SMP: RNA

Project: Identification of Gene Expression Signatures in Breast Cancer

Holger Sültmann - **German Cancer Research Center (DKFZ), Heidelberg** - **h.sueltmann@dkfz.de**

Introduction

With about 46,000 new cases per year breast cancer is the most frequent type of malignancy in women. The incidence has drastically increased over the past 20 years (German Cancer Registry, 2002). Genetic factors predisposing to breast cancer are mutations in BRCA1 and BRCA2. However, these contribute to less than 10% of the breast cancer cases. In the recent past, initial progress was achieved in correlating gene expression signatures with prognostic and genetic factors in breast cancer. Yet accurate diagnostic and prognostic markers are lacking, and the confirmation of genetic patterns is still missing.

Macroscopic and histomorphological appearances of the tumors indicate that the category "breast cancer" is a heterogeneous mixture of different tumor entities. This may be the inherent cause for the widely varying degree of therapeutic success, as exemplified by the fact that about 80% of breast tumors will ultimately develop resistance against chemotherapeutic agents. Thus, new diagnostic and therapeutic markers are urgently needed.

The goal of this subproject is to comprehensively analyse one of the largest and best characterised tumor sample collections from breast cancer patients worldwide. By gene expression profiling of breast cancer samples, we establish comprehensive relationships between molecular signatures, prognostic factors (invasion, metastases, chemoresistance, survival) and pathological data (staging, grading, hormone receptor status, Her-2 expression).

Results/Project Status

The division of Molecular Genome Analysis at the DKFZ (SMP-RNA) and the Departments of Pathology and Clinical Oncology at the University of Graz (Prof. Dr. Kurt Zatloukal and Prof. Dr. Hellmut Samonigg) have been running a highly active collaboration on the analysis of a large set of breast cancer samples. The tumor bank in Graz has collected 2,600 frozen breast cancer samples and corresponding patient sera as well as 17,000 paraffin embedded breast tumor samples over the past 20 years. The sample collection in Graz is unique in Europe with respect to sample size (2.8 million paraffin-embedded tumor samples and 29,000 frozen samples) and quality for RNA-, DNA- and protein based studies. The advantage of this collection is the complete and long-term patient follow-up data (including estrogen receptor status, ERBB2 status, tumor morphology), which allows a comprehensive testing for correlations with the gene expression signatures. This offers a unique starting position which renders the project highly competitive at the international level.

Due to the heterogeneity of the tumors, all samples are histopathologically re-evaluated (Figure 1). All relevant sample data are raised and associated with the gene expression data.
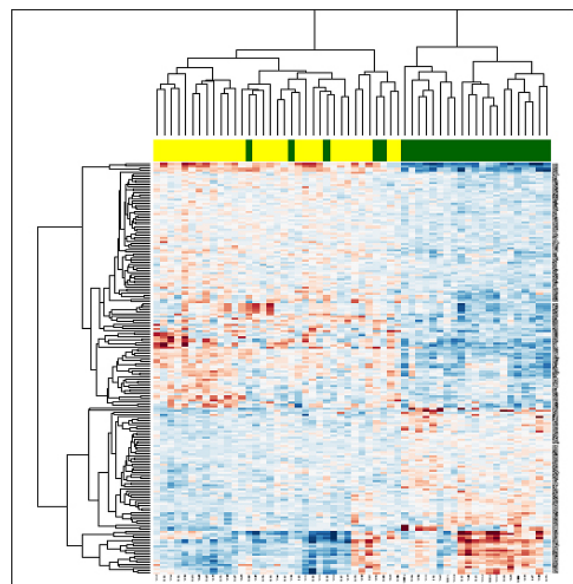
cDNA microarray analysis

WithIn the core area DKFZ, the division of Molecular Genome Analysis has analyzed 90 microdissected tumor samples from this collection using breast cancer specific microarrays containing 8,200 clones. These analyses revealed gene signatures associated with tumor type and tumor grading, estrogen receptor expression and morphological parameters indicating that the established protocols and quality controls work effectively. In particular, estrogen expression status of tumors revealed a clear signal

of differential gene expression (Figure 2). The genes in this signature may be useful for a better diagnosis of breast cancer patients and for the initiation of novel antioncogenic therapies.
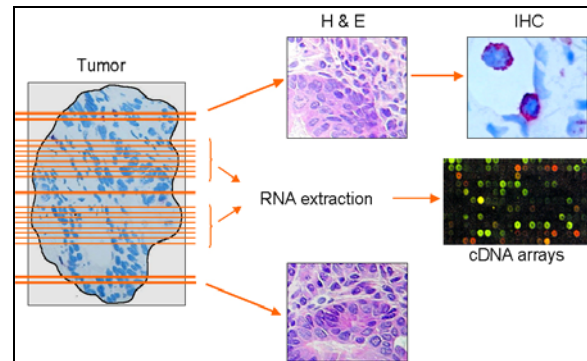
Fig. 1: Workflow of sample preparation: Tumors are sliced in a cryostat, and tissue sections from both ends are stained with HE as well as selected antibodies in order to assess the percentage of tumor and stromal cells in the tissue block. Subsequently, sections from the central region are macrodissected and subjected to microarray analysis.

Complementary to these experiments, hybridizations of 126 breast tumor samples on human whole genome cDNA microarrays encompassing 37,000 clones were performed. In agreement with the previous study, an estrogen receptor gene expression signature was identified. In addition, the expression of certain genes was associated with the amplification of chromosomal regions in breast cancer.


Fig. 2: Heatmap showing the association of gene expression (red: upregulated, blue: downregulated) with estrogen receptor expression (yellow and green samples) in breast cancer patients.

Oligonucleotide microarray analysis

A subset of tumor samples will be used for hybridization onto 70mer oligonucleotide microarrays (see also Core Unit DKFZ). These data will be compared to the cDNA microarray data obtained with the same set of samples. With this approach, we will be able to interrogate alternative splicing variants that are associated with tumor specific properties (e. g. histopathology, hormone receptor status) and clinical parameters (e. g. stage, grade, metastasis formation).

Validation by tissue arrays and functional assays

The validation of gene expression data will be extended to the analysis of protein expression using tissue arrays. The partners at the company Oridis Biomed (Graz) have set up the facilities for the automated construction of tissue arrays (Figure 3).

donor block   recipient block   TMA block   TMA section
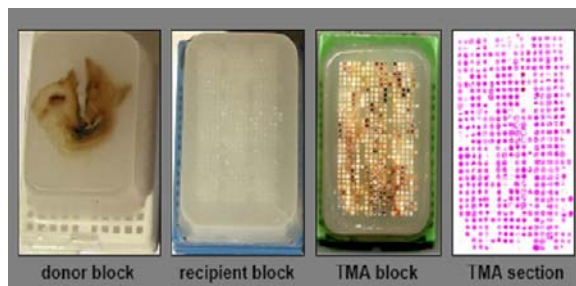
Fig 3: *Principle of tissue array technology. Hollow-needle samples are taken from paraffin-embedded tumor samples (donor block) and arranged in a regular grid (recipient block). After completion of the tissue array (TMA) block, sections are transferred to glass slides and stained with HE or antibodies of interest.*

Tissue blocks with tumor samples from the same patients whose gene expression data were obtained in microarray experiments will be used for the validation of protein expression. Tissue arrays will be constructed and incubated with antibodies directed against the proteins encoded by the differentially expressed genes. This will enable the spatial analysis of protein expression in breast cancer tissues.
Promising marker genes will be further characterized in cellular assays (SMP-Cell) and the EP "Standardized functional cell arrays" for functional validation as potential drug targets.

Outlook

The integration of data sets will provide the largest complementary record on gene and protein expression profiles in human cancer. This will make considerable contributions to the diagnosis and treatment of breast cancer and will represent a model for future integrated genomic and proteomic studies in human disease. Importantly, this cooperation will also be a model for future European research in that strong partner institutes from the German (NGFN) and the Austrian (GEN-AU) Genome Research networks will combine their expertises.

Lit.: **1.** *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.* **2.** *Haedicke W et al., Automated evaluation and normalization of immunohistochemistry on tissue microarrays with a DNA microarray scanner. Biotechniques 2003 35(1): 164-168. .*