SMP: RNA

Project: Core Unit DKFZ: High throughput Gene Expression Analysis

Annemarie Poustka - **German Cancer Research Center (DKFZ), Heidelberg – a.poustka@dkfz.de**

Introduction

The analysis of gene expression patterns has been demonstrated to be one of the most powerful and direct ways to systematically generate data for functional studies, including the understanding of gene function in development and differentiation. Within less than 10 years, gene expression profiling with microarrays has become a key technology in biomedical research. Gene expression analyses of large numbers of patient samples have been providing significant advances to the classification of diseases and the identification of target genes relevant to human disease for therapeutic, diagnostic, and prognostic applications. In addition, microarray data have made important contributions to our understanding of the regulation of biological networks and gene function. This knowledge is fundamental to the elucidation of the entire network of changes causing and caused by a disease, and an absolute requirement for the knowledge-driven medicine of the future.
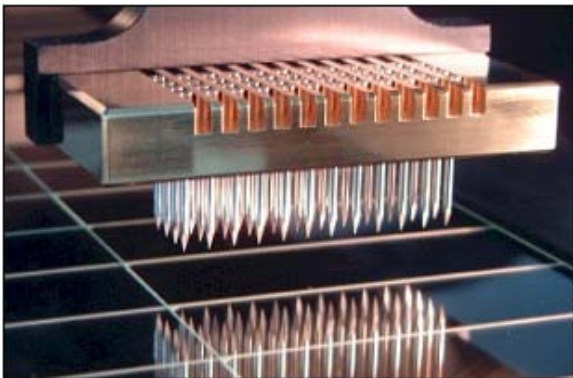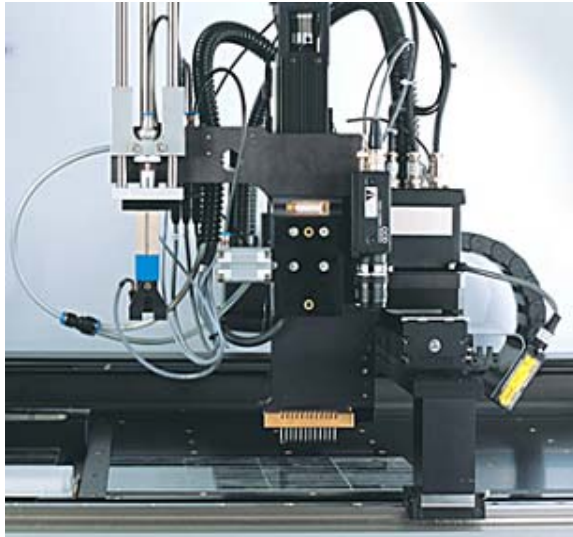

Fig 1: Spotter and spotting printhead for the production of microarrays in the core unit at the DKFZ.

Results/Project Status

The gene expression profiling platform at the DKFZ is focusing on the development of technologies for high throughput gene expression analyses and their application to fight human diseases. The common goal of all activities is the elucidation of disease pathomechanisms. The core unit performs collaborative projects with the Disease-oriented Genome Networks (KGs) and other Systematic Methodological Platforms (SMPs) as well as separate research projects. The strengths of the platform have been documented in numerous scientific collaborations and publications (see Lit.).

Fig 2: False colour image of a microarray hybridisation (section of an array)

On the platform, the experimental protocols for standardized array production (handling of large clone collections, PCR amplification of clone inserts by robotic pipetting systems, spotting of microarrays in a controlled environment; Figure 1) sample treatment (isolation, amplification and labeling), internal controls (negative controls, spiking with *Arabidopsis thaliana* reference RNA samples), as well as automated hybridization routines are established. From each microarray batch, 10% of the microarrays are routinely scanned after spotting. All slides are checked for remaining background before hybridization. DNA on two selected microarrays from each batch is stained with fluorescent DNA-binding dyes (Syto 61) to control for the presence of target. All relevant protocols are published and accessible for the scientific community on the NGFN website (www.ngfn.de). We routinely apply human and mouse whole genome microarrays for microarray analyses in the NGFN. Specifically, our arrays include:

Human whole genome cDNA microarrays (36,000 fully resequenced and bioinformatically characterized human genes/ESTs; RZPD Unigene 3.1). This clone set is currently the most comprehensive cDNA resource for microarray

construction. It includes the complete collection of the currently 20,000 ENSEMBL-annotated cDNA clones and approx. 16,000 ESTs. It has been distributed to multiple users, among them partners in the EU.

Human whole genome 70mer oligonucleotide microarrays (35,000 genes; Operon).

Mouse whole genome cDNA microarrays (NIA 23,000 and LION 20,000 clones).

Apart from microarray analyses, the core unit is increasingly working on the validation of gene expression data (e.g. by quantitative RT-PCR). New technologies (e. g. RNA and signal amplification) are constantly developed and implemented into the laboratory routines. Most importantly, the integration of strong bioinformatic know-how working in immediate collaboration with the experimental projects has proven to be highly successful in the past and is being continued in SMP-RNA.

Many collaborative projects with clinical partners have led to a wealth of information on gene expression in various tumors and other diseases. An example is mentioned below:

Diagnostic gene expression profiles in kidney cancer

Adult renal cell carcinoma (RCC) is one of the ten most common human malignancies in developed countries. Its global incidence has been increasing continuously during the past 30 years. RCC is divided into clear cell (ccRCC; 80 % of all cases), papillary (pRCC, 10 %), chromophobe (chRCC; 5 %), and several other rare types. Although the histopathological diagnosis of kidney cancer is well established in the clinical routine, the molecular basis for the distinction of RCC types is poorly understood.

We constructed RCC specific cDNA microarrays and hybridized these with labeled cDNA derived from tumor samples of the three major RCC types. By using the microarray data of 35 RCC samples, we identified an 18-gene-signature for the diagnosis of kidney tumor types (Figure 3). The genes in this set may also be important for a differential antioncogenic therapy.
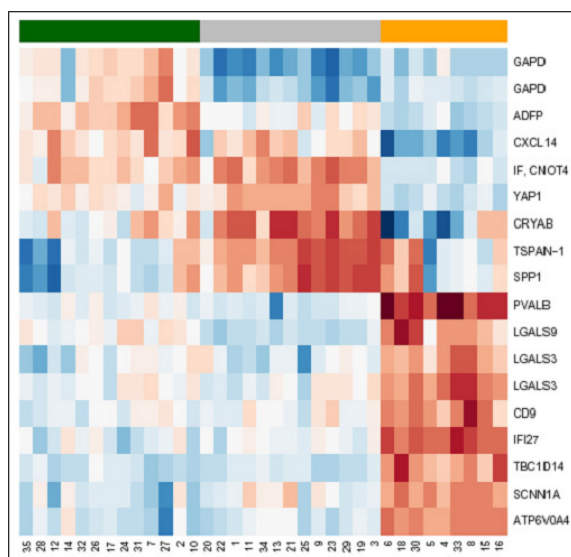

Fig. 3: *Diagnostic gene expression profile in kidney cancer. Three different kidney tumor types (green, grey, orange) can be distinguished by the expression pattern of only 18 genes.*

Outlook

During the past years, the core unit at the DKFZ has generated a large amount of data on differentially expressed genes in human diseases. The challenge for the future is to use these data for the translation into the clinical setting. To this end, the gene expression signatures are screened for novel diagnostic and prognostic applications. At the same time, in collaboration with partners (e. g. SMP-Cell), our gene expression data give rise to high throughput cellular functional analyses and proteomics assays in order to understand disease pathomechanisms. Furthermore, the technology is used for the identification of novel gene expression networks (see "Systematic Analysis of Functional Networks and Key Mediators"). Currently, biomedical research is at the stage of using microarrays as a routine tool and from this will give a considerable impact on the generation of hypotheses concerning disease pathomechanisms.

Lit.: **1.** *Fritz B et al. Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. Cancer Res 2002 62: 2993-2998.* **2.** *Huber W et al. Transcription Profiling of Renal Cell Carcinoma. Proceedings of the German Society for Pathology 2002 86: 153-164.* **3.** *Huber W et al. Variance stabilisation applied to Microarray Data Calibration and to the Quantification of Differential Expression. Bioinformatic 2002 Suppl 1: 96-104.* **4.** *Kokocinski F et al. QuickLIMS: facilitating the data management for DNA microarray fabrication. Bioinformatics 2002 19: 283-284.* **5.** *Korzhunov A et al. Gene expression patterns in ependymomas are correlated to tumor location and patient age. Am J Pathol 2004 163:1721-1727.* **6.** *Schlingemann et al., Profile of gene expression induced by the tumor promotor TPA in murine epithelial cells. Int J Cancer 104: 699-708.* **7.** *Schneider J et al., Systematic analysis of T7 RNA polymerase based in vitro RNA amplification for use in microarray experiments. BMC Genomics 2004 5:29.* **8.** *Haller F et al. Normalisation in quantitative RT-PCR with reference genes: Equivalence test in quantitative RT-PCR: Confirmation of reference genes suitable for normalization. Analyt. Biochem. 2004 335: 1-9.* **9.** *Sültmann H, Poustka A. Recent advances in transcription profiling of human cancer. Curr. Op. Mol. Therap. 2004 6, 593-599.* **10.** *Sültmann H et al. Gene expression in kidney cancer is associated with cytogenetic abnormalities, metastasis formation, and patient survival. Clin. Canc. Res. 2005 11(2): 646-655.* **11.** *Buness A et al. ArrayMagic: two-colour cDNA microarray quality control and preprocessing. Bioinformatics, 21(4), 554-556.* **12.** *Haller F et al. Prognostic role of E2F1 and members of the CDKN2A network in gastrointestinal stromal tumors. Clin. Can. Res. 2005. In Press.* **13.** *Schlomm T et al. Extraction and processing of high quality RNA from impalpable and macroscopically invisible prostate cancer for microarray gene expression analysis. Int. J. Oncol. 2005 27(3), 713-720.*