Implementation of postgenomic technologies for cancer research. Institutional experience

Martina Rudgalvytė1, 2, Diana Stančiūtė1, Kristina Bielskienė1, Violeta Jonušienė2, Vida Kirvelienė2, Kęstutis Sužiedėlis1, 2, Vydmantas Atkočius1, Konstantinas Povilas Valuckas1

1 Institute of Oncology, Vilnius University, Vilnius, Lithuania
2 Department of Biochemistry and Biophysics, Vilnius University, Vilnius, Lithuania

Advance in molecular biology and the new technologies for biomedical research are being rapidly introduced into the research of complex pathologies worldwide. Implementation of these technologies, however, needs substantial financial resources for the equipment and for training the specialists. The rapid development of biomedical research over the past decade increases the risk of moral ageing of the implemented technologies and raises doubts as to whether countries with limited financial resources could afford them. In this article, we share our institutional experience in the implementation of post-genomic technologies in cancer research in Lithuania and stress the need of modern infrastructure in biomedical research, despite the needed efforts and associated risks.

Key words: postgenomic technologies, cancer research

INTRODUCTION

The completion of the Human Genome Project (1, 2) and the onset of the postgenomic era have resulted in the development of new genomic and proteomic research technologies. Application of these technologies resulted in numerous global research projects which became a characteristic feature of biomedical research in the postgenomic era (3–5). The exploratory design of global research projects is an optimal choice in the research of complex pathologies (such as cancer); therefore, it is not surprising that high-throughput screening techniques have been extensively used in cancer projects, aiming to identify markers for early cancer detection and molecular-targeted treatments (3–11). Due to the higher efficiency of the exploratory versus the hypothesis-driven design, the global research projects monitoring the differential gene expression have rapidly resulted in identification of the genes and pathways that are dysregulated in a variety of human cancers, allowing a better understanding of malignization mechanisms and creating molecular tools for cancer prognosis (12–14). In brief, advances in molecular biology have set new standards for biomedical research in the postgenomic era and formulated the necessity of the postgenomic research infrastructure.

In this paper, we review postgenomic technologies and the available facilities dedicated to cancer research at worldwide recognized cancer research centers, share our experience in the implementation of postgenomic research infrastructure at the Institute of Oncology, Vilnius University and at Vilnius University, and evaluate the potential of the implemented infrastructure in the light of the rapidly developing research in biology.
GENOMICS AND PROTEOMIC TECHNIQUES FOR CANCER RESEARCH

New genomic and proteomic high-throughput technologies are major achievements in biology and medicine in general, contributing to the understanding of carcinogenesis. New molecular insights and technologies in oncology have given clues to early detection and prevention of cancer, prognostic and predictive markers, and targets for therapy (6). Genomics, based on high-throughput techniques, refers to a comprehensive analysis of gene number variations or the expression of a large number of genes or genome as a whole, i.e. to the analysis of the genetic content of an organism (7). For years, scientists have studied one gene at a time, selecting the genes by the hypothesis-driven approach and analyzing the impact of this gene in an isolation from the larger context of other genes. Unfortunately, such project design results in neither confirmation nor rejection of the hypothesis. In the past several years, a new technology, called DNA microarray or DNA chip, has been developed. This technology is based on an orderly arrangement of a great number of specific probes in a reduced space, allowing large-scale studies and monitoring of the whole genome on a single chip. In this way, researchers can have a better picture of simultaneous interactions among several genes, and this project design eliminates the need for the hypothesis-driven research, resulting in the exploratory research projects which are much more efficient (5, 7). Under pathological conditions, the cellular activity is dynamically regulated through specific changes in gene expression. Hence, microarray-generated expression data can be used as a standard gene expression contents of a particular cell phenotype. Such defined and specific profiles of gene expression are referred to as molecular signatures, which could be useful to identify certain clinical conditions, the class or phase of diseases (8, 9).

Microarray technologies in profiling tumors at the DNA, RNA and protein levels have led to the discoveries of disease susceptibility genes, therapeutic targets and expression profiles related to disease outcomes, drug sensitivity and resistance. A detailed molecular classification of individual patient tumor will provide a means for the selection of individual therapy best suited to an individual patient and based on his molecular profile. This detailed molecular classification will open the way for molecular medicine.

In the last few years, new methods allowing the analysis of cellular proteins in a new perspective (as a whole set), i.e. proteomics, have emerged. Not only complexes of proteins, their structure, interaction, expression in a biological system, but also post-translational modifications can be analyzed (6). Genomics, based on high-throughput techniques, refers to a comprehensive analysis of gene number variations or the expression of a large number of genes or genome as a whole, i.e. to the analysis of the genetic content of an organism (7). For years, scientists have studied one gene at a time, selecting the genes by the hypothesis-driven approach and analyzing the impact of this gene in an isolation from the larger context of other genes. Unfortunately, such project design results in neither confirmation nor rejection of the hypothesis. In the past several years, a new technology, called DNA microarray or DNA chip, has been developed. This technology is based on an orderly arrangement of a great number of specific probes in a reduced space, allowing large-scale studies and monitoring of the whole genome on a single chip. In this way, researchers can have a better picture of simultaneous interactions among several genes, and this project design eliminates the need for the hypothesis-driven research, resulting in the exploratory research projects which are much more efficient (5, 7). Under pathological conditions, the cellular activity is dynamically regulated through specific changes in gene expression. Hence, microarray-generated expression data can be used as a standard gene expression contents of a particular cell phenotype. Such defined and specific profiles of gene expression are referred to as molecular signatures, which could be useful to identify certain clinical conditions, the class or phase of diseases (8, 9).

There are two steps in the proteomic analysis: separation of the protein mixture and identification of the separated proteins by various analytic methods. The most efficient and most widely used protein identification method in proteomics is the multi-dimensional liquid chromatography, coupled to electrospray ionization tandem mass spectrometry (LC-ESI MS / MS) exploiting time of flight (TOF) or ion trap mass spectrometry. (6, 7, 10). Also, the new emerging proteomic technology such as protein microarrays allows to investigate a several targets in a signaling pathways simultaneously. Protein microarray technology is used to study antibody-antigen, protein-protein, protein-nucleic-acid, protein-lipid and protein-small-molecule interactions (6, 10). Finally, the bio-informatics tools allow to analyze the results and to identify the possible biological targets (7, 10).

DEVELOPMENT OF INSTITUTIONAL INFRASTRUCTURES AND TRANSLATIONAL CANCER RESEARCH

The last 25 years were rewarding in a biomedical research. The new molecular biology technologies contributed much to a better understanding of the mammalian, especially human, systems. However, these advances in basic biomedical research showed only a few examples of a successful application of the findings of such research in medical practice. Such applications are commonly described as 'translation', as the "process of translating discoveries in the laboratory into clinical interventions for the diagnosis, treatment, prognosis, or prevention of disease with a direct benefit to human health" (15, 16).

The development and maintenance of infrastructures of scientific laboratories is considered a major goal for academic centers promoting translational cancer research programs. Among infrastructures favoring translational research, centralized facilities characterized by a shared, multidisciplinary use (by different departments, divisions, research units) of expensive laboratory instrumentation, or by a complex computer hardware and software, and / or by high professional skills are necessary to maintain or improve institutional scientific competitiveness (16, 17).

According to the National Cancer Institute (NCI) Cancer Centers program (18), the majority of cancer centers possess at least the following shared resources: centralized equipment; general and specialized animal colonies; nucleic acid sequencing / synthesis labs; amino acid analysis HPLC facilities; cell sorting (flow cytometry) facilities; chemical and drug synthesis labs; mass spectrometry labs; electron microscope facilities; media preparation; microarrays and proteomics facilities and services; histology and pathology services; experimental radiation facilities and services; biostatistics. The additional, less represented, shared resources are quite heterogeneous and depend on the scientific orientation of each center (i.e. more clinical- or basic research-oriented) (17). In the following paragraphs, will be briefly summarize some of the mentioned shared resources and their necessity for developing effective translational research programs in oncology.
Genomic technologies. Its an important tool for analyzing a large number of genes. They allow analyzing their structure (wild type, mutations, deletions) and regulation by assessing the RNA content. Also, epigenetic and karyotypic analysis could be made (7). In particular, DNA microarrays are widely used for the diagnostics, prognostics and predictions of response to therapy in various cancers (17, 20).

Proteomic technologies. The methods of proteomics includes detection, identification, and measurement of proteins and/or peptides, protein modification analysis, the study of protein–protein or protein–DNA interactions and regulation. Proteomic application to cancer provides an important information on biomarkers for an early detection of tumor development, tumor profiling for diagnostic and staging purposes, and on the mapping of cancer signaling pathways aimed at developing new treatments (17, 21).

Confocal microscopy. Confocal microscopy permits collection of three-dimensional images from living or fixed cells and tissues by the use of laser scanning technology. This technique is popular in biomedical cancer research because it allows an analysis of several processes of tumorigenesis, such as angiogenesis and its inhibition by biological molecules (17, 19).

Laser capture microdissection. By the laser capture microdissection (LCM) technique it is possible to separate tumor, stromal, and normal cells within a single biopsy specimen. High-throughput analysis of microdissected specimens allows for a clean discrimination of events occurring in and between each of these tissue microcompartments. Application of these technologies to patient samples allows dissection of genomic changes, expression events, and the differential expression, activation, and signaling of a variety of proteins in tumor samples (4, 22).

Cell sorting (flow cytometry). This method can provide rapid, quantitative, multivariable analyses of a single living or dead cell, based on the measurement of fluorescent light emission. Flow cytometry can be used for a simultaneous analysis of surface and intracellular molecules. DNA analysis is the second most important application of flow cytometry. By measuring the DNA content in individual cells, information about their ploidy and the distribution of cells across the cell cycle is obtained. Increasing the knowledge of various cell cycle kinetic parameters is of considerable importance for several purposes, including tumor diagnostics and treatment (23, 24).

Biobanking. Biobanking is an emerging activity that includes the collection and preservation of biological samples (tissues, cells, serum, plasma, nucleic acids). The collection of human specimens is situated at the beginning of the chain of translational research; therefore, biobanks are actively contributing to advances in translational research by offering opportunities to safely collect and store these samples and link laboratory research to clinical practice, ultimately accelerating the development of personalized medicine (17, 25).

Bioinformatics. Bioinformatics emerged as an essential discipline with the development of high throughput genomic technologies a few years ago. Nowadays, skills in bioinformatics are highly required in the institutions that develop research programs based on high throughput technologies and obtain results based on large quantities of data, such as genomics and proteomics (17, 26).

Research infrastructure is an essential tool in developing successful programs in translational research. Each center needs clear policies as regards the development and the rules governing the establishment of shared resource and the availability of financial resources to set up and maintain these facilities. It is expected that the overall advances in genomic and proteomic technologies will help to guide our judgment as to the best treatment of each individual patient.

ESTABLISHMENT OF POSTGENOMIC RESEARCH INFRASTRUCTURE AT THE INSTITUTE OF ONCOLOGY (VILNIUS UNIVERSITY) AND AT VILNIUS UNIVERSITY

The development of strategies of a more efficient treatment of cancer patients is the goal of researchers of the Laboratory of Molecular Oncology of the Institute of Oncology, Vilnius University. In order to better understand the tumor cell response to the treatment, it is necessary to identify the molecular mechanisms that determine the response. Identification of the global gene expression signatures is the most straightforward way to characterize the cell and its response; therefore, the necessity of the implementation of postgenomic technologies is obvious.

To benefit from the advantage of postgenomic research technologies, researchers of the Institute of Oncology of Vilnius University in collaboration with Vilnius University and the Institute of Biochemistry of Vilnius University have settled the minimal genomics and proteomics research infrastructure implementing two EU SF projects. The established infrastructure included a DNA microarray spotter, a DNA microarray hybridization station, DNA microarray scanner and a terminal for the preparation of samples for mass spectrometry analysis.

The terminal of the Proteomics Centre allows the initial separation of protein samples, which could be further characterized by mass spectrometry at the main laboratory of the Centre. The Proteomics Centre, consisting of the main laboratory (located at the Institute of Biochemistry of Vilnius University) and five terminals for sample preparation (one of them is at Institute of Oncology of Vilnius University) were established by the Institute of Biochemistry of Vilnius University and partner institutions during the implementation of EU SF projects. The main laboratory of the Proteomics Centre is equipped with liquid chromatography systems for protein separation. A two-dimensional electrophoresis system for protein separation, mass spectrometers for protein identification and analysis, a laser scanning confocal microscope with a microinjection system for protein microinjection and investigation of protein localization in the cell. Terminals
for sample preparation are equipped with liquid chromatography systems for protein separation. The infrastructure of the Proteomics Centre allows protein identification, purity and stability analysis; peptide de novo sequencing; protein–protein complex formation analysis; identification of protein chemical modification; proteome differential screening and hypothesis-driven protein analysis; protein microinjection and investigation of protein localization. Therefore, this infrastructure enables a discovery and verification of biomarkers as well as investigation of protein function, which is crucial in cancer research.

An important feature of the established infrastructure for the genomic research is the flexibility of the suitable DNA microarrays. In fact, various commercial DNA microarrays (for example, produced by Agilent) and DNA microarrays produced using a DNA microarray printer are suitable for the analysis. The possibility to select the scale of genome coverage for laboratory-printed DNA microarrays and to prepare the arrays dedicated not for initial screening but for a secondary selection of the initially preselected genes (arrays for the analysis of 2000–5000 genes) is an important advantage of the established infrastructure and makes it valuable, despite the rapid development in the arrays technology.

The established infrastructure, however, was not complete and was lacking a few important structural components and the qualification of personnel to use the infrastructure. This problem was solved during the implementation of the project “Dynamics of prognostic and predictive markers for oncology” (No. 2004-LT0057-1P-1NOR). The implementation of this project allows installing a set for cell cultivation, a laser capture microdissection system, a real-time PCR cycler, an automatic nucleic acid (RNA and DNA) isolation system and a microcapillary electrophoresis unit to evaluate the quality of the isolated nucleic acids.

At the Department of Biochemistry and Biophysics of Vilnius University, which is the partner of the project, a system for flow cytometry was established. It allows determination of differential phenotypes of cells by measuring their apoptosis at different stages, analysis of surface molecules and intracellular cytokines at a single-cell level, and determination of proliferative responses by cells.

Training of specialists at the Norwegian Microarray Consortium, supported by the project, resulted in the optimal transfer of research technology.

Impact of the established infrastructure in the light of rapidly developing post-genomic technologies. Recently, methods like Western and Southern blotting, DNA sequencing or polymerase chain reaction (PCR) have been used in a clinical diagnostic laboratories as modern an effective molecular tools. Nowadays, more-parallel molecular tests that involve a simultaneous analysis of thousands of genes are required. Microarray analysis is a new technology that has evolved from Southern blotting and Comparative genomic hybridization. For the last two decades, the microarray technology has developed into a powerful molecular diagnostic tool which already allows analyzing genes not only at the mRNA level, but also at DNA methylation and protein levels. First works with microarray technology came along with the BAC (bacterial artificial chromosome) plasmids which have been used to sequence the genome of organisms in different genome projects. These plasmids can contain inserts from 100 to 700 kbp and serve as a basis for hybridization analysis. Mostly BAC-arrays were used for comparative genomic hybridization (27). Later, modern platforms emerged, such as cDNA arrays in which PCR products are placed on an array surface and on an oligonucleotide array, in which chemically synthesized short fragments of nucleic acids of a definite chemical structure are attached on an array surface using robotic arrayer (28).

Revolutionary changes in the microarray technology were made in the last few years by several companies (Illumina, Affymetrix, Applied Biosystems). The advance in gene expression analysis tools enables a rapid profiling of the transcriptome. One of the most widely used platforms is arrays by Affymetrix Inc., in which DNA oligonucleotides are chemically synthesized in situ on an array surface, using photolithography (29). The other producer of next-generation arrays is Illumina which offers an innovative BeadChip technology allowing to get sharp signals and to characterize all transcriptional activity, both coding and non-coding. These new-generation array platforms are more suitable for routine clinical laboratory testing (30).

For more than 10 years, the majority of technical platforms for gene expression measurements have been based on the microarray. The DNA microarray technology has given rise to several successful clinical applications, such as the Amsterdam 70-gene breast cancer gene MammaPrint signature (31) which was cleared by the U. S. Food and Drug Administration (FDA) in February 2007 (32). Global analysis of gene expression is reliant on RNA hybridization on high-density arrays, and this allows the profiling of many tissues, although detecting only specific sequences. Whole-genome arrays theoretically allow capturing much of the complexity of the transcriptome, but they ignore splice-junction information. From the technical perspective, however, the reliance of the microarray on nucleic-acid hybridization results in several limitations, such as background and cross-hybridization problems. More and more researchers begin taking notice of the next-generation sequencing technology (33).

Next-generation sequencing (NGS) platforms provide a possibility to expand the genomic methods to a new scale. Depending on the technology, these platforms can produce gigabases of sequences per day. Due to its resolution and sensitivity, NGS is increasingly used to replace array technologies, particularly in gene expression profiling experiments. Finding many more significant genes not interrogated by the microarrays stresses the necessity of a more comprehensive transcript profiling by the next-generation sequencing-based methods (34).

NGS technologies have an impressive range of applications including: a) full-genome resequencing or a more
targeted discovery of mutations or polymorphisms; b) mapping of structural rearrangements, which may include translocation breakpoints and chromosomal inversions; c) ‘RNA-Seq’, analogous to expressed sequence tags (EST) or a serial analysis of gene expression (SAGE); d) large-scale analysis of DNA methylation by deep sequencing of bisulfite-treated DNA; e) ‘ChIP-Seq’, or the genome-wide mapping of DNA–protein interactions (35).

These applications can be used to characterize the evolutionary relationships of ancient genomes and to elucidate the role of noncoding RNAs in health and disease. In the not too distant future, it is foreseeable that NGS technologies could be used to obtain high-quality sequence data from a single cell, which would be an achievement, particularly for cancer genomics. However, analysis of these massive and heterogeneous deep sequencing data offers several challenges, including effective data mapping, annotation and visualization, efficient data storage, integration and interpretation of data from multiple technological platforms, tissues and cell lines. For example, different platforms exhibit different sequencing errors at different rates, and currently these errors are only partially understood. It is also clear that the processing of the input samples to prepare the final libraries that are sequenced and currently poorly understood will take time, and likely many technical improvements will be made over the coming years to reduce these errors (36).

However, whereas on microarrays there has been a large amount of work done with the analysis of data, including the issues of normalization, noise analysis, sequence-composition errors, background corrections and etc., the deep-sequencing-based expression analysis is still under development, and no standardized analysis protocols have been composed yet (36). It is important to compare the data quality of microarray and NGS as well as to reveal the advantages and disadvantages of each platform. The development of the NGS is not the ultimate goal of the microarray technology; a the combination of microarray and sequencing-based platforms will definitely push forward the development of research on gene expression and regulation.

An overview of the development of technologies for genomic research indicates a tremendous change in the field upon the introduction of microarray techniques. Microarray techniques have increased the productivity, because, instead of doing gene-hunting efforts, scientists may identify many differently expressed genes (from tens of thousands) in one experiment.

The benefits from the implementation of post-genomic research technologies at Lithuanian institutions are not just the availability of these technologies for Lithuanian researchers, but rather the gain of competence in the field, which is associated with the new possibilities for upgrading the infrastructure and the further development of research in the field in Lithuania with the further developments of technologies. The decision to implement the infrastructure is nevertheless associated with a risk that the implemented equipment will become outdated immediately after implementation due to the rapid development of technologies. Evaluation of the level of the current development of technologies for genomic research allows predicting the efficient use of the microarray technology infrastructure, at least for the next decade, indicating that the risk of selecting the equipment for implementation has been well managed by the consideration that microarray technology is a milestone dividing the biomedical research into pre and postgenomic research.

Therefore, we can conclude that the modern and complete postgenomic research infrastructure (allowing isolation of a single cell and a well-controlled automatic investigation and global characterization of a tumor cell) has been established at the Vilnius University and its institutes (Institute of Oncology, Institute of Biochemistry). It is already used for investigating genes involved in changes of cellular sensitivity to the treatment as well as in the search for predictive and prognostic biomarkers in different cancers. Its infrastructure is available for different applications of genomic and proteomic analysis and, despite the risk of moral ageing, it has the potential to speed up the development of postgenomic research in Lithuania.

ACKNOWLEDGEMENTS

This project was supported by the EEA grant of the Norwegian financial mechanism, No. 2004-LT0057-IP-1NOR.

Received 10 November 2010
Accepted 7 December 2010

References

Implementation of postgenomic technologies for cancer research. Institutional experience

105


VĖŽIO TYRIMŲ POGENOMINIŲ TECHNOLOGIJŲ İDIEGIMAS LITUVOJE

Santrauka

Ką tik sukurtos naujos bei pažangios biomedicininių tyrimų tyrimo technologijos tuoja pat yra naudojamos kompleksinių patologijų tyrimams visame pasaulyje. Šioms technologijoms ir jų infrastruktūrų kūrimui reikia ženklių investicijų (aparatūra, specialistų parengimas), todėl stengiamės neprasidėti ir visai iš anksto apsvarstyti šias technologijas. Šiame straipsnyje aptariamame savo patirtį kūrimui tikėtis ir planuoti kompetencijos ir technologijos įgyvendinimą Lietuvoje.

Raktažodžiai: pagrindiniai ir pogenominiai tyrimų, vėžio tyrimai