PERIODICA POLYTECHNICA SER. EL. ENG. VOL. 50, NO. 1–2, PP. 91–102 (2006)

DNA-CHIPS IN PRACTICE AND THEIR CLINICAL RELEVANCE

Tamás MICSIK

Department of Control Engineering and Information Technology
Budapest University of Technology and Economics
H–1117 Budapest, Magyar Tudósok krt. 2, Hungary
e-mail: kefir@seeger.iit.bme.hu

Received: April 28, 2005

Abstract

The DNA-chips are almost a decade old and by now have brought an explosion in genetics. With their ability to investigate the biological systems in their complexity microarray experiments resulted in enormous amount of complex data. It will be a challenge for biologists, statisticians and for information technology engineers to develop capable methods for coping with the analysis, validation and interpretation. This revolutionary new field of biology demands absolutely new algorithms, analytical methods and new ways of thinking combined with effective team-work. No doubt that DNA-chips will shortly be very useful and essential technology in the diagnostic and therapeutic processes. In this review we shortly introduce the most important types of DNA-chips and describe their capability and limitations to be clear how to position them in the future scientific ways of thinking.

Keywords: microarray, diagnostic methods, gene expression, validation, statistical methods, clusters.

1. Introduction

DNA-chips are a revolutionary new method in genetics and brought breakthrough in the diagnostic and therapeutic regimen. It makes possible to perform thousands of experiments parallel which gave hope for gaining some insight into the enormous complexity of the living biological systems and processes [1]. Our life is made of thousands of processes working simultaneously and before DNA-chips the investigators were able to follow only a limited number of various factors. Usually a distinct change of a gene does not make manifesting effect, but joining other concomitant changes it could lead to significant illnesses [2]. More diseases, originating from quite different ways can look very alike, and therefore the adequate therapy can be very different. Colorectal cancers might cause the same symptoms, but studies revealed more different ways and genes of tumor development in bowels [3].

The growing palette of molecular investigating methods can bring more insight and better understanding into the development of special conditions and in turn discovering novel potential therapeutic targets. Investigating of signal transduction processes recently led to a targeted blockage of a specific kinase, which is proved to be very successful in the treatment of chronic myelogenous leukemia and gastrointestinal stromal tumors, and resulted in much better prognosis [4].

2. DNA-Chips

DNA-chips is a phrase appertaining a little device with lots of individual areas covered with millions of uniform DNA-sequences. It works upon the characteristic and inevitable feature of hybridization of nucleic acids. During hybridization two complementary sequences (they match to each other, like the key fits the lock) will find and join each other. Hybridization is a key and very accurate process in biology in duplicating the DNA and in functioning and controlling the genes, and that is why this approach is very suitable for a broad kind of experiments.

The inventor of DNA-chips is the Affymetrix corporation in California [5]. They use photolithography (like in case of integrated circuits, hence comes the name) to produce their patent pended Genechips. For today there are very many kinds and ways of producing DNA-chips which led to many names: DNA-arrays, DNA-matrices, microarrays, macroarrays and also to a confusing variety of devices and literature.

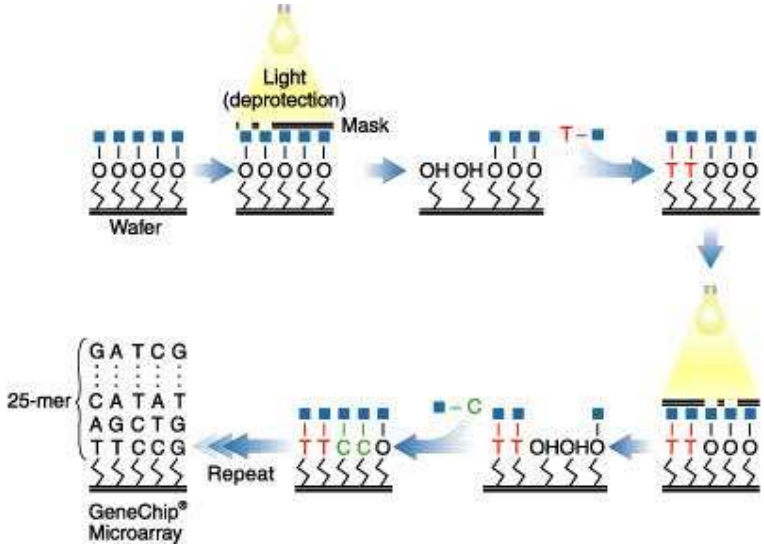

Fig. 1. Affymetrix uses photolithography for manufacturing high density oligonucleotide GeneChips, of which resolution is in nanometer range (www.affymetrix.com)

The basic idea of the DNA-chips: make a very small device of individual affinity sensors and cover it with different sequences. If it is small enough it is capable of putting even hundreds thousands of sensors on a little surface and so it would be possible to perform a huge number of various experiments at the same time. The DNA-chip contains the investigatory sequences (probes) bound to its surface at well-known and circumscribed places. The unknown sample which should be investigated can be hybridized onto it and relevant data should be read indirectly.

The probes on the DNA-chip may be of DNA, RNA, specially modified sequences or even protein and antibodies and depending on the purpose of the DNA-chip. In vast majority the chips are made of DNA-sequences. The quality and quantity of hybridization depends on environmental conditions and factors (pH, temperature, pK, ion-strength etc.) but the most important thing is the complementarity of the two sequences. The stable double helix of DNA is made up of two antiparallel strands of desoxyribonucleic acids and this structure makes them very appropriate for their function of containing all information of life in a secure and easy readable way. Hybridization happens according to thermodynamic rules and has optimal length and base-content. Short sequences are bound to each other less strongly than longer ones, while too long sequences are too robust and can form the stable double helix inside their own single stranded chain. Base-sequence is crucial in determining the number of possible bounds and the strength and optimal temperature of hybridization. All these issues should be cared very cautiously in assembling a DNA-chip.

The investigated sample should be prepared, amplified and labelled in a way to be able to hybridize and afterwards detected on the chip. Among many possibilities fluorescent labelling is the most wide-spread: dyes bounded covalently to the sample can be excited and detected with laser-scanners. In the enzyme-linked method the production of a special substrate (chemiluminescent) is used without laser. In the knowledge of the sharing out of the chip, after hybridization and detection we can determine the quality and quantity of the hybridized investigated sequences indirectly.

According to the surface density of a DNA-chip we used to talk about macroarrays and microarrays. *Macroarrays* are less surface density chips for investigating around several dozens or hundreds of sequences and offer a simpler structure and instrumentation and therefore an easier way to perform investigations. The generating fewer data is usually still enough for investigating a certain condition, disease or question so macroarrays might therefore be more widely spread. The high density DNA-chips, the *microarrays* might contain probes for determining the expression profile of all human genes (round 30000) simultaneously. These microarrays are really sophisticated devices and need high accuracy, expensive and very reliable instrumentation and a specialized team of experts to cope with the enormous raw data generated. Therefore the high-density microarrays can be exploited only in limited number of laboratories.

DNA-chips can be also sorted according to the manufacturing type of sequences. Affymetrix and a couple of another manufacturer build up the nucleic acid sequences right on the surface of the chips via the so-called *in situ synthesis*. This technology yields high accuracy, quality control and enormous surface density but is limited in size and sequence length due to the high instrumentation needs. Cheaper, simpler and more widely used method is the *postsynthetic immobilization* of ready nucleic acid sequences. Many research laboratories have built own plotter machines to put their in-house-made probe sequences on the surface of a glass-slide and then use UV-crosslinking technique to immobilize them on the surface of the chips. This method is simpler to perform and more variable and adjustable to the

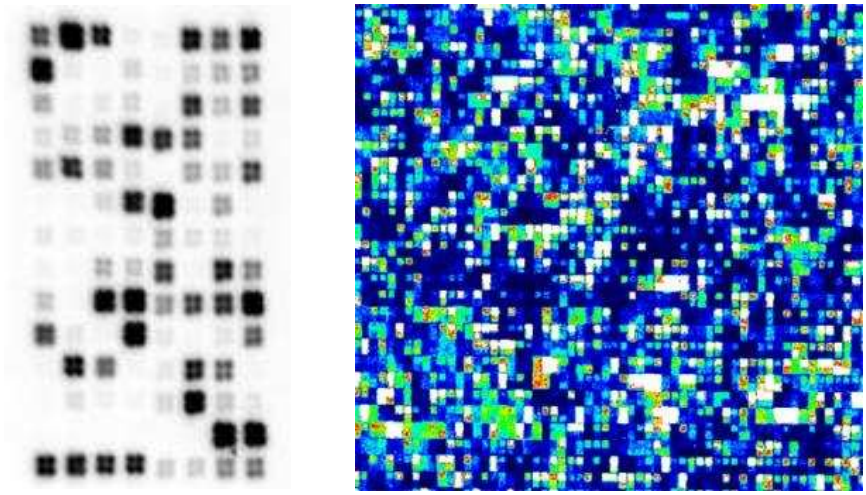

Fig. 2. Comparison of a macroarray (left, 112 quadraspots) with a microarray (right, about 3600 spots). (www.superarray.com, www.affymetrix.com).

special needs of the laboratory and has its benefits for being so widely used, but has its crude limitations in quality control and reproducibility. A lots of special, home-made and company-manufactured chips are in use and the most striking experience with them is the lack of reproducibility and coincidence of results coming from the various laboratories. This is the main drawback of further utilization of DNA-chips.

DNA-chips are able to perform thousands of processes parallel and automated, but can not think instead of us - just like computers. Hybridization is just part of investigation and for best utilization it is inevitable to ask good questions and think over the investigations, layout of the chip carefully to bring real new results into science. Listing the findings is useless: the collected raw data should be analysed and interpreted in useful and cautious ways to make sense of the sequence [6].

2.1. Diagnostic DNA-chips

A certain condition which can be caused by several mutations or alterations in the DNA sequence can be pointed out with a series of experiments, although this can be cost, time and labour-consuming. If we have already learnt all the numerous possible causative changes for a condition and want to find quickly the actual one then it is high time for using a diagnostic DNA-chips. We use all the possible investigating probes and perform all experiments simultaneously in several hours instead of weeks. The otherwise long and labour-intensive work can be cut out and better and adequate therapy might be started much sooner, which in turn can lead to

better efficacy of treatments with less costs and side-effects. In these cases we are looking for quality information, for a presence or absence of a special alteration. It is most suitable, when there are a lot of possible causes, but we can combine tests, too. First applications are with HIV, cystic fibrosis, Ct-P450 enzymes, BRCA-mutations etc.[7] Diagnostic DNA-chips can be used also for research tools, like metilation detection.

Lots of infectious diseases can look very similar, and better treatment can be achieved by diagnostic applications of DNA-chips[8]. Furthermore, if we would like to be very cautious, we might test the pathogen and the host body of the infected people to find out which medicamentation would be most useful and effective with the least side-effects[9]. Huge expenses in health care go for curing harmful side-effects, which could be avoided with better diagnostic tools and individualized therapy. This means that curing a patient would not only mean the following of an accepted and quite effective, but still empiric algorithm, but making special investigations to find the best possible and effective therapy from the lots of possible ones. DNA-chips are very advantageous in speeding up this development [10].

2.2. Gene-expression Monitoring

In biological systems DNA is the source of all information, which is transcribed to messengerRNA to travel out from the nucleus and being translated into working proteins. Proteins are the very engines of cells and they are catalyzing and controlling all biological processes, so cell functions can be traced by the amounts of present proteins and indirectly mRNA. DNA-chips are highly perfect for determining simultaneously thousands of various mRNA levels in certain space and time[11]. The sensitive gene expression monitoring needs more caution and strict environmental conditions, but brought a very useful and promising way for obtaining insight in disease and cancer development. The inestimable benefit of the application of DNA-chips is the view of the gene-network in its complexity and with the capacity of analysing thousands of regulating factors and proteins simultaneously [12].

Rarely a single mutation can cause severe illnesses, but more often many factors with a slight imbalance lead to the development of a disease. DNA-chips make it possible to investigate all the 30000 genes simultaneously of the human body. This enormous amount of data is really admirable and results in quite different possibilities of investigations. Analysing the network of the working and controlling genes or proteins of the systems can reveal key-processes and factors. This so-called ‘gene-fishing’ in the whole genome is made by statistical and mathematical methods and cooperation between the different workgroups of biologist, statisticians, information technologists. Determining key-processes and proteins could result in gaining better understanding of biological systems and in finding new therapeutic targets [13].

For gene expression monitoring we take the sample of the patient immediately after removal from the living body. After isolating RNA a special labeling

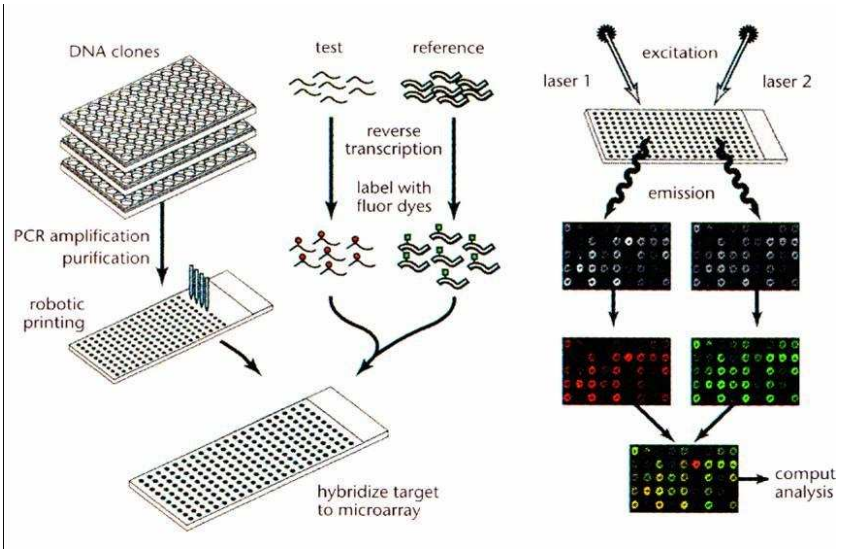

Fig. 3. The process of gene expression monitoring [13].

and amplification method should be used to gain detectable amounts of nucleic acid sequences for hybridization. Gene expression monitoring can be made by cDNA (RNA isolated from cells are reverse transcribed to cDNA) or by directly synthesized oligonucleotides. Both techniques are widely spread but now it seems that oligonucleotides offer a better way because of the less uncertainty in the hybridization process.

The quality of the very sensitive and rapidly degraded mRNA is very crucious because of the need for quantitative data. Furthermore, a biological sample contains several kinds of cells, which could be not all involved in disease development. For example cancerous sample contains comparable amount of stromal and other cells, which have quite other functions and in turn gene expression levels, therefore data could be not relevant for only the cancer cells. In some extent it's not a problem, because in cancer growth other cell types are also involved, and cancer can not grow without the parallel growing of feeding stroma: the growth of stroma is also a possible issue in cancer therapy. With bulk sample we will have enough mRNA to work with, but could be afraid of gaining not true and relevant data about involved genes because of the contaminating cells. Most convincing works and results have been conducted on pure samples containing only one type of cells, but it is more or less accepted, that an involved genes altered expression could be investigated also with bulk samples. The interaction of cancer cells and their environment is also essential in the spreading of cancer and this could also be investigated only with bulk samples. If we cut out only the investigated tumor cells of a bulk sample with laser capture microdissection in a long and tiring process with an expensive machine we will get so few cells and mRNA that should be amplified very strongly

and this in turn can cause false data because of the nonlinearity of the amplification methods. Both methods have their backs and fors and the literature is not sure about the proper way of investigating.

An emerging way of investigating bulk samples without microdissection is the post-processing of the scanned data via mathematical and statistical methods. With the power of computers and access to various databases it is possible to effectively search for new target genes even from bulk tissue samples [14]. If we build up a statistical database of the various gene expression data of pure cells and tissues in advance, we will be able to deal with the gene expression data coming from bulk tissues. This could be simply made by substracting the gene activities coming from the contaminating cells from the bulk results, but with Boolean operators it is possible to search databases for expression patterns of interest in a more sophisticated method. This method is also applicable in tissues, where the pathological pathway, the cause and consequence are not really clear: for example various inflammations, chronic diseases. The generation of the pure gene expression databases of various tissues (muscle-, soft tissue , mucosa-associated genes) could be used in separating of samples into their constituents and analysing simultaneously and also separately each cell type and the various pathological pathways (and their network) to reveal interactions of the various cells, tissues and their roles in disease development and propagation [15]. Another successful approach of searching for new genes in cancer and disease development is the use of neural networks [16].

3. Data Interpretation and Validation

The real brain-work begins just after detecting the hybridized sequences on the chip to find out the meaning of the results. All experiments should be made at least in duplications and should be reproducible. DNA-chip technology is relatively new and can not be used as only diagnostic tool yet, but needs strict validation process, in which the experiment is repeated with a different and independent investigation method and concludes to the same results.

In case of diagnostic DNA-chips it is quite straightforward to perform a PCR or some blotting with the same sample but with gene expression studies the situation is more complicated and validation is a real problem. Gene expression monitoring DNA-chips now are mostly used as research tools for gaining new insights and 'fishing' for new key-factors. This means that these chips are working in a relatively unexplored and uncertain area and therefore relatively few background and reliable information is available. Furthermore, gene expression is valid for a cell/tissue at a moment and at a certain place: it is just a snapshot from the cell. Because the many uncertainty only the significantly altered genes are to deal with and normalization, standardization should be made very carefully. Normalization happens usually against equally expressed, 'housekeeping genes'. Significantly altered expressions should be validated with other investigation methods (blotting, real-time quantitative PCR) and here we are facing a doubtful issue: the data com-

ing from the various microarray experiment are not always prone to be validated by other methods. Often the data are controversary and only a small percent of data coincides with other technics or with the results of other laboratories investigating the same disease [17]. The biggest problem with microarray studies remains reproducibility and validability [18].

Microarray technique is a rather expensive one and therefore many laboratories made up their own custom-made spotted cDNA-microarrays[19], what resulted in a lot of investigatory methods and lots of kinds of used arrays, without the ability to compare results. A cDNA sequence is about 100-1000 nucleic acid long and more instable and therefore huge crowd and conformity changes can occur disturbing the hybridization process[20]. A lot of interesting and important gene-functions could be revealed with cDNA-s, but for the more reliable and comparable data it would be advantageous to use shorter sequences, the 25-70 nucleic acids long oligonucleotides.

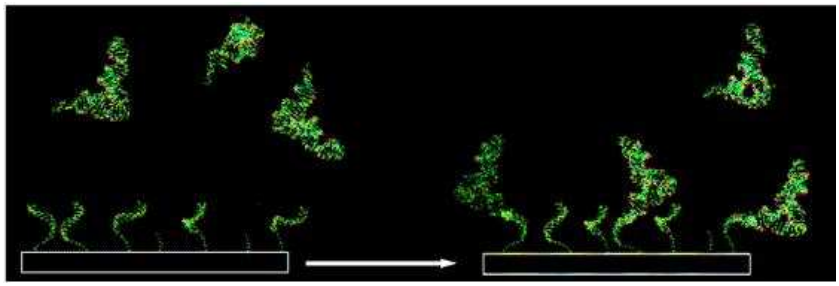

Fig. 4. 3D conformation and molecular interactions influence hybridization with cDNA [20].

Affymetrix uses sets of short oligonucleotides for testing a single gene and usually two dozens of 25-mer oligonucleotides are pointing on different sequences of the same gene. Data analysis is restricted to the genes which concluded to the same intensity at all of the investigating sequences, which is a very effective way of quality assurance. Furthermore Affymetrix continuously builds up a database from all the experiments performed with GeneChips making it possible for every user to compare data and to instantly improve data analysis methods [21]. Quality assurance issues are managed somehow by the other DNA-chip manufacturers as well, but they are rather concentrating on the clear and reliable manufacturing methods.

The other possibility to build up reliable databases and compare various experiments leads through mathematical and statistical methods and these postprocessing algorithms could improve a lot in the reliability of microarray. Gathering all the data and systematically grouping them is a huge work but inevitable in progressing in scientific means [22]. The generated huge amount of data coming from microarray studies would also need new methods and algorithms for viewing and systematizing

them and to cope with the meanings of parallel gained data and ways of interpreting them [23]. There are several companies dealing with these issues and many people working in statistical and computation/information technology field could find new challenges in this data-mining and in silico techniques.

4. Clinical Use

The first *diagnostic DNA-chip* was made in 1996 to HIV-infections to find the most effective therapy in less possible time [24]. For today the list is very long: www.affymtrix.com, www.superarray.com. In these cases DNA-chips accelerates the otherwise long diagnostic process, which is advantageous for sure but the cost for a microarray remained rather high and so the spreading of the technology is still a bit slow, and will speed up when manufacturing costs become less, especially in the fields of infectious diseases, predicting adverse events and in forensic science where rapidity is an important issue and cost is more affordable [25].

Another possibility is the use of macroarrays, which bear less sequences on their surface resulting less data but they are much cheaper and easier to use. With their capacity of hundreds of investigating probes tailored and focused on a special condition or disease they are highly capable of using for clinical and diagnostical purposes, since differential diagnostic problems usually concern several possibilities and high density microarrays would mean to fire at a sparrow with a huge gun [26].

No question about that the literature and the investigators are very enthusiastic about *gene expression monitoring microarrays*. The huge capacity, parallelism and automation undoubtly brought a revolution into the investigation methods and understanding of cancer and various diseases. Although reproducibility and validation remains still a problem, clustering methods with gene expression studies already resulted in huge leaps forward. Clustering methods can group the diseases upon their gene expression data into various groups with different prognosis. There are also gene expression monitoring macroarrays existing, which are manufactured according to strict rules and therefore can mean an alternative and better standardizable method of studying several targeted levels of genes instead of the house made spotted arrays.

One of the very first clinical application of gene expression microarrays was the differentiation between acut lymphocytic and myelogenous leukemia using a nearest –neighbour clustering method. The chip was not only able to distinguish between the two kinds of leukemia neccesitating quite different therapy but even the prognosis of the diseases could be estimated. This has huge clinical relevance, because before microarray studies the so called borderline leukemia were not to be grouped well phenotipically, and therefore adequate therapy could not have been promising [27].

This first success was rapidly followed by other studies and effective clustering methods upon the gathered data in the case of various cancers: breast[28],

prostate[29], soft tissue sarcomas[30], colorectal carcinomas[31], lymphomas[32]. These investigations were successful and brought clinical use and relevance of gene expression data. The data were not only good for clustering the various malignant diseases into separate groups but these clusters had also differences in patient survival, disease state, prognosis and so gene expression information really revealed new insights in cancer research. New genes came to spotlight in cancer development and potential new targets for therapy could be found. This is a real breakthrough in microarray studies which of course mean a lot of further work and analysis to be performed but these data could speed up cancer therapy development. Diseases which could not have been effectively cured before are now more capable to find the best effective and personalized treatment resulting in better prognosis.

5. Conclusion

The DNA-chips were introduced just about a decade ago and since then a lot of application was manufactured with them. No doubt they brought a detonation-like change and improvement in scientific thinking and led to the development of genomics, proteomics and a lot of other 'mics', because of their high-parallelism and capability to look into the biological processes in their complexity. Lots of various application methods and technology has been invented but the technology is a rather expensive one even today and therefore it needs more improvement, especially in the field of cost effectiveness.

There are several fields of direct usage of microarrays but the validity and reproducibility is still an issue. Data-mining is also a good field of improvement and newer algorithms, analytical methods are welcome to be able to cope with all the generated data and to find clinically relevant meanings and connections. Statistical methods, clustering methods, generating databases and working with neural networks can offer new ways of thinking about biological processes. For today we just can not cope with all the data generated by microarray studies. For these reasons DNA-chips remain rather a researchers' tool than a pure diagnostic technology for several years more. The modern medical thinking is based on evidences now and therefore it is essential to perform prospective clinical studies for the fortification an determination of the position of DNA-chips in diagnostic processes.

The most benefit will come from the gene expression microarray studies. The new findings are very promising in determining new ways of thinking about various illnesses, especially in the case of malignant diseases. Recent investigations have already brought new molecular classification, grading and staging methods for numerous cancer types. However direct usage of gene-expression DNA-chips will be not right beside the patient's bed, but instead of performing gross studies and developing screening tests which will bring new aspects, investigations and methods in to therapy and treatment. Upon microarray data simpler and easy to perform screening methods and investigation could be carried out for determining the more proper diagnoses and in turn to find the most effective therapy in less time.

Acknowledgements

I would like to thank the invitation of Prof. Zoltán Benyó to write this article and the members of the Bioinformatic Laboratory of the Budapest Technical University for their kind help.

References

[1] PENNISI, E.: DNA Chips Give New View of Classic Test; Science vol 283 January 1999: 17-18
[2] LYNCH, T.J.-BELL, D.W.-SORDELLA, R.-GURUBHAGA, V.S.-OKIMOTO, R.A.-BRANNIGAN, B.W.-HARRIS, P.L.-HASERLAT, S.M.-SUPKO, J.G.-HALUSKA, F.G.-LOUIS, D.N.-CHRISTIANI D.C.-SETTLEMAN, J.-HABER, D.A.: Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non-small-cell Lung cancer to Gefitinib., N Engl J Med. 2004 May 20;350(21):2129-39. Epub 2004 Apr 29.
[3] ILYAS, M.-STRAUB, J.-TOMLINSON, I.P.– BODMER, W. F.: Genetic Pathways in Colorectal and Other Cancers.; Eur J Cancer. 1999 Dec;35(14):1986-2002. Review.
[4] BLAY, J.Y.– LE CESNE, A.– ALBERTI, L.– RAY-COQUART, I.: Targeted Cancer Therapies.; Bull Cancer. 2005 Feb;92(2):E13-8. Review.
[5] FODOR, S.P.A.– READ, J.L.– PIRRUNG, M.C.– STRYER, L.-LU, A.T.-SOLAS, D.: Light-Directed, Spatially Addressable Parallel Chemical Synthesis; *Science* **251**, 15 February 1991, pp. 767–773.
[6] ZVARA, A.– HACKLER, L. JR. – NAGY, Z.B. – MICSIK, T.– PUSKÁS, L.G.: New Molecular Methods for Classification, Diagnosis and Therapy Prediction of Hematological Malignancies.;Pathol Oncol Res. 2002;8(4):231-40. Epub 2003 Feb 11. Review.
[7] Affymetrix: GeneChip^{TM} Assays; (www.affymterix.com)
[8] HELLER, M. J.:An Active Microelectronic Device for Multiplex DNA Analysis, *IEEE Engineering in Medicine and Biology March/April 1998*, pp. 100–105.
[9] GINGERA, T.R.– GHANDOUR, G.– WANG, E.– BERNO, A.– SMALL, P.M.– DROBNIEWSKI, F.– ALLAND, D.– DESMOND, E.– HOLODIY, M. – DENKOV, J.: Simultaneous Genotyping and Species Identification Using Hybridization Pattern Recognition Analysis of Generic *Mycobacterium* DNA Arrays; Genome Research 1998, 8: pp. 435–448.
[10] WALLACE, R. W.: DNA on a Chip: Serving Up the Genome for Diagnostic and Research; Molecular Medicine Today, September 1997 pp. 384–389.
[11] LOCKHART, D. J. – DONG, H.– BYRNE, M. C. – FOLLETIE, M. T. – GALLO, M. V. – CHEE, M. S. – MITTMANN, M. – WANG, C. – KOBAYASI, M. – HORTON, H. – BROWN, E.L.: Expression Monitoring by Hyridization to High-density Coligonucleotide Arrays, *Nature Biotechnology* **14**, December 1996 pp. 1675–1680.
[12] DERISI, J.– PENLAND, L. – BROWN, P. O. – BITTNER, M. L. – MELTZER, P.– RAY, M.– CHEN, Y.– SU, Y. A.– TRENT, J. M. : Use of cDNA Microarray to Analyse Gene Expression Patterns in Human Cancer; *Nature Genetics*, **14**, December 1996, pp. 457–460.
[13] DEBOUCK, C.– GOODFELLOW, P. N.: DNA Microarrays in Drug Discovery and Development; Nature Genetics Supplement **21**, January 1999, pp. 48–50.
[14] TURECI, O.– DING, J.– HILTON, H. – BIAN, H.– OHKAWA, H.– BRAXENTHALER, M.– SEITZ, G.– RADDRIZZANI, L.– FRIESS, H. – BUCHLER, M. – SAHIN, U. – HAMMER, J.: Computational Dissection of Tissue Contamination for Identification of Colon Cancer-specific Expression Profiles FASEB J. 2003 Mar;17(3) pp. 376–85.
[15] VENET, D. – PECASSE, F. – MAENHAUT, C. – BERSINI, H.: Separation of Samples into their Constituents Using Gene Expression Data; *Bioinformatics*, **17** Suppl.1 2001, pp. 279–287.
[16] LIU, B.– CUI, Q. – JIANG, T. – MA, S.: A Combinational Feature Selection and Ensemble Neural Network Method for Classification of Gene Expression Data, BMC Bioinformatics. 2004 Sep 27, **5(1)**, p. 136.

[17] SHIMKETS, R. A.: Gene Expression Quantitation Technology Summary, *Methods Mol Biol.*, 2004, **258**, pp. 7–12.
[18] SHI, L.– TONG, W. – GOODSAID, F.– FRUEH, F. W. – FANG, H. – HAN, T. – FUSCOE, J. C. – CASCIANO, D. A.: QA/QC: Challenges and Pitfalls Facing the Microarray Community and Regulatory Agencies, *Expert Rev Mol Diagn.*, 2004 Nov;4(6) pp. 761–77.
[19] BROWN, P.O.-BOTSTEIN, D.: Exploring the New World of the Genome with DNA Microarrays, *Nature Genetics Supplement* **21**, January 1999 pp. 33–37.
[20] SOTHERN, E. – MIR, K. M. – SHSCHEPINOV, M.: Molecular Interactions on Microarrays, *Nature Genetics Supplement* **21**, January 1999 pp. 5–9.
[21] Affymetrix GeneChip^{TM} systems (www.affymetrix.com)
[22] TULSI, B.: Converting Protein Data into Knowledge for Discovery, Computational Biology; Genomics and Proteomics **4**, No.9, December 2004, pp. 28–32.
[23] BASSETT, D. E. JR. – EISEN, M. B. – BOGUSKI, M. S.: Gene Expression Informatics — it's all in your mine, *Nature Genetics Supplement* **21**, January 1999 pp. 51–55.
[24] Affymetrix: GeneChip HIV PRT Assay; (part #700116)
[25] ANTHONY, R. M. – BROWM, T. J. – FRENCH, G. L.: DNA Array Technology and Diagnostic Microbiology; Expert Rev Mol Diagn. 2001 May;1(1) pp. 30–8.
[26] ZENG, X.: The Making of a Portrait – Bringing it into Focus, *Curr Pharm Biotechnol.* 2003 Dec;4(6), pp. 397–9.
[27] GOLUB, T. R. – SLONIM, D. K. – TAMAYO, P. – HUARD, C. – GAASENBEEK, M. – MESIROV, J. P. – COLLER, H. – LOH, M. L. – DOWNING, J. R. – CALIGIURI, M. A. – BLOOMFIELD, C. D. – LANDER, E. S.: Molecular Classification of Cancer: class discovery and class prediction by gene expression Monitoring.; Science. 1999 Oct 15, **286**(5439) pp. 531–7.
[28] SORLIE, T. – PEROU, C. M. – TIBSHIRANI, R. – AAS, T. – GEISLER, S. – JOHNSEN, H. – HASTIE, T.– EISEN, M. B. – VAN DE RIJN, M. – JEFFREY, S. S. – THORSEN, T. – QUIST, H. – MATESE, J.C. – BROWN, P.O. – BOTSTEIN, D. – EYSTEIN, L. P. – BORRESEN – DALE, A.: Gene Expression Patterns of Breast Carcinomas Distinguish Tumor Subclasses with Clinical Implications; *Proc Natl Acad Sci U S A.* 2001 Sep 11;98(19):10869-74.
[29] DHANASEKARAN, S. M. – BARRETTE, T. R. – GHOSH, D. – SHAH, R. – VARAMBALLY, S.– KURACHI, K. – PIENTA, K. J. – RUBIN, M. A. – CHINNAIYAN, A.M.: Delineation of Prognostic Biomarkers in Prostate Cancer, *Nature*, 2001 Aug 23;412(6849) pp. 822–6.
[30] SCHOFIELD, D. – TRICHE, T. J.: cDNA Microarray Analysis of Global Gene Expression in Sarcomas, Curr Opin Oncol. 2002 Jul;14(4) pp. 406-11. Review.
[31] BERTUCCI, F.– SALAS, S. – EYSTERIES, S. – NASSER, V.– FINETTI, P. – GINESTIER, C. – CHARAFE-JAUFFRET, E. – LORIOD , B. – BACHELART, L. – MONTFORT, J. – VICTOREO, G. – VIRET, F. – OLLENDORFF, V. – FERT, V. – GIOVANINNI, M. – DELPERO, J. R. – NGUYEN, C. – VIENS, P. – MONGES, G. – BIRNBAUM, D. – HOULGATTE, R., Gene Expression Profiling of Colon Cancer by DNA Microarrays and Correlation with Histoclinical Parameters; Oncogene. 2004 Feb 19;23(7) pp. 1377-91.
[32] ALIZADEH, A. A. – EISEN, M. B. – DAVIS, R. E. – MA, C. – LOSSOS, I. S. – ROSENWALD, A. – BOLDRICK, J. C. – SABET, H. – TRAN, T. – YU, X. – POWELL, J. I. – YANG, L. – MARTI, G. E. – MOORE, T. – HUDSON, J. JR.– LU, L. – LEWIS, D. B. – TIBSHIRANI, R. – SHORLOCK, G. – CHAN, W. C. – GREINER, T. C. – WEISENBURGER, D. D. – ARMITAGE, J. O. – WARNKE, R. – LEVY, R. – WILSON, W.– GREVER, M.R. – BYRD, J.C. – BOTSTEIN, D. – BROWN, P. O. – STAUDT, L. M. : Distinct Types of Diffuse Large B-cell Lymphoma Identified by Gene Expression Profiling, *Nature*, 2000 Feb 3; 403 (6769) pp. 503–11.