Ruprecht-Karls-Universität Heidelberg

Medizinische Fakultät Mannheim

Dissertations-Kurzfassung

Bioinformatic analysis of RNA regulatory networks involving transcription factor regulation in prostate cancer.

Autor:	Prashant Kumar Srivastava
Institut / Klinik:	Medical Research Center, Medical Faculty Mannheim, University of Heidelberg. and German Cancer Research Centre (DKFZ), Heidelberg
Doktorvater:	Prof. Dr. med. Norbert Gretz

Despite much progress in PC management, new diagnostic, prognostic, and therapeutic tools are needed to predict disease severity, choice among available treatments and establish more effective therapies for advanced PC. In this thesis we dealt with this issue by: 1) developing methods for the microarray analysis of FFPE tissues making it comparable to the microarray data obtained by FF tissues. 2) Characterizing the implications of PC field cancerization and potential role of miRNA in this phenomenon. 3) A statistical tool was developed for the characterization of context specific modules for a given expression dataset.

Microarray analysis of FFPE tissue is of importance for the detection of molecular marker-sets in PC. The compromised RNA integrity of FFPE tissue results in a high degree of variability at the probe level of microarray data as shown by a degradation plot. We tested methods that reduce the variability by including all probes within 300 nucleotides, 600 nucleotides or up to a calculated breakpoint with reference to the 3'-end. Accepted PC pathways such as the Wnt signaling pathway could be observed to be significantly regulated within FFPE microarray datasets. The best representation of PC gene expression, as well as better comparability to meta-analysis and fresh-frozen microarray data, could be obtained with a 600 nucleotide cut-off. Beyond the specific impact for PC microarray data analysis we propose a cut-off of 600 nucleotides for samples where the integrity of RNA cannot be guaranteed.

Field cancerization is a phenomenon where the adjacent tissues to the tumor become genetically altered. Microarrays have been proven to be essential for the qualitative as well as quantitative analysis of relative expression for thousands of genes at the same time. However, relative gene expression can be rather misleading if one does not compare it to a suitable control. In PC, two types of tissues can be used as a control: 1) tissues from prostate free of any pathological alterations and 2) tissues adjacent to the tumor tissue. Most of the genome-

wide microarray mRNA expression profiling of PC have been performed using control tissue adjacent to the tumor. In many epithelial tumors, genetic alterations in histologically normal adjacent cells are described as "field cancerization". Although the phenomenon of field cancerization is well established in oral cancers, it is still to be established in PC. In this study we demonstrate that the choice of control tissues can lead to different (reverse) interpretations for certain sets of genes and pathways in PC.

PCA led to a set of 15 genes that have the potential of classifying metastasis in contrast to cancer and normal tissues. Some of them were previously reported to play an important role in metastasis. We also performed statistical analysis between the two types of control tissues and have characterized a set of genes that might have a potential to act as putative early stage marker genes for PC.

In addition, we are first to report the involvement of miRNAs in the PC field cancerization phenomenon and have shown that the hsa-miR-155 may play an important role in PC initiation, as well as its targets, oncogenes *ETS1* and *JUN*. Further we have developed a statistical method to characterize relevant modules between significantly and non-significantly altered sets of genes. The method would help in the identification of relevant modules that might be involved in the transition from PC to metastasis.

ETS1 target genes were identified from a published ChIP-chip experiment. Using this set of genes we demonstrate efficacy of the method at two levels: 1) with a set of genes that has a controlled number of *ETS1* target genes compared to an artificially constructed set containing varying numbers of false negatives. We have observed that even at 50% noise level our method can characterize relevant transcription factor pairs with a specificity of 88%. 2) We have analyzed knock-down experiments of *ETS1* and interacting transcription factors (*YY1* and *SP1*) and applied the method to characterize the difference between significantly and non-significantly altered *ETS1* target genes. Qualitative measures showed that, in this situation also a high specificity is maintained.

The method was applied to data from various stages of prostate cancer progression and has identified previously reported *ETS1* interaction partners as relevant for the modules in the significantly regulated genes. Some of the modules were observed to be unique to the metastatic state only. These modules consisted of TFs such as *STAT*, *YY1* and *P300* that have been proven to have metastatic potential. In addition we have also identified several novel potential factors interacting with *ETS1* such as *KID3* and *ZF5*.