DNA sequencing:Layout 1 17/1/08 19:13 Page 74

Next Generation DNA Sequencing technologies are facilitating new approaches for drug discovery and development

Human genetics is the foundation of disease as well as the response to pharmaceutical agents. Today, promising drugs are abandoned due to the lack of significant efficacy in broad patient populations. Recently, blockbuster drugs have been removed from the market due to unexplained toxicity not revealed in clinical trials. A need exists to economically address individual human variation in disease and drug response. With the latest DNA sequencing technologies ushering in an era of personalised medicine the application of these technologies could provide the flexible genetic tools needed to support the drug discovery and development process.

By Dr Bruce E. Taillon and Brendon Hill

Over the past 10 years, R&D budgets in pharmaceutical and biotech companies have steadily increased. The rising cost of bringing a drug to market has created an environment in which frequent blockbuster drugs are needed to sustain the modern pharmaceutical industry. Today, promising drugs are not pursued because they lack in the broad patient efficacy required to achieve this blockbuster status. To grow profitably, pharmaceutical companies must reduce the cost of getting drugs to market and enable the commercialisation of drugs that could be viable for segmented patient populations. Evolving from the current climate of blockbuster or bust will require flexible solutions that can address the needs of the drug discovery process and support downstream clinical trials. 'Next Generation Sequencing' is one technical solution that will change the DNA sequencing and genetic characterisation used in drug discovery, drug development, the clinical setting and patient management.

The Human Genome Project was an initiative that promised to address some of the needs of the medical community in terms of a more detailed understanding of human genes and their roles in disease processes. The past half century has seen a dramatic increase in the understanding of genetics in disease. This increased understanding has been spurred by the technological improvements in the study of DNA that began with the solving of the structure of DNA and the resolution of how genetic information is transmitted. The ability to determine the order of nucleotides in a fragment of DNA was pioneered by Sanger and it was the Sanger principle of sequencing via base exten-

DNA sequencing:Layout 1 17/1/08 19:13 Page 75

sion that became the standard. The Human Genome Project took advantage of the fact that sequencing reads length had increased to beyond 400 base pairs. Furthermore, the addition of automated sample preparation allowed for core facilities to significantly increase their overall sequence yield. The technical and productivity advances resulted in the completion of the Human Genome Project ahead of schedule. The project, which began in 1990 as a 15-year project, was completed in 2003 after 13 years. The Human Genome Project had as its goals to identify all the human genes, determine the sequence of DNA that make up the human chromosomes, provide storage for this information, develop tools for the analysis of the data, transfer of technologies into the private sector and the addressing of ethical issues regarding genetic information.

One of the promises of the Human Genome Project was to change the face of biomedicine. This promise has been satisfied in many ways through the study of genomics and proteomics – new ways of studying biology that arose directly from the Human Genome Project. While we have a significantly better understanding of human genetics there is still much more to be done in the areas of improving diagnosis of disease, early detection of genetic predispositions to disease, rational drug design, improved drug target discovery and pharmacogenomics (or 'custom drugs'). The promise of DNA sequencing as a tool to understand disease, as well as other fundamental biological problems, established this need for new DNA sequencing technologies that could deliver more information, in a shorter time period, for less money. In 2005, 454 Life Sciences commercialised the first new, or 'next generation' sequencing technology. This technology is described as next generation because it has reduced the cost of DNA sequencing over conventional Sanger sequencing by at least 10-fold while reducing the time to complete a typical project by 100-fold. In addition to improving efficiency, the massively parallel sequencing of hundreds of thousands of individual DNA molecules facilitates projects that were previously impossible with Sanger technology. Since the launch of the Genome Sequencer FLX there have been two additional sequencing products that have entered the market through Illumina and Applied Biosystems.

Accelerating discoveries

As R&D spending continues its steady march upwards industry-wide, the flexibility and cost

DNA sequencing:Layout 1 17/1/08 19:13 Page 76

effectiveness of DNA sequencing provides a unique opportunity to obtain the current level of data at a significantly reduced cost or to generate additional and novel data at current costs. In order to make a material impact on spending, DNA sequencing technology must be flexible enough to address a broad range of therapeutic areas and meet variable needs throughout the discovery and development process. The next generation sequencing technologies allows for applications across many different experimental areas.

The importance of this new technology is exemplified by the fact that more than 120 studies have been published since the introduction of the first next generation sequencing technology. Publication of novel information is a significant validation for the technology and many of the applications have direct relevance to drug discovery and development. Importantly, the next generation sequencing has facilitated new research approaches including the whole genome analysis of disease causing organisms, the comprehensive study of small and micro-RNA populations, the extremely deep analysis of drug resistance mutations in complex genetic settings such as HIV, the genomic characterisation of a community of micro-organisms (metagenomic) from diseased versus healthy individuals to elucidate a causative agent, the broad genetic analysis of drug targets across a population and the better understanding of chromatin organisation and epigenetic regulation of gene expression. The new sequencing platforms' capacity to address many therapeutic areas will materially affect the price of drug discovery across organisations.

The large number of individual sequencing reads produced by the various Next Generation Sequencing systems has ushered in a revolution in the study of small RNA molecules. This application has contributed significantly to the understanding of RNA interference – a biological process discovered only eight years ago. Since then, RNAi has taken the biology and pharmaceutical world by storm illustrated by the fact that the 2006 Nobel Prize in Medicine was awarded to the discoverers of RNAi. The application of high throughput, highly parallelised next generation sequencing has been at the forefront of many recent advances in RNAi research. These technologies allow for the deeper understanding of the population of small RNAs expressed in a given tissue or cell type. These applications have lead to identification of novel mechanisms of small or microRNA synthesis as well as the discovery of novel classes of small RNAs. Since 2005, when the first next generation sequencing system was introduced, there have been nearly 30 publications regarding small RNAs.

Advancements in molecular medicine due to next generation sequencing is not limited to the new study of small RNAs but also includes other areas in which the deep sequencing of a population of molecules is essential to the understanding of the molecular processes at work. Drug resistance is one area in particular that will benefit from the application of next generation sequencing. The application of deep sequencing for drug resistance has been applied to infectious agents such as Mycobacterium tuberculosis, Staphylococcus aureus and HIV. The next generation sequencing platforms make the sequencing of whole bacterial genomes a routine practice. In fact several hundred different bacteria have been partially or completely sequenced using these new technologies. Applications to antimicrobial drug development include the use of sequencing to pinpoint the precise mechanism of action of a new compound and the mechanism of resistance to that compound. In addition to the application of sequencing to the drug development process, next generation sequencing can also be applied to the clinical development of the compound. Sequencing can also be applied prior to prescribing a particular drug to determine if the patient is harbouring any bacteria resistant to that class of drug. Next generation sequencing can also be used in the healthcare setting to search for pockets of bacterial that are of particular interest due to patient risk such as Clostridium difficile. These applications can be made cost-effective because a typical bacterial genome can be sequenced to significant depth in a partial sequencing run on the next generation platforms.

The application of next generation sequencing to HIV research is extremely powerful because the virus rapidly mutates as a part of its normal biology. The massive throughput enabled by these platforms has allowed researchers to dig deeply into the metagenome of a viral population and identify all subtypes of virus present. The ability to sequence a viral genome thousands of times on a single sequencing run makes them an ideal tool for anti-viral research. As an example, at the 16th International HIV Drug Resistance Workshop, several presentations were made demonstrating that 454 Life Sciences' Ultra-Deep™ sequencing application identified previously undetectable rare drug resistant HIV variants in patient samples. One of these studies

DNA sequencing:Layout 1 17/1/08 19:13 Page 77

(Simens et al) was a retrospective analysis of drug resistance and associated virologic failure. All of the patients tested were naïve to the class of compounds under study and yet a significant portion of these patients were harbouring viruses with drug resistant mutations. The patients had been previously screened for viral genotype using conventional sequencing but the next generation sequencing found that twice as many of the patients had these drug resistant mutations. When the clinical outcomes were compared to the sequencing results it was found that a significant number of the patients with a drug resistant mutation showed virologic failure versus those patients that did not harbour a virus with drug resistance mutations. Results from this study and similar studies demonstrate that next generation sequencing has a tremendous potential to enable the prediction of early drug treatment failure and thus allow for a more patient-specific course of treatment.

Drug resistance is not the only area impacted by the application of new sequencing technologies. Human genetics in general holds the key to a better understanding of many disease processes. Currently, there is a reference human genome and a collection of more than 2 million SNPs that make an excellent toolbox of the characterisation of inherited diseases or susceptibility to disease. With these tools many research groups have been able to identify regions of importance in the human genome that may be linked to a particular disease. The problem remains how to characterise these regions, which are typically 100s of kb in size, across a large enough population to focus in on those sequences that may play a causal role. The next generation sequencing technologies are likely to provide the mechanism to complete these types of studies at a reasonable cost. Recently, two studies were published that made use of the capture of targeted sequences on a high density, programmable microarray followed by the sequencing of the captured DNAs on next generation sequencing platforms. Both of the studies demonstrated that the combination of these genomic technologies can be a powerful mechanism to carry out targeted analysis of broad regions of the human genome for population studies of disease. Overall, it is important to note that the Human Genome Project did not reveal all the secrets of the human genome but rather revealed the tip of a very large iceberg and revealed the requirement for the continued analysis of many more genomes. The next generation sequencing technologies available today and those that are promised in the future will provide the technical solutions that will make these studies possible.

DNA sequencing:Layout 1 17/1/08 19:13 Page 78

Next Generation Sequencing in the clinical setting

There are two competing forces in drug development – the desire to produce drugs with broad applicability (blockbuster drugs) and the need for increased specificity and safety of new drugs. New technologies are needed that increase specificity without dramatically increasing drug development costs to help mediate these two forces. There are excellent examples of compounds with good economic profiles that are effective for patients with specific genetic profiles. Genentech's Herceptin, AstraZeneca's Iressa and Novartis' Gleevec are among these.

Integration of DNA sequencing into the drug discovery process will allow the identification of specific patient populations as well as identifying diagnostic and/or theranostic markers. DNA sequencing offers the most reliable and accurate method of grouping individuals into characteristic genetic profiles. Sequencing of disease-associated regions enables the differentiation of genetic profiles, regardless of the underlying genetic changes. Up until now, DNA sequencing has been of limited use in clinical trials because of the prohibitive cost and amount of time associated with sequencing the hundreds of individuals enrolled in a single trial. Today, the economics are shifting in favour of sequencing because of the massive throughput of next generation sequencing. Parallelisation of samples processing can further facilitate the stratification of patient populations via the analysis of specific genomic regions. The pharmaceutical industry is on the cusp of experiencing clinical trials stratified by the genetic profile of its subjects. Next Generation Sequencing is poised to enable this paradigm shift through fast, inexpensive, and accurate sequencing. The ability to stratify clinical trials will reduce the cost of getting drugs to market and will make drugs more specific to their target populations. Furthermore, sequencing will also enable a more sophisticated choice of drugs based on a patient's or infectious organism's genetic background. For example, the choice of anti-retroviral treatment based on the genetic make-up of the viruses infecting an individual is a real possibility through the deep sequencing of the viral genomes present. Information regarding potential drug resistant viruses can be used by the physician to choose the most effect treatment course.

The next generation sequencing technologies offer the combination of speed, cost and accuracy demanded to meet the growing need for the genetic analysis of patient populations. The application of these technologies in the future during the drug discovery and development processes, as well as follow-on patient care, should provide significant improvement to the process including patient stratification and identification of disease specific genetic markers. DDW

Dr Bruce Taillon *is currently Director of Healthcare Business Development for 454 Life Sciences. His career spans more than 20 years and has focused on the application of molecular technologies to solve biotechnological problems. Prior to his present position, Dr Taillon was Director of Technology at Curagen Corporation where he was responsible for the establishment of technological solutions to drug discovery problems.*

Brendon Hill *is presently Manager of Global Marketing and Communications for 454 Life Sciences. Mr Hill graduated from Yale University with a degree in Molecular Biology.*

ADVERTISEMENT INDEX

BD Biosciences	25	Cyprotex Plc	31	MDS Analytical Technologies	59
Beckman Coulter, Inc	6	Eppendorf AG	41	Society for Biomolecular Sciences	IBC
BioFocus DPI Plc	53	GE Healthcare	4,16	Tecan Schweiz AG	13
Biotek Instruments, Inc	15	Genetix Ltd	21	Thermo Fisher Scientific	61,OBC
BMG Labtech Gmbh	68	HTStec Ltd	11	V&P Scientific, Inc	45
CISBIO International SA	57	ID Business Solutions Ltd	IFC-3,26		
Covaris, Inc	34	Millipore Corporation	8		