
REVIEWS

UDC 577.2.616-006;577.2.08

Methods of Searching for Markers for Serological Serum Diagnosis of Tumors

Yu. A. Bukurova, G. S. Krasnov, I. G. Nikitina, V. L. Karpov, N. A. Lisitsyn, and S. F. Beresten

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia;

e-mail: bukurova@eimb.ru

Received October 4, 2011; in final form, November 11, 2011

Abstract—This review describes the most popular methods of searching for serological markers of tumors that are used in a clinical setting, as well as a comparison of their efficiency.

DOI: 10.1134/S0026893313010032

Keywords: proteomics, tumor markers, bioinformatics, microRNA, noncoding RNAs

INTRODUCTION

In a report by the United Nations, it was noted that oncological diseases hold the first place (followed by diabetes, heart disease, and lung) in the statistics of mortality rates in developed countries. It is conventionally thought that the early diagnosis of oncological diseases, including serum diagnostics, largely determines the success of treatment and allows the survival rate of patients to be improved significantly. However, in clinical practice, to diagnose tumors, only a few dozen markers are still used, which appears to be associated with the lack of efficiency of the existing search methods. Modern methods are based on the comparative analysis of proteomes, transcriptomes, or genomes of normal and tumor tissues (groups of methods are listed according to the frequency of their use in clinical practice, Table. 1). This review describes the most popular methods of searching for serological markers of tumors that are used in a clinical setting, as well as a comparison of their efficiency.

PROTEOMICS METHODS

Proteomics methods allow one to analyze the changes in protein content, as well as processing and post-translational modifications. The most commonly used in vitro methods of proteomics, which are based on the fractionation of proteins of normal and tumor tissue, followed by the mass-spectrometric identification of proteins, the content of which changes the most noticeably, often in tumors. The second group includes methods of identifying proteins that can bind to autoimmune antibodies, which are present in the blood serum of cancer patients, but not in the blood of healthy donors (reverse proteomics). Finally, the methods of proteomics in silico are designed to iden-

tify variations in the content (or processing) of the mRNA of tumors compared with normal tissues. These methods are related to the field of proteomics, since only proteins encoded by mRNA are used for serological diagnostics (mRNA rapidly degrades in blood and other body fluids).

The first and third approaches allow one to detect tumor-specific proteins that can enter the bloodstream (e.g., carcinoembryonic antigen), or secretory proteins, the content of which in the blood increases dramatically upon the occurrence of a tumor (such as specific for prostate cancer antigen). These markers are typically detected using high-affinity antibodies and immunofluorescence analysis in sandwich format. A promising approach to the serological diagnostics of cancer is the analysis of exosomes (which are surrounded by the bilayer membrane mini-cells without a nucleus and mitochondria approximately 100 nm in size), which are secreted into the blood by tumor cells. The use of antibodies to tissue-specific membrane proteins of tumor exosomes can significantly increase the sensitivity of the test by removing exosomes secreted by blood cells and dramatically increase the sensitivity of the test as a result of the amplification of the signal since, on the surfaces of exosomes, there are thousands of molecules of the protein marker.

PROTEOMICS in vitro

Two-Dimensional Gel Electrophoresis (2DE)

The method of two-dimensional gel electrophoresis (2DE) includes two main procedures, i.e., the separation of proteins using isoelectric focusing in tubes in one direction (separation on the basis of charge), and the electrophoretic separation of the contents of the tube in a denaturing polyacrylamide gel electrophoresis in the perpendicular direction (separation based on weight in the range of 10–200 kDa). For the

Abbreviations: ncRNA, noncoding RNA.

first time, the technique was described in 1970 by Koenig et al. [13], after which O'Farrell [14] proposed that it be used for the isoelectric focusing of proteins of the nonequilibrium pH gradient, which is formed during electrophoresis by ampholines (Fig. 1, top). The subsequent identification of protein spots with different intensities on the received pairs of 2DE-electrophoregrams of normal and tumor tissues is performed using mass spectrometry. Although 2DE is very laborious and has a low sensitivity (only 1000–3000 high-copied proteins of the proteome are detected), for decades, it has been the most highly effective and commonly used method of analyzing proteomes.

Two-Dimensional Differential Gel Electrophoresis (2D DIGE)

The method of differential gel electrophoresis (DIGE) is the method of the simultaneous separation of proteins of healthy and neoplastic tissues, as well as internal control (a mix of healthy and neoplastic tissue in a certain ratio), to identify protein spots only typical of tumor tissue. The proteins are labeled with fluorescent cyanine dyes with different wavelengths of the emitted radiation (the norm is usually Cy3, the tumor is Cy5, and the control is Cy2) [15, 16]. Labeled proteins are combined and separated by 2DE. 2DE analysis allows one to reveal tumor-specific proteins (Fig. 1, bottom) and identify them in selected spots by mass spectrometry. The method also allows one to estimate the amount of protein in a spot on the ratio of fluorescence intensities of spots in normal, tumor, and control tissues.

Surface-Enhanced Laser Desorption/Ionization (SELDI)

In 1993, the method of the surface-enhanced laser desorption/ionization (SELDI) was described for the first time [17]. SELDI is the most frequently used modification of MALDI technology [18]. SELDI combines the separation of proteins by adsorption to spots on the surface of a microchip treated in different ways with the subsequent removal of unbound proteins and the mass-spectrometric identification of proteins bound to one of the surfaces of the microchip (Fig. 2) [19]. The identification of proteins bound to the surfaces of the microchip is based on the definition of masses and charges of the produced positively charged peptides that fall on the surface of the negatively charged detector. The main advantage of SELDI is the possibility to quickly analyze proteomes of biological and clinical samples [20] with hardly any preliminary preparation. [21] The use of SELDI has not yet led to the discovery of an important marker of tumors due to a number of unresolved technical problems [22].

METHODS OF REVERSE PROTEOMICS

Serological Identification of Antigen by Recombinant cDNA Expression Cloning (SEREX)

The method of serological identification of antigen by recombinant cDNA expression cloning (SEREX) is based on comparing the results of the immunological analysis of two identical copies of the bacterial expression library of tumor cDNA using the blood serum of cancer patients and healthy donors. Tumor-specific clones are sequenced to identify tumor antigens (Fig. 3) [4, 23, 24]. The method allows one to identify a subset of tumor antigens that react with the antibodies of cancer patients, but do not react with the serum of healthy donors. An analysis of the structure, expression profile, and frequency of serological reactivity of identified antigens allows one to select the most promising diagnosis of tumors markers for the serum, as well as for the prediction of post-operative monitoring and the creation of antineoplastic vaccines [6].

ProteomeX Method

The method is based on the 2DE separation of proteins of normal and tumor tissues (two gels). Proteins are transferred to membranes and incubated with antisera of cancer patients, then with secondary antibody conjugates to human immunoglobulin, the revealed tumor-specific antigens were identified by mass spectrometry (Fig. 3) [5, 25–28]. Like other methods of proteomics in vitro, ProteomeX can detect the tumor-specific post-translational modification of proteins that play an important role in the development of cancer and autoimmune reactions, such as glycosylation, sumoylation, acetylation, and phosphorylation. However, this analysis requires additional technology and sophisticated software. For example, in the analysis of phosphorylation, it is necessary to supplement phosphorylated proteins with cation-exchange chromatography or conduct a comparative analysis of mass spectra before and after the removal or modification of the phosphate groups [29–32].

Autoantibody-Mediated Identification of Antigens (AMIDA)

Autoantibody-mediated identification of antigens (AMIDA technology) is based on the parallel immunoprecipitation of proteins of the tumor tissue using serum antibodies from healthy donors and sera from cancer patients. Immunoprecipitates are separated by 2DE, electrophoregrams are stained with silver salts, and tumor-specific proteins are identified by mass spectrometry (Fig. 3) [33–38].

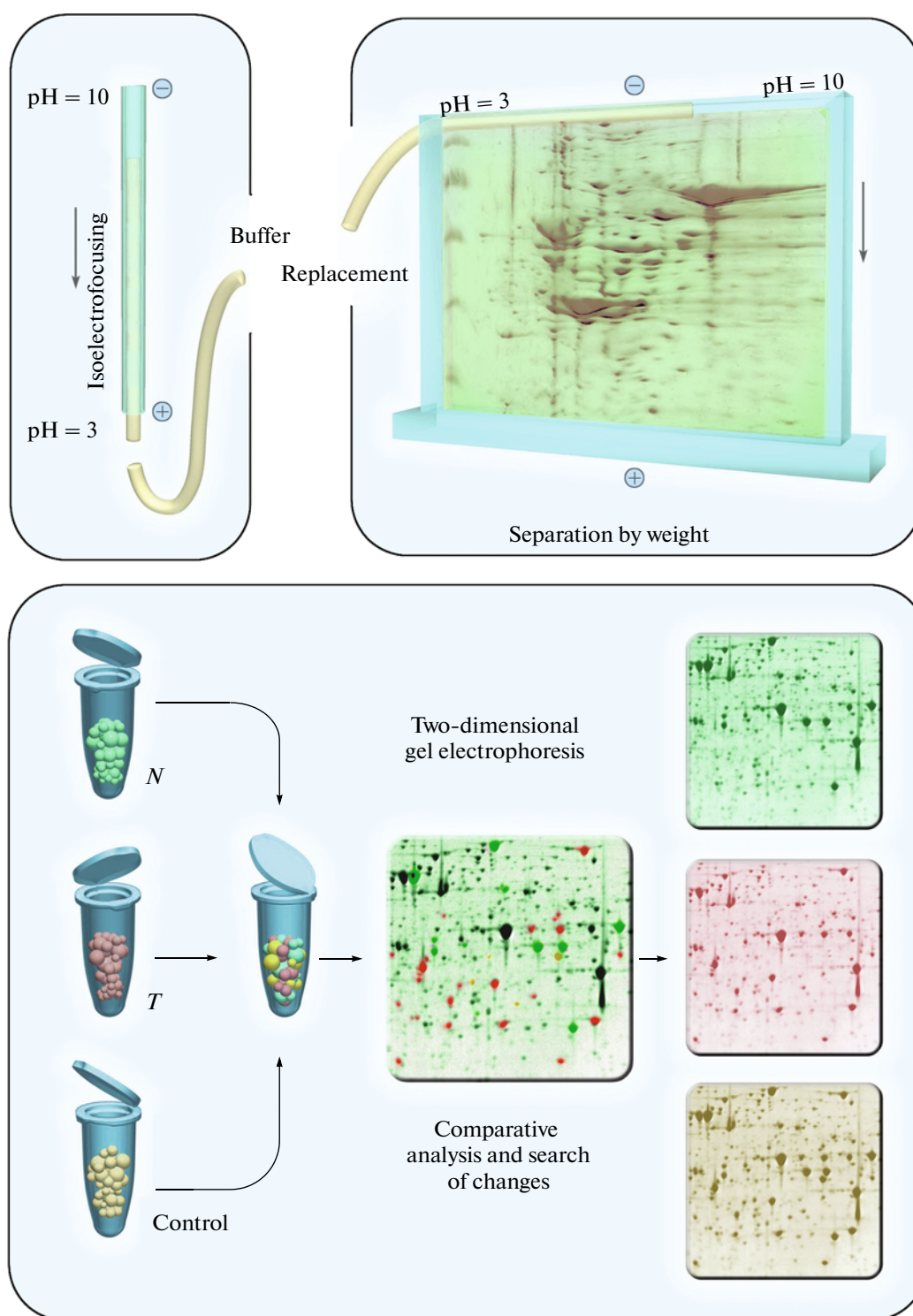


Fig. 1. Scheme of proteomics in vitro methods. Top: 2DE. Bottom: 2D DIGE. Color patches do not reflect the wavelength of the emitted radiation and used to demonstrate the principle.

PROTEOMICS in silico METHODS

Bioinformatics Search in Oncomine Database

Oncomine database contains information on gene expression profiles obtained by hybridization of total

mRNA of normal and tumor tissue with probes of commercially distributed microarrays. Each microchip contains short sequences (probes) that are complementary to most human mRNA [7, 39]. Using an appropriate web service [8], it is possible to estimate

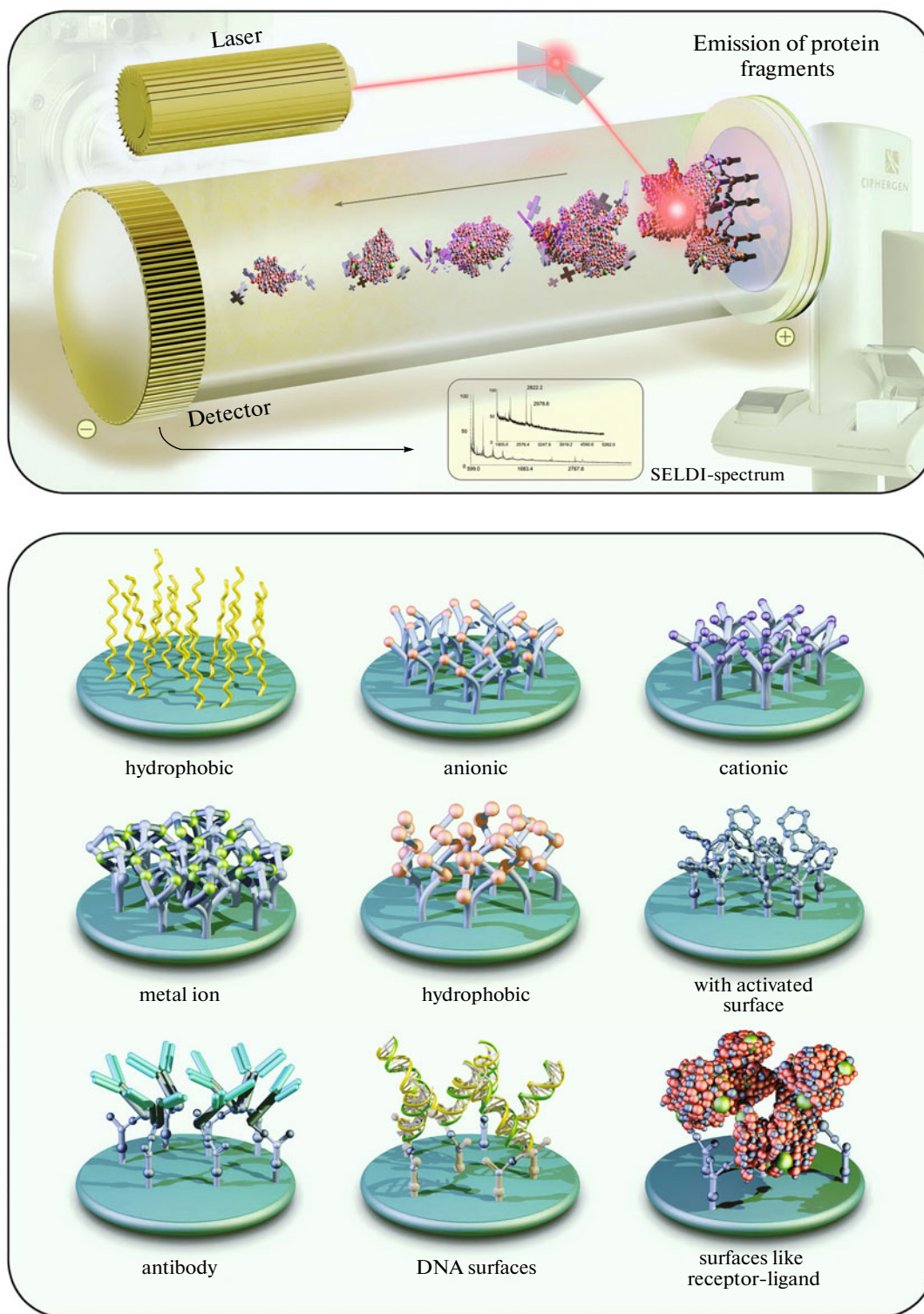


Fig. 2. Scheme of SELDI method. Top: SELDI mass spectrometer, where the argon laser beam is refracted by a lens and directed to be bound with protein-activated surfaces (bottom).

the level of expression of almost any gene in hundreds of tissues and to quickly detect genes, the expression of which differs statistically significantly in certain types

of normal and tumor tissues [40]. The similar information results in the Human Proteome Atlas database [41], which contains the results of a detailed immuno-

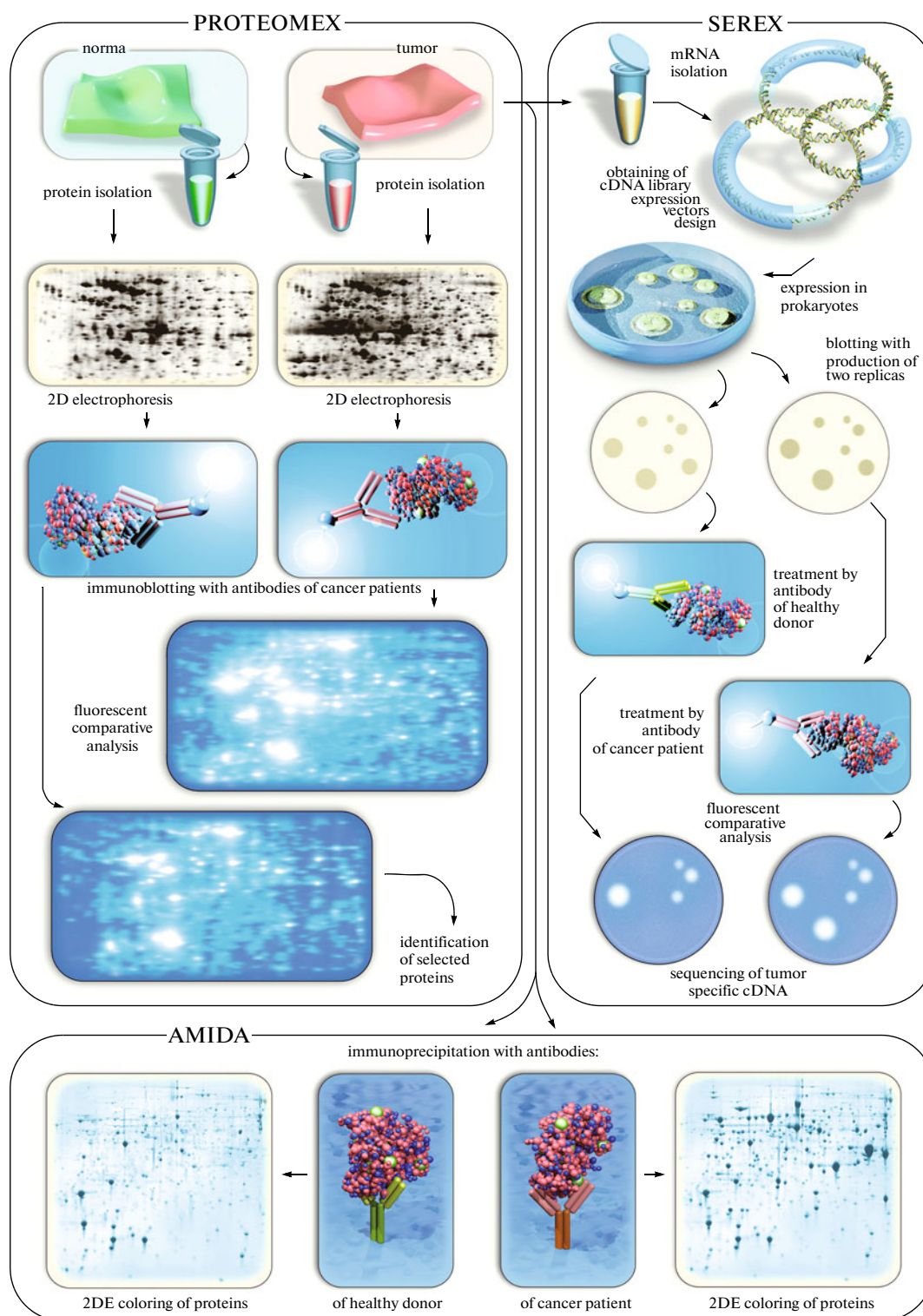


Fig. 3 Scheme of reverse proteomics (SEREX, ProteomeX, and AMIDA).

histological analysis of 60 pairs of normal and tumor samples.

Upon relative cheapness (300–1000 USD per experiment), the method lacks a number of things.

First, during the analysis, there is a cross hybridization of homologous members of the related mRNA family. Furthermore, due to the low concentration of mRNA in the sample, the complexity of the test (great total

length of transcriptomes sequences and significant variations in mRNA concentration) often leads to the absence of the detected signal from a rather large part of mRNA. The analysis is complicated by the significant discrepancy between the results obtained using microarrays from different companies, and different methods of a statistical analysis of the results. Finally, the method cannot detect most of the structural changes in the mRNA (mutations, products of alternative splicing, etc.). Currently, the use of microchips can only analyze mRNA and microRNA, whereas human transcriptome apparently contains thousands of noncoding RNAs (ncRNAs), and a large number of antisense RNAs. Thus, the design of microchips must improve continuously.

Bioinformatics Search of EST Databases

Expressed sequence tags (EST technique) include the sequencing of 200–800-bp cDNA nucleotide fragments [42, 43] (Fig. 4). The method is based on the assumption that the number of sequenced EST clones reflects the content of each mRNA in the sample. In the 1990s, a number of EST libraries of normal and tumor tissues were obtained that contained over 100 000 clones; it allows one to determine the level of synthesis of the most high- and middle-copy-number mRNA. A dbEST database was constructed based on the results of long-term studies. To conduct a bioinformatics analysis of this database, a wide range of approaches and tools was proposed. However, after the creation of a high-performance sequence of transcriptomes (see below), this database will be used less and less.

Bioinformatics Search of SAGE Databases

The serial analysis of gene expression (SAGE) yields short (10 bp) fragments of each cDNA sample (nucleotide tags), the ligation of tags in one concatemer sequence, and the amplification and sequencing of the concatemer [44]. To identify the gene, the sequence tag is mapped to a genomic nucleotide sequence (Fig. 4). The SAGE sensitivity depends on the depth of tag sequencing and the representativeness of the library [45, 46]; furthermore, the evaluation of the level of expression of transcripts is based on the frequency of the tag. The method provides an adequate assessment of the absolute mRNA content, which allows one to directly compare the quantity of various mRNAs in the same sample or in different samples. A comparison of the efficiency of the search for tumor markers using the SAGE and 2DE methods has shown that the method of the serial analysis of gene expression exceeds 2DE by more than three times with regard to informativeness [47]. This is related to two problems, i.e., (1) the presence of tags for which there is no appropriating site in the genome (in this case, mistakes occur in sequencing or amplification) and (2)

the existence of several sites in the genome with sequences identical to the tag. Both problems are eliminated by extending the tag in LongSage [48] and Tag-seq [49, 50] modifications. Currently, the production and sequencing of the Tag-seq libraries for samples is conducted on a commercial basis (mRNA tag profiling service). The main advantage, as well as the shortcoming, is that, as in the hybridization on microchips, this approach is only part of the transcriptome analysis; on one hand, it makes the analysis cheaper and, on the other, less informative.

METHODS OF TRANSCRIPTOMICS

Transcriptome analysis methods were originally aimed at identifying the mRNA, the level of synthesis of which dramatically changes during carcinogenesis. This approach can detect mRNA prognostic markers used for the PCR analysis of biopsy materials in which mRNAs are fairly stable. However, it is not possible to use mRNA markers for diagnoses because they quickly degrade in the serum. A promising alternative approach to searching the transcriptome for diagnostic markers is the analysis of noncoding RNA (ncRNA), which is divided into two groups of short (less than 200 bp) and long (more than 200 bp). There are several thousand ncRNAs encoded in the human genome. The most numerous family of ncRNA is microRNA (more than 800 types have currently been identified), the main function of which is to regulate the expression of nearly all genes that code human proteins. Regulation is carried out at the post-transcriptional level as a result of the specific hybridization of miRNA to its target mRNA [51–53]. Recently, it has been shown that miRNA is secreted into the bloodstream in the exosome by tumors and, after the degradation of exosomes, accumulate in the blood in the form of highly stable complexes with Argonaut protein family [54]. Profiling the tissue-specific microRNA content in blood serum allows one to precisely establish the localization of tumors (often with the definition of its subclass) and predict its further progression [55].

Unlike the genes that encode short ncRNA, long ncRNA genes occupy the largest part of the genome, including areas previously called “junk DNA.” The total length and, therefore, the total number of these genes is many times greater than the length of the genes that encode the proteins (more than 90% and 2% of the genome, respectively) [56]. In view of the uncertainty of the position of the 5'- and 3'-ends of mRNA, as well as their processing, alternative splicing, and editing, and the synthesis of antisemantic sequences, the number of individual RNAs in the transcriptome can reach one million or more. The greater part of this RNA does not contain poly(A)-sequences and, apparently, is transcribed with the partici-

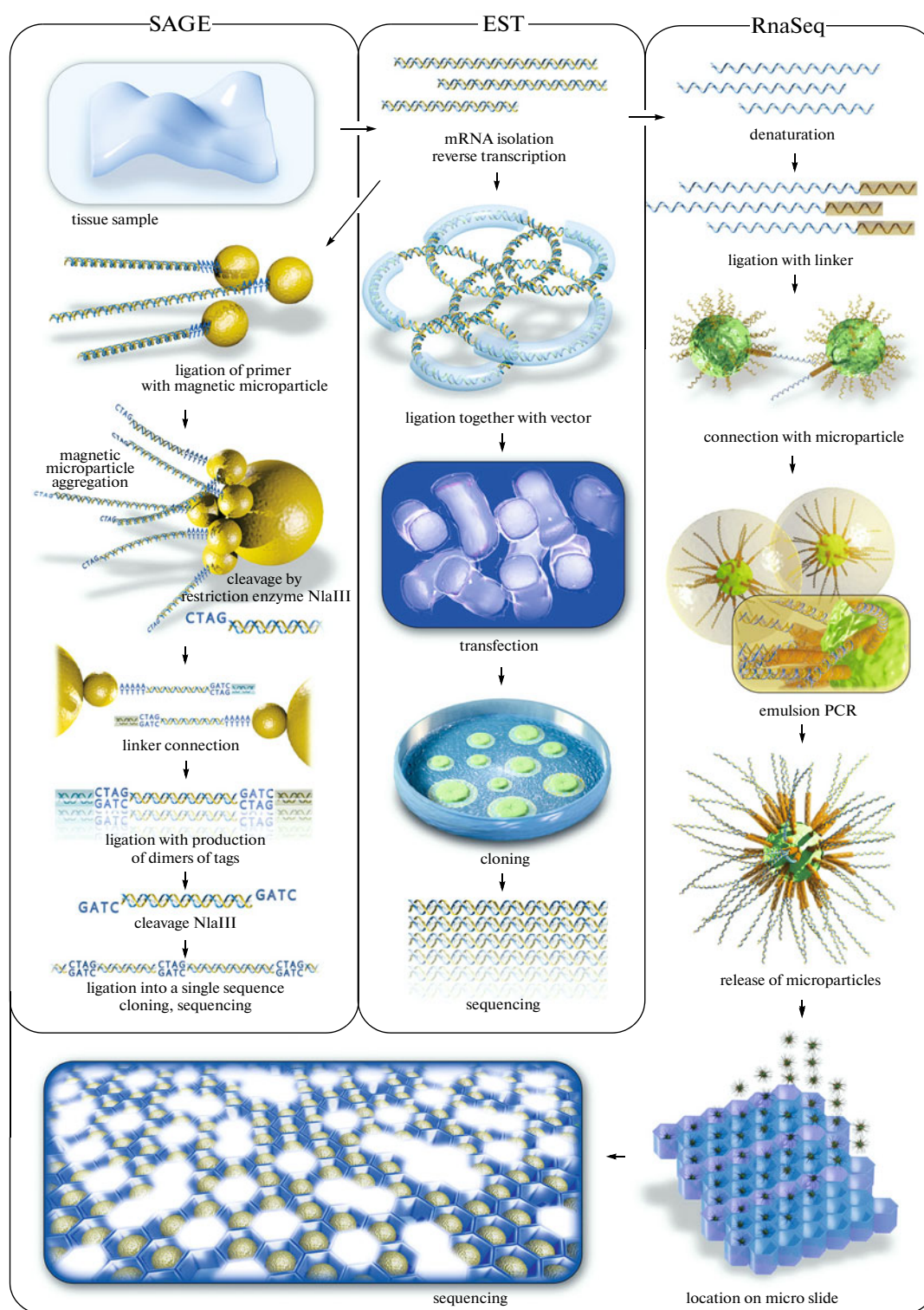


Fig. 4. Schemes of proteomics in silico methods based on comparison of normal and tumor tissues transcriptomes (SAGE, EST, and RnaSeq).

pation of RNA-polymerase III. Preliminary data indicate that the further analysis of this terra incognita can lead to revolutionary changes in molecular biology [57]. While some long ncRNA is fairly stable in the blood serum and

can be used in serological tumor diagnostics, some are secreted in blood flow in stable protein complexes [58]. Thus, the number of tumor markers used in clinical practice can be greatly increased in the future.

Table 1. Methods of searching for tumor markers

Approach	Method	Analysis of protein or cDNA	Marker identification	Reference
Proteomics in vitro	2DE, 2D DIGE	2DE electrophoresis	Mass spectrometry	[1]
	SELDI	Adsorption to surfaces of microchips	Mass spectrometry	[2]
Reverse proteomics	SEREX	Analysis of two copies of expression cDNA libraries of tumor using autoantibodies	Sequencing	[4]
	Proteomex	Immunoblotting of 2DE cancer electrophoregram of normal and tumor tissues using autoantibodies	Mass spectrometry	[5]
	AMIDA	2DE immunoprecipitates of normal and tumor tissues obtained using autoantibodies of cancer patient	Mass spectrometry	[6]
Proteomics in silico	Oncomine	Analysis of cDNA hybridization databases with probes of microchip	On microchip probe	[7]
	EST	Analysis of sequencing cDNA fragment databases	On sequence	[8]
	SAGE	Analysis of databases of cDNA expression	On tag	[8]
Transcriptomics	Analysis of microRNA	Analysis of microRNA expression profiles of tumor exosomes	On miRNA content	[9]
	WTSS	Analysis of high-performance database transcriptome sequencing	On ncRNA content	[10]
Genomics	GM	High-performance sequencing of exome of normal and tumor tissues	PCR	[11]
	EGM	High-performance sequencing libraries of DNA fragments of normal and tumor tissue after treatment with bisulfite	PCR	[12]

Profiling of MicroRNA Expression

Due to the tissue specificity and mass secretion of many miRNA in the exosomes by tumors in blood, the recently developed method of profiling miRNA expression is most promising for the sensitivity and accuracy of cancer diagnoses. This method is not yet used in clinical practice. In this case, the application of hybridization on microchips for detecting a marker is not so effective (the length of microRNA does not exceed 23 bp, and some miRNA differ by only one base), while the RT-PCR elongation of the amplified fragment during the reaction (mirQ method) provides highly reproducible results [51].

Whole Transcriptome Shotgun ncRNA Sequencing

The method of whole transcriptome shotgun sequencing (WTSS) basically allows one to determine the nucleotide sequences of mRNA and ncRNA of tissue transcriptome using commercially available kits for preparing samples and poly (A)-neutral amplification (Fig. 4) [59–61]. Illumina sequencers, which are the most commonly used for WTSS, have a capacity of up to 300 billion nucleotides, the length of the “read,” i.e., the fragments read, is ~50–100 nt, and the sequencing costs is 1 USD per 1 million nucleotides (\$12000 per run) [62]. The subsequent computer analysis provides a way to define and connect

sequences of all transcribed exons to identify all RNA isoforms, i.e., to reveal the distinctions in the position of the 5'-and 3'-ends of all gene transcripts, as well as to detect previously unidentified RNAs [55]. Over time, this approach can lead to the complete reconstruction of all tissue and cell transcriptomes (using microdissection technology) that would require the depth of sequencing in some terabases to ensure adequate coverage for the quantitative analysis of all ncRNA.

GENOMIC METHODS

There are two major groups of genomic tumor markers [63], i.e., (1) genetic markers, i.e., DNA fragments of tumor cells, at the point where mutations were found, as well as deletions, amplifications, or translocations that arise at different stages of carcinogenesis, and (2) epigenetic marks, including areas of the promoter regions, the hypermethylation of which in tumor cells leads to the inhibition of gene transcription (silencing). Another rarely used epigenetic approach is based on an analysis of the global hypomethylation of DNA of tumor cells.

Search for Genetic Markers

The largest number of genetic markers was found before the genomic era in the analysis of retroviruses in

Table 2. Analysis methods of searching for tumor markers on PubMed database*

Method	First reference (year)	Number of publications			Trend***
		total number	per year	2010–2011**	
2DE	1972	1680	43	114	2.7
2D DIGE	2002	180	20	29	1.5
SELDI	2001	376	37	20	0.5
SEREX	1998	53	4	3	0.8
Proteomex	2003	4	0	0	0.0
AMIDA	2004	2	0	0	0.0
Oncomine	2007	16	4	4	1.0
EST	1996	38	2	1	0.5
SAGE	1998	134	10	4	0.4
miRNA	2002	838	93	275	2.9
WTSS	2005	103	17	22	1.3
GM	1960	24097	473	1450	3.1
EGM	1978	1354	41	220	5.4

* Search was performed using the following algorithm (the method name) AND (cancer markers). ** The number of publications per year for 2010–2011 (current efficiency). *** The ratio of the current productivity of the number of publications per year over the lifetime of the method (trend >1 indicates an increase in popularity, <1 indicates a decrease in popularity).

mice, cotransfection experiments, and positional cloning. Currently, these markers are identified using high-sequencing methods of exome (coding part of the genome or parts of it) of normal and tumor cells of cancer patients. The analysis of genetic markers is a basis for consulting families with hereditary predisposition to the occurrence of cancer (these cases average about 20–30% of the general number of patients) [11]. Finally, genetic markers are used for predictions, but not for serological diagnostics, since for this purpose, it is necessary to clear tumor cells circulating in blood, which is only possible at the late stages of the disease or in postoperative monitoring.

Search for Epigenetic Markers

The CpG-island hypermethylation of promoter regions of genes in tumor cells is a much more common event than DNA mutation. The library of epigenetic markers are obtained by treating DNA with restriction enzyme, introducing adapters to the ends of all the fragments, DNA bisulfite treatment (which allows one to split most of the fragments of islands with unmethylated cytosine residues), and genome amplification of fragments that remain intact. Libraries of normal and tumor tissue obtained from the same cancer patients are subjected to high-performance sequencing with subsequent identification of tumor-specific epigenetic markers by computer analysis [12]. The high stability of DNA can detect tumor-specific epigenetic markers in biological fluids and stool using PCR. The specificity of the assay significantly increases upon the amplification of hypermethylated

fragments longer than 200 bp, since it allows one to eliminate the DNA of normal cells that enter into the bloodstream because of apoptosis [64]. However, despite the high specificity and sensitivity of analysis of epigenetic markers (especially when their panels is using), in the clinical diagnosis of tumors, they are seldom used, due to the identification of common, but nondangerous pre-neoplastic tissues and adenomas, as well as the age-related changes of the genome [65, 66].

CONCLUSIONS

An analysis of the frequency of the use of methods described in the review for identifying tumor markers using the PubMed database showed that genomic techniques are the most popular, followed by methods of analyzing miRNA and the two-dimensional electrophoresis of proteins (Table 2). The popularity and effectiveness of the search for transcriptomic tumor marker by high-performance sequencing methods is growing rapidly. The possibility of the complete reconstruction of the cell and tissue transcriptome in the future could lead to the identification of new diagnostics of noncoding RNA that is stable in sera in protein complexes and suitable for tumors.

It is important to note that, despite the long history of genomic methods of the search for diagnostic markers of tumors, their use in clinical practice is limited. Thus, the review of currently in use or undergoing clinical and preclinical testing of diagnostic markers of colon cancer shows that most of them (70%) are proteins, and 30% are DNA fragments [67]. This is probably due to the fact that, despite the lower sensitivity of

immunochemical methods of analyzing protein markers compared to the results of DNA amplification in protein tests provide an opportunity for a clear diagnosis, whereas DNA testing requires the use of large panels of markers and the laborious sequencing of DNA products. At the same time, one must radically increase the sensitivity of modern methods of protein detection. This will make it possible to analyze low-copy-number proteins and overcome an enormous dynamic range of concentrations when testing biological samples (up to a trillion times).

ACKNOWLEDGMENTS

This work was supported by the State Contracts Nos. 14.740.11.0757 and 16.552.11.7034 of the Russian Ministry of Education and Science and the Russian Foundation for Basic Research (11-04-01023).

REFERENCES

- Wong S.C., Chan C.M., Ma B.B., et al. 2009. Advanced proteomic technologies for cancer biomarker discovery. *Exp. Rev. Proteomics*. **6**, 123–134.
- Garrisi V.M., Abbate I., Quaranta M., et al. 2008. SELDI-TOF serum proteomics and breast cancer: Which perspective? *Expert Rev. Proteomics*. **5**, 779–785.
- Balestrieri M.L., Giovane A., Mancini F.P., et al. 2008. Proteomics and cardiovascular disease: An up-date. *Curr. Med. Chem.* **15**, 555–572.
- Song M.H., Ha J.C., Lee S.M., et al. 2011. Identification of BCP-20 (FBXO39) as a cancer/testis antigen from colon cancer patients by SEREX. *Biochem. Biophys. Res. Commun.* **408**, 195–201.
- Seliger B., Kellner R. 2002. Design of proteome-based studies in combination with serology for the identification of biomarkers and novel targets. *Proteomics*. **2**, 1641–1651.
- Rauch J., Gires O. 2008. SEREX, Proteomex, AMIDA, and beyond: Serological screening technologies for target identification. *Proteomics Clin. Appl.* **2**, 355–371.
- <https://www.oncomine.org/resource/login.html>
- <http://www.genecards.org/>
- <http://www.hptaa.org/miRNA/index.htm>
- <http://rna-seqblog.com/>
- Lichtenstein P., Holm N.V., Verkasalo P.K., et al. 2000. Environmental and heritable factors in the causation of cancer: Analyses of cohorts of twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **343**, 78–85.
- Mill J., Petronis A. 2009. Profiling DNA methylation from small amounts of genomic DNA starting material: Efficient sodium bisulfite conversion and subsequent whole-genome amplification. *Methods Mol. Biol.* **507**, 371–381.
- Koenig R., Stegemann H., Francksen H., et al. 1970. Protein subunits in the potato virus X group: Determination of the molecular weights by polyacrylamide electrophoresis. *Biochim. Biophys. Acta.* **207**, 184–189.
- O'Farrell P.H. 1975. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250**, 4007–4021.
- Issaq H.J., Veenstra T.D. 2008. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): Advances and perspectives. *BioTechniques*. **44**, 697–700.
- Collet B., Guitton N., Saïkali S., et al. 2011. Differential analysis of glioblastoma multiforme proteome by a 2D-DIGE approach. *Proteome Sci.* **9**, 16.
- Hutchens T.W., Yip T. 1993. New desorption strategies for the mass spectrometric analysis of macromolecules. *Rapid Commun. Mass Spectrom.* **7**, 576–580.
- Kusmann M., Roepstorff P. 2000. Sample preparation techniques for peptides and proteins analyzed by MALDI-MS. *Methods Mol. Biol.* **146**, 405–424.
- Caputo E., Moharram R., Martin B.M. 2003. Methods for on-chip protein analysis. *Anal. Biochem.* **321**, 116–124.
- Issaq H.J., Veenstra T.D., Conrads T.P., et al. 2002. The SELDI-TOF MS approach to proteomics: Protein profiling and biomarker identification. *Biochem. Biophys. Res. Commun.* **292**, 587–592.
- Bischoff R., Luider T.M. 2004. Methodological advances in the discovery of protein and peptide disease markers. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **803**, 27–40.
- Liu J.Y., Jin L., Zhao M.Y., et al. 2011. New serum biomarkers for detection of tuberculosis using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry. *Clin. Chem. Lab. Med.* **49**, 1727–1733.
- <http://www.licr.org/SEREX.html>
- Liggins A.P., Guinn B.A., Banham A.H. 2005. Identification of lymphoma-associated antigens using SEREX. *Methods Mol. Med.* **115**, 109–128.
- Klade C.S., Voss T., Krystek E., et al. 2001. Identification of tumor antigens in renal cell carcinoma by serological proteome analysis. *Proteomics*. **1**, 890–898.
- Klade C.S. 2002. Proteomics approaches towards antigen discovery and vaccine development. *Curr. Opin. Mol. Ther.* **4**, 216–223.
- Le Naour F., Hohenkirk L., Grolleau A., et al. 2001. Profiling changes in gene expression during differentiation and maturation of monocyte-derived dendritic cells using both oligonucleotide microarrays and proteomics. *J. Biol. Chem.* **276**, 17920–17931.
- Lichtenfels R., Kellner R., Bukur J., et al. 2002. Heat shock protein expression and anti-heat shock protein reactivity in renal cell carcinoma. *Proteomics*. **2**, 561–570.
- Mann M., Jensen O.N. 2003. Proteomic analysis of post-translational modifications. *Nature Biotechnol.* **21**, 255–261.
- Gallego M., Virshup D.M. 2007. Post-translational modifications regulate the ticking of the circadian clock. *Nature Rev. Mol. Cell. Biol.* **8**, 139–148.
- Kim K.I., Baek S.H. 2006. SUMOylation code in cancer development and metastasis. *Mol. Cells*. **22**, 247–253.
- Krueger K.E., Srivastava S. 2006. Posttranslational protein modifications: Current implications for cancer

- detection, prevention, and therapeutics. *Mol. Cell. Proteomics*. **5**, 1799–1810.
33. Gires O., Munz M., Schaffrik M., et al. 2004. Profile identification of disease-associated humoral antigens using AMIDA, a novel proteomics-based technology. *Cell. Mol. Life Sci.* **61**, 1198–1207.
 34. Rauch J., Ahlemann M., Schaffrik M., et al. 2004. Allogenic antibody-mediated identification of head and neck cancer antigens. *Biochem. Biophys. Res. Commun.* **323**, 156–162.
 35. Scanlan M.J., Gout I., Gordon C.M., et al. 2001. Humoral immunity to human breast cancer: Antigen definition and quantitative analysis of mRNA expression. *Cancer Immun.* **1**, 4.
 36. Scanlan M.J., Welt S., Gordon C.M., et al. 2002. Cancer-related serological recognition of human colon cancer: Identification of potential diagnostic and immunotherapeutic targets. *Cancer Res.* **62**, 4041–4047.
 37. Scanlan M.J., Simpson A.J., Old L.J. 2004 The cancer/testis genes: Review, standardization, and commentary. *Cancer Immun.* **4**, 1.
 38. <http://www2.licr.org/CancerImmunimeDB>
 39. Rhodes D.R., Kalyana-Sundaram S., Mahavisno V., et al. 2007. Oncomine 3.0: Genes, pathways, and networks in a collection of 18000 cancer gene expression profiles. *Neoplasia*. **9**, 166–180.
 40. Yang Y., Pospisil P., Iyer L.K., et al. 2008. Integrative genomic data mining for discovery of potential blood-borne biomarkers for early diagnosis of cancer. *PLoS ONE*. **3**, e3661.
 41. <http://www.proteinatlas.org/>
 42. Adams M.D., Kelley J.M., Gocayne J.D., et al. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science*. **252**, 1651–1656.
 43. Wilcox A.S., Khan A.S., Hopkins J.A., et al. 1991. Use of 3' untranslated sequences of human cDNAs for rapid chromosome assignment and conversion to STSs: Implications for an expression map of the genome. *Nucleic Acids Res.* **19**, 1837–1843.
 44. Velculescu V.E., Zhang L., Vogelstein B., et al. 1995. Serial analysis of gene expression. *Science*. **270**, 484–487.
 45. Yamamoto M., Wakatsuki T., Hada A., et al. 2001. Use of serial analysis of gene expression (SAGE) technology. *J. Immunol. Methods*. **250**, 45–66.
 46. Ruijter J.M., Van Kampen A.H., Baas F. 2002. Statistical evaluation of SAGE libraries: Consequences for experimental design. *Physiol. Genomics*. **11**, 37–44.
 47. Bukurova Yu.A., Khankin S.L., Krasnov G.S., et al. 2010. Estimation of the efficiency of 2D analysis and bioinformatics search in identification of protein markers for colon tumors. *Mol. Biol. (Moscow)*. **44**, 334–340.
 48. Saha S., Sparks A.B., Rago C., et al. 2002. Using the transcriptome to annotate the genome. *Nature Biotechnol.* **20**, 508–512.
 49. Morrissy A.S., Morin R.D., Delaney A., et al. 2009. Next-generation tag sequencing for cancer gene expression profiling. *Genome Res.* **19**, 1825–1835.
 50. Morrissy S., Zhao Y., Delaney A., et al. 2010. Digital gene expression by tag sequencing on the illumina genome analyzer. *Curr. Protoc. Hum. Genet.* Chapter 11: Unit 11.11.1–36.
 51. Lu J., Getz G., Miska E.A., et al. 2005. MicroRNA expression profiles classify human cancers. *Nature*. **435**, 834–838.
 52. Naqvi A.R., Islam M.N., Choudhury N.R., et al. 2009. The fascinating world of RNA interference. *Int. J. Biol. Sci.* **5**, 97–117.
 53. Lee S.K., Calin G.A. 2011. Non-coding RNAs and cancer: New paradigms in oncology. *Discov. Med.* **11**, 245–254.
 54. Patel D.J., Ma J.B., Yuan Y.R., et al. 2006. Structural biology of RNA silencing and its functional implications. *Cold Spring Harbor Symp. Quant. Biol.* **71**, 81–93.
 55. Schetter A.J., Leung S.Y., Sohn J.J., et al. 2008. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *J. Am. Med. Assoc.* **299**, 425–436.
 56. Gibb E.A., Brown C.J., Lam W.L. 2011. The functional role of long non-coding RNA in human carcinomas. *Mol. Cancer*. **10**, 38.
 57. Guttman M., Garber M., Levin J.Z., et al. 2010. Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature Biotechnol.* **28**, 503–510.
 58. Panzitt K., Tschernatsch M.M., Guelly C., et al. 2007. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. *Gastroenterology*. **132**, 330–342.
 59. Castle J.C., Armour C.D., Löwer M., et al. 2010. Digital genome-wide ncRNA expression, including SnoRNAs, across 11 human tissues using polyA-neutral amplification. *PLoS ONE*. **5**, e11779.
 60. Tariq M.A., Kim H.J., Jejelowo O., et al. 2011. Whole-transcriptome RNAseq analysis from minute amount of total RNA. *Nucleic Acids Res.* **39**, e120.
 61. <http://www.454.com/applications/transcriptome-sequencing.asp>
 62. Wang Z., Gerstein M., Snyder M. 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nature Rev. Genet.* **10**, 57–63.
 63. Taby R., Issa J.P. 2010. Cancer epigenetics. *CA Cancer J. Clin.* **60**, 376–392.
 64. Kalimutho M., Del Vecchio Blanco G., Cretella M., et al. 2011. A simplified, non-invasive fecal-based DNA integrity assay and iFOBT for colorectal cancer detection. *Int. J. Colorect. Dis.* **26**, 583–592.
 65. Widschwendter A., Müller H.M., Fiegl H., et al. 2004. DNA methylation in serum and tumors of cervical cancer patients. *Clin. Cancer Res.* **10**, 565–571.
 66. Wong T.S., Kwong D.L., Sham J.S., et al. 2004. Quantitative plasma hypermethylated DNA markers of undifferentiated nasopharyngeal carcinoma. *Clin. Cancer Res.* **10**, 2401–2406.
 67. Tanaka T., Tanaka M., Tanaka T., et al. 2010. Biomarkers for colorectal cancer. *Int. J. Mol. Sci.* **11**, 3209–3225.