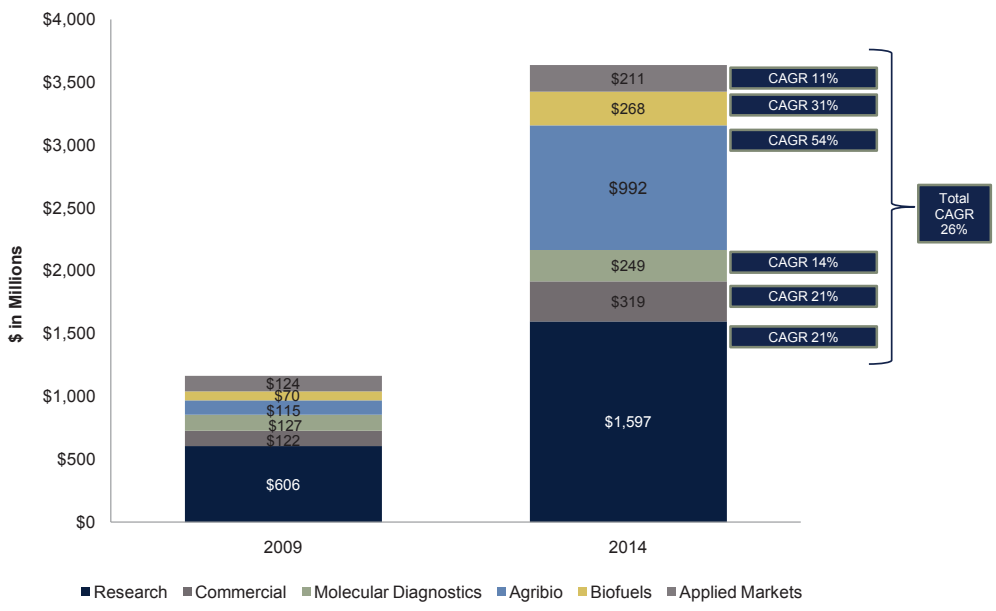
Contents

Investment Summary

Complete Genomics is a DNA sequencing service provider based in Mountain View, California. While most vendors sell sequencing instruments via a razor/razorblade model, Complete Genomics has chosen a full-service approach, allowing customers to outsource sample prep through sequencing and data analysis. Customers send samples to Complete Genomics' lab in Mountain View; in return, they receive a research-ready, sequenced genome with fully annotated variant analysis. The whole process takes place within 90-120 days and is getting ever shorter. Complete Genomics specifically focuses on the sequencing of whole human genomes (the company requires a minimum of eight whole genomes per order) with one input, human DNA, and one output, a standardized genome report. The company was incorporated in 2005, began commercial operations with the launch of its Complete Genomics Analysis Service in 2010, and raised $54 million in an initial public offering in November 2010.

DNA sequencing is used to determine the order of nucleotide bases—adenine (A), cytosine (C), thymine (T), and guanine (G)—of a DNA strand, which carries information for building and maintaining an organism. Human DNA contains 3 billion base pairs, 99.9% of which are identical among individuals and across populations. The 0.1% that varies drives unique genetic traits or phenotypes (e.g., hair color and height), as well as predispositions to medical conditions. As a result, researchers have a strong interest in identifying and comparing genetic variation across individuals and populations as well as within and between other species of plants and animals to better understand the heritability and mechanism of disease. *The sequencing market has become increasingly important,* both from a life sciences research perspective and in applied markets (including agricultural biotechnology and diagnostics). Scientia Advisors, a life sciences consulting firm, estimates the sequencing market to be roughly $1.2 billion and *growing at an annual compound rate of 20% to 25%.*

Figure 1
Sequencing Market (2009 to 2014)



Source: Scientia Advisors

DNA sequencing methods were first developed in the late 1970s by two labs, Allan Maxam and Walter Gilbert's lab at Harvard and Frederick Sanger's lab at Cambridge University. Applied Biosystems (ABI, now part of Life Technologies) was the first to commercialize sequencing technology in 1986, leveraging Sanger's chain-termination sequencing methodology

to introduce an automated sequencing platform based on slab electrophoresis. Later, ABI replaced the slab gel with capillary electrophoresis (CE) and remains the primary vendor of Sanger/CE-based machines, which labs continue to use to this day.

By leveraging better chemistry and massive parallelism, next-generation sequencing platforms first introduced by 454 Life Sciences (now a subsidiary of Roche) in 2005 dramatically improved the throughput (the amount of DNA sequenced per unit of time) relative to Sanger/CE sequencing. Over the past five years, scientists have improved these technologies and platforms, leading to a dramatic reduction in sequencing cost. As an example, the Human Genome Project, which sought to sequence the human genome for the first time, took 13 years and cost more than $3 billion (the sequencing component was estimated to cost roughly $500 million). Now, Illumina and Complete Genomics offer whole-genome sequencing to consumers for less than $10,000.

Despite the reagent cost improvements, however, sequencing instruments involve significant capital investment (from a few hundred thousand dollars to close to $1 million) and ongoing operational and maintenance costs (including depreciation; human resources, including technicians and bioinformaticians; ongoing reagent/consumable costs; and IT storage and processing expenses, to name a few). Typically, when the cost to sequence a genome is quantified (e.g., the race to offer a $1,000 genome), estimates only include consumables expenses, such as flow cells and reagents. In reality, however, the cost per genome ends up closer to $30,000 after taking into account labor and overhead, data analysis, and storage and project management costs, in addition to the up-front capital investment of a sequencer (which is typically depreciated over a period of three years or less).

Ideally, researchers would sequence the entire genome of the population of interest, which would provide the most comprehensive set of information. While the cost to sequence a genome is orders of magnitude lower than even a couple of years ago, the up-front investment and all-in production expenses are still cost prohibitive for many players. Many labs still use Sanger/CE-based sequencing, which requires a lower capital investment and provides lower throughput but produces highly accurate data that is much more manageable. These machines allow researchers to sequence small pieces of DNA. For example, labs might use CE machines to resequence candidate markers identified through other means (e.g., polymerase chain reaction), which involves sequencing 100 base pair DNA fragments versus the 3 billion nucleotides of the human genome.

When next-generation platforms came online, cost per base sequenced was reduced even further. With the introduction of the HiSeq 2000 by Illumina in early 2010, the cost to sequence a genome dropped by 62% in merely 12 months, from around $47,000 to $29,000. Larger labs that had funding to purchase new sequencing platforms and the capacity to manage the bioinformatics requirements of data produced were able to expand the amount of sequence targeted. As an example, although whole-human-genome sequencing is still relatively expensive on a wide scale, researchers have used the next-generation platforms to perform whole-exome sequencing (which includes just the protein coding regions of the genome, representing about 1% of the total but including most of disease-causing mutations). There are no solid estimates of the number of exomes currently being sequenced, but we know that whole-exome sequencing is still predominantly performed in the genome centers and large core labs. Estimates of the number of exomes expected to be sequenced in 2011 measure in the tens of thousands. Although whole-human-genome sequencing is still mostly completed in the large labs, the sequencing market has proved to be highly elastic—as the ASP continues to fall, researchers have sought to sequence more and more of the genome.

Given the rapid growth in the sequencing space and the desire to sequence samples from microbes to humans, a number of service providers have emerged to address the more capital- and cost-sensitive segment of the market. Illumina, for example, established its Illumina Genome Network in 2010, which offers whole-genome sequencing among other services through a network of service providers. BGI in China purchased 128 Illumina HiSeqs in early

2010 and has a total of 137 Illumina HiSeq 2000 machines, according to its website, as well as 27 SOLiD 4 sequencers, and has become the largest service provider in the world. These providers enable researchers to pay on a fee-for-service basis for DNA sequencing without having to invest in the capital equipment and resources to run the machines.

While the service provider market has become increasingly competitive, Complete Genomics' focus on whole-human-genome sequencing and use of its proprietary sequencing technology allows the company to provide a differentiated and lower-cost model to customers. Because CGI is a service provider and operates its machines internally (with 16 instruments in production), it has not had to invest in self-contained automation, miniaturization, and user friendliness of its machines as compared with vendors that sell instruments directly to users. Rather, Complete Genomics has been able to invest research and development in further optimizing and enhancing its platform specifically for whole-human-genome sequencing.

Ultimately, this has enabled CGI to build an instrument and workflow that is relatively inflexible but highly economical in terms of reagent usage and instrument efficiency, while providing accuracy levels in line with or better than other instruments on the market. As a service provider, it is able to run many genomes in batches, which drives operating efficiencies and leverage from increased sample flow given the high fixed-cost nature of its business model. Ultimately, this leads to the ability to continually lower the company's production cost per genome, and in turn, ASP. The company's unique model has enabled it to lower its reagent cost per genome, which is now less than $1,000 a genome—significantly lower than Illumina, at $5,000 per genome. Complete Genomics' current ASP per genome is less than $10,000 (in line with BGI and the Illumina Genome Network), and with further improvements to its arrays, it expects to exit 2011 at an ASP of $5,000.

While there was initially some skepticism around the adoption and sustainability of the company's business model, our recent channel checks with researchers suggest that Complete Genomics has become an accepted player in the sequencing community. With data now published from early adopters and a building customer list and backlog, the company's platform has been increasingly validated and endorsed by the research community. A number of peer-reviewed journal articles have leveraged Complete Genomics' platform, which has helped support the accuracy (at 99.999% and above, which is at least in line with sequencing platforms on the market) and coverage (the average number of times a nucleotide is represented in the data, which is important for accuracy) of its data. The company has also amassed a notable customer list, including the primary research-and-development contractor for the National Cancer Institute, as well as the Institute for Systems Biology (ISB), which has placed three orders with CGI for a total of 719 genomes. Lastly, Complete Genomics has a backlog that has been consistently building (almost 2,000 genomes as of early March). Although the company has sequenced fewer than the initial 5,000 genomes promised in 2010 (it has sequenced 1,000 genomes as of March 2011), its model is clearly beginning to gain traction.

We believe there are three primary markets that Complete Genomics can penetrate with its whole-human-genome sequencing services. Given that its ASP remains above $5,000, in the shorter term, research-based whole-genome sequencing is likely to remain the most meaningful market for Complete Genomics. In our view, once Complete Genomics drives its ASP below $5,000 (estimated by the end of 2011 into 2012), its sequencing services could become appealing to pharma and biotech companies looking to leverage sequencing in drug discovery and later in clinical trials without significant capital outlays. Based on the company's high accuracy rate, its recent addition of copy number and structural variant analysis (important in cancer research), and its cancer assembly software (with a commercial version expected in mid-2011), in our view, Complete Genomics' solution could be very attractive to pharma and biotech as sequencing becomes a larger component of clinical trials. Pharma continues to increasingly outsource research and development; although this has particularly been the case in the clinical side of the market, we also expect outsourcing of preclinical research to reaccelerate over the next five years. For example, we know that

Pfizer, Genentech, and Eli Lilly have used Complete Genomics' services; Merck has an agreement with BGI to collaborate on biomarker discovery and validation; and Illumina has suggested that its whole-human-genome service backlog includes a large, multi-hundred genome order from a pharmaceutical company.

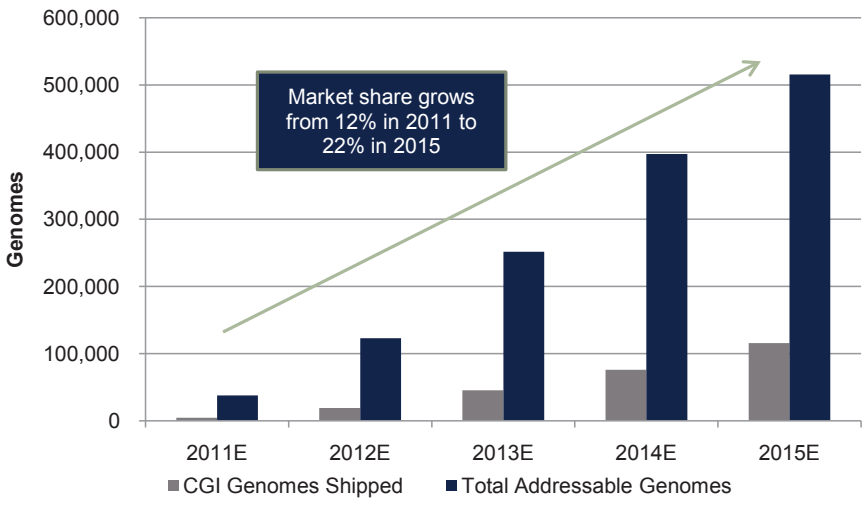
We believe the ultimate opportunity is to transition sequencing into the diagnostic realm. CGI's advantage in this market could be its high accuracy rates. With its new long fragment read technology (currently in development and expected to be commercialized by the end of 2012), the company has the ability to drive its accuracy rate to an unprecedented 99.99999%, which would reduce its error rate from 30,000 errors to 300 errors per genome. This accuracy level would be the highest available and ultimately necessary for use in the clinic. Given the current focus on cancer genomes, sequencing of tumors either to identify appropriate therapeutics or as a monitoring mechanism could become a clinical reality in the next three to five years, in our view. If sequencing technology is incorporated into cancer diagnosis and therapeutic decision-making and monitoring, markets would be sizable and defined by annual cancer incidence rates (1.5 million people are diagnosed with cancer in the United States annually) and could extend to prevalence estimates if tests are obtained on an ongoing basis.

Quantifying exactly how many genomes might be completed over the next five years is difficult given the elasticity of the demand curve. Our estimates for the market potential and the number of genomes Complete Genomics will ship through 2014 admittedly are somewhat speculative. Given the dramatic growth and adoption of sequencing, we believe the opportunity for outsourced whole-human-genome sequencing should continue to accelerate. We gain near-term comfort from the fact that Complete Genomics has a backlog of 1,900 genomes as of the end of March (which represents 40% of the genomes we estimate the company will ship in 2011). In addition, *Nature* magazine interviewed 90 genomics centers and labs to identify how many genomes they expect to sequence. From this, the magazine estimated 1,000 genomes at high coverage and 1,756 at low coverage would be completed by the end of October 2010, and another 30,000 should be completed by the end of 2011 ("Human genome: Genomes by the thousand," *Nature* 467, 1026-1027 [2010], published online 27 October 2010, doi: 10.1038/4671026a).

As discussed in more detail below, we assume the company will be able to garner 22% of the whole genome market by 2015, which we define as: 1) the base research market focused on sequencing whole human genomes; 2) the whole human exome market, which we assume converts to whole-human-genome sequencing at a certain price point; and 3) the oncology clinical trials market, which we assume becomes a viable opportunity once ASP drops below a certain price point (around 2013). We believe the oncology clinical trials market (which is estimated to be growing at a compound annual rate of close to 10%) could be a significant opportunity for the company, totaling around $300 million in revenue potential by 2015 (assuming that by that point, whole-human-genome sequencing is used for roughly 55% of patients enrolled). Given the uncertainty around the clinical diagnostic opportunity, we have not included this market in our estimates. We are modeling that the company ships roughly 4,600 genomes in 2011 (growth close to 470%), 19,000 genomes in 2012 (growth close to 310%), and around 45,000 genomes in 2013 (growth of 138%).

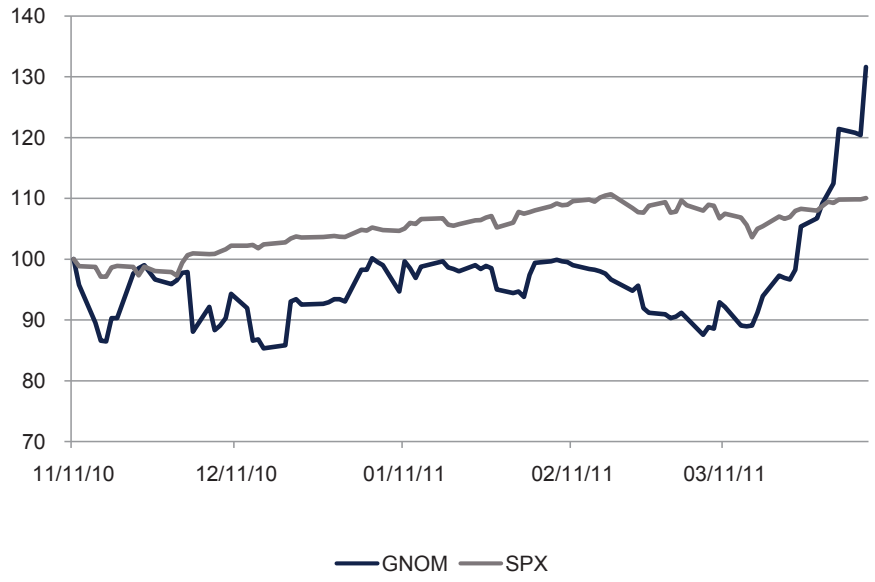
Although the company is in early stages and therefore carries significant commercialization risk, there are a number of indications that its model is gaining traction (including a building repertoire of peer-reviewed publications, customers, and backlog). We believe Complete Genomics offers a differentiated model that facilitates a continued reduction in cost for researchers, and thus is positioned to gain share of the developing whole-human-genome sequencing market. In our view, Complete Genomics represents a compelling way for small-cap investors to participate in the high-growth sequencing market and/or a diversification mechanism for investors concerned about the ability of researchers to secure funding for large-box purchases. Therefore, we are initiating coverage with an Outperform rating and an Aggressive Growth company profile.

Figure 2
Complete Genomics, Inc.
Addressable Market Penetration



Source: William Blair & Company, L.L.C. estimates

Figure 3
Complete Genomics, Inc.
Stock Performance (Normalized to 100)



Source: Thomson One

As illustrated in figure 3, the stock has performed well in recent trading (up 46% since March 14, 2011) and thus appears to reflect some of this excitement. We evaluated the stock on two different valuation metrics, enterprise-value-to-revenue ratio and our own discounted-cash-flow analysis.

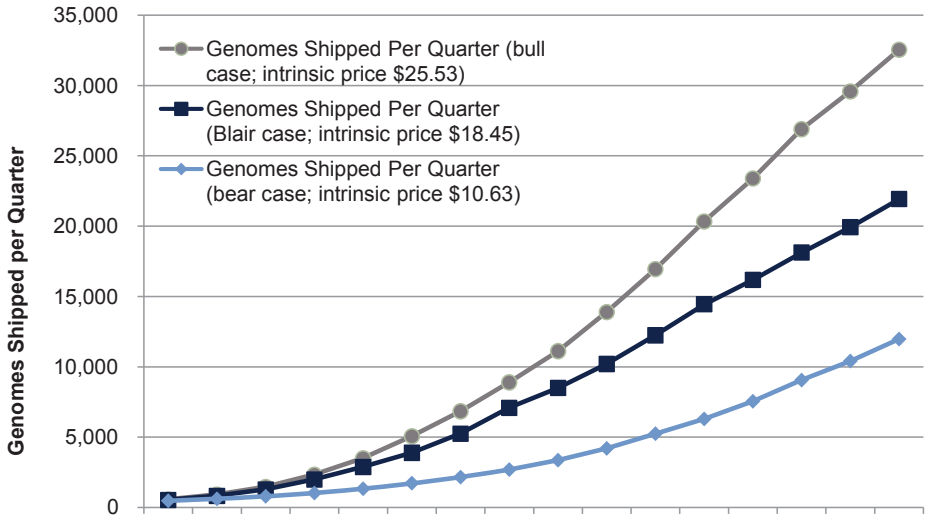
- In terms of enterprise value to revenue, it is difficult to identify a comparable peer group given Complete Genomics' sequencing competitors operate with an instrument/consumable model (e.g., Illumina and Pacific Biosciences) and are at much different stages in

terms of commercialization and maturity. Service-based companies in our coverage universe that have a proprietary offering service such as Myriad Genetics and Genomic Health represent one potential comparable group, while companies that service pharma such as the CROs represent another. On an EV/revenue basis, the stock is trading at 2.8 times our 2012 revenue estimate of $90 million. This is well below Illumina, which has the largest component of its business attributable to sequencing and is trading at 6.5 times our 2012 revenue estimate. Complete Genomics trades somewhat below Life Technologies, at 2.9 times, but Life also has multiple operating segments with varying growth rates. Myriad Genetics and Genomic Health trade at an average of 3.1 times our 2012 revenue estimates—also slightly above Complete Genomics. Companies that service pharma, including the CROs and tech companies such as Medidata, trade at an average 1.9 times, but many of these companies are more mature and thus have lower earnings growth potential. See table 7, at the end of this report, for more detailed information.

- Given the difficulty of evaluating Complete Genomics on a multiple basis because of a lack of true comparables and the company's early stage, we also used an intrinsic value approach. Our DCF analysis (based on a weighted average cost of capital of 14.8% and a terminal growth rate of 2%) suggests an intrinsic value of $18.45, or a 58% premium to where the stock was trading on April 14.

  With the high fixed-cost nature of this business, the company's ability to drive efficiency improvements on its hardware and software (e.g., successfully rolling out higher-density arrays) is key to driving its gross margin into positive territory. We estimate the company will begin to make money (gross profit) on a per-genome basis at the end of 2011. If this is delayed, it would increase the company's cash burn rate and delay our cash-flow-positive target of 2013 (and thus reduce our DCF valuation). In addition, there is clearly risk around the company's adoption curve, particularly as it relates to Complete Genomics' ability to drive down ASP and operating leverage from increased volume. Thus, we include a sensitivity analysis to our DCF model. See table 8, at the end of this report, for more detail around our DCF analysis.

Figure 4
Complete Genomics, Inc.
Adoption Sensitivity Analysis
(Intrinsic price based on discount rate of 14.8% and terminal growth rate of 2%)



Source: William Blair & Company, L.L.C. estimates

Investment Highlights

I. Unique Service Model Allows the Company to Drive Cost Reductions by Leveraging Scale Efficiencies

Most sequencing vendors sell instruments via a razor/razorblade model—benefiting from machine sales, which then translate into a recurring revenue stream of higher-margin reagent sales. Complete Genomics, however, has opted to pursue a full-service model, selling sequencing on a fee-for-service basis via its Complete Genomics Analysis (CGA) Service. The company runs its own proprietary sequencing platform based on cPAL (combinatorial probe-anchor ligations) chemistry on DNA nanoball arrays, which is protected by exclusive patent license, some non-exclusive licenses, and 106 pending patent applications (owned or licensed). The company has optimized its instrumentation and workflow specifically to sequence whole human genomes, which provides both longer-term cost and accuracy advantages.

From a workflow standpoint, as illustrated in figure 5, Complete Genomics receives samples, completes the sequencing and analysis, and then delivers sequencing data and variant calls to the customer via the cloud (Amazon Web Services) and/or hard drives delivered by FedEx. The company delivers research-ready data, which includes the sequence at high coverage (referring to the number of times each base is represented in the final sequence data, which helps improve the accuracy of the finished sequence), with more than 95% of the genome fully identified, or "called," as well as annotated variant analysis. Complete Genomics also stores the data for 30 days in addition to delivering the hard drive.

Figure 5
Complete Genomics, Inc.
Whole-Genome Sequencing Workflow

Source: Company reports

Cost advantage. The benefit of Complete Genomics' service model is that the company does not bear the cost of service infrastructure support (including field installation and service support staff). In addition, Complete Genomics has not had to invest in the self-contained automation, miniaturization, and user friendliness of its machines as do the vendors that

sell instruments directly to users. Instead, the company has been able to invest research and development in further enhancing its platform specifically for whole-human-genome sequencing, which has resulted in a relatively inflexible but highly economical model.

Complete Genomics operates 16 instruments, which have the capacity to sequence one genome a day, or 150 gigabases per day. This throughput is greater than Illumina's powerhouse, the HiSeq, which will be able to process 75 gigabases a day with its most recent upgrade (expected in the second quarter). At the same time, although sequencing by ligation is not a new technology, Complete Genomics has developed proprietary sample preparation (DNA nanoballs) and sequencing (cPAL) technology, in addition to improved array density and imaging speed, which has led to increased productivity and lower reagent costs, as well as improved data analytics capabilities.

The company benefits from batching both sample prep and sequencing. Genomes are prepared in batches of 88 loaded into a 96-well plate. Two plates can be run at a time for a total of 196 genomes. Based on current sample prep capacity, the company can complete a sample prep run per week, or four per month, suggesting a total capacity of 704 genomes per month. Note this is higher than the company's current sequencing capacity of 400 genomes per month and points to the fact that its sample prep can keep pace with throughput improvements. Additional lines of sample prep can be added to further increase sample prep capacity (from 704 to 1408 genomes per month, etc.) Complete Genomics' instruments are also optimized to batch genomes; the machines have 18 flow cells (versus Illumina's 2 flow cells), which ultimately yields 150 gigabases of sequencing data per 12-day run. Thus, via its service model, the company is able to benefit from scale efficiencies, allowing fixed costs to be spread over a number of genomes and ultimately lowering reagent usage per genome.

Figure 6
Complete Genomics, Inc.
Complete Genomics Sequencing Instrument

Source: Company report

As a result, CGI can offer whole-genome sequencing with at least 40-fold coverage (i.e., each nucleotide is represented on average 40 times in the sequence data) and thus high accuracy (with each nucleotide represented 40 times, the final assembled sequence is

99.999% accurate), all at a lower cost. Complete Genomics currently charges less than $10,000 per genome and roughly $5,000 per genome for large-scale customers as of our latest data point. The company can sequence a genome now for less than $1,000 in reagent costs and is targeting exiting 2011 at an all-in ASP of $5,000 per genome—which is impressive, considering the cost to sequence a genome was in the hundreds of thousands of dollars a couple of years ago.

See the "Platform Overview" section for more details about Complete Genomics' platform.

At a high level, sequencing by ligation (used by Complete Genomics and serving as the basis for Life Technologies' SOLiD platform) leverages the enzyme ligase, which catalyzes the joining of two DNA molecules and subsequently identifies the series of nucleotides. On the other hand, sequencing by synthesis (used in Illumina's and 454's technology) uses a polymerase enzyme, which identifies the nucleotide sequence as the polymerase assists fluorescent-labeled reversible terminators to form the complementary strand of the DNA sample.

The following are examples (also illustrated in figure 7, on the following page) of how Complete Genomics has been able to optimize its platform for whole human genomes:

- At Complete Genomics, the reaction is conducted on one-inch by three-inch silicon substrate, referred to as "flow slides," that allows highly parallel sequencing (in addition to batching of genomes, increasing throughput).

- The technology generates short reads (2 × 35 base pairs); however, as illustrated in figure 7, CGI has been able to leverage the benefits of sequencing by ligation (higher accuracy, lower reagent usage) while overcoming some of the limitations (super-short read lengths).

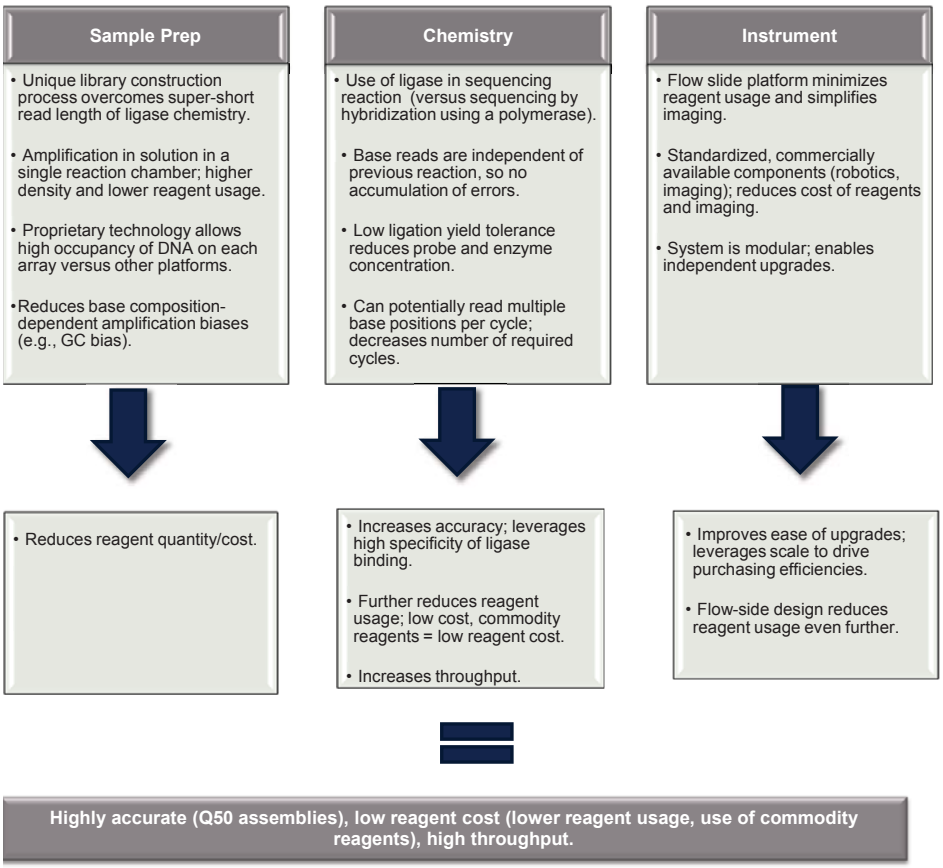
- Complete Genomics has also made a number of improvements to its consumables. For example, the company has successfully increased the density of its arrays, increasing throughput. In addition, it is one of the only companies to successfully achieve high occupancy levels—exceeding 90%, which is another key mechanism to increase throughput/reduce reagent usage. Occupancy refers to the ability to ensure only one amplified DNA fragment is present per reaction chamber.

- Currently, fluids are delivered to the flow slides via basic robotic pipetting. The company is in process of developing a specially designed flow cell platform (which houses the 1 × 3 arrays discussed above), with microchannels that facilitate efficient reagent delivery and imaging.

There are number of key levers the company can use to further push down the cost barrier, including additional chemistry and hardware improvements, more automation in the sample prep to reduce labor requirements, and further leveraging scale to lower marginal cost.

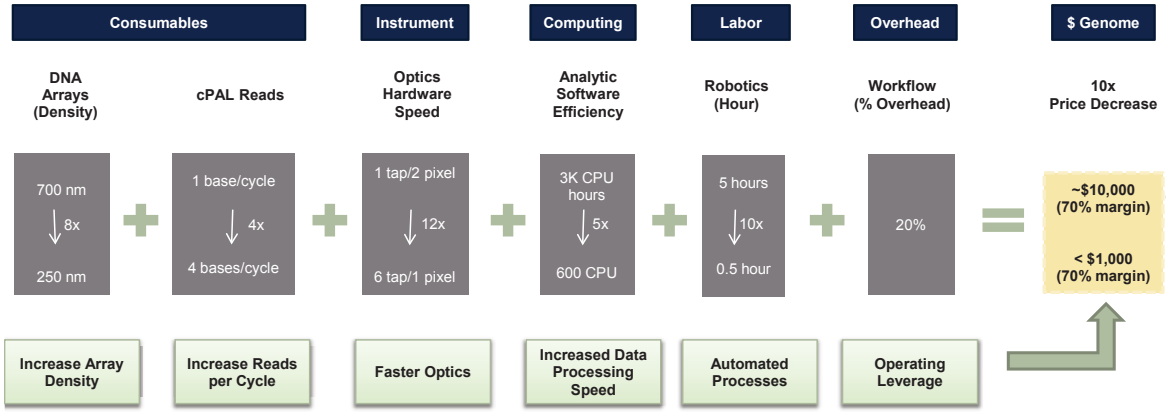
In the nearer term, the company is focused on increasing imaging throughput by improving the architecture of its array. The improved chip design reduces the number of pixels per spot from two pixels to one, effectively improving array density from 700 nanometers (nm) to about 500 nm. Thus, CGI will be able to increase the number of genomes sequenced per instrument per day from the current 1 genome to 1.5 to 2 genomes for the same reagent cost—yielding a similar reduction in reagent cost per genome. In addition, CGI is adding 8 instruments, which equals a 50% increase to its current fleet of 16. In combination with the increased imaging throughput, CGI expects to increase monthly capacity by 2 to 3 times (from 400 genomes per months to 800 to 1200 genomes per month) by the end of 2011.

Figure 7
Complete Genomics, Inc.
Advantages of DNB and cPAL technology



Source: Company reports

Figure 8
Complete Genomics, Inc.
Long-Term Potential to Process Optimization



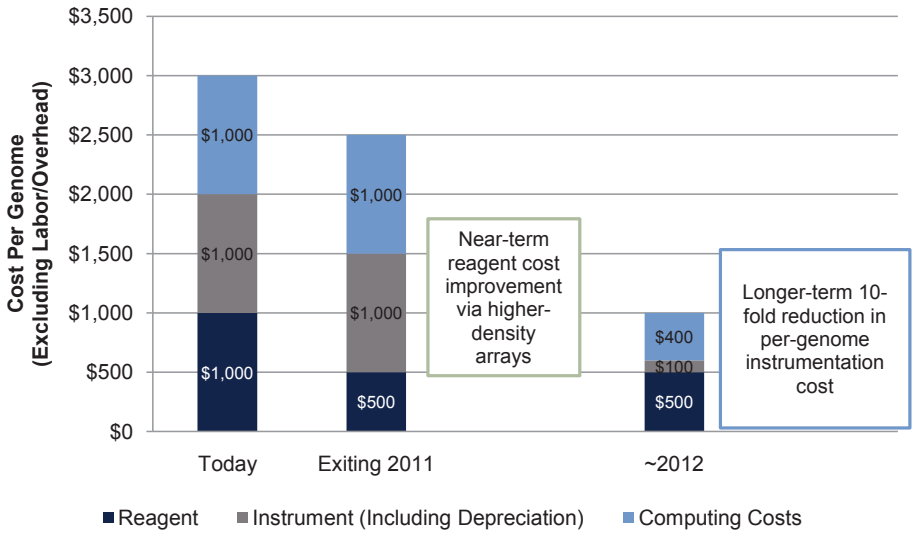
Source: Company reports

The company is also working on a second generation of the sequencing instrument, which should be able to capture signals 10 times faster (by switching to faster cameras and installing four cameras instead of the current two) than the current platforms. In theory, a 10-fold

increase in throughput would drive the instrument cost down from $1,000 today to $100 per genome. The company is targeting completion of prototypes for these machines in 2011 and expects them to be ready for commercial use in 2012.

CGI has also been exploring the possibility of further decreasing the distance between spots on the arrays from 700 nm to 400 nm, which would increase the number of spots per chip from 3 billion to 9 billion. These arrays would work on the newer instruments and could further reduce reagent cost/genome if implemented.

Figure 9
Complete Genomics, Inc.
Cost Improvements



Sources: William Blair & Company, L.L.C. estimates

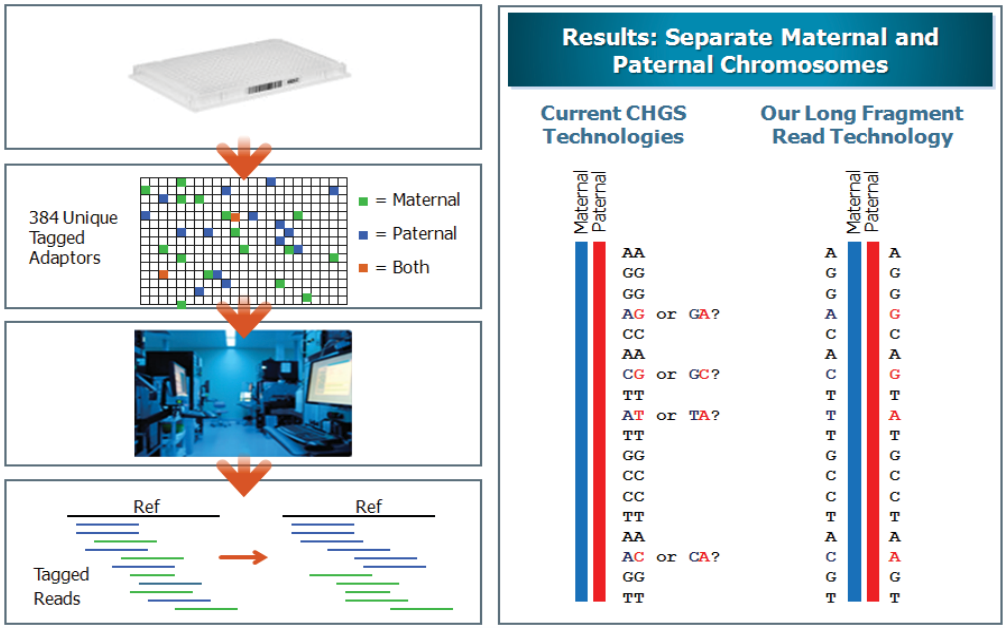
Accuracy. In peer-reviewed publications, Complete Genomics has demonstrated 99.999% accuracy (or 1 error in 100,000 bases called), which is compatible with Sanger sequencing. Via its proprietary library-construction process, the company has been able to overcome one of the limitations to sequencing by ligation chemistry (or read lengths that are typically restricted to six base pairs). First, CGI has been able to increase the read length to 10 base pairs from the point of ligation. Also, the company inserts four known adaptor sequences into each DNA fragment during the sample prep process, which allows the ligase enzyme to read the 10 base pair segments adjacent to each adapter. This ultimately yields read lengths of 35 base pairs.

As a result, despite its relatively shorter read length, the company is able to take advantage of the high accuracy of the ligase (versus the polymerase used in sequencing-by-synthesis methods of other sequencing platforms). In addition, it leverages its high throughput to increase coverage of each nucleotide base sequenced, thereby driving consensus accuracy of 99.999%.

In the future, we expect Complete Genomics' long-fragment read technology and haplotyping service offering will increase accuracy by 100-fold, to 99.99999% (1 error in 10,000,000 bases). The improvement suggests a decrease from roughly 30,000 errors per whole human genome sequenced to roughly 300 errors. This accuracy rate would surpass other available sequencing platforms and is expected to be in beta testing with two early-access customers by the end of 2011; it is expected to be introduced commercially in 2012.

Haplotyping is a method used to identify the DNA sequence from each chromosome in a pair. Chromosome pairs are made up of two copies of DNA strands—one strand inherited from the maternal side and the other from the paternal side. With existing technology, researchers are unable to identify from which parent the specific mutations originated. Through Complete Genomics' haplotyping method, DNA samples are separated and indexed in sample prep steps. The indexed DNA then follows Complete Genomics' regular sequencing protocol. Additional software functionality will be added to the data analysis step to identify the parental indices and thus assemble genome sequences with parental information.

Figure 10
Complete Genomics, Inc.
Haplotyping



Source: Company reports

II. Validated Approach and Large Orders From Major Research Institutes Lend Credibility to Complete Genomics

With data now published from early adopters, and with a building customer list and backlog, the company's platform has been increasingly validated and endorsed by the research community. While there was initially some skepticism around the adoption and sustainability of the company's business model, our recent channel checks with researchers suggest that Complete Genomics has become an accepted player in the sequencing community.

A number of peer-reviewed journal articles have leveraged Complete Genomics' platform, which has helped support the accuracy (at 99.999% accuracy and above, which is at least in line with other sequencing platforms on the market) and coverage of its data (averaging greater than 40-fold). In addition, Complete Genomics has seen more than 350 downloads of the 60 genomes that it made public in early January; thus, we would be not be surprised to see more presentations, abstracts, and publications evaluating Complete Genomics' data quality over the next year.

The company has also amassed a notable customer list, which includes the primary research-and-development contractor for the National Cancer Institute as well as the Institute for Systems Biology (ISB), which has placed three orders for whole-genome sequencing. The section in this report titled "Customer Case Studies Point to Advantages of Complete Genomics' Technology" provides more detail on the company's key customers; as of the end of March 2011, Complete Genomics had more than 45 past and present customers (up from more than 35 as of its IPO in November 2010).

Table 1
Complete Genomics, Inc.
Customer Analysis

	2009	1Q10	2Q10 (cumulative six months)	3Q10	2010
	Percent of Revenue				
Children's Hospital of Boston	**x**	**<10%**	**x**	**14%**	**x**
Pfizer	**16%**	**<10%**	**20%**	**<10%**	**11%**
The Institute for Systems Biology	**13%**	**<10%**	**<10%**	**x**	**<10%**
Sichuan	**<10%**	**x**	**15%**	**x**	**x**
The Broad Institute	**16%**	**<10%**	**<10%**	**x**	**<10%**
University of Missouri Kansas City	x	x	x	13%	x
Erasmus Medical Center	x	x	x	12%	x
Johns Hopkins University	16%	<10%	<10%	x	<10%
The Flanders Institute for Biotechnology	16%	<10%	<10%	x	<10%
University of Texas Southwestern Medical Center	17%	<10%	<10%	x	<10%
The International Mesothelioma Program at Brigham	<10%	30%	11%	x	x
Ontario Institute for Cancer Research	x	30%	x	x	x
Scripps Genomic Medicine	x	30%	x	x	x
Customers	7	5	13	27	40
Total Revenue	$623	$336	$1,425	$4,161	$4,161

Marquee clients have become repeat orders—overall, CGI has seen a 50%-plus reorder rate...

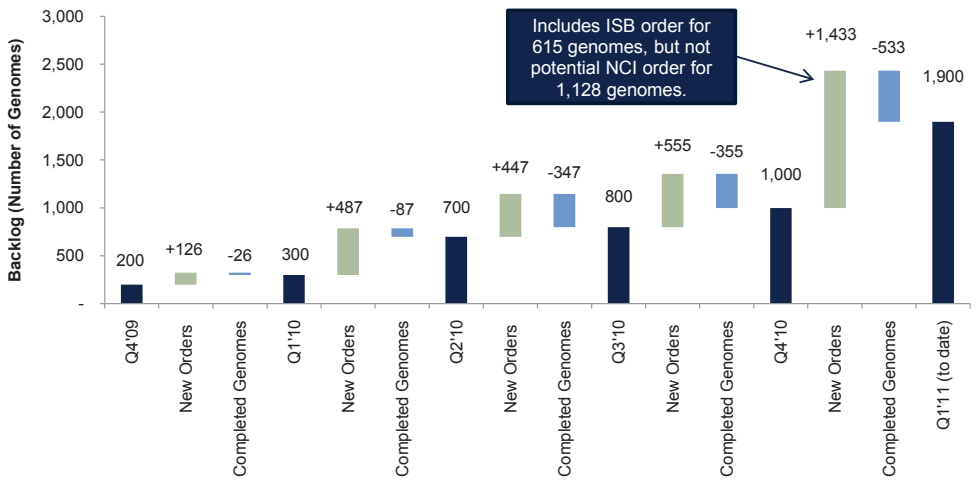
...and a building customer list; as of March 2011, CGI had more than 45 customers.

x - not referenced in filings for that period (assumed not to be a customer in the quarter)
Note: based on fiscal quarters

Sources: Company reports and William Blair & Company, L.L.C. estimates

Complete Genomics has also been consistently building its backlog of whole human genomes. Although the company has sequenced fewer than the initial 5,000 genomes promised in 2010 (it has sequenced an estimated 1,000 genomes as of March 2011), clearly its model is beginning to gain traction. We note that roughly 2,000 whole human genomes were sequenced in 2010, which helps put the magnitude of 1,000 completed genomes in perspective. In addition, the company expects to add eight new instruments this year to meet projected demand in the second half, which provides further support for its ongoing traction. Lastly, on its last earnings call, management indicated that it has had more than 50% repeat orders, which points to a high rate of satisfaction with results.

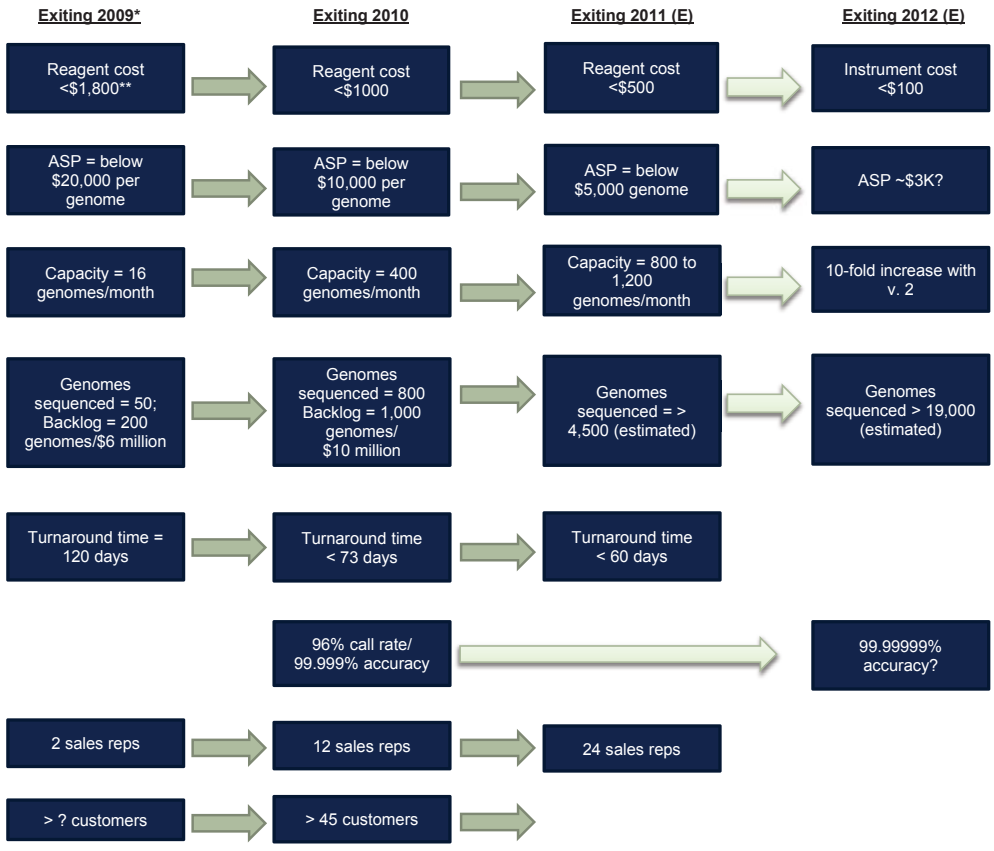
Figure 11
Complete Genomics, Inc.
Historical Backlog Build



Source: William Blair & Company, L.L.C. estimates

As we discuss in more detail in the following section, we believe Complete Genomics' ability to drive its ASP below $5,000 per whole human genome (down from roughly $10,000 currently) is key to its ability to open up what could be meaningful markets for whole-human-genome sequencing (i.e., clinical trials and the whole-exome sequencing market). Thus, we are encouraged to see that the company has continued to make progress, as illustrated in figure 12, since its commercial launch in 2010 from a turnaround time, pricing, cost, capacity, backlog, customer base, and salesforce perspective.

Figure 12
Complete Genomics, Inc.
Business Progress



* Note: Commercial launch was not until May 2010

** As of January 10, 2010, publication in *Science*

E- Willliam Blair estimates

Sources: Company reports and William Blair & Company, L.L.C. estimates

III. Sequencing Is a Large, Growing Market; Human Genome Sequencing Could Benefit From the Wave of Personalized Medicine

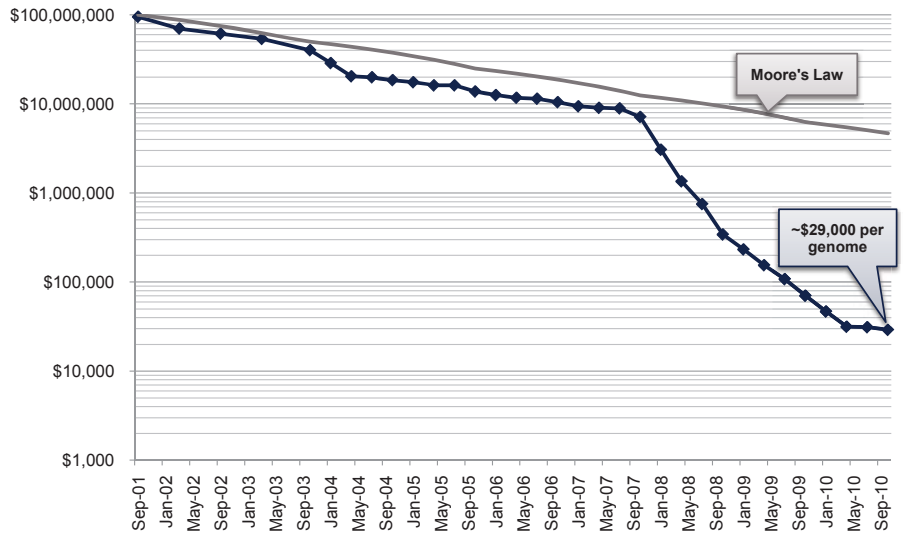
In theory, every person could have his or her genome sequenced. In the case of cancer tumors, whole-genome sequencing (WGS) could be performed multiple times. Basing market size calculations on population and cancer incidence numbers leads to impractical estimates, however.

Still, the sequencing market is highly elastic. Thus, as the ASP continues to fall, sequencing is becoming more viable for more capital-sensitive researchers. To put some context around the opportunity, we believe there are three primary markets that Complete Genomics can penetrate: in the short term, the WGS research market, and in the longer term, drug-discovery/clinical trials and clinical diagnostics.

Whole-genome sequencing research market. Sequencing has evolved to be a critical technology employed in scientific research and in the clinical market to further understand, diagnose, and treat disease. The pace of innovation in the sequencing industry has been astounding as companies and scientists continue to improve sequencing technologies and platforms, which has led to a dramatic increase in throughput and reduction in sequencing cost. At the recent Advances in Genome Biology and Technology (AGBT) conference, for example, some genome centers/large core labs spoke to data generation in the magnitude of terabases per week—which, to provide a reference, is more than the whole world was generating a few years ago. As an another example, according to National Human Genome Research Institute (NHGRI), the cost to sequence one whole human genome has decreased to $29,000 in January 2011 (excluding up-front instrument investment, project management, and IT management costs), as compared with close to $100 million per genome in 2001, representing more than a 10,000-fold cost reduction in just a decade.

The pace of cost reduction has exceeded Moore's Law, which was presented by Intel co-founder Gordon Moore in 1965 to predict the pace of silicon technology. Moore recognized that the number of transistors that can be placed per square inch on integrated circuits had doubled every year since their invention. This trend of increasing performance and declining costs has continued to this day, and the law is now commonly used as an illustration of technology advancement in the sequencing space.

Figure 13
Sequencing Cost Per Genome (September 2001 to October 2010)*



* Cost includes an estimate of production expenses: labor, adminstration, management, utilities, reagent/consumables, amortization, informatics, sample prep. Nonproduction costs (QC, technology improvements to workflows, project management, downstream analysis, informatics equipment, and tools) are excluded.

Source: www.genome.gov/sequencingcosts

Despite the reagent cost improvements, sequencing instruments involve significant capital investment and ongoing operational and maintenance costs (including depreciation expense; human resources, including technicians and bioinformaticians; ongoing reagent/consumable costs; and IT storage and data analysis expenses). Typically, when the cost to sequence a genome is quantified (e.g., the race to the $1,000 genome), estimates only include consumables (e.g., flow cells and reagents) expenses. David Dooling, assistant director of informatics from the Genome Center at Washington University, provided a framework to estimate the all-in cost of sequencing a human genome (with the caveat that every lab is different in terms of salaries paid, overhead allocations, etc.) in the blog *PolITiGenomics*. In sum, once factoring in all the production and overhead costs, what is currently quoted as a $10,000 genome ends up being closer to a $30,000 genome, as illustrated in table 2, on the following page.

Table 2
All-in Cost to Sequence a Genome

Operating Costs	Per Flow Cell
Reagents/Consumables	$10,000
Depreciation	$2,500
Sample Acquisition Cost/Sample	$2,000
Sample Prep (Consumables and FTE)	$200
Total Operating Costs	$14,700
Labor	
Sample Acquisition	$10
IT	$375
LIMS	$375
Analysis	$480
Project Management	$400
Total Labor	$1,640
IT	
Storage	$1,000
Alignment	$100
Variant Analysis	$145
Validation	$2,000
Total IT Resources	$3,245
Total	$19,585
Institutional Overhead	$9,793
Total Cost Per Genome	**$29,378**
Assumptions	
FTE (Salary Per Year)	$60,000
Number of Samples	50
Machine Useful Life	3
Capacity (Runs Per Year)	40
Number of Flow Cells	2
Instrument Cost	$600,000

Source: www.politigenomics.com/2010/06/the-cost-of-doing-sequencing.html, accessed April 6, 2011

One of our key conclusions from the AGBT conference in February was that with the dramatically improved throughput on the newer sequencing platforms, labs are generating a lot of sequencing data. The community appears to have vastly underestimated the need for bioinformatics support, and a lack of standard data analysis tools, data analysis, and storage have become big issues for users (and can cost at least as much as reagents). In addition, it is our understanding that to qualify for NIH funding, projects have to present a viable data management solution. To be fair, some of the costs included above (such as data storage) would be still be incurred in the Complete Genomics model; however, in our view, the data analysis bottleneck, as well as high capital outlays, may ultimately drive researchers to a more full-service model.

Given that the ASP remains above $5,000, we believe that in the shorter term, research-based whole-genome sequencing is likely to remain the most meaningful market for Complete Genomics. Quantifying exactly how many genomes might be completed over the next five years is difficult given the elasticity of the demand curve. According to life sciences consulting firm Scientia Advisors, the sequencing market is roughly $1.2 billion and is expected to grow at a compound annual rate of 20% to 25% through 2014. Assuming the entire research market is whole human genomes (which it clearly is not), at an average cost per genome of $160,000 throughout 2009, the market would imply 4,000 human genomes in 2011 growing to 160,000 genomes in 2014 (assuming a cost of $10,000 per genome). The article “Genomes by the Thousand,” published in *Nature* in October 2010, suggested that roughly 30,000 genomes should be sequenced by the end of 2011.

With the introduction of Illumina's HiSeq in early 2010, next-generation instrumentation broke through a new throughput threshold, which drove a meaningful acceleration in the number of WGS cases completed last year. Still, WGS is predominantly performed at the large genome centers. Many genome centers and core labs are also leveraging whole-exome sequencing, whereby just the exomes (representing about 1% of the genome and the protein coding regions) are sequenced. The benefit of this approach is that it captures the region of the genome where most disease-causing mutations are found, but represents roughly one-seventh the cost of whole-genome sequencing.

Thus, one other way to gauge the size of the nearer-term opportunity is to assume that at some cost point, researchers currently performing exome sequencing would switch to WGS. Based on commentary at the recent AGBT meeting, the Broad Institute of MIT and Harvard plans to sequence roughly 5,000 exomes this year. Extrapolating this to the broader market (using the Broad market share of next-generation sequencing instrumentation and assuming most exome sequencing is still done at genome centers) would suggest roughly 49,400 exomes are sequenced globally on an annual basis. Interestingly, also at AGBT, Dr. Deborah Nickerson of Genome Sciences at the University of Washington School of Medicine predicted 10,000 exomes would be sequenced in 2011, although during our recent channel checks, we have heard estimates as high as 75,000 this year.

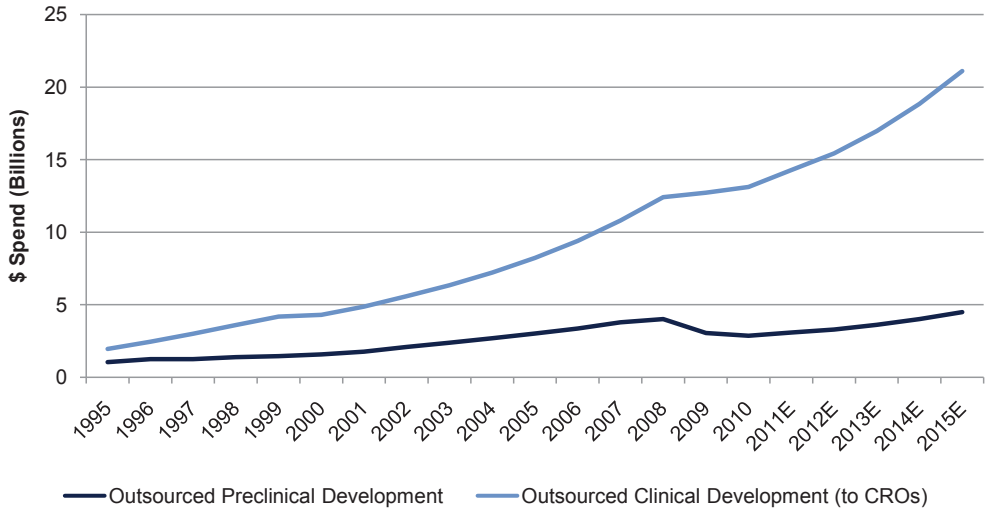
Pharma outsourcing/cancer clinical trials. In our view, once Complete Genomics is able to drive its ASP below $5,000, its sequencing services could become appealing to pharma and biotech companies looking to leverage sequencing in drug discovery and later in clinical trials while reducing capital outlays.

Accuracy is important in analysis of cancer tumors given the need to detect minor variants in heterogeneous samples and ensure fewer false positives (helping manage project cost by reducing the need for validation). Thus, based on: 1) the company's high accuracy rate (and potential for an even lower error rate in the future), 2) its recent addition of copy number and structural variant analysis (important in cancer research), and 3) its cancer assembly software (with a commercial version expected in mid-2011), we believe Complete Genomics' solution could be very attractive to pharma and biotech as sequencing becomes a larger component of clinical trials. Complete Genomics' model is beginning to gain traction in cancer; for example, in September 2010, the company received an order to sequence 50 pediatric cancer normal/tumor pairs, which, if successful, could be extended to a total of 500 normal/tumor samples.

As illustrated in figure 14, on the following page, pharma continues to increasingly outsource research and development; although this has particularly been the case in the clinical side of the market, we also expect outsourcing of preclinical research to reaccelerate over the next five years. Anecdotally, it appears pharma and biotech are more likely to outsource sequencing projects, although thus far these projects have primarily been on the research versus the clinical side. We know, for example, that Pfizer, Genentech, and Eli Lilly have used Complete Genomics' services; Merck has an agreement with BGI to collaborate on biomarker discovery and validation; and Illumina has suggested that its whole-human-genome service backlog includes a large, multi-hundred genome order from pharma.

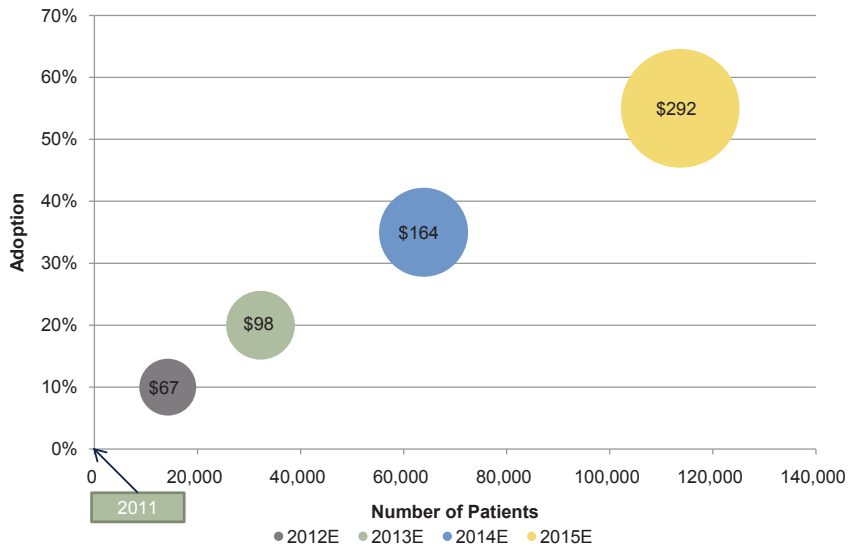
In terms of the opportunity for Complete Genomics, there were roughly 760 cancer-related clinical trials conducted in 2010, which is projected to grow at a compound annual rate of 10% over the next five years. As illustrated in figure 15, also on the following page, whole-human-genome sequencing for oncology clinical trials could be a significant opportunity for the company, totaling around $300 million by 2015 (assuming by that point whole-human-genome sequencing is used for roughly 55% of patients enrolled).

Figure 14
CRO Market Model



Note: Estimates are those of William Blair analyst John Kreger
Sources: Company reports, EFPIA, JPMA, PhRMA, Pharmaprojects, and William Blair & Company, L.L.C. estimates

Figure 15
Potential Cancer Clinical Trial Market Size
(bubble size indicates market size in millions)



Sources: clinicaltrials.gov and William Blair & Company, L.L.C. estimates

Clinical diagnostics. We believe the ultimate opportunity is to transition sequencing into the diagnostic realm. The wide-scale incorporation of whole-genome sequencing into the clinical realm is 10 or more years away, in our view, given a number of existing hurdles: the need for clinical relevance, continued cost reductions, physician addition, payer reimbursement, and FDA regulatory uncertainty. As more cancer genomes are sequenced, however, sequencing of tumors either to identify appropriate therapeutics or as a monitoring mechanism could become a clinical reality in the next three to five years, in our view.

In addition to traditional diagnosis methods (e.g., morphology and staging), oncologists are increasingly relying on the molecular profiling of tumors. For example, in colorectal cancer, screening for mutations in KRAS, BRAF, and PI3KCA genes can identify up to 70% of tumors unlikely to respond to anti-EGFR therapies (a common treatment strategy). From a clinical application standpoint, the goal is to enhance molecular profiling of tumors to stratify patients who might benefit from specific therapeutics, and ultimately, to identify novel drug targets.

Washington University published the first cancer genome (of acute myelogenous leukemia) in November 2008 and has since sequenced a second AML sample. Thus far, according to the COSMIC (Catalogue of Somatic Mutations in Cancer) database, 385 full cancer genome sequences have been published (up from almost 30 in early 2011), with thousands of others completed or in progress. A number of projects are under way to facilitate sequencing of cancer genomes or tumor/normal pairs, such as the International Cancer Genome Consortium, the Cancer Genome Project, and the Cancer Genome Atlas, as well as Illumina's own efforts to analyze 25 tumor/normal pairs for ovarian cancer and 25 tumor/normal pairs for gastric cancer.

Through exome and full-genome sequencing initiatives, scientists have begun to catalog both germline and somatic single nucleotide polymorphism (SNP), copy number, and structural mutations within tumors as well as to identify the pathways affected. To date, many tumors have been found to be genetically diverse across similar phenotypes, highly rearranged, and rapidly mutating, while others have a similar mutational profile (AML and glioblastoma, for example). Thus, it is not clear whether specific cancers will be found to have consistent, identifiable mutations or whether each tumor will need to be profiled on an individual basis (to track rearrangements over time). For example, researchers were able to track chromosomal rearrangements in patients with colorectal cancer, which enabled the treatment's effectiveness to be monitored (e.g., biomarker levels dropped after the tumor was removed).

Again, Complete Genomics' advantage is accuracy; the ability to deliver 99.99999% accuracy in 2012 when sequencing a human genome would be unprecedented and ultimately necessary for use in the clinic. In addition, as discussed in more detail in the "Accuracy" section of the first investment highlight, with its long read technology, Complete Genomics will be able to sequence both maternal and paternal chromosomes. Essentially, this would allow clinicians to determine whether the mutation were present in only one copy of the gene (from either parent) or both, and would provide another advantage in the diagnostic realm.

If sequencing technology is incorporated into cancer diagnosis, therapeutic decision-making, and monitoring, markets would be sizable and defined by annual cancer incidence rates (1.5 million people are diagnosed with cancer in the United States annually) and could extend to prevalence estimates if tests are obtained on an ongoing basis.

Table 3
Cancer Incidence Rates

Cancer Type	Incidence Rate	No. of New Cases[1]	Prevalence[2]
Bone	0.9	2,763	NA
Breast	122.9	188,656	2,591,855
Cervix Uteri	8.1	12,434	247,180
Colon and Rectum	47.9	147,056	1,112,493
Kidney and Renal Pelvis	14.1	43,288	281,490
Acute Lymphocytic Leukemia	1.6	4,912	60,783
Liver and Intrahepatic	6.9	21,183	27,753
Lung and Bronchus	62.5	191,879	370,617
Ovary	12.9	19,802	177,162
Pancreas	11.7	35,920	32,993
Prostate	156.9	240,847	2,276,112
Stomach	7.8	23,947	65,639

[1] New case calculation is based on U.S. Census Bureau data of the total U.S. population at 307 million as of July 2009
[2] Prevalence (total number of people living with that cancer) estimates are based on SEER estimates as of January 2007

Sources: SEER, U.S. Census Bureau, and William Blair & Company, L.L.C. estimates

In addition, although Complete Genomics appears committed to human DNA sequencing, management has suggested that transcriptome and methylome analysis are within the company's strategic footprint, which could further expand Complete Genomics' research market opportunity.

Although as mentioned it is difficult to precisely size the addressable market, we attempted to use several available data points we have garnered from both industry conferences and our recent channel checks. Admittedly the capacity and the nature of work done at large genome centers are vastly different from smaller labs; thus, we have assumed labs that have greater or equal to six Illumina sequencers are most likely to sequence whole genomes and whole exomes. In addition, we have assumed these genome centers will convert all the existing Genome Analyzers to HiSeqs, and they either have started using the latest Illumina TruSeq reagent kits (which increase sequencing capacity from 200 GB to 600 GB), or will have the kit by sometime this year. With these assumptions, we arrived at the estimate of around 35,200 whole genomes, and around 49,400 exomes to be sequenced this year, or a market of $459 million in 2011.

For the following years, we have assumed continued growth in sequencing capacity and sequencing cost decline. We have also incorporated the oncology clinical trial market beginning in 2013. In addition, we believe a portion of the whole-exome sequencing market will convert to whole-genome sequencing when the price of WGS drops to certain point. Based on these assumptions, we arrived at the estimate of around 201,000 whole genomes, 200,800 whole exomes converted to whole genome sequencing, and 206,700 clinical trial participants, which translates to a market of around $1.8 billion in 2015.

Please see table 4 for more detail.

Table 4
Complete Genomics, Inc.
Addressable Market

	2011E	2012E	2013E	2014E	2015E	
Whole-Exome Market						
Qualified Illumina HiSeqs	457	594	772	1,004	1,305	
Exome sequenced per HiSeq per year	2,700	2,700	2,700	2,700	2,700	
Capacity allocated to WES[1]	4%	20%	30%	20%	10%	Assume gradual decline due to conversion to WGS
Exomes per year	49,356	320,814	625,587	542,176	334,793	
Sequencing Cost	$1,667	$1,250	$938	$750	$600	
Implied Market ($ million)	$82	$401	$586	$407	$201	
Base Research Market						
Qualified Illumina HiSeqs	457	594	772	1,004	1,305	
Capacity allocated to WGS[2]	35%	45%	65%	65%	70%	
Genome sequenced per HiSeq per year	220	220	220	220	220	Assumes no change in HiSeq throughput
Genomes per year (research)	35,189	58,816	110,443	143,576	201,007	
Exomes converted to WGS	2,468	64,163	125,117	189,761	200,876	
Total whole genomes	37,657	122,979	235,561	333,338	401,883	
Sequencing Cost	$10,000	$7,500	$5,625	$4,219	$3,164	
Implied Market ($ million)	$377	$922	$1,325	$1,406	$1,272	
Clinical Trial Market						
Patient Enrolled	125,908	142,418	161,167	182,468	206,676	
Number Using WGS	0%	0%	10%	35%	55%	
Blair estimated ASP	$7,387	$4,703	$3,049	$2,573	$2,573	
Implied Market ($ million)	$0	$0	$49	$164	$292	
Total Market ($ million)	**$459**	**$1,323**	**$1,961**	**$1,977**	**$1,765**	
Total Addressable Genomes	**37,657**	**122,979**	**251,677**	**397,201**	**515,554**	
Growth		*227%*	*105%*	*58%*	*30%*	CGI accounts for 22% of the market by 2015
Complete Genomics Penetration	**12%**	**16%**	**18%**	**19%**	**22%**	

[1]WES: Whole-exome sequencing
[2]WGS: Whole-genome sequencing

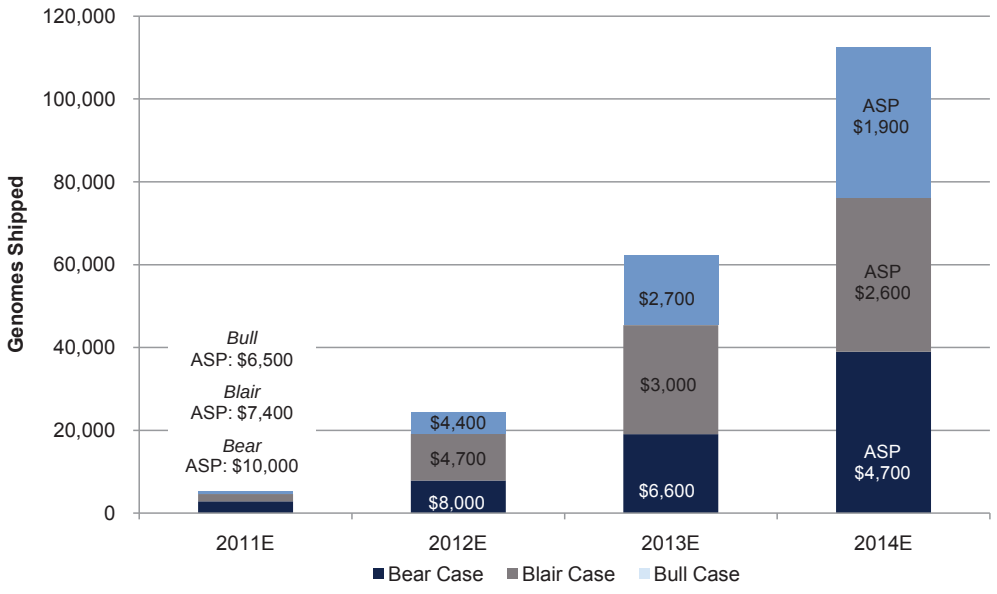
Source: William Blair & Company, L.L.C. estimates

Investment Risks

I. Ability to Drive Down Cost Is Key to Company Longevity and Ability to Drive Enough Long-Term Demand to Support Its Valuation

As we have discussed, we believe one of the key drivers of adoption for Complete Genomics after 2011 is its ability to continue to drive down ASP to open up potentially large market opportunities. Therefore, there is a risk that CGI will not be able to lower its ASP quickly enough to garner sample flow beyond what we would consider to be the base research market. Therefore, figure 16 provides a sensitivity analysis of our adoption assumptions, incorporating various scenarios for ASP reductions.

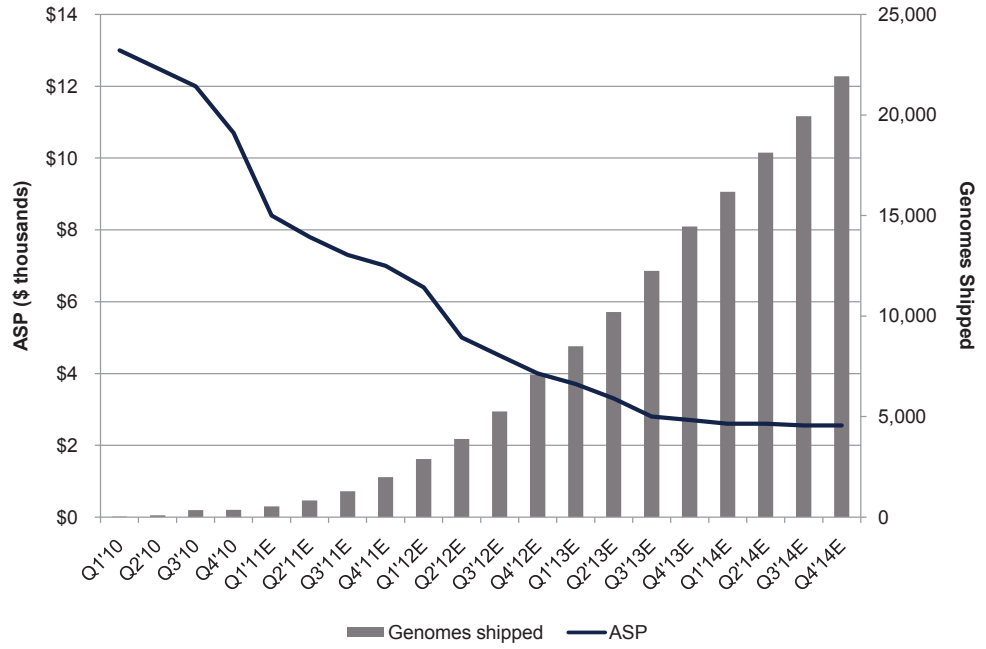
Figure 16
Complete Genomics, Inc.
Estimated Genomes Shipped Based on Different Scenarios



Source: William Blair & Company, L.L.C. estimates

Complete Genomics should be able to drive margin leverage from incremental scale and may be able to differentially price for premium service. Still, as illustrated in figure 17, on the following page, by nature of its model, it will be fighting a declining ASP to a price point at which the market becomes relatively inelastic. Therefore, if CGI is not able to drive enough volume to offset the ASP reductions we have modeled, it will have a negative impact on our assumptions for the company's cash flow burn rate and cash flow breakeven target of early 2013, as well as our P&L breakeven target of 2013.

Figure 17
Complete Genomics, Inc.
Estimated ASP Versus Genomes Shipped



Sources: Company reports and William Blair, L.L.C. estimates

II. A Number of Competing Service Providers Could Erode Complete Genomics' Target Market

Given the rapid growth in the sequencing market, a number of sequencing providers have emerged that offer full-service next-generation sequencing outsourcing services (including sample prep and downstream data analytics). In addition to Complete Genomics, core labs associated with academic centers, genome centers, and commercial operations offer sequencing services at a wide range of service levels (full service to just sequencing with no downstream analytics). Examples include the Beijing Genomics Institute, which purchased 128 Illumina HiSeqs in early 2010 (now with a total of 137 HiSeqs) and 27 Life Technologies SOLiD machines in late 2010; EdgeBio, which operates SOLiD platforms; Cofactor Genomics, which operates SOLiD 4s and GAIIxs; and PerkinElmer.

While Complete Genomics' solitary focus on whole-human-genome sequencing provides cost advantages, other competing service providers—in addition to the various platform vendors who are also seeking share of global whole-genome samples—represent potential competitive threats and continue to attract market share.

Other service providers have avoided ties to specific platforms and focused on optimizing workflow (e.g., sample prep and data analytics), which has benefits for researchers that may be more comfortable with existing platforms. For example, Illumina's SBS chemistry has been cited in almost 1,500 publications and has well-established software and bioinformatic support. In addition to selling instrument platforms, Illumina also recently launched whole-human-genome sequencing services (via certified service providers), which are currently on par with Complete Genomics from a pricing standpoint. BGI in China (which is by far the largest genome center in the world) also offers whole-human-genome sequencing (among many other services) based on Illumina and Life Technologies platforms. In addition, many large genome centers offer sequencing services, predominantly on Illumina platforms, that can be purchased by outside customers.

As we have discussed, the ability to continue driving down ASP is key to unlocking potential large market opportunities for whole-human-genome sequencing. Thus, while we believe the market for whole-human-genome sequencing is large enough for many players to succeed, any one of these providers in addition to Complete Genomics could successfully "break the cost barrier," providing a first-mover advantage.

Lastly, there is great research interest in the ability to capture and sequence rare or single cells circulating in the body (e.g., tumor cells and fetal cells in maternal blood), because the single-cell method requires less starting sample and provides high sensitivity. A number of players are working to develop and improve this technology. Among these is Illumina, which is developing its proprietary rare-cell capture technology and is in beta testing with four customers; it plans on adding this capability to its sequencing services for analyzing fetal and circulating tumor cells. While we expect ultimately there will be commercial kits available for single-cell sample prep, the first player to successfully offer the ability to sequence genomes from a single cell could also garner a first-mover advantage from a know-how standpoint.

Key service provider competitors

BGI (formerly known as Beijing Genomics Institute). With a total of 137 HiSeqs from Illumina and 27 SOLiD sequencers in late 2010, BGI has become the largest genome center in the world, with the capacity to generate 5 terabases of data per day—or more than 50 genomes at 30-times coverage, implying capacity of roughly 12,000 genomes per year. BGI has its hand in a number of projects; however, many of them do not focus on whole-genome sequencing. Thus, its actual human genome output is likely to be somewhat lower.

BGI was founded in 1999 as Beijing Genomics Institute and has sequencing labs in Shenzhen and Hong Kong (which houses the majority of its HiSeqs). BGI is funded by both the Chinese government and private investments and has presence in Asia, America, and Europe. Since its inception, BGI has been involved in a number of major sequencing projects, including the Human Genome Project; the HapMap Project; the sequencing of the SARS virus; the full sequence of the rice, silkworm, and giant panda genomes, which BGI completed in eight months; and the Earth Microbiome Project, among others. It can also boast more than 179 publications at last count.

With roughly 180 sequencers in total, BGI is positioning itself to be the service provider for sequencing projects globally. The organization offers a collaborative model (BGI funds all or a component of the project, and therefore, shares in costs, authorship, and any associated IP) or a fee-for-service model where data can be delivered based on client specs similar to CGI. For example, Merck and BGI recently announced they are working together to leverage sequencing in pharmaceutical development. Their agreement includes "master agreements" and joint projects; both plan on building infrastructure to collaborate in biomarker identification, validation, and target development. As another example, BGI will be sequencing a single human genome on a fee-for-service basis for the National Institute of Diabetes and Digestive and Kidney Diseases at 90-fold coverage on the Illumina platform. The deliverable will include consensus sequence assembly, as well as variant detection and comparison with ethnic genomics with a two-and-a-half-month turnaround and data delivery via either BGI's FTP site or a hard disk sent through postal mail.

BGI does not focus only on human genomes, but is also sequencing a number of other organisms, including bacteria, plants, microbes, and complex diseases, which could give Complete Genomics an advantage given its sole focus on human genomes. There are some questions as to whether researchers will want to share authorship and IP rights with BGI, and ultimately, the market is likely to be large enough for a number of players. Still, we view BGI as a serious contender and potential competitive threat to Complete Genomics.

Illumina Genome Network. In early 2010, Illumina announced the creation of its Human Genome Network, which is the company's version of whole-human-genome sequencing services. In an effort to avoid competing with its customers, Illumina subcontracts with providers of marketing and sales services as well as load-balancing services across its network (and in some cases, network partners are providing excess capacity if needed). For the service partners, it provides another marketing channel without incremental costs and allows Illumina to market each provider based on their specific strengths. In addition to whole-human-genome sequencing services, Illumina also offers whole-exome sequencing, RNA sequencing, and methylation sequencing across multiple organisms.

Although we estimate Illumina's service offering does not account for a large portion of revenue,during its fourth-quarter call in February, Illumina indicated it had 1,000 genomes in its backlog, which includes one order for several hundred genomes from a major pharmaceutical company.

Through the Genome Network, Illumina provides data analysis services similar to Complete Genomics', including genome assembly and variant calls (SNP, indels [insertions and deletions], copy-number variations, and other structural arrangements). Illumina's reagent cost is now $5,000 for whole-genome sequencing, which is above Complete Genomics at less $1,000, and its ASP is roughly $10,000 per genome (or less for high-volume orders)—about in line with Complete Genomics. Currently, however, Illumina only has two partners listed on its website: Macrogen and the National Center for Genome Resources (NCGR). Illumina is a powerhouse in the sequencing world, and thus we believe Illumina's whole-genome service offering should not be discounted as a potential threat.

PerkinElmer. In early 2011, PerkinElmer announced that it launched a sequencing and data analysis service. The company is using Illumina's HiSeq 2000 and intends to focus specifically on exome sequencing, delivering data to clients via the cloud. As of the announcement, the company had 20 customers and had delivered hundreds of exomes. PerkinElmer has not disclosed its ASP.

III. In Early Phases of Commercial Launch, Thus Significant Execution Risk Exists
As an early-stage company, Complete Genomics carries significant execution risk, including: 1) high customer consolidation, 2) potential service disruptions as the company scales production, and 3) dependence on improving margins to reach profitability.

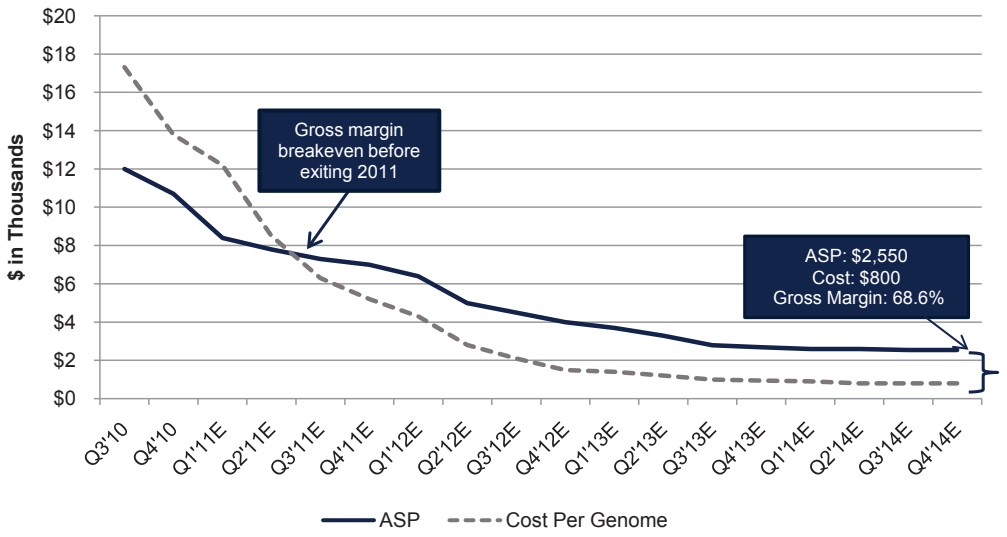
***Customer consolidation*.** Complete Genomics has continued to diversify its customer base (with 45 customers as of the middle of March 2011); however, the company's revenue is still generated by relatively few customers.

***Potential service disruptions*.** In addition, the company depends on some single-source suppliers (such as that for the wafers that comprise the base of its arrays) and has customized some parts for its consumables and instrumentation. Vendor supply issues could limit its ability to service sequencing demand, add machines as necessary, or introduce the second version of its platform. The addition of eight new instruments this year is critical for the company to be able to meet whole-genome sequencing demand in the back half of 2011. Successful commercial launch of version 2 of its instrument is important for the company to be able to further break the cost barrier to penetrate some of the larger, more cost-sensitive areas of the market (e.g., pharma). As it is a service provider, the company's reputation is critical, and any dramatic increase in turnaround time could cause a dampening of demand. In addition, use of single-source suppliers potentially exposes the company to underlying cost risk.

***Margin improvements*.** Given the high fixed-cost nature of this business, the ability for the company to drive efficiency improvements on its hardware and software (e.g., successfully rolling out higher-density arrays) is key to driving its gross margin into positive territory. We

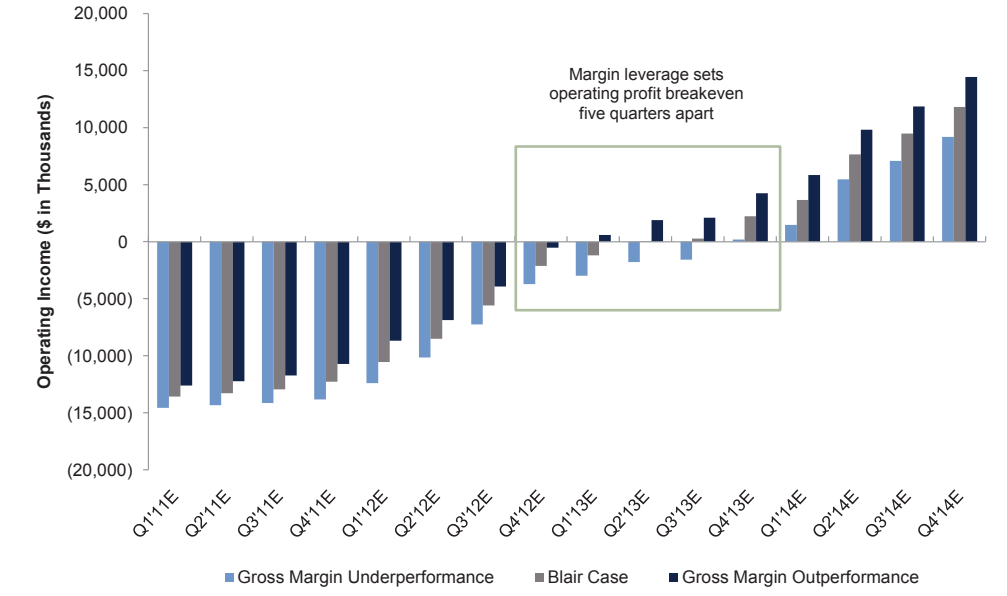
estimate the company will begin to make money (gross profit) on a per-genome basis at the end of 2011. If this is delayed, it would increase the company's cash burn rate and delay our cash-flow-positive target of 2013.

Figure 18
Complete Genomics, Inc.
Margin Leverage



Sources: Company reports and William Blair, L.L.C. estimates

Figure 19
Complete Genomics, Inc.
Cost Per Genome Versus Operating Profitability



Source: William Blair & Company, L.L.C. estimates

Company Overview

Brief History

Complete Genomics was incorporated in Delaware in June 2005 and began operations in March 2006. The company is based in Mountain View, California, and has 185 employees as of the end of 2010. Among its employees, 56 hold doctoral degrees and 71 are involved in full-time research and development.

Complete Genomics was co-founded by Chief Executive Officer Clifford Reid, Chief Scientific Officer Radoje Drmanac, and Vice President of Financial Operations Robert J. Curson, and is a spin-off of Callida Genomics. Callida Genomics, co-founded by Dr. Drmanac, was owned by biotech companies Hyseq (renamed Nuvelo after acquiring Variagenics in 2003) and Affymetrix. Callida Genomics focused on developing sequencing-by-hybridization technology (an early prototype of cPAL technology, which Complete Genomics adopted and further refined) and DNA sequencing on microarrays.

Complete Genomics recognized its first revenue in the fourth quarter of 2009. As of early March 2011, it had more than 45 customers and shipped more than 1,000 whole human genomes. The company currently operates its second-generation sequencing instruments, which produce 150 gigabases per datum, per instrument (or a genome per day); it plans to commercialize the second generation of instruments in the near future, which should run 10 times as fast as its previous iteration. Complete Genomics has around 16 instruments installed in its facility and expects to add another 8 by the end of the year.

Management Team

The company's management team has extensive experience in genomics and sequencing, biopharmaceuticals, and software development. With the leadership of CEO Clifford Reid, the company has successfully launched a commercial-sequencing service operation that was responsible for an impressive 1,000 of the roughly 2,000 whole human genomes sequenced in the world over the past year.

Table 5
Complete Genomics
Management Profiles

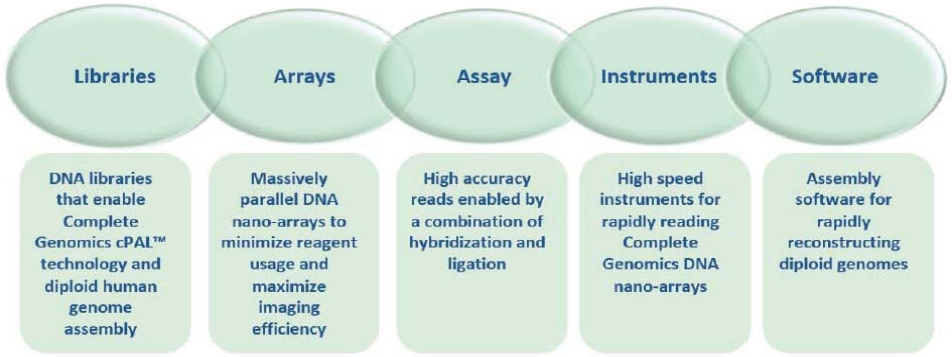
Name	Title	Tenure	Experience
Clifford A. Reid	Chairman, President, and Chief Executive Officer	6	Clifford Reid is one of the co-founders of Complete Genomics and has served as member of the board, president, and chief executive officer since July 2005. Prior to Complete Genomics, Dr. Reid was vice president of collaborative solutions at OpenText Corporation, a software company that targets mainly enterprise customers on enterprise content management. Dr. Reid co-founded and worked at Eloquent, Inc., a digital video communications company, which was acquired by OpenText Corporation in 1995, and Verity, Inc., an enterprise text search engine company, in 1988. Dr. Reid has a B.S. in physics from Massachusetts Institute of Technology, an M.B.A. from Harvard University, and a Ph.D. in management science and engineering from Stanford University.
Ajay Bansal	Chief Financial Officer	1	Mr. Bansal has served in his current role since May 2010 and brings to Complete Genomics decades of experience in finance and the biopharmaceutical industry. Prior to Complete Genomics, Mr. Bansal worked at several senior roles in biopharmaceutical companies, such as Tercica, Inc., Nektar Therapeutics, Novartis, and financial companies, such as Capital One Financial and Mehta Partners, L.L.C. Mr. Bansal also has experience in the consulting industry, having worked at companies such as McKinsey & Company and ZS Associates. Mr. Bansal has a B.S. in mechanical engineering from the Indian Institute of Technology (Delhi), as well as an M.S. in operations management and an M.B.A. from Northwestern University.
Radoje Drmanac	Chief Scientific Officer	6	Dr. Drmanac is one of the founders of Complete Genomics. Before that, Mr. Drmanac co-founded Callida Genomics in 2001 and Hyseq (later renamed Nuvelo) in 1994. Callida Genomics developed and commercialized sequencing-by-hybridization technology and DNA sequencing on a chip, whereas Hyseq, the majority owner of Callida Genomics, developed DNA array technology. Prior to these business ventures, Dr. Drmanac served as a group leader at Argonne National Laboratory. Dr. Drmanac has a B.S., M.S., and Ph.D. in molecular biology from the University of Belgrade.
Bruce Martin	Senior Vice President of Product Development	6	Mr. Martin has worked in senior roles at Complete Genomics since 2005. Prior to Complete Genomics, he was vice president of product strategy at PSS Systems, a software company specializing in corporate information governance; it was acquired by IBM. Prior to PSS Systems, Mr. Martin served as chief technical officer of Openwave System, a software company that specializes in mobile data management. Mr. Martin has a B.S. in computer science and electrical engineering from the University of California, Davis.
Mark J. Sutherland	Senior Vice President of Business Development	1	Mr. Sutherland joined Complete Genomics in March 2010. He has more than 20 years' experience in the genomics field. Prior to Complete Genomics, he was senior vice president of business development at GenVault Corporation, a biotechnology company specialized in providing biological sample storage; it was acquired by IntegenX in February 2011. Mr. Sutherland had been with Molecular Dynamics and its successor companies (Amersham Biosciences and GE Healthcare) since 1998; there, he served several senior roles, such as vice president of genomics at Amersham and vice president of discovery systems at GE Healthcare. Mr. Sutherland has a B.S. in chemistry with honors from Stanford University.

Sources: Company reports and Thomson One

Platform Overview

The whole-genome sequencing service at Complete Genomics encompasses five steps, starting with library construction and ending with downstream data analysis and report generation (which contains variant information and an explanation of its medical significance). The following section provides more background on Complete Genomics' cPAL and DNA nanoball sequencing platform, as outlined in the company's technology white paper. See http://www.completegenomics.com/knowledge-center/whitepapers for more information.

Figure 20
Complete Genomics, Inc.
Overview of Complete Genomics' Sequencing Technology

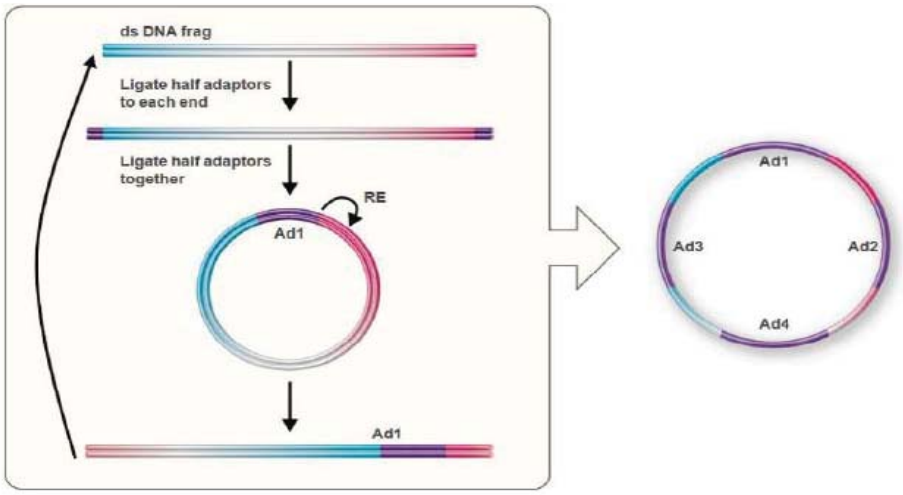


Source: Company reports

Library Construction

The customized workflow at Complete Genomics begins with library construction. A library is essentially randomly generated DNA fragments representing the entire genome of an individual such that it can be manipulated experimentally.

Figure 21
Complete Genomics, Inc.
Library Construction



Source: Company reports

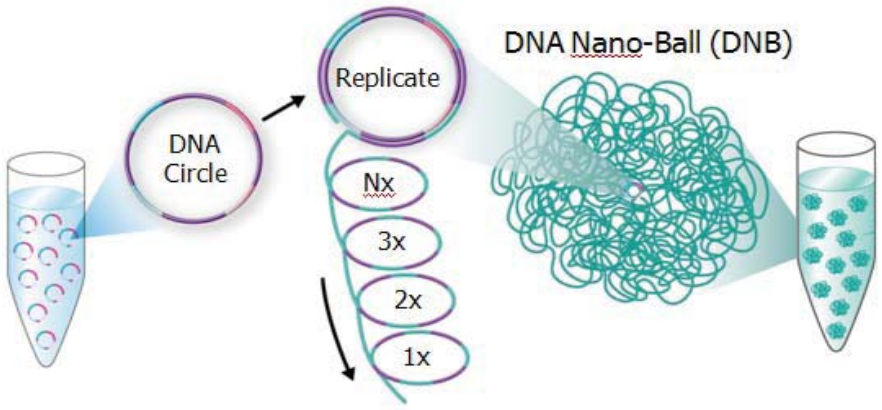
To create the library, genomic DNA is first broken down into smaller fragments (roughly 500 bases per fragment). Complete Genomics then uses four known DNA sequences (called adaptors) that are inserted into each fragment. As illustrated in figure 21, half of the first

adaptor is attached to one end of the denatured double-strand DNA fragment, and the other half is attached to the other end. The complementary adaptor ends then hybridize to form a closed loop. The loop is broken down again at a different site, and the process repeats with a second adaptor until all four adaptors have been inserted into the DNA fragment. The adaptors serve as a starting point for reading the DNA sequence later in the process.

Amplification

Before sequencing, the DNA sample is amplified to improve signal-to-background ratio and increase throughput. As compared with other amplification technologies used in second-generation sequencing, Complete Genomics conducts the amplification process in solution and in a single chamber (versus on a substrate or in an emulsion), thus reducing reagent usage and creating a higher density of clonal amplicons. The process starts with the 80 bases of genomic DNA and the four adaptors in circular form described above, which is then replicated to create more than 200 copies of each circular template. This amplified synthetic DNA is then condensed into DNA nanoballs (DNBs) via Complete Genomics' proprietary method.

Figure 22
Complete Genomics, Inc.
DNA Nanoball

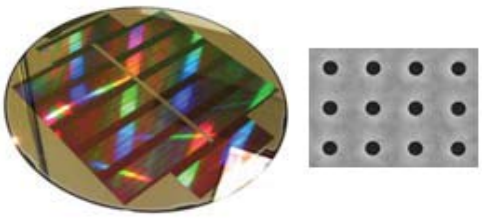


Source: Company reports

Array

The sequencing reaction is performed on two-dimensional one-inch by three-inch arrays. The array manufacturing process leverages photolithography using standard silicon processing techniques to create (currently) 3 billion "sticky" spots on each array. DNBs are attracted to the spots and do not stick to the space in between the spots. When the DNBs described above are introduced to the array, they self-assemble such that each spot contains one DNB (which includes 180 billion bases of DNA). Once a DNB sticks to a spot, it repels other DNBs. Thus, the company is able to ensure that over 90% of the spots are occupied by only one DNB; occupancy (where one DNA fragment can be found in one reaction chamber) has been a difficult problem for other vendors to solve and is one of the key reasons Complete Genomics has been able to drive down cost via increased throughput. The three-dimensional nature of the DNB further reduces the quantity of reagent required and results in brighter spots (and more efficient imaging).

Figure 23
Complete Genomics, Inc.
Silicon Substrate



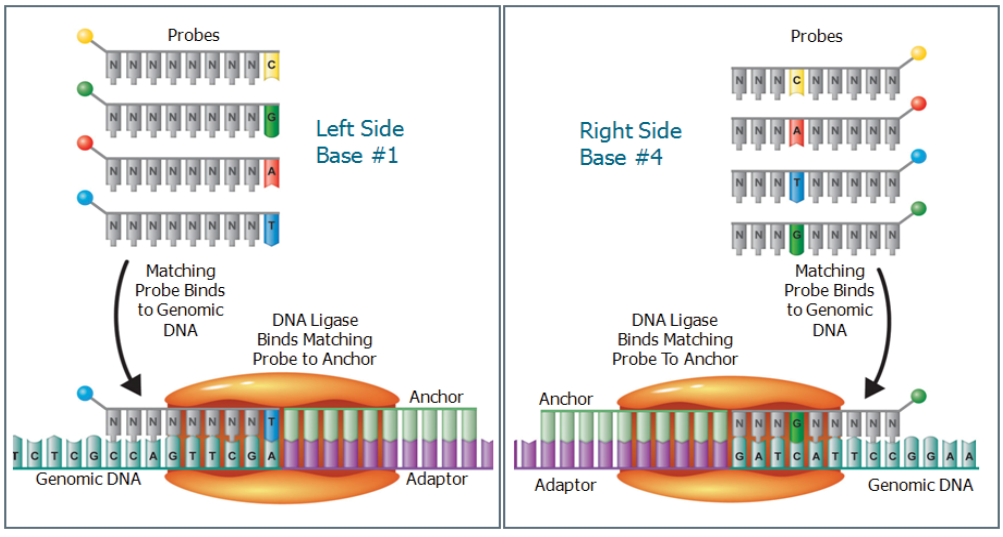
~3 Billion Spots per 3" × 1" Silicon Chip

Source: Company reports

Sequencing Reaction

Sequencing at Complete Genomics uses what is called combinatorial probe-anchor ligation (cPAL) technology. To read the DNA sequences between the four adaptors, pools of probes are added to the mix, which are labeled with four distinct dyes (one for each base). Each read position has its own set of probes (for example, the first base adjacent to the adapter has four specific probes). Depending on which probe binds to the sample, the system can detect the base because of the fluorescent signal. The probe/ligase complex is washed away, and four probes are introduced for the next base position (e.g., two bases away from the adaptor). The process is repeated until 10 bases are read adjacent to each adapter.

Figure 24
Complete Genomics, Inc.
cPAL Sequencing



Source: Company reports

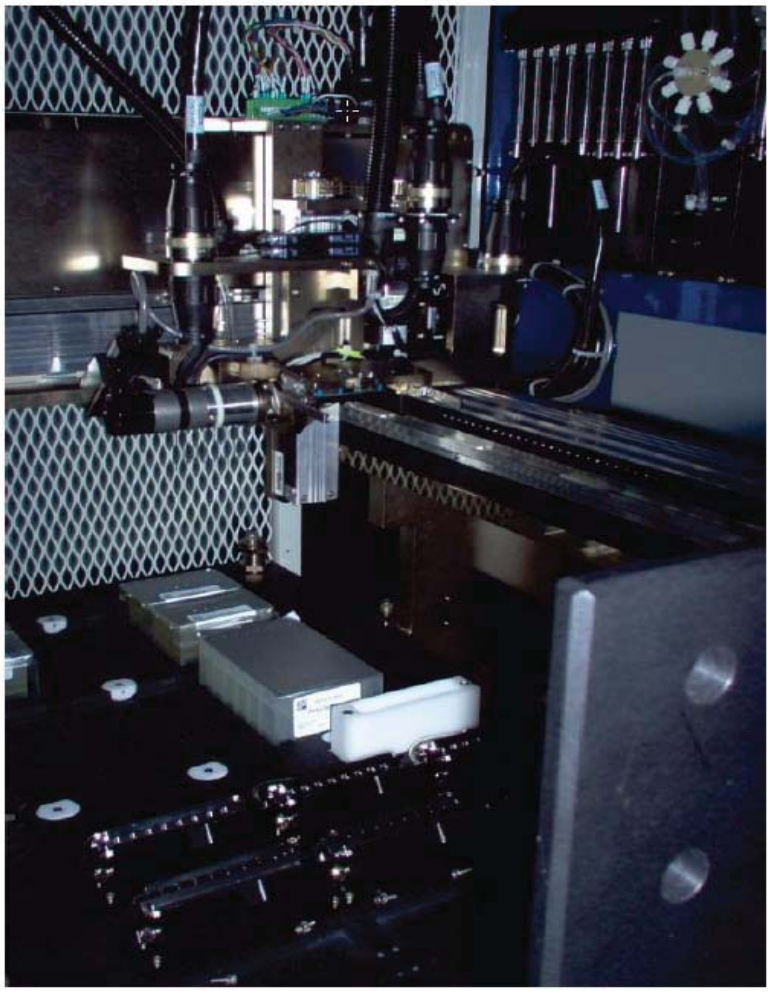
The ligase is highly specific, thus the raw read error rate is low (it can be less than 0.1%). The reads are independent from each other and not dependent on the read completed prior, as is the case with sequencing by synthesis—meaning that each base does not have to be

read after the one before it. Thus, errors do not accumulate. In addition, the independent nature of the reads results in tolerance for low ligation yield, reducing probe and ligase concentrations needed to complete the reaction (and therefore reagent usage). Lastly, multiple base positions can be read during the same cycle, which is not possible with sequencing by synthesis. This reduces the number of cycles needed (and further reduces reagent usage and imaging requirements).

Instruments

Because Complete Genomics only operates its instruments in-house, the company is able to manage its instruments in a relatively open platform, which facilitates ease of software/hardware upgrades. The system is modular, and thus components can be independently upgraded. In general, each sequencing instrument contains a DNA nanoarray (silicon flow slide), standard liquid-handling robot, and high-speed imager.

Figure 25
Complete Genomics, Inc.
Sequencing Instrument



Source: Company reports

Within the module, the DNA nanoarray (silicon flow slide) is optimized for minimal reagent use and simple fluorescent imaging. The instrument can be scaled by adding more flow slides to the liquid-handling deck; this flexibility helps process capacity (DNA spots measured per cycle) keep up with increased imaging speed (e.g., of the version 2 instrument). Standard off-the-shelf liquid-handling robots draw reagents to the flow slides for the sequencing reaction; when the reaction is complete, robots transfer each flow slide to the imaging deck. Current

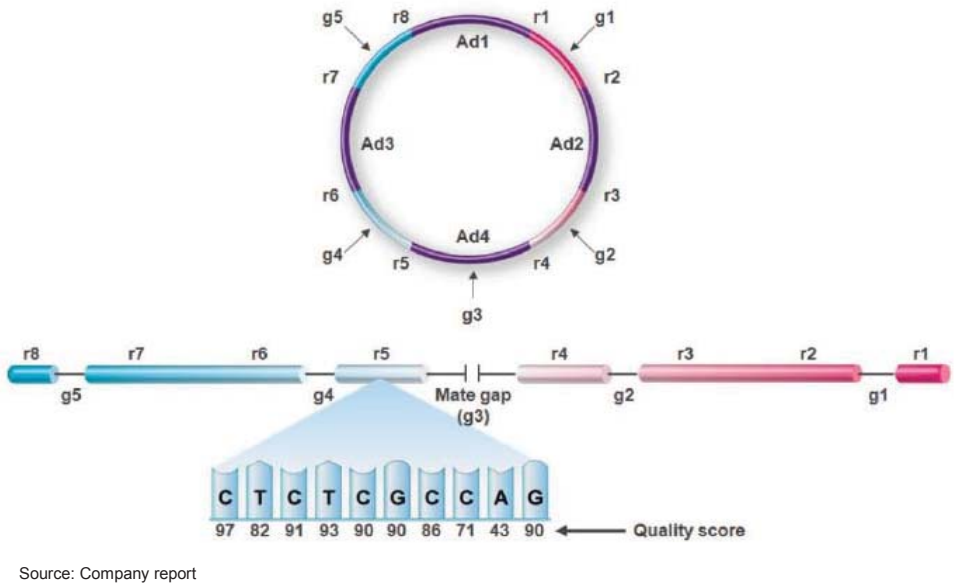
capacity allows running 2-16 slides in parallel (while one slide is imaged, the other slides are also being imaged). Lastly, the imager is constructed with off-the-shelf components and is essentially a four-color fluorescence microscope equipped with an illuminator, filter changer, microscope objective, tube lens, and other common parts from an optical microscope.

Software

Complete Genomics has developed its own software to assemble paired-end reads from its sequencing results and has optimized the software to be especially efficient for human-genome data analysis. According to the company, the software can align reads and assemble human genomes in less than a day.

Base-calling software captures data from the imager after each reaction cycle. The system uses these images to determine the base at each position of the DNA fragment. Four images (one for each color dye) are generated for each base position, and the software assigns a quality score based on its signal strength. The quality score will be used at the assembly step to identify each nucleotide. As figure 26 indicates, although there are gaps between each read of a circular template (due to enzyme digestion), the cPAL chemistry has the benefit of independent reads. Thus, the system has a high tolerance for sequencing errors (even if the calling of one or several bases is missed, it has no impact on the correct calling of the next base); the gap will not affect the accuracy of base calling.

Figure 26
Complete Genomics, Inc.
Base Calling



Source: Company report

In the last steps, base call data is gathered and assembled. Complete Genomics' assembly software is able to identify SNPs, short indels, and block substitutions up to roughly 50 base pairs on over 90% of the genome. Paired-end reads support identification of larger variants (e.g., structural variants). The variant calls are then matched with reference genomes available from the National Center for Biotechnology Information (NCBI) to generate reports that feature variant information and their medical significance.

Each complete genome sequenced will pass a series of quality-control processes, and the data generated will be automatically transferred via a secure network connection using cloud computing technology to an offsite location for storage after completion. The data is generated at 40-times coverage and accuracy of 99.999%, or one false variant in every 100 kilobases

(which will improve to 99.99999% in the near future), with roughly 3.2 million-4.5 million variants identified per human genome. The end-product includes summary statistics; variant calls; quantitative CNV calls; and reads, scores, and mappings.

Deliverables
Complete Genomics has two product offerings:

1. ST001V: Around 40 Gb in data size. Includes a list of all identified sequence variants (including SNPs and small indels, copy number variants, and structural variants), as well as annotation of these variants and data summary reports.

2. ST001RM: Around 400 Gb in data size. Includes all offerings listed above, plus initial mapping and reads data.

The product is delivered virtually via Amazon Web Services or physically via hard drive. A summary of data deliverables can be found in table 6. For customers that request the complete data set (ST001RM), Complete Genomics provides information on individual reads, base calls from single DNBs, initial mapping to the reference genome (NCBI database), and quality score. Customers may also receive the raw sequence data, if desired.

Table 6
Complete Genomics, Inc.
Data Deliverables

Summary Metrics	Details on Variants	Copy Number Variants and Structural Variants
Summary statistics of the genome	Summary of known variants identified	*For tumor genomes:* Segmentation of reference genomes into regions of discrete coverage levels
Number of bases	Annotation of variants in known protein coding genes	Estimated ploidy for the genome and average coverage for every 2 kb of the genome
Coverage depth	Assembled sequence for each variant allele	Information for detected junctions
Read length of small insertion/deletion and substitution identified	Supporting reads for each assembled sequence	Alignments of the individual discordant mate pairs supporting each junction
Library	Base-level coverage and scores	Information and annotation of identified junctions

Source: Company reports

In addition, Complete Genomics offers an open-source platform, the Complete Genomics Analysis Tools (CGA Tools), which enables the comparison of genomes from the customer end.

Customer Case Studies Point to Advantages of Complete Genomics' Technology

In the following section, we discuss some key customer projects, which in our view highlight three important points: 1) use of Complete Genomics' high inherent accuracy for cancer-based clinical trials, 2) the propensity on the part of pharma to outsource whole-genome sequencing, and 3) the ability of Complete Genomics to garner repeat orders.

SAIC-Frederick (National Cancer Institute): Complete Genomics Gaining Traction in Cancer Sequencing
In May 2010, Complete Genomics received an order from SAIC-Frederick (one of the National Cancer Institute's prime research-and-development contractors) to sequence 50 tumor/normal pairs (100 total genomes) over a six-month period with an option to sequence an additional 564 tumor/normal pairs over 18 more months. The sequencing project is a

component of the National Cancer Institute's initiative to identify potential drug targets for pediatric cancers, specifically acute lymphoblastic leukemia, acute myeloid leukemia, neuroblastoma, osteosarcoma, and Wilms tumor.

While the company has already shipped the 100 genomes, SAIC-Frederick is still completing its validation processes, and thus Complete Genomics will not likely recognize revenue for these genomes until the second or third quarter of 2011. We expect Complete Genomics should hear whether SAIC-Frederick intends to move forward with the additional tumor sequencing by the second quarter. In our view, this project suggests sequencing cancer tumors is feasible with Complete Genomics' technology. Securing the second piece of this order would provide further validation for the company's data quality and analysis capabilities in cancer sequencing, in our view, particularly of its newly added copy number and structural variant analysis capabilities, which has been difficult with shorter-read technology.

Genentech, Inc.: Non–small-cell Lung Cancer Study Points to Pharma's Willingness to Outsource Whole-Genome Sequencing

Genentech (a Roche subsidiary) was looking for ways to reduce sequencing costs and therefore leveraged Complete Genomics' services to sequence the first primary lung tumor and evaluate somatic variation.

The sample came from a 51-year-old male who had smoked 25 cigarettes a day for 15 years. Complete Genomics sequenced both the tumor and normal sample at 60-times and 46-times coverage, respectively, and covered more than 90% of the reference genome in both cases. The high coverage enabled the researchers to identify a number of novel as well as known mutations. The study identified 50,000 somatic mutations, or 17.7 mutations per megabase—double the frequency demonstrated in any previously published study, suggesting that perhaps smoking causes an elevated somatic mutation rate. The results of this study were published in *Nature* (William Lee, Zhaoshi Jiang, et al., "The mutation spectrum revealed by paired genome sequences from a lung cancer patients," *Nature* 465, 473-477 [May 27, 2010]).

Although Genentech has in-house Illumina machines that the company uses for small-scale studies, Genentech has continued to work with Complete Genomics for full-human-genome sequencing; this points to the propensity on the part of pharma to outsource whole-genome sequencing.

Institute for Systems Biology (ISB): A Multi-Order Customer

The ISB has placed three orders for whole human genomes in increasingly larger increments. The initial order was for four genomes (specifically a nuclear family with two unaffected parents and two children suffering from Miller syndrome and primary ciliary dyskinesia). These genomes were sequenced with 85% to 92% of the bases called. Analysis identified rare mutations in four genes as candidate genes for Miller syndrome.

The project used 15 µg of DNA sample from each individual, sequenced to a depth of 51-times to 88-times coverage. Among the four genomes sequenced, the call rate (percentage of nucleotide bases being identified as reference from a complete genome) ranged from 85% to 92%, and more than 4.47 million SNPs were identified, with more than 3.67 million SNPs varied among the family members. The accuracy rate, as compared with previous independent studies, was 3.3×10-6 for the called bases, compared with 8.16×10-6 from previous exome sequencing results published (sequencing of two children).

The ISB has since placed two more orders for 100 genomes (delivered in 2010) and 615 genomes for family-based neurodegenerative disease, respectively; the company expects to receive samples for the latter study this year.

Financial and Forecasts

Given its early stage of commercialization, management does not yet provide full-year guidance. The company has indicated it expects to ship more than 500 genomes in the first quarter of 2011. On a GAAP basis, we estimate losses per share of $1.97 in 2011 and $0.96 in 2012, followed by earnings per share of $0.01 in 2013 and $1.03 in 2014. This compares with consensus estimates of -$1.97 for 2011, -$0.80 for 2012, $0.18 for 2013, and $1.12 for 2014.

Revenue
The company recognizes revenue once the genome is shipped and received by the customer, based on contract terms, and collectability can be assured. Any up-front payments received before shipment are recorded as deferred revenues.

Revenue is determined by the number of genomes shipped in the quarter (assuming the recognition criteria are met), multiplied by the price of the genomes as defined by the purchase order or contract. The company has a minimum order of eight genomes and gives volume discounts for large orders. We note that because an order has to actually translate into samples received by the company, revenue recognition for specific orders can be lumpy.

Given the early stage of the company, management does not provide guidance on the number of genomes shipped, other than its expectation to ship more than 500 genomes in the first quarter. In addition, management has suggested its ASP averaged less than $10,000 exiting 2010. It has targeted an ASP of $5,000 exiting 2011.

Given the ability to increase capacity to 800-1,200 genomes per month and its current backlogs, we are modeling more than 4,600 genomes shipped in 2011. In addition, the upgrade to second-generation instruments allows Complete Genomics to increase throughput by 10 times; therefore, we are modeling around 19,000 genomes shipped in 2012, around 45,000 genomes in 2013, and around 76,000 genomes in 2014. We note that although the company has shipped the 122 SAIC-Frederick genomes, it will not recognize revenue for these genomes until the customer has completed its validation process (expected in the second or third quarter of 2011).

We assume it takes a couple of quarters for the ASP the company is currently signing orders for to flow through the P&L line. Therefore, we model an average ASP of $7,400 in 2011, $4,700 in 2012, $3,000 in 2013, and $2,600 in 2014.

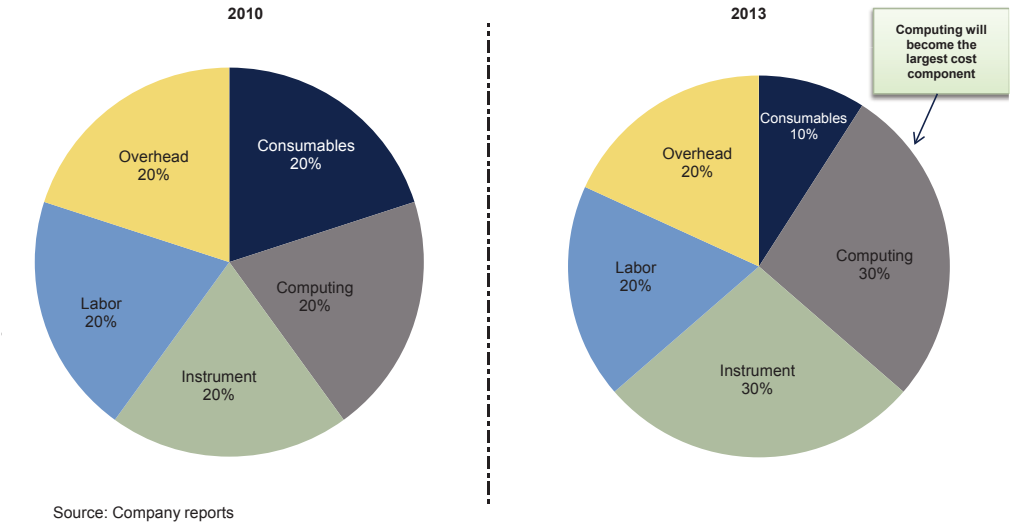
That translates to revenues of $34 million in 2011, $90 million in 2012, $138 million in 2013, and $196 million in 2014.

Cost of Goods Sold
We are assuming that Complete Genomics will be able to drive down its cost per genome sequenced, which contains consumable, instrument operating computing, and labor and overhead costs. Although we model ASP to exceed COGS beginning in the fourth quarter of 2011, we model negative margins for the first half of 2011. The trend then reverses and will gradually improve in 2012 and reach 68% by 2013.

In addition, given the company's plan to introduce the second generation of its sequencing instrument in 2012, which could in theory reduce COGS per genome by 10 times, we model a steeper decline in costs throughout that year.

As shown in figure 27 and according to Complete Genomics' estimate, by 2013, computing, including data storage and CPU power, will become its largest cost component—pointing to the benefits of leveraging its scale to drive down per-genome reagent cost.

Figure 27
Complete Genomics, Inc.
Cost Structure, 2010 and 2013



Source: Company reports

Research-and-Development Expense
Complete Genomics spent around $21.7 million in research and development in 2010, which is roughly the same level as in 2009. Over the past few months, the company reclassified a number of research-and-development employees to its operations group, which led to flat research-and-development spending.

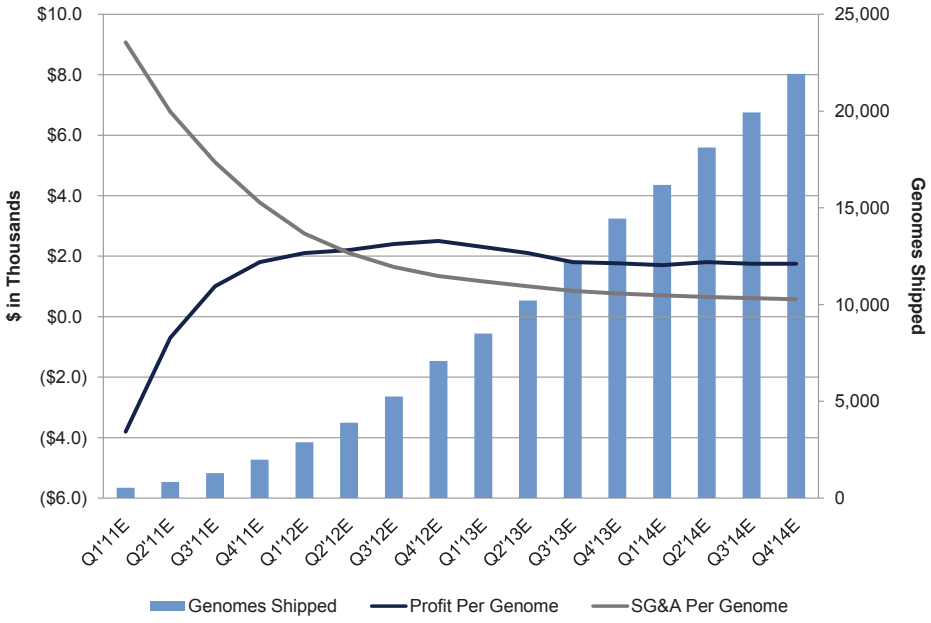
Complete Genomics has several products in its pipeline, such as analytical software for cancer genomes. In addition, as the company continues to optimize its sequencing instruments and prepares for the rollout of the next-generation instrument in 2012, we expect a steep ramp-up in research-and-development expense in the following years.

We model research-and-development expense approaching long-term company guidance of 20% to 25% by 2014.

Selling, General, and Administrative Expense
From its commercial launch in May 2010, the company had increased the number of its sales representatives from 2 to 12 by the end of 2010, and plans to raise the number to 24 when exiting 2011. In addition, management estimates it takes around eight months of training for a sales rep to achieve the company's sales target per rep. We expect the addition of sales reps and early commercialization marketing will increase SG&A expense significantly.

In the long run, we expect SG&A expense of sales reps to be a relative fixed cost; i.e., as the genome orders grow significantly, the operating costs per genome will be optimized. We expect SG&A expense to be roughly on par with research-and-development expense by 2013. We model SG&A expense normalized to 24.6% by 2014.

Figure 28
Complete Genomics, Inc.
SG&A Expense Minimizes as Volume Increases



Source: William Blair & Company, L.L.C. estimates

Balance Sheet and Cash Flow

Through its series E funding in October 2010 and initial public offering in November 2010, the company ended the year with $69 million in cash and equivalents on its balance sheet. In addition, Complete Genomics raised $20 million in debt (although it has since paid down $7 million) in March 2011 to fill the gap in capital it intended to raise through the equity offering.

We expect the company's cash burn to increase in 2011 as Complete Genomics increases its inventory balance and places orders for additional sequencing instruments. Based on the projected cash burn rate, we have modeled an additional cash infusion in 2012 of $60 million. We expect the company to reach positive operating cash flow by 2013.

In terms of capital expenditures, management guidance suggests between $22 million and $25 million of spending in 2011, which includes costs of additional sequencing instruments and other facility expansion costs.

Additional information is available upon request.

This report is available in electronic form to registered users via R*Docs™ at www.rdocs.com or www.williamblair.com.

Please contact us at (800) 621-0687 or consult http://www.williamblair.com/pages/eqresearch_coverage for all disclosures.

DJIA:	12,285.15
S&P 500:	1,314.52
NASDAQ:	2,760.22

Amanda Murphy 312.364.8951

Table 7
Complete Genomics, Inc,
Comparable Valuation Summary

Company	Ticker	Price 4/14/2011	Shares Outstanding	L-T Mean EPS Growth	Market Cap	Enterprise Value	Revenue 2011E	Revenue 2012E	EV/Revenue 2011E	EV/Revenue 2012E
Diagnostics Services Labs										
Bio-Reference Labs	BRLI	23.19	28	20%	$652	$676	$559	$662	1.2x	1.0x
Genomic Health	GHDX	25.92	23	20%	$592	$697	$205	$232	3.4x	3.0x
Laboratory Corporation of America	LH	95.06	105	12%	$9,934	$11,996	$5,531	$5,834	2.2x	2.1x
Myriad Genetics	MYGN	20.10	21	20%	$430	$1,508	$411	$468	3.7x	3.2x
Quest Diagnostics	DGX	57.76	172	10%	$9,923	$12,588	$7,425	$7,591	1.7x	1.7x
Average				***16%***	***$4,306***	***$5,493***	***$2,826***	***$2,957***	***2.4x***	***2.2x***
Diagnostic Products										
Luminex	LMNX	19.25	41	27%	$795	$716	$166	$197	4.3x	3.6x
Qiagen NV	QGEN	20.00	233	12%	$4,657	$3,207	$1,176	$1,290	2.7x	2.5x
Sequenom	SQNM	7.09	99	30%	$702	$554	$54	$82	10.3x	6.7x
Average				***23%***	***$2,051***	***$1,492***	***$465***	***$523***	***5.8x***	***4.3x***
Life Sciences Tools										
Affymetrix	AFFX	5.62	34	8%	$191	$397	$312	$319	1.3x	1.2x
Complete Genomics	GNOM	11.70	26	20%	$303	$247	$34	$90	7.2x	2.8x
Life Technologies	LIFE	52.71	180	11%	$9,491	$11,783	$3,819	$4,049	3.1x	2.9x
Illumina	ILMN	67.28	140	30%	$9,425	$8,923	$1,111	$1,375	8.0x	6.5x
Pacific Biosciences	PACB	12.52	53	30%	$662	$367	$32	$107	NM	3.4x
Average				***20%***	***$4,014***	***$4,368***	***$1,062***	***$1,188***	***4.9x***	***3.4x***
HCIT										
Medidata Solutions	MDSO	25.82	24	15%	$623	$529	$184	$212	2.9x	2.5x
Pharmaceutical Outsourcing (CROs)										
Charles River Laboratories	CRL	41.97	52	13%	$2,166	$2,657	$1,137	$1,175	2.3x	2.3x
Covance	CVD	60.77	60	16%	$3,671	$3,416	$2,040	$2,168	1.7x	1.6x
ICON plc	ICLR	23.15	60	12%	$1,396	$1,144	$956	$1,033	1.2x	1.1x
Parexel International	PRXL	25.52	59	17%	$1,496	$1,721	$1,229	$1,346	1.4x	1.3x
Pharmaceutical Product Development	PPDI	30.75	113	17%	$3,477	$2,921	$1,523	$1,678	1.9x	1.7x
Average				***15%***	***$2,441***	***$2,372***	***$1,377***	***$1,480***	***1.7x***	***1.6x***
Market Index			YTD Change							
S&P 500	SPX	1314.52	3%							
NASDAQ	COMP	2760.22	3%							

(in millions, except per share data)

Sources: Thomson One and William Blair & Company, L.L.C. estimates

Table 8
Complete Genomics, Inc.
Discounted Cash Flow Analysis

DCF Assumptions

Company information		**Q4'10**
Most recent share price	$	11.70
Shares outstanding (thousands)		52,900
Cost of capital inputs:		
Levered beta		0.794
Unlevered beta		0.78
Risk Free Rate (10 Year T-Bill)		*3.56%*
Equity Risk Premium (Ibbotson)		*6.70%*
Size Premia (Ibbotson)		*7.80%*

Capital structure:			
$ thousands		**Q4'10**	**% of total**
Net debt:			
Gross debt	$	13,301	*2.1%*
Cash and equivalents	$	-	*0.0%*
Marketable securities	$	-	*0.0%*
Net debt	$	13,301	*2.1%*
Market capitalization	$	618,930	*97.9%*
Net invested capital	**$**	**632,231**	***100.0%***

Weighed average cost of capital:	
Cost of equity (based on CAPM formula)	*15.1%*
Equity as % of total invested capital	*97.9%*
Weighted average cost of equity	***14.8%***
Cost of debt (estimated)	*6.5%*
Representative tax rate	*37.0%*
Net debt as % of total invested capital	*2.1%*
Weighted average cost debt	***0.1%***
Weighted average cost of capital	***14.8%***

Free Cash Flow Forecast

Periods to cash		0.50	1.50	2.50	3.50	4.50	5.50	6.50	7.50	8.50	9.50	10.50	CAGR
$ thousands, unless noted	**2010A**	**2011E**	**2012E**	**2013E**	**2014E**	**2015E**	**2016E**	**2017E**	**2018E**	**2019E**	**2020E**	**2021E**	**'10 - '21**
Revenue	**9,389**	**34,171**	**89,795**	**138,366**	**195,868**	**254,628**	**318,285**	**375,577**	**443,180**	**509,657**	**586,106**	**674,022**	***47.5%***
% growth/(decline)	*1407.1%*	*263.9%*	*162.8%*	*54.1%*	*41.6%*	*30.0%*	*25.0%*	*18.0%*	*18.0%*	*15.0%*	*15.0%*	*15.0%*	
Cost of revenue	19,895	31,921	44,901	49,952	62,528	81,287	101,609	119,898	139,264	160,153	184,176	211,803	***NM***
Gross profit	**(10,506)**	**2,250**	**44,893**	**88,414**	**133,340**	**173,341**	**216,677**	**255,679**	**303,917**	**349,504**	**401,930**	**462,219**	***NM***
% margin	*-111.9%*	*6.6%*	*50.0%*	*63.9%*	*68.1%*	*68.1%*	*68.1%*	*68.1%*	*68.6%*	*68.6%*	*68.6%*	*68.6%*	
Operating expenses	37,147	54,307	71,652	87,089	100,767	108,829	117,535	126,938	137,093	148,060	159,905	172,697	*15.0%*
% growth/(decline)	*27.3%*	*46.2%*	*31.9%*	*21.5%*	*15.7%*	*8.0%*	*8.0%*	*8.0%*	*8.0%*	*8.0%*	*8.0%*	*8.0%*	
Operating income	**(47,653)**	**(52,057)**	**(26,759)**	**1,325**	**32,572**	**64,513**	**99,142**	**128,741**	**166,824**	**201,444**	**242,025**	**289,522**	***NM***
% growth/(decline)	*41.9%*	*9.2%*	*-48.6%*	*-105.0%*	*2358.1%*	*98.1%*	*53.7%*	*29.9%*	*29.6%*	*20.8%*	*20.1%*	*19.6%*	
(-) income taxes	-	-	-	-	-	-	-	-	-	-	-	-	*NM*
(+) depreciation and amortization	8,428	5,221	8,860	11,512	15,465	16,702	18,038	19,481	21,040	22,723	24,541	26,504	*11.0%*
(-) (incr)/decr in working capital	644	(2,784)	5,535	11,326	3,842	9,639	10,410	11,243	12,142	13,114	14,163	15,296	*33.4%*
(-) capital expenditures	(18,802)	(23,600)	(28,320)	(18,408)	(18,408)	(19,881)	(21,471)	(23,189)	(25,044)	(27,047)	(29,211)	(31,548)	*4.8%*
(+) stock-based comp	1,751	2,903	4,537	6,226	8,814	9,519	10,281	11,103	11,991	12,951	13,987	15,106	*21.6%*
Total free cash flow (equity)	**(55,632)**	**(70,317)**	**(36,147)**	**11,982**	**42,286**	**80,492**	**116,400**	**147,379**	**186,954**	**223,184**	**265,504**	**314,879**	***NM***
% growth/(decline)	50.5%	26.4%	*-48.6%*	*-133.1%*	*252.9%*	*90.4%*	*44.6%*	*26.6%*	*26.9%*	*19.4%*	*19.0%*	*18.6%*	

Output Tables

Implied Share Price

		Perpetual growth rate used in terminal value calculation						
		-1.0%	***0.0%***	***1.0%***	***2.0%***	***3.0%***	***4.0%***	***5.0%***
	12.8%	$ 20.60	$ 21.66	$ 22.91	$ 24.38	$ 26.15	$ 28.32	$ 31.05
Discount	***13.8%***	$ 18.16	$ 19.01	$ 19.98	$ 21.13	$ 22.48	$ 24.11	$ 26.11
rate	***14.8%***	$ 16.09	$ 16.77	$ 17.55	**$ 18.45**	$ 19.50	$ 20.75	$ 22.25
(WACC)	***15.8%***	$ 14.31	$ 14.86	$ 15.49	$ 16.21	$ 17.04	$ 18.01	$ 19.15
	16.8%	$ 12.78	$ 13.23	$ 13.74	$ 14.32	$ 14.98	$ 15.74	$ 16.64

% Premium/(Discount) to Current Share Price

	Perpetual growth rate used in terminal value calculation						
	-1.0%	***0.0%***	***1.0%***	***2.0%***	***3.0%***	***4.0%***	***5.0%***
12.8%	*76.1%*	*85.1%*	*95.8%*	*108.4%*	*123.5%*	*142.1%*	*165.4%*
13.8%	*55.2%*	*62.4%*	*70.8%*	*80.6%*	*92.2%*	*106.1%*	*123.2%*
14.8%	*37.5%*	*43.3%*	*50.0%*	***57.7%***	*66.7%*	*77.3%*	*90.1%*
15.8%	*22.3%*	*27.0%*	*32.4%*	*38.5%*	*45.6%*	*53.9%*	*63.7%*
16.8%	*9.2%*	*13.1%*	*17.4%*	*22.4%*	*28.0%*	*34.6%*	*42.2%*

Sources: Company reports and William Blair & Company, L.L.C. estimates

Table 9
Complete Genomics, Inc.
Income Statement

	2008	2009	2010	Q1'11E	Q2'11E	Q3'11E	Q4'11E	2011E	Q1'12E	Q2'12E	Q3'12E	Q4'12E	2012E	2013E	2014E
Revenue:															
Total Revenue	**-**	**$623**	**$9,389**	**$4,478**	**$6,446**	**$9,350**	**$13,897**	**$34,171**	**$18,424**	**$19,431**	**$23,609**	**$28,331**	**$89,795**	**$138,366**	**$195,868**
Cost of Revenue	-	5,033	19,895	6,504	7,024	8,069	10,324	31,921	12,378	10,881	11,018	10,624	44,901	49,952	62,528
Gross Profit	-	($4,410)	($10,506)	($2,026)	($578)	$1,281	$3,574	$2,250	$6,045	$8,550	$12,591	$17,707	$44,893	$88,414	$133,340
R&D	23,633	22,424	21,691	6,717	7,090	7,667	8,338	29,813	8,659	8,900	9,562	10,341	37,461	45,353	52,545
SG&A	4,224	6,751	15,456	4,837	5,608	6,545	7,504	24,494	7,922	8,161	8,617	9,491	34,191	41,736	48,222
Total Operating Expenses	27,857	29,175	37,147	11,554	12,698	14,212	15,843	54,307	16,581	17,061	18,179	19,832	71,652	87,089	100,767
Operating Income	(27,857)	(33,585)	(47,653)	(13,580)	(13,276)	(12,931)	(12,269)	(52,057)	(10,536)	(8,511)	(5,587)	(2,125)	(26,759)	1,325	32,572
Interest Expense (Income) and Other, net	537	2,364	10,034	130	130	138	162	560	211	211	160	181	763	894	991
Pretax Income	($28,394)	($35,949)	($57,687)	($13,710)	($13,406)	($13,070)	($12,431)	($52,617)	($10,747)	($8,722)	($5,747)	($2,306)	($27,522)	$431	$31,581
Income Tax Expense (Benefit)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Net Income (excl. non-recurr items)	($28,394)	($35,949)	($57,687)	($13,710)	($13,406)	($13,070)	($12,431)	($52,617)	($10,747)	($8,722)	($5,747)	($2,306)	($27,522)	$431	$31,581
Stock-Based Comp	336	1,410	1,751	537	645	748	973	2,903	1,013	1,069	1,180	1,275	4,537	6,226	8,814
Tax Adjustments		-	-	-	-	-	-	-	-	-	-	-	-	-	-
Stock-Based Comp (net of tax)	336	1,410	1,751	537	645	748	973	2,903	1,013	1,069	1,180	1,275	4,537	6,226	8,814
Adjusted Net Income (excl. stock-based comp)	(28,058)	(34,539)	(55,936)	(13,173)	(12,762)	(12,321)	(11,459)	(49,714)	(9,734)	(7,653)	(4,567)	(1,031)	(22,985)	6,657	40,395
Non Recurring Items	-	-	(405)	-	-	-	-	-	-	-	-	-	-	-	-
Tax Adjustments	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Non Recurring Items (net of tax)	-	-	(405)	-	-	-	-	-	-	-	-	-	-	-	-
Net Income (GAAP)	**($28,394)**	**($35,949)**	**($58,092)**	**($13,710)**	**($13,406)**	**($13,070)**	**($12,431)**	**($52,617)**	**($10,747)**	**($8,722)**	**($5,747)**	**($2,306)**	**($27,522)**	**$431**	**$31,581**
EPS (excl. non-recurr. Items)	($305.32)	($386.56)	($13.51)	($0.52)	($0.51)	($0.49)	($0.45)	($1.97)	($0.38)	($0.31)	($0.20)	($0.08)	($0.96)	$0.01	$1.03
Adjusted EPS (excl. stock-based comp)	($301.71)	($371.40)	($13.10)	($0.50)	($0.48)	($0.46)	($0.42)	($1.86)	($0.35)	($0.27)	($0.16)	($0.03)	($0.80)	$0.22	$1.31
EPS (GAAP)	**($305.32)**	**($386.56)**	**($13.60)**	**($0.52)**	**($0.51)**	**($0.49)**	**($0.45)**	**($1.97)**	**($0.38)**	**($0.31)**	**($0.20)**	**($0.08)**	**($0.96)**	**$0.01**	**$1.03**
W. Avg. Shares Outstanding (Diluted)	93	93	4,271	26,200	26,500	26,900	27,400	26,750	27,950	28,500	29,050	29,600	28,775	30,350	30,800
MARGIN ANALYSIS:															
Total Gross Profit	**NM**	(707.9%)	(111.9%)	(45.2%)	(9.0%)	13.7%	25.7%	6.6%	32.8%	44.0%	53.3%	62.5%	50.0%	63.9%	68.1%
R&D	NM	3599.4%	231.0%	150.0%	110.0%	82.0%	60.0%	87.2%	47.0%	45.8%	40.5%	36.5%	41.7%	32.8%	26.8%
SG&A	NM	1083.6%	164.6%	108.0%	87.0%	70.0%	54.0%	71.7%	43.0%	42.0%	36.5%	33.5%	38.1%	30.2%	24.6%
EBIT	**NM**	**NM**	**(507.5%)**	**(303.2%)**	**(206.0%)**	**(138.3%)**	**(88.3%)**	**(152.3%)**	**(57.2%)**	**(43.8%)**	**(23.7%)**	**(7.5%)**	**(29.8%)**	**1.0%**	**16.6%**
Tax Rate	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Stock-Based Comp	NM	226.3%	18.6%	12.0%	10.0%	8.0%	7.0%	8.5%	5.5%	5.5%	5.0%	4.5%	5.1%	4.5%	4.5%
Net Income (GAAP)	**NM**	**NM**	**(614.4%)**	**(306.1%)**	**(208.0%)**	**(139.8%)**	**(89.5%)**	**(154.0%)**	**(58.3%)**	**(44.9%)**	**(24.3%)**	**(8.1%)**	**(30.7%)**	**0.3%**	**16.1%**
GROWTH METRICS:															
Total Revenue Growth	**NA**	**NM**	**1407%**	**1233%**	**492%**	**125%**	**265%**	**264%**	**311%**	**201%**	**152%**	**104%**	**163%**	**54%**	**42%**
COGS	NA	NM	295.3%	60%	43%	34%	111%	60.4%	90%	55%	37%	3%	41%	11%	25%
Gross Profit	NA	NM	NM	NM	NM	NM	NM	NM	NM	NM	883%	395%	NM	97%	51%
R&D	NA	(5%)	(3%)	9%	44%	55%	48%	37%	29%	26%	25%	24%	26%	21%	16%
SG&A	NA	60%	129%	12%	82%	67%	82%	58%	64%	46%	32%	26%	40%	22%	16%
Operating Income	**NA**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**2358%**
Net Income (excl. nonrecur.)	NA	NM	NM	NM	NM	NM	NM	66%	NM	NM	NM	NM	NM	NM	7228%
Stock Based Comp	NA	320%	24%	(0%)	85%	77%	121%	66%	89%	66%	58%	31%	56%	37%	42%
EPS (GAAP)	**NA**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	7120%
EPS (excl. stock-based comp)	**NA**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	**NM**	498%
Diluted Shares Outstanding	NA	-	4493%	513%	520%	530%	85%	526%	7%	8%	8%	8%	8%	5%	1%

Sources: Company reports and William Blair & Company, L.L.C. estimates

Table 10
Complete Genomics, Inc.
Balance Sheet

	2008	2009	2010	Q1'11E	Q2'11E	Q3'11E	Q4'11E	2011E	Q1'12E	Q2'12E	Q3'12E	Q4'12E	2012E	2013E	2014E
Cash and Cash Equivalents	**$6,186**	**$7,765**	**$68,918**	**$62,892**	**$44,091**	**$25,030**	**$7,890**	**$7,890**	**$51,493**	**$36,887**	**$23,272**	**$14,118**	**$14,118**	**$10,037**	**$37,849**
Accounts Receivable, Net	-	1,288	4,943	4,478	6,446	8,831	13,125	13,125	17,400	17,704	21,510	25,183	25,183	32,494	46,270
Inventory, net	-	354	3,980	5,420	5,853	6,993	8,029	8,029	9,628	8,342	8,324	7,673	7,673	9,654	11,691
Prepaid Expenses and Other Current Assets	1,034	5,612	1,179	1,191	1,203	1,215	1,227	1,227	1,239	1,252	1,264	1,277	1,277	1,329	1,382
Total Current Assets	7,220	15,019	79,020	73,981	57,592	42,068	30,271	30,271	79,760	64,185	54,371	48,250	48,250	53,513	97,193
PP&E, net	8,023	14,864	23,843	28,906	33,874	38,662	43,262	43,262	48,776	54,287	59,661	64,815	64,815	74,614	81,541
Goodwill	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Intangible Assets & Other	511	395	297	149	(15)	(211)	(441)	(441)	(717)	(994)	(1,295)	(1,635)	(1,635)	(3,154)	(5,180)
Total Assets	**$15,754**	**$30,278**	**$103,160**	**$103,036**	**$91,451**	**$80,518**	**$73,092**	**$73,092**	**$127,819**	**$117,478**	**$112,737**	**$111,430**	**$111,430**	**$124,973**	**$173,554**
Accounts Payable	745	4,281	3,066	4,192	4,683	5,559	6,882	6,882	8,940	7,859	7,957	8,027	8,027	10,866	14,321
Accrued Liabilities	1,744	2,032	3,102	4,228	4,566	5,164	6,607	6,607	7,675	6,747	6,721	6,374	6,374	7,874	9,470
Long-Term Debt, Current Portion	3,990	4,440	5,780	5,780	5,780	5,780	5,780	5,780	5,780	5,780	5,780	5,780	5,780	5,780	5,780
Other Current Liabilities	-	1,302	5,739	3,903	4,214	4,035	5,162	5,162	6,313	5,441	4,958	4,675	4,675	5,837	7,014
Total Current Liabilities	6,479	12,055	17,687	18,102	19,243	20,538	24,431	24,431	28,708	25,826	25,416	24,856	24,856	30,357	36,585
Other Long-term Liabilities	-	5,017	4,316	4,316	4,316	4,316	4,316	4,316	4,316	4,316	4,316	4,316	4,316	4,316	4,316
Long-Term Debt, Less Current Portion	7,707	3,510	7,521	7,555	7,591	7,684	7,823	7,823	8,007	8,202	8,438	8,721	8,721	10,105	12,063
Convertible Preferred Stock Warrant Liability	1,100	1,553	-	-	-	-	-	-	-	-	-	-	-	-	-
Total Liabilities	**$15,286**	**$22,135**	**$29,524**	**$29,973**	**$31,149**	**$32,538**	**$36,571**	**$36,571**	**$41,031**	**$38,344**	**$38,169**	**$37,893**	**$37,893**	**$44,778**	**$52,965**
Preferred Stock	45,622	85,833	-	-	-	-	-	-	-	-	-	-	-	-	-
Common Stock	-	-	26	26	26	26	26	26	26	26	26	26	26	26	26
Additional Paid-in Capital	58	3,471	212,458	225,058	225,058	225,058	225,058	225,058	285,058	285,058	285,058	285,058	285,058	285,058	285,058
Accumulated Deficit	(45,212)	(81,161)	(138,848)	(152,021)	(164,782)	(177,104)	(188,562)	(188,562)	(198,296)	(205,950)	(210,516)	(211,547)	(211,547)	(204,890)	(164,495)
Other	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Total Stockholder's Equity	**$468**	**$8,143**	**$73,636**	**$73,063**	**$60,302**	**$47,980**	**$36,522**	**$36,522**	**$86,788**	**$79,134**	**$74,568**	**$73,537**	**$73,537**	**$80,194**	**$120,589**
Total Liabilities and Stockholder's Equity	**$15,754**	**$30,278**	**$103,160**	**$103,037**	**$91,451**	**$80,518**	**$73,092**	**$73,092**	**$127,819**	**$117,478**	**$112,737**	**$111,430**	**$111,430**	**$124,973**	**$173,554**
Key Metrics:															
Debt-to-Capital Ratio	90%	54%	12%	12%	14%	17%	22%	22%	10%	11%	12%	12%	12%	11%	8%
Net Debt-to-Capital Ratio	(49%)	10%	(70%)	(63%)	(48%)	(26%)	5%	5%	(43%)	(30%)	(16%)	(5%)	(5%)	0%	(21%)
Working Capital (excl. cash and ST investments)	(5,445)	(4,801)	(7,585)	(7,013)	(5,741)	(3,499)	(2,050)	(2,050)	(441)	1,472	5,683	9,276	9,276	13,119	22,758
Debt to EBITDA	(0.5x)	(0.3x)	(0.3x)	(0.3x)	(0.3x)	(0.4x)	(0.4x)	(0.3x)	(0.3x)	(0.3x)	(0.3x)	(0.3x)	(0.8x)	1.2x	0.4x
A/P days	NA	306	55	58	60	62	60	78	65	65	65	68	64	78	82
A/R DSO	NA	744	190	90	90	85	85	138	85	82	82	80	101	85	85
Inventory Days	NA	25	72	75	75	78	70	91	70	69	68	65	62	70	67
Other Current Assets (% of Sales)	NA	3603.2%	50.2%	26.6%	18.7%	13.0%	8.8%	14.4%	6.7%	6.4%	5.4%	4.5%	5.7%	3.8%	2.8%
Accrued Liabilities (% of COGS)	NA	161.5%	62.4%	65.0%	65.0%	64.0%	64.0%	82.8%	62.0%	62.0%	61.0%	60.0%	56.8%	63.1%	60.6%
Other Current Liabilities (% of COGS)	NA	25.9%	115.4%	60.0%	60.0%	50.0%	50.0%	64.7%	51.0%	50.0%	45.0%	44.0%	41.6%	46.7%	44.9%
ROE (excluding one time items)	NA	(835.0%)	(141.1%)	(74.8%)	(80.4%)	(96.6%)	(117.7%)	(95.5%)	(69.7%)	(42.1%)	(29.9%)	(12.5%)	(50.0%)	0.6%	31.5%
ROA (excluding one time items)	NA	(156.2%)	(86.5%)	(53.2%)	(55.1%)	(60.8%)	(64.7%)	(59.7%)	(42.8%)	(28.4%)	(20.0%)	(8.2%)	(29.8%)	0.4%	21.2%
Cash per Share (including ST investments)	$66.52	$83.50	$16.14	$2.40	$1.66	$0.93	$0.29	$0.29	$1.84	$1.29	$0.80	$0.48	$0.49	$0.33	$1.23
Book Value per share (including goodwill)	$5.03	$87.56	$17.24	$2.79	$2.28	$1.78	$1.33	$1.37	$3.11	$2.78	$2.57	$2.48	$2.56	$2.64	$3.92
NOPAT (fully taxed)	($27,857)	($33,585)	($47,653)	($13,580)	($13,276)	($12,931)	($12,269)	($52,057)	($10,536)	($8,511)	($5,587)	($2,125)	($26,759)	$1,325	$32,572
Invested Capital	$7,079	$14,898	$22,335	$27,822	$33,898	$40,731	$46,551	$46,551	$53,398	$60,545	$69,829	$78,236	$78,236	$90,358	$104,899
ROIC	NA	(327%)	(310%)	(219%)	(174%)	(140%)	(114%)	(153%)	(86%)	(61%)	(35%)	(12%)	(44%)	1%	32%

Sources: Company reports and William Blair & Company, L.L.C. estimates

Amanda Murphy 312.364.8951

Table 11
Complete Genomics, Inc.
Cash Flow Statement

	2008	2009	2010	Q1'11E	Q2'11E	Q3'11E	Q4'11E	2011E	Q1'12E	Q2'12E	Q3'12E	Q4'12E	2012E	2013E	2014E
Cash Flow from Operating Activities:															
Net Income from Continuing Operations	(28,394)	(35,949)	(57,687)	(13,710)	(13,406)	(13,070)	(12,431)	(52,617)	(10,747)	(8,722)	(5,747)	(2,306)	(27,522)	431	31,581
Adjustments to reconcile Net Income	3,798	7,965	20,972	1,556	1,776	2,151	2,640	8,124	3,040	3,109	3,423	3,825	13,397	17,739	24,279
Depreciation & Amortization	3,021	5,499	8,428	1,019	1,132	1,403	1,668	5,221	2,027	2,040	2,243	2,550	8,860	11,512	15,465
Stock-Based Compensation Expense	336	1,410	1,751	537	645	748	973	2,903	1,013	1,069	1,180	1,275	4,537	6,226	8,814
Change in Convt. Preferred Stock Warrant Liability Fair Value	(65)	(1,088)	7,211	-	-	-	-	-	-	-	-	-	-	-	-
Others	506	2,144	3,582	-	-	-	-	-	-	-	-	-	-	-	-
Change in Assets and Liabilities	293	1,322	1,697	(572)	(1,272)	(2,242)	(1,449)	(5,535)	(1,609)	(1,913)	(4,211)	(3,593)	(11,326)	(3,842)	(9,639)
Change In Assets	(394)	(6,363)	(43,878)	(987)	(2,412)	(3,537)	(5,343)	(12,279)	(5,885)	969	(3,801)	(3,034)	(11,751)	(9,344)	(15,867)
Change In Liabilities	687	7,685	32,589	415	1,141	1,295	3,893	6,744	4,276	(2,881)	(410)	(560)	425	5,501	6,228
Net Operating Cash Flow	**($24,303)**	**($26,662)**	**($35,018)**	**($12,726)**	**($12,902)**	**($13,161)**	**($11,240)**	**($50,028)**	**($9,317)**	**($7,526)**	**($6,535)**	**($2,075)**	**($25,452)**	**$14,327**	**$46,221**
% of Net Income	*86%*	*74%*	*60%*	*93%*	*96%*	*101%*	*90%*	*95%*	*87%*	*86%*	*114%*	*90%*	*92%*	*3324%*	*146%*
% Growth y-y			*31%*	*201%*	*30%*	*18%*	*16%*	*43%*	*-27%*	*-42%*	*-50%*	*-82%*	*-49%*	*-156%*	*223%*
Cash Flow from Investing Activities:															
Investment in PP&E (CAPEX)	(7,419)	(9,654)	(18,802)	(5,900)	(5,900)	(5,900)	(5,900)	(23,600)	(7,080)	(7,080)	(7,080)	(7,080)	(28,320)	(18,408)	(18,408)
Y/Y Growth		*30%*	*95%*	*NM*	*NM*	*NM*	*NM*	*26%*	*20%*	*20%*	*20%*	*20%*	*20%*	*-35%*	*0%*
Purchase of Investments	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Sales and Maturities of Investments	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Others	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Net Cash Used in Investing Activities	**($7,419)**	**($9,654)**	**($18,802)**	**($5,900)**	**($5,900)**	**($5,900)**	**($5,900)**	**($23,600)**	**($7,080)**	**($7,080)**	**($7,080)**	**($7,080)**	**($28,320)**	**($18,408)**	**($18,408)**
Cash Flow from Financing Activities:															
Proceeds from Issuance of Stock & Notes Payable	38,593	41,881	122,820	12,600	-	-		12,600	60,000	-	-	-	60,000	-	-
Proceeds from Exercise of Options	25	4	796	-	-	-	-	-	-	-	-	-	-	-	-
Repayment of Notes Payable	(4,970)	(3,990)	(8,643)	-	-	-	-	-	-	-	-	-	-	-	-
Net Cash Generated in Financing Activities	**$33,648**	**$37,895**	**$114,973**	**$12,600**	**-**	**-**	**-**	**$12,600**	**$60,000**	**-**	**-**	**-**	**$60,000**	**-**	**-**
Net increase (decrease) in cash and equivalents	**$1,926**	**$1,579**	**$61,153**	**($6,026)**	**($18,802)**	**($19,061)**	**($17,140)**	**($61,028)**	**$43,603**	**($14,606)**	**($13,615)**	**($9,155)**	**$6,228**	**($4,081)**	**$27,813**
Beginning Cash (including short-term investments)	4,260	6,186	7,765	68,918	62,892	44,091	25,030	68,918	7,890	51,493	36,887	23,272	7,890	14,118	10,037
Ending Cash (including short-term investments)	**$6,186**	**$7,765**	**$68,918**	**$62,892**	**$44,091**	**$25,030**	**$7,890**	**$7,890**	**$51,493**	**$36,887**	**$23,272**	**$14,118**	**$14,118**	**$10,037**	**$37,849**
Free Cash Flow (=Oper. Cash Flow - CAPEX)	**($31,722)**	**($36,316)**	**($53,820)**	**($18,626)**	**($18,802)**	**($19,061)**	**($17,140)**	**($73,628)**	**($16,397)**	**($14,606)**	**($13,615)**	**($9,155)**	**($53,772)**	**($4,081)**	**$27,813**
% of Net Income	*112%*	*101%*	*93%*	*136%*	*140%*	*146%*	*138%*	*140%*	*153%*	*167%*	*237%*	*397%*	*195%*	*-947%*	*88%*
% Growth y-y			*48%*	*29%*	*21%*	*41%*	*64%*	*37%*	*-12%*	*-22%*	*-29%*	*-47%*	*-27%*	*-92%*	*-782%*
Free Cash Flow per Share	**($341.10)**	**($390.50)**	**($12.60)**	**($0.72)**	**($0.72)**	**($0.72)**	**($0.64)**	**($2.79)**	**($0.60)**	**($0.52)**	**($0.48)**	**($0.32)**	**($1.91)**	**($0.14)**	**$0.90**
% of EPS	*112%*	*101%*	*93%*	*137%*	*142%*	*148%*	*140%*	*142%*	*156%*	*171%*	*241%*	*405%*	*199%*	*-956%*	*88%*

Sources: Company reports and William Blair & Company, L.L.C. estimates

The prices of the common stock of other public companies mentioned in this report follow:

Affymetrix	$5.62
Amazon.com, Inc. (Outperform)	$125.10
Bio-Reference Labs	$23.19
Charles River Laboratories	$41.97
Covance	$60.77
Eli Lilly and Company	$24.60
FedEx Corporation (Outperform)	$92.94
Genomic Health, Inc. (Market Perform)	$25.92
ICON plc	$23.15
Illumina, Inc. (Outperform)	$67.28
Laboratory Corporation of America	$95.06
Life Technologies	$52.71
Luminex	$19.25
Medidata Solutions	$25.82
Merck & Company, Inc.	$33.86
Myriad Genetics, Inc. (Outperform)	$20.10
Pacific Biosciences of California, Inc. (Market Perform)	$12.52
Parexel International	$25.52
PerkinElmer, Inc.	$26.24
Pfizer, Inc.	$20.49
Pharmaceutical Product development	$30.75
Roche Holding AG	$132.9
SAIC, Inc. (Market Perform)	$16.92
Sequenom	$7.09
Qiagen NV	$20.00
Quest Diagnostics	$57.76

William Blair & Company, L.L.C. is a market maker in the security of this company and may have a long or short position.

Current Ratings Distribution (as of 3/31/11)

Coverage Universe	Percent	Inv. Banking Relationships*	Percent
Outperform (Buy)	58%	Outperform (Buy)	9%
Market Perform (Hold)	31%	Market Perform (Hold)	2%
Underperform (Sell)	1%	Underperform (Sell)	0%

* Percentage of companies in each rating category that are investment banking clients, defined as companies for which William Blair has received compensation for investment banking services within the past 12 months.

Amanda Murphy attests that 1) all of the views expressed in this research report accurately reflect her personal views about any and all of the securities and companies covered by this report, and 2) no part of her compensation was, is, or will be related, directly or indirectly, to the specific recommendations or views expressed by him/her in this report. We seek to update our research as appropriate, but various regulations may prohibit us from doing so. Other than certain periodical industry reports, the majority of reports are published at irregular intervals as deemed appropriate by the analyst.

Stock Rating: William Blair & Company, L.L.C. uses a three-point system to rate stocks. Individual ratings reflect the expected performance of the stock relative to the broader market over the next 12 months. The assessment of expected performance is a function of near-term company fundamentals, industry outlook, confidence in earnings estimates, valuation, and other factors. Outperform (O) – stock expected to outperform the broader market over the next 12 months; Market Perform (M) – stock expected to perform approximately in line with the broader market over the next 12 months; Underperform (U) – stock expected to underperform the broader market over the next 12 months; Not Rated (NR) – the stock is currently not rated.

Company Profile: The William Blair research philosophy is focused on quality growth companies. Growth companies by their nature tend to be more volatile than the overall stock market. Company profile is a fundamental assessment, over a longer-term horizon, of the business risk of the company relative to the broader William Blair universe. Factors assessed include: 1) durability and strength of franchise (management strength and track record, market leadership, distinctive capabilities);

2) financial profile (earnings growth rate/consistency, cash flow generation, return on investment, balance sheet, accounting); 3) other factors such as sector or industry conditions, economic environment, confidence in long-term growth prospects, etc. Established Growth (E) – Fundamental risk is lower relative to the broader William Blair universe; Core Growth (C) – Fundamental risk is approximately in line with the broader William Blair universe; Aggressive Growth (A) – Fundamental risk is higher relative to the broader William Blair universe.

The ratings and company profile assessments reflect the opinion of the individual analyst and are subject to change at any time.

The compensation of the research analyst is based on a variety of factors, including performance of his or her stock recommendations; contributions to all of the firm's departments, including asset management, corporate finance, institutional sales, and retail brokerage; firm profitability; and competitive factors.

Our salespeople, traders, and other professionals may provide oral or written market commentary or trading strategies—to our clients and our trading desks—that are contrary to opinions expressed in this research. Our asset management and trading desks may make investment decisions that are inconsistent with recommendations or views expressed in this report. We will from time to time have long or short positions in, act as principal in, and buy or sell the securities referred to in this report. Our research is disseminated primarily electronically, and in some instances in printed form. Electronic research is simultaneously available to all clients. This research is for our clients only. No part of this material may be copied or duplicated in any form by any means or redistributed without the prior written consent of William Blair & Company, L.L.C.

THIS IS NOT IN ANY SENSE A SOLICITATION OR OFFER OF THE PURCHASE OR SALE OF SECURITIES. THE FACTUAL STATEMENTS HEREIN HAVE BEEN TAKEN FROM SOURCES WE BELIEVE TO BE RELIABLE, BUT SUCH STATEMENTS ARE MADE WITHOUT ANY REPRESENTATION AS TO ACCURACY OR COMPLETENESS OR OTHERWISE. OPINIONS EXPRESSED ARE OUR OWN UNLESS OTHERWISE STATED. PRICES SHOWN ARE APPROXIMATE.

THIS MATERIAL HAS BEEN APPROVED FOR DISTRIBUTION IN THE UNITED KINGDOM BY WILLIAM BLAIR INTERNATIONAL, LIMITED, REGULATED BY THE FINANCIAL SERVICES AUTHORITY (FSA), AND IS DIRECTED ONLY AT, AND IS ONLY MADE AVAILABLE TO, PERSONS FALLING WITHIN COB 3.5 AND 3.6 OF THE FSA HANDBOOK (BEING "ELIGIBLE COUNTERPARTIES" AND "PROFESSIONAL CLIENTS"). THIS DOCUMENT IS NOT TO BE DISTRIBUTED OR PASSED ON TO ANY "RETAIL CLIENTS." NO PERSONS OTHER THAN PERSONS TO WHOM THIS DOCUMENT IS DIRECTED SHOULD RELY ON IT OR ITS CONTENTS OR USE IT AS THE BASIS TO MAKE AN INVESTMENT DECISION.

Equity Research Directory

Bob Newman, CFA, Principal Manager and Director of Research 312.364.8783
Kyle Harris, CFA, Principal Operations Manager 312.364.8230

BUSINESS SERVICES

Global Services

Brandon Dobell, Principal 312.364.8773
Group Head–Business Services
Educational Services, Real Estate Services

Timothy McHugh, CFA 312.364.8229
Consulting, Employer Services, Staffing

Robert Riggs, CFA 312.364.8610
Business Process Outsourcing, Information Services

Christopher Shutler, CFA 312.364.8197
FinTech and Payment Services

Industrial

Nate Brochmann, CFA 312.364.5385
Commercial Services, Logistics/Transportation

Brian Drab, CFA 312.364.8280
Diversified Industrials, Filtration and Water Management, Industrial Technology

Ryan Merkel, CFA 312.364.8603
Commercial Services, Industrial Distribution

CONSUMER

Sharon Zackfia, CFA, Principal 312.364.5386
Group Head–Consumer
Apparel, Leisure, Restaurants

Jon Andersen, CFA, Principal 312.364.8697
Consumer Products

Meggan Friedman 312.364.8664
Marketing Services

Daniel Hofkin 312.364.8965
Broad Assortment and Hardlines, Health, Beauty, and Convenience

Mark Miller, CFA, Principal 312.364.8498
Broad Assortment and Hardlines, E-commerce, Health, Beauty, and Convenience

FINANCIAL

Adam Klauber, CFA 312.364.8232
Group Head–Financial
Insurance Brokers, Property & Casualty Insurance

CHINA

Liping Cai, CFA +86 21 2327 2260
China-Based Companies

HEALTHCARE

Ben Andrew, Principal 312.364.8828
Group Head–Healthcare
Medical Devices

Ryan Daniels, CFA, Principal 312.364.8418
Healthcare Information Technology, Healthcare Services

John Kreger, Principal 312.364.8597
Distribution, Outsourcing, Pharmacy Benefit Management

Amanda Murphy, CFA 312.364.8951
Diagnostic Services, Life Sciences, Pharmacy Benefit Management

Matthew O'Brien 312.364.8582
Medical Devices

John Sonnier, Principal 312.364.8224
Biotechnology

Brian Weinstein, CFA 312.364.8170
Diagnostic Products

Y. Katherine Xu, Ph.D. 212.237.2758
Biotechnology

TECHNOLOGY

Jason Ader, CFA, principal 617.235.7519
Co-Group Head–Technology
Data Networking and Storage

Laura Lederman, CFA, Principal 312.364.8223
Co-Group Head–Technology
Business Software & Services, IT Services, Software as a Service

Jim Breen, CFA 617.235.7513
Communication Services

Anil Doradla 312.364.8016
Communication Semiconductors, Electronic Components, Wireless Communications

Jonathan Ho 312.364.8276
Cybersecurity, Security Technology

Ralph Schackart III, CFA, Principal 312.364.8753
Digital Media and Internet

Bhavan Suri 312.364.5341
Business Software & Services, IT Services

EDITORIAL

Steve Goldsmith, Head Editor 312.364.8540
Maria Erdmann 312.364.8925
Beth Pekol +44 20 7868 4516
Lisa Zurcher 312.364.8437

William Blair & Company®

CHICAGO BOSTON LONDON NEW YORK SAN FRANCISCO SHANGHAI WILMINGTON ZURICH