

Molecular changes in androgen-independent prostate cancer

Abstract

Many prostate carcinomas are initially responsive to androgen ablation therapy. However, resistance to this initial androgen blockade usually develops and correlates with progression to androgen-independent disease. In this study, microarray data from the Gene Expression Omnibus (GEO) was used to examine the molecular changes between androgen-dependent and -independent primary prostate tumors. The data was generated by a study examining gene expression changes in androgen-independent prostate cancer (Best, et al., 2005). This dataset was alternatively analyzed with GeneSifter® microarray analysis software (VizX Labs, Seattle, WA). This analysis system was used to discover differentially regulated genes, and map them to candidate gene ontology terms and pathways. Several clear patterns of gene expression were discovered which correlated with genes involved in RNA metabolism, cell cycle, macromolecular biosynthesis, and apoptosis.

Introduction

Prostate cancer is the second most common malignancy in American men. Despite the number of studies on carcinoma of the prostate, there is little data describing the progression from androgen-dependent to androgen-independent disease, and development of aggressive androgen-independent prostate cancer (AIPC). In particular, microarray analysis of these tissues can be helpful in understanding the molecular changes between androgen-dependent and -independent prostate cancer. We examined previously published microarray results from 10 primary

untreated androgen-dependent tumors and 10 androgen-independent tumors. The analysis methodology can be broken down into two major steps: identification of differentially expressed genes, and the determination of biological significance of both individual genes and groups of genes. GeneSifter accomplishes this by creating both Gene Ontology (GO), and KEGG pathway reports. Both reports make use of z-score statistics to identify terms or pathways that are significantly over- or under-represented than expected by chance. This study describes the use of these methods to identify biological themes during the transition from androgen-dependent to -independent prostate cancer.

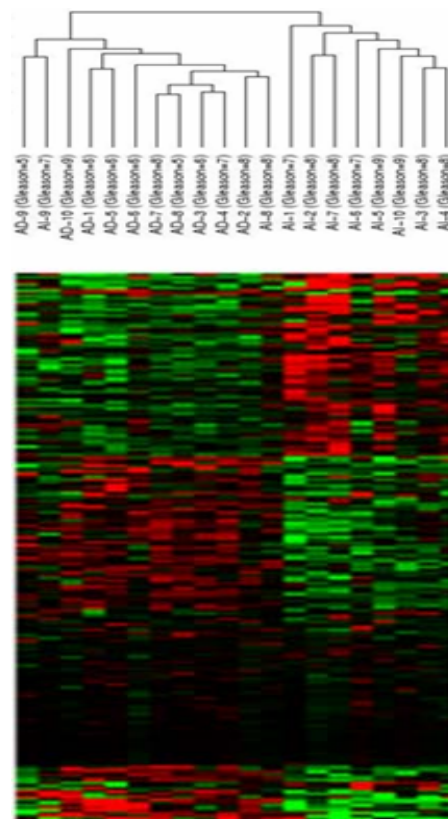


Figure 1. Hierarchical cluster analysis of androgen-dependent (AD1-10) and -independent (AI 1-10) samples (Gleason scores associated with each sample are also listed). The filtered data was subjected to hierarchical cluster analysis both by samples (columns) and genes (rows). Genes displayed in red are up-regulated and those in green are down-regulated (as compared to the row mean).

Microarray data

Gene expression profiles were measured in laser capture microdissected prostate tissue samples using the Affymetrix® GeneChip® HG-U133A containing probe sets for ~22,000 transcripts. A total of 20 samples were used: 10 androgen-dependent, and 10 androgen-independent (all from primary tumors). Data for each array was downloaded from the NCBI Gene Expression Omnibus (GEO), GSE2443. Data from the CEL files was loaded to GeneSifter for analysis.

Identification of differentially expressed genes

The probe-level data was compiled, normalized, and transformed using GC-RMA (Wu et al., 2004). Using filtering criteria of at least 1.5 fold change and a Wilcoxon rank sum test where $p < 0.05$, the raw data was reduced to a list of 785 genes. Applying the method of Benjamini and Hochberg (Reiner, et al., 2003) to estimate a false discovery rate of 5% reduced the list to 468 (analysis parameters are summarized in Table 1). Of this set, 222 genes were up-regulated, while the remainder (246) were down-regulated.

Normalization	GC-RMA
Statistical test	Wilcoxon rank sum test ($p < 0.05$)
Fold change	At least 1.5
Multiple testing correction	Benjamini-Hochberg (estimates FDR)

Table 1. Analysis methodology. A summary of the analysis parameters used to filter the raw data. Application of these methods resulted in a list of 468 genes, which were analyzed further for biological significance.

Biological significance

Those genes found to be differentially expressed were clustered using hierarchical cluster analysis to visualize the patterns of gene expression for each sample. Despite clustering only a filtered subset of the raw data, there remained differences in

expression between the groups such that several androgen-independent samples (AI-8, AI-9) were associated more closely with the androgen-dependent samples (Figure 1).

A.

Biological process GO term	Up	Array	z-up
Cell cycle	37	577	5.5
nuclear transport	10	91	4.8
RNA processing	20	283	4.4
ubiquitin cycle	23	350	4.4
RNA export from nucleus	5	35	4.1
biopolymer metabolism	77	1996	3.3
M phase of mitotic cell cycle	9	131	2.9
protein localization	23	499	2.5
Cell adhesion	21	489	2.1
regulation of signal transduction	9	169	2

B.

Biological process GO term	Down	Array down	z-down
macromolecule biosynthesis	45	473	9.7
protein targeting	11	128	4.3
cellular protein metabolism	79	2192	3.4
establishment of protein localization	22	484	2.8
intracellular protein transport	15	316	2.5
alcohol metabolism	11	211	2.4
protein transport	20	469	2.3
primary metabolism	144	4885	2.3
cellular carbohydrate metabolism	12	258	2.1
regulation of kinase activity	6	105	2

Table 2. Significant gene ontology terms affected by androgen ablation therapy. Part A contains the GO terms most prevalent among the up-regulated genes. Part B displays the z-scores associated with the down-regulated genes. Up/Down column contains the number of genes in each ontology term, "Array" is the number of genes on the array (U133A) that are in the specified ontology. The last column, z-up or z-down, displays the z-score statistic.

The biological process GO terms and KEGG pathways associated with the differentially expressed genes were examined using a z-score report. The z-score report identifies ontology terms or pathways that are significantly over- or under-represented (Doniger et al., 2003).

Among the genes which increased in expression, the most significant terms were those associated with RNA metabolism and cell cycle. Also highly over-represented were genes involved in cell cycle, cell adhesion, and angiogenesis. Among the genes decreased in the androgen-independent samples, several GO terms were also significantly over-represented including macromolecule (protein) biosynthesis, and protein localization and transport. Cell death and apoptosis genes both decreased. The most significant gene was GREB1, an estrogen-responsive gene thought to contribute to the invasiveness of breast cancers (Nagaraja, et al., 2005). A selected list of significant GO terms for both groups of genes can be found in Table 2 (for a complete list, see supplemental material).

We also examined pathway information for these results using data from the Kyoto Encyclopedia of Genes and Genomes (KEGG). Of particular interest was ribosomal metabolism with all pathway members universally decreasing in expression in AIPC (19 of 19 genes). A complete list of the affected KEGG pathways can be found in the supplemental material.

Summary

Microarray analysis of androgen-independent prostate cancer identified 468 genes that were differentially expressed following androgen ablation therapy. Our initial filtering criteria reduced the raw data set to 468 genes, which were subjected to further analysis. Hierarchical clustering allowed us to visualize these genes across the androgen-dependent and -independent samples. Despite pre-filtering the data prior to clustering, samples AI-8 and AI-9 appear to be associated with the androgen-dependent group, rather than -independent.

To identify gene families associated with our filtered gene list, we used the GO terms and KEGG pathways. Consistent with the current understanding, genes involved in

cell cycle and cell adhesion were over-represented in the androgen-independent group. Ribosomal genes universally decreased in expression, indicating that translation in general is profoundly affected. Of particular interest was GREB1, a protein which differentiates invasive breast cancer from non-invasive. It is possible that this protein also contributes to invasiveness in prostate cancer, as it appears to be responsive to androgen ablation therapy. These findings help to define a signature of gene expression changes and contribute to the understanding of androgen-independent prostate cancer.

Supplemental Material

Raw data is available from the Gene Expression Omnibus (GSE2443). Data can also be viewed and analyzed using the GeneSifter Data Center (www.genesifter.net/web/DC).

References

- Aimone et al., 2004. Spatial and temporal gene expression profiling of the contused rat spinal cord. *Experimental Neurology* 189(2): 204-221.
- Best, et al., 2005. Molecular alterations in primary prostate cancer after androgen ablation therapy. *Clin Cancer Res* 11(19): 6823-6834.
- Doniger et al., 2003. MAPPfinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biology* 4(1): R7.
- Nagaraja, et al., 2005. Gene expression signatures and biomarkers of noninvasive and invasive breast cancer cells: comprehensive profiles by representational difference analysis, microarrays and proteomics. *Oncogene* Nov. 28 (Epub ahead of print)
- Reiner et al., 2003. Identifying differentially expressed genes using false discovery rate controlling procedures. *Bioinformatics* 19(3): 368-375.
- Wu et al., 2004. Preprocessing of oligonucleotide array data. *Nature Biotechnology* 22(6): 656-658.