BioTechnologia vol. 92(1) • pp. 11-12 • 2011
Journal of Biotechnology, Computational Biology and Bionanotechnology PERSPECTIVES

Progress of medically-related biotechnology in the post-genomic era

PIOTR KOZLOWSKI
Institute of Bioorganic Chemistry, Polish Academy of Sciences, Poznań, Poland
e-mail: kozlowp@ibch.poznan.pl

Each day, the enormous progress in life sciences brings about new discoveries which substantially influence the medical biotechnology nowadays and will shape its directions in the near future.

Genome

A good illustration of the above statement is the progress made in the field of genetics and genomics accomplished within just one decade after the first draft of the human genome sequence was announced. In the following part of the article I will refer to this decade as the post-genomic era (PGE). Before the PGE, we had known the genomic sequences of only four Eukaryotes yeast, fly, nematode, and *Arabidopsis*) and not more than 40 Prokaryotes. Today, we have sequenced 250 genomes of eukaryotes and about 4,000 genomes of prokaryotes and viruses as well as many metagenomes of different environments (e.g. ocean, human gut). What is more, as a result of the 1000 Genomes Project, over 500 individual human genomes have been re-sequenced, and the ideas such as 'a $1000 genome' or the sequencing of additional 10,000 genomes of vertebrates are not pure science fiction any longer. The reference sequence of the human genome serves as a scaffold on which any piece of information, including the results of our everyday experiments can be laid. This facilitates the creation of new maps (with nucleotide resolution) of different genetic features. As of now, we have numerous maps and these include conservation maps, repeat element maps (e.g. RepeatMasker), maps of non-coding RNA, maps of cancer genes, cancer aberration maps (e.g. Cancer Genome Atlas), structure conservation map (EvoFold), maps of genetic variants (e.g. dbSNP, HapMap, DGV), epigenome maps, 3D folding map and many others. We can easily access all this information through user-friendly genome browsers such as UCSC at the University of California, Santa Cruz, USA.

During the PGE, a number of cloned genes associated with mendelian disorders was doubled to the result that today there are almost 3,000 of them. Even a more significant progress has been made in terms of the identification of the common genetic variants associated with complex phenotypes/diseases. Before the PGE about 20 such variants had been known and the *APOE* alleles associated with the Alzheimer's disease was one of them. Now, there are over 1,000 of genetic variants reproducibly correlated with different human diseases. Over the period of the PGE the number of known SNPs has increased dramatically from about 1 million to over 20 millions. A substantial fraction of the currently known SNPs is well characterized in terms of their frequency, LD relationships and ethnic specificity, which substantially facilitate their use as genetic markers and help to understand their role in the modification of the phenotype. All the achievements mentioned above would not have been possible without the development of the breakthrough technologies of which the most important ones are: the massively parallel hybridization (microarrays) and the next generation sequencing (NGS). Microarray technologies allow highly multiplexed genotyping, expression profiling, precise analysis of aberrations in cancer genomes and genome-wide identification of copy number variants (CNVs). Even higher expectations are associated with the recent advances in sequencing technologies and with the development of several platforms of NGS. All these platforms take advantage of the massively parallel sequencing of individual molecules of DNA, either directly or after their clonal amplification (e.g. emulsion PCR, bridge PCR) which allows to generate hundreds of thousands or even millions of sequencing reads in one ex-

periment. Initially, NGS platforms were applied mostly for high throughput sequencing (e.g. *de novo* sequencing of bacterial genomes and complex genomes re-sequencing). Presently, due to enormously high multiplicity and very low (per read) cost, the array of NGS applications has substantially widened and includes expression profiling, epigenetic modification profiling, RNA structure analysis, short RNA identifications, SNPs genotyping, identification of chromosomal breakpoints and many others. Direct consequences of technological progress are an enormous increase (100,000 times comparing to pre-genomic era) in the throughput and a reduction of cost of sequencing. The same tendencies apply to the throughput and cost of SNP genotyping. For example, while 10 years ago genotyping of a single or several SNPs at a time was a standard, the nowadays routine is parallel genotyping of hundreds of thousands or even millions of SNPs.

Non-coding RNA

Another great accomplishment of the last decade that can influence medical biotechnology is discovery of transcriptional activity of a vast majority of human genome, which gives rise to many new classes of non-protein-coding (non-coding) RNAs, including small interfering RNAs, microRNAs, PIWI-interacting RNAs and various types of long non-coding RNAs (e.g. lincRNAs). It has been estimated that human genome encodes about thousand of microRNAs and several thousands of long non-coding RNAs. Long non-coding RNAs account for at least 80% of transcripts in mammalian cells. Although the function of non-coding RNAs is far from being ultimately explained, it is clear nowadays non-coding RNAs play an important role in the transcriptional and post transcriptional regulation of gene expression as well as in the regulation of epigenetic modification. Although our current knowledge about non-coding RNA is only the tip of the iceberg, it has substantially changed our conventional definition of a gene as a region of genome encoding mRNA that is translated into protein.

Stem cells

The recent progress in the stem cell research also opens new perspectives for medicine. The use of stem cells gives an opportunity to regenerate or repair tissues affected by diseases or traumas and to develop therapies for diseases such as spinal cord injury, diabetes, macular degeneration, myocardial infarction, neurodegenerative disorders and inflammatory diseases. Widespread ongoing clinical trials carefully evaluate potential benefits and drawbacks of treatments with the use of stem cells. Particularly great interest is associated with the use of induced pluripotent stem cells (iPS cells) which have the potential of embryonic stem cells but their utilization is free of ethical controversy.

Biomaterials

Finally, a great progress has recently been made in the area of biomaterials and their applications in regenerative medicine, drug delivery, construction of mechanical component used in surgery, and construction of functional nanoparticles. One of such fascinating examples is spider silk whose unusual mechanical properties, biocompatibility, biodegradability and modular structure open endless possibilities for engineering and generation of new materials with very specific and predictable properties.

All the achievements mentioned above are only a fraction of the recent progress in life sciences which can help to understand the mechanism of diseases, to develop new diagnostic tools, and to search for new and more effective treatments and therapies to ultimately improve the quality of human life.

Acknowledgements: Ministry of Science and Higher Education, Grant No. N N302-278937