Macquarie (USA)
Equities Research

The Equities Research Specialist

UNITED STATES

Inside

Jon Groberg
1 212 231 2612 jon.groberg@macquarie.com

6 April 2010

Life Science Tools

The coming sequencing tsunami

A seismic shift is occurring in genome technology

What is shaking the foundation of genomics is the fact that by the end of this year it will be possible to sequence a complete human genome in about a week for around US$5,000. This will trigger a wave of sequencing information that even a year ago would have seemed impossible. The purpose of this report is to help investors understand the potentially sea-changing implications of these genome technology innovations in order to position themselves appropriately.

This shift will lead to the *real* genomic revolution

A genomic revolution has been predicted for years. However, a revolution can never be driven from the top, rather, it must flow from the bottom, and those engaged need to feel that they both have something significant for which to fight and a chance to win. In our opinion, that inflexion point is near as relevant biological discoveries are being made and the cost per genetic data point is falling at an exponential rate.

Genomics will help “heal, feed, and fuel” the world

Our minds tend to immediately associate genomic information, technologies, and breakthroughs with understanding and improving human health and human genomes. However, demand for genetic information is also being driven by agriculture, energy, and industrial end markets.

The winners and losers might not be who you expect

We think it is highly likely that the raw genetic data itself will fairly quickly become a commodity, meaning that no one will care where the data comes from or how it is generated. More value will ultimately accrue to those who create easy-to-use, relevant applications or those who enable the simple manipulation, interpretation, and adoption of the massive amounts of information.

Don’t be fooled by the secular theme; stay nimble

We estimate that the total sequencing market will grow at a three-year CAGR of ~11%, with most of the growth occurring in 2010/11 driven by the ongoing legacy technology restatement. However, in order to avoid a repeat of the 2001-2006 bust, new technologies will need to emerge. Investors need to be nimble with their investment allocations given the rate of technological change.

LIFE currently offers the best risk reward

In the near-term most genome technology firms will benefit from the wave of genetic information demand that will engulf us; however, there is likely to be a shift back to downstream work once the technology restatement in sequencing has taken place. In this context, we think that Life Technologies currently offers the best risk reward.

The coming sequencing tsunami

The purpose of this report is to help the reader understand the potentially sea-changing implications of genome technology innovations.

When you see the ocean drawback and a large wave gathering on the horizon you feel compelled to warn those exploring the exposed sea bed. The purpose of this report is to help investors understand the potentially sea-changing implications of ongoing genome technology innovations in order to position themselves appropriately as we forecast that the world will soon be inundated with genetic information.

Looking back to understand the future

Diving into the world of genome information and technology is a daunting task, so let me begin with a couple of stories.

England: The Middle Ages

The first took place during the Middle Ages, around the fourteenth century. Back then it was forbidden to translate or read the Bible in English – in fact, it has been reported that the first question ever asked a "heretic" during the Inquisition was whether he knew any part of the Bible in his own tongue[1]. Only the most learned men had access to books (which were hand written) and even fewer had access to, or could read, the Vulgate (the then-accepted Latin Bible). No one living in the 1300s would have ever been able to imagine a world in which people could own and read the Bible in their native tongue.

An Oxford theologian named John Wycliffe, however, decided to attempt to translate the Bible from Latin to Middle English. There was such demand for his work that the translation was painstakingly copied and recopied by hand throughout England (this was, again, before the printing press). After Wycliffe died, his followers, called Lollards, were heavily persecuted and, ultimately, the authorities in Rome ordered that Wycliffe's Bible be destroyed and his remains be dug up in England, burned, and scattered.

Into this environment in 1494 William Tyndale was born. He dedicated his life to translating the Bible into English and more than any other man he is the father of the English Bible translation most in use today. The story is told that in a conversation with a learned authority who opposed his efforts he said, "If God spare my life...I will cause a boy that driveth the plow shall know more of the Scripture than thou dust." Tyndale, too, could not avoid persecution and was ultimately forced to flee from England to Europe where he was finally captured and burned at the stake.

However, nearly 100 years later, in 1611, King James amazingly authorized an English translation of the Bible and with the help of the printing press, in time, nearly every home owned a copy. The impact was profound. Once people were free to interpret the Bible, according to their own understanding, they began to question the authority of their inherited institutions, which led to a political and religious revolution. It was a turning point in history.

Idaho; 1988

The second story is less dramatic, though more personal. It took place in Idaho in 1988. I was waiting for my friend to pick me up to go to a high school dance. I was excited. It was one of my first dates. The plan had been for my friend's date to first pick up both my friend and my date and then for all of them to pick me up before heading to the dance.

I waited anxiously for the group to show up. I paced back and forth in our living room, often checking the driveway. I noticed that it was getting pretty late. I started to get nervous. What if my date didn't show? What if something had happened? What if I had been stood up?! Suddenly the phone rang. It was my friend. He sounded depressed. He told me that he was still at home and that he had gotten a call from his date's mom saying the girls could not make it. There was a long pause on both ends. I was still trying to orient myself when all of a sudden he burst out laughing. He shouted, "I'm just kidding, Jon, we're turning the corner right now!"

[1] *Wide as the Waters,* Benson Bobrick, 2001

It is our assertion that ten years from today words and phrases such as "polymorphism," "copy number," "insertion," "deletion," and "methylation," will be common place.

When he said that, I knew he was toying with me and I was not happy. He was talking to me on the phone, and that, by definition, meant that he, too, had to be on the phone. And phones were not mobile. So there was no way he could be turning around the corner. I was about to lash into him when a car pulled up into our driveway and out jumped my friend. I could not believe it.

How had he called me on the phone? Of course today the answer is obvious given the ubiquity of mobile devices (it turned out that his date's car – her dad's – had a car phone), but back then they were almost unheard of given their expense, size, and form factor.

The lesson to be drawn from both of these stories is that technologies[2] that appear too expensive, clumsy, and complex today will evolve if customers value what it is that they produce. And what appears to be nothing but a dream or science fiction today can become reality tomorrow.

Your town; 2010

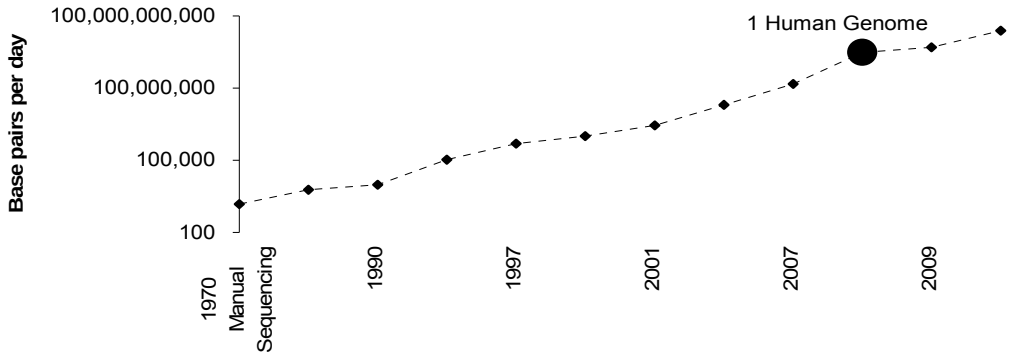
Now, let me ask you the following: When was the last time that you or someone you know went to the doctor and the conversation or the treatment plan took into account your genetic profile? How many people do you know who have had their cancer tumors sequenced to guide their therapy? How much do you know about your own genetic code? When is the last time you used any kind of genetic jargon in your day-to-day conversation? It is our assertion that ten years from today words and phrases such as "polymorphism," "copy number," "insertion," "deletion," and "methylation," will be as common place as "Genesis," "GSM," "Gmail," glucose," etc.

A seismic shift is occurring in genome technology

A seismic shift is occurring in genome technology and we are issuing a tsunami warning for a number of areas that could be impacted. In the near term, this genomic wave is likely to lift all boats, but once the water recedes, we think many firms and business models will be weakened and others will simply be gone.

Some might argue the tremors were first felt in 1953 when James Watson and Francis Crick proposed the spiral staircase structure of DNA, the building block of life. Others might convincingly put forward that the quake intensified in 2001 when the initial working draft of the human genome sequence was first published[3]. However, in our view these were merely foreshocks to the major quake ahead. What is shaking the foundation of genomics is the fact that by the end of this year, it will be possible to sequence a complete human genome in about a week for around US$5,000.

Fig 1 Data generated per day per instrument (Human DNA has ~3bn base pairs)



Source: Industry literature, Macquarie Capital (USA), April 2010

[2] We define *technology* most similarly to Clayton Christensen in *The Innovator's Dilemma*: the processes by which labor, capital, materials, and information are transformed into products and services of greater value.

[3] See http://www.ornl.gov/sci/techresources/ Human_Genome/project/timeline.shtml

Illumina's new instrument is likely to soon be able to produce ~45Gb of sequence data per day; enough to generate all the Pilot 1 data of the 1000 Genomes Project in 6 weeks.

Let me provide one example of how the technology has advanced. A Broad Institute process development manager recently discussed the performance of the Institute's technology at a recent talk at a genome technology conference. According to her, a **single Broad machine**[4] could produce the entire sequence produced in Pilot 1 of the 1000 Genomes Project[5] in 9 months. Furthermore, once Illumina's new instrument, the HiSeq, has been tuned, it should be able to produce 43 billion bases (Gb) per day; enough to generate all the Pilot 1 data in 6 weeks. One machine!

This will enable the *real* genomic revolution

A genomic revolution has been predicted for years. However, a revolution can never be driven from the top, rather, it must flow from the bottom. Those engaged need to feel that they both have something significant for which to fight and a chance to win. In our opinion, that inflexion point is near as relevant biological discoveries are being made and the cost per genetic data point is falling at an exponential rate.

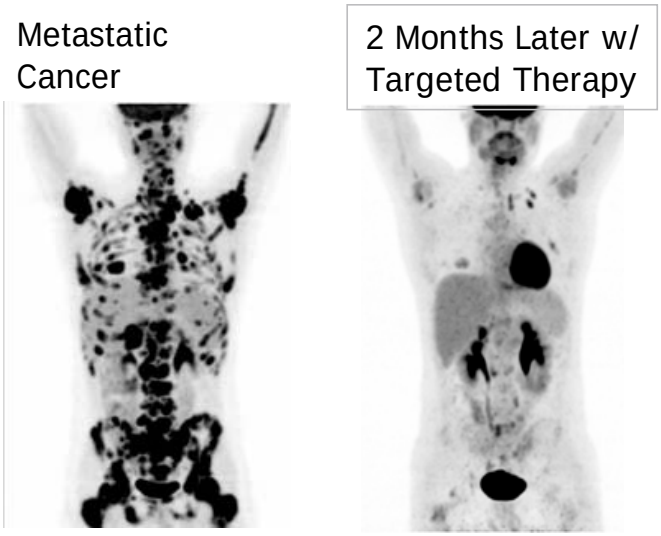
Biologically relevant discoveries are occurring

Let us take cancer as an example. Today, oncologists (physicians specializing in cancer) have limited tools when determining the best treatments for refractory cancers (e.g. cancers that no longer respond to initial treatments). In fact, it often simply comes down to an oncologist's "best guess." Sadly, studies have estimated that these clinician-selected therapies are only successful in ~4% of patients[6].

Matching refractory cancer patients with therapies based on their genetic profile increases the success rate from 4% to 27%.

Dr. Daniel Van Hoff at The Translational Genomics Research Institute (TGen), however, has recently shown that matching patients with therapies based on their genetic profile (e.g. molecular profile) increases this success rate to 27%. Still lower than we all would like, but a large and statistically significant improvement.

Fig 2 Treatment successful in 27% of patients profiled molecularly



Source: Life Technologies and Dr. Van Hoff at TGen, April 2010

[4] The Broad primarily uses Illumina instruments.

[5] Multiple sites with multiple instrument platforms around the world are participating in the 1000 Genomes Project, the goal of which is to generate a nearly complete catalogue of human variation.

[6] TGen and Life Technologies.

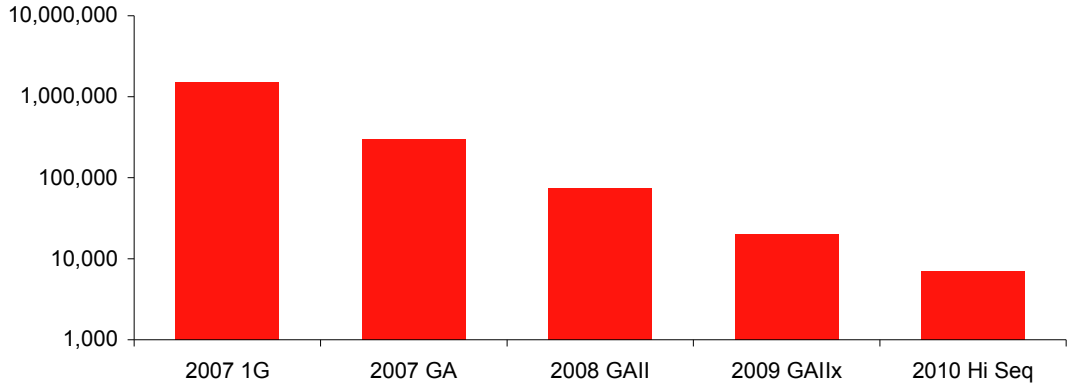
And costs per genetic data point are plummeting

A sequencer costs about the same as it did 4 years ago despite generating 5,000 times more data.

It is estimated that it cost between **US$1-3bn**[7] in order to deliver the first nearly-complete reference human genome (circa 2001). By the end of this year, it will cost **~US$5,000** to sequence a human genome.

The reason the throughput described earlier has led to cost declines is that instrument and reagent manufacturers have been able to keep the prices relatively stable while generating an exponential increase in output; analogous to what happened with computer memory. For example, a high-throughput CE (capillary) sequencer cost ~US$325k in 2005. That is about the same cost as a second generation platform despite the latter's ability to generate 5,000 times more data.

Fig 3 ~100x reduction in cost per whole human genome in last 3 years alone



Note: Analysis uses advancements on ILMN's sequencing platform as representative of cost declines.
Source: Company data, Macquarie Capital (USA), April 2010

Which we believe will ultimately help heal, feed, and fuel the world

Demand for genetic information is also driven by agriculture, energy, and industrial end markets.

Our minds tend to immediately associate genomic information, technologies, and breakthroughs with understanding and improving human health and human genomes. However, demand for genetic information is also driven by agriculture, energy, and industrial end markets.

Human health

We discussed one potential application of genome knowledge in human health earlier: Cancer. But a more precise understanding of how genomes work will also help us better tackle infectious diseases (such as HIV and malaria), inherited diseases, and perhaps even common diseases. It will help us better predict how a pharmaceutical compound will interact with the body and how the body will interact with the drug (i.e. toxicology and ADME[8]). It could lead to better pain management, better diagnostics, more precise treatment, along with a host of other human health applications.

But this demand for relevant genetic information is by no means driven only by our quest for better healthcare. It is also driven by our quest to solve some of humanity's biggest challenges.

[7] Estimates vary widely, according to www.genome.gov the Human Genome Project cost an estimated US$2.7 billion.

[8] Absorption, distribution, metabolism, and excretion.

Agriculture

In order to keep up with the growth in human population, more food will have to be produced worldwide over the next 50 years than has been during the past 10,000 years combined.

It has been estimated that in order to keep up with the growth in human population, more food will have to be produced worldwide over the next 50 years than has been during the past 10,000 years combined[9].

Modern advances in plant genetics and cultivation of plant-associated microbes are allowing rapid improvements to be made in crops, even in those that have seen little past enhancement. For example, a better understanding of molecular plant biology has already allowed researchers to create seeds that are more drought resistant, produce higher yielding crops, or reduce the application of chemical fertilizers and pesticides.

In addition, plant-associated microbe discoveries are being made, such as their ability to improve the availability and uptake of nutrients by plants and promote plant health and growth. The best strains of microbes could be developed as biofertilizers and as disease-control agents, reducing the application of chemical fertilizers and pesticides and enabling effective disease control.

While many individuals and groups have rightfully expressed concern surrounding many of these prospects, technological advances are arguably all that will allow us to continue to feed the increasing human population at a reasonable cost.

Energy

There are many potential applications of genomics in Energy. Let us take biofuels as one example. A better genomic understanding could help us improve the yield of first generation biofuels, such as corn-based ethanol. Of course, second generation biofuels, such as cellulose, are arguably even more desirable as they side-step the fuel versus food debate, and genomic knowledge could allow us to genetically alter enzymes to better break down the cellulose. Third generation biofuels, such as algae, may hold the most promise of all. These could be modified such that they produce more oil and are more stable or resistant to disease.

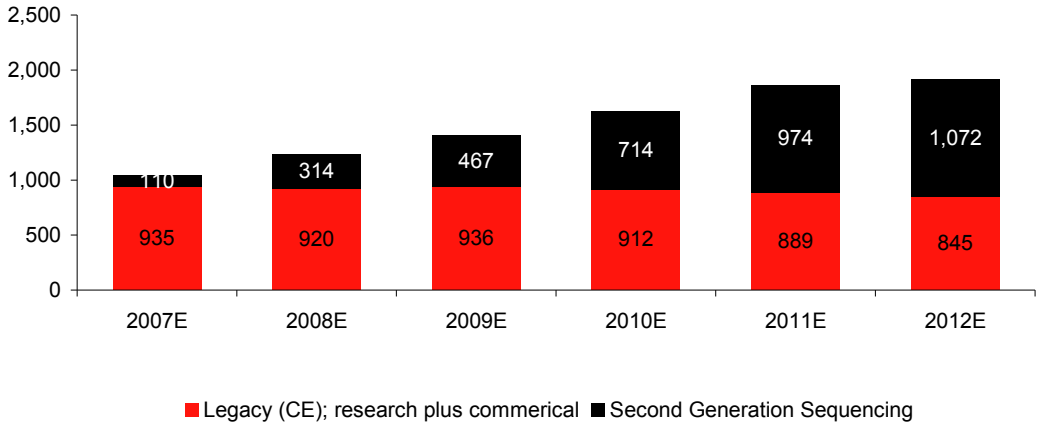
As an example of the potential of genomics in Energy, Exxon Mobil recently signed a US$300m agreement with Craig Venter's Synthetic Genomics (SGI) to develop next generation biofuels using photosynthetic algae. Biocrude is SGI's targeted first product.

Market growth implications

A US$1.1bn second gen market implies ~4,500 installed instruments.

Second generation sequencers should continue to restate the legacy technology market, with most growth occurring in 2010 and 2011 due to ARRA. See Figure 4. A US$1.1bn second gen market implies ~4,500 installed instruments.

Fig 4 Sequencing market forecast: 11% CAGR 2010E-2012E (US$m)

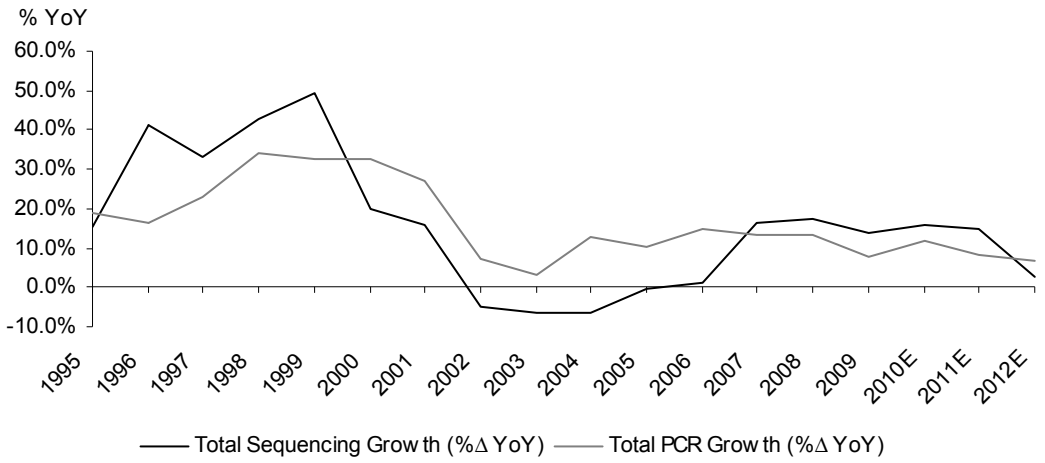


Source: Company data, industry literature, Macquarie Capital (USA), April 2010

[9] Global Policy Forum

In the near term, as the second generation sequencing market grows it will drag many other tools' markets along with it, such as PCR and microarrays (see Figure 5).

Fig 5 The PCR market continued to grow after the last sequencing boom



Note: The PCR market is estimated to be roughly twice as big as the sequencing market.
Source: Company data, industry literature, Macquarie Capital (USA), April 2010

The winners and losers might not be who you expect

The knee-jerk reaction when one considers this report is to think of those companies that are providing the raw genetic horsepower. And they will surely benefit. However, second gen sequencing technologies appear more **sustaining** than disruptive, and we think it is highly likely that the raw genetic data itself will ultimately become un-differentiable, meaning that no one will care where the data comes from or how it is generated (accuracy will be assumed). In our view, more value will ultimately accrue to those who create easy-to-use, relevant applications or those who enable the simple manipulation and interpretation of the massive amounts of information.

Researchers are already intimating that the horsepower game is nearing an end.

In fact, researchers are already intimating that the horsepower game is nearing an end. For example, Eric Green, the director of the NHGRI, recently stated, "Data analysis is the rate-limiting factor in genomics. Not data generation[10]." This is a truly astounding statement to us.

Data analysis is the rate limiting factor simply because there is so much of it. And the amount of data will only keep growing. For example, on the leading platform today, it takes over 600GB to align and map a human genome at 30x coverage. And that is just the genomic information, which without a large database of phenotypic and other medical information, is fairly useless. Thus, the level of data intensity for "precision medicine" will be enormous as the data demands will likely be measured in Terabytes (TB, ~1 trillion bytes) or potentially even Petabytes (PB, TB x 1,000). Depending on how the technology and its associated IT infrastructure demands evolve, this could create huge data management challenges and could be the largest barrier to entry for new laboratories and wider adoption.

In addition, within Human Health medical practice moves at a glacial pace so until clinicians understand why the genomic information is relevant, how to protect it, and why their patients will benefit from it, clinical adoption will lag the research. This creates an opportunity for those in a position to influence physician and clinical decision making, such as PBMs.

[10] *Current Topics in Genome Analysis 2010* lecture series

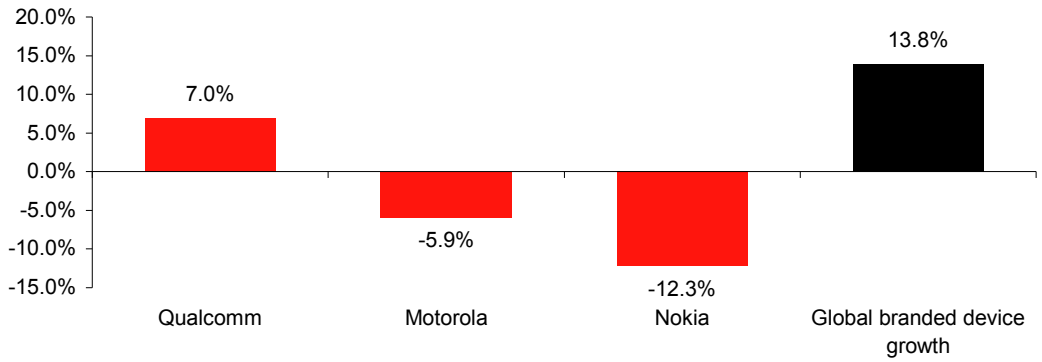
Don't be fooled by the secular theme; stay nimble

All one has to do is look at the popular press to see that genetics is already becoming a big theme. Recently, the top two articles in Bloomberg's Science section were, *Personalized Genetic-Based Medicine Spurred by Medco's Cost-Saving Tests*, and *Family DNA maps May Speed Discoveries of Rare Disease Links, Doctor's Say*. The same day in the New York Times there was an article entitled, *Disease Cause is Pinpointed with Genome*.

So genomics is a big theme, so what?

So genomics is a big theme, so what? As anyone who invested in the major cell phone manufacturers knows, stocks generally anticipate large markets well before they actually emerge. See Figure 6. We think that in the near term most genome technology firms will benefit from the wave of genetic information that will engulf us; however, we think investors need to be nimble with their investment allocations given the rate of technological change.

Fig 6 Stock performance versus growth in branded cell phones (CAGR 2001-2009)



Source: FactSet, Macquarie Capital (USA), April 2010

In our view there is one key principle and one key unanswered question that will go a long way in determining the fate of many of these companies.

In our view there is one key principle and one key as-of-yet unanswered question that will go a long way in determining the fate of many genomic tools companies. The principle is this: Customers care about the genetic information, not the technology. In as much as companies confuse the two, they are at risk of becoming obsolete. The key question is: In what applications will the information of the entire genomic sequence be needed versus only targeted regions of interest and what is the best way to access and deliver that information?

Scope of this report

Many reports have been written describing in painful detail the science and technology of next generation genome sequencing platforms. That is not the purpose of this report.

We do provide a sequencing primer and discuss existing technologies and technical specifications and trends at the back of this report; however, its purpose is to help the reader understand the potentially sea-changing implications of these innovations rather than the nitty-gritty detail of the science itself.

In addition, while this report is focused on genome sequencing technologies, investors should recognize that any technology involved in genomic understanding and analysis is likely to be affected. As the cost-per-genetic data point falls and the technologies become easier to use, we expect the genome technology landscape to be permanently altered over time. Thus, we discuss how the sequencing market might impact other genomic tools markets as well.

The building blocks of life

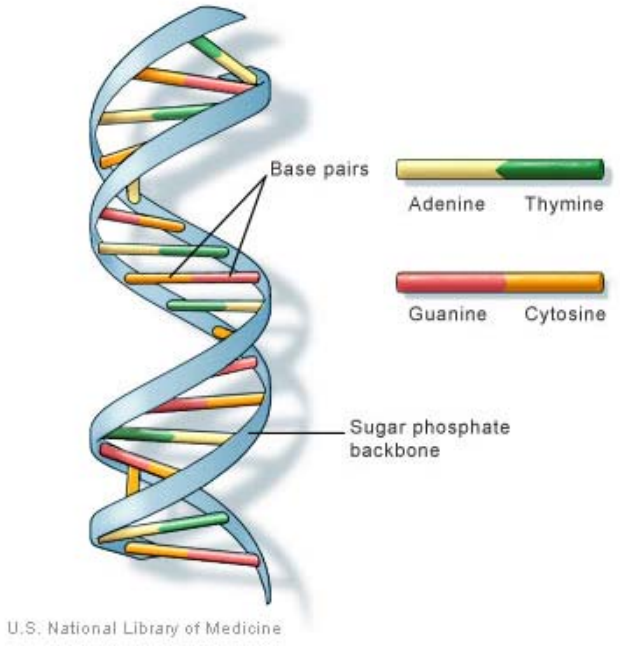
Nucleic acid is the fundamental unit of life, and, therefore, understanding its structure and how it regulates cells is critical.

Before we dive into our analysis of the market, we need to briefly establish what nucleic acid sequencing is and why it is the foundation for furthering our genetic understanding. We then argue why there is, indeed, demand for this knowledge and, by extension, the relevant tools.

What is DNA and DNA sequencing?

DNA is found in the nucleus of each cell and has been called the building block of life. The building blocks of DNA (a molecule that looks a lot like a spiral staircase) are called nucleotides or bases. Each of the four types of nucleotides is typically represented by the letters A, T, C or G. Thus, a DNA sequence would be represented as TATCAGC or some other series of these four letters. DNA is also double stranded, meaning that each base has its complementary pair (A goes with T and G goes with C) such that the above sequence's compliment would read ATAGTCG. The haploid human genome (humans are diploidy, or have two copies of DNA) is made up of approximately 3 **billion** base pairs. *H. Influenzae,* on the other hand, has 1.8 **million** base pairs.

Fig 7 DNA is a double stranded molecule linked together by base pairs



Source: US National Library of Medicine, April 2010

Nucleic acid sequencing is the process of identifying the specific order of nucleotides in a given strand of DNA – or RNA, a single stranded complement to DNA and used to carry out DNA's instructions. Nucleic acid is critical in dictating how cells function, and, therefore, understanding its structure and how it regulates cells has become a key focus for researchers.

For example, a specific sequence of DNA (called a gene) leads to a specific sequence of RNA (messenger RNA, or mRNA), which leads to the creation of a specific protein. Proteins (e.g. enzymes, hormones, antibodies etc.) are the chief actors within a cell.

Technology advances are dramatically reducing the cost

Knowing the precise order of DNA (and RNA) nucleotides, or sequencing, has long been considered a key to advancing our understanding of life, as well as disease and death. However, it has remained prohibitively expensive. For example, it is estimated that it costs between US$1-3bn in order to deliver the first nearly-complete reference human genome (The Human Genome Project; largely completed in 2003). Today, that equation is changing.

From Sanger to SBS; nearing the US$5,000 genome

Sequencing output is generally measured in terms of the number of base-pairs generated (remember, if you know one base you automatically know its complement). In the early 1970s researchers used a manual process that was incredibly labor intensive – and hazardous – to generate, at best, around 1,500 bases per day. Then in 1977 Frederick Sanger developed a radical new protocol that ultimately led to the development of sequencing machines.

The first commercial DNA sequencer was launched in 1985 by Applied Biosystems (now a part of Life Technologies) and could generate ~6,000 bases a day. In 1995 AB released its first capillary electrophoresis (CE) sequencer that could generate 5,000-15,000 bases per day. By 1998 AB's, and other's, instruments were generating 500k-1 million bases per day.

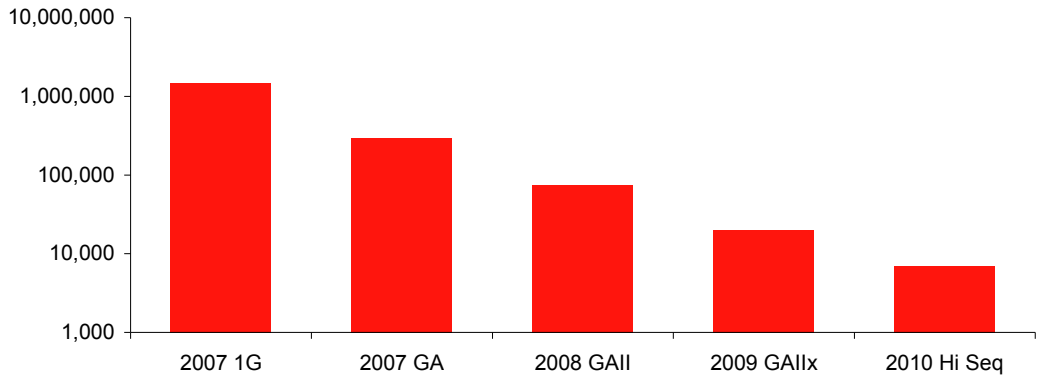
In 2005, 454 Life Sciences (now a part of Roche) introduced the era of second, or "next," generation sequencing by introducing a sequencing system that could generate ~20 million bases per day (or 20Mb). In 2007 new "short read" next generation sequencers were generating ~150Mb per day. By the beginning of this year (2010) platforms such as Illumina's Genome Analyzer (based on Sequencing by Synthesis, SBS, technology) were already generating ~5 billion bases (5Gb) per day, and as has already been mentioned are on a roadmap to be generating 25Gb, or more, per day by the end of the year.

Fig 8 Milestones in DNA sequencing technology

	Output (base pairs per day)
1970's Manual sequencing using radioactive isotopes for tagging DNA	1,500
1985 First Automated Sequencer ABI model 370A DNA Sequencer	6,000
1990 ABI PRISM Model 373 DNA Sequencer	9,600
1995 First Capillary Electropheresis Sequencer (ABI PRISM 310)	15,000
1995 ABI Prism 377 DNA Sequencer	105,000
1997 MegaBACE 1000	500,000
1998 PE Biosystems Prism 3700	1,000,000
2001 MegaBACE 4000	2,800,000
2005 First 'Next Gen Sequencer" GS20 454 Life Sciences	20,000,000
2007 Illumina Genome Analyzer	150,000,000
2009 Genome Analyzer IIx and SOLiD 3	5,000,000,000
2010E HiSeq 2000 and SOLiD 4	25,000,000,000

Source: Company data, Macquarie Capital (USA), April 2010

Fig 9 ~100x reduction in cost per whole human genome in last 3 years alone



Note: Analysis uses advancements on ILMN's sequencing platform as representative of cost declines.
Source: Company data, Macquarie Capital (USA) April 2010

How second generation sequencing works: The 'Cliffs Notes'

With first generation sequencing technologies, the way throughput was radically increased, once capillaries were introduced, was by simply adding more and more capillary lanes or more and more instruments, along with better automation. Genome centers had rooms filled with AB 3730s, and other platforms, running around the clock.

With second generation sequencing the opposite became true, sort of. While each second generation technology differs in details, the principles are broadly the same:

1. Because millions of sequencing reactions take place at the same time, the process is often called massively parallel.
2. The basic technological premise that drives current next generation sequencing platforms is 1) millions of copies of DNA strands must be made in order for existing optic technologies to be able to see and read the signal from the dyes that represent the individual bases as they bind to, and extend, the target DNA; and 2) it is possible to shrink the distance between these clusters of reactions in order to allow more of them to occur at the same time.
3. The second generation sequencing process involves a number of steps.

At some point in the process the signal fades and the optics can no longer accurately see or distinguish between the dyes as there are not enough copies of DNA.

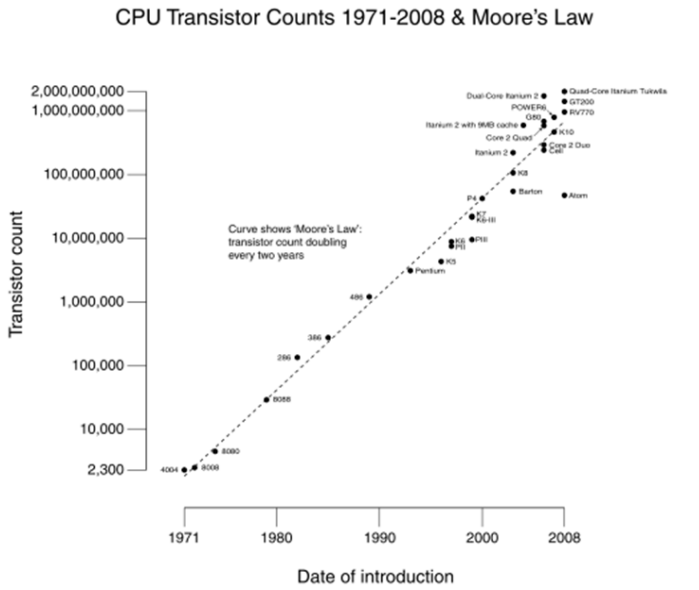
- ⇒ Before sequencing can begin, the sample DNA must be prepared. This is called sample preparation, or sample prep for short. The target DNA (also called template DNA) is first randomly fragmented (or sheared) into millions of fragments. This is because the longer the template, the harder it is to work with, generally speaking.
- ⇒ Once the fragments are made and cleaned up, various adaptors are added. These adaptors allow the template strands of DNA to bind to an anchored complementary template and to extend (e.g. grow the complementary strand of DNA) when prompted.
- ⇒ The template DNA is then amplified (e.g. millions of copies are made). These copies are typically on slides, or flow cells, of some sort.
- ⇒ Once the slide is prepared, it is put into the sequencing instrument. The instrument then goes through a number of cycles in which bases are added, pictures are taken, dyes are removed, and the process starts all over again. Because this is happening at millions of points on the slide at the same time the process is called, again, "massively parallel."
- ⇒ It is important to note that usable template DNA is lost at every step in the sample preparation process and also during each instrument cycle. This is because at each step chemistries inevitably fail (DNA itself is not an easy molecule to work with). This means that at some point in the process the signal fades and the optics can no longer accurately see or distinguish between the dyes as there are not enough copies of DNA. The accuracy (of the raw read) falls off and the read length has reached its limit.
- ⇒ Because biases can be introduced at multiple steps throughout the process (for example, when the template strands are copied), read-lengths are short (50-100 base pairs), and bases are inevitably called incorrectly, the process is typically repeated multiple times in order to be able to make consensus determinations.
- ⇒ The user must then figure out how to interpret the data.

Moore's law on steroids; output doubling every 0.7 years, on average

As many of those involved with sequencing technologies like to point out, there are clearly similarities between current sequencing technology advancements and the early days of the computer industry – when Moore crushed Grosch. Grosch's Law stated that a computer's performance increased as the square of the cost (e.g. larger computers were more economic, so the larger the mainframe the better). The microprocessor and Moore's Law (Figure 10) turned the computer industry on its head by instead predicting that transistor density would double every two years making the price-performance of smaller servers and PCs ultimately superior.

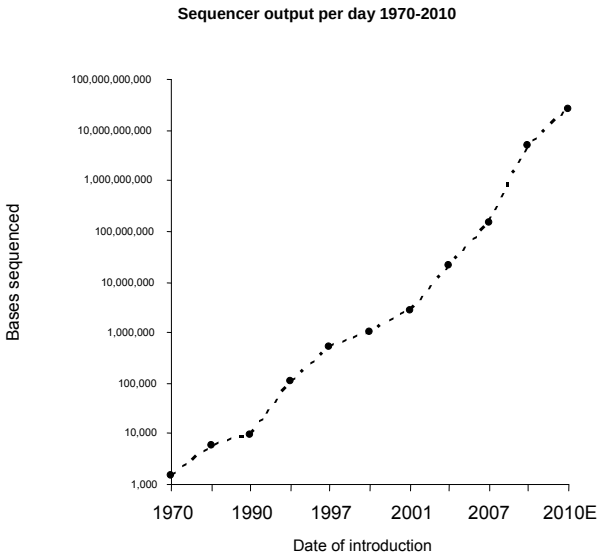
In sequencing, however, the amount of Gb per day has doubled, on average, every 0.7 years since the introduction of the first CE instrument in 1995 (Figure 11).

Fig 10 Moore's law in pictures



Source: Google Annual Report 2008, April 2010

Fig 11 Moore's law on steroids: DNA output



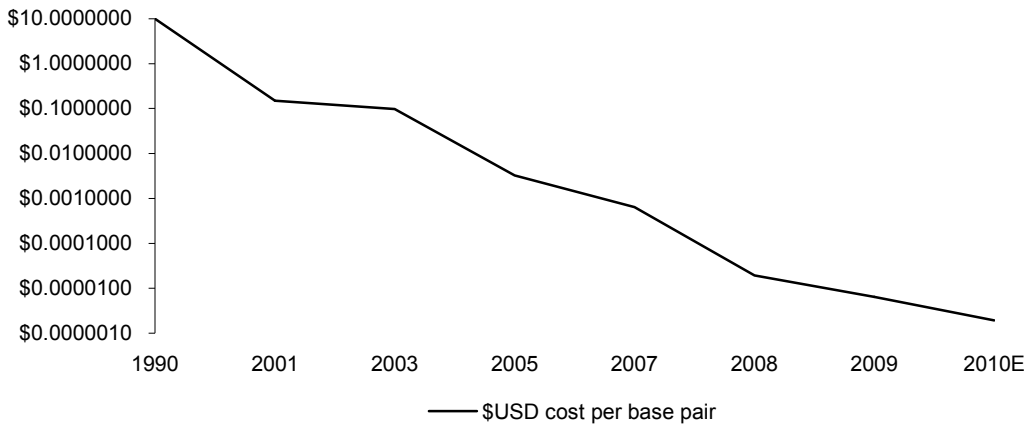
Source: Company data, Macquarie Capital (USA), April 2010

A high-throughput CE sequencer cost ~US$325k in 2005. That is about the same cost as a next generation platform today which generates 5,000 times more data.

More bases for the same buck

The reason these throughput advances have led to dramatic cost declines is similar to another phenomenon in the computer industry: prices remained stable or even declined (similar to memory). In sequencing, instrument and reagent manufacturers have been able to keep prices relatively stable while generating an exponential increase in output. For example, a high-throughput CE (capillary) sequencer cost ~US$325k in 2005. That is about the same cost as a next generation platform today; however, today's instrument generates 5,000 times more data.

Fig 12 The cost per called base is 10^{-7} what it was in 1990



Source: Industry literature, Macquarie Capital (USA), April 2010

Let me provide a vivid example of how these new technologies have radically advanced throughput. According to a Broad Institute process development manager, 60% of the Broad's Illumina instruments now run the 2X100bp protocol (this means the instrument is reading both ends of a strand of DNA, 100 base pairs in each direction) and produce over 5Gb per day per machine. Other machines are actually running a 2X150bp protocol along with a higher cluster density, which produces 7Gb/day. For the latter, **a single machine** could produce 30X coverage of a genome (i.e. sequence the same genome 30 times to improve accuracy) in a single 2-week run, or the entire sequence produced in Pilot 1 of the 1000 Genomes Project in 9 months. Once Illumina's new instrument, the HiSeq, has the same cluster density as existing platforms, it should be able to produce 43Gb per day; enough to generate all the Pilot 1 data in six weeks.

Both restating the old market and opening some that are new

In some ways second (or "next") generation sequencing instruments are disruptive to existing platforms and in other ways they could be considered important sustaining technologies[11], a debate we will tackle in more detail later on. They are disruptive in that they do alter the nucleic acid sequencing game as they allow researchers to do things that they simply could not do before due to time and cost constraints. They are sustaining in that they primarily satisfy existing customer's demands for more sequence, as fast as possible, at the lowest price.

To provide an example of how second generation technologies have both disruptive and sustaining elements, consider the following:

4. Applied Biosystems *3730 DNA Analyzer* was the gold standard in high-throughput DNA sequencing technology. We estimate that ~3,000 of these machines were placed in genetic research labs around the world and that AB had at least two-thirds of the market. Re-sequencing a human genome using a single 3730xl would take over three years. To lower these time constraints a researcher would need multiple instruments and complex sample preparation capabilities, making it cost and resource prohibitive for a single researcher to compete against larger genome centers.

In many ways these new instruments do democratize the genetic research playing field.

5. Second generation technologies now allow this same research to be done in a matter of weeks or days on one instrument, allowing smaller research laboratories to get in on the game, if they can afford an instrument. As these technologies advance, this re-sequencing will, in all likelihood, ultimately be done at a cost of ~US$1,000. So in many ways these new instruments do democratize the genetic research playing field.

However, the technology remains primarily research- and lab-based.

6. However, while smaller research labs can now ostensibly compete against the larger labs, the up-front instrument cost (Illumina's latest product offering is listed at US$690k, for example) and the sample prep and informatics requirements mean that the technology remains prohibitive to many and primarily research- and lab-based. In addition, large genome centers have simply scaled up on the new platforms effectively widening the gap between smaller labs.

Until the cost, ease of use, read-length, footprint, informatics and other barriers are removed, we think second generation sequencing will struggle to move outside of existing high-throughput markets.

[11] Here we again use terms from *The Innovator's Dilemma.*

And allowing us to answer some fundamental questions

We truly are at the extreme tip of the proverbial iceberg of genetic understanding.

Despite the significant discoveries that have occurred in genomics over the years, our knowledge of DNA, and its many cellular corollaries, remains surprisingly sparse. We truly are at the extreme tip of the proverbial iceberg of genetic understanding. However, the rate at which the cost of sequencing has fallen over the last year alone makes us think that we are nearer to an inflection point in genomic knowledge than many realize.

What are some of the fundamental questions that we hope sequencing can help us answer? There are too many to list, but examples include:

7. Interpreting the human genome: What parts of it are functional and which parts are not functional – it is estimated that ~5% is functional, but is that right? What part of the genome sequences for proteins? Again, it is estimated that there are ~20K genes, but could we be wrong? What does the non-coding functional sequence do? Gene regulation, chromosomal packaging, segregation, and replication? Non-coding RNA (e.g. micro RNA)? Others we don't know about? Do we have a poor inventory of genetic information generally?
8. Comparative sequencing (inter-species). What can we learn about how our DNA evolved?
9. Implicating genetic variants in human disease: 1,000 Genomes project, Cancer Genome Atlas, Human Microbiome project.
    - ⇒ The 1,000 Genomes Project: Launched to uncover genetic differences of at least 1% frequency across 2000 unidentified people from about 20 distinct populations globally. Sequenced sample data from the main study of 2000 genomes has been released intermittently between 2009 and 2010 and is projected to conclude in 2011.
    - ⇒ Cancer Genome Atlas: The National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) launched The Cancer Genome Atlas (TCGA) program to create a comprehensive atlas of the genomic changes involved in more than 20 common types of cancer. This large-scale, high-throughput effort is being carried out by a network of more than 100 researchers at many organizations across the United States.
    - ⇒ Human Microbiome project: Within the body of a healthy adult, microbial cells are estimated to outnumber human cells by a factor of ten to one. These microbial communities, however, remain largely unstudied, meaning we know almost nothing about their influence upon human development, physiology, immunity, and nutrition. The NIH Roadmap has initiated the Human Microbiome Project (HMP) with the mission of "generating resources enabling comprehensive characterization of the human microbiota and analysis of its role in human health and disease."
10. Functional research: What genes code for converting sunlight into fuel?

Demand drivers of genetic information

What we strive to convey in this section of the report is simply that there is tremendous demand for better genetic information and knowledge.

Our minds tend to immediately associate genomic information, technologies, and breakthroughs with understanding and improving human health and human genomes. However, demand for genetic information is also driven by agriculture, energy, and industrial end markets.

Attempting to detail each and every potential genomic demand driver is not practical. The point we hope to drive home in this section of the report is simply that there is tremendous demand for better genetic information and knowledge and since our understanding of the genome and its associated pathways remains nascent, knowledge of the complete genetic blueprint is critical.

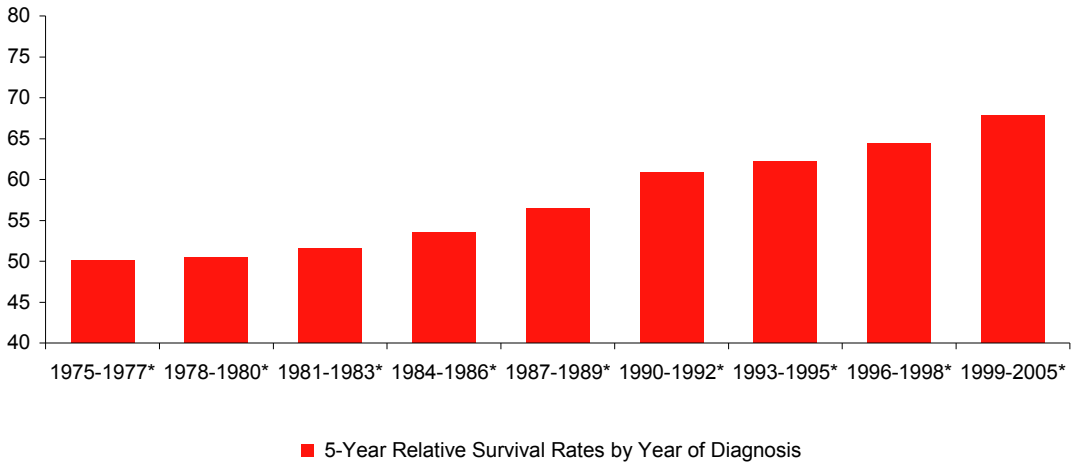
Human Health

It is estimated that genetic factors play at least some part in ~90% of the deaths that occur in the world. Here we discuss only a few of many potential examples where better genetic information could improve human health. These examples are, of course, representative and not exhaustive.

Cancer, the new chronic disease?

The stark reality is that today approximately 50% of us who live in the United States will get cancer and ~50% of those who do will die[12] . Yet, many experts believe that a better understanding of the genomic complexities of cancer will lead to discoveries that will soon (within our lifetime) make cancer a chronic and manageable disease, similar to diabetes.

Fig 13 Trends in 5-year survival rates post diagnosis of Cancer



Source: National Cancer Institute, Macquarie Capital (USA), April 2010

The first melanoma genome sequenced found 33,345 single base mutations, 1,018 small indels, 37 large structural changes, and 198 copy number changes.

What makes cancer particularly exciting to those in the genomics space is that cancer is, in fact, a disease of the genome and, perhaps even more interesting, researchers are finding that every cancer is a **new** genome. This genomic complexity is demanding a genome-wide approach toward its understanding, which means that rather than simply look at one part of the genomic pathway, researchers need to understand the entire genome[13].

11. As an example of the complexities associated with cancer genomes, the first melanoma genome sequenced found 33,345 single base mutations, 1,018 small indels, 37 large structural changes, and 198 copy number changes[14].

[12] American Cancer Society: Lifetime probability of developing cancer for women is 1:3 and Men 1:2 in the US, cancer accounted for 23.1% of all US deaths in 2006

[13] Note that not everyone agrees that a genome-wide approach is warranted.

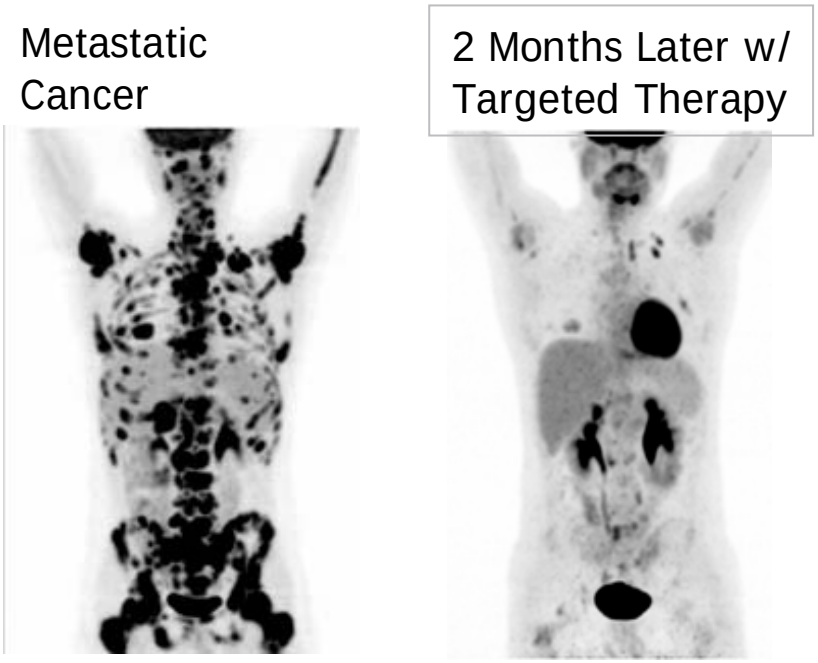
[14] From Pleasance, Cheetham et al. 2009

Generally speaking, oncologists (physicians specializing in cancer) have limited tools when determining best treatments for cancer patients, but this is especially true in refractory cancers (e.g. cancers that no longer respond to initial treatments). In fact, prescribed treatments often simply come down to an oncologist's "best guess." Sadly, studies have estimated that these clinician-selected therapies are only successful in ~4% of patients[15].

Matching refractory cancer patients with therapies based on their genetic profile increases the success rate from 4% to 27%.

Dr. Daniel Van Hoff at The Translational Genomics Research Institute (TGen), however, has recently shown that matching patients with therapies based on their genetic profile (e.g. molecular profile) increases this success rate to 27%. Still lower than we all would like, but a large and statistically significant improvement, nonetheless, and one which gives us hope that further discoveries will continue to radically improve the success rate.

Fig 14 Treatment successful in 27% of patients profiled molecularly



Source: Life Technologies and Dr. Van Hoff at TGen, April 2010

One area where we think radical change will be needed in order to truly move forward with the nation's goal of defeating cancer is in clinical trials.

One area where we think radical change will be needed in order to truly move forward with the nation's goal of defeating cancer is in clinical trials. Researchers are discovering that cancers are best defined by their mutations, or molecular pathways, rather than by their physical locations. However, the way clinical trials are currently required to be designed makes the application of this knowledge problematic.

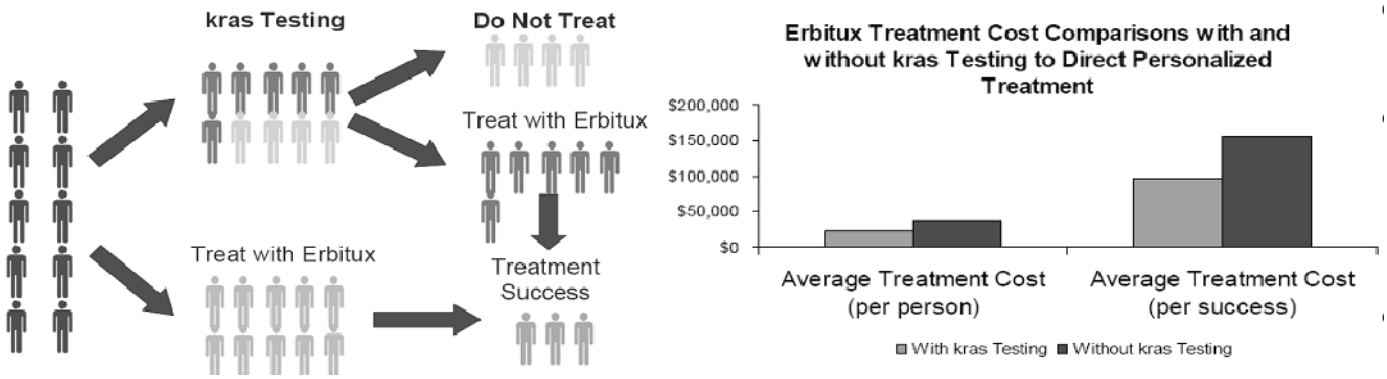
Today, if a medical researcher has a lung cancer patient that has mutation *X* and wants to enroll them in a clinical trial for a pancreatic cancer drug that has been designed to treat mutation *X*, they cannot. The researcher would need to instead start up a new trial for lung cancer designed to treat the specific mutation. There are a couple of problems with this requirement: 1) the patient is likely to be dead by the time the trial is designed and 2) the patient could have a very rare mutation for a lung cancer making it impossible to fill the trial.

Another key to the successful adoption of genetic profiling in cancer is, of course, the cost. Here we expect the science to help. The NIH estimated that US$89bn was spent on cancer care in 2007, or roughly ~5% of all health care spending in the United States, up from US$41bn in 1995. However, evidence suggests that this number could be reduced, or at least its growth rate could be slowed, if patients were to receive the appropriate therapy the first time.

[15] TGen and Life Technologies

For example, a study published in the New England Journal of Medicine in 2008 concluded that advanced colorectal cancer patients that had a certain genetic mutation (the KRAS- mutation) did not benefit from a cutting edge cancer drug called cetuximab (Erbitux) [16]. Another study estimated that knowing whether patients had this mutation would save the US health care system US$604m annually, or, said differently, a genetic test priced at around US$400 could save ~US$61,000 in treatment costs per patient[17]. See Figure 15.

Fig 15 Patients with KRAS mutations will not benefit from Erbitux



Source: Luminex; Macquarie Capital (USA), April 2010

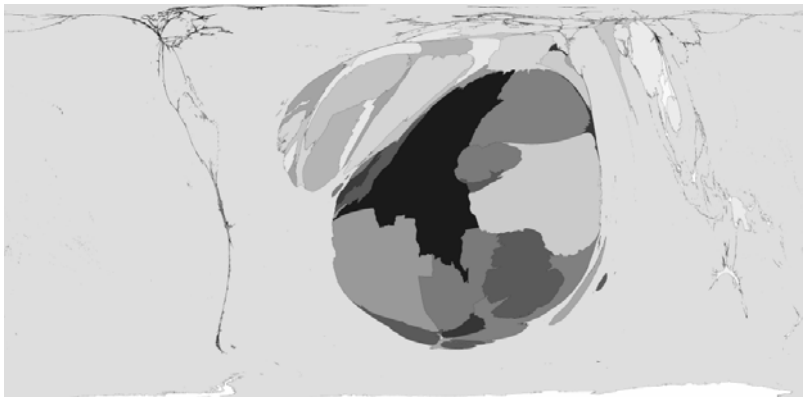
Richard P. Lifton, chairman of the genetics department at Yale University School of Medicine in New Haven, Connecticut, recently wrote in an editorial in the New England Journal of Medicine that "plummeting costs mean DNA sequencing will soon become routine in medical diagnoses, shedding light on the genetic characteristics of [certain] diseases." We think cancer is a natural first disease category.

Malaria (as a proxy for infectious diseases)

Every year, more than a million people – primarily children in Africa – die of malaria.

Another example where better genetic information could improve human health is dealing with malaria. According to the World Health Organization (WHO) a child dies every 30 seconds of malaria while every year more than a million people die of the disease and hundreds of millions more are infected but survive. There is a cruel irony in these statistics as several decades ago it was thought that malaria had been, for all intents and purposes, eradicated from the world. Ever since then, unfortunately, cases have been on the rise.

Fig 16 The majority of malaria deaths are concentrated in Africa and SE



Note: Territory size shows the proportional burden of mortality. ~1 million deaths are reported each year.

Source: Worldmapper.org, Copyright 2009 SASI Group (University of Sheffield), April 2010

[16] New England Journal of Medicine, October 23, 2008, Vol. 359: 1757-1765, "K-ras Mutations and Benefit from Cetuximab in Advanced Colorectal Cancer"

[17] American Society of Clinical Oncology, Provisional Clinical Opinion January 13, 2009

This is because malaria, we have found, evolves. And fast. Today, resistance to chloroquine, the cheapest and most widely used drug, is widespread and resistance to the more powerful combination treatment of sulfadoxine-pyrimethamine has spread from South America and Southeast Asia to Wast Africa. There is also evidence that resistance to the newest drug, artemisinin, could be emerging. Unfortunately, there are currently no other drugs available.

Sequencing the genomes of 16 parasites from different parts of the world helped researchers see what regions of the parasite's genome had evolved and where it could be more vulnerable.

Malaria evolves so quickly because every time the parasite is transmitted (malaria is transmitted via mosquito bites), it essentially undergoes a mating event. This means that its evolutionary time-table mimics the behaviour of viruses more than that of other single-cell organisms.

One way to begin to understand malaria is to understand exactly how it has evolved over time. A team at the Broad created a comprehensive map of the genetic diversity of *P. falciparum,* one of the common parasites in Africa, by sequencing the genomes of 16 parasites from different parts of the world. This helped them to see what regions of the parasite's genome had evolved and where the parasite could, potentially, be more vulnerable[18].

Of course, to truly understand malaria we also need to better understand the genomic evolution of humans and mosquitoes. For example, mosquitoes have evolved resistance to insecticides, which has made them more abundant, which increases the rate at which the malaria parasite is transmitted, which increases its rate of evolution, which increases the rate at which malaria develops resistance to existing drugs.

The genetic research to date has led to some potentially powerful ideas. The hope, of course, is that further genomic discoveries and a better applied understanding could lead to vaccines (there has been some recent exciting news on this front from a team at the University of Maryland), earlier prediction of drug resistance in order to contain it, or newer and more effective anti-malarial drugs and/or insecticides.

Precision (often called personalized) medicine

With $770bn in global drug sales in 2008, that means that ~US$300bn was potentially spent on ineffective medicines globally.

We touched on this topic in our discussion on cancer, but it is bigger than one disease. It is estimated that, on average, drugs generally only work in 30–50% of the people to whom they are given (see Figure 17). With US$770bn in global drug sales in 2008, that means that ~$300bn was potentially spent on ineffective medicines globally.

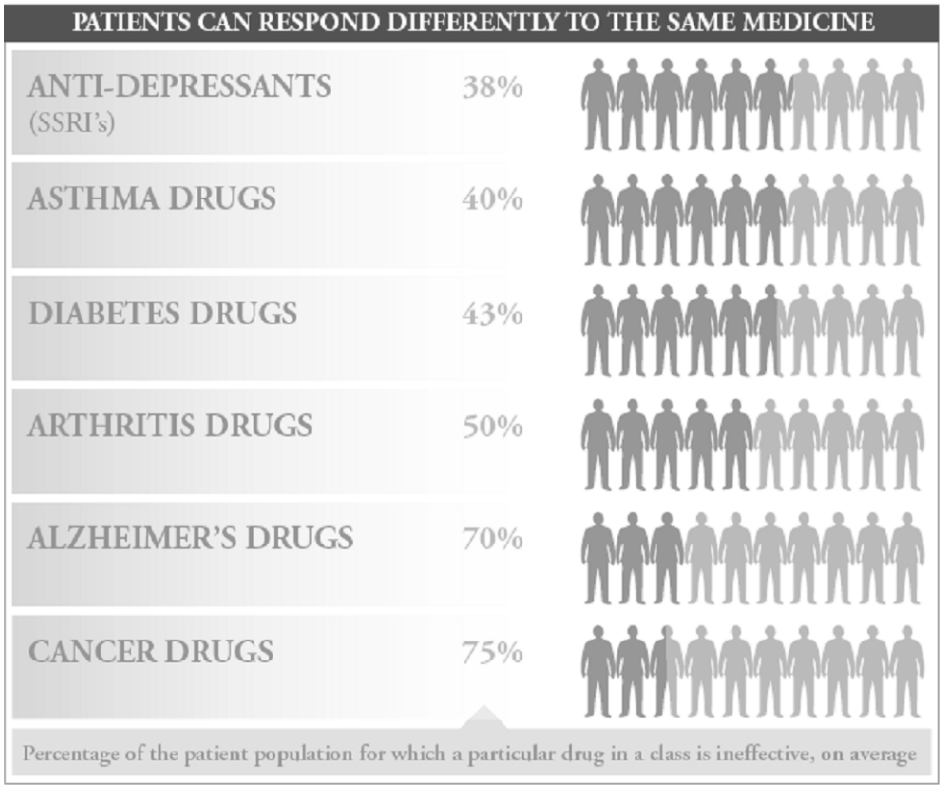
Some argue that genetic differences between people are one reason almost half of the US$292 billion spent on prescription drugs in the U.S. in 2008 likely went to medications that did not help patients[19].

Thus, there is growing enthusiasm that new genetic discoveries will lead to new companion diagnostics, identifying those patients who are more likely to respond to a given therapy. This could herald a new era of precision medicine (in the past this has often been called personalized medicine, but many doctors rightfully, in our view, argue that they are already practicing "personal" medicine).

[18] Volkman et al., (2006)*; Nature Genetics*

[19] McKinsey & Company at Harvard Medical School, as quoted in Bloomberg

Fig 17 Drugs are ineffective in 30–50% of patients, on average



Source: Personalized Medicine Coalition, April 2010

The annual market for diagnostic tests and drugs tailored to individuals already totals ~US$24 billion.

While we often think of precision medicine as being a future trend, according to the American Society of Human Genetics there are currently more than 50 companies in the genetic testing business, including lab equipment companies. In fact, the annual market for diagnostic tests and drugs tailored to individuals already totals ~US$24 billion, according to a report last October from PricewaterhouseCoopers LLP, and the market is likely to grow 10 percent annually, reaching US$42 billion by 2015. Figure 18 provides examples of existing precision therapies.

Fig 18 Selected examples of personalized medicine in action

Therapy	Biomarker/Test	Indication
Herceptin	Her-2/neu receptor	Breast cancer: "…for the treatment of patients with metastatic breast cancer whose tumors over-express the HER2 protein and who have received one or more chemotherapy regimens for their metastatic disease."
Coumadin	CYP2C9 and VKORC1	Cardiovascular disease: Determines CYP2C9 and VKORC1 genotypes to predict likelihood of adverse events with warfarin therapy.
Camptosar	UGTIA1	Colon cancer: "Variations in the UGT1A1 gene can influence a patient's ability to break down irinotecan, which can lead to increased blood levels of the drug and a higher risk of side effects."
Immunosuppressive drugs	AlloMap® gene profile	Heart transplantation: Monitors patient's immune response to heart transplant to guide immunosuppressive therapy.
Ziagen	HLA-B*5701	HIV: "Patients who carry the HLA-B*5701 allele are at high risk for experiencing a hypersensitivity reaction to abacavir. Prior to initiating therapy with abacavir, screening for the HLA-B*5701 allele is recommended."
Gleevec	BCR-ABL	Leukemia: "Gleevec® (imatinib mesylate) is indicated for the treatment of newly diagnosed adult and pediatric patients with Philadelphia chromosome positive [indicated by presence of BCR-ABL] chronic myeloid leukemia (CML) in chronic phase."

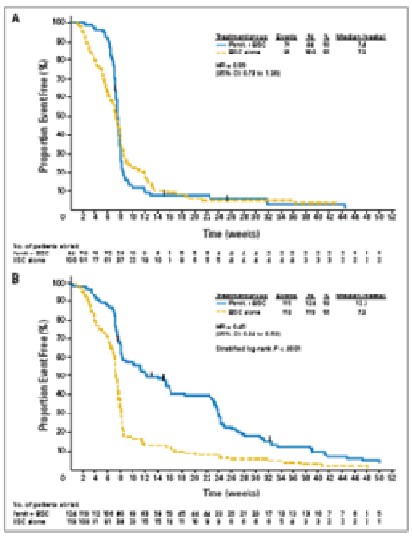
Source: Industry literature, Macquarie Capital (USA), April 2010

The era of precision medicine arguably began with Herceptin, a breast cancer drug from Genentech (now Roche AG) that was found to only work in patients whose tumors overproduced a protein called Her2/neu. In approving the drug in 1998, the U.S. FDA required the drug's label to say that it should only be used in Her2 positive patients.

Vectibix and Erbitux were recently found to not work in the 40 percent of colon cancer patients with a K-RAS mutation.

The trend has since continued, if at a plodding pace. For example, in July 2009, the FDA, after reviewing the evidence that Vectibix (Amgen) and Erbitux (BMS/Merck KGaA) did not work in the 40 percent of colon cancer patients with a K-RAS mutation (see Figure 19) changed both of these drugs' labels as well. K-RAS is the gene associated with the Epidermal Growth Factor receptor.

Fig 19 Vectibux and Erbitux showed no benefit in KRAS mutated patients



Source: New England Journal of Medicine, April 2010

Qiagen projects that the market for K-RAS tests will climb from US$25-$30 million today to US$100 million annually within five years driven by its compelling value proposition (according to press reports Vectibix costs around ~US$8,400 a month). As mentioned earlier, it has been estimated that these tests, which cost ~US$400, would save ~US$604 million a year in drug costs in the U.S.

Recent Warfarin study offers further support

Another example is with the widely prescribed blood thinner Warfarin (also marketed as Coumadin and Jantoven). Determining the appropriate dose of Warfarin can be challenging given the variable drug response between individuals and the fact that serious complications can arise from either doses that are too high – bleeding – or too low – clotting.

At a recent ACC meeting, results were published from a warfarin trial. Those in the genetic testing group were 28 percent less likely to be hospitalized for bleeding or thromboembolism, and 31 percent less likely to be hospitalized for any reason.

At the recent American College of Cardiology meeting (ACC), researchers from Medco and the Mayo Clinic reported on a comparative-effectiveness study that looked at the utility of genetic testing for improving the safety and effectiveness of Warfarin. The research was funded by Medco and the Mayo Clinic's Center for Individualized Medicine. (Note that the recent passing of Healthcare reform in the US could increase the number and size of these types of comparative effectiveness studies.)

The study began in July 2007 and enrolled patients from 49 states who were insured by dozens of health plan sponsors managed by Medco. Patient ages ranged from 40 to 75; the average age was 65. Around 60 percent of participants in the study and control groups were men.

The team compared 896 individuals who received genetic testing for CYP2C9 and VKORC1 genes early in their warfarin treatment with 2,688 control individuals, selected from the same group of health insurance sponsors the previous year, who had received treatment without genetic testing. Genetic testing for the study was performed at the Mayo Clinic, which also gave doctors guidelines for applying genetic information to drug dosing and management.

Those in the genetic testing group were 28% less likely to be hospitalized for bleeding or thromboembolism, and 31% less likely to be hospitalized for any reason, than individuals in the control group, based on medical claim data.

Medco's chief medical officer, Robert Epstein, said in a statement, "These results show that we can greatly reduce hospitalizations, and their significant costs, by making genetic testing routine early in a patient's therapy with warfarin. If it costs a few hundred dollars for the genetic test but avoids the $13,500 hospital bill, it very quickly pays for itself."

And more findings are rolling in; healthcare reform (hopefully) will help

A professor of molecular and human genetics at Baylor College of Medicine in Houston, Richard Gibbs, recently published a report describing single-gene mutations responsible for Charcot-Marie-Tooth disease, an inherited nerve disorder. The authors studied 10 family members to understand the genetic cause of the disease, previously linked with dozens of errant genes. This along with a host of other recent studies shows that the era of precision medicine is near.

In addition, the recently passed US Healthcare reform bill could also help provide important data as it establishes the Patient-Centered Outcomes Research Institute (POCRI to "identify priorities for and provide for the conduct of comparative outcomes research."

Fig 20 Patient-Centered Outcomes Research Institute funding (US$m)

2010	2011	2012	2013	2014-2019
10	50	150	150+ ($1XAvg lives covered)	150+ ($2XAvg lives covered)

Source: H.R. 3590, Macquarie Capital (USA), April 2010

"Tailoring medicine is likely to be one of the most important themes in healthcare"

According to recent reports, the FDA today requires genetic testing for six drugs, and also recommends testing before prescription for more than two dozen medicines. It also mentions diagnostic tests in the labels of more than 150 others. "Tailoring medicine, so that the right therapeutic is delivered to the right person, is likely to be one of the most important themes in healthcare," FDA Commissioner Margaret Hamburg said in a Feb. 25 speech.

Leroy Hood, the founder of the Institute for Systems Biology (ISB) recently stated in an interview with *The New York Times* that, "There will be an explosion of family sequencing that will identify disease genes. My prediction is that in 10 years or so, most of us will have our genome sequences done as part of our medical records and it will be an important part of predictive medicine."

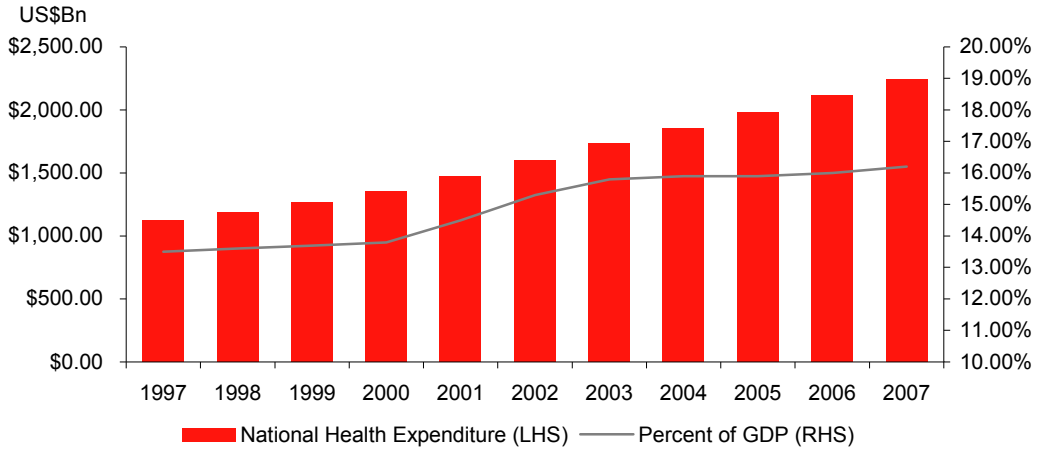
"This is just one of a series of events that indicates a revolution in the way that we are doing genetics," Georg M. Church, professor of genetics at Harvard Medical School, was recently quoted as saying in response to a recent sequencing-disease study. "It's signaling the early arrival of inexpensive DNA sequencing."

"We are finally about to turn the corner, and I suspect that in the next few years human genetics will finally begin to systematically deliver clinically meaningful findings," David B. Goldstein, a Duke University geneticist, was quoted as saying in the same article.

General healthcare trends in the US and the world

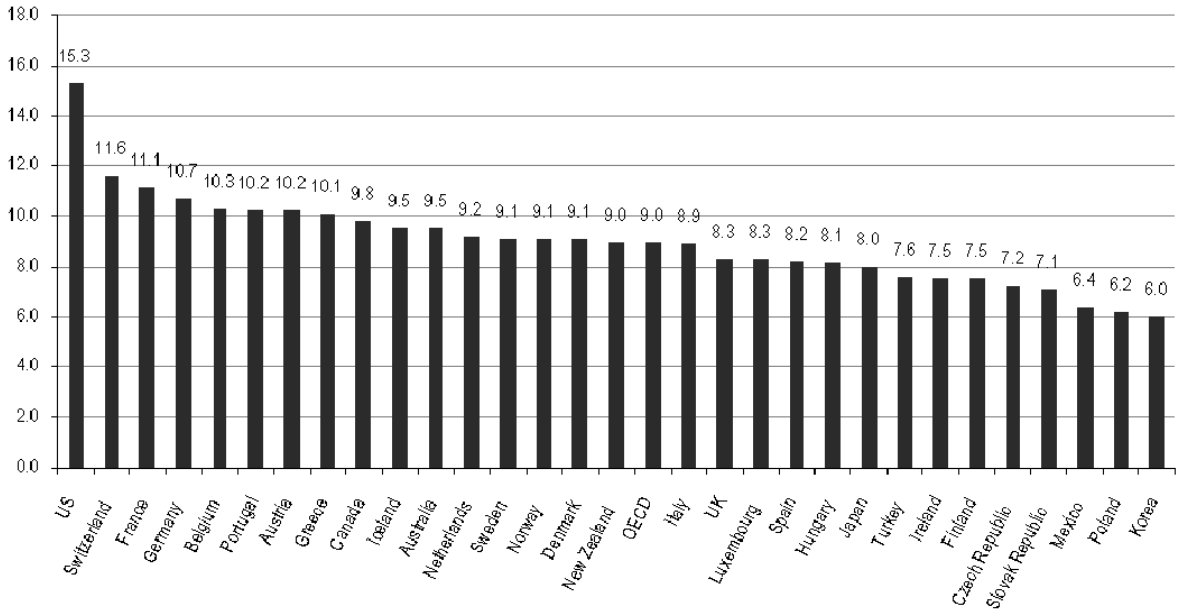
All of this, of course, must be put in the context of current healthcare trends. The United States spends more money on health care than any other country in the world (nearly 17% of GDP; see Figures 21 and 22) and new discoveries must help lower that burden.

Fig 21 US national health expenditures (1997–2007)



Source: CMS, Macquarie Capital (USA), April 2010

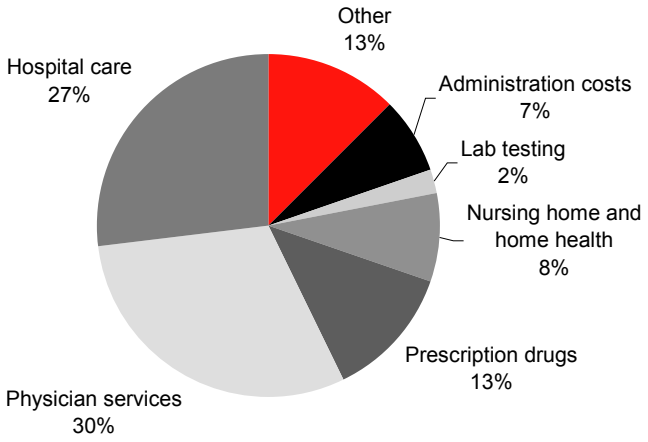
Fig 22 Worldwide healthcare expenditure as a percentage of GDP, 2005



Source: WHO, Macquarie Capital (USA), April 2010

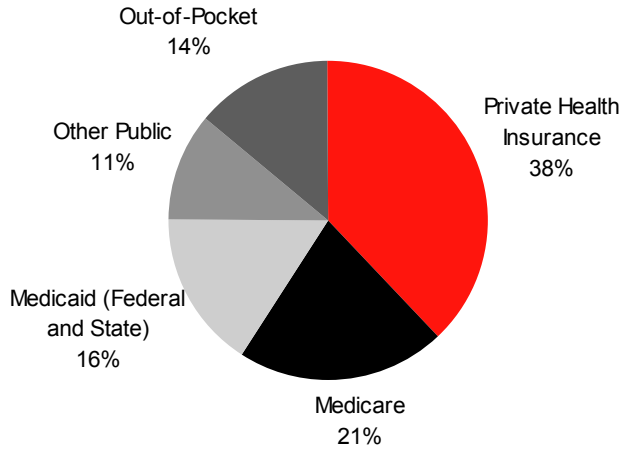
The US lags in key health indicators such as life expectancy, where, according to 2009 estimates from the CIA Factbook, the US ranks 50th out of 224 nations with an average life span of 78.1 years. In addition, according to the Office of Management and Budget (OMB), if costs per enrollee in Medicare and Medicaid grow at the same rate over the next four decades as they have over the past four, those two programs will increase from 5% of GDP today to 20% by 2050.

Fig 23 2008 US health expenditure detail (US$2.4tr)



Source: CMS, Company data, Macquarie Capital (USA), April 2010

Fig 24 US healthcare spending by sponsor



Source: Washington Research Council, Macquarie Capital (USA), April 2010

Emerging markets' healthcare expenditures likely to continue to grow

Research shows that the use of medical services tends to increase with the level of income; eg, as Americans became wealthier, they became willing to spend more on safer and better medical services. This suggests that we may see similar trends in developing countries as per-capita wealth continues to grow.

Government officials in other parts of the world are also more committed to providing citizens with better healthcare. For example, most residents in China are paying for medical services themselves (in 2004 out-of-pocket expenses accounted for 64% of healthcare spending). The Chinese government wants to build up state healthcare spending and has already started rebuilding community health centers in China's major cities. The Chinese government has also proposed a medical insurance system for both urban and rural workers.

In India, total healthcare expenditures are just over 5% of GDP, with public spending around 1%. Public sector services are not only weak, but they are also underutilized and inefficient. In rural areas, clinics are badly maintained and ill-equipped. It is reported that only one in ten Indians has any form of health insurance, and thus out-of-pocket payments for medical services amount to 98.4% of total healthcare expenditures. However, according to the Public Health Foundation of India, the government has promised more money for rural health through its National Rural Health Mission. More specifically, officials say they will increase public health spending from the current 1% of GDP to 3% by 2010.

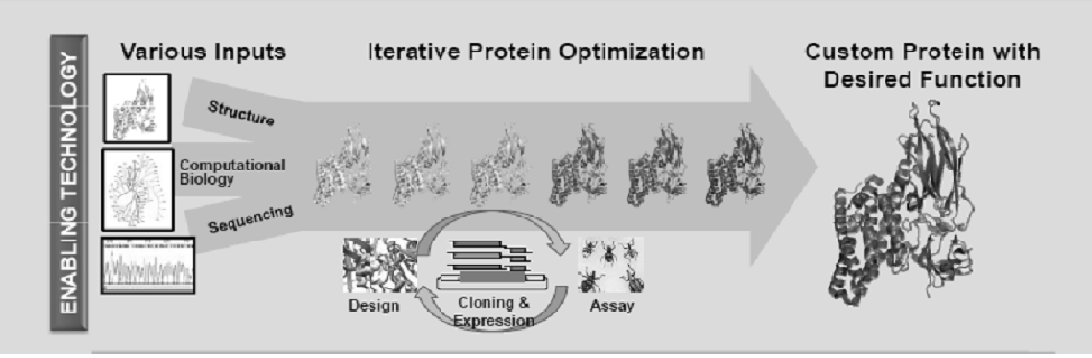
Agriculture

More food will have to be produced worldwide over the next 50 years than has been during the past 10,000 years combined.

It has been estimated that in order to keep up with the growth in human population, more food will have to be produced worldwide over the next 50 years than has been during the past 10,000 years combined[20].

Modern advances in plant genetics and cultivation of plant-associated microbes are allowing rapid improvements to be made in crops, even in those that have seen little past enhancement. In fact, it has been estimated that today over ~70% of processed conventional foods consumed in the US contain some genetically modified ingredients.

Fig 25 Process of genetically engineering seed crops



Source: Monsanto, April 2010

For example, a better understanding of molecular plant biology has already allowed researchers to create seeds that are more drought resistant, produce higher yielding crops, or reduce the application of chemical fertilizers and pesticides. See Figure 25.

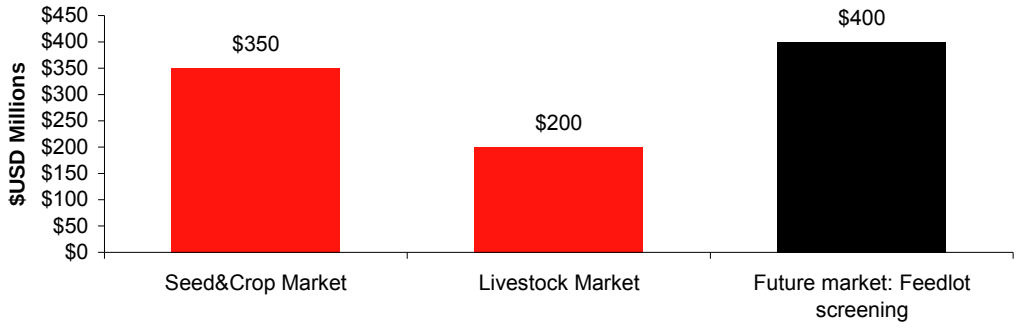
In addition, plant-associated microbe discoveries are being made, such as their ability to improve the availability and uptake of nutrients by plants and promote plant health and growth. The best strains of microbes could be developed as biofertilizers and as disease-control agents, reducing the application of chemical fertilizers and pesticides and enabling effective disease control.

In the field of livestock, farmers could improve breeding of cattle to ensure leaner beef or better dairy production. Of course, the utility curve of agrigenomic genotyping will be highly variable depending on the applied commodity value and application. For example, the genotyping of a prize bull could easily be worth US$350 to a farmer whereas genotyping a chicken would be worth much less given the number of required tests and the lower relative value of a chicken.

Illumina recently estimated that the entire agrigenomics market was around US$550m with the potential to grow another US$400m over time should feedlot screening take off. For this market to grow beyond some of the obvious markets, however, we estimate that an extremely low cost per sample test would be required (perhaps as low as cents on the dollar).

[20] Global Policy Forum

Fig 26 Agrigenomics estimated potential market opportunity (~US$1bn)



Source: Illumina, Macquarie Capital (USA), April 2010

While many individuals and groups have rightfully expressed concern surrounding the genetic modification of agricultural products, technological advances are arguably all that will allow us to continue to feed the increasing human population at a reasonable cost.

Energy

Biofuel companies received over US$500m in VC investments in 2009.

There are many potential applications of genomics in Energy. Let us take biofuels as just one example. A better genomic understanding could help us improve the yield of first generation biofuels, such as corn-based ethanol. Of course, second generation biofuels, such as cellulose, are arguably even more desirable as they side-step the fuel versus food debate, and genomic knowledge could allow us to genetically alter enzymes to better break-down the cellulose. Third generation biofuels, such as algae, may hold the most promise of all. These could be modified such that they produce more oil and are more stable or resistant to disease.

Money has been flowing into biofuel research and development. Within the venture capital community biofuel companies received over US$500m in investments in 2009. See Figure 27.

Fig 27 Global VC investments in cleantech through 2009

Technology Sector	Amount Invested (US$m)	% of Total
Solar	1,200	*21%*
Transportation	1,100	*20%*
Energy Efficiency	1,000	*18%*
Biofuels	**554**	***10%***
Smart Grid	414	*7%*
Water	117	*2%*

Source: The Cleantech group, Macquarie Capital (USA), April 2010

As an example of the potential of genomics in Energy, Exxon Mobil recently signed a US$300m agreement with Craig Venter's Synthetic Genomics Institute (SGI) to develop next generation biofuels using photosynthetic algae.

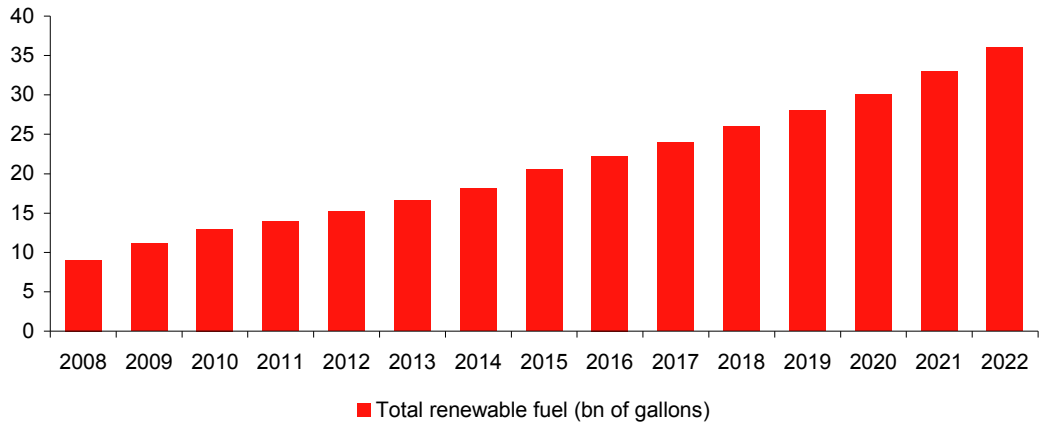
SGI is harnessing photosynthetic algae to produce a range of liquid fuels and chemicals directly from sunlight and carbon dioxide. Algae produce significantly higher amounts of biomass and oil as compared to terrestrial crops; in addition, algae can be grown on land that is not suitable for agriculture and can thrive in sewage or other types of waste water, and are efficient at capturing and recycling carbon dioxide, a major greenhouse gas.

Current methods to produce fuel from algae include processes that resemble farming. Algal cells are grown, harvested, and then bio-processed to recover the lipids from within the cells. In contrast, in one of SGI's solutions, it has engineered algal cells to secrete oil in a continuous manner through their cell walls, thus facilitating the production of algal fuels and chemicals in large-scale industrial operations. Biocrude is SGI's targeted first product.

The 2007 Energy Act created renewable fuel standard requirements that are driven primarily by cellulosic and advanced biofuels.

In the 2007 Energy Act, the US government created renewable fuel standard requirements. These requirements are primarily driven by cellulosic and advanced biofuels. See Figure 28.

Fig 28 US renewable fuel standard requirements; 2007 Energy Act



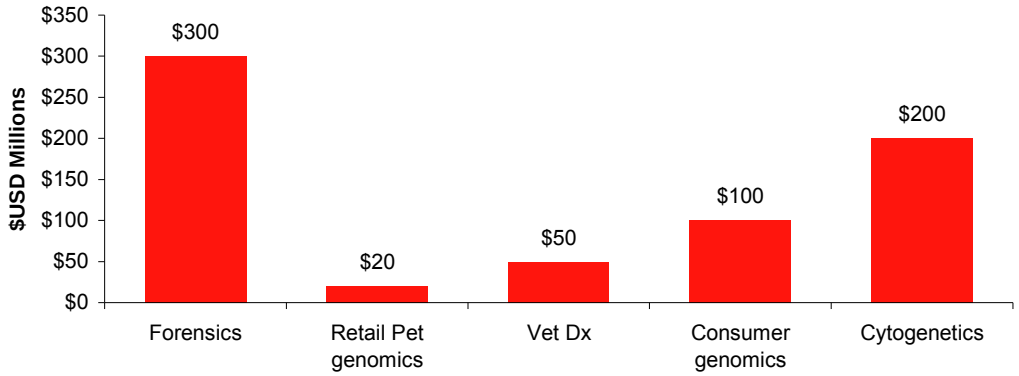
Source: EPA, Macquarie Capital (USA) April 2010

For an informative Joint Genome Institute (JGI) video on what Energy Genomics is and why it will be so important go to http://www.youtube.com/watch?v=qchN5FX_QN0.

Other

There are, of course, numerous other potential applications of increased genomic knowledge that could benefit many other applied markets. For example, one of the largest existing applied markets in genetic tools is forensics, where DNA profiling is used in criminal justice. This market is estimated to be ~US$300m and dominated by Life Technologies. See Figure 29.

Fig 29 Other applied market opportunities for genetic analysis (~US$700m)



Source: Illumina, Macquarie Capital (USA), April 2010

Translating demand into market size

Hopefully, we have made a compelling case that there is a tremendous amount of demand for relevant genetic information. Since sequence knowledge is the foundation of better genetic understanding, we now attempt to quantify the current sequencing market and also provide our view as to how the market is likely to develop over time. We also address how the sequencing market might impact other genomic tools markets.

Customers and funding sources

We estimate that ~5,000 labs have at least one sequencing instrument.

Sequencing today occurs in laboratories and primarily in research laboratories. Industry data suggests there are over 200,000 laboratories scattered throughout the world and around 50,000 of these could be categorized as labs with molecular biology capabilities. Most of these would be research labs; however, in the US there are approximately 2,300 clinical labs that have molecular capabilities. In addition, there are estimated to be approximately 1,000 forensics labs globally. We estimate that ~5,000 labs have at least one sequencing instrument. See Figure 30.

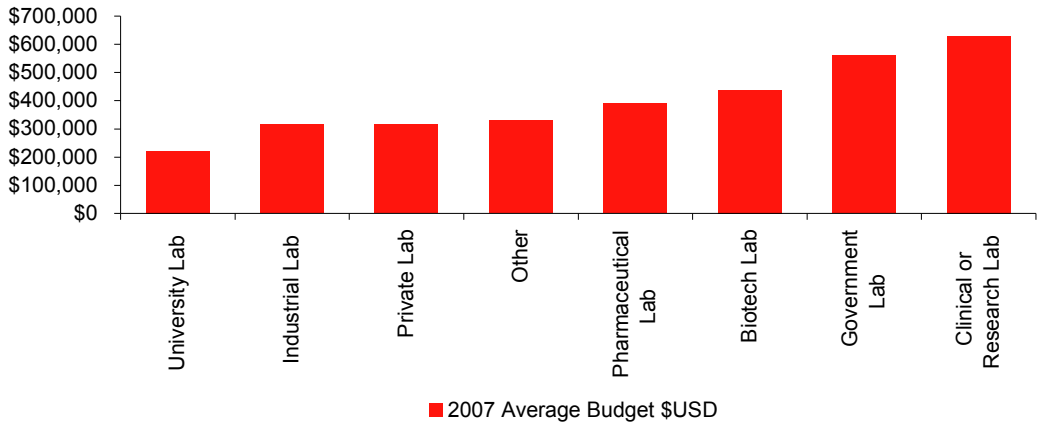
Fig 30 Number of labs globally

	Approximate # of labs
Total laboratories	200,000
Molecular Biology labs	50,000
Clinical with molecular capabilities (US)	2,300
Forensics with molecular capabilities (global)	1,000
Estimated number of labs with at least one sequencer	5,000

Source: Company data, Macquarie Capital (USA), April 2010

According to a 2007 survey by Lab Manager Magazine, the average lab budget in the US is around US$400K. However, 41% of labs reported that their budget was over US$1m and ~45% said that their budget for purchasing equipment was over US$100K. It is important to note that 75% of the respondents stated that 25-75% of their budget was for salary and compensation.

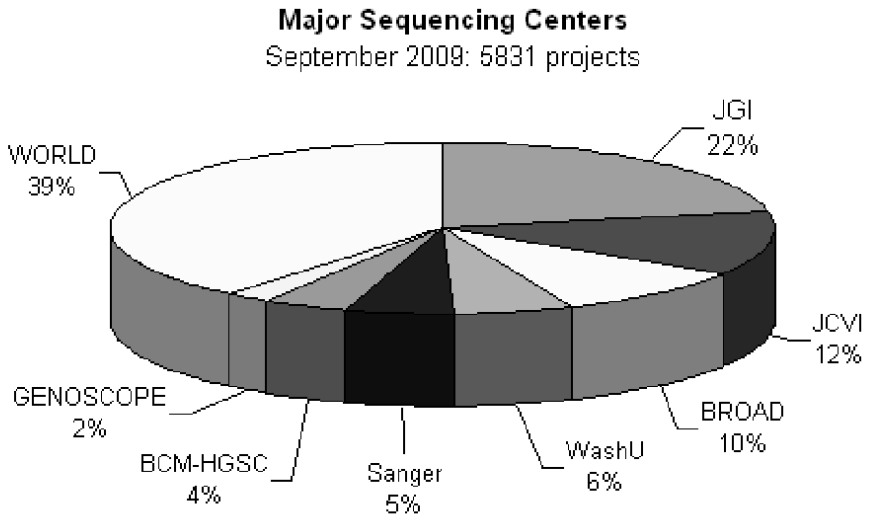
Fig 31 Average lab budget by industry pre recession



Source: *Lab Manager* magazine, Macquarie Capital (USA), April 2010

Because smaller labs are budget constrained, major sequencing centers remain an important market, especially for high-throughput sequencing instrumentation. The major sequencing centers remain primarily US based, including: The Broad, Washington University, Baylor College of Medicine, JGI, and JCVI. The Wellcome Trust Sanger Institute and the Beijing Genome Institute have become the major ex-US Genome centers. See Figure 32.

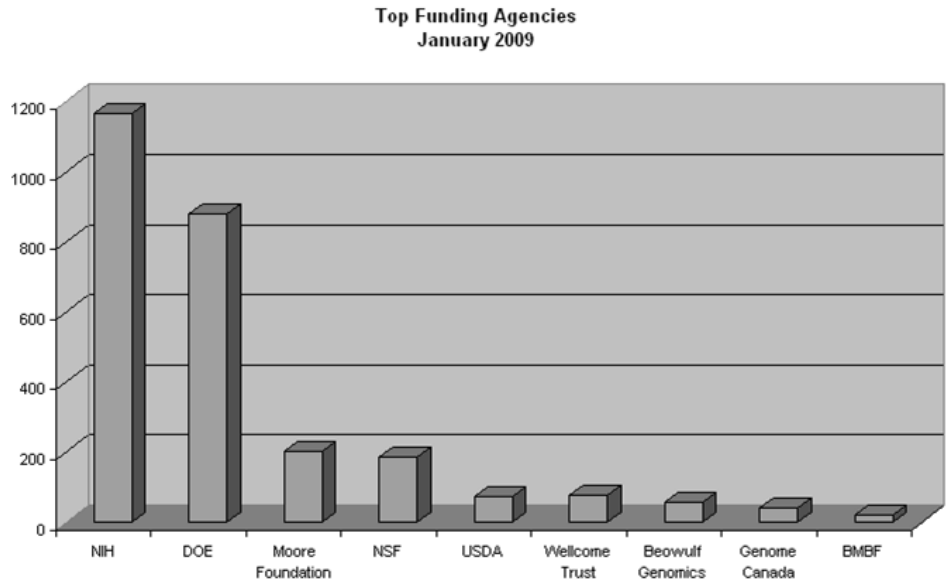
Fig 32 Most major sequencing centers are in the US



Source: GOLD statistics, Macquarie Capital (USA), April 2010

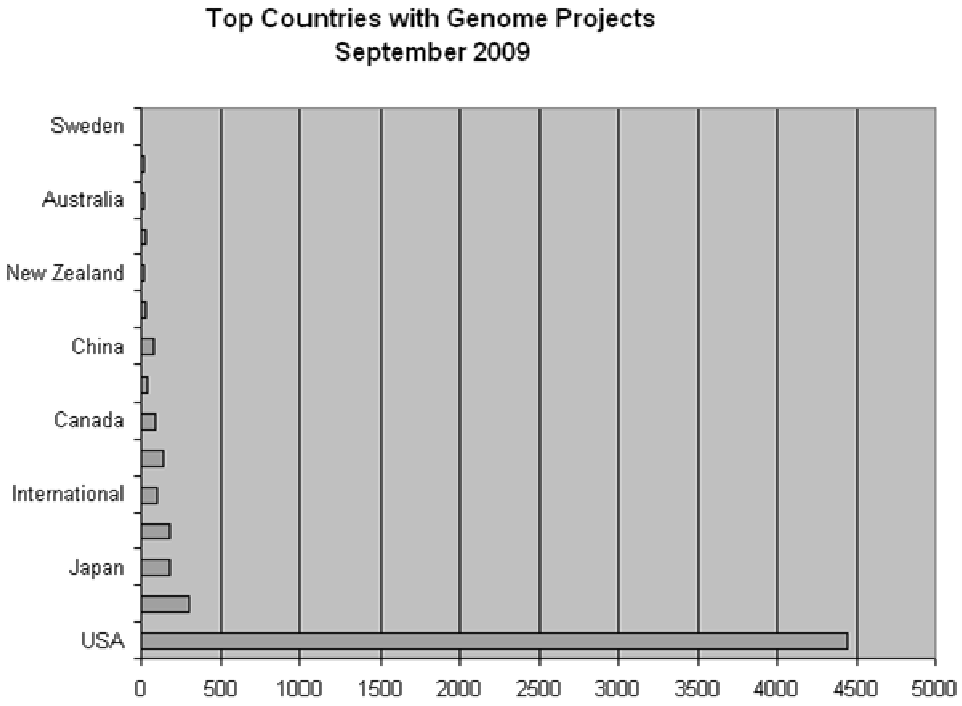
In addition, the US government is the primary funding source for sequencing projects via the National Institutes of Health and the Department of Energy. See Figure 33. Thus, most sequencing projects are based in the US. See Figure 34.

Fig 33 The US Government is the primary source of sequencing project funding



Source: GOLD statistics, Macquarie Capital (USA), April 2010

Fig 34 The US also has the most sequencing projects, by far



Source: GOLD statistics, Macquarie Capital (USA), April 2010

Competition and rivalry

The current sequencing market is served by two relatively distinct technologies: Capillary Electrophoresis (CE) and massively parallel sequencers (e.g. second generation)[21]. As the first highly automated and high throughput DNA sequencers, CE systems were broadly adopted in the final sprint of the Human Genome Project in the late 1990's. There are an estimated ~13,500 total CE units in the field today, with the lions share (~70-75%) supplied by Life Technologies (formerly ABI). GE Healthcare and Beckman Coulter are the other major players.

We estimate that Illumina has around 2/3 of the market with Life Technologies second and Roche third.

Massively parallel sequencers can be either "short" or "long-read." There are ~1,500 second gen instruments installed. The main players in second generation sequencing are Life Technologies with its SOLiD platform, Illumina with its Genome Analyzer and Hi Seq offerings, and Roche with its 454 sequencer. We estimate that Illumina has around 2/3 of the market with Life Tech second and Roche third (despite being first to market in 2005).

Fig 35 Second generation installed base market share estimates

Firm/Instrument	# of instruments in survey	Market share
Illumina/Genome Analyser or Hi Seq	635	62%
Roche/454	153	15%
Life Tech/SOLiD	235	23%
	1023	

Note: Illumina and ABI estimates adjust for recent announcements with BGI and Ignite, respectively

Source: http://pathogenomics.bham.ac.uk/hts/stats, Macquarie Capital (USA), April 2010

[21] Each of the current systems on the market relies on a camera optically reading colors emitted from a DNA sequence reaction.

As Figure 36 shows, most large genome centers have already decommissioned their legacy high-throughput CE instruments and adopted second generation sequencers.

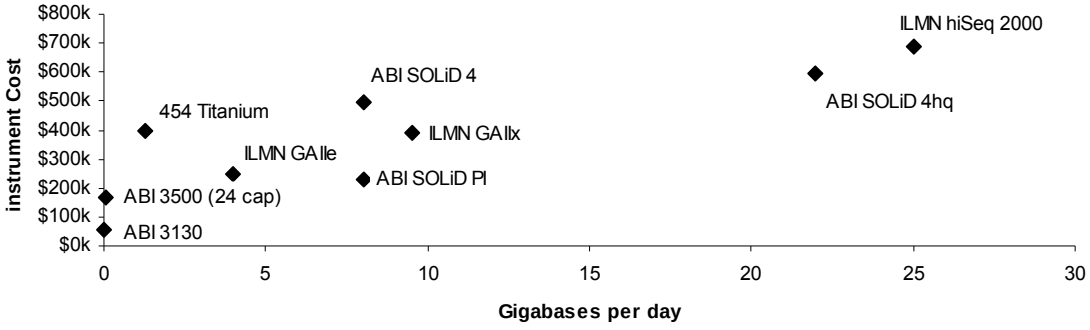
Fig 36 Top 20 Genome Centers and current second generation installed base

Top 20 Genome Centers	Number of 2nd gen instruments
Broad Institute	109
BGI (formerly Beijing Genomics Institute)	128
Ignite Institute	100
The Genome Center at Washington University	58
Wellcome Trust Sanger Institute	42
DOE Joint Genome Institute	20
Baylor College of Medicine	18
Michael Smith Genome Sciences Centre	16
Ontario Institute for Cancer Research	15
Beijing Institute of Genomics	15
Centro Nacional de Análisis Genómico (CNAG)	10
Genome Institute of Singapore	9
Cold Spring Harbor Laboratory	8
Centre for Genomic Research	7
Beckman Coulter Genomics (formerly Agencourt)	6
UCL Genomics	6
JCVI	6
Cambridge Research Institute	5
GATC	5
Duke IGSP Sequencing Core Facility	5
ErasmusMC	5
DNAVision Agrifood	5

Source: http://pathogenomics.bham.ac.uk/hts/stats, Macquarie Capital (USA), April 2010

Although CE technology cannot compete with second generation sequencers on the basis of speed or output, CE is still finding a growing market in forensics and other applied markets where a lower amount of genetic information is demanded and where cost, accuracy, read-length, and ease-of-use are the main considerations. See Figure 37.

Fig 37 Current positioning of DNA sequencing instruments



Source: Company data, Macquarie Capital (USA) April 2010

New technologies are being developed that may be able to make sequencing even simpler and more accessible.

Of course, new technologies are being developed that may be able to make sequencing even simpler and more accessible. Some are aiming to accurately read a DNA sequence electronically (e.g. via an electronic current rather than optics) while others are attempting to remove the need for making millions of copies of DNA (which is expensive and can ad bias). Some are attempting to combine both of these attributes. Most experts believe this is the next logical evolution in sequencing speed and cost reduction. For more details see the Appendix.

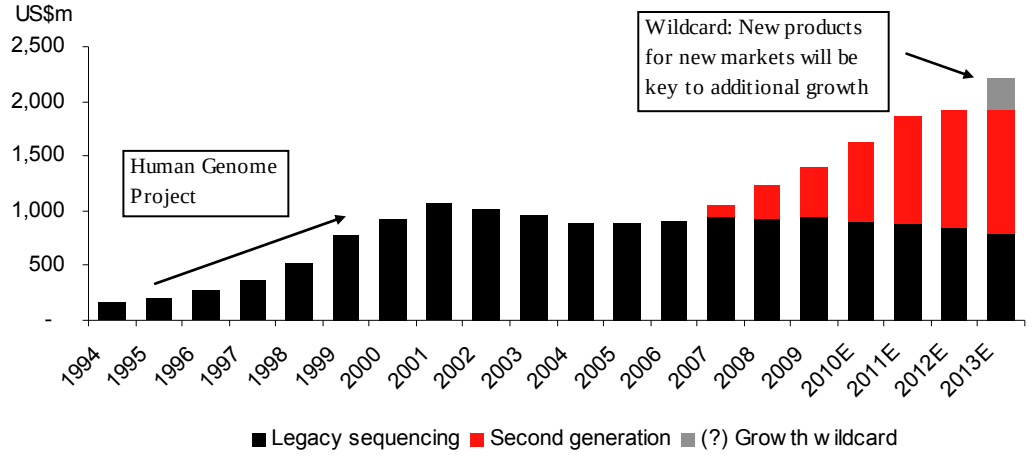
Sequencing: Historical growth (past as prologue)

Ronald Reagan has been quoted as saying, "I do not want to go back to the past; I want to go back to the past way of facing the future[22]." Let us use the past, then, as a guide as to what may happen in the future.

The legacy (CE) DNA sequencing market is currently around US$950m, roughly US$100m below what it was in 2001, the year we estimate to be its peak.

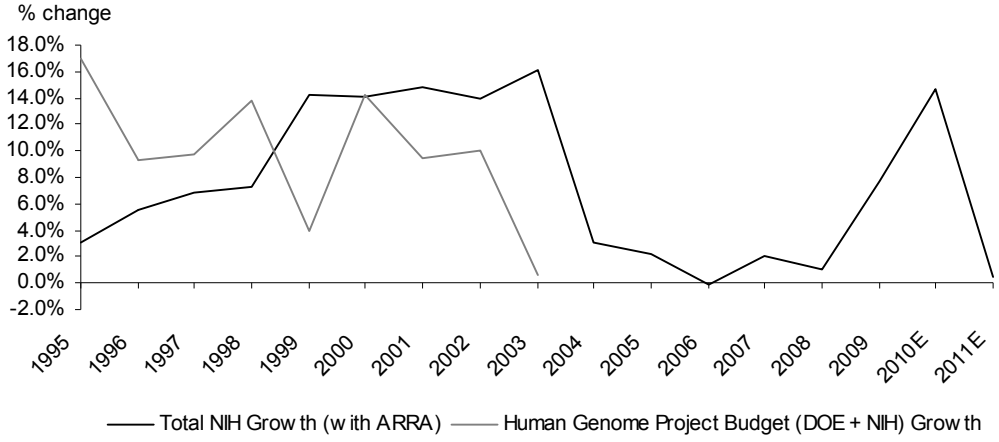
The legacy (CE) DNA sequencing market is currently around US$950m, roughly US$100m below what it was in 2001, the year we estimate to be its peak. The sequencing market experienced rapid growth in the late 90s as there was a mad rush between both the public and private sectors to be the first to sequence the human genome. Then, between 2001 and 2006, the market contracted before again levelling out. Second generation sequencing was introduced in 2005 with 454's (now Roche) instrument but it was not until 2007 when Solexa (via its new parent Illumina) introduced its Genome Analyzer that the market again began to truly grow. See Figures 38 and 39.

Fig 38 The life, death, and re-birth of the sequencing market (1994-2013E)



Source: Company data, industry literature, Macquarie Capital (USA), April 2010

Fig 39 NIH and Human Genome Project funding accelerated growth in the 90s



Source: NIH, DOE, Macquarie Capital (USA), April 2010

[22] 1984 address to the nation

Sequencing: Growth forecasts through 2013

We will do our best to learn from history but accept that we are likely destined to become someone else's future punch-line.

The only thing certain about market forecasts is that they are bound to be wrong. For example, in examining a host of old market research reports we found that they inevitably looked at the future simply by extrapolating from the most recent past. This is human nature. For example, look at Figures 40 and 41, which are from a 1999 market research report. As we stated previously, the market experienced rapid growth in 1999-2001 and then began to decline.

Fig 40 1999 forecast of the US sequencing market

Automated DNA Sequencer Market: Revenue Forecasts (U.S.), 1995-2005

Year	Revenues ($ Million)	Revenue Growth Rate (%)
1995	94.7	---
1996	101.8	7.5
1997	109.7	7.7
1998	122.1	11.3
1999	134.9	10.5
2000	149.0	10.5
2001	155.6	4.4
2002	167.4	7.6
2003	180.5	7.8
2004	194.8	7.9
2005	204.4	4.9
Compound Annual Growth Rate (1998-2005): 7.6%		

Note: All figures are rounded; the base year is 1998. Source: Frost & Sullivan

Source: Frost & Sullivan, April 2010

Fig 41 1999 forecast of the US sequencing market (cont'd)

Measurement Name	Measurement	Trend
Market age	Mature	
Revenues	$114.0 million	Increasing
Potential revenues (maximum future market size)	$181.6 million	Increasing
Current (base year) market growth rate	7.4%	Decreasing
Forecast period market growth rate	7.0%	Decreasing
Saturation (current/potential users)	70%	Increasing
Average price	$118,000	Increasing
Price range	$30,000 to $220,000	
Price-sensitivity	Medium	
Number of products	8	Decreasing
Competitors (active market competitors in the base year)	4	Decreasing
Concentration (percent of base year market controlled by top competitor)	90%	Stable

Note: All figures are rounded. Source: Frost & Sullivan

Source: Frost & Sullivan, April 2010

Then, in 2006, when the sequencing market had been in decline for four years market researchers failed to adequately account for the boost from new second generation technologies. Consider this description from an SDI report in 2006:

"In 2005, the initial systems market accounted for $310 million or 37% of the total market. It is expected to decline to $241 million by 2010. The market is currently saturated with roughly 11,000 automated DNA sequencers in the installed base, most of which were sold by Applied Biosystems or GE Healthcare. ***The demand for high throughput de novo sequencing has probably matured.****"*

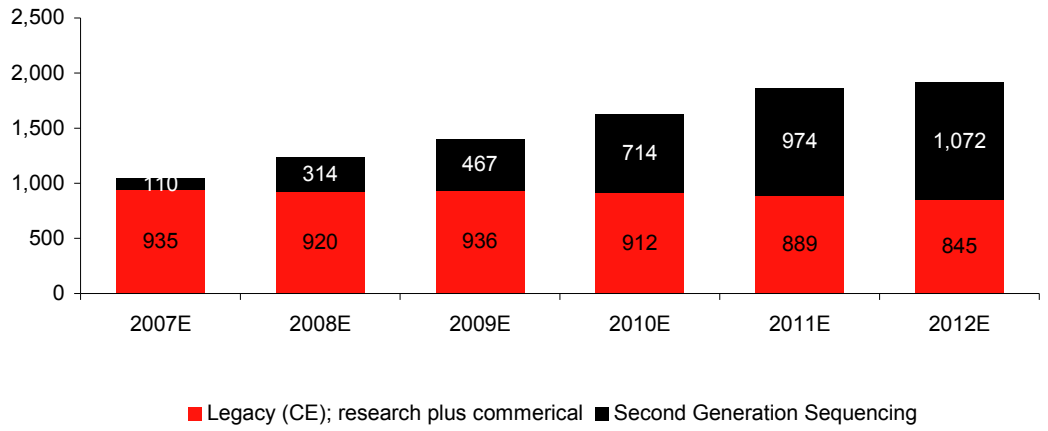
With that as background, we will do our best to learn from history but willingly accept that we are likely destined to become someone else's punch-line at some point in the future.

Total market expected to grow at a 3-year CAGR of ~11%

We estimate that the total sequencing market will grow from roughly US$1.4bn in 2009 to around US$1.9bn in 2012, with most of the growth occurring in 2010 and 2011.

We estimate that the total sequencing market will grow from roughly US$1.4bn in 2009 to around US$1.9bn in 2012, with most of the growth occurring in 2010 and 2011. The CE market is expected to decline 3-5%, on average, each year, while the second generation market is expected to grow 50% in 2010, 36% in 2011, and 10% in 2012. The wildcard in 2012 and beyond, in our view, is the potential impact of new technologies.

Fig 42 Sequencing market forecast: ~11% CAGR 2010E-2012E (US$m)



Source: Company data, industry literature, Macquarie Capital (USA), April 2010

We estimate that the second generation sequencing market needs to grow to roughly US$1.1bn by 2012 in order to hit our US$1.9bn market forecast. A US$1.1bn market would imply ~4,500 next-generation sequencing instruments placed in labs.

Fig 43 A $1.9bn sequencing market implies ~4,500 next gen instruments

	2009	2012
Legacy CE forecast (US$m)	936	803
Second generation forecast (US$m)	467	1,100
Second generation market size (US$m)	468	1,093
Total instruments placed	1,500	4,500
Consumable US$'s/instrument/year	170,000	175,000
% new/replacement instruments	*38%*	*17%*
New/Replacement instrument cost (US$)	380,000	400,000

Source: Macquarie Capital (USA), April 2010

Our estimates are based on the assumption that both the annual consumable stream and the average instrument cost remain fairly constant with 2009 (e.g. higher priced high-throughput products are offset by lower-priced lower-throughput ones). To put this number in perspective, we estimate that there were ~4,000 high-throughput CE instruments prior to the beginning of the recent second generation restatement (the Human Genome Project itself used ~900). So we see the installed base of high-throughput instrumentation expanding by 10-15% given the broader capabilities of second gen instrumentation.

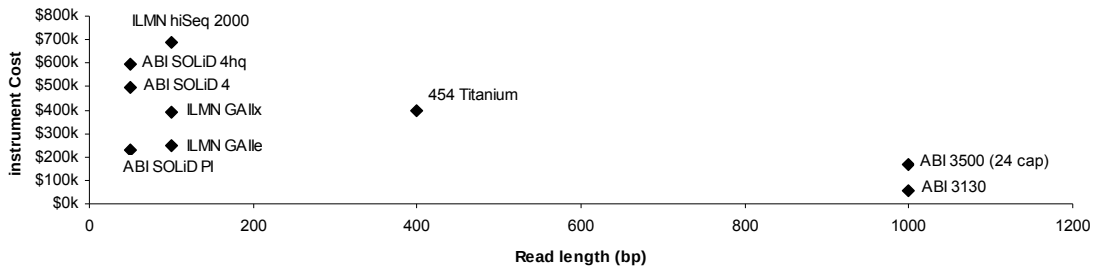
CE market expected to decline 3-5%, but not linearly

We expect the CE market to decline on average 3−5% per year, though it will likely not be linear.

We expect the CE market to decline on average 3−5% per year as researchers continue to transition to second generation technologies. However, it is important to note that further displacement will take time and will almost certainly not be linear. In 2009, for example, the research market, which has been in decline since 2002, actually grew.

CE remains an important tool for a few reasons. It is the gold standard in sequencing, meaning that it is used to QA/QC (e.g. quality check and validate) most all of the second generation work. In addition, CE is able to read long lengths of sequence while second generation sequencing is still limited in the length of sequence it can read, meaning that CE is often used to fill in its gaps. Finally, CE has become the standard in commercial sequencing applications (for example diagnostic tests and forensics) which continue to grow and are not likely to migrate to new technology platforms anytime soon given the cost and risk involved. See Figure 44.

Fig 44 Sequencer platforms: Read length versus cost



Source: Company data, Macquarie Capital (USA) April 2010

ARRA to accelerate second generation technology restatement

We forecast that second generation sequencing will grow ~50% in 2010, ~35% in 2011, and only ~10% in 2012.

As can be seen from Figure 42, we forecast that second generation sequencing will grow ~50% in 2010, ~35% in 2011, and only ~10% in 2012 for a three year CAGR of ~32%. We had expected the second generation market to grow rapidly of its own accord; however, we think that the US$10bn in NIH stimulus funds will accelerate the legacy technology restatement in 2010/11.

Over time, most high-throughput CE systems should migrate to second generation instruments. There are still 1,500-2,000 high-throughput instruments in the field, though it is impossible to know how actively they are being used. See Figure 45. Replacing these systems could double the ~1,500 second generation systems that are currently installed.

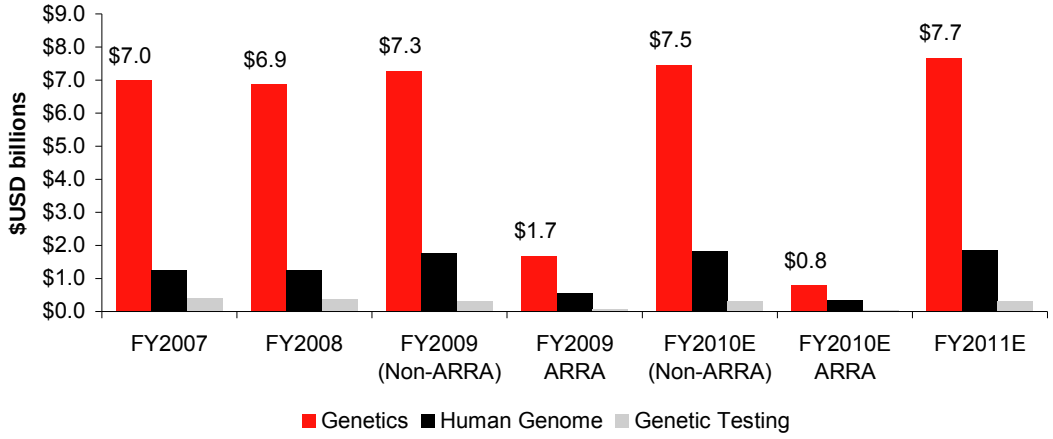
Fig 45 Nearly 10% of molecular biology labs currently have CE instrumentation

	Units	Total Molecular Labs Globally	% of labs with sequencers	Estimated labs
Low-to-medium throughput	12,000		9.1%	4,571
High Throughput	1,500		0.4%	188
Total	13,500	50,000	9.5%	4,759

Source: Macquarie Capital (USA), April 2010

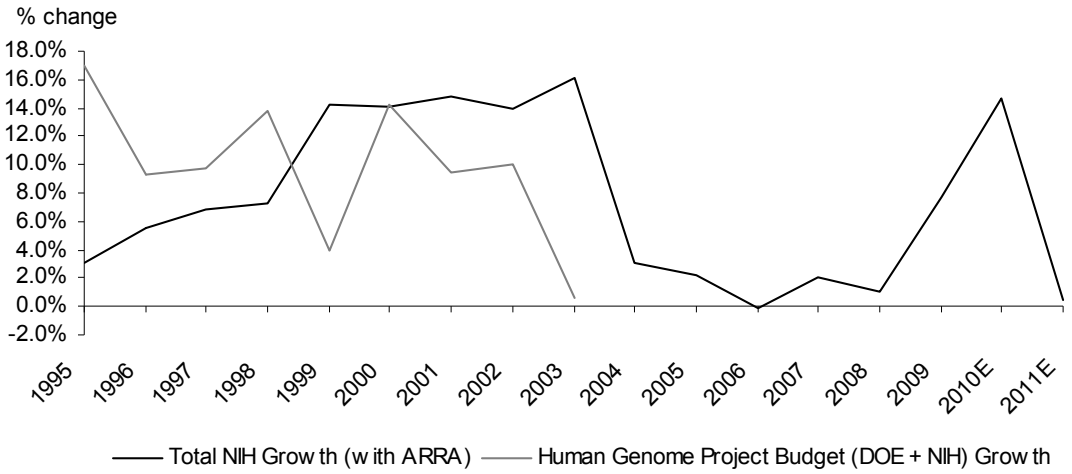
While it is virtually impossible to dissect with any precision, we estimate that the ARRA will likely double what would have been the underlying growth rate of the second generation sequencing market. Excluding the ARRA, the underlying NIH budget would have grown at right around 3%. With the ARRA the three-year CAGR will be close to 8%, with much of the additional funding going toward genomics and other leading edge projects. See Figures 46, 47 and 48.

Fig 46 A disproportionate amount of ARRA funding is going to genetics and we estimate over US$130m in sequencing projects alone were approved in FY2010



Source: NIH, Macquarie Capital (USA), April 2010

Fig 47 NIH stimulus dollars are assisting the technology restatement



Source: NIH, DOE, Macquarie Capital (USA), April 2010

Fig 48 Second gen sequencing is the driver behind revised sequencing goals

Quarter	Total Q20* Bases (Billions)			Q20* Bases (Billions) by			Operating Hours**		
	Goal	Actual Total	Actual % of Goal	Sanger	454	Illumina	Goal	Actual Total	Actual % Goal
Q1 2009	39.9	124.21	311	6.02	23.01	95.18	2100	2088	99.4
Q2 2009	60.1	196.829	328	5.849	38.48	152.5	2100	2146	102
Q3 2009	71.2	236.566	332	5.251	63.127	168.188	2100	2184	104
Q4 2009	81.8	446.278	546	3.452	45.96	396.866	2100	2208	105
FY 2009 Total	**253**	**1003.887**	**397**	**20.58**	**170.58**	**812.73**	**8400**	**8626**	**103**
Quarter	Total Q20* Bases (Billions)			Q20* Bases (Billions) by			Operating Hours**		
	Goal	Actual Total	Actual % of Goal	Sanger	454	Illumina	Goal	Actual Total	Actual % Goal
Q1 2010	643	640.365	99.5	1.223	45.865	593.278	2100	2160	103
Q2 2010	1285						2100		
Q3 2010	1285						2100		
Q4 2010	1285						2100		
FY 2010 Total	**4498**	**640.365**	**0.1423**	**1.223**	**45.865**	**593.278**	**8400**		

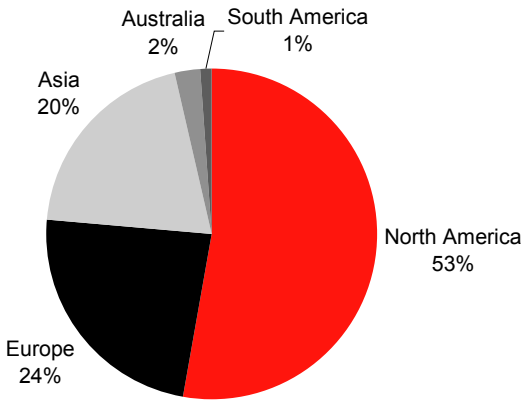
*Q20 indicates good confidence in the assignment of a base. **Number of hours a week that sequencing machines are producing data.

Source: JGI (Joint Genome Institute) sequencing statistics, Macquarie Capital (USA), April 2010

Funding and investment outside of the U.S. could also begin to drive more growth.

Funding and investment from outside of the U.S. could also begin to drive more growth. Just as Asia and Latin America are the growth engines of other parts of the economy, they will continue to ramp up their respective investment in cutting edge technologies as well. For example, BGI (formerly called The Beijing Genomic Institute) did not exist ten years ago. Now, with its new capacity of 128 HiSeqs, which would allow BGI to sequence ~11,000 human genomes a year (an average of 30 per day), it claims it will surpass the sequencing output of the entire US.

Fig 49 Geographic distribution of second generation sequencers



Source: http://pathogenomics.bham.ac.uk/hts/stats, Macquarie Capital (USA), April 2010

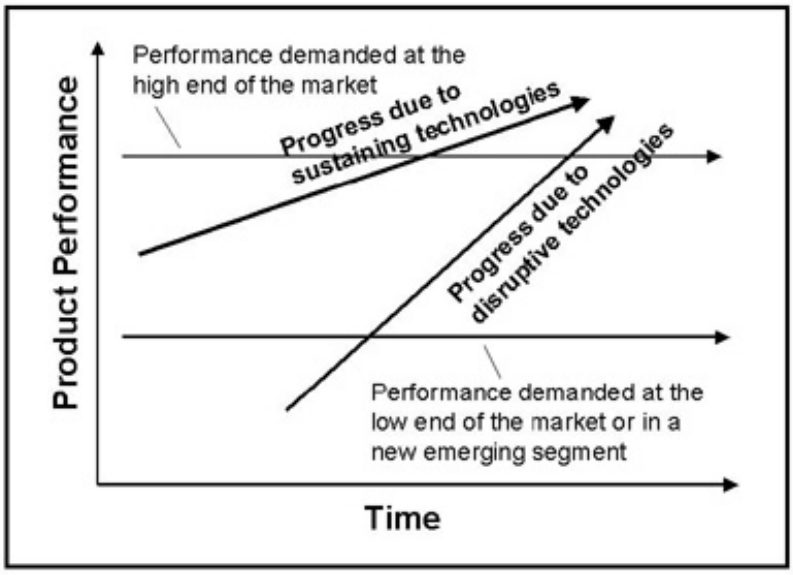
Second gen technologies appear more sustaining than disruptive

In his book *The Innovator's Dilemma,* Clayton Christensen defined sustaining technological innovations as those which improve the performance of established products along the dimensions of performance that mainstream customers in major markets value.

Our conclusion is that second generation sequencing is, at its core, a sustaining rather than disruptive innovation.

Disruptive technological innovations, he noted on the other hand, tend to bring to the market a different value proposition than what had been available previously. These technologies generally **under-perform** established products in the mainstream markets on the dimensions existing customers care about; however, they offer other features that a few fringe, and generally new, customers value. Typically, but again not always, the product offerings are also cheaper, simpler, smaller, and more convenient to use, and initially have a worse profitability profile.

Fig 50 Second gen has both disruptive and sustaining technology characteristics



Source: Clayton Christensen's *The Innovator's Dilemma,* April 2010

Whether second generation sequencers are disruptive or not is an important question for us to answer, in our view, as Christensen found that rarely do even the most radically difficult sustaining technologies lead to the failure of incumbent firms.

To be fair, as we evaluate second generation sequencing instruments, we find elements of both types of technological innovations. They are disruptive in that they do alter the nucleic acid sequencing game by allowing new researchers to do things that they simply could not do before due to time and cost constraints. However, they are sustaining in that they primarily satisfy an existing customer's demands for more sequence, as fast as possible, at the lowest price.

After much consideration, however, our conclusion is that second generation sequencing at its core is a sustaining innovation.

When a user was recently asked about a new technology platform he replied, "All I want is the most data, in the shortest amount of time, at the lowest cost."

This hit home to us when we recently met with a researcher from the Mayo Clinic. When asked about new technology platforms he replied, "All I want is the most data, in the shortest amount of time, at the lowest cost." He assumed it would be accurate.

With the exception of read-length, second generation sequencing in many ways outperforms existing sequencing technologies along the dimensions traditional high-throughput customers care about. In addition, it is sold to the same customers and is funded by the same agencies. As discussed in great detail, the instruments do lower the sequencing cost tremendously, but are also still large, expensive and complex. Perhaps most importantly, second generation sequencing is just as, if not more, profitable than legacy technologies.

By comparison, in a survey of sequencing customers that included many smaller labs (including applied markets) the most influential features and factors in choosing which instrument to purchase were accuracy, reliability, read-length, ease of use, company reputation and price.

Until the cost, ease of use, read-length, footprint, informatics and other barriers are removed, we think second gen sequencing will struggle to move outside of existing high-throughput markets.

Second generation technologies likely to ultimately compete on price

An important corollary to all of this is that when the performance of two or more competing products has improved beyond what the market demands, customers are no longer able to base their choice on performance. The basis of product choice then often evolves from functionality to reliability, then to convenience, and, ultimately, to price.

Second generation sequencing markets will likely be forced to compete on price in the not too distant future.

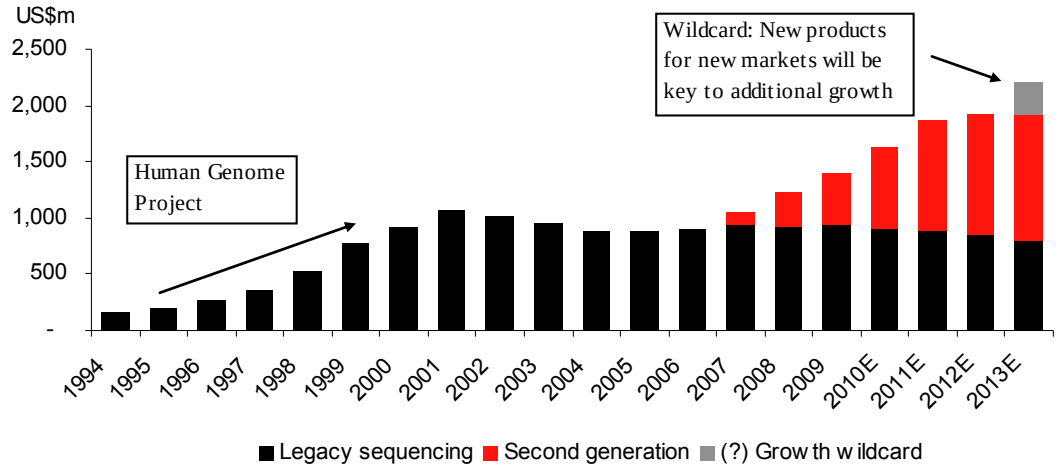
As we wrote earlier, we think that data analysis will soon be the rate-limiting factor in genomic analysis, not data generation. This means that unless they can somehow differentiate themselves via accuracy, informatics, or some other feature, second generation sequencers will likely be forced to compete on price in the not too distant future.

This is one reason why we think firms are already aggressively investing in third-generation technologies, though we suspect that existing companies will go after markets that help, rather than hurt, their margins, potentially leaving room for new disruptive entrants.

How to avoid a post-ARRA hangover

In order to avoid a repeat in 2012 of the 2001-2006 bust, new technologies will need to emerge that allow sequencing to move beyond its current customer base, in our view.

Fig 51 The life, death, and re-birth of the sequencing market (1994-2013E)



Source: Company data, industry literature, Macquarie Capital (USA) April 2010

A second gen sequencer in every lab? Not likely with existing offerings

Typical second generation sequencing instruments cost US$400-US$700k, though recent introductions are targeting the ~US$250K range. The question is, how many labs can actually buy these sequencers with their existing budgets (e.g. ex government or outside funding)? The answer: not many.

As mentioned earlier, the average lab budget in the US is around US$400K, though 41% of labs reported that their budget was over US$1 million and ~45% said that their budget for purchasing equipment was over US$100K. It is important to note that 75% of the respondents stated that 25-75% of their budget was for salary and compensation (see Figure 53).

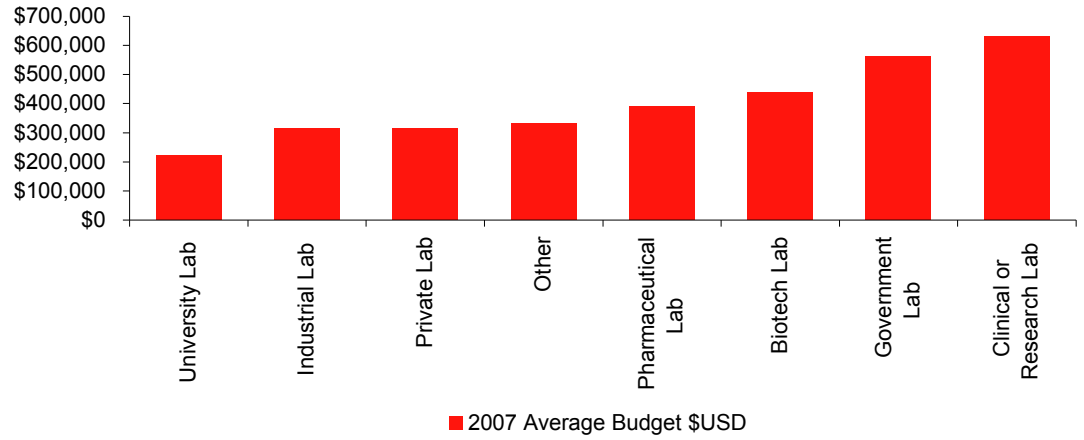
Fig 52 ~10% of molecular biology labs currently have sequencing instrumentation

	Units	Total Molecular Labs Globally	% of labs with sequencers	Estimated labs
Low-to-medium throughput	12,000		9.1%	4,571
High Throughput CE	1,500		0.4%	188
High Throughput Next Gen	1,500		0.4%	188
Total	15,000	50,000	9.9%	4,946

Source: Company data, Macquarie Capital (USA), April 2010

Of course, there are other barriers beyond the initial instrument cost. Second generation workflows require significant lab space, informatics capabilities, and technical expertise. Smaller labs simply may not be able, or want, to make the necessary investments.

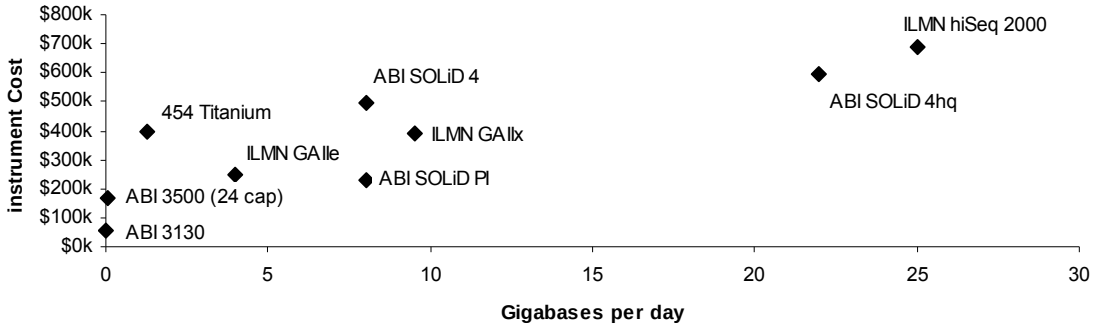
Fig 53 Average lab budget by industry pre recession



Source: *Lab Manager* magazine, Macquarie Capital (USA), April 2010

This is one reason why lower cost and lower throughput instruments have remained unscathed by the second generation innovations. However, both existing and new companies are actively looking to introduce a new generation of technologies aimed at these smaller labs. In our view, these technologies will need to have lower instrument costs, lower reagent costs, fewer sample prep and informatics requirements, and superior accuracy in order to be adopted. See our Appendix for details on some technologies under development.

Fig 54 Current competitive DNA sequencing instruments



Source: Company data, Macquarie Capital (USA), April 2010

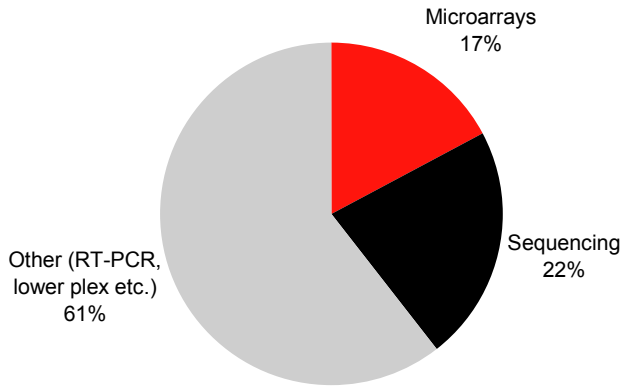
Other genomic tools likely to benefit from sequencing discoveries

By this point it should be clear that genomes are read out as linear sequences and that this sequence information is the foundation of genomic understanding and discovery. However, in the cell there are many complex interactions and mechanisms that operate around DNA in order to translate genetic information into biological function.

Sequencing drives use of sample prep, QA/QC, concordance, validation, and cell culture products, to name a few.

As the current second generation sequencing market grows, we believe it will drag many other tools' markets along with it. For example, sequencing drives use of sample prep, QA/QC, concordance, validation, and cell culture products, to name a few. This bodes well for more diversified genomic firms like Life Technologies, in our view.

Fig 55 Total estimated current genetic analysis market (~US$6bn)



Source: Company data, industry literature, Macquarie Capital (USA) April 2010

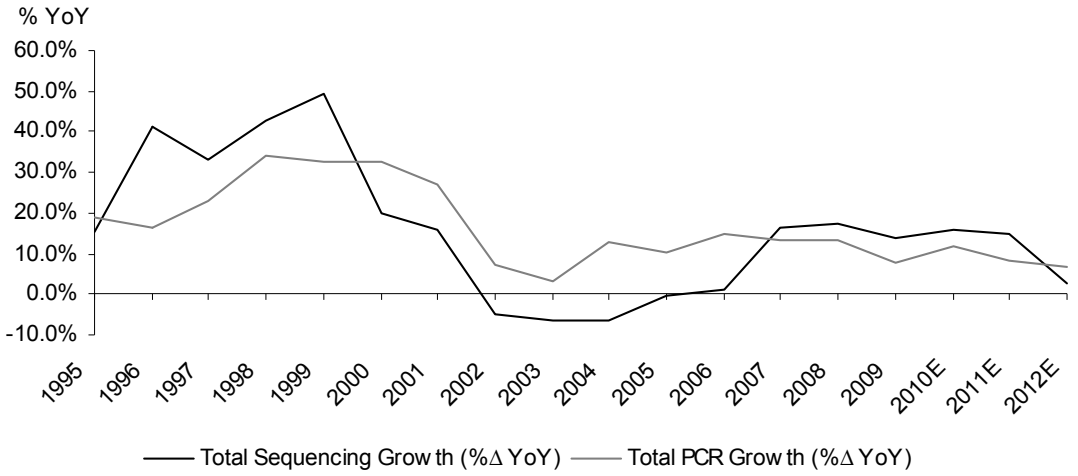
Let us just take an example. Depending on the second generation sequencing application a scientist wanted to run he might use library construction kits, SAGE combo kits, RNA-Seq kits, or enrichment kits along with transcriptome multiplexing kits, whole transcriptome analysis kits, or small RNA expression kits.

Once a sequencing discovery is thought to be made the researcher will want to validate the discovery. To do so he may use microarray technology to study thousands of markers across multiple samples or he may use a lower-plex (meaning looking at fewer genetic markers per sample) technology like qPCR. He may want to look at how the sequence evolves in stem cells or grow and manipulate other cells with the same mutation. He may want to look at proteins.

The last time the sequencing market stalled, the PCR market continued to grow double digits.

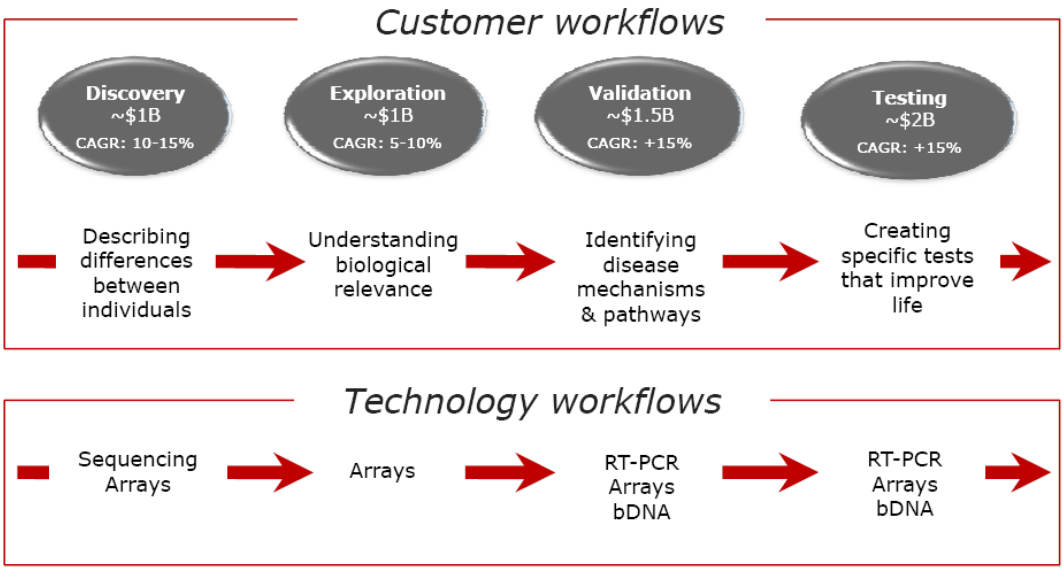
As an example of how sequencing drives downstream work one only needs to look at the last time the sequencing market stalled. As mentioned previously, once CE sequencing machines were in place by 2001, the sequencing market contracted. However, the PCR market, on the other hand, continued to grow at double-digit rates (see Figure 56). PCR is a technology that allows for the amplification and identification of a targeted genomic region of interest.

Fig 56 The PCR market continued to grow after the last sequencing boom



Note: The PCR market is estimated to be roughly twice as big as the sequencing market.
Source: Company data, industry literature, Macquarie Capital (USA), April 2010

Fig 57 Illustrative growth rates in the genetic analysis market (~US$6bn)



Source: Affymetrix, April 2010

Microarrays should actually benefit as well

Despite most predicting its death, another market that is likely to see accelerated growth in the near term is the microarray market. Researchers are likely to continue to take content from the many ongoing sequencing studies and put it onto microarrays so that they can power their studies with larger numbers of samples.

GWAS results have been both encouraging and disappointing, but given the number of samples that need to be run, arrays remain the only practical solution, in our view.

These studies, such as genome wide association studies (GWAS), have been tried before. In GWAS in particular, scientists initially thought that the mutations that caused common diseases would themselves be common so they first identified what they thought were the common mutations in the human population, put them onto arrays, and then compared patients' genomes with those of healthy genomes. The results have been both encouraging and disappointing.

About 2,000 sites on the human genome have been statistically linked with various diseases; however, in many cases the sites are not inside working genes, suggesting there may be some conceptual flaw in the GWAS statistics. In addition, the problem DNA was most often linked to only a small portion of all the cases of the disease.

Now, many researchers believe that it is not the common but rather the rarer variants that are important, believing, for example, that natural selection may have weeded out any disease-causing mutation before it was able to become common. Given the number of samples that need to be run in order to get any kind of statistical power from these studies, microarrays remain the only practical solution, in our view. So sequencing should drive a new round of GWAS studies.

Fig 58 Microarrays are still 1/10th the cost and 1/25th the time of sequencing

	Genotyping (microarray)	Sequencing
Samples	**4000**	**4000**
Dataset	**5M array**	**10x shotgun**
Unit Cost	**$0.5k**	**$5k**
Project Cost	**$4M**	**$40M**
Samples / machine day	**48 / day**	**0.5 / day**
Machine days	**83**	**2000**

Source: Illumina, Macquarie Capital (USA), April 2010

A downstream shift is likely post sequencing technology restatement

In the research community there is also likely to be a shift back to downstream work once the restatement in sequencing technology has taken place. This work would include RNA, protein analysis (e.g. enzymes, metabolites, ligand/receptor sites etc.), stem cells, cell signalling, signal transduction, cell structure, and targeted DNA analysis, to name a few.

Fig 59 Human body and disease mechanisms are extremely complex

Complexity of the human body
100 trillion human cells
216 stem cell lineages
6 billion base pairs of DNA
30,000 genes encoding proteins
10 million total distinct proteins in a person
2000 distinct proteins functioning in a cell
60,000 reactions/cell/minute
100,000s of molecular events
50 or so organs and organ systems

Source: Luminex, NIST, Macquarie Capital (USA), April 2010

Implications: Near term vs long term

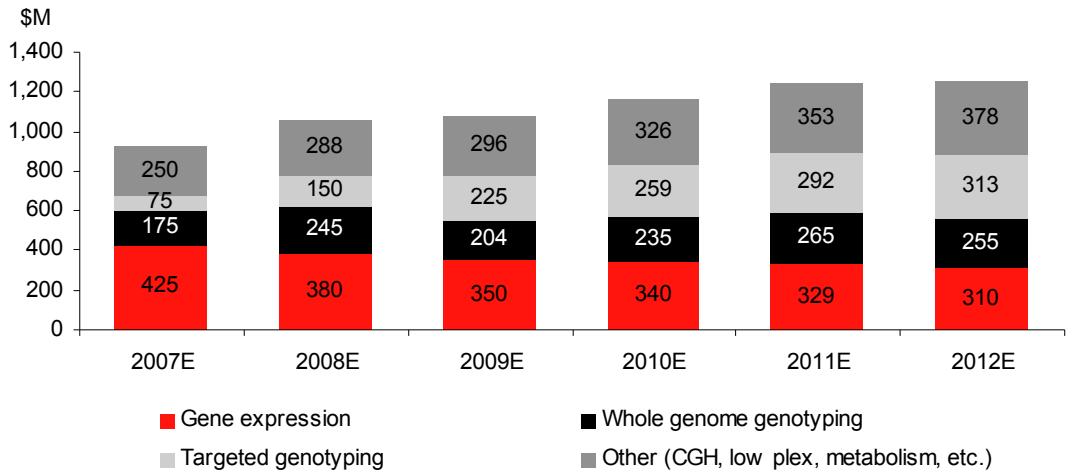
Near term: A rising tide lifts all boats

NIH and other global stimuli will make genetic analysis dollars less competitive than they would otherwise have been.

What does this tsunami of sequencing demand and information mean for genomic technology companies? In the near term, of course, second generation sequencing will be the fastest-growing segment. But growth in sequencing will also drive upstream sample prep, concurrent QA/QC and concordance work, and downstream validation. In addition, extraordinary money flows from the NIH and other global stimuli will make funding more readily available generally, meaning that genetic analysis dollars will be less competitive than they would otherwise have been. Thus, the rising tide should indeed lift all boats, as discussed in the previous section.

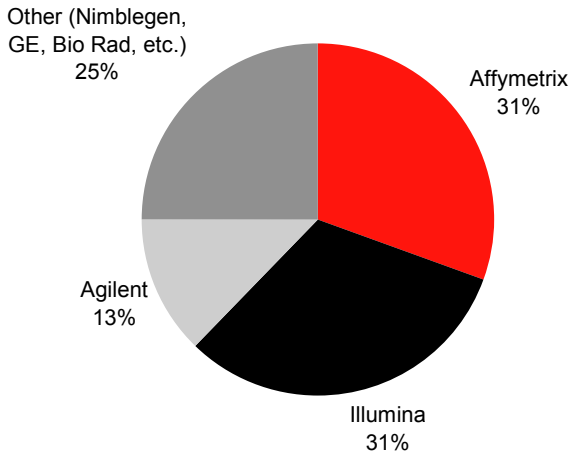
As an example, we estimate that the total microarray market was ~US$1.1bn in 2009, and will grow at a 6% CAGR through 2012, far from disappearing. See Figure 60.

Fig 60 We forecast the total microarray market will grow ~8% in 2010



Source: Company data, industry literature, Macquarie Capital (USA), April 2010

Fig 61 Estimated microarray market share



Source: Company data, industry literature Macquarie Capital (USA), April 2010

Implications: AFFX looks undervalued

If investors are paying the same price for ILMN's array business as they are for AFFX, then that means the sequencing business is being valued at just under 10x sales.

If our forecast is correct, then we think AFFX shares remain undervalued. As a thought exercise we look at what investors would be paying for ILMN's sequencing business if we applied AFFX's multiple to ILMN's array business. We recognize that investors may argue as to whether the analysis is fair given differing views on which array platform has more longevity, but still think the analysis is worthwhile; if nothing else, we believe it is thought provoking.

As Figure 62 depicts, if investors are paying the same price for ILMN's array business as they are for AFFX, then that means the sequencing business is being valued at just under 10x sales. We think 9.2x sales is a steep price for a sequencing technology business and instead believe that the array business is likely worth more than 1.2x sales, which bodes well for AFFX.

Fig 62 ILMN SOTP analysis with array business valued at AFFX multiple

	EV/Sales	2010E expected revenue (US$m)	Enterprise Value (US$m)
Array	1.2	389	467
Sequencing	**9.2**	466	4,278
Total	5.5	855	4,744

Note: AFFX currently trades at 1.2x our sales estimate.

Source: Macquarie Capital (USA), April 2010

Long term: Sink or swim (or float and be pulled out to sea)

This report has shown just how quickly the technology in the genomics space can change, which makes it extremely difficult to predict what will happen over the long term to the many different firms and technologies currently participating or attempting to participate within the genetic analysis space. How technologies evolve, what the actual underlying biological discoveries turn out to be, and how end markets evolve (e.g. how clinicians are reimbursed for tests or implications from the Myriad patent decision) will play a large role in determining their fate. What we do know is that the technologies **will** need to evolve.

One key principle and one key (unanswered) question

In our view there is one key principle and one key question that together will go a long way in determining the fate of many of these companies.

In our view, there is one key principle and one key as-of-yet unanswered question that will go a long way in determining the fate of many of these companies. The principle is this: Customers care about the genetic information, not the technology. In as much as companies confuse the two, they are at risk of becoming obsolete. The key question is this: In what applications will the information of the entire genomic sequence be needed versus only targeted regions of interest and what is the best way to access and deliver that information?

In many ways the two are interrelated as ultimately customers will want the technology that provides the relevant information with the simplest workflow. For example, it may turn out that we only need to sequence an individual once, store the information electronically, and then simply scan the individual's genomic database as needed. However, it also may turn out that other genomic information is more important than simply the sequence and that other technologies are better suited to obtain the relevant information easily and reliably.

The winners and losers ultimately might not be who you expect

The knee-jerk reaction when one considers this report is to think of those companies that are providing the raw genetic horsepower. And for now, as has been highlighted, they will surely benefit. However, as we have highlighted, we think it is highly likely that the raw genetic data itself will ultimately become undifferentiable, meaning that no one will care where the data comes from or how it is generated (accuracy will be assumed), and that more value will ultimately accrue to those who create easy-to-use, relevant applications or those who enable the simple manipulation and interpretation of the massive amounts of information.

Researchers are already intimating that the horsepower game is nearing an end.

As mentioned earlier some are already intimating that the horsepower game is nearing an end. Remember the quote from Eric Green, "Data analysis is the rate-limiting factor in genomics. Not data generation[23]."

Data analysis is the rate limiting factor simply because there is so much of it. And the amount of data will only keep growing. For example, on the leading platform today it takes over 600GB to align and map a human genome at 30x coverage. And that is just the genomic information, which on its own is fairly useless. Thus, the level of data intensity for "precision medicine" will be enormous as the data demands will likely be measured in Terabytes (TB, ~1 trillion bytes) or potentially even Petabytes (PB, TB x 1,000). Depending on how the technology and its associated IT infrastructure demands evolve, this could create huge data management challenges and could be the largest barrier to entry for new laboratories and wider adoption.

In addition, within Human Health medical practice moves at a glacial pace so until clinicians understand why the genomic information is relevant, how to protect it, and why their patients will benefit from it, clinical adoption will surely lag the research. This creates an opportunity for those in a position to influence physician and clinical decision making.

Looking beyond tools: Are PBMs best positioned to drive precision medicine?

While the promise of precision medicine is compelling, there are plenty of challenges that must be overcome to make it a reality. To begin with, compelling, relevant data must exist and getting that data will not be easy. In addition, the various stakeholders within healthcare must work together. However, pharmaceutical companies want to sell more, not fewer, drugs; payers want to save, not spend money; providers want better results, but do not want to get sued; and patients want better care, but are wary of new tests. While targeted medicine could be aligned with each of these interests, no one wants to move first.

A key intermediary that has emerged over the years between these groups is the pharmacy benefit manager, or PBM. PBMs emerged from the smoke in the days of rapidly rising prescription drug costs and largely took over the prescription drug industry 15 years ago. They became the key link in the process of getting the right drug, to the right patient, at the right cost. Effectively, PBMs make money by saving their customers money.

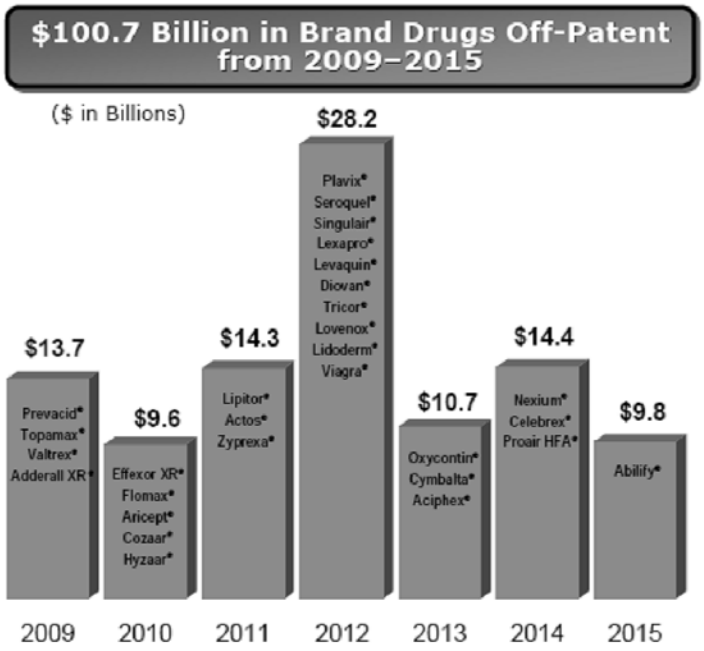
PBMs have become an integral part of the healthcare system and, in particular, have been instrumental in driving a change in prescription methodologies.

PBMs negotiate cheaper drug prices and develop tools to improve the prescription process. They have become an integral part of the healthcare system and, in particular, have been instrumental in driving the use of generic drugs, which typically sell for 20-80% of the price of branded drugs. Importantly, they have access to a tremendous amount of drug and outcome data.

However, PBMs could face growth challenges beyond 2012 as generic sales peak (~US$100bn in branded drugs come off patent from 2009-2015, with the peak being in 2012). See Figure 63.

[23] *Current Topics in Genome Analysis 2010* lecture series

Fig 63 2012 will mark the peak in Brand drugs coming off patent



Source: Medco, April 2010

Genetic testing may be the next growth opportunity PBMs have as they seek to continue to guide prescriptions and save their customers money. In fact, two of the largest PBMs recently moved in this direction. Medco recently acquired DNA Direct Inc., a genetic testing company in San Francisco, and in December, Medco's larger rival, CVS Caremark, increased its investment in Generation Health.

We think that PBMs may be able to use their unique position within the healthcare ecosystem to both drive and benefit from genetic testing.

As an example of how PBMs could direct this process let us take a look at the blood thinner Warfarin. As mentioned earlier, Warfarin is difficult to dose. In 2005 an FDA advisory committee recommended that genetic information be considered in making treatment decisions with Warfarin. Medco decided to explore personalized drug treatments that year. As Medco looked into its database, it found that of its million patients on the drug, a quarter of them ended up in the hospital within six months of starting on Warfarin.

The company kicked off its program in earnest in May 2008 and has reportedly already signed up more than 200 employers whose clients cover the health care of 7 million people. Medco is able to direct doctors who are prescribing Warfarin to have their patient genetically tested as to whether they under- or over-metabolize the drug. According to the firm, avoiding one hospitalization could underwrite the cost of the test for 100 patients.

Currently Medco offers tests for just two drugs where it expects the best results – the blood thinner Warfarin, as mentioned, and the breast cancer drug Tamoxifen. However, the company aims to eventually expand the program to more medicines, for example, potentially Plavix.

Given the early experience of Medco, we think that PBMs may be able to use their unique position within the healthcare ecosystem to both drive and benefit from genetic testing.

Investment opportunities

Don't be fooled by the secular theme; stay nimble

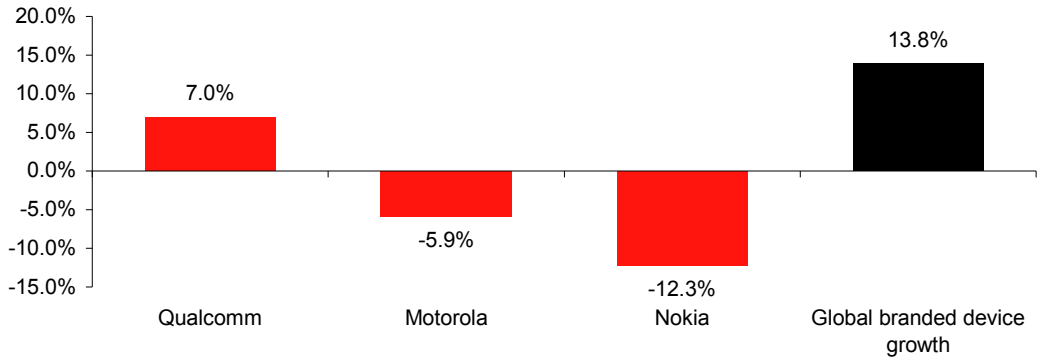
While we think that in the near term most genome technology firms will benefit, we think long-term investors need to be nimble with their investment allocations.

All one has to do is look at the popular press to see that genetics is already becoming a big theme. Recently, the top two articles in Bloomberg's Science section were, *Personalized Genetic-Based Medicine Spurred by Medco's Cost-Saving Tests*, and *Family DNA maps May Speed Discoveries of Rare Disease Links, Doctor's Say*. The same day in the *New York Times* there was an article entitled, "*Disease Cause is Pinpointed with Genome*."

As has already been discussed, in the near term second generation sequencing will be the fastest-growing segment within genomic tools. This growth in sequencing will drive upstream sample prep, concurrent QA/QC and concordance work, and downstream validation. In addition, extraordinary money flows from the NIH and other global stimuli will make funding more readily available generally, meaning that the demand for genetic analysis dollars will be less competitive than it would otherwise have been. We have laid out a number of investment ideas that we think make sense in the near term.

Long term we think it is clear that there is tremendous demand for relevant genomic information. However, as anyone who invested in the major cell phone manufacturers knows, stocks generally anticipate large markets well before they actually emerge. See Figure 64.

Fig 64 Stock performance versus growth in branded cell phones (CAGR 2001-2009)



Source: FactSet, Macquarie Capital (USA), April 2010

Thus, while we think that in the near term most genome technology firms will benefit from the wave of genetic information that will engulf us, we think long-term investors need to be nimble with their investment allocations given the rate of technological change.

Appendix

In the following appendix we discuss the science of DNA and DNA sequencing technologies in greater detail and in a manner that we hope the casual reader can understand. We also include a glossary of technical terms that have been used within the paper.

DNA sequencing technology

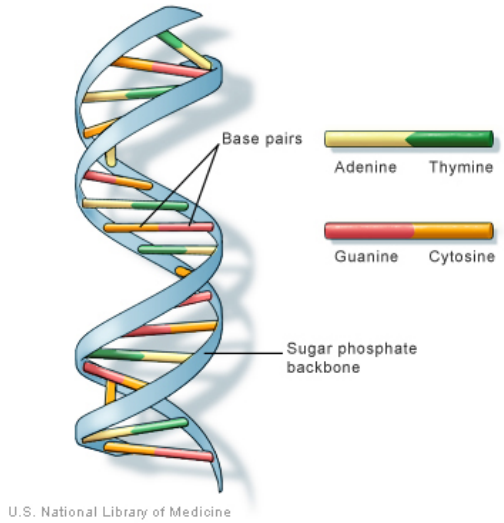
A brief explanation of the human genome

To understand DNA sequencing we need to first start with an explanation of DNA itself. DNA (Deoxyribonucleic Acid) resides in the nucleus of every cell and can most easily be thought of as the instruction manual of life.

How does DNA provide the information necessary for a living organism to grow and live? DNA is analogous to the binary code, a series of 0s and 1s, software programs use to run a computer. Instead of 0s and 1s, the DNA molecule uses sequences of four nucleotide bases: Adenine, Cytosine, Guanine, and Thymine or commonly denoted A, C, G, or T. There are an estimated ~6 billion nucleotide base pairs in the full human genome.

A critical element in understanding how DNA works is in its physical structure, known as a double helix (first proposed by James Watson and Francis Crick in 1953), where two single strands of nucleic acids wrap around each other. Each strand can be thought of as a linear ~3.1 billion nucleotide base (A, C, G, or T) chain where individual nucleic acid molecules are attached together by a sugar-phosphate backbone or spine. The two individual strands that wrap around each other are bound together by aligning mirror-image nucleotides, where an A on one strand will only bond to a complementary T on the opposing strand and G will only bond to a complementary C (See Figure 65 for an illustration).

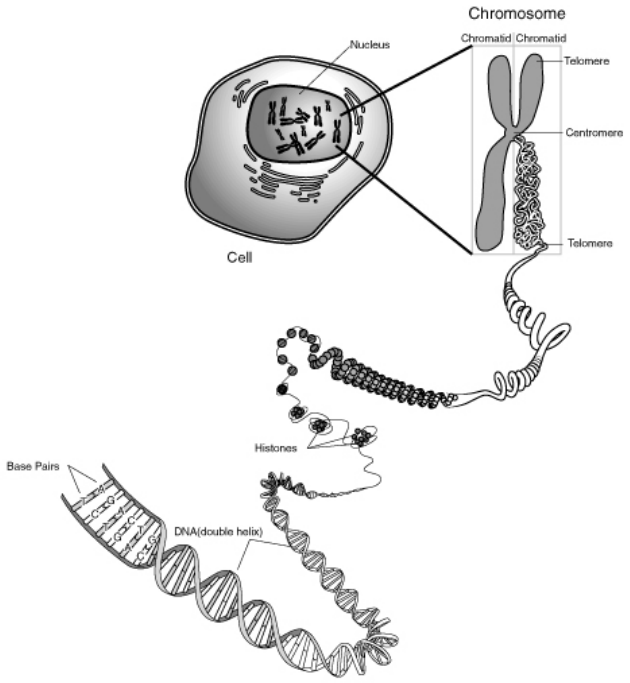
Fig 65 Illustration of the DNA molecule structure (two strands creating a double helix)



Source: U.S National Library of Medicine, April 2010

DNA found within the human body is actually a packaged collection of 46 individual double helix strands, known as chromatids, contained within the nucleus of each cell. Each chromatid is paired with a mate-strand, forming a chromosome, and each individual strand of a chromosome is adopted from one parent (one from the mother and one from the father). There are 23 mate-pairs (46 divided by 2), with 22 homologous pairs consisting of a chromosome from each parent (autosomes) and one mate-pair that determines gender (X & Y sex chromosomes). Within each chromosome there are large defined pieces of the DNA molecule that are known to code for proteins and are called genes, it is estimated humans have ~23,000 genes (the number is even now debated). There are also non-gene components of the DNA molecule that influence overall activity or maintain structure, such as a centomere, which is where a pair of chromatids 'connect'. Figure 66 is an illustration of the DNA molecular 'package' found in humans.

Fig 66 Structure of the DNA 'package' found in humans



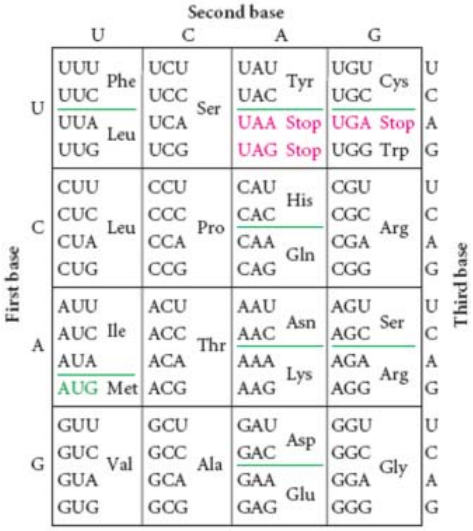
Source: American Mathematical Society, April 2010

How DNA works, in a nutshell

The human body runs off of a class of molecules called proteins that are involved in structural formation, muscle movement, chemical reactions such as digestion and hormones (cellular communication) to name a few. Proteins are basically long chains of amino acids manufactured by cells. Each amino acid corresponds to a string of three nucleotides (nucleic acids A, C, G and T) found in the DNA molecule, called a codon. In this way, a given A, C, G and T nucleotide sequence of DNA can instruct a cell to create proteins that ultimately define the basic biology of the human body.

Figure 67 below is a look-up table used to relate 3 nucleotide base strings (codons) to specific amino acids; you may notice an absence of the letter T, this is because T is replaced with a U (Uracil) in the actual message (RNA) sent to the protein manufacturing process within a cell. We discuss this process in greater detail in the following paragraph.

Fig 67 Codons and related amino acids



Source: www.nature.com, April 2010

The actual process of DNA coding for proteins is complex, and is still being studied. When a segment of DNA begins the process of creating a message for the cell to build a new protein, a process known as transcription, the bonds between a section of two individual strands that form the double helix are broken apart and a molecule called RNA (Ribonucleic Acid) is formed. RNA is a string of nucleotide bases that are mirror images of the template DNA, where A in the template DNA will code for a corresponding U (Uracil replaces Thymine in a RNA molecule) in the RNA molecule and G in the template will code for C in the RNA molecule (please note the base coding can happen in reverse order as well, except that the T in the template DNA would code for an A in the RNA molecule). When the single-stranded RNA molecule coding (transcription) is finished, the molecule detaches from the template DNA and the parted strands of DNA rejoin to create the double-helix structure.

Once transcription is completed, a process within the cell called translation occurs, whereby a cellular structure known as the ribosome builds (by linking amino acids) a protein based on the 'code' of the single-stranded RNA molecule. Formed proteins influence and are used in a broad range of biological processes, as described earlier.

Fig 68 The central dogma of biology

Source: American Mathematical Association, April 2010

Examples of genetic mutation

A mutation is any change in the DNA sequence of a cell's genome. Mutations are important because they can impact the cellular processes described above.

There are two types of genetic variation that impact an individual's genetic make-up. Germline variations occur at conception and are inherited from parents. Somatic variations are those mutations that occur as cells replicate and accumulate over a person's lifetime.

There are many ways to think about mutations, but one helpful way is to consider the scale at which the mutation occurs.

12. Small-scale mutations are those that affect a small gene in one or a few nucleotides. Examples include:

⇒ **Point mutations:** The exchange of a single nucleotide base for another. Often caused by chemicals or malfunction of DNA replication, these changes are classified as transitions or transversions. Most common is the transition that exchanges a purine for a purine (A ↔ G) or a pyrimidine for a pyrimidine, (C ↔ T). Less common is a transversion, which exchanges a purine for a pyrimidine or a pyrimidine for a purine (C/T ↔ A/G). An example of a transversion is Adenine (A) being converted into a Cytosine (C).

⇒ **Insertions:** The addition of one or more extra nucleotides into the DNA. They are often caused by transposable elements, or errors during replication of repeating elements (e.g. AT repeats). Insertions in the coding region of a gene may alter splicing of the mRNA or cause a shift in the reading frame, both of which can significantly alter the gene product.

⇒ **Deletions:** The removal of one or more nucleotides from the DNA. Like insertions, these mutations can alter the reading frame of the gene.

13. Large-scale mutations affect the larger structure of the chromosome. Examples include:

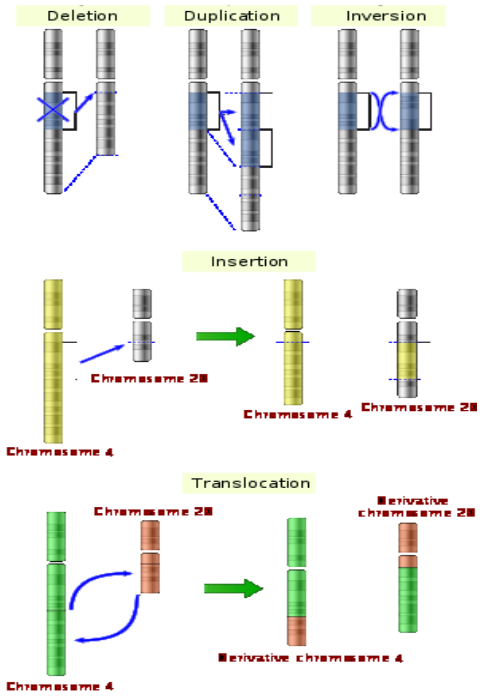
⇒ **Amplifications (or gene duplications):** Multiple copies of large or all chromosomal regions.

⇒ **Deletions:** The loss of large chromosomal regions, leading to loss of the genes within those regions.

⇒ Mutations whose effect is to juxtapose previously separate pieces of DNA (potentially bringing together separate genes to form functionally distinct fusion genes) such as:

- **Chromosomal translocations:** The interchange of genetic parts from nonhomologous chromosomes.
- **Interstitial deletions:** An intra-chromosomal deletion that removes a segment of DNA from a single chromosome.
- **Chromosomal inversions:** The reversal of the orientation of a chromosomal segment.
- **Loss of heterozygosity (LOH):** The loss of an allele, either by a deletion or recombination event, in an organism that previously had two different alleles.

Fig 69 Examples of DNA mutations



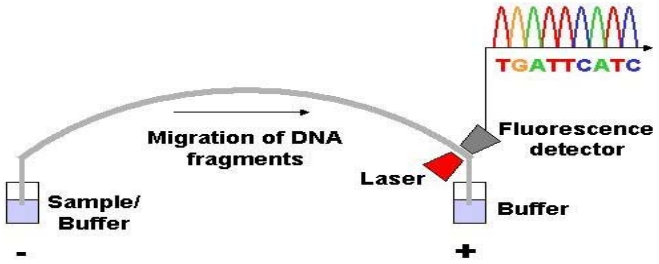
Source: U.S National Library of Medicine, April 2010

Legacy sequencing technologies: Capillary Electrophoresis (CE)

CE is considered to be the gold standard in sequencing technology. It was developed in the mid 1990s and first introduced by Applied Biosystems on an instrument called the ABI PRISM 310. Although having a much lower throughput compared to newer sequencing technologies available on the market today, CE sequencing is valued for its longer read lengths and the greater genomic structure accuracy they provide. Below are the very basic technological steps of the process:

1. **Creating a library of smaller fragments:** Multiple copies of the target DNA are broken into millions of small pieces through shearing (needle) or sonification (sound). The DNA fragments are then separated by size, and the fragments of desired size are packaged into loops of bacterial DNA, known as plasmids. As the bacteria cells grow and divide, they replicate the plasmid (and target) DNA as well as their own DNA.
2. **Purification:** Detergents plus physical separation techniques are used to release and purify the target DNA from the plasmid into sample vials.
3. **Amplification (PCR):** The sequencing mixture of plasmid and target DNA, DNA nucleotides, fluorescently labeled (modified) nucleotides, enzyme and 'primer' sequences is then put onto a heating block. A cycle of heating and cooling (Polymerase Chain Reaction) is repeated many times, generating a large number of fragments, of different lengths, that end in fluorescently labeled bases.
4. **Reading the sequences:** Vials containing many copies of each target DNA fragment are then loaded into a sequencing machine. Inside the machine, the samples are transferred into thin capillaries. An electrical charge migrates the negatively charged DNA molecules through a gel matrix. As the DNA fragments move along the gel matrix, longer DNA fragments are slowed down more by the gel than the shorter fragments. A laser at the end of the capillary is used to excite the final fluorescent base, which is recorded as a colored peak or bar.
5. **Analysis:** Since the computer only has fragments of the total template DNA strand, a mathematical algorithm is applied to estimate the most likely sequence based on areas of overlapping DNA code between the fragments. The construction of the template DNA sequence may employ a reference genome, if a similar target DNA has been sequenced, validated, and recorded by a previous experiment (the Human Genome Project provided researchers with the first human reference genome).

Fig 70 Illustration of CE sequencing



Source: DNA Sequencing Service, April 2010

Massively parallel sequencing or "Second" generation sequencing

As you will see below, steps 1 & 2 in 'Next generation' sequencing are conceptually the same as CE sequencing, the term coined for this sequencing method is 'shotgun' sequencing, due to the process of somewhat randomly blowing apart (through shearing, sonification or digesting) the target DNA and then sequencing the smaller fragments. The main difference between 'Next generation' sequencing and CE sequencing is that the smaller fragments of the template DNA are sequenced at the same time (e.g. in parallel) attached to a planar surface (ILMN) or bead (Life and 454) versus singularly being fed through capillaries (the number of capillaries used by a CE machine directly relates to throughput capability).

1. **Creating a library of smaller fragments:** Multiple copies of the target DNA are broken into millions of small pieces through shearing (needle) or sonification (sound). The DNA fragments are then separated by size and the fragments of desired size are packaged into loops of bacterial DNA, known as plasmids. As the bacteria cells grow and divide, they replicate the plasmid (and target) DNA as well as their own DNA.
2. **Purification:** Detergents plus physical separation techniques are used to release and purify the target DNA from the plasmid into sample vials.
3. **Amplification (PCR):** Adaptors and primers are attached to the DNA fragments, that are then attached to a surface such as a slide or bead coated with matching adaptors. Once the millions of DNA fragments are connected to the respective surface, amplification exponentially increases the sample, creating large colonies. Amplification of the DNA fragments is necessary to create colonies of the sample DNA large enough to produce a visible light signature for the camera during the 'reading' process.

   ⇒ Current next generation sequencing technologies are based on the fact that with current optic technologies you need large numbers of copies of DNA in order to see the dyes that represent the individual bases as they bind to, and extend, the target DNA.
4. **Reading the sequences:** Several different methods are used in 'Next generation' sequencing to detect the addition or removal of a single nucleotide (A, C, G, or T) to/from the template DNA fragment. However, the process is similar across platforms in that a laser excited fluorescent label denoting a single nucleotide (or di-nucleotide in the case of ABI's SOLiD platform) is read by a camera (optically) to record the DNA sequence.

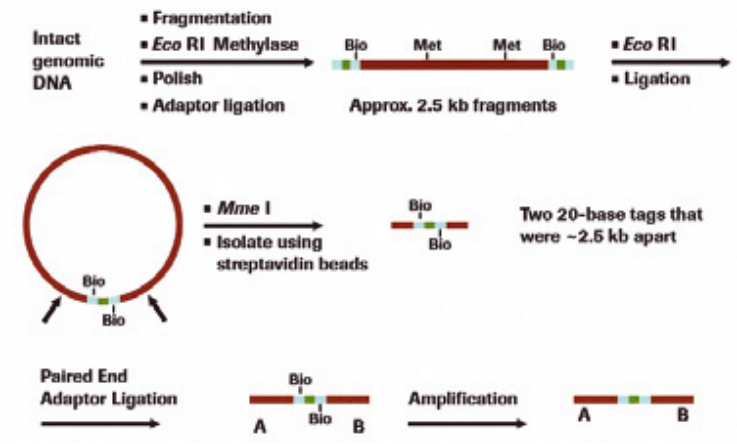
   ⇒ Once the slide is ready, it is put into the sequencing instrument. The instrument then goes through a number of cycles in which bases are added, pictures are taken, dyes are removed, and the process starts all over again. Because this is happening at millions of points on the slide at the same time the process is called massively parallel.
5. **Analysis:** Since the computer only has fragments of the total template DNA strand, a mathematical algorithm is applied to estimate the most likely sequence based on areas of overlapping DNA code between the fragments. Again, the construction of the template DNA sequence may employ a reference genome, if a similar target DNA has been sequenced, validated, and recorded by a previous experiment. The data analysis process is more onerous for 'Next generation' sequencing versus CE sequencing as the template DNA fragments are much shorter.

   ⇒ Because DNA is lost at every step in the process and during each sequencing cycle, at some point the sequencer optics can no longer see the dyes and accuracy begins to diminish. That becomes the limit of your read length.

Paired-end reads

Due to the difficulty of constructing a DNA sequence with shorter template DNA fragments read by 'Next generation' sequencers, a process called paired-end reads is often used. This technique enables *de novo* sequencing and structural variation analysis to be performed with shorter read sequencers (ABI and ILMN in particular). The basic premise is that you can determine the location of a DNA fragment within the larger template DNA sample by only sequencing both ends of a given DNA fragment. This technique allows sequencers to increase the linear coverage, and thereby increase the theoretical accuracy of the template DNA sequence, without increasing the physical number of bases read per fragment. All current 'Next generation' sequencing platforms along with CE sequencers offer protocols for paired-end reads.

Fig 71 Example of paired-end read sample prep for Roche 454 system



Source: Roche 454, April 2010

How the competing instruments stack up

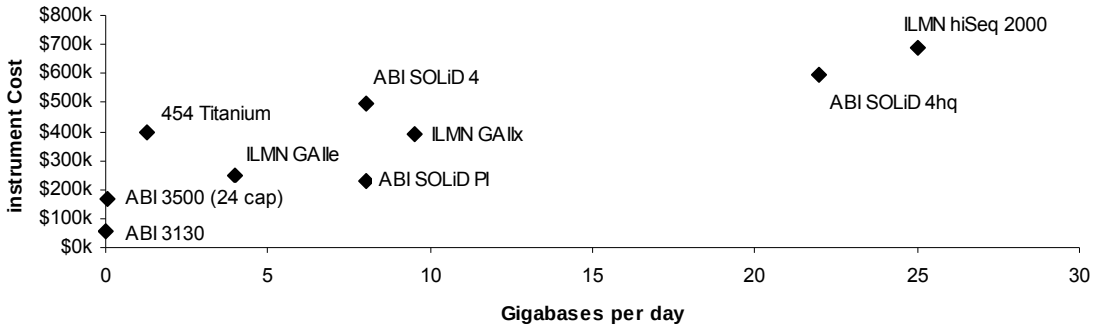
Figures 72-75 compare the most broadly adopted DNA sequencer instruments currently on the market (used in the major research labs). We have omitted several machines, either due to low adoption rates or lack of comparable information.

Fig 72 Second generation sequencing machines

	ILLUMINA			LIFE Technologies			454 XLR
	GAIIe	GAIIx	HiSeq 2000	SOLiD PI	SOLiD 4	SOLiD 4hq	Titanium
Price	US$250k	US$390k	US$690k	US$230k	US$490k	Upgrade kit for SOLiD 4	US$400k
Yield per run (Paired-end run):	35Gb	50Gb	200Gb	50Gb	100Gb	300Gb	~0.5Gb
Time per run (Paired-end run):	~10 days	~10 days	~8 days	~14 days	~14 days	~14 days	~10 hours
Accuracy:	*	~99%	*	*	~99.99%	*	*
Read length (Paired-end):	2x100	2x100	2x100	2x50	2x50	2x50	400
Cost per human genome:	~US$20k	~US$20k	~US$10k	*	~US$6k	~US$3k	~US$150k
Annual consumable usage:	100-125	150-200	300-400	*	$120 - $125K	$120 - $125K	~US$750k

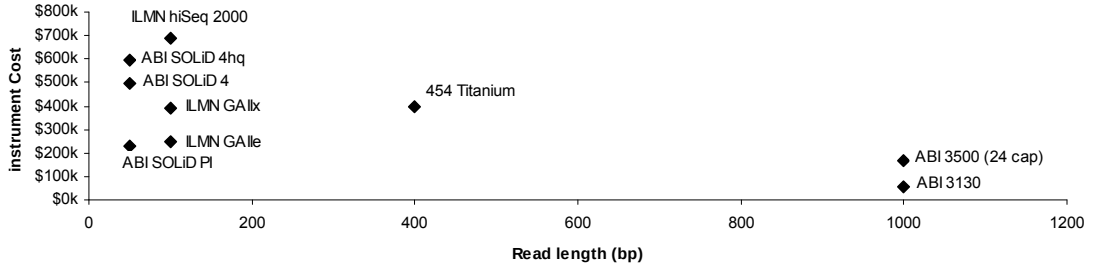
Source: Company data, Macquarie Capital (USA), April 2010

Fig 73 Sequencer instrument cost versus output per day



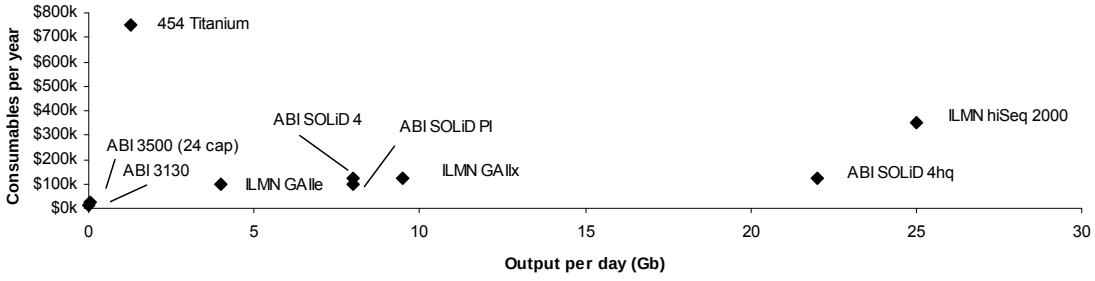
Source: Company data, Macquarie Capital (USA), April 2010

Fig 74 Sequencer instrument cost versus read length



Source: Company data, Macquarie Capital (USA), April 2010

Fig 75 Sequencer consumable usage per year versus output per day



Source: Company data, Macquarie Capital (USA), April 2010

Single-molecule and other "third" generation sequencing technologies

Even with the ongoing rapid adoption of second generation sequencers (and many technological question marks), the excitement surrounding 'Single-molecule' sequencers is palpable. Third generation DNA sequencing technology has been loosely defined as a sequencer that has a dramatically longer read length, and reduces or eliminates the need for DNA fragment amplification. Another point of differentiation between some third generation technologies will be the use of a camera as some instruments are vying to electronically, rather than optically, read nucleotide bases in the DNA sequence via an electrical current. See Figure 76 for a list of known third generation sequencers.

Fig 76 Third generation 'Single-molecule' sequencers

Company	Sequencer	Mechanism of Action	Output/Cost	Instrument Cost	Launch date
Pacific BioScience	**SMRT DNA sequencer**	**Optically read base identification through emitted wave-length during nucleotide incorporation.**	**?**	**~US$750k**	**Beta systems and early commercialization 2010**
Ion Torrent	**Ion Personal Genome Machine**	**Electronic detection of the release of hydrogen ions following an incorporation of a nucleotide in the DNA sequence.**	**~1 run per hour US$300-500**	**~US$50k**	**No date**
Illumina	**Oxford Nanopore**	**Electronic detection as DNA passes through tiny protein holes on a silicon chip.**	**?**	**?**	**No date**
Life Technologies	**Quantum dot**	**Optically read base identification from light emitting nanocrystals attached to DNA polymerase molecules.**	**~1k - 1.5k base-pairs read length**	**?**	**Beta systems ~2010**
Helicos	**Heliscope**	**Sequencing process similar to 'Next generation' sequencing (sequencing by synthesis) although no amplification is required.**	**~US$48k per genome**	**~US$1m**	**~2007, only 10 placements estimated to date**

Source: Company data, Macquarie Capital (USA), April 2010

Sequencing as a service

Sequencing as a service is also available today and is offered by well over 50 labs and companies worldwide using second generation or CE sequencers. However, a privately funded company named Complete Genomics has developed its own sequencing technology and designed a business model to solely provide sequencing services (not instrument sales). The company expects to be fully operational within the first half of 2010, has ~30 customers and plans to sequence 10,000 genomes in the first 12 months. Currently, Complete Genomics offers whole genome sequencing at a rate of ~US$5k per genome for bulk orders.

Fig 77 Valuation, risks and catalysts

AFFX, US$7.64, Outperform	Our US$12 target price is based on a target multiple of 2x our CY2010 sales estimates.	Risks to our rating and target price are management's inability to execute according to plan, including a failure to adequately cut costs and integrate recent acquisitions; a faster-than-expected decline in its core gene expression products; and dilutive corporate initiatives.	Quarterly earnings results, scientific publications/other announcements.
ILMN, US$39.26, Neutral	Our US$38 target price is based on a target multiple of 28x our C2011 earnings estimates.	Risks to our rating and target price are decreased demand for genetic analysis solutions leading to lower-than-expected revenue growth, management's failure to manage growth appropriately, increased competition, the introduction of new and potentially disruptive technologies and potential volatility around reported quarterly results.	Quarterly earnings results, scientific publications/other announcements.
LIFE, US$52.67, Outperform	Our US$60 target price is based on a target multiple of 16x our C2011 earnings estimates.	Risks to our rating and target price are a greater-than-expected decline in pharma and biotech research spending, failure to execute on mergers and acquisitions, market share losses and the potential negative consequences of LIFE's relatively highly leveraged balance sheet.	Quarterly earnings results, scientific publications/other announcements.

Note: Priced as of April 5, 2010.
Source: Macquarie Capital (USA), April 2010

Glossary of terms

Fig 78 Definitions of scientific and other terms used in this report

1000 Genomes Project	A project to discover genetic differences of at least 1% in occurrence across 2000 unidentified people from about 20 distinct populations globally. Next-generation sequencers were used, and results of the study would be freely accessible to researchers worldwide. Pilot studies were conducted prior to the main study and were concluded in 2008 and 2009. Sequenced sample data from the main study of 2000 genomes has been released intermittently between 2009 and 2010, and is projected to conclude in 2011.
Adaptor	A DNA fragment attached to the ends of the DNA targeted to be sequenced that promote binding (for example to a flow cell), amplification, or sequencing.
Adenine	A purine nucleobase that is one of the four main building blocks of DNA, and is complementary to Thymine in DNA or Uracil in RNA.
Agri-genomic	Genomic analysis applied to the development of agriculture and livestock.
Algae	Photosynthetic organisms that occur abundantly in nature ranging from single-celled to multi-cellular species.
Allele	An allele is one of two or more versions of a gene. An individual inherits two alleles for each gene, one from each parent. If the two alleles are the same, the individual is homozygous for that gene. If the alleles are different, the individual is heterozygous. Though the term "allele" was originally used to describe variation among genes, it now also refers to variation among non-coding DNA sequences.
American Society of Human Genetics	A large (~8,000 member) professional membership organization for human geneticists in North America, founded in 1948.
Amino acid	Amino acids are a set of 20 different molecules used to build proteins. Proteins consist of one or more chains of amino acids called polypeptides. The sequence of the amino acid chain causes the polypeptide to fold into a shape that is biologically active. The amino acid sequences of proteins are encoded in the genes.
Amplification (or gene duplication)	An increase in the number of copies of a gene sequence, and commonly associated with mutations found in cancer cells.
Antibody	An antibody is a protein component of the immune system that circulates in the blood, recognizes foreign substances like bacteria and viruses, and neutralizes them. After exposure to a foreign substance, called an antigen, antibodies continue to circulate in the blood, providing protection against future exposures to that antigen.
Artemisinin	A malaria drug derived from a herb.
Bacterial Artificial Chromosome (BAC)	An engineered DNA molecule used to produce multiple copies of a desired DNA sequence by using replication of bacterial cell growth.
Base Pair	Two chemical bound together by a hydrogen bond that create a "rung of the DNA ladder", Adenine A pairs with Thymine T and Cytosine C pairs Guanine G.
Biofuel	An energy source at least 80% derived from renewable materials.
Biomarker	A biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease. A biomarker may be used to see how well the body responds to a treatment for a disease or condition. Also called molecular marker and signature molecule.
Bioprocess	Manufacturing a product through the use of living cells. Most commonly used within the biotech industry to produce drugs.
Broad Institute	Created in 2004 through the philanthropic investment of Eli and Edythe Broad as a collaboration of Harvard and its affiliated hospitals, MIT, and the Whitehead Institute.
Byte	A unit of digital information used in computing, comprised of 8-bits.
Cancer Genome Atlas	The National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) launched The Cancer Genome Atlas (TCGA) program to create a comprehensive atlas of the genomic changes involved in more than 20 common types of cancer. This large-scale, high-throughput effort is being carried out by a network of more than 100 researchers at many organizations across the nation. The overarching goal of TCGA is to further scientific understanding of the genomic changes in cancer, thereby improving our ability to diagnose, treat and prevent this devastating disease.
Capillary electrophoresis (CE)	A process using fused-silica capillaries and an electrical current to separate DNA fragments of different sizes for the purpose of sequencing.
Cellular pathway	Complex sequences of proteins and other molecules that can change some aspect of cell behavior, potentially causing disease. A focus of genetic cancer research.
Cellulose	Main component of the cell wall in green plants and partially digestible by humans.
Cetuximab (Erbitux)	A monoclonal antibody that is an epidermal growth factor receptor inhibitor.
Charcot-Marie-Tooth disease	A group of genetic nerve disorders that result in problems with movement and sensation.
Chemotherapy	Taking various drugs to treat cancer, with or without radiation treatment or surgery.
CHiP-seq	Used to associate protein production with specific parts of the genome.
Chloroquine	Drug used in the treatment or prevention of malaria.
Chromosomal inversions	A segment of a chromosome is reversed end to end.
Chromosomal translocation	Rearrangement of pieces of DNA between non-homologous (non-identical) chromosomes.
Chromosome	A chromosome is an organized package of DNA found in the nucleus of the cell. Different organisms have different numbers of chromosomes. Humans have 23 pairs of chromosomes--22 pairs of numbered chromosomes, called autosomes, and one pair of sex chromosomes, X and Y. Each parent contributes one chromosome to each pair so that offspring get half of their chromosomes from their mother and half from their father.
Clinical lab	Perform tests to detect, diagnose, and help with the treatment of disease. Commonly found in hospitals.
Cluster density	The proximity of DNA colonies placed on a flow cell.
Coding	Coding DNA are regions of the genome that have been found to code for amino acids.
Codon	A codon is a trinucleotide sequence of DNA or RNA that corresponds to a specific amino acid. The genetic code

Fig 78 Definitions of scientific and other terms used in this report

	describes the relationship between the sequence of DNA bases (A, C, G, and T) in a gene and the corresponding protein sequence that it encodes. The cell reads the sequence of the gene in groups of three bases. There are 64 different codons: 61 specify amino acids while the remaining three are used as stop signals.
Colorectal cancer	Cancer that starts in the large intestine (colon) or the rectum.
Companion diagnostic	A diagnostic test that characterizes a patient's potential response to a specified drug based on their genome. In 1998, the first 'Companion Diagnostic' test was released to identify breast cancer patients that would respond to Genentech's Herceptin.
Comparative-effectiveness	Comparing two or more treatments for a health problem to determine the optimal treatment based on the costs and benefits.
complementary DNA	DNA in which the sequence of the molecules (G, A, C, T) on one strand of the double stranded structure chemically matches the sequence on the other strand.
Consensus determination	The process of using multiple samples during sequencing to come to a best estimate as to the true identity of a given base.
Copy number change	Gains or losses involving the copy number of specific DNA sequences.
Corn-based ethanol	A type of fuel derived from corn.
Cost-per-genetic data point	The monetary cost of defining one base within a DNA sequence.
Coverage	The number of times a full DNA genome is sequenced during an experiment. Due to the degree of uncertainty surrounding current sequencing technologies in constructing the linear order of bases of a full human genome, the process is generally repeated multiple times and then an algorithm is applied to determine the most likely structure.
Cytogenetic	The study of the structure and function of the cell and specifically chromosomes. Cytogenetic testing is used to identify individuals with chromosomal disorders such as Down's syndrome.
Cytosine	Cytosine (C) is one of four chemical bases in DNA, and forms chemical bonds with Guanine bases on the opposite strand of DNA.
De novo sequencing	Indicates that an organism or DNA sample is being sequenced without a reference genome, usually for the first time.
Deletion	Deletion is a type of mutation involving the loss of genetic material. It can be small, involving a single missing DNA base pair, or large, involving a piece of a chromosome.
Diabetes	Diabetes mellitus is a disease characterized by an inability to make or use the hormone insulin. Insulin is needed by cells to metabolize glucose, the body's main source of chemical energy. Type I diabetes, also called insulin-dependent diabetes mellitus, is usually caused by an autoimmune destruction of insulin-producing cells. Type II diabetes, also called non-insulin-dependent diabetes mellitus, occurs when cells become resistant to the effects of insulin.
Diploid	Diploid is a cell or organism that has paired chromosomes, one from each parent. In humans, cells other than human sex cells, are diploid and have 23 pairs of chromosomes. Human sex cells (egg and sperm cells) contain a single set of chromosomes and are known as haploid.
DNA	DNA (Deoxyribonucleic Acid) is the chemical name for the molecule that carries genetic instructions in all living things. The DNA molecule consists of two strands that wind around one another to form a shape known as a double helix. Each strand has a backbone made of alternating sugar (deoxyribose) and phosphate groups. Attached to each sugar is one of four bases--Adenine (A), Cytosine (C), Guanine (G), and Thymine (T). The two strands are held together by bonds between the bases; Adenine bonds with Thymine, and Cytosine bonds with Guanine. The sequence of the bases along the backbones serves as instructions for assembling protein and RNA molecules.
Drought resistant	Generally refers to an agricultural crop that can survive on very little water.
Dye	A label (usually florescent) attached to DNA that helps to distinguish the position of a specific base (G, A, C, T), usually by laser excitement.
Enzyme	An enzyme is a biological catalyst and is almost always a protein. It speeds up the rate of a specific chemical reaction in the cell. The enzyme is not destroyed during the reaction and is used over and over. A cell contains thousands of different types of enzyme molecules, each specific to a particular chemical reaction.
Forensic	The application of science to provide answers to a legal system.
Gene	The gene is the basic physical unit of inheritance. Genes are passed from parents to offspring and contain the information needed to specify traits (the color of your hair). Genes are arranged, one after another, on structures called chromosomes. A chromosome contains a single, long DNA molecule, only a portion of which corresponds to a single gene. Humans have approximately 23,000 genes arranged on their chromosomes.
Gene expression	Gene expression is the process by which the information encoded in a gene is used to direct the assembly of a protein molecule. The cell reads the sequence of the gene in groups of three bases (G, A, C, T). Each group of three bases (codon) corresponds to one of 20 different amino acids used to build the protein.
Genetic Information Nondiscrimination Act (GINA)	The Genetic Information Non-discrimination Act (GINA) is US federal legislation that makes it unlawful to discriminate against individuals on the basis of their genetic profiles in regard to health insurance and employment. These protections are intended to encourage Americans to take advantage of genetic testing as part of their medical care. President George W. Bush signed GINA into law on May 22, 2008.
Genetic Profile	Referring to an individual's genetic code.
Genome	The genome is the entire set of genetic instructions found in a cell. In humans, the genome consists of 23 pairs of chromosomes, found in the nucleus, as well as a small chromosome found in the cells' mitochondria. These chromosomes, taken together, contain approximately 3.1 billion bases of DNA sequence.
Genotype	A genotype is an individual's collection of genes. The term also can refer to the two alleles inherited for a particular gene. The genotype is expressed when the information encoded in the genes' DNA is used to make protein and RNA molecules. The expression of the genotype contributes to the individual's observable traits, called the phenotype.
Germline variation	Genetic variation passed by parent to offspring.
Gigabase (Gb)	1 billion bases of DNA (G, A, C, T).
Gigabyte (GB)	1 billion bytes, units of digital information used in computing.

Fig 78 Definitions of scientific and other terms used in this report

Grosch's Law	The added economy (usually referred to computing) is the square root of the increase in speed. For example if you wanted a computer 8 times as fast, you would need to pay twice what the current computer costs.
Guanine	Guanine (G) is one of four chemical bases in DNA, with the other three being Adenine (A), Cytosine (C), and Thymine (T). Within the DNA molecule, Guanine bases located on one strand form chemical bonds with Cytosine bases on the opposite strand. The sequence of four DNA bases encodes the cell's genetic instructions.
GWAS (Genome Wide Association Studies)	A genome-wide association study (GWAS) is an approach used in genetics research to associate specific genetic variations with particular diseases. The method involves scanning the genomes from many different people and looking for genetic markers that can be used to predict the presence of a disease. Once such genetic markers are identified, they can be used to understand how genes contribute to the disease and develop better prevention and treatment strategies.
Haploid	A cell or organism that only contains a single-set of chromosomes. Usually only found in organisms that reproduce asexually or egg and sperm cells in humans.
Herceptin	Also known as Trastuzumab, a monoclonal antibody used in the treatment of breast cancer.
Homologous chromosome	Homologous chromosomes are the same size, their centromeres are in the same position and they have the same number of genes, arranged in the same order.
Hormone	A chemical messenger used by cells to communicate.
Human Genome Project	The Human Genome Project was an international project that mapped and sequenced the first entire human genome. Completed in April 2003, data from the project are freely available to researchers and others interested in genetics and human health.
Human Microbiome project	Within the body of a healthy adult, microbial cells are estimated to outnumber human cells by a factor of ten to one. These communities, however, remain largely unstudied, leaving almost entirely unknown their influence upon human development, physiology, immunity, and nutrition. To take advantage of recent technological advances (DNA sequencing) and to develop new ones, the NIH Roadmap has initiated the Human Microbiome Project (HMP) with the mission of generating resources enabling comprehensive characterization of the human microbiota and analysis of its role in human health and disease.
Immunosuppressive drugs	Drugs that suppress the body's immune system and are usually used for treatment of auto-immune diseases where the body essentially attacks itself.
Insertion	Insertion is a type of mutation involving the addition of genetic material. An insertion mutation can be small, involving a single extra DNA base pair, or large, involving a piece of a chromosome.
Institute for Systems Biology (ISB)	ISB was co-founded in 2000 in Seattle, Washington by Dr. Leroy Hood, an immunologist and technologist; Dr. Alan Aderem, an immunologist and Dr. Ruedi Aebersold, a protein chemist. It has since grown to more than 220 staff members, including eleven faculty members and laboratory groups. It has established facilities for DNA sequencing, genotyping, DNA arrays and cell separations.
International HapMap project	The International HapMap Project is a partnership of scientists and funding agencies from Canada, China, Japan, Nigeria, the United Kingdom and the United States to develop a public resource to help researchers find genes associated with human disease and response to pharmaceuticals. The project officially started in 2002 and the third data set was released in 2009. The DNA sequence of any two people has been estimated to be 99.5 percent identical. The variations, however, may greatly affect an individual's disease risk. Sites in the DNA sequence where individuals differ at a single DNA base are called single nucleotide polymorphisms (SNPs). Sets of nearby SNPs on the same chromosome are inherited in blocks. This pattern of SNPs on a block is a haplotype. Blocks may contain a large number of SNPs, but a few SNPs are enough to uniquely identify the haplotypes in a block. The HapMap is a map of these haplotype blocks and the specific SNPs that identify the haplotypes are called tag SNPs.
Interstitial deletions	Deletion that does not involve the terminal parts of a chromosome.
KRAS mutation	A gene that is associated with cell growth and cancer.
Large structural change	When large parts of the genetic code from a chromosome shift locations or order.
Lipid	Small molecules commonly associated with fats and waxes that provide energy storage within a cell as well structure and communication.
Loss of heterozygosity	When an allele of one gene becomes activated in the same manner as the corresponding allele of a second gene.
Malaria	Malaria is a parasitic disease that involves high fevers, shaking chills, flu-like symptoms, and anemia.
Mayo Clinic	Mayo Clinic is a leading not-for-profit medical practice dedicated to the diagnosis and treatment of virtually every type of complex illness. Mayo provides clinic and hospital services at its locations in Rochester, Minn.; Jacksonville, Fla.; and Phoenix and Scottsdale, Ariz. and employs ~3,400 physicians and scientists.
Medicaid	Medicaid is available only to certain low-income individuals and families who fit into an eligibility group that is recognized by United States federal and state law. Medicaid does not pay money to you; instead, it sends payments directly to your health care providers. Depending on your state's rules, you may also be asked to pay a small part of the cost (co-payment) for some medical services. Medicaid is a state administered program and each state sets its own guidelines regarding eligibility and services.
Medicare	Medicare is a United States government run health insurance program for people age 65 or older, people under age 65 with certain disabilities, and people of all ages with End-Stage Renal Disease (permanent kidney failure requiring dialysis or a kidney transplant).
messenger RNA	A molecule of RNA that carries protein construction instructions to the part of the cell where proteins are synthesized (the ribosome)
Metastatic cancer	A state when cancer has spread throughout the body, outside of the original local organ (breast/colon/prostate etc.)
Methylation	A chemical state of a Cytosine and Guanine repeating stretch of DNA that influences the activity of genes. A state of methylation within the DNA is associated with inactivity.
Microarray	Typically a glass slide or silicon chip to which molecular probes are attached with dyes that can be excited by laser. Microarrays are used within the genetic research lab to compare genotypes (DNA) or gene expression (RNA) between genetic samples.
Microbe	A micro organism that is too small to be seen without a microscope.
Molecule	A group of at least two atoms bound together.

Fig 78 Definitions of scientific and other terms used in this report

Moore's law	Describes the timing of advances in computer technology, predicting that the number of transistors that can be placed on an integrated circuit has doubled every ~2 years.
Mutation	Changes in the DNA sequence of a cell's genetic code caused by environmental damage and or error in replication.
Next-generation (genome/genetic) sequencer	An instrument that estimates a contiguous arrangement of genetic code from a sample by sequencing millions of DNA fragments from the sample at the same time (in-parallel) using a camera(s) that takes pictures of the excitement of dyes by a laser.
NHGRI	The National Human Genome Research Institute (NHGRI) began as the National Center for Human Genome Research (NCHGR), which was established in 1989 to carry out the role of the National Institutes of Health (NIH) in the International Human Genome Project (HGP). The HGP was developed in collaboration with the United States Department of Energy and begun in 1990 to map the human genome. In 1993, NCHGR expanded its role on the NIH campus by establishing the Division of Intramural Research to apply genome technologies to the study of specific diseases. In 1996, the Center for Inherited Disease Research (CIDR) was also established (co-funded by eight NIH institutes and centers) to study the genetic components of complex disorders. In 1997 the United States Department of Health and Human Services renamed NCHGR the National Human Genome Research Institute (NHGRI), officially elevating it to the status of research institute - one of 27 institutes and centers that make up the NIH.
Non-coding	Non-coding DNA sequences do not code for amino acids. Most non-coding DNA lies between genes on the chromosome and has no known function. Other non-coding DNA, called introns, is found within genes. Some non-coding DNA plays a role in the regulation of gene expression.
Nucleotide	A nucleotide is the basic building block of nucleic acids. RNA and DNA are polymers made of long chains of nucleotides. A nucleotide consists of a sugar molecule (either ribose in RNA or deoxyribose in DNA) attached to a phosphate group and a nitrogen-containing base. The bases used in DNA are Adenine (A), Cytosine (C), Guanine (G), and Thymine (T). In RNA, the base Uracil (U) takes the place of Thymine.
Nucleus	A membrane enclosed part of a cell that contains DNA.
Oncologist	A medical professional that is concerned with the diagnosis and treatment of cancer.
Paired-end read	Two opposing ends of a DNA fragment are sequenced to determine the location of the fragment within the overall sequence. A method optimal for de novo sequencing and detection of structural variation within a DNA sample.
Parasite	An organism that interacts with a different species (a host) and usually benefits at the expense of the host.
PCR (Polymerase Chain Reaction)	A process used to exponentially increase the amount of DNA sequence present in a sample.
Petabyte	10^15 bytes, units of digital information used in computing.
Pharmaceutical Benefit Manager (PBM)	A third-party administrator of prescription drug programs, generally responsible for processing and paying prescription drug claims. A middle-man operation between a payer (the insurance company), a drug manufacturer and the provider (a pharmacy) that relies on administrative efficiency and other various services to reduce costs to the payer.
Phenotype	A phenotype is an individual's observable traits, such as height, eye color, and blood type. The genetic contribution to the phenotype is called the genotype. Some traits are largely determined by the genotype, while other traits are largely determined by environmental factors.
Photosynthetic algae	Algae that can convert light into energy.
Point mutation	A point mutation is when a single base pair is altered. Point mutations can have one of three effects. First, the base substitution can be a silent mutation where the altered codon corresponds to the same amino acid. Second, the base substitution can be a missense mutation where the altered codon corresponds to a different amino acid. Or third, the base substitution can be a nonsense mutation where the altered codon corresponds to a stop signal.
Protein	Proteins are an important class of molecules found in all living cells. A protein is composed of one or more long chains of amino acids, the sequence of which corresponds to the DNA sequence of the gene that encodes it. Proteins play a variety of roles in the cell, including structural (cytoskeleton), mechanical (muscle), biochemical (enzymes), and cell signaling (hormones). Proteins are also an essential part of diet.
Purine	A nitrogen base group including nucleic acids Adenine and Guanine.
Pyrimidine	A nitrogen base group including nucleic acids Thymine and Cytosine.
QA/QC	Quality Assurance/Quality Control is a process used in drug development and other chemical manufacturing that samples the product for accuracy, i.e. to make sure what the manufacturer thinks is being produced, really is being produced. This term has also been adopted within the DNA sequencing market with respect to using "long" read DNA sequencers (usually CE sequencers) to validate and check the quality of output from a "short" read sequencer.
Rare variant	Genetic variation that occurs in a small percentage of the population, hypothesized to be correlated to various diseases. The 1000 Genomes Project data has been used to associate 'rare variants' with disease.
Raw read accuracy	The rate at which a genetic sequencer incorrectly identifies a nucleotide base within a DNA sequence compared to correct identifications.
Read-length	The number of contiguous base reads a DNA sequencer performs. A higher read-length is associated with a higher level of accuracy.
Reagent	A chemical used to facilitate a reaction.
Reference genome	The Human Genome Project was an international project that mapped and sequenced the first entire human genome, creating a reference DNA sequence that is used by researchers as a means to construct and compare additional sequenced human genomes.
Refractory cancer	Cancer that failed to respond to an initial treatment therapy.
Re-sequencing	Sequencing a previously sequenced DNA sample to scan for known or unknown mutations, and can be performed by microarrays as well as DNA sequencers.
RNA	Ribonucleic acid (RNA) is a molecule similar to DNA. Unlike DNA, RNA is single-stranded. An RNA strand has a backbone made of alternating sugar (ribose) and phosphate groups. Attached to each sugar is one of four bases--Adenine (A), Uracil (U), Cytosine (C), or Guanine (G). Different types of RNA exist in the cell: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). More recently, some small RNAs have been found to

Fig 78 Definitions of scientific and other terms used in this report

	be involved in regulating gene expression.
RNA-seq	Sequencing a sample to identify and quantify RNA content, to understand the genetic activity of the sample. The analysis has been broadly adopted in cancer research. Also known as Whole Transcriptome Analysis.
SAGE combo kits	SAGE (Serial Analysis of Gene Expression) is a method that generates sequences from cDNA fragments for the discovery of new genes and the quantification of their expression levels in a certain tissue.
Sequence	An order of nucleotide bases Adenine, Cytosine, Guanine, and Thymine (Uracil in RNA)
Sequencing by Synthesis	A chemistry method used by next generation DNA sequencers that use a repeated stepped process of removing an inhibitory enzyme that allows the addition of one dye/fluorescently tagged nucleotide base which can be read by a camera and placed into a computer algorithm to determine a longer DNA sequence.
Single-base mutation	A single nucleotide mutation where one of the four base nucleotides has been replaced by a different nucleotide within the DNA sequence.
Small indel	Small insertions/deletions of genetic code within a sequence.
Somatic variation	Genetic mutation associated with an environmentally induced change.
Statistically significant	A conclusion that is highly probable.
Targeted genotyping	An experiment used to identify commonality of nucleotide variation across samples of DNA sequence.
Targeted therapy	A therapy designed to treat a patient population with a specific genetic profile.
template DNA	The single strand of DNA that is copied during DNA replication.
Terabyte	10^12 bytes, units of digital information used in computing.
Thromboembolism	The formation of a blood clot inside a blood vessel.
Thymine	Thymine (T) is one of four chemical bases in DNA, the other three being Adenine (A), Cytosine (C), and Guanine (G). Within the DNA molecule, Thymine bases located on one strand form chemical bonds with Adenine bases on the opposite strand. The sequence of four DNA bases encodes the cell's genetic instructions.
Vaccine	A drug used to stimulate the body's immune system to develop a defense response against a particular foreign agent.
Warfarin	An anticoagulant drug used to prevent the formation of blood clots.
Whole transcriptome analysis	Sequencing a sample to identify and quantify RNA to understand the genetic activity of the sample. The analysis has been broadly adopted in cancer research. Also known as Whole Transcriptome Analysis.

Source: National Human Genome Research Institute “Glossary of Genetic Terms”, National Cancer Institute, Macquarie Capital (USA), April 2010

Important disclosures:

Recommendation definitions

Macquarie - Australia/New Zealand
Outperform – return >5% in excess of benchmark return
Neutral – return within 5% of benchmark return
Underperform – return >5% below benchmark return

Macquarie – Asia/Europe
Outperform – expected return >+10%
Neutral – expected return from -10% to +10%
Underperform – expected return <-10%

Macquarie First South - South Africa
Outperform – expected return >+10%
Neutral – expected return from -10% to +10%
Underperform – expected return <-10%

Macquarie - Canada
Outperform – return >5% in excess of benchmark return
Neutral – return within 5% of benchmark return
Underperform – return >5% below benchmark return

Macquarie - USA
Outperform (Buy) – return >5% in excess of Russell 3000 index return
Neutral (Hold) – return within 5% of Russell 3000 index return
Underperform (Sell)– return >5% below Russell 3000 index return

Recommendations – 12 months

Note: Quant recommendations may differ from Fundamental Analyst recommendations

Volatility index definition*

This is calculated from the volatility of historical price movements.

Very high–highest risk – Stock should be expected to move up or down 60–100% in a year – investors should be aware this stock is highly speculative.

High – stock should be expected to move up or down at least 40–60% in a year – investors should be aware this stock could be speculative.

Medium – stock should be expected to move up or down at least 30–40% in a year.

Low–medium – stock should be expected to move up or down at least 25–30% in a year.

Low – stock should be expected to move up or down at least 15–25% in a year.
* Applicable to Australian/NZ/Canada stocks only

Financial definitions

All "Adjusted" data items have had the following adjustments made:
Added back: goodwill amortisation, provision for catastrophe reserves, IFRS derivatives & hedging, IFRS impairments & IFRS interest expense
Excluded: non recurring items, asset revals, property revals, appraisal value uplift, preference dividends & minority interests

EPS = adjusted net profit / efpowa*
ROA = adjusted ebit / average total assets
ROA Banks/Insurance = adjusted net profit /average total assets
ROE = adjusted net profit / average shareholders funds
Gross cashflow = adjusted net profit + depreciation
*equivalent fully paid ordinary weighted average number of shares

All Reported numbers for Australian/NZ listed stocks are modelled under IFRS (International Financial Reporting Standards).

Recommendation proportions – For quarter ending 31 March 2010

	AU/NZ	Asia	RSA	USA	CA	EUR	
Outperform	50.55%	62.20%	42.25%	42.39%	62.16%	46.74%	(for US coverage by MCUSA, 6.53% of stocks covered are investment banking clients)
Neutral	36.63%	19.02%	47.89%	50.35%	31.89%	34.78%	(for US coverage by MCUSA, 9.62% of stocks covered are investment banking clients)
Underperform	12.82%	18.78%	9.86%	7.27%	5.95%	18.48%	(for US coverage by MCUSA, 0.00% of stocks covered are investment banking clients)

Company Specific Disclosures:

Important disclosure information regarding the subject companies covered in this report is available at www.macquarie.com/research/disclosures.

Analyst Certification:
The views expressed in this research accurately reflect the personal views of the analyst(s) about the subject securities or issuers and no part of the compensation of the analyst(s) was, is, or will be directly or indirectly related to the inclusion of specific recommendations or views in this research. The analyst principally responsible for the preparation of this research receives compensation based on overall revenues of Macquarie Group Ltd ABN 94 122 169 279 (AFSL No. 318062) (MGL) and its related entities (the Macquarie Group) and has taken reasonable care to achieve and maintain independence and objectivity in making any recommendations.

General Disclaimers:
Macquarie Securities (Australia) Ltd; Macquarie Capital (Europe) Ltd; Macquarie Capital Markets Canada Ltd; Macquarie Capital Markets North America Ltd; Macquarie Capital (USA) Inc; Macquarie Capital Securities Ltd; Macquarie Capital Securities (Singapore) Pte Ltd; Macquarie Securities (NZ) Ltd; and Macquarie First South Securities (Pty) Limited are not authorized deposit-taking institutions for the purposes of the Banking Act 1959 (Commonwealth of Australia), and their obligations do not represent deposits or other liabilities of Macquarie Bank Limited ABN 46 008 583 542 (MBL) or MGL. MBL does not guarantee or otherwise provide assurance in respect of the obligations of any of the above mentioned entities. MGL provides a guarantee to the Monetary Authority of Singapore in respect of the obligations and liabilities of Macquarie Capital Securities (Singapore) Pte Ltd for up to SGD 35 million. This research has been prepared for the general use of the wholesale clients of the Macquarie Group and must not be copied, either in whole or in part, or distributed to any other person. If you are not the intended recipient you must not use or disclose the information in this research in any way. If you received it in error, please tell us immediately by return e-mail and delete the document. We do not guarantee the integrity of any e-mails or attached files and are not responsible for any changes made to them by any other person. MGL has established and implemented a conflicts policy at group level (which may be revised and updated from time to time) (the "Conflicts Policy") pursuant to regulatory requirements (including the FSA Rules) which sets out how we must seek to identify and manage all material conflicts of interest. Nothing in this research shall be construed as a solicitation to buy or sell any security or product, or to engage in or refrain from engaging in any transaction. In preparing this research, we did not take into account your investment objectives, financial situation or particular needs. Before making an investment decision on the basis of this research, you need to consider, with or without the assistance of an adviser, whether the advice is appropriate in light of your particular investment needs, objectives and financial circumstances. There are risks involved in securities trading. The price of securities can and does fluctuate, and an individual security may even become valueless. International investors are reminded of the additional risks inherent in international investments, such as currency fluctuations and international stock market or economic conditions, which may adversely affect the value of the investment. This research is based on information obtained from sources believed to be reliable but we do not make any representation or warranty that it is accurate, complete or up to date. We accept no obligation to correct or update the information or opinions in it. Opinions expressed are subject to change without notice. No member of the Macquarie Group accepts any liability whatsoever for any direct, indirect, consequential or other loss arising from any use of this research and/or further communication in relation to this research. Clients should contact analysts at, and execute transactions through, a Macquarie Group entity in their home jurisdiction unless governing law permits otherwise.

Country-Specific Disclaimers:
Australia: In Australia, research is issued and distributed by Macquarie Securities (Australia) Ltd (AFSL No. 238947), a participating organisation of the Australian Securities Exchange. **New Zealand**: In New Zealand, research is issued and distributed by Macquarie Securities (NZ) Ltd, a NZX Firm. **Canada**: In Canada, research is prepared, approved and distributed by Macquarie Capital Markets Canada Ltd, a participating organisation of the Toronto Stock Exchange, TSX Venture Exchange & Montréal Exchange. Macquarie Capital Markets North America Ltd., which is a registered broker-dealer and member of FINRA, accepts responsibility for the contents of reports issued by Macquarie Capital Markets Canada Ltd in the United States and sent to US persons. Any person wishing to effect transactions in the securities described in the reports issued by Macquarie Capital Markets Canada Ltd should do so with Macquarie Capital Markets North America Ltd. The Research Distribution Policy of Macquarie Capital Markets Canada Ltd is to allow all clients that are

entitled to have equal access to our research. **United Kingdom**: In the United Kingdom, research is issued and distributed by Macquarie Capital (Europe) Ltd, which is authorised and regulated by the Financial Services Authority (No. 193905). **Hong Kong**: In Hong Kong, research is issued and distributed by Macquarie Capital Securities Ltd, which is licensed and regulated by the Securities and Futures Commission. **Japan**: In Japan, research is issued and distributed by Macquarie Capital Securities (Japan) Limited, a member of the Tokyo Stock Exchange, Inc., Osaka Securities Exchange Co. Ltd, and Jasdaq Securities Exchange, Inc. (Financial Instruments Firm, Kanto Financial Bureau (kin-sho) No. 231, a member of Japan Securities Dealers Association and Financial Futures Association of Japan). **South Africa**: In South Africa, research is issued and distributed by Macquarie First South Securities (Pty) Limited, a member of the JSE Limited. **Singapore**: In Singapore, research is issued and distributed by Macquarie Capital Securities (Singapore) Pte Ltd (Company Registration Number: 198702912C), a Capital Markets Services license holder under the Securities and Futures Act to deal in securities and provide custodial services in Singapore. Pursuant to the Financial Advisers (Amendment) Regulations 2005, Macquarie Capital Securities (Singapore) Pte Ltd is exempt from complying with sections 25, 27 and 36 of the Financial Advisers Act. All Singapore-based recipients of research produced by Macquarie Capital (Europe) Limited, Macquarie Capital Markets Canada Ltd, Macquarie First South Securities (Pty) Limited and Macquarie Capital (USA) Inc. represent and warrant that they are institutional investors as defined in the Securities and Futures Act. **United States**: In the United States, research is issued and distributed by Macquarie Capital (USA) Inc., which is a registered broker-dealer and member of FINRA. Macquarie Capital (USA) Inc, accepts responsibility for the content of each research report prepared by one of its non-US affiliates when the research report is distributed in the United States by Macquarie Capital (USA) Inc. Macquarie Capital (USA) Inc.'s affiliate's analysts are not registered as research analysts with FINRA, may not be associated persons of Macquarie Capital (USA) Inc., and therefore may not be subject to FINRA rule restrictions on communications with a subject company, public appearances, and trading securities held by a research analyst account. Any persons receiving this report directly from Macquarie Capital (USA) Inc. and wishing to effect a transaction in any security described herein should do so with Macquarie Capital (USA) Inc. Important disclosure information regarding the subject companies covered in this report is available at www.macquarie.com/research/disclosures, or contact your registered representative at 1-888-MAC-STOCK, or write to the Supervisory Analysts, Research Department, Macquarie Securities, 125 W.55th Street, New York, NY 10019.

Auckland Tel: (649) 377 6433	**Bangkok** Tel: (662) 694 7999	**Calgary** Tel: (1 403) 218 6650	**Hong Kong** Tel: (852) 2823 3588	**Jakarta** Tel: (62 21) 515 1818	**Johannesburg** Tel: (2711) 583 2000	**Kuala Lumpur** Tel: (60 3) 2059 8833
London Tel: (44 20) 3037 4400	**Manila** Tel: (63 2) 857 0888	**Melbourne** Tel: (613) 9635 8139	**Montreal** Tel: (1 514) 925 2850	**Mumbai** Tel: (91 22) 6653 3000	**Perth** Tel: (618) 9224 0888	**Seoul** Tel: (82 2) 3705 8500
Shanghai Tel: (86 21) 6841 3355	**Singapore** Tel: (65) 6231 1111	**Sydney** Tel: (612) 8232 9555	**Taipei** Tel: (886 2) 2734 7500	**Tokyo** Tel: (81 3) 3512 7900	**Toronto** Tel: (1 416) 848 3500	**New York** Tel: (1 212) 231 2500

Available to clients on the world wide web at www.macquarieresearch.com and through Thomson Financial, FactSet, Reuters, Bloomberg, CapitalIQ and TheMarkets.com.

Macquarie (USA)
Equities Research

MACQUARIE

Research

Heads of Equity Research

John O'Connell (Global Co – Head)	(612) 8232 7544
David Rickards (Global Co – Head)	(44 20) 3037 4399
Mark Little (US)	(1 212) 231 2577
Stephen Harris (Canada)	(1 416) 848 3655

Consumer Discretionary

Gaming & Leisure

Joel Simkins (New York)	(1 212) 231 2635
Chad Beynon (New York)	(1 212) 231 2634

Retailing

David Pupo (Toronto)	(1 416) 848 3505

Homebuilding & Materials

Ken Zener (New York)	(1 212) 231 2479

Energy

Jason Gammel (New York)	(1 212) 231 2633
Chris Feltin (Calgary)	(1 403) 539 8544
Cristina Lopez (Calgary)	(1 403) 539 8542
Leon Knight (Calgary)	(1 403) 303 8655
Chris Theal (Calgary)	(1 403) 539 4349
David Popowich (Calgary)	(1 403) 539 8529
Scott Treadwell (Calgary)	(1 403) 539 8530
Waqar Syed (Denver)	(1 303) 952 2753
Joe Magner (Denver)	(1 303) 952 2751
Ryan McCormick (Denver)	(1 303) 952 2752

Alternative Energy

Kelly Dougherty (New York)	(1 212) 231 2493

Financials

Asset Managers/Financial Technology

Roger Smith (New York)	(1 212) 231 8016
William Broomall (New York)	(1 212) 231 8039

Banks/Trust Banks

David Trone (Head of US Commercial and Investment Banks Research)	(1 212) 231 8051
Andrew Marquardt (New York)	(1 212) 231 8037
Albert Savastano (New York)	(1 212) 231 8046
Adam Klauber (Chicago)	(1 312) 660 9187
John Pancari (New York)	(1 212) 231 8014
Thomas Alonso (New York)	(1 212) 231 8047
Jonathan Elmi (New York)	(1 212) 231 8065
Bill Young (New York)	(1 212) 231 8052
Sumit Malhotra (Toronto)	(1 416) 848 3687

Financial Strategies

Adam Klauber (Chicago)	(1 312) 660 9187

Investment Banks & Brokers

David Trone (New York)	(1 212) 231 8051
Steven Fu (New York)	(1 212) 231 8049
Bimal Shah (New York)	(1 212) 231 8053

Financials – cont

Life Insurance

Mark Finkelstein (Chicago)	(1 312) 660 9179

Market Structure

Edward Ditmire (New York)	(1 212) 231 8076

Mortgage & Consumer Finance

Bill Carcache (New York)	(1 212) 231 8034
Matthew Howlett (New York)	(1 212) 231 8063

Mortgage REITs

Matthew Howlett (New York)	(1 212) 231 8063

Property & Casualty Insurance

William Yankus (West Hartford)	(1 860) 380 2003
Matthew J Carletti (Chicago)	(1 312) 660 9134
Dan Farrell (New York)	(1 212) 231 8044
Amit Kumar (New York)	(1 212) 231 8013
Dan Schlemmer (Chicago)	(1 312) 660 9182

Industrials

Aerospace & Defense

Rob Stallard (New York)	(1 212) 231 2486
Rama Bondada (New York)	(1 212) 231 2481

Capital Goods

Steven Song (New York)	(1 212) 231 2455

Infrastructure Services

Avi Dalfen (Toronto)	(1 416) 628 3934

Construction and Engineering

Sameer Rathod (New York)	(1 212) 231 2474

Transportation & Logistics

Scott Flower (New York)	(1 212) 231 2537

Materials

Global Metals & Mining

Curt Woodworth (New York)	(1 212) 231 2482
Pierre Vaillancourt (Toronto)	(1 416) 848 3647
Duncan McKeen (Montréal)	(1 514) 925 2856

Real Estate

Property Trusts & Developers

Robert Stevenson (New York)	(1 212) 231 8068
Ki Bin Kim (New York)	(1 212) 231 6386
Michael Levy (New York)	(1 212) 231 2626
Dave Wigginton (New York)	(1 212) 231 6380
Michael Smith (Toronto)	(1 416) 848 3696

TMET

Telecommunications

Phil Cusick (New York)	(1 212) 231 6376
Glenn Jamieson (Toronto)	(416) 848 3658

Software

Brad Zelnick (New York)	(1 212) 231 2618

IT Hardware

Richard Choe (New York)	(1 212) 231 6370

Media

Ben Stretch (New York)	(1 212) 231 2574

Technology

Shawn Webster (New York)	(1 212) 231 2539
Glenn Jamieson (Toronto)	(416) 848 3658

Business Services

Kevin McVeigh (New York)	(1 212) 231 6191

Utilities

Angie Storozynski (New York)	(1 212) 231 2569
Matthew Akman (Toronto)	(1 416) 848 3510
Stephen Harris (Toronto)	(1 416) 848 3655

Commodities & Precious Metals

Metals & Mining

Jim Lennon (London)	(44 20) 3037 4271
Max Layton (London)	(44 20) 3037 4273
Kona Haque (London)	(44 20) 3037 4334

Oil & gas

Jan Stuart (New York)	(1 212) 231 2485

Emerging Leaders (Small/Mid Cap)

Al Kabili (New York)	(1 212) 231 2473
Jon Groberg (New York)	(1 212) 231 2612
Cooley May (New York)	(1 212) 231 2586

Economics and Strategy

Stephen Harris (Toronto)	(1 416) 848 3655
Jan Stuart (Global Oil Economist)	(1 212) 231 2485

Find our research at

Macquarie:	www.macquarie.com.au/research
Thomson:	www.thomson.com/financial
Reuters:	www.knowledge.reuters.com
Bloomberg:	MAC GO
Factset:	http://www.factset.com/home.aspx
CapitalIQ	www.capitaliq.com
TheMarkets.com	www.themarkets.com

Contact Gareth Warfield for access (612) 8232 3207

Email addresses

FirstName.Surname@macquarie.com

eg. David.Rickards@macquarie.com

Sales

Equities

Stevan Vrcelj (Head of Global Sales)	(612) 8232 5999
Alex Rothwell (Toronto)	(1 416) 848 3677

US Sales

Greg Coleman (New York)	(1 212) 231 2567

US Financial Specialist Sales

Scott Barishaw (New York)	(1 212) 231 8026

US Sales Trading

Austin Graham (New York)	(1 212) 231 2494

Canada Sales

Tim Sorensen (Toronto)	(1 416) 848 3623

Canada Trading

Perry Catellier (Toronto)	(1 416) 848 3619