The development and utilization of a novel DNA microarray platform for biomarker and target identification in advanced prostate cancer

S. M. Walker[1], F. Mcdyer[1], G. R. Oliver[1], C. Greenan [1], S. Deharo[1], N. McCabe[1], P. G. Johnston[2], D. P Harkin[1,2], R. D. Kennedy[1,2].

[1]Almac Diagnostics, Craigavon, United Kingdom; [2]Queens University Belfast, Belfast, United Kingdom.

Background and Results

Prostate cancer is the second leading cause of cancer-related deaths in men; specifically one in six men are diagnosed with prostate cancer in their lifetime. The use of serum prostate-specific antigen (PSA) levels has been recognised as a huge step forward in the diagnosis and treatment of prostate cancer, however since its implementation as a biomarker it has become apparent that the numbers of deaths from prostate cancer has only decreased slightly. Therefore the identification of new biomarkers and treatments for highly invasive/metastatic prostate cancers is of a high priority.

We sought to identify new biomarkers and drug targets by performing DNA microarray analysis of high Gleason score prostate tumour samples and normal prostate tissue samples. To further these goals we developed a specific array platform, this array was based upon extensive sequencing of prostate tumour samples and contains approximately 90,000 probesets, many of which are specific to prostate cancer.

Methodology

-Generation of a prostate tumour specific microarray;
The Cancer DSA™ research tool generation process has been described in Tanney, et al., (2008) BMC Med Genomics 1(1):20. Briefly RNA was extracted from a series of 100 tumour and 50 normal prostate tissue samples, this was compiled into cDNA libraries representing the cancer and normal prostate transcriptomes. These libraries were sequenced and the transcripts compiled and multiple probes per transcript were generated upon Affymetrix GeneChip® technology (see technical specifications table).

PROSTATE CANCER DSA™ Technical Specifications	
Total Probesets on Array	~121,563
Control Probesets	2,567
Number of Transcripts	~93,000
Feature Size	11 Micron
Probe Length	25-MER
Probe Pairs / Probeset	11

-Patient Samples;
In total 15 samples were selected for profiling onto the prostate tumour specific microarray. All samples were reviewed by a pathologist and the percentage tumour and Gleason score were recorded where appropriate. Specifically these samples were enriched for tumours of high Gleason score with all samples being greater than 7 (see patient sample information table).

-Application of the platform to tumor samples;
Total RNA was isolated from frozen samples using Stat-60. Following RNA isolation, the frozen samples were subjected to DNase treatment using the RNase-free DNase set and then purified and concentrated using the RNeasy MinElute Cleanup Kit. Target preparation was performed using the WT-Ovation™ RNA Amplification System. The amplified cDNA was fragmented and labeled using the FL-Ovation™ cDNA Biotin Module V2. 5 µg of cDNA hybridized to DSA arrays. Arrays were then washed and stained using Affymetrix fluidics script EukGE-WS2-v4 and scanned using the Affymetrix GeneChip Scanner 3000 for data acquisition. All kits, reagents and equipment were used according to manufacturer's instructions.

-Bioinformatic analysis;
The raw data was preprocessed using the RMA algorithm and differentially expressed transcripts were selected using a combination of variance, fold change and t-test filtering. GeneGo Metacore was used to functionally characterise the resulting differentially expressed genes.

Patient Sample Information

SAMPLE NAME	CASE DIAGNOSIS	GLEASON SCORE	% TUMOUR
T1	Adenocarcinoma of prostate	Gleason Score: 4+4=8/10	100
T2	Adenocarcinoma of prostate	Gleason Score: 4+4=8/10	95
T3	Adenocarcinoma of prostate	Gleason Score: 3+4=7/10	60
T4	Adenocarcinoma of prostate	Gleason Score: 4+5=9/10	75
T5	Adenocarcinoma of prostate	Gleason Score: 4+5=9/10	80
T6	Adenocarcinoma of prostate	Gleason Score: 3+4=7/10	70
T7	Adenocarcinoma of prostate	Gleason Score: 3+4=7/10	85
T8	Adenocarcinoma of prostate	Gleason Score: 3+4=7/10	80
T9	Adenocarcinoma of prostate	Gleason Score: 3+5=8/10	75
T10	Adenocarcinoma of prostate	Gleason Score: 4+4=8/10	90
N1	N/A	N/A	N/A
N2	N/A	N/A	N/A
N3	N/A	N/A	N/A
N4	N/A	N/A	N/A
N5	N/A	N/A	N/A

Figure 1 Content of the Prostate Specific Array Platform

The Prostate Cancer specific DNA microarray contains 90,000 biologically relevant transcripts available for interrogation. The content was compared to the NCBI's Reference Sequence (RefSeq) database and it was determined that 62% was not present in the RefSeq database and 14% represented antisense sequences to annotated transcripts. This represents unique information which is believed to be specifically relevant to prostate cancer.

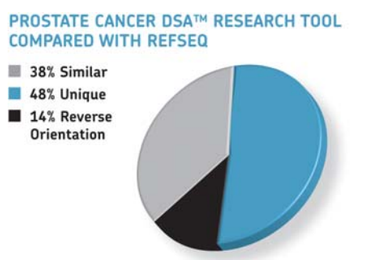


Figure 2 Analysis of protein homology of the unique content

To assess the relevance of the unique content we performed protein homology analysis. As shown in the table, this analysis revealed that a large proportion of the unique transcripts are likely to be protein coding and furthermore are involved in processes closely related to cancer initiation and development. In particular these unique transcripts have high homology to proteins which are involved in the cell cycle, cell motility, DNA damage response, control of apoptosis and DNA replication processes.

A similar proportion are transcripts with antisense homology to these cancer-related proteins and represent sequences with important regulatory potential.

Furthermore a large proportion of these unique transcripts are likely to be enzymes which may have relevance in the development of new drug targets for advanced prostate cancer. The annotation of these novel enzymes to known processes would assist in the validation of these targets.

PROTEIN HOMOLOGY ANALYSIS	SENSE	ANTISENSE
Uncharacterised protein count	1773	2118
Hypothetical protein count	213	475
Cell-cycle-related protein count	40	4
Intercellular signaling-related protein count	303	48
Metabolism-related protein count	645	77
Cytoskeleton, motility and spindle assembly-related protein count	453	113
Proteasome-related protein count	131	16
DNA damage/repair-related protein count	181	23
Cellular signalling-related protein count	378	81
Apoptosis-related protein count	21	15
Chromatin remodelling-related protein count	39	0
Kinase of unknown function count	43	6
Cell surface and intracellular receptors/ channel-related protein count	211	42
Transcription-related protein count	295	153
Cancer-related proteins of unknown function count	81	36
Immunoglobulin-related protein count	30	16
Translation-related protein count	390	76
Cellular stress response-related protein count	29	8
Cell-adhesion & extracellular matrix-related protein count	51	14
DNA replication-related protein count	17	2
Other proteins of known function	1399	491
	6723	3814

Figure 3 Microarray analysis of a series of prostate normal and tumour tissue samples

To validate the utility of this array platform to address the need for novel biomarkers and drug targets for advanced prostate cancer, we profiled a series of prostate tumour and normal tissue samples as detailed in the methodology sections. During QC we performed Principal Component Analysis to assess the quality of the information retrieved and the ability of our platform to distinguish normal & tumour tissue. The figure demonstrates that the data generated separated the samples efficiently into tumour and normal sample clusters, this indicates that the data generated is of biological significance. Thereby giving confidence that the differentially expressed transcripts identified are related to disease rather than technical factors.

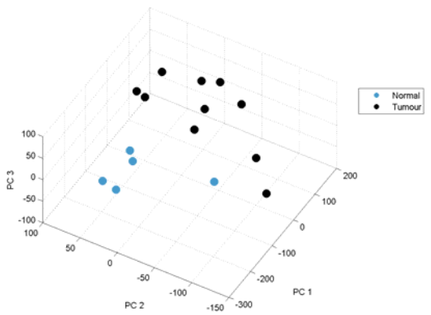


Figure 4 Selection of the top differentially expressed transcripts

We identified approximately 1600 differentially expressed transcripts between tumour and normal samples. In the adjacent table some of the significantly overexpressed transcripts present in the tumour samples when compared to the normal samples. Furthermore the identification of several transcripts which are known to be specifically expressed in prostate cancer including prostate cancer antigen 3 (PCA3), α-methylacyl-coA-racemase AKA 2-methylacyl-CoA 2-epimerase (AMACR) and members of the olfactory receptor family 51. This highlights the utility of this approach to identify both novel and well established biomarkers and targets of importance in prostate cancer. Additionally a significant proportion of the most highly expressed transcripts were annotated as antisense and these could represent novel cancer biomarkers.

A Selection of the Highest Overexpressed Transcripts

GENE SYMBOL	PROTEIN	PROTEIN NAME	ANNOTATION	FOLD CHANGE
FOLH1	FOLH1_HUMAN	Glutamate carboxypeptidase 2	Generic protease	8.6
CLDN8	CLDN8	claudin 8	Cell adhesion	13.5
RAD21	RAD21_HUMAN	Double-strand-break repair protein rad21 homolog	DNA Repair	13.7
SPOCK1	TICN1_HUMAN	Testican-1	Generic binding protein	13.9
AMACR	AMACR_HUMAN	Alpha-methylacyl-CoA racemase	Generic enzyme	15.2
PCA3	N/A	prostate cancer antigen 3 (non-protein coding)	RNA	18.4
PTPLB	PTPLB	Homo sapiens protein tyrosine phosphatase-like member b (PTPLB) mRNA,complete cds	AntiSense Transcript	19.1
PDLIM5	PDLIM5	Homo sapiens PDZ and LIM domain 5 (PDLIM5), transcript variant 2,	AntiSense Transcript	23.9
OR51E2	O51E2_HUMAN	Olfactory receptor 51E2	GPCR	30.5
PDLIM5	PDLIM5	Homo sapiens BAC clone RP11-554D13 from 4, complete sequence.	Uncharacterised Protein	40.3

Figure 5 Functional Analysis

Subsequent to the identification of differentially expressed transcripts between tumour and normal samples we sought to identify which molecular pathways were differentially expressed in the tumour samples.

We employed GeneGo Metacore software for these aims. Represented in the table are a number of the most statistically significant pathways identified with the associated p-values.

Interestingly a number of the pathways are known to be deregulated in prostate cancer including the WNT pathway, Insulin-like Growth Factor Receptor-1 (IGF-R1) signalling and Bcl-2-associated death promoter (BAD) apoptotic pathways.

Figure 6 A Deregulated pathway

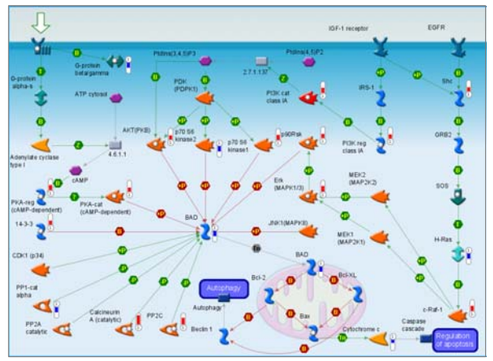
The attached figure represents one of the most statistically significant pathways identified, BAD regulation of apoptosis.

The temperature bars beside selected targets highlight the fold change in expression detected in the tumour samples when compared with normal. The colour red indicates overexpressed and blue indicates under expressed with the size of the bar highlighting the level of expression scales to the maximum in that pathway.

This pathway highlights that BAD which is a pro-apoptotic protein is significantly down-regulated in our series of prostate tumours. Furthermore the Metacore pathway highlights that a number of known down regulators of BAD are up-regulated (AKT, p70S6K etc). These targets maybe of significance as biomarkers but also novel drug targets. This analysis highlights the utility of this integrated approach to biomarker and target identification.

Top Pathways Identified Using GeneGo Metacore Analysis

GENEGO PATHWAY NAME	pVALUE
Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	1.303e-8
Protein folding_Membrane trafficking and signal transduction of G-alpha (i) heterotrimeric G-protein	2.799e-8
Regulation of lipid metabolism_Insulin signaling:generic cascades	6.242e-8
Apoptosis and survival_BAD phosphorylation	1.300e-7
Translation_Insulin regulation of translation	2.094e-7
Development_IGF-RI signaling	3.433e-7
NGF activation of NF-kB	4.640e-7
Cytoskeleton remodeling_Cytoskeleton remodeling	5.971e-7
Development_EDG1 signaling via beta-arrestin	5.99E-07
Translation _Regulation activity of EIF4F	7.409e-7



Conclusions

- **We developed a prostate tumour specific microarray with over 90,000 biologically relevant transcripts.**
- **A series of fresh high Gleason score primary prostate tumour and normal samples were profiled.**
- **We identified approximately 1,600 differentially expressed transcripts.**
- **Several transcripts were identified that were already reported in prostate cancer.**
- **544 transcripts were identified that were unique to the prostate cancer microarray platform and may represent novel drug targets and biomarkers.**
- **Functional analysis of cancer related transcripts identified apoptotic processes, DNA repair and cellular proliferation, amongst others.**
- **The prostate cancer specific DNA microarray may be a valuable tool for biomarker and target development in advanced prostate cancer.**