Article is available online at http://www.webio.hu/por/2002/8/4/0231

ARTICLE

New Molecular Methods for Classification, Diagnosis and Therapy Prediction of Hematological Malignancies

Ágnes ZVARA, László HACKLER Jr., B Zsolt NAGY, Tamás MICSIK, László G PUSKÁS

Laboratory of Functional Genomics, Biological Research Center, Hungarian Academy of Sciences, Szeged, Hungary

Normal functions of the cell are based on the precise regulation of various genes. If this strict regulation and the hierarchy of genes becomes upset due to flaws in this system, the result will be cellular dysfunction which eventually may lead to carcinogenic transformation. Two basic challenges of the classification of cancers are the discovery of new molecular markers characteristic to defined disease groups and the classification of already diagnosed or new cases into existing groups. This precise classification may open the door to tailored treatment or project the expected outcome of the disease. Today there is unlimited access available to the databases containing sequences and localization of the genes within the confines of Human Genome project. It provides significant help for the discovery of chromosome abnormalities and systematic analysis of gene expression patterns. This is important not only to understand normal functions of the cells, but it also contributes to the identification of new genes that are characteristic to given disease groups as markers and that are potential drug targets. Until the second half of the twentieth century the study of the function and regulation of genes was based on step-by-step investigation of individual genes. Regarding the fact, that the genomes of an increasing number of organisms have become known in whole or in part, numerous new techniques have been developed that facilitated the systematic analysis of gene functions. The aim of this study is to summarize the new, molecular based possibilities for classification, diagnosis and prognosis of hematological malignancies, as well as to summarize the main results of these areas. (Pathology Oncology Research Vol 8, No 4, 231–240)

Keywords: molecular methods, DNA-chip, oncohematology

Classification of human cancers traditionally based on tissue of origin or histological characteristics. However, genetic defects that basically determine the abnormal behavior of tumor cells, can be analyzed more effectively using the knowledge of almost the total sequence of the human genome. New, high throughput molecular methods have been developed (mutation analysis of the whole genome, comparative genomic hybridization, gene expression monitoring with cDNA microarrays, array-based protein expression and interaction analysis) to obtain more information about different malignant transformations at different system levels (genome, transcriptome, proteome) (*Table 1*). Results obtained with these methods can help to answer the questions arising in connection with prevention, diagnosis and classification, therapy and outcome of the disease.[8,31,61] The most studied disease among the carcinogenic transformations, from the point of view of functional genomics, is leukemia.

Acute leukemia is a complex disease that can be classified into individual sub-groups based on morphological, immunological and cytogenetic characteristics and their response to chemotherapy. Classification of the disease and therapeutic grouping of the patients are currently quite difficult and costly tasks, which require serious laboratory (e.g. immuno-phenotyping), cytogenetic and molecular diagnostic tests as well as a wide-ranging medical specialist background (hematologist, oncologist, pathologist, cytogeneticist). Individual and fast testing cannot provide sufficient basis for diagnosis. Despite the thorough and laborious diagnostic processes, the classification of leukemias is usually not perfect and error-prone.[52] From a clinical standpoint, the classification of

Received: Oct 18, 2002; *accepted:* Dec 15, 2002

Correspondence: László PUSKÁS, Laboratory of Functional Genomics, Biological Research Center, Hungarian Academy of Sciences, Temesvári krt. 62, H-6726 Szeged, Hungary

Table 1. **Comparative table of different methods for DNA, mRNA and protein analysis at different system levels.**

System level	*Technique*	*Application*	*Throughput*		*Biological requirement*
			Biological sample Sample No.	*Gene, transcript or protein No.*	
Genome (genes)	Sequencing	Detection deletions / amplifications, mutations and polymorphic sites (SNP) analysis	medium	low	Genomic DNA from fresh or archive biological sample
	Chromosome band analysis (FISH)	Detection of chromosome abnormalities, deletions / amplifications	low	low	Genomic DNA from fresh or archive biological sample
	Microsatellite analysis	Detection of chromosome abnormalities, deletions / amplifications	medium	low	Genomic DNA from fresh or archive biological sample
	Array-based comparative genomic hybridization (CGH)	Detection of chromosome abnormalities, deletions / amplifications	medium	high	Genomic DNA from fresh or archive biological sample
	Microarray-based methylation pattern analysis	Analysis of methylation pattern of the genome	medium	high	Genomic DNA from fresh biological sample
	DNA or oligonucleotide microarray	Mutation and polymorphic sites (SNP) analysis, detection of chromosome abnormalities, deletions / amplifications	medium	high	Genomic DNA from fresh or archive biological sample
	Northern blot analysis	Gene expression analysis	low	low	RNA from fresh biological sample
	Differential display	Differential gene expression analysis	low	high	RNA from fresh biological sample
	Subtractive hybridization	Differential gene expression analysis	low	high	RNA from fresh bio logical sample
Transcriptome (mRNA)	cDNA Representational Difference analysis (RDA	Differential gene expression analysis	low	high	RNA from fresh bio logical sample
	Serial analysis of gene expression (SAGE)	Differential gene expression analysis	low	high	RNA from fresh biological sample
	Real –time PCR	Comparative gene expression analysis	high	low	RNA from fresh biological sample
	cDNA microarray	Expression profiling, differential gene expression analysis	medium	high	RNA from fresh biological sample
Proteome (protein)	Two-dimensional gel-electrophoresis	Analysis of protein profile	low	high	Denatured protein lysates
	Two-hybrid system	Protein-protein interactions	low	high	Cloned cDNA of protein of interest and cDNA library from given organism
	Protein arrays	Analysis of protein profile, protein-protein, protein-DNA interactions	medium	high	Native protein mixtures or denatured protein lysates

leukemias is important for two reasons: firstly to understand the initial causes of the disease and to project a view on the final processes; and secondly for diagnosis supported by molecular results toprovide optimal therapy and recovery.

Acute myeloid leukemia (AML) is classified into eight morphological subgroups (M0-M7), which all show a correlation with clinical appearance, cytogenetic markers and severity of the disease.[56] Classification of acute myeloid leukemias is currently based on cytogenetic differences instead of the earlier FAB system based on morphology (existence or lack of certain chromosome translocations). Leukemias can also be classified based on special chromosome abnormalities, which can be identified in more than 90 % of the cases.[27] A few chromosome abnormalities have therapeutic and prognostic significance.[5,26,2]

Acute lymphoid leukemias (ALL) can be classified into T and B cell types, and further subgroups identified based on cell surface markers, with the help of immunhistochemistry. The classification has prognostic significance in cases of both leukemias, but only in the statistical sense, because patients (especially AML) with the same cytogenetic classification may show individual variations. It is a frequent observation that leukemias classified into the same group based on various criteria follow a completely different clinical course, and react completely different ways to the same therapeutic intervention.

The first classification of ALL reflected various clinical outcomes of the disease and fine cytomorphological differences. At the end of the 1960's researchers tried to classify the disease with the help of a new method – enzyme based histochemistry. They managed to show that some of the leukemias proved to be acid-Schiff positive, while others proved to be myeloperoxidase positive with the periodate- acid- dying method. This gave the first basis for the molecular level classification of leukemias. According to this, we can differentiate between ALL with either thymus-T-cell-precursor origin (T-ALL) or bone-marrow-B-cell precursor origin (B-ALL) and AML. This classification gained further confirmation in the 1970's using antibodies recognizing cell surface antigens.[57] This review gives a short overview of the new, molecular-based, high throughput methods applied successfully for classification and therapy prediction of leukemias.

High throughput analysis of chromosome abnormalities based on comparative genomic hybridization (CGH)

Changes within the chromosome: deletion or amplification is quite frequent in neoplasias.[48] The specific rearrangements, in many cases, are characteristic to the individual tumor types and states. Genes mapped to the locations of these rearrangements can play roles in the formation of tumor. Their investigation can contribute to the characterization of the different tumors and tumor stages. Clinically distinct subgroups of ALL patients are characterized by chromosome translocations and chromosome imbalance. These abnormalities are the defining molecular features of ALL. In case of B-ALL the presence of a fusion oncoprotein may frequently be observed.[22] Patients with t(12;21)/TEl-AML1 and t(1;19)/E2A-PBX1 translocations is usually expected to give relatively good treatment outcome, especially by raising the dosage of chemotherapy, however the prognosis for patients with translocations t(9;22)/BCR-ABL and t(4;11)/MLL-AF4 are much worse.[6] Patients with B-ALL characterized by high chromosome number (hyperdiploid chromosome content, 50) have also good treatment outcome. T-ALL is characterized by repetitive but relatively rare chromosome translocation, which usually affects genes encoding for transcription factors (LYL1, HOX11, HOX11L2, TAL1), their abnormal expression impairs early thymocyte differentiation and finally leads to leukemia.[22]

Techniques based on *chromosome band analysis, fluorescence in situ hybridization (FISH)* have proven to be useful in detecting rearrangements and imbalances, but less informative in recognizing potentially amplified or deleted regions. Using this method, the detection of nucleotide sequences on DNA molecule is performed indirectly, by first hybridizing the seeker nucleotide sequences (probes) with the DNA (also called target). If the probes are synthesized with incorporated fluorescent molecules or antigenic sites that can be recognized with fluorescent antibodies, the direct visualization of the relative position of the probes is possible.[73] In addition, these techniques require analysis of metaphases after cell culture, which are time-consuming and misleading in the aspect of selecting subclones having growth advantage. The technique based on *microsatellite analysis* is a higher resolution possibility for chromosome region instability than FISH. Microsatellites are short, highly polymorphic tandem repeat sequences suitable for detection of chromosome imbalances. PCR amplification of these certain region give direct information about the copy number alteration of a given chromosome segment. The technique was applied successfully to detect submicroscopic chromosomal deletions in hematological malignancies, as well.[49, 65]

Comparative genomic hybridization (CGH) is a rapid method that does not require cell cultures and provide more accurate information about the possible occurrence of chromosome amplifications than do cytogenetic methods.[34] It provides a lot of information about the genomic balance of tumor cells, mono- or trisomies, amplifications and deletions in a simple experimental procedure. Using this technique, DNA purified from the tumor-tissue is labeled with green-fluorescent dye (Cy3),

while the DNA purified from the healthy control is labeled with another, red fluorescent dye (Cy5). They are hybridized together onto a DNA chip containing high number of genomic DNA fragments or cDNAs of already sequenced and identified genes immobilized on glass surface. After reading with a confocal laser scanner, differences in color can be measured and analyzed with computer software; the determined Cy3/Cy5 ratio indicates deletion, or amplification of the given gene (*Figure 1*).

Pinkel et al. performed an initial study in this area. They immobilized bacterium artificial chromosomes (BAC) and genomic fragments from human chromosome 20 on glass surface and demonstrated the feasibility of detecting both gains and losses with single copy sensitivity using array CHG.[50] Despite the fact, that the CGH method based on DNA microarrays is not suitable for the identification of small mutations (SNPs or deletions/amplifications of few nucleotide within genes), it offers a new solution for amplification/deletion analysis applied so far, because it is extremely well applicable and high throughput way for the overall analysis of the whole genome. It has the significant advantage of being less sensitive to cell contamination: a single gene-copy change may be detected from a sample containing up to 60% of normal, healthy cell contamination.[33] A great advantage is that good quality labeled probe can be obtained even from a small amount of paraffin-archived material. The results can be easily analyzed with the help of currently available databases that contain significant amount of information about the chromosomal location of the genes identified in the experiments.

In cases of adult T-cell leukemia, both deletion and amplification were successfully determined by this method. Testing of 64 patients showed amplification in 14q, 7q, and 3q chromosome regions, while in the regions 6q and 13q, deletions were observed. These chromosome changes were much more frequent in patients with an aggressive form of leukemia than in the indolent form. An increased number of chromosome imbalances were detected in patients, where the chance of survival was significantly lower.[68] In cases of acute myeloid leukemia, a high level of DNA amplification was detected at the chromosome region 8q24, the locus of the *c-myc* proto-oncogene using CGH.[10] The authors could subdivide the reported AML cases into two groups: cases with *c-myc* amplification together with further complex chromosome aberrations were associated with rapid disease progression and short survival time; while cases with *c-myc* amplification accompanied by a single other chromosome aberration (e.g. loss of one X chromosome or trisomy of chromosome 4) showed good response to chemotherapy and had prolonged survival time.

In these cases, data obtained from CGH experiments provided more information about the outcome of the disease, and distinguished between AML cases for diagnosis and therapy. The clonal identity of unusual metastases and the suspected primery tumor can be also successfully comfirmed by the CGH technique using cDNA microarray.[24]

In case of childhood ALL, deletion of the 12p chromosome region was also successfully shown with the CGH technique, which may offer favorable prognosis.[39] DNA chip-based CGH method may provide significant help in the identification of new genes that may play a role in tumorigenesis (novel oncogenes and tumor suppressors) including leukemia, in the follow up examination of metastases during the course of tumor development, and in classification of different cases. CGH has been used to determinate DNA copy number alterations in many other cancer types, like renal cancer,[74] breast cancer[7] or thyroid cancers.[76.]

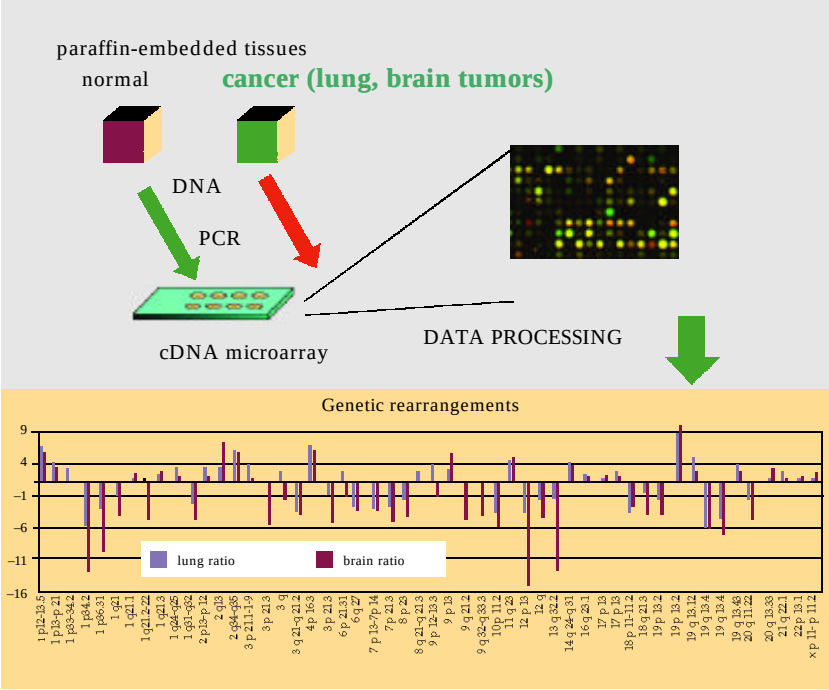

Figure 1. Comparative Genomic Hybridization for the detection of genetic rearrangements. Genetic rearrangements of two tumor tissues (brain and lung) have been compared in order to confirm that the brain metastasis derived from the lung cancer

Classification based on DNA-methylation-pattern analysis by DNA microarray

Aberrant DNA methylation pattern of CpG islands is one of the earliest and most common alterations in human malignancies. Several studies confirm the hypothesis that aberrant methylation occurs in a tumor-specific manner. Gain or loss of function of wide range of genes was shown in these cases.[38] Abnormal methylation patterns can be detected in CpG-rich regulatory elements either in introns or in exons.[12] Tumor cells are generally characterized by the hypermethylation of tumor suppressor genes and, in contrast, hypomethylation of the whole DNA molecule. This general hypomethylation can be detected relatively early, before the development of the actual tumor. Correlation between hypomethylation and increased gene expression can be detected in cases of large number of oncogenes.[21,18] With a genome-wide approach suitable for whole genome screening, it was shown that the methylation pattern may vary among different tumors.[15,35] The methylation pattern of the as whole genome, similarly to the gene expression pattern, is characteristic to the given cell, tissue, or disease as a molecular fingerprint. Analysis of genome wide methylation profiles enables the caracterization of new tumor classes, or the classification newly diagnosed cases into already existing groups based on their methylation pattern. In recent years a new method had emerged for the analysis of methylation pattern extending to the whole genome that is suitable for analyzing large number of genes simultaneously.[1,25,67]

Classification based on gene expression changes

A primary goal of expression profiling study is to characterize genes that expressed abnormally due to the different chromosomal abnormalities mentioned before. Currently there are two main ways to analyze molecular expression patterns: (1) generating mRNA-expression profile, the “transcriptome” and (2) analyzing the complete protein population, the “proteome”. Tools for studying gene-expression at the transcript level can be divided into three major groups: (1) hybridization-based techniques, such as Northern blotting, subtractive hybridization, DNA microarrays or macroarrays, (2) PCR-based techniques such as differential display, RDA (representational difference analysis) (3) Sequence-based techniques such as SAGE (serial analysis of gene expression)[42]

The breakthrough in studying gene-expression was *Northern blot analysis* that made the identification of different mRNA in a given sample possible. In this technique, labeled RNAs or cDNA probes hybridized to RNA fragments separated by gel electrophoresis and immobilized onto nylon filter.[4] This technique is still a useful method for detecting a transcript in a given mRNA population and confirming expression data obtained with other experiments.

The first technique that enabled identification of differentially expressed genes was based on *subtractive hybridization*. This technique based on development of subtractive cDNA libraries, which are generated by hybridizing an mRNA pool of one origin to an mRNA pool of a different origin.[32] Transcripts that do not have a complementary strand in the hybridization step are then used for the construction of a cDNA library. A variety of refinements of this method have been developed to identify specific mRNAs.[64] Despite the fact that numerous genes were successfully identified with this method, it had serious disadvantages: 1. only a small fraction of the gene expression differences were successfully discovered, 2. it required a large amount of RNA sample, and 3. it proved to be quite laborious and time-consuming.

In 1992 it was replaced by a new technique, called *differential display PCR (DD-PCR).* This technique was the first one-tube method to compare differentially expressed genes systematically.[44] mRNA obtained from two different cell populations or tissues was amplified by PCR after a reverse transcription (RT) step, and then the generated fragments, that reflected the expression pattern of the given cell populations or tissues, were separated by denaturing gel electrophoresis. Differentially expressed genes could be isolated from the gel, sequenced and identified. Numerous studies have been published, which - despite the serious disadvantages (maintenance the quantitative correlation after RT and PCR reactions, repeatability, and the elimination of false positive signals) - applied this method successfully.[51,71]

cDNA Representational Difference analysis (RDA) became widely known as an alternative method, and is a PCR based subtractive hybridization procedure. It was originally developed for the comparison of two complex genomes, and then it was later adapted for cloning differentially expressed genes.[46] mRNA of different origins (“tester” for treated or diseased sample and “driver” for control sample) is reverse transcribed, digested with a frequently cutting restriction endonuclease and then amplified according to the ligated linker sequence. The linker sequences of both driver and tester pools are removed and a new sequence is linked to the tester cDNA. After hybridization of both cDNAs and subsequent PCR, only those cDNAs present in the tester population are amplified. The advantage of this technique is the specific amplification of different fragments presented only in one DNA population. Both differential display and the RDA are suitable techniques for analyzing small number of transcripts.

Serial analysis of gene expression (SAGE) technique uses a sequence-based strategy that allows parallel analysis of a large number of transcripts.[70] The method is based on two principles: 1, a short, 8-9 base pair long nucleotide sequence tag contains enough information for the identification of the transcript, 2. concatenation of these short tags

allows the efficient analysis of transcripts in a serial manner by the sequencing of multiple tags within a single clone. Results obtained with this technique allow the determination of significant quantitative relationship between mRNA populations derived from various experimental procedures. In contrast to other methods mentioned before, it is much more sensitive in detecting low copy number transcripts. The aforementioned techniques are material-intensive and time consuming. For these reason, efforts had been undertaken to develop methods for high-throughput screening.

In recent years, a new technique, the *DNA microarray technology* (DNA-chip) has emerged offering the possibility of high-throughput systematic analysis of the "transcriptome". It is the most informative and most effective functional method of techniques mentioned so far, regarding the time and work necessity.[16,8,20,62] The arrays are constructed of thousands of DNA fragments either spotted or synthesized (Affymetrix) onto chemically activated glass slides. DNA fragments can be collections of short or long oligonucleotides or cDNAs of variable length. Oligonucleotide microarrays are suitable for the simultaneous detection of several thousands of single nucleotide mutations polymor, phismus (SNPs) in an amplified genomic DNA sample. This has a great significance in searching for polymorphic loci, or in detection of single or multiple mutations in medical samples. The most important and most informative application of DNA-chips is the parallel study of gene expression from different biological samples that focuses on the functionally active parts of the genome. DNA microarrays with sets of cDNA fragments on their surface can be used to obtain a molecular fingerprint of gene expression of cells.[9,41,45] The method has enabled large numbers of genes, from specific cell populations, to be studied in a single experiment. An important difference between the oligonucleotide and cDNA chips is the hybridization step. While oligonucleotide chips are hybridized by labeled RNA (the different sample RNAs hybridized onto different chip), in the case of cDNA chips, labeled cDNA samples are hybridized together onto one chip. This method is a reverse blotting technique, where mRNA populations gained from diverse biological sample (tissues or cell cultures) converted to cDNA in the presence of fluorescence dye (Cy3 or Cy5) labeled nucleotides. Using a co-hybridization strategy, with Cy3-labeled cDNA from the test sample and Cy5-labeled cDNA from the control sample, the relative intensity ratio on the microarrays can be determined and the expression pattern can be analyzed (*Figure 2.*). The quantity of the starting mRNA depends on the type of tissue (e.g. liver or brain tissues). The amount of RNA is strongly limited in those cases, where the amount of the sample tissue is small, for example in cases of biological sample obtained with laser microdissection or other operative methods, or

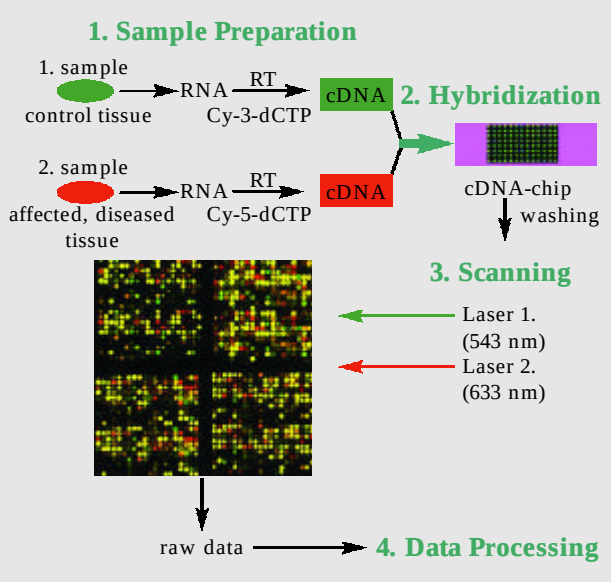

Figure 2. Application of cDNA arrays for the follow up detection of gene expression changes

in experimental systems where 1000–5000 cells are the object of the investigation. In these cases, amplification of the RNA sample or signal is necessary. During the amplification steps, however, it is very important to keep the quantitative ratios presented in the original mRNA population. The proper use of exponential (PCR) and linear (*in vitro* transcription, IVT) amplification can solve this problem.[53,54,40] By this technique, changes in cells generated by various effects (e.g. pharmaceutical treatment, pathological processes) can be traced, new biochemical markers and genes responsible for pathological phenotype can be discovered, drug effects can be followed and the treatment can be optimized. The differences in gene expression of the treated and untreated cells or tissues provide information about the regulation of the enzymatic pathways influenced by drugs, about the enzymes, transporters playing a role in drug resistance. Identification of gene expression patterns may provide vital information for the understanding of the pathological processes and may contribute to diagnostic decisions and therapies tailored to the individual patient.

The discovery of gene expression changes associated with the chromosome imbalances described in the previous section can have crucial diagnostic and therapeutic values. The molecular basis of T-cell acute lymphoid leukemias was largely discovered by the analysis of chromosome translocations and intrachromosomal rearrangements. These abnormalities typically reflect such arrangements where strictly regulated genes that important for cell-maturation (e.g. HOX11, LMO1, LMO2) become regulated by strong promoters or enhancers (e.g. T-cell receptor promoter). These rearrangements greatly increase the expression of these genes

and cause the abnormal maturation of precursor cells. Little is known about the downstream mechanisms that maintain the T-ALL phenotype. By using oligonucleotide chips, Ferrando et al. have shown that in case of T-ALL, five different T-cell oncogenes (HOX11, TAL1, LYL1, LMO1, LMO2) are frequently expressed abnormally in the absence of chromosome abnormalities. During the maturation of thymocytes, expression of these oncogenes is characteristic to specific developmental stages of the cells (e.g. LYL1+ [pro-T], HOX 11+ [early cortical thymocyte], TAL1+ [late cortical thymocyte]). They identified several gene expression signatures that were indicative of leukemic arrest at specific stages of thymocyte development. Samples could be grouped according to their shared oncogenic pathways by hierarchical clustering analysis.[23] These findings can help to form subgroups that are also important from the point of view of therapy.

Golub and colleagues applied successfully the DNA microarray technology to test whether the differencies in gene expression profile of ALL and AML can be used to distinguish between patients suffering from the two cancer types. They identified numerous genes that were associated specifically with either AML or ALL. Using statistical methods sets of genes were defined as "class predictors" helping the accurate classification. They also showed that this approach could be useful for characterization of new classes of tumors ("class discovery").[28]

In those cases where the applied intensity of therapy and possible side effects are the object of more serious consideration (e.g. children), classification of leukemias based on gene expression changes play a very important role in tailoring the therapy to patients and to define optimal therapeutic groups. Pediatric ALL were classified by Yeoh et al. based on gene expression differences using DNA microarrays.[69] Cases were classified into seven groups based on gene expression with hierarchical clustering. Using gene expression patterns subgroups of childhood ALLs can be accurately identified with an overall diagnostic accuracy of 96% and used in the class prediction and therapeutic decisions. The identification of molecular signatures characteristic to distinct subgroups was based on those genes that were differentially expressed and clustered into distinct groups by statistical methods. In case of B-ALL the gene expression pattern not only facilitates the classification into different subgroups, but provides explanation for occasional failure of therapy. Different leukemia subgroups followed different oncogene activation mechanisms, have different signal-transduction pathways, and therefore respond to therapies in different ways. Chemotherapy-induced secondary AML can also be identified by using distinct expression profiles (molecular signatures) despite the fact that secondary AML is thought to arise from a hematopoetic stem cells different from that giving rise the primary leukemia.[69]

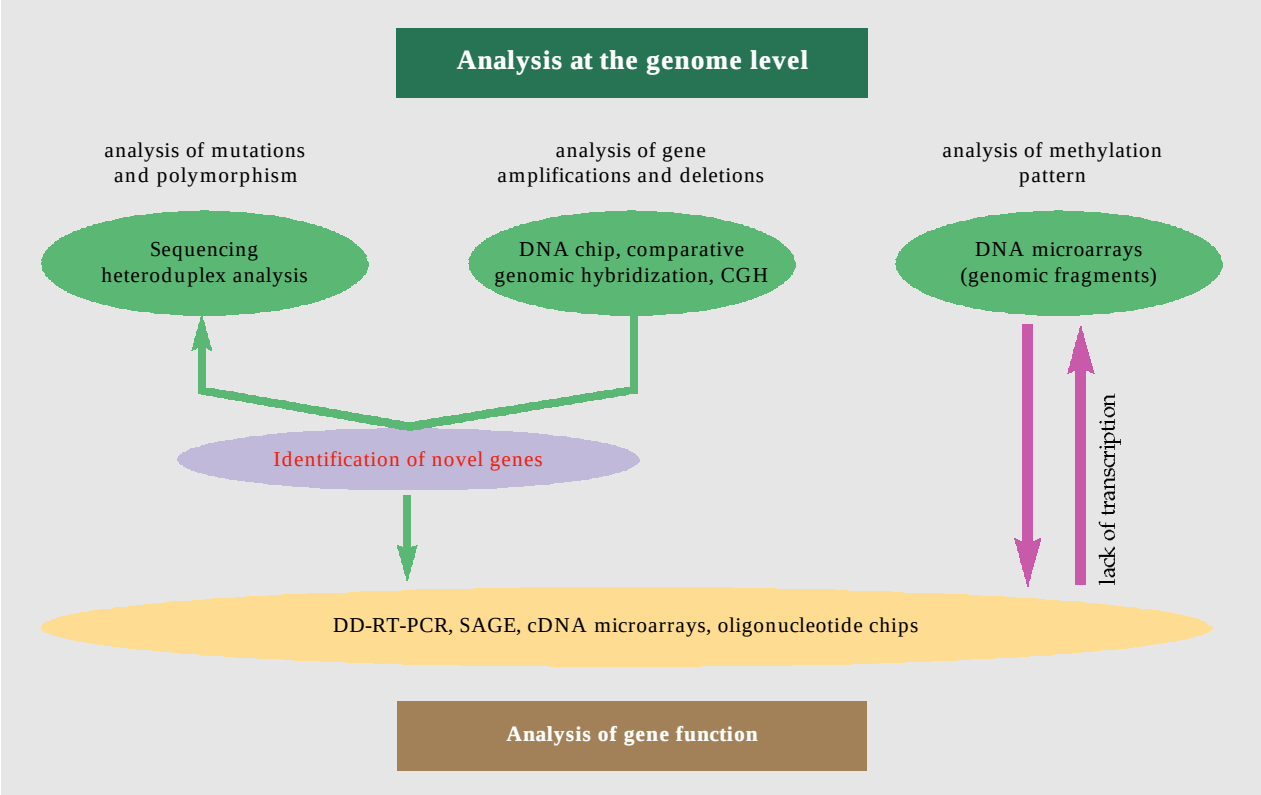

Figure 3. Analysis of genome and gene function with various techniques

Expression profiling with DNA-chip offers a viable and applicable alternative method in comparison to traditional methods (chromosome band analysis or *in situ* fluorescent hybridization) for diagnosis, treatment, and prediction of outcome of cancers especially in cases where the chromosome abnormalities cannot be detected with cytogenetic methods.[28,55,59] A general review by Kozian et al. is an excellent overview of different techniques used in differential gene expression analysis studies.[42]

Proteome level: Array-based analysis of protein expression pattern and protein function

Although transcript profiling offers good opportunity to identify genes that play a role in cancer formation, even complete mRNA fingerprints have their limitations. It is obvious that cellular functions are carried out by proteins, not by DNA or RNA. Numerous protein modifications, such as RNA splicing and posttranslational modification (e.g. phosphorylations, glycosylations), are known to influence protein function. The genomic or the transcript sequences do not give information about the different protein-protein interactions, or how and where these interaction occur inside the cells under various conditions. To obtain detailed information about a complex biological sample, information about many proteins, and protein-protein interaction is required. A new field of cell function analysis, the field of proteomics, has emerged based on the complete set of proteins expressed by the genome (called proteome). From clinical standpoint, the field of proteomics has great potential because the drug treatment and therapy directed at protein level rather than genes. The global analysis of protein profile has many difficulties, since the proteins tertiary, native structure should be maintained for successful binding assays. Protein regulation often based on reversible modification instead of synthesis or degradation. Strategies to solve these problems have been reviewed by Weinbarger et al.[72]

Until recently, global protein expression analysis was performed using two-dimensional (2D-PAGE) gel electrophoresis, while methods for the analysis of protein-protein interactions were limited to the yeast two-hybrid system. Although thousands of proteins can be displayed in a gel by 2D-PAGE, and can be tested for potential interaction by two-hybrid systems, these techniques are labor intensive and not suitable for high-throughput applications. New, high-throughput approaches, such as ProteinChip® proteomic platforms had been developed to override these limitations.[2,11,19] Haab et al adapted the dual fluorescent labeling that used for comparative expression analysis to antigen-antibody microarrays.[29] Either antibodies or antigens were immobilized, while the corresponding targets obtained from complex biological sample such as serum were fluorescently labeled. The two samples were mixed and incubated simultaneously on the same microarray. Dual color detection system revealed the quantitative differences between the two samples. Since monoclonal antibody production is labor intensive, development of alternative methods, such as high-throughput generation of recombinant proteins have become crucial. The recently developed microarray-based protein-protein interaction (protein-protein, small synthetic molecules-protein, enzyme-substrate, receptor-ligand) assays give the possibilities for high-throughput protein function assays. Some of the current technologies for proteome profiling and the application of proteomics to the analysis of leukemias have been reviewed by Hannash et al. [30]

The technical developments of the last 15 years have revolutionized the molecular understanding, the discovery of functional genomic backgrounds, and the identification of diagnostic markers of leukemias and other cancers. A part of these techniques strive to detect genome wide abnormalities (mutations, gene amplifications-deletions, methylation pattern), while another part aims to analyze the products of active genes, the mRNA populations (transcriptome) (gene expression analyses, SAGE, DD-RT-PCR, DNA chip technology) and protein composition and function (proteome) of cells (*Figure 3*). The number of studies analyzing the functional genomics of cancers are growing exponentially, for example melanoma,[8,1,75] colon cancer,[45,17,63,62] thyroid tumor,[66,13,36] breast cancer,[37,14,43] prostate cancer.[60,58,47] In the future, hopefully, these techniques will be available for clinical use helping diagnosis, prognostication the outcome of the disease, and choosing the most appropriate therapy.

References

1. *Adorján P, Distler J, Lipscher E, et al:* Tumour class prediction and discovery by microarray-based DNA methylation analysis. Nucleic Acids Res 30:e21, 2002.
2. *Albala JS and Humphrey-Smith I*: Array-based proteomics: high-throughput expression and purification of IMAGE consortium cDNA clones. Curr Opin Mol Ther 680–684, 1999.
3. *Alizadeh AA, Ross DT, Perou CM, et al:* Towards a novel classification of human malignancies based on gene expression patterns. J Pathol 195: 41-52, 2001.
4. *Alwine JC, Kemp DJ, Stark GR:* Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci USA 74: 5350-5354, 1977.
5. *Appelbaum FR:* Molecular Diagnosis and Clinical Decisions in Adult Acute Leukemia, Seminars in Hematology 36: 401-410, 1999.
6. *Arico M, Valsecchi MG, Camitta B et al:* Outcome of treatment in children with Philadelphia chromosome-positive acute lymphoblastic leukemia. N Engl J Med 342: 998-1006, 2000.
7. *Aubele M, Auer G, Braselmann H et al:* Chromosomal imbalances are associated with metastasis-free survival in breast cancer patients. Anal Cell Pathol 24: 77-87, 2002.

8. *Bittner M, Meltzer P, Chen Y et al:* Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406: 536-540, 2000.
9. *Blohm DH, Guiseppi-Elie A*: New developments in microarray technology. Curr Opin Biotechnol 12: 41-47, 2001.
10. *Bruckert P, Kappler R, Scherthan H et al:* Double minutes and c-MYC amplification in acute myelogenous leukemia: Are they prognostic factors? Cancer Genet Cytogenet 120: 73-79, 2000.
11. *Cahill DJ:* Protein arrays: a high-throughput solution for proteomics research? Proteomics: A Trends Guide, 47–51, 2000.
12. *Chan MF, Liang G and Jones PA:* Relationship between transcription and DNA methylation. Curr Top Microbiol Immunol 249: 75–86, 2000.
13. *Chen KT, Lin JD, Chao TC et al:* Identifying differentially expressed genes associated with metastasis of follicular thyroid cancer by cDNA expression array. Thyroid 11: 41-46, 2001.
14. *Clark J, Edwards S, John M et al:* Identification of amplified and expressed genes in breast cancer by comparative hybridization onto microarrays of randomly selected cDNA clones. Genes Chromosomes Cancer 34: 104-114, 2002.
15. *Costello JF, Fruhwald MC, Smiraglia DJ et al:* Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. Nature Genet 24: 132–138, 2000.
16. *DeRisi J, Penland L, Brown P et al:* Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nature Genet 14: 457-460, 1996.
17. *Dunican DS, McWilliam P, Tighe O et al*: Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. Oncogene 21: 3253-3257, 2002.
18. *Eads CA, Danenberg KD, Kawakami K et al:* MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 28-c32, 2000.
19. *Emili AQ and Cagney G:* Large-scale functional analysis using peptide or protein arrays. Nat. Biotechnol 18: 393–397, 2000.
20. *Ernst T, Hergenhahn M, Kenzelmann M et al:* Decrease and gain of gene expression are equally discriminatory markers for prostate carcinoma: a gene expression analysis on total and microdissected prostate tissue. Am J Pathol 160: 2169-2180, 2002.
21. *Esteller M, Corn PG, Baylin SB et al:* A gene hypermethylation profile of human cancer. Cancer Res 61: 3225–3229, 2001.
22. *Ferrando AA, Look AT:* Clinical implications of recurring chromosomal and associated molecular abnormalities in acute lymphoblastic leukemia. Semin Hematol 37: 381-95, 2000.
23. *Ferrando AA, Neuberg DS, Staunton J, et al:* Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. Cancer Cell, 1: 75-87, 2002.
24. *Furák JI, Troján T, Szõke L et al:* Development of Brain Metastasis 5 Years Before the Appearance of the Primary Lung Cancer: Messenger Metachronous Metastasis. Annals Thoracic Surg (in press).
25. *Gitan RS, Shi H, Chen CM et al:* Methylation-specific oligonucleotide microarray: a new potential for high-throughput methylation analysis. Genome Res 12: 158-164, 2001.
26. *Gilliland DG:* Molecular Genetics of Human Leukemia. Leukemia 12: S7-S12, 1998.
27. *Golub TR:* The Genetics of AML: An Update in Proceedings of the American Society of Hematology, pp. 102-111, 1999.
28. *Golub TR, Slonim DK, Tamayo P et al.:* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 286: 531-537, 1999.
29. *Haab BB Dunham MJ, Brown PO:* Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. Genome Biol 2: Research 0004.1–0004.13, 2001.
30. *Hanash SM, Madoz-Gurpide J, Misek DE:* Identification of novel targets for cancer therapy using expression proteomics. Leukemia 16: 478-485, 2002.
31. *Hedenfalk I, Duggan D, Chen Y et al:* Gene-expression profiles in hereditary breast cancer. N Engl J Med 344: 539-548, 2001.
32. *Hedrick SM, Cohen DI, Nielsen EA et al:* Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. Nature 308: 149-153, 1984.
33. *Hodgson G, Hager JH, Volik S et al:* Genome scanning with array CGH delineates regional alterations in mouse islet carcinomas. Nat Genet 29: 459-464, 2001.
34. *Houldsworth J, Chaganti RS:* Comparative genomic hybridization: an overview. Am J Pathol 145: 1253-1260, 1994.
35. *Huang TH-M, Perry MR and Laux DE:* Methylation profiling of CpG islands in human breast cancer cells. Hum Mol Genet 8: 459–470, 1999.
36. *Huang Y, Prasad M, Lemon WJ et al:* Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. Proc Natl Acad Sci USA 98: 15044-15049, 2001.
37. *Jiang Y, Harlocker SL, Molesh DA et al:* Discovery of differentially expressed genes in human breast cancer using subtracted cDNA libraries and cDNA microarrays. Oncogene 21: 2270-2282, 2002.
38. *Jones PA:* DNA methylation errors and cancer. Cancer Res 65: 2463–2467, 1996.
39. *Kanerva J, Niini T, Vettenranta K et al:* Loss at 12p detected by comparative genomic hybridization (CGH): association with TEL-AML1 fusion and favorable prognostic features in childhood acute lymphoblastic leukemia (ALL). A multi-institutional study. Med Pediatr Oncol 37: 419-425, 2001.
40. *Kitajka K, Puskás LG, Zvara Á et al:* The role of n-3 polyunsaturated fatty acids in brain: Modulation of rat brain gene expression by dietary n-3 fatty acids. Proc Natl Acad Sci USA 99, 2619-2624, 2002.
41. *Kopper L, Tímár J:* Gene expression profiles in the diagnosis and prognosis of cancer. Magy Onkol 46:3-9, 2002.
42. *Kozian DH, Kirschbaum BJ*: Comparative gene-expression analysis. Trends Biotechnol 17: 73-78, 1999.
43. *Kroll T, Odyvanova L, Clement JH et al:* Molecular characterization of breast cancer cell lines by expression profiling. J Cancer Res Clin Oncol 128: 125-134, 2002.
44. *Liang P, Pardee A et al:* Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science 257: 967-971, 1992.
45. *Lin YM, Furukawa Y, Tsunoda T et al:* Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. Oncogene 21: 4120-4128, 2002.
46. *Lisitsyn N, Wigler M et al:* Cloning the differences between two complex genomes. Science 259: 946-951, 1993.
47. *Luo JH, Yu YP, Cieply K et al:* Gene expression analysis of prostate cancers. Mol Carcinog 33: 25-35, 2002.
48. *Mitelman F, Mertens F, Johansson B:* A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. Nat Genet Spec No: 417-474, 1997.
49. *Novak U, Oppliger Leibundgut E, Hager J et a:* A high-resolution allelotype of B-cell chronic lymphocytic leukemia (B-CLL). Blood 100: 1787-1794, 2002.
50. *Pinkel D, Segraves R, Sudar D et al:* High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 20: 207-11, 1998.

51. *Prashar Y, Weissman S et al:* Analysis of differential gene expression by display of 3' end restriction fragments of cDNAs. Proc Nat Acad Sci USA 93: 659-663, 1996.
52. *Pui CH, Campana D, Evans WE:* Childhood acute lymphoblastic leukaemia current status and future perspectives. Lancet Oncol 10: 597-607, 2001.
53. *Puskás LG, Zvara Á, Hackler JrL et al:* RNA amplification results in reproducible microarray data with slight ratio biases. Biotechniques 32: 1330-1342, 2002.
54. *Puskás LG, Zvara Á, Hackler JrL et al:* Production of bulk amounts of universal RNA for DNA-microarrays. Biotechniques 33: 898-900, 902, 904, 2002.
55. *Ramaswamy S, Tamayo P, Rifkin R et al:* Multiclass cancer diagnosis using tumor gene expression signatures. Proc Natl Acad Sci USA 98: 15149-15154, 2001.
56. *Rubnitz JE and Pui C-H:* Leukemias, in *Principles of Molecular Medicine*, pp. 233-239. (J.L. Jameson ed.) Humana Press, NJ, 1998.
57. *Schlossman SF, Chess L, Humphreys RE et al:* Distribution of Ia-like molecules on the surface of normal and leukemic human cells. Proc Natl Acad Sci USA 73: 1288-1292, 1976.
58. *Schwarze SR, DePrimo SE, Grabert LM et al:* Novel pathways associated with bypassing cellular senescence in human prostate epithelial cells. J Biol Chem 277: 14877-14883, 2002.
59. *Shipp MA, Ross KN, Tamayo P et al:* Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med. 8: 68-74, 2002.
60. *Shou J, Soriano R, Hayward SW et al:* Expression profiling of a human cell line model of prostatic cancer reveals a direct involvement of interferon signaling in prostate tumor progression. Proc Natl Acad Sci USA 99: 2830-2835, 2002.
61. *Sidransky D:* Emerging molecular markers of cancer. Nature Rev Cancer 2: 210-219, 2002.
62. *Stremmel C, Wein A, Hohenberger W et al:* DNA microarrays: a new diagnostic tool and its implications in colorectal cancer. Int J Colorectal Dis 17: 131-136, 2002.
63. *Suzuki H, Gabrielson E, Chen W et al:* A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. Nat Genet 31: 141-149, 2002.
64. *Swaroop A, Xu JZ, Agarwal N et al.:* A simple and efficient cDNA library subtraction procedure: isolation of human retina-specific cDNA clones. Nucleic Acids Res 19: 1954, 1991.
65. *Sweetser DA, Chen CS, Blomberg AA et al:* Loss of heterozygosity in childhood de novo acute myelogenous leukemia. Blood 98: 1188-94, 2001.
66. *Takano T, Hasegawa Y, Matsuzuka F et al:* Gene expression profiles in thyroid carcinomas Br J Cancer 83: 1495-502, 2000.
67. *Toyota M, Ho C, Ahuja N et al:* Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification. Cancer Res 59: 2307–2312, 1999.
68. *Tsukasaki K, Krebs J, Nagai K et al:* Comparative genomic hybridization analysis in adult T-cell leukemia/lymphoma: correlation with clinical course. Blood 97: 3875-3881, 2001.
69. *Yeoh EJ, Ross ME, Shurtleff SAA et al:* Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer Cell 1: 133-143, 2002.
70. *Velculescu VE, Zhang L, Vogelstein B et al:* Serial analysis of gene expression. Science 270: 464-467, 1995.
71. *Vogeli-Lange R, Burckert N, Boiler T et al:* Rapid selectionand classification of positive clones generated by mRNA differential display. Nucleic Acids Res 24: 1385-1386, 1996.
72. *Weinberger SR, Morris TS, Pawlak M:* Recent trends in protein biochip technology. Pharmacogenomics 1: 395-416, 2000.
73. *Wieser R:* Rearrangements of chromosome band 3q21 in myeloid leukemia. Leuk Lymphoma 43: 59-65, 2002.
74. *Wilhelm M, Veltman JA, Olshen AB et al:* Array-based comparative genomic hybridization for the differential diagnosis of renal cell cancer. Cancer Res 62: 957-960, 2002.
75. *de Wit NJ, Burtscher HJ, Weidle UH et al:* Differentially expressed genes identified in human melanoma cell lines with different metastatic behaviour using high density oligonucleotide arrays. Melanoma Res 12: 57-69, 2002.
76. *Wreesmann VB, Ghossein RA, Patel SG et al:* Genome-wide appraisal of thyroid cancer progression. Am J Pathol 161: 1549-1556, 2002.