BRIEFING No.5

HEALTH

NEXT GENERATION SEQUENCING

DNA sequencing technologies over the last three decades have been based on principles first developed by Frederick Sanger in the mid-1970s. Subsequent improvements have led to an increase in the length of DNA fragments able to be read and a move away from radioactive labelling towards detection using fluorescent markers. Next generation sequencing (NGS) has taken some of the later developments of the Sanger method and has added another dimension in the form of high-throughput technologies that can parallelize the process, that integrate reactions at the micro- or nanoscale on chip surfaces, and that produce thousands or millions of sequences at once. These high-throughput sequencing technologies are intended to lower the costs of DNA sequencing far beyond that possible with earlier methods.

This dramatic change in throughput has been facilitated by the application of nanotechnology, allowing the miniaturisation of processes, although challenges remain in handling and storing the large amounts of data produced. These rapid advances in output along with falling costs are expected eventually to allow for rapid and cost-effective sequencing of single or multiple personal genomes. It is hoped that this will have a strong impact in terms of developing personalised medical treatments based on the genetic background of the patient. Such developments in sequencing technology are therefore likely to have a considerable impact on society in general.

The 'grand' challenge

A key goal is the routine sequencing of the personal genome which will form a prerequisite to matching therapies to individual patients' genotypes in the future, leading to the development of medicine at a personalised level. Applications of nanotechnology will increase sequencing throughput, speed and accuracy, and will lead to a reduction in sample size.

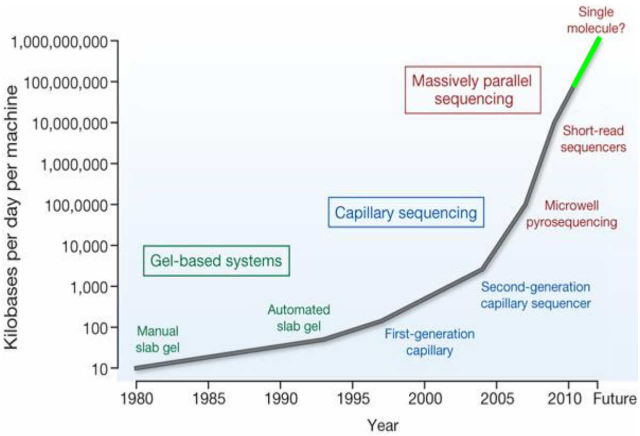

***Figure 1:** Development of DNA sequencing throughput[1].*

Background to nano development

Nanopores and nanoparticles

The novel elements that nanotechnology has brought to sequencing lie in the extreme miniaturisation of the process that becomes possible. In the case of nanopores this is downsized to the single molecule level[2]. The pore can be formed of protein that is inserted into a membrane or made by nanofabrication into a silicon-based chip. The nanopore works as a gate which is occluded by threaded DNA strands. The current passing through this gate is dependent on the base composition of the DNA. This allows the sequencing by studying the characteristic ionic current blockades.

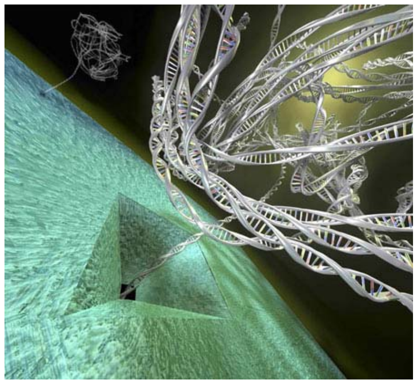

***Figure 2:** Representation of DNA passing through a nanopore for analysis (image courtesy of Dekker Group, University of Delft)*

Nanoparticles such as magnetic "nanobeads" are able to regulate the flow of DNA strands through multiple nanopores slowly enough for the base pairs to be read very accurately thereby opening the way for large-scale applications.[3]

Other methods

Other new nanoscale technologies are being applied that facilitate the sequencing of single molecules. These include diffusion of fluorophore-labelled nucleotides through a zero-mode waveguide (ZMW), the use of quantum dots (nanometre-sized semiconductor crystals) attached to proprietary DNA polymerase molecules as the core sequencing engine, and FRET- (fluorescence resonance energy transfer) based approaches [4, 5, 6].

These new approaches will contribute to a much faster and more economic sequencing of genomes. This will be useful for the sequencing of unknown genomes of species (or subspecies of interest like pathogens), domesticated animals and, most importantly, individual variations within the human genome itself.

Impacts

Genomic sequencing of pathogens will contribute to a better understanding of their biology and help to identify the "Achilles heel" that might be used as a route to develop better treatments for infectious diseases. The sequencing of domesticated animals and crop plants will help to boost their productivity, disease resistance, and robustness by improving breeding.

However, the greatest impact is likely to be the sequencing and comparison of human genomes. Already, genomic differences are being studied to explain metabolic, physiological and behavioural differences between people. So far this has been aimed at identifying genomic traits leading to certain predispositions to disease and also variations that are targets or indicators for improved disease treatment. These data will eventually be useful in predicting the health of individuals and potentially influencing their lifestyle choices.

Affordable sequencing technologies could have potential impacts for European healthcare systems; a significant proportion of drugs are not effective for certain groups of patients due to genetic differences. Sequencing services may also lead to new business sectors servicing the collection, storage, evaluation, and protection of data.
There are potentially complex and interrelated problems to be addressed here, including:

- What legal restrictions should apply to such data?
- For whom, for what purposes and with what specific criteria?
- How will such restrictions and safeguards be enforced?

Economic/Industry

According to BCC Research, the worldwide market for sequencing products is predicted to grow from an estimated $1.3 billion in 2010 to more than $3.3 billion by 2015, a compound annual growth rate (CAGR) of 20.5% over the next five years.
Life science research, and drug discovery and de-

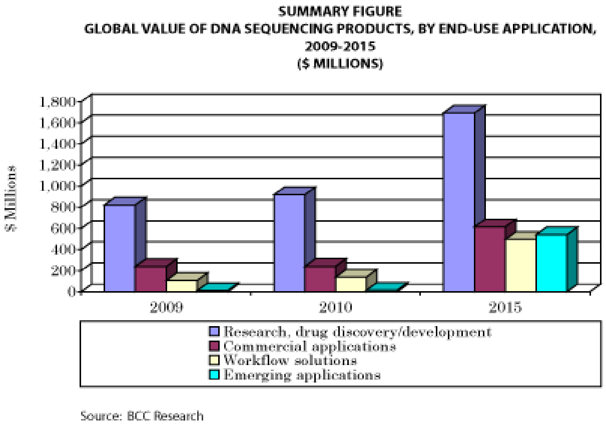

Figure 3: *Global value of sequencing products* [7]

velopment applications represent the two largest markets for DNA sequencing revenues, accounting for an estimated $920.1 million in 2010. These markets are forecast to grow at a compound annual growth rate (CAGR) of 13% to reach nearly $1.7 billion by 2015. Emerging applications, including personal genomics and clinical diagnostics, are forecast to account for $541.4 million by the year 2015, an increase from $15.5 million in 2010 and representing a 103.5% compound annual growth rate. NGS may also contribute towards the growth of personalised medicine which will rely on the availability of suitable genomic biomarkers. [8]

The majority of companies providing sequencing technologies are US based; exceptions include Roche Diagnostics and Oxford Nanopore Technologies, UK. There are also a number of companies offering NGS as a service in many European countries with the UK, France, and Germany leading this list.

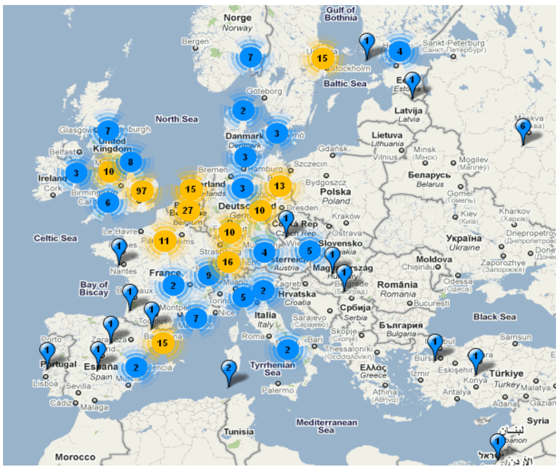

Figure 4: *Distribution of NGS facilities in the EU.*

Some of these companies have previously offered classical sequencing services and have now adapted to the new technologies that are becoming available. Statistics on the global distribution of sequencers and other facilities have been compiled by the Centre for Systems Biology, University of Birmingham, UK [9]. According to these statistics, Illumina is currently providing the majority of sequencing machines with ABI and Roche also contributing significantly.

Technology readiness levels

So-called second generation sequencing technologies are now widely available. However, relatively few research centres and companies can currently afford to purchase the newest sequencers due to their high costs.

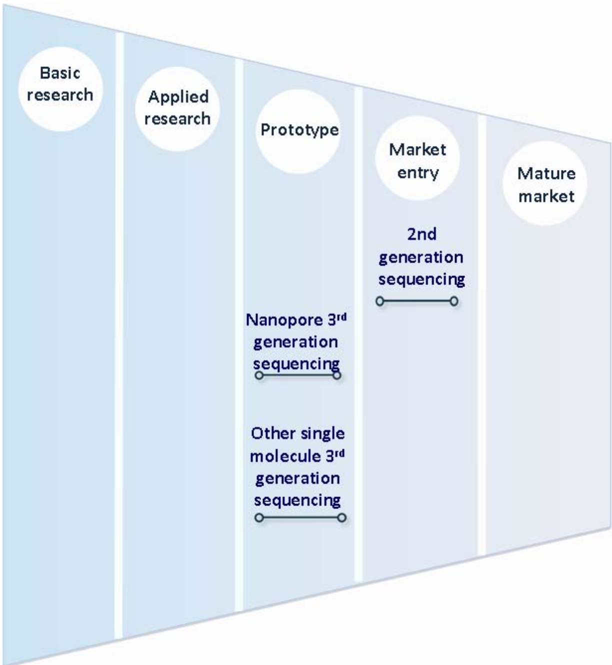

Figure 5: *Technology Readiness Levels for next generation sequencing technologies*

At the same time, other third generation sequencing technologies are reaching maturity and will also become available by the end of 2010/11. They offer faster sequencing and much longer read lengths without any previous amplification steps which are very useful for *de novo* sequencing of entirely unknown genomes.

Societal/Impact on European citizens

The fields of personal genomics and clinical diagnostics will be strongly impacted by these developments. The costs for sequencing a human genome are likely to drop to less than €1000 in the next two years. This has the potential to reduce healthcare and societal costs due to better diagnosis of predisposition towards certain diseases and, subsequently, the adaptation of treatments or lifestyles.

The sequencing of cancerous tissues will help to customise therapies, to reduce the quantities of highly toxic drugs used and their unwanted side effects, and increase the effectiveness of novel (and possibly costly) targeted drugs.

Attention must be paid also to potential unwanted or unintended consequences of the technologies. There is a potential for misuse of the data generated by NGS technologies by business, health insurers, employers, social services or individuals. In addition, some of the expected impacts may also have knock-on effects and consequences that are, as yet, unquantified including the unclear boundary with future genomic therapies.

Challenges

Reducing costs

A key challenge for NGS is the reduction of cost. The current goal is to reduce the cost of sequencing a human genome to $1000 and it seems likely that this goal will be achieved in the next two years. To this end, a variety of different approaches are being developed to increase speed, throughput and read lengths, to reduce consumables, and to simplify equipment.

Patent applications

A World Intellectual Property Organization (WIPO) search using the keywords 'nano' and 'high throughput sequencing' identified the following patent developments (see ***Figure 6***)

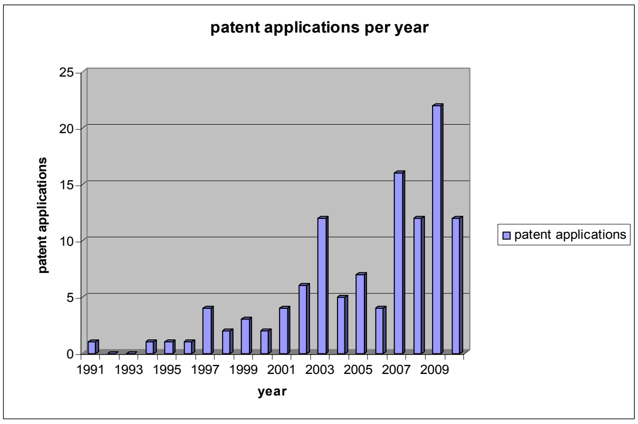

Figure 6: *Patent applications for NGS per year .*

The same WIPO search shows that NGS technologies are under rapid development internationally, illustrated by a continuing rise in patent applications. Main players are currently the US and the EU, followed by China.

Intellectual property disputes

Given the number of participants aiming for market share in the second- and third-generation sequencing arena, it is not surprising that overlaps in features of some technologies have raised concerns that intellectual property (IP) disputes will hinder the efforts of some contenders. The DNA microarray sector experienced considerable IP conflict during its first decade of commercial activity and, indeed, several contenders were forced by litigation, or threat of litigation, to pay substantial license fees, and in several cases to exit the arena (or to not enter it in the first place).

HEALTH: NEXT GENERATION SEQUENCING

PacBio and Life Technologies, which owns VisiGen, are developing potentially rival real-time, single-molecule sequencing platforms and the two companies are currently involved in a patent interference case. Similar disputes based on the surface amplification of DNA have also occurred between the companies Illumina and Life Technologies.

Ethical concerns

A number of ethical issues are raised by the availability of personal genomic information. Employers may be interested in their employees health status as influenced by their genetic background; this is also true for health insurers, whether public or private. Additionally, affected persons may be subjected to discrimination because of certain characteristics of their genome or certain variations within it. Such ethical concerns call for a clear legal framework regarding access to sequence data. However, even the best and most comprehensive laws will not completely eliminate the possible dual use of data.

EHS aspects

General and workplace environments could benefit from new sequencing methods as fewer reagents are consumed and thus fewer waste products (like radioactivity or toxic dyes) are likely to be released. Significant direct environmental exposure is unlikely but, in the waste stage, environmental exposure and effects may occur. It is hard to predict what kind of materials may enter waste streams, thus hampering the assessment of any (potential) risk. Regarding human safety, applications will be *in vitro*, and nanomaterials will be embedded in the devices, both implying there is likely to be little significant human exposure during use. There could, however, be potential worker exposure during product manufacturing and environmental exposure resulting from the process.

Communication

The prospect of personal genomic sequencing is likely to provoke high hopes, and fears, amongst the public. In order to maintain a firm factual base in debate and to avoid undue speculation, it is important not to over-hype the expected impacts of NGS on society. Communication on the topic should therefore remain prudent and address principally current scientific knowledge, developments and applications, rather than hypothesising on possible future scenarios.

EU Competitive Position

In July 2010, it was announced that EU based Roche and IBM would join forces with the aim of reducing the cost of DNA sequencing to a few hundred Euros by combining a nanopore-based sequencing approach with IBM's "DNA transistor" technology[10]. According to the report, Roche Applied Science will develop and market all products based on the technology providing EU industry with an opportunity to operate in the rapidly-expanding market for rapid genomic sequencing.

Further European activity includes the EU FP7 Project READNA www.cng.fr/READNA/ which is developing methods and devices that will boost the possibilities of genetic research by closing in on the target of €1000 for complete human genome sequencing.

Summary

- Next generation sequencing has matured and is entering the market.
- Companies may soon be able to sequence entire human genomes as a service.
- The progress in sequencing throughput is exponential.
- Current developments in the technologies are reducing sequencing costs.
- Information on the personal genome will facilitate personalised medicine, strongly influencing future healthcare provision.
- There are countless other possible applications, which European countries will have to choose those to allow and those to not.
- Regulation of access to personal genomic data will require thorough legal reflection.

Contact Information

Technical: Richard Moore, Institute of Nanotechnology, richard.moore@nano.org.uk
Economic: Jörn Gloekler, NMTC, gloekler@nmtc.de

References

[1] Stratton, M.R., Campbell, P.J. & Futreal, P.A. The cancer genome. Nature 458, 719-724 (2009).
[2] Branton, D. The potential & challenges of nanopore sequencing. Nat Biotech 26, 1146-1153 (2008).
[3] Hongbo Peng and Xinsheng Sean Ling, Nanotechnology 20 (2009) http://iopscience.iop.org/0957-4484/20/18/185101/pdf/nano9_18_185101.pdf
[4] Rothberg, J.M. & Leamon, J.H. The development and impact of 454 sequencing. Nat Biotech 26, 1117-1124 (2008).
[5] Pacific Biosciences' SMRT Technology. http://www.pacificbiosciences.com/index.php?q=overview1
[6] DNA sequencing: generation next-next, Nature Methods - 5, 267 - 274 (2008), http://www.nature.com/nmeth/journal/v5/n3/full/nmeth0308-267.html
[7] DNA Sequencing: Emerging Technologies and Applications (BIO045C). http://www.bccresearch.com/report/BIO045C.html
[8] Diagnostics 2009 - Moving towards personalised medicine, PricewaterhouseCoopers, 2009, http://www.pwc.com/diagnosticsRx
[9] Genomics: High-throughput "Next-Generation" Sequencing Facilities Statistics, University of Birmingham, UK, http://pathogenomics.bham.ac.uk/hts/stats
[10] European Biotechnology Science and Industry News, 01.07.10, Biocom AG, http://www.eurobiotechnews.eu/index.php?id=11535&no_cache=1&tx_ttnews[tt_news]=12393&cHash=6668c79611