Peer reviewed article

Functional genomics and gene microarrays – the use in research and clinical medicine

Ladina Joos, Emel Eryüksel, Martin H. Brutsche

Division of Respiratory Medicine and Pulmonary Research Laboratory, University Hospitals, Basel, Switzerland

Summary

In the year 2000, the Human Genome Project Consortium presented the first complete draft of the human genome together with Celera Genetics. Since then, the so-called "post-genome era" has started. Microarrays are capable of profiling gene expression patterns of tens of thousands of genes in a single experiment and thus allow a systematic analysis of DNA and RNA variation. They seem likely to become a standard tool of both molecular biology research and clinical diagnostics. These prospects have attracted great interest and investment from both the public and private sectors. This review introduces the principle of microarray technology and gives an overview of its current and future potential in clinical medicine.

Key words: genomics; genetics; pharmacogenomics; pharmacogenetics; microarrays

Introduction

Why analyse human biology on a genomic scale – Homage to biological complexity

In the year 2000, the first draft of the human genome was presented [1, 2] and made the headlines worldwide. However, the current version of the human genome, consisting of 3×10^9 nucleotides, is only a rough draft, and the full reliable version of the human genome is not expected before 2004. It is presumed to contain approximately 50'000 genes.

Each gene can be present in different variants, called polymorphisms (mostly single nucleotide polymorphisms SNPs). To date, 3×10^6 gene polymorphisms are described – a number which increases daily and is estimated to be greater than 11×10^6 for the whole human genome. Each individual is composed of a pair of inherited combinations of variants for each of the 50'000 genes explaining the enormous genetic diversity. Only a fraction of these gene polymorphisms is important to human health and disease, and to find the important ones is a major challenge for the next decades.

Not all of these genes are used at a time. Depending on the developmental stage, age of the individual, cell type, organ, and environmental factors, a different set of genes is used or transcribed. Genes are transcribed into messenger RNA (mRNA). Analysis of the expression of mRNA under defined conditions is called functional genomics (table 1). It allows comparison of different sets of genes used in different conditions, eg, healthy and diseased. Most studies applying microarrays are functional genomic studies.

Not all of the transcribed genes will result in a protein. Also, practically all proteins are modified after the first assembly of amino acids. It is estimated that a protein derived from the same gene strand can be altered in 10–20 different splits and 3-dimensional forms. Some proteins interact directly with the DNA, leading to expression or silencing of genes. Practically all of the proteins interact with other proteins within pathways to form complex multidimensional related networks. The analysis of the proteins is called proteomics.

The long way from genotype to phenotype

If the genotype would automatically lead to a specific condition, so-called phenotype, all identical twins would have exactly the same diseases. Although they considerably resemble each other, they also differ in many ways. Thus, there is not a 100% match between genotype and phenotype due to environment-gene interactions. Also, many conditions result from a variety of pathogenetic mechanisms, thus, despite a different genetic makeup, the same phenotype arises, such as asthma or hypertension. In these, so-called complex diseases, a constellation of different susceptibility and disease-modifying genes need to be present. Research has therefore failed to identify specific disease genes, eg, an "asthma gene". To better understand the functional aspects of disease and to

No financial support declared.

Table 1
Definitions.

Term	definition
Genetics	analysis of the genome structure and its variations
Functional genomics	analysis of gene expression of a cell, tissue or organ under given conditions
Proteomics	analysis of protein molecules of a cell, tissue or organ under given conditions
Pharmacogenetics	study of variability in drug responses attributed to hereditary factors in different populations
Pharmacogenomics	determination and analysis of the genome and its products (RNA and proteins) as they relate to drug response

Table 2
Levels of gene – function analysis.

Level of analysis	definition	status	method of analysis
Genome	complete set of genes of an organism	context-independent	high-throughput sequencing
Transcriptome	complete set of mRNA molecules present in a cell, tissue or organ	context-dependent (varies with changes in physiology, development or pathology)	cDNA or oligonucleotide microarrays Serial analysis of gene expression (SAGE) Northern analysis RT-PCR
Proteome	complete set of protein molecules present in a cell, tissue or organ	context-dependent	protein arrays Two-hybrid analysis 2-D gel electrophoresis Peptide mass fingerprinting

bridge the long way between genotype and phenotype, it is necessary to combine genetic, functional genomic and proteomic analyses (table 2).

Hypothesis-driven vs. hypothesis-generating research

Classical hypothesis-driven research, often analysing a single or several genes or proteins, was and is a very successful and reliable scientific strategy. However, classical research is not able to cope with the number of newly discovered genes, proteins and potential interactions. In the light of the complexity of genes, gene expression, proteins, functional networks and pathways, a conclusion must be reached: *for complex questions, eg, what are the causes of, and factors influencing a specific disease, there might be complex answers.* Thus, it is necessary to apply techniques which permit complex answers. Therefore, novel techniques are necessary to screen thousands of genes more rapidly and generate new hypotheses. This is the role of high-throughput technology, like microarrays. As a hypothesis-generating approach, high-throughput methods can lead to the identification of a set of potentially interesting genes associated with a certain condition, so called candidate genes. *Microarray techniques, however, will not replace the classical hypothesis-driven research.* Identified candidate genes need to be tested for their function and relevance by classical approaches.

Technicalities of microarrays

What is a microarray and how does it work?

A microarray – or "gene chip" – measures the expression level of a gene by determining the amount of messenger RNA that is present (mRNA abundance). The company Affymetrix® owns a registered trademark, GeneChip®, which refers to its high density, oligonucleotide-based DNA arrays. However, in some articles appearing in professional journals, popular magazines, and on the world wide web, the term "gene chip(s)" has been applied generally referring to microarray technology. Unlike a conventional Northern blot which analyses one, two, or up to 20 mRNAs, a microarray allows the simultaneous analysis of the expression levels of hundreds, thousands, or even tens of thousands of genes in a single experiment. The latest chips carry up to 450'000 spots for the analysis of more than 20'000 genes and gene sequences on a small glass slide (1.2 × 1.2 cm, figure 1).

Production of arrays begins with the selection of the *probes* to be printed on the array. These are often chosen directly from gene databases (eg, GenBank, http://www.ncbi.nlm.nih.gov). Two main methods are used to generate microarrays. In the first – cDNA array technology – DNA probes, ie, PCR products representing the complementary DNA (cDNA) code of specific genes, are spotted onto a glass slide. In the second – oligonucleotide array technology – oligonucleotides consisting of nucleic acids are synthesised on to a silica slide by a process known as photolithography. The *target*, ie, purified mRNA from the biological sample (eg,

blood or tissue) is labelled by fluorescence or radioactivity and then hybridised for several hours to the microarray. mRNA consists of a sequence built up of 4 different oligonucleotides (thymidine, guanine, cytosine, and uracil), which is specific for each gene. Such a specific mRNA sequence is able to bind to a single complementary sequence of oligonucleotides only. This specificity of binding is used in microarrays. If the sequence of oligonucleotides matches with the sequence of oligonucleotides on a specific spot of the microarray, hybridisation occurs. On each of the up to 450'000 spots, different binding intensities occur depending on the concentration of the different genes in the biological sample tested. Thus, the concentration of mRNA (= gene expression) can be measured quantitatively [3].

A different strategy is used to select probes for genotyping arrays. These arrays rely on multiple probes to detect individual nucleotides in a sequence. The identity of a target base can be deduced using four identical probes that vary only in the target position, each containing one of the four possible bases. Alternatively, the presence of a consensus sequence can be tested using one or two probes representing specific alleles. To genotype heterozygous or genetically mixed samples, arrays with many probes can be created to provide redundant information, resulting in unequivocal genotyping (http://www.affymetrix.com).

What can be measured by microarrays

Microarray technology can be used for three main applications:

1. *Gene expression profiling* – mRNA extracted from a biological sample is applied to the microarray. The result reveals the level of expression of tens of thousands of genes in that sample. This result is known as a gene expression "profile" or "signature" [4].

2. *Genotyping* – DNA, extracted from a biological sample, is amplified by a polymerase chain reaction and applied to the microarray. The genotype for hundreds or thousands of genetic markers across the genome can be determined in a single experiment. This approach has considerable potential in disease risk assessment, both in research and clinical practice [5].

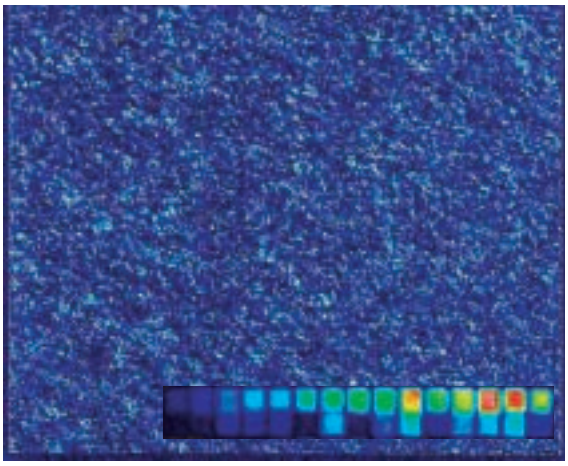
3. *DNA sequencing* – DNA extracted from a biological sample is amplified and applied to specific "sequencing" microarrays. Thousands of base pairs of DNA can be screened on a single microarray for polymorphisms in specific genes whose sequence is already known. This greatly increases the scope for precise molecular diagnosis in single gene and genetically complex diseases [5, 6].

Microarray data analysis

The analysis of microarray data is complex and involves several steps. After hybridisation, microarrays are scanned and images representing the intensity of the fluorescence signal are generated. After image processing, it is necessary to normalise the fluorescence intensities. The normalisation is done for each microarray. Typically, the signal for each gene is divided by the median gene signal. This process is called per chip normalisation. Often, a second normalisation, called per gene-normalisation, is applied, where the signals for a specific gene throughout the different microarrays are divided by the median gene signal. Such normalised signal intensities of different microarrays, representing different conditions, can be compared. Genes, for which the mRNA is over-represented, or under-represented, are called upregulated, or downregulated, respectively. Most published studies have used a post-normalisation cut-off of two-fold increase or decrease in measured level to define differential expression. Although it is generally accepted that one can not rely on a single gene chip experiment, there is no clear consensus about the number of experimental replicates that are needed for robust results. And, indeed, the answer would differ from gene to gene, as each gene has a different gene expression variability and expression range. Clearly, the trend is "the more the better". Apart from identifying genes by the arbitrary two-fold increase or decrease, the expression variability of the genes can

Figure 1

Example of a hybridised GeneChip® with enlarged probe set of a gene as an insert bottom right. With this chip, 12'626 genes and expressed sequence tags can be assayed in a single experiment. Messenger RNA is copied into labelled cRNA with reverse transcriptase so that the relative abundance of individual mRNAs is reflected in the cRNA product. Thus, the intensity of the hybridisation signal for a given gene product is a result of its relative abundance in the target sample. This method has proven to provide excellent specificity and reproducibility. Messenger RNA species comprising 1:10'000–100'000 of the mass of the target poly(A)+RNA, which corresponds to approximately 1 transcript per 100'000, can readily be detected. The intensity of the hybridisation signal (red = high intensity, dark blue = no signal) for a given gene is a result of its relative abundance in the RNA-derived DNA probe.



be compared using classical statistical significance testing. Before interpreting the data, p-values need to be corrected with specific algorithms (eg, Bonferroni corrections, Benjamini and Hochberg false discovery rate) due to multiple comparisons. However, the true power of microarray analysis is to identify common patterns of gene expression associated with a specific experimental condition. There exists a large group of statistical methods for pattern recognition, such as principal component analysis, cluster analysis, and intelligent maps [7].

Another important component of expression data exploration are powerful data visualisation methods and tools. Such visualisation techniques, combined with integrated links to annotated sequence databases, provide very valuable tools that allow biologists to examine large expression data sets and develop new insights into and models of genome-wide transcriptional regulation [8].

Clinical application of microarrays

In the clinical application, microarrays have strengths in four areas:

1. It is possible to identify individuals at risk for certain diseases by looking for *disease susceptibility* genes. Such patients can be included in specific disease prevention programs.

2. Microarrays can help to establish the *correct diagnosis efficiently and early* in the disease process.

3. Microarrays can be used to measure *reliable prognostic markers and gene expression scores.*

4. It is possible – in the future – to apply *an individualised treatment* according to the patient's gene expression profile. Therefore, the treatment with an optimised effect to side effect potential in each patient will be chosen on an individual basis.

The following section gives a rough overview of microarray studies with clinical relevance in different medical disciplines.

Oncology

Regulation of the cell division cycle is crucial to most biological processes including gametogenesis and wound healing. Processes involved in cell division are characteristically aberrant in cancer. These processes and their regulation have been extensively studied at the molecular level in primary human fibroblasts [9] and human HeLa cells [10] using cDNA microarrays.

Acute leukaemia is an example of a successful diagnostic approach using microarray technology. Independent of any histological or histochemical diagnosis, the analysis of expression profiles of 6'817 genes was used to distinguish between acute lymphoid leukaemia and acute myeloid leukaemia. The results correctly classified 36 out of 38 "unknown" leukaemia samples derived from either bone marrow or peripheral blood [11]. The technique successfully divided acute lymphoid leukaemia into T-cell derived or B-cell derived leukaemia and raised new and intriguing insights into the role of the 50 most predictive genes in disease pathogenesis. However, the study was not able to allow the definition of new prognostic groups.

In contrast, a similar study of diffuse large B-cell lymphoma (DLBCL) detected a prognosis-relevant pattern of gene expression by microarray profiling [12]. DLBCL is the most common subtype of non-Hodgkin's lymphoma and its prognosis is notoriously difficult to classify on morphological or clinical grounds. It is therefore of considerable clinical relevance that gene expression analysis defined two new prognostic subgroups in diffuse large B cell lymphoma, with a five year survival of 80% and 40%, respectively. Similar studies were performed in metastatic and non-metastatic melanoma [13], breast cancer [14], and colorectal cancer [15].

It seems likely that major new insights will be derived for a wide range of cancers, leading to the prospect of better targeted treatment and, in the longer term, new treatments based on improved understanding of the molecular pathogenesis of these diseases.

Infectious diseases

Microarrays that detect gene sequences in the genomes of *Mycobacterium tuberculosis*, HIV, and other pathogens have been developed to provide a diagnostic tool that detects expression of antibiotic resistance genes or specifies viral subtypes [16]. A major advantage is that these tests can be undertaken in less than 24 hours without the need for bacterial or viral cultures. If such tests are brought into clinical practice, they will lead to earlier, more targeted treatment based on antibiotic or antiviral sensitivities: Microarrays will be particularly valuable for organisms such as *Mycobacterium tuberculosis* and HIV, for which sensitivity profiles can presently be determined only after lengthy analysis by other methods.

Lung diseases

Kaminski et al. [17] investigated bleomycin-induced pulmonary fibrosis in mice. While bleomycin induces lung inflammation and fibrosis in wild-type mice, mice deficient in epithelium-restricted integrin beta 6-subunit (beta $6^{-/-}$) develop exaggerated inflammation but are protected from pulmonary fibrosis. Comparative analysis of gene expression profiles during bleomycin-induced pulmonary disease in wild-type and beta $6^{-/-}$ mice distinguishes gene clusters involved in the inflammatory response from gene clusters that mediate

specifically fibrotic responses. Luzina et al. [18] analysed gene expression in bronchoalveolar lavage cells from scleroderma patients with and without interstitial lung disease and found indications for T-cell recruitment and macrophage activation in scleroderma patients at greater risk for lung fibrosis. Own gene chip studies found a phenotype-specific reduction in apoptosis signals in atopy, asthma [4] and sarcoidosis [19] as compared to healthy controls in vivo. In another study in patients with sarcoidosis, the expression of specific growth factor-related genes was associated with progressive disease and therefore predicted outcome [20].

We designed an array-based composite gene expression score (CAGE) to quantify atopy and asthma [21]. So far, an individual is considered as either being atopic or not. However, as experienced in clinical practice, some patients seem to be more allergic then others. In this situation, the composite atopy gene expression score is helpful to decide upon the best therapy. The CAGE score was better than total IgE in differentiating atopic from non-atopic subjects (sensitivity 96%, specificity 92%; figure 4). A further in vivo-study investigated the differences in B-cell isotype control mechanisms in atopy and asthma compared to healthy control subjects [22].

The pathological distinction between malignant pleural mesothelioma and adenocarcinoma of the lung can be cumbersome using established methods. Gordon et al. used a simple technique, based on the expression levels of a small number of genes, for an early and accurate diagnosis of mesothelioma and lung cancer. They found that the differential diagnoses of mesothelioma and

Figure 2

Scatter plot of the mean gene expression in sarcoidosis patients compared to controls. Colour-codes are given for the p-value (Mann-Whitney U-test between phenotypes). Gene expression in both phenotypes is very similar for most of the genes tested resulting in a R-square of 0.98. Genes with different expression are likely to be involved in the disease process. Therefore, the statistical analysis aims to identify genes with phenotype-specific alterations in gene expression within the relatively small differences in gene expression observed.

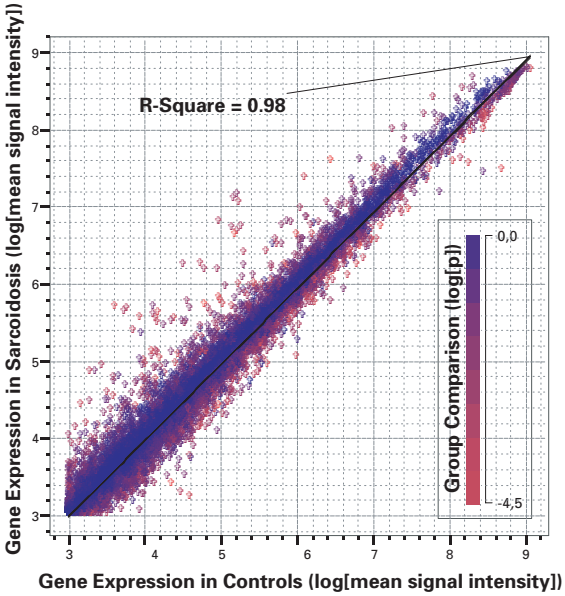


Figure 3

In a double-blind controlled trial 16 healthy individuals were randomised (3:1) to either placebo or 0.5 mg/kg of body weight oral prednisone daily for 14 days. At baseline, and on days 1, 7, and 14, broad-spectrum gene expression (GeneChip® HGU 95A, Affymetrix Inc.) in purified blood mononuclear cells was evaluated. Of the 12'605 genes measured, 3'520 or 28% were expressed and 450 of these or 13% were significantly dysregulated 8 h after first prednisone dose. The figure shows the significantly up- (red) and downregulated (green) genes. We could identify acute and long-term steroid-responsive genes in different metabolic pathways, which might help to characterise differences in the individual susceptibility to steroid-related side-effects [34].

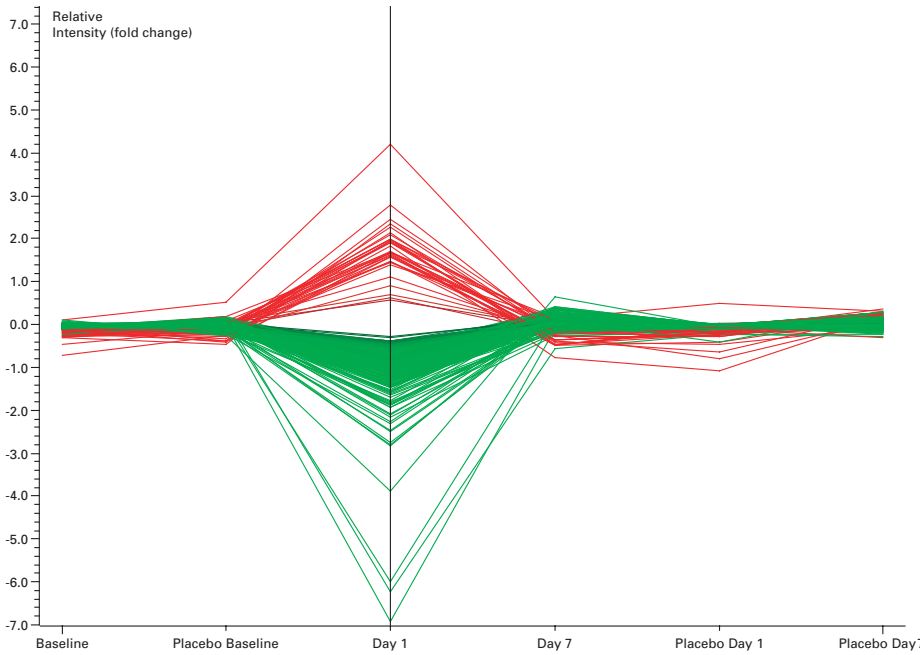
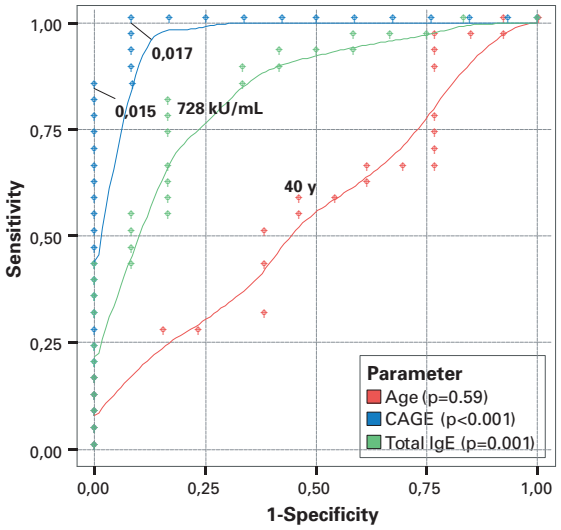


Figure 4

Receiver-operating characteristic (ROC) analyses to compare the composite atopy gene expression (CAGE) score, total IgE and age for diagnosing atopy. The CAGE score is a microarray-based diagnostic tool which allows to quantify atopy. It is composed of 10 disease-relevant genes which are differentially expressed in atopy and asthma. The CAGE score (area = 0.96 [0.90–1.00]; p <0.001) was better in differentiating atopic from non-atopic individuals than total IgE (area = 0.85 [0.73–0.98]; p = 0.001). The ROC curve for age fluctuated around the diagonal and was not significant (area = 0.50 [0.30–0.71]; p = 0.97). The dots show the relationship between sensitivity and specificity at the respective cut-off values. A smoothing iteration was performed to show continuous ROC curves [21].



adenocarcinoma were 95% and 99% accurate, respectively [23].

Histopathology is insufficient to predict disease progression and clinical outcome in lung adenocarcinoma. Beer et al. identified a set of genes that predicted survival in early-stage lung adenocarcinoma [24].

Cardiology

Microarray technology was used to evaluate the expression of >4'000 genes in a rat model of myocardial infarction [25]. More than 200 genes were identified that showed differential expression in response to myocardial infarction. Unique patterns were revealed within the transcriptional responses that illuminate changes in biological processes associated with myocardial infarction.

Recent studies suggest that the treatment of dyslipidemia and the prevention of coronary artery disease should be genotype-specific, as the genetic make-up can determine the outcome of a pharmacological intervention [26]. Several genetic risk factors for the development of coronary artery disease have been identified including elevated levels of lipoprotein (a), the DD genotype of the angiotensin converting enzyme and elevated levels of homocysteine. Furthermore, clinically validated genotype analysis for dyslipidemia is available and in some cases recommended in case of familial dysbetalipoproteinemia [27], hypo-alphalipoproteinemia [28], Tangier' disease [29], lecithin-cholesterol acyltransferase (LCAT) deficiency [30], Fish Eye disease [30] and elevated triglycerides [31]. The presence or absence of the Taq1B polymorphism in the cholesteryl ester transfer protein (CETP) gene can predict response to statin treatment [32]. Using microarray technology, all genes can be tested at once to establish a comprehensive risk profile and an individualised intervention strategy.

Pharmacogenomics by microarrays

Microarrays can be used to investigate the pharmacodynamic effects of drugs at a genomic level in order to predict effects and side effects of drugs in a dose- and/or time-dependent fashion. Drugs, even with a known mechanism of action, often have unknown collateral effects, which become apparent through comprehensive microarray studies. In the future, analysis of a patient's sample should permit prediction of the therapeutic ratio of a particular treatment before the drug has been applied. Thus, the pharmacological treatment could be chosen based on the genetic background of a specific patient. This approach is called *individualised or tailored treatment.*

Microarrays have been used to study the effect of psychoactive agents like delta 9-tetrahydrocannabinol, the primary psychoactive component of marijuana, in the rat brain [33]. Under high-stringency conditions, several differentially expressed genes were detected, including those involved in cannabinoid synthesis and receptor-effector systems.

In a pharmacogenomic study, we investigated the effects of a 14-days course of oral prednisone therapy using Affymetrix GeneChips® (12'626 genes) (figure 3). We identified acute and long-term steroid-responsive genes in different relevant metabolic pathways. These findings will allow screening of patients on long-term oral corticosteroid treatment at high risk for relevant side effects in future [34].

Outlook – Medicine in 2010

It is likely that some *disease phenotypes will be categorised differently* to better match with genotypes and specific gene expression signatures. This will be particularly useful in clinical situations where different therapeutic strategies apply. Such diseases could include malignant diseases, chronic multi-organ disorders, chronic obstructive lung disease, heart failure, and many others – diseases with different pathogenetic mechanisms but similar phenotypic presentation. Along these lines, *microarrays could be used to identify individuals at risk for certain conditions*, *to establish the exact and early diagnosis*, *to establish reliable prognosis and to give guidance for therapy*. It is possible that the need for some classic diagnostic procedures will be reduced. The increased clinical information provided by microarrays should assure their *entry into routine clinical practice within the next three to five years*, although the added costs will have to be justified by the clinical benefit.

Due to the complexity of the matter, microarray-facilitated medicine will first happen in specialised centres before being introduced broadly. *Physicians need to be trained in molecular biology* for a successful introduction of microarrays in clinical medicine.

Conclusion

Functional genomics will undoubtedly help to improve screening, early detection/diagnosis, prognostic markers and individualised treatment strategies. Most microarray-based tests are still in the developmental stage, although substantial progress towards commercialisation has occurred in some cases. Like any new diagnostic tool, microarrays will have to be rigorously appraised for sensitivity, specificity, and predictive value. The high costs of microarray-based tests will inevitably limit the speed of introduction into clinical practice and initially restrict their use to specialised centres. However, given the huge potential gain in clinically relevant information for individual patients and their diseases, the technology is likely to reach most large hospitals within the next 10 years.

Correspondence
Martin H. Brutsche, MD, PhD
University Hospital Basel
Petersgraben 4
CH-4031 Basel
E-Mail: mbrutsche@uhbs.ch

References

1 Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860–921.
2 Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, et al. The sequence of the human genome. Science 2001;291:1304–51.
3 Duggan DJ, Bittner M, Chen Y, Meltzer P, Trent JM. Expression profiling using cDNA microarrays. Nat Genet 1999;21:10–4.
4 Brutsche MH, Brutsche IC, Wood P, Brass A, Morrison N, Rattay M, et al. Apoptosis signals in atopy and asthma measured with cDNA arrays. Clin Exp Immunol 2001;123:181–7.
5 Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 1998;280:1077–82.
6 Takahashi Y, Ishii Y, Nagata T, Ikarashi M, Ishikawa K, Asai S. Clinical Application of Oligonucleotide Probe Array for Full-Length Gene Sequencing of TP53 in Colon Cancer. Oncology 2003;64:54–60.
7 Quackenbush J. Computational analysis of microarray data. Nat Rev Genet 2001;2:418–27.
8 Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.
9 Cho RJ, Huang M, Campbell MJ, Dong H, Steinmetz L, Sapinoso L, et al. Transcriptional regulation and function during the human cell cycle. Nat Genet 2001;27:48–54.
10 Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Mol Biol Cell 2002;13:1977–2000.
11 Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531–7.
12 Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000;403:503–11.
13 Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. Nature 2000;406:532–5.
14 Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature 2000;406:747–52.
15 St Croix B, Rago C, Velculescu V, Traverso G, Romans KE, Montgomery E, et al. Genes expressed in human tumor endothelium. Science 2000;289:1197–202.
16 Kozal MJ, Shah N, Shen N, Yang R, Fucini R, Merigan TC, et al. Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. Nat Med 1996;2:753–9.
17 Kaminski N, Allard JD, Pittet JF, Zuo F, Griffiths MJ, Morris D, et al. Global analysis of gene expression in pulmonary fibrosis reveals distinct programs regulating lung inflammation and fibrosis. Proc Natl Acad Sci U S A 2000;97:1778–83.
18 Luzina IG, Atamas SP, Wise R, Wigley FM, Xiao HQ, White B. Gene expression in bronchoalveolar lavage cells from scleroderma patients. Am J Respir Cell Mol Biol 2002;26:549–57.
19 Rutherford RM, Kehren J, Staedtler F, Chibout SD, Egan JJ, Tamm M, et al. Functional genomics in sarcoidosis—reduced or increased apoptosis? Swiss Med Wkly 2001;131:459–70.

20 Eryuksel E, Rutherford R, Bihl M, Joos L, Kehren J, Staedtler F, et al. Specific pattern of growth factor gene expression in stage I versus stage II/III sarcoidosis. Eur Respir J 2002:A 1414.

21 Brutsche MH, Joos L, Carlen Brutsche IE, Bissinger R, Tamm M, Custovic A, et al. Array-based diagnostic gene-expression score for atopy and asthma. J Allergy Clin Immunol 2002;109:271–3.

22 Brutsche MH, Brutsche IC, Wood P, Mogulkoc N, Custovic A, Egan J, et al. B-cell isotype control in atopy and asthma assessed with cDNA array technology. Am J Physiol Lung Cell Mol Physiol 2001;280:L627–37.

23 Gordon GJ, Jensen RV, Hsiao LL, Gullans SR, Blumenstock JE, Ramaswamy S, et al. Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. Cancer Res 2002;62:4963–7.

24 Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 2002;8:816–24.

25 Stanton LW, Garrard LJ, Damm D, Garrick BL, Lam A, Kapoun AM, et al. Altered patterns of gene expression in response to myocardial infarction. Circ Res 2000;86:939–45.

26 Doevendans PA, Jukema W, Spiering W, Defesche JC, Kastelein JJ. Molecular genetics and gene expression in atherosclerosis. Int J Cardiol 2001;80:161–72.

27 de Knijff P, van den Maagdenberg AM, Boomsma DI, Stalenhoef AF, Smelt AH, Kastelein JJ, et al. Variable expression of familial dysbetalipoproteinemia in apolipoprotein E*2 (Lys146–>Gln) Allele carriers. J Clin Invest 1994;94:1252–62.

28 Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K, van Dam M, et al. Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. Nat Genet 1999;22:336–45.

29 Rust S, Rosier M, Funke H, Real J, Amoura Z, Piette JC, et al. Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. Nat Genet 1999;22:352–5.

30 Kuivenhoven JA, van Voorst tot Voorst EJ, Wiebusch H, Marcovina SM, Funke H, Assmann G, et al. A unique genetic and biochemical presentation of fish-eye disease. J Clin Invest 1995;96:2783–91.

31 Jukema JW, van Boven AJ, Groenemeijer B, Zwinderman AH, Reiber JH, Bruschke AV, et al. The Asp9 Asn mutation in the lipoprotein lipase gene is associated with increased progression of coronary atherosclerosis. REGRESS Study Group, Interuniversity Cardiology Institute, Utrecht, The Netherlands. Regression Growth Evaluation Statin Study. Circulation 1996;94:1913–8.

32 Kuivenhoven JA, Jukema JW, Zwinderman AH, de Knijff P, McPherson R, Bruschke AV, et al. The role of a common variant of the cholesteryl ester transfer protein gene in the progression of coronary atherosclerosis. The Regression Growth Evaluation Statin Study Group. N Engl J Med 1998;338:86–93.

33 Kittler JT, Grigorenko EV, Clayton C, Zhuang SY, Bundey SC, Trower MM, et al. Large-scale analysis of gene expression changes during acute and chronic exposure to [Delta]9–THC in rats. Physiol Genomics 2000;3:175–85.

34 Brutsche MH, Herrmann M, Huber P, Kehren J, Staedtler F, Tamm M, et al. Effects of intake and withdrawal of a 14–day corse of systemic prednisone on adreno-cortical function, exercise capacity, and gene expression. Am J Respir Crit Care Med 2002:A 768.