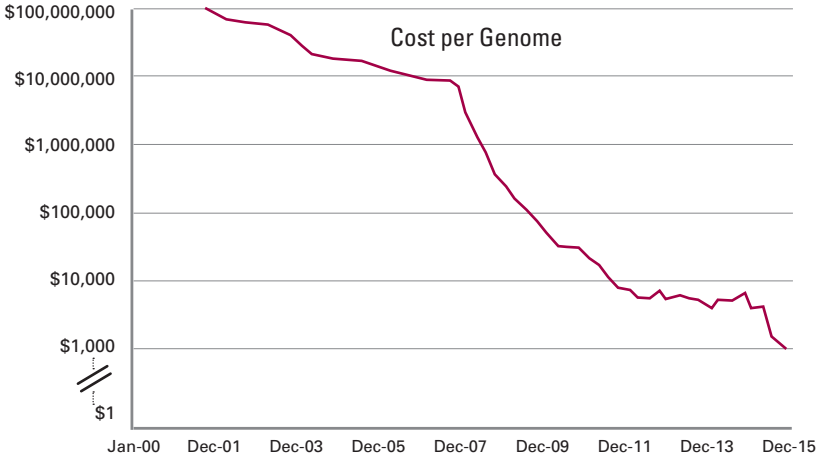
Mining DNA: businesses built on genomics

Over the past decade, DNA sequencing costs have been collapsing exponentially. Today, there is an abundance of sequencing facilities and having a person's full genome sequenced now only costs US$ 1,000 (see Graph 1). The pace of technology development was so fast that corporate stars rose and fell swiftly, with one wave of disruptive innovators following the other and a sizeable number of high-profile corporate failures. Challengers continue to emerge, which again may reshape the market for sequencing instruments (see Box 1); however, the main instrument providers seem more or less set.

GRAPH 1 / SEQUENCING COSTS



Source: NHGRI.

Already here: personalized medicine

Today, DNA-sequence based tests (or molecular diagnostics) are an integral part of medical practice. For example, molecular diagnostic tests (e.g. for EGFR and KRAS) can determine for certain cancers whether a patient will benefit from a specific, highly targeted treatment. It is now common that companion diagnostics are developed together with new cancer drugs. These are the prime examples of "personalized medicine".

There are many other areas in which molecular diagnostics supports medicine. For example, testing the predisposition of a patients' body to clear the drug substance from the blood helps physicians to adjust doseage. This is highly relevant if drugs are taken over an extended period of time, such as in psychiatry, pain management, or for cardiovascular conditions. For certain cancer patients, a DNA test can help decide whether they should undergo chemotherapy after surgery. Establishing whether a person has inherited a predisposition for a disease enables the physician to have an informed discussion on preventive measures. All these applications bring a great medical and health economic benefit and increasingly health insurances are rewarding the development of tests with reimbursement.

Consumer genomics in a changing environment

The early success and big media attention enjoyed by 23andMe, one of the first consumer genomics companies, showed that there is great demand from consumers to learn about their genetic makeup, be it to answer relatively minor (ear wax type), personal (ancestry) or serious (cancer risk) questions. However, the regulatory requirements for consumer genomic testing are not yet clear: in late 2013, FDA stopped 23andMe from providing medical information.

Another determinant for the future market of consumer genomic testing are legal restrictions for the access to and the use of genomic data, as Ellen Matloff points out when speaking to us. Ellen is a renowned cancer genetics expert and CEO of My Gene Counsel, a US-based firm providing digital tools to ensure accurate understanding of genetic test results. Prior to the Genetic Information Nondiscrimination Act (GINA) of 2008, she points out that it was not uncommon for patients to use aliases and pay cash for genetic testing, out of fear of potential insurance or employment repercussions in case

of adverse results. Despite certain loopholes, such as not covering life insurance, long-term care or disability insurance, GINA has alleviated the fears of many genetics customers.
Thus, consumer-driven genomic tests are here to stay, as there is clearly a demand. However, it is likely that they will be offered as a package together with counselling, as done by Kailos Genetics (see Box 2), a combination which may even become mandatory under certain legislations.

Building the Google of genomics

The revolution in DNA sequencing is also the basis for a long-term vision: to build and mine large-scale genomic databases to find new links between genes and medical conditions.
However, there is still a lot of work to be done before such visions become reality. The reason for this is that most diseases are complex long-term processes, involving multiple genes (to varying degrees) and dependent on such factors as lifestyle, nutrition and the environment.
In order to convert genomic data to knowledge about disease mechanisms, three challenges need to be met: i) sequencing of a massive number (hundreds of thousands) of complete genomes; ii) collection of data on medical conditions and relevant traits of each sequenced individual; and iii) development of powerful algorithms and computational methods to process and integrate the various data types.
There are several ongoing efforts to create large-scale genomic databases, both privately (deCODE, Invitae) and publicly funded (100,000 Genomes Project of Genomics England). One of the most prominent examples is, again, 23andMe (mentioned above).

BOX 1 / SEQUENCING TECHNOLOGY PROVIDERS

	Established players			Challengers	
Criteria	Illumina	Ion Torrent (Thermo Fisher)	PacBio	Oxford Nanopore	Genia (Roche)
Accurate reads	+	+	-	-	
Long reads	-	-	+	+	
Single molecule	-	-	+	+	+
No complicated optics (fast and robust)	-	+	-	+	
Low price		+	-	+	
Wide selection of instruments	+	-	-		
Handheld	-	-	-	+	

BOX 2

KAILOS GENETICS – A SERVICE PROVIDER EXAMPLE

As the medical applications of molecular diagnostics are more widely used, new types of businesses emerge. Privately held Kailos Genetics, based in Alabama USA, for example, could be described as a "genome analysis service provider".

Kailos performs tests to predict patients' responses to common drugs, such as antidepressants or birth control treatment. Even though consumers drive the demand for testing, Kailos' CEO, Brian Pollock, emphasizes that physicians and other medical professionals are closely involved in the entire process, from approving each test to interpreting and discussing the results. "Sequencing today is a cheap commodity" he tells us.

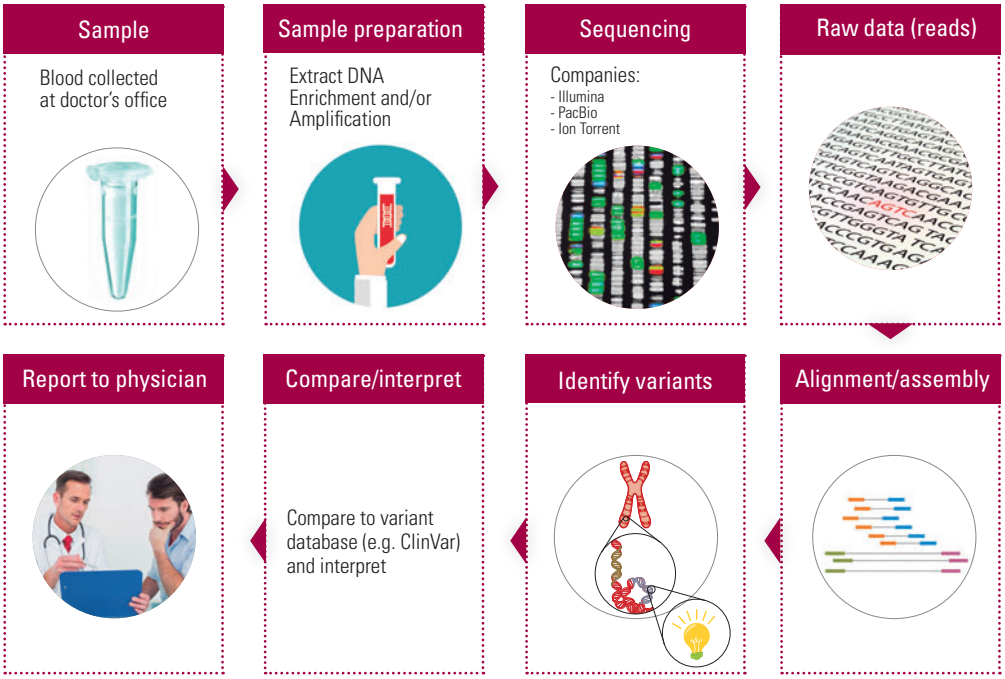
"Sample preparation and data analysis are really value-adding". Kailos simplified the wet lab component by developing a one-step sample enrichment method, TargetRich.

The company's Kailos Blue analysis software is based on open source algorithms and tools such as Broad Institute's Genome Analysis Toolkit (GATK), and can be used with sequencing data from any platform, in contrast to other software.

FROM SAMPLE TO INFORMATION – TYPICAL GENOMIC WORKFLOW

A typical genomic workflow consists of several wet lab and computational components. After sequencing, the raw data (reads) are assembled into a whole genome (or genomic fragments).

Algorithms and databases of known genetic variants are used to evaluate the customer's data and to create a report.



The company, fittingly backed by Google (the prime collector of massive amounts of data), cleverly marketed its genetics tests to customers. This allowed 23andMe to collect huge amounts of data – its database currently contains genomic information (albeit not full genomes) of approximately one million people. In addition, 23andMe has managed to create a highly engaged user community, which regularly responds to questionnaires and provides additional information.
To exploit that trove, the company has partnered with a number of Pharma companies for specific diseases and is building up capabilities to start in-house drug development.
The continued interaction with its customers allows the 23andMe team to constantly adjust their hypotheses and to probe for new gene/trait connections.
While 23andMe has developed an excellent method to collect data, other initiatives focus on developing sophisticated algorithms to extract the knowledge from the data glut, including both open source projects such as Bioconductor and private entities (Sophia Genetics and Veritas Genetics). The key question for owners of commercial large-scale genomic databases is how to protect their knowledge. It is not clear how intellectual property regarding variant-trait combinations will be handled. For example, genes cannot be patented, as the outcome of the epic battle against Myriad Genetics' BRCA patent (the famous "breast cancer gene") has shown. Ownership of personal data will also be a point to consider, particularly in Europe. Both issues set boundary conditions on how the "Google of Genomics" can develop.

On the back of the sequencing revolution, an evolving system of business models

The speed of technical development in DNA sequencing remains high, but the competitive field of instrument providers will more or less remain stable (see Box 2), although Illumina will probably lose part of its 70% market share as the market matures. Sequencing technology has laid the basis for the evolution of a variety of down-stream business models, including test service providers such as Kailos, big data-collectors such as 23andMe, data miners such as Switzerland-based Sophia Genetics and of course, a large number of novel diagnostic tests. Which model will get the highest share of the value created remains unclear and will be interesting to follow.

CONTRIBUTED BY

External Expert
Olga Nikolayeva
(olga.nikolayeva@gmail.com)
Biomedical Engineer
and Bioinformatician

IMAP Switzerland
Christoph Bieri
(christoph.bieri@imap.com)