BJUI
BJU International

Identification of novel oncogenic events occurring early in prostate carcinogenesis using purified autologous malignant and non-malignant prostate epithelial cells

Pavel Sluka*, Carmel Pezaro*†, Hady Wardan*, Shomik Sengupta*‡§ and Ian D. Davis*†

*Eastern Health Clinical School, Faculty of Medicine, Nursing & Health Sciences, Monash University, †Eastern Health, ‡Department of Surgery, University of Melbourne, and §Department of Urology, Austin Health, Melbourne, Vic., Australia

Objective

To interrogate enriched prostate cancer cells and autologous non-malignant prostate epithelial cells from men with localized prostate cancer, in order to identify early oncogenic pathways.

Patients and Methods

We collected malignant and matched non-malignant prostatectomy samples from men with adenocarcinoma involving two or more contiguous areas in only one lobe of the prostate. Tissue samples from both lobes were subjected to digestion and single-cell suspensions were prepared. Epithelial cell adhesion molecule-positive cells from cancerous and contralateral non-malignant (control) samples were isolated using magnetic beads, ensuring uniform populations were obtained for each donor. Unbiased RNA sequencing analysis was used to measure gene expression and for detection of transcribed mutations or splice variants that were over- or under-represented in malignant prostate epithelial cells relative to autologous control prostate epithelial cells.

Results

From five patient samples we identified 17 genes that were altered in prostate cancer epithelial cells, with 82% of genes being downregulated. Three genes, *TDRD1*, *ANGTL4*, and *CLDN3*, were consistently upregulated in malignant tissue. Malignant cells from three of the five patients showed evidence of upregulated ERG signalling, however, only one of these contained a *TMPRSS2-ERG* rearrangement. We did not identify mutations, gene rearrangements, or splice variants that were consistent amongst the patients.

Conclusions

Events occurring early in prostate cancer oncogenesis in these samples were characterized by a predominant downregulation of gene expression along with upregulation of *TDRD1, ANGTL4* and *CLDN3*. No consistent mutations or splice variants were observed, but upregulation of *ERG* signalling was seen both in the presence and absence of the classic *TMPRSS2-ERG* rearrangement.

Keywords

prostate cancer, carcinogenesis, gene rearrangement, #ProstateCancer, #PCSM

Introduction

Prostate cancer is one of the commonest cancers worldwide and is increasing in incidence [1]. Most men with organ-confined prostate cancer are either cured by definitive surgery or radiotherapy, or eventually die from non-cancerous causes. Some prostate cancers are indolent and do not require any form of treatment [2]. Risk groupings such as the D'Amico score help to identify high-risk men and those who do not need immediate treatment [3]. Recent refinements of risk stratification include the introduction of Gleason grade grouping [4], but these clinical variables have limited ability to predict the clinical course for a specific cancer.

Interrogation of advanced prostate cancer reveals numerous defects that have been accumulated by the cancer as it has evolved under pressure from its microenvironment. The Cancer Genome Atlas (TCGA) has undertaken comprehensive genomic analyses of localized and metastatic prostate cancer [5]. This work showed that 74% of primary

BJU International © 2019 BJU International | doi:10.1111/bju.14695
Published by John Wiley & Sons Ltd. www.bjui.org

prostate cancers could be defined as one of seven subtypes based on characteristic gene fusions or mutations, with epigenetic heterogeneity and variable androgen receptor activity. These broad subtypes probably represent different early driver pathways that steered the individual cancers towards utilization and dependence on molecular events, which allowed the cancer to thrive in its microenvironment. The identification of changes that occur early on in the malignancy will assist in providing new diagnostics, risk stratification approaches, or therapeutic targets that are tailored to the cancer subtype.

Our current understanding of the molecular events that occur early on in prostate cancer have been hindered by the fact that prostate cancer cells are intermixed with normal prostate epithelial cells. Previous studies examining the molecular events of early prostate cancer have been contaminated by the presence of normal cells within the samples such as non-malignant epithelial, stromal, and other cell types. The events that have been identified in early prostate cancer, such as the *TMPRSS2-ERG* translocation [6], have probably been found because of their complete absence in normal tissue. More subtle changes such as altered gene expression levels, and in particular reduced gene expression levels, splice and epigenetic variants, and miRNA levels, would be more difficult to identify and could have been missed by previous studies.

These genome-wide studies used tissue samples that were contaminated by non-malignant stromal and other cells, and did not compare to autologous normal prostate epithelium. These additional tissue components may mask signals from relevant or targetable pathways that are up- or downregulated in cancer. We therefore aimed to use autologous matched epithelial cell adhesion molecule (EpCAM)-enriched cells to undertake deep global RNA sequencing (RNA-Seq) analysis, in order to identify up- and downregulation of gene expression and their associated signalling pathways, RNA splice variants, and transcribed mutations that are differentially expressed between exclusive populations of malignant and non-malignant prostate epithelial cells.

Materials and Methods

Specimen Collection and Processing

Primary tissue specimens were collected from patients undergoing radical prostatectomy for adenocarcinoma at Austin Health in Melbourne, Australia. The study was approved by the Austin Health Human Research Ethics Committee. Five tissue donors were selected based on TRUS-guided prostate biopsy reports that showed cancer involving at least two contiguous regions, confined to one prostate lobe. All patients had no clinical evidence of metastasis, and were treated with curative intent. No patient had received prior treatment for prostate cancer, and although some had locally advanced disease, all were considered to be early in the disease course because there had been no selection pressure through treatment such as androgen deprivation therapy. After prostatectomy, 8 mm diameter cores (~100–200 mg) were taken from areas in the involved lobe, as well as apparently uninvolved regions from the contralateral lobe, and stored in RPMI medium (Life Technologies, Waltham, MA, USA) for a maximum of 2 h at 4°C until processed. The entirety of each sample was processed in order to maximize cell yields. Quality assurance involved histological examination of the archival tissue blocks for evidence that the samples were derived from areas of malignancy or non-malignant tissue, respectively. A finding that at least 50% of the tissue surrounding the sampled malignant region was shown to contain cancer gave confidence that the sample was highly likely to contain significant numbers of malignant cells. Similarly, collection of non-malignant tissue was considered to be confirmed when the tissue surrounding the sampled region was completely devoid of cancer and the remainder of that lobe contained no cancer on routine pathological examination.

Samples were digested to generate single cell suspensions in a method adapted from Niranjan et al. [7] as follows. Reagents were obtained from Life Technologies unless otherwise stated. Samples were cut into 2-mm^3 pieces, washed twice with PBS containing 100 μg/mL gentamicin under unit gravity sedimentation, and digested overnight in RPMI medium containing 100 U/mL penicillin, 100 μg/mL streptomycin, 100 μg/mL gentamicin, 25 mM HEPES, 10% fetal calf serum (FCS), 225 U/mL collagenase, and 125 U/mL hyaluronidase, at 37°C with gentle rotation. Tissue pieces were then washed with RPMI medium containing 25 mM HEPES, 10% FCS, 100 U/mL penicillin and 100 μg/mL streptomycin, and single cells were collected by passing the digest through a 70-μm filter. Unfiltered tissue clumps were digested in 0.25% trypsin-EDTA for 30 min at 37°C and gently passed through 18-G, 20-G, and then 23-G needles to form a single cell suspension. Any remaining undigested material was removed by filtering the suspension through a 70-μm filter and combining cells with the first cell filtrate.

The EpCAM-positive cells within the generated suspension were enriched using EpCAM-coated Epithelial Enrich Dynabeads (Thermo Fisher Scientific, Waltham, MA, USA) according to the manufacturer's instructions. EpCAM-positive cell yields from each sample ranged from 80 000 to 460 000 cells per isolation.

The purity of EpCAM enrichment was assessed using flow cytometry. Cells were stained with PE-conjugated anti-human EpCAM antibody (Miltenyi Biotec, Sydney, NSW, Australia; 1:100) in the presence of Sytox Blue viability dye (1:1000).

Flow cytometry was performed using a FACSCanto II flow cytometer (BD Biosciences, Franklin Lakes, NJ, USA) equipped with FACSDiva software (v6.0; BD Biosciences) and the data were analysed using FlowJo software (v9.0; FlowJo, LLC, Franklin Lakes, NJ, USA). Epithelial (EpCAM-positive) cells comprised 30% of the live cell population that was generated post enzymatic digestion, which increased to 97% after enrichment with EpCAM antibody-coated beads (data not shown).

For RNA extraction, EpCAM-positive cells were lysed with RNA lysis buffer while still attached to magnetic beads. After lysis, beads were removed from the lysate using a magnet and the lysate was processed for RNA extraction using RNeasy Mini Kits (Qiagen, Hilden, Germany) according to the manufacturer's instructions and eluted in 30 μL of RNase-free water.

RNA-Seq

The RNA-Seq was performed by the Australian Genome Research Facility (Melbourne, Australia). RNA quality was first assessed using an Agilent 2100 BioAnalyser (Agilent Technologies, Santa Clara, CA, USA). All RNA Integrity Number scores were > 9, indicating the extracted RNA was of excellent quality. RNA was enriched using poly(A) selection and the entire RNA extract was used to construct an Illumina stranded RNA library and processed using the Illumina HiSeq for 100 bp stranded paired end reads. During sequencing, image analysis was performed in real time by the HiSeq Control Software v2.2.38 and Real Time Analysis v1.18.61.

Bioinformatics

The primary sequence data were generated using the Illumina CASAVA1.8.2 pipeline. The per-base sequence for all samples showed excellent quality with all bases being above Q30. Reads were screened for the presence of any Illumina adaptor/over-represented sequences and cross-species contamination. The cleaned sequence reads were aligned to the Homo sapiens genome (Build HG19). Reads were mapped to genomic sequences using the Tophat aligner (v1.3.1), including identifications of gene fusions. Transcripts were assembled with the Cufflinks tool (v2.2.1) using the RABT option and UCSC annotation.

Differential Expression

Differential gene expression was determined using the paired samples test with the Generalized Linear Model likelihood ratio test [8]. This test is similar to the traditional paired *t*-test, where changes in gene expression in malignant cells were compared to those in autologous control cells. Genes were deemed to be differentially expressed if they showed a false discovery rate, a method of testing for type I error in large-scale multiple comparisons, of ≤0.05. All genes that showed this false discovery rate had a magnitude of log fold change of 1.4 (equivalent to 2.5-fold in absolute terms) or greater, and were not filtered any further.

Validation of differential gene expression was performed using independent real-time RT-PCR on RNA from the same samples as were used for RNA-Seq. From each sample, 1 μL total RNA was reverse transcribed using Sensiscript III reverse transcriptase (Invitrogen, Waltham, MA, USA) using both random hexamer and oligo(dT) primers, according to the manufacturer's instructions. Real-time PCR was performed using SYBR Select Master Mix for CFX (Applied Biosystems, Waltham, MA, USA) with gene-specific primers (see Table S1 for primer sequences) on a CFX Connect Real-Time PCR Detection System controlled by CFX Manager Software (v2.1; Bio-Rad, Hercules, CA, USA), using the manufacturer's recommended PCR cycling conditions. 'No RT' controls were included to confirm absence of contaminating genomic DNA.

Hierarchical Clustering

The relationship of samples to one another was assessed in an unbiased and unsupervised fashion using hierarchical clustering, which was performed in R Studio (v3.4.3) with limma (v3.30.13), edgeR (v3.16.5), vegan (v2.4-3), and gplots (v3.0.1) libraries. A heatmap was generated using the heatmap.2 function of gplots after normalizing log read counts and by selecting the 100 genes with the highest absolute (modulus) variation between cancer and control samples.

Mutation Analysis

We reasoned that activating mutations could be drivers of oncogene overexpression or, conversely, that inactivating mutations could downregulate the expression of tumour-suppressor genes. The upstream signalling regulators of differentially expressed genes were examined for coding mutations that would result in changes to the amino acid sequence (i.e. missense, nonsense, insertion, deletion, or frameshift mutations) and consequently altered protein function and downstream gene expression.

For each gene that was differentially expressed, we consulted the literature to identify the signalling factors that regulated the transcription of each gene and then manually compared the sequences of these upstream genes in malignant and autologous control samples using the Integrative Genomics Viewer (v2.3; Broad Institute, Cambridge, MA, USA) to detect changes in the coding sequence of upstream genes. As a result of the polyA

selection of our samples, DNA regulatory elements and epigenetic mechanisms of gene regulation were not able to be assessed in this study.

Gene Set Enrichment Analysis

We further reasoned that in the absence of identifiable coding mutations, the biological pathways that were dysregulated in malignant cells could be identified using gene set enrichment analysis (GSEA [9]), which was performed using GSEA software (v2.2.2; Broad Institute). The hallmark gene sets from the Molecular Signatures Database were used for this analysis.

Results

Patients

Samples were collected from patients who underwent prostatectomy at Austin Health between April and June 2012. Patient features of the cohort are described in Table 1. Between 8.0×10^4 and 4.6×10^5 epithelial cells were obtained for each sample, from which total RNA was isolated and sent for RNA-Seq. Approximately 50 million 100 bp paired-end reads were obtained for each sample with quality scores above 30 across all bases. Over 93% of reads were successfully mapped to the human HG19 genome, with a unique mapping rate of between 70% and 75%.

Gene Expression Analysis

Differential expression analysis revealed significant changes in 17 genes, of which three were upregulated in malignant prostate epithelial cells relative to autologous non-malignant prostate epithelial cells, whilst 19 were downregulated (Table 2).

TDRD1 exhibited the greatest increase in expression in malignant epithelial cells. This increase was seen in three of the five patients examined (Fig. 1A). *TDRD1* expression (Fig. 1B) is known to be directly regulated by ERG (but not other ETS family members such as ETV1 [10]). The same three samples also showed upregulation of *ERG* mRNA (Fig. 1C), which is most commonly thought to be attributable to *TMPRSS2-ERG* gene arrangement; however, a *TMPRSS2-ERG* rearrangement was found in only one patient (patient 3). This implies that the upregulation of *ERG* and consequently of *TDRD1* in the other two samples occurred in the absence of a detectable *TMPRSS2* translocation; an observation that has not yet been reported in the literature. Upregulation of *ANGPTL4* and *CLDN3* expression (Fig. 2) was seen in all five and four of the five patients, respectively. These gene expression changes were validated by independent real-time RT-PCR (see Fig. S1).

Unbiased and unsupervised hierarchical clustering analysis of the 100 genes with greatest absolute variation between cancer and control samples showed that most cancer and control samples segregated into discrete groups (Fig. 3). Of particular interest, the three cancer samples with evidence of functional ERG signalling (patients 1–3) segregated furthest from the non-malignant samples. The matched cancer and control samples from patients 4 and 5 were more closely located, suggesting a more subtle distinction between them.

The GSEA was performed to identify pathways that were over- or under-represented in malignant prostate epithelial cells relative to non-malignant prostate epithelial cells. Genes involved in the MYC (NES = 1.23, qFDR = 0.145) and PI3K/AKT/mTOR (NES = 1.43, qFDR = 0.153) signalling pathways showed the greatest over-representation in malignant prostate epithelial cells (Fig. 4A,B), suggesting these pathways were most likely being exploited by malignant cells and therefore biologically relevant. KRAS (NES = −0.36, qFDR = 1.00) and ALK (NES = −1.17, qFDR = 1.00) signalling pathways were the most under-represented (Fig. 4C,D), suggesting potential downregulation of these pathways.

Table 1 Patient details.

Mean ± SD age at diagnosis, years	61.4 ± 8.9
Mean ± SD preoperative PSA level, μg/L	8.1 ± 7.4
Pathological stage, *n*	
pT2c	3
pT3b	1
pT4	1
Gleason score, *n*	
3 + 3	1
3 + 4	3
4 + 3	1
Lymph node metastases	None

Table 2 Genes differentially expressed between matched malignant and control prostate epithelial cells.

Gene	LogFC	False discovery rate	LogCPM
TDRD1	5.533074	0.000136	1.760846
ANGPTL4	1.830622	0.005588	8.042724
CLDN3	1.668238	0.021476	6.079898
GM2A	−1.42721	0.023844	5.928331
NDRG4	−1.8231	0.005588	6.037927
MMP9	−1.8395	0.008335	6.141663
SPP1	−1.948	0.013281	4.101353
IL1B	−1.96667	0.003978	7.026739
CD163	−2.0724	0.021476	3.866201
CCL4	−2.24481	0.001538	4.851115
CCL3	−2.28006	0.000966	5.455539
TRIM9	−2.39006	0.021476	3.347681
SIGLEC1	−2.43893	0.002621	3.397178
CLCA2	−2.73336	0.026676	6.004356
WSCD2	−2.9294	0.021476	0.237088
PIK3C2G	−3.11731	0.039265	1.118883
FGG	−4.05948	0.017377	0.552109

LogFC, log fold change; LogCPM, log counts per million.

Fig. 1 ERG **s**ignalling is upregulated in malignant prostate epithelial cells. Three of the five matched samples examined showed upregulation of *TDRD1* (**A** and **B**) and *ERG* (**C** and **D**; the upstream regulator of *TDRD1*) mRNA expression. Data are shown as absolute values for individual patients (**A** and **C**) and mean ± SEM for combined (**B** and **D**).

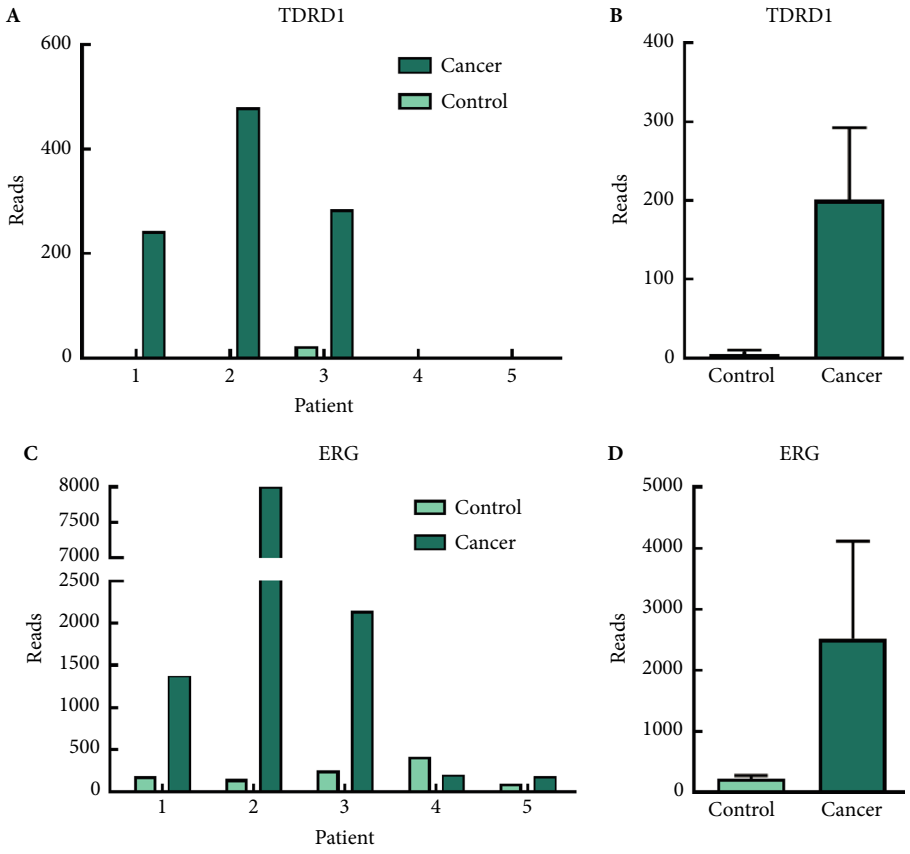
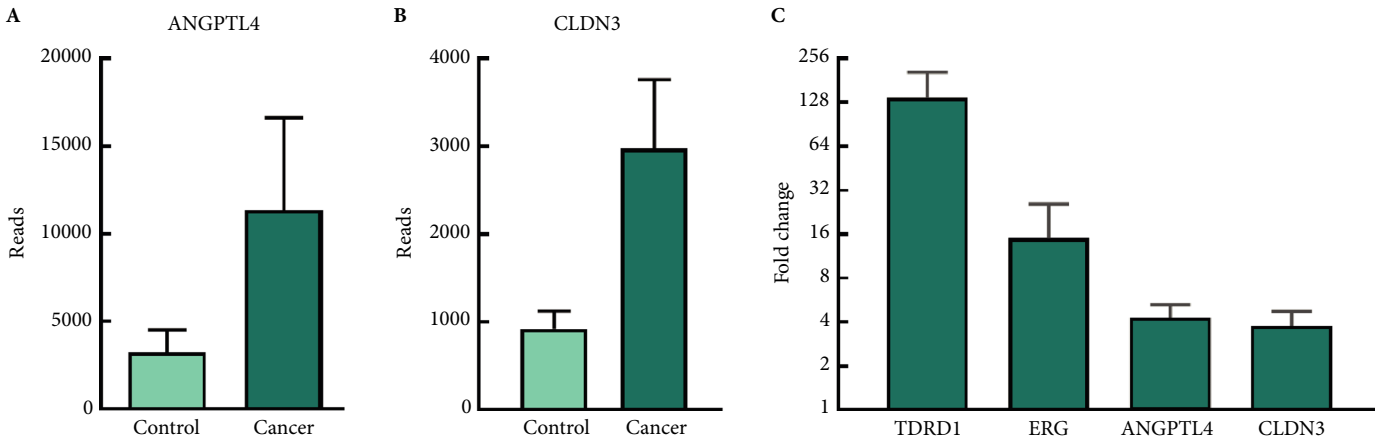


Fig. 2 ANGPTL4 and CLDN3 are upregulated in malignant prostate epithelial cells. *ANGPTL4* (**A**) and *CLDN3* (**B**) were upregulated in malignant prostate epithelial cells ($P < 0.01$). The fold increase in mRNA expression in malignant vs nonmalignant epithelial cells is shown in panel **C**. All data are shown as mean ± SEM.

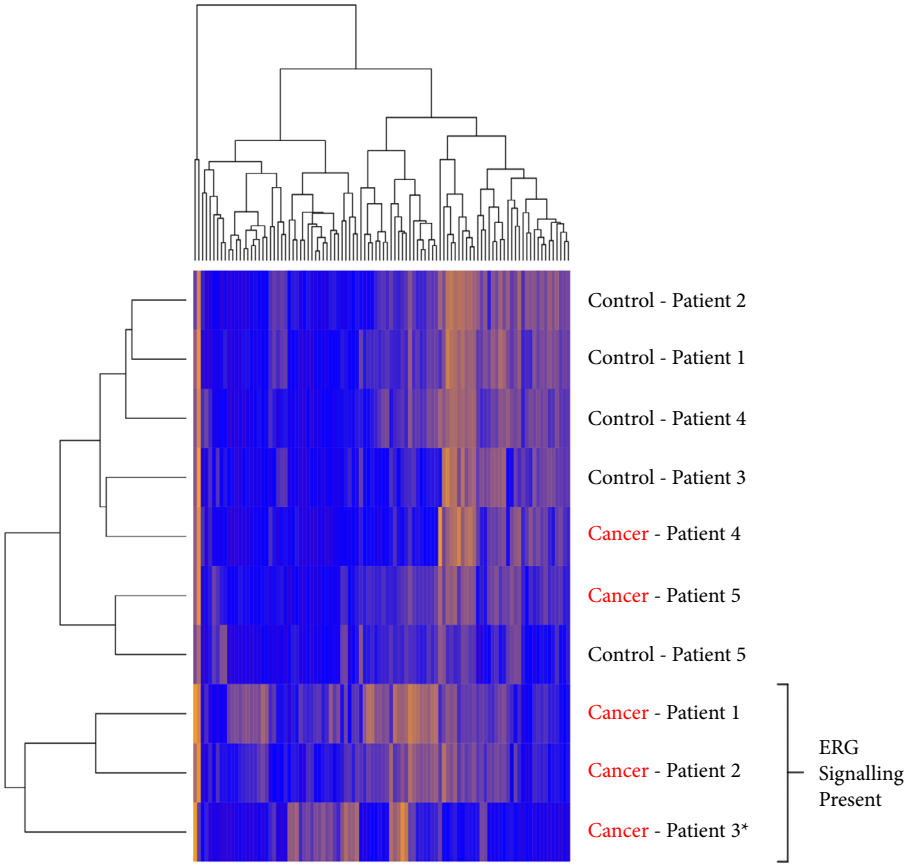


Splice Variant and Mutational Analysis

Splice variant analysis did not identify variants that reached statistical significance.

Given that the differentially expressed genes were not themselves mutated, we reasoned that the true driver of overexpressed but non-mutated genes was likely to be upstream in the regulatory pathway. The known upstream

Fig. 3 Hierarchical clustering. Hierarchical clustering showed segregation of three cancer samples (patients 1, 2, and 3), all of which contained evidence of functional ERG signalling. These samples were most distant to their matched controls. Of note, the cancer sample from patient 3 (asterisk) was the only sample in which a *TMPRSS2-ERG* rearrangement was observed.



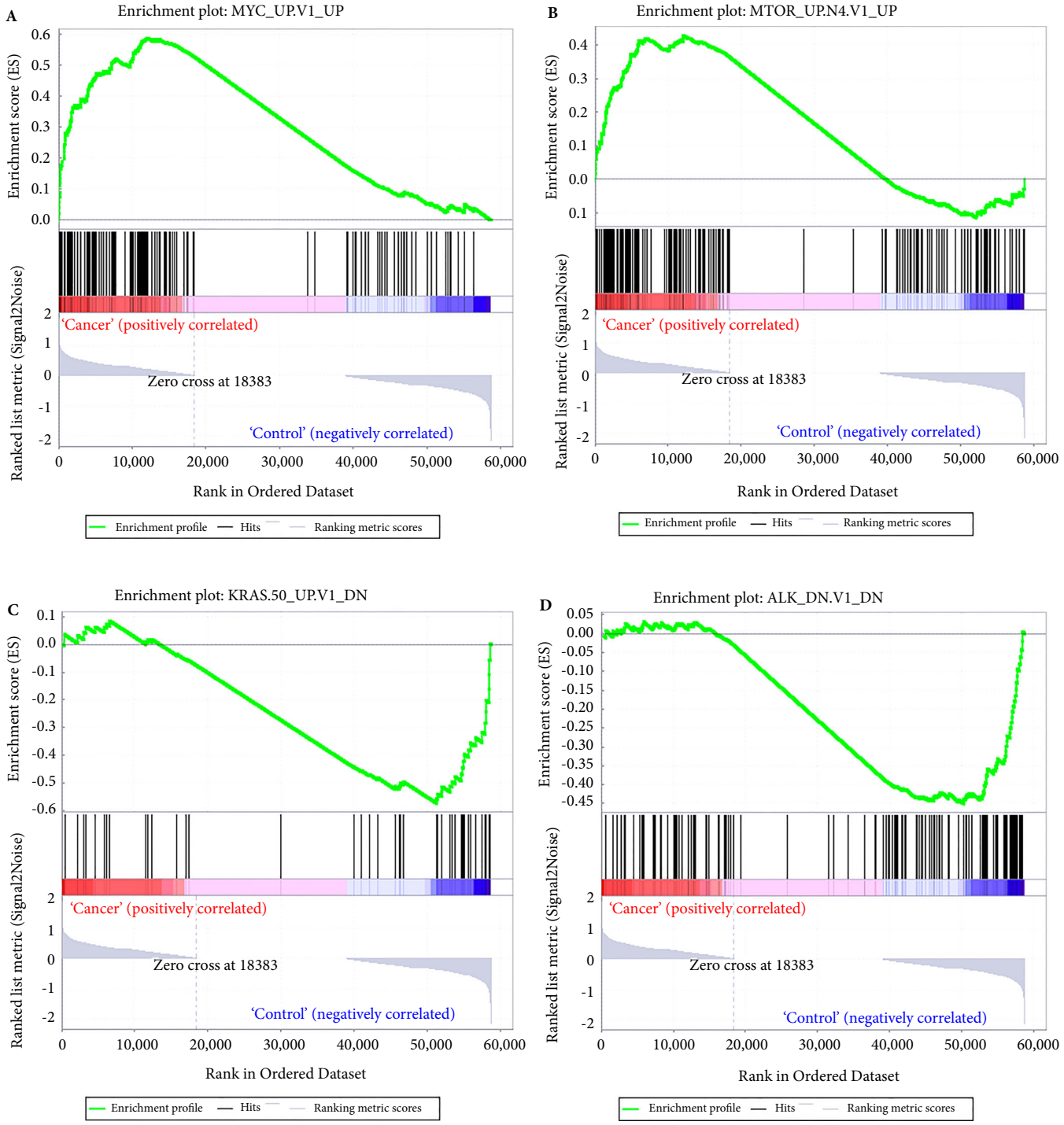
regulators of each of the differentially expressed genes were examined in an attempt to identify coding mutations that would result in a change in function of the native protein and could therefore explain the changes in expression observed. However, no coding mutations were identified. Other genes commonly found to be mutated in prostate cancer, specifically *PTEN*, *AR*, *TP53*, and *SPOP* were also examined and found not to be mutated in this set of samples from localized prostate cancer. We were unable to exclude epigenetic mechanisms as regulators of gene expression.

Discussion

In the present study we interrogated highly enriched prostate adenocarcinoma and matched autologous benign prostate epithelial cells, with a view to identifying early oncogenic events that distinguish malignant cells from non-malignant counterparts and that may represent early drivers of carcinogenesis. In contrast to previous studies, interference caused by contaminating stromal cells was minimized by using EpCAM purification of epithelial cells. Comparison against autologous non-malignant prostate epithelial cells removed other potential sources of confounding. The analysis identified the predominant finding of marked and statistically significant downregulation of gene expression. Three genes were consistently upregulated: *TDRD1*, *CLDN3* and *ANGPTL4*. The analysis also detected upregulation of *ERG* expression and signalling, both in the presence and absence of *TMPRSS2* fusion events, suggesting that early ERG signalling may be important in prostate oncogenesis but that mechanisms other than *TMPRSS2* rearrangements may be involved.

Activation of the ERG signalling pathway was confirmed by demonstration of concomitant upregulation of ERG and its downstream gene *TDRD1* [10,11]. The common upstream driver of upregulation of this pathway was not able to be identified; however, these findings suggest that upregulation of the ERG pathway and specifically of *TDRD1* might be an early marker of prostate carcinogenesis and, in some patients, a key driver event. Active ERG signalling in the absence of a *TMPRSS2* rearrangement is a novel finding that will require prospective validation in order to determine if it is of diagnostic or prognostic significance. Gene rearrangements

Fig. 4 Altered signalling pathways. Gene set enrichment analysis indicated that genes associated with MYC (**A**) and PI3K/AKT/mTOR (**B**) signalling were gained in malignant prostate epithelial cells. In contrast, KRAS (**C**) and ALK (**D**) signalling genes were reduced in malignant prostate epithelial cells. Gene set enrichment analysis takes a list of genes that are proposed to be related, in this case in their role in a common biological pathway. A positive association is seen as a sharp rise in the curve, which shows genes that were upregulated in the experiment were found in the pre-defined gene list. A negative association is seen as a sharp fall from the right hand side of the graph. A lack of association would be seen as a line that hovers around the *x*-axis across its entire length.



other than *TMPRSS2-ERG* were not identified in the patients examined, consistent with previous reports indicating a low level of gene rearrangements in early prostate cancer.

The upregulation of *ANGPTL4* expression was seen in all five patient samples. ANGPTL4 is a pro-angiogenic gene [12] whose expression has been correlated with prostate cancer tumorigenicity *in vitro*, where its overexpression correlates with colony formation, and *in vivo*, where it enhances tumour formation in mice [13,14].

CLDN3 was found to be overexpressed in four of five tumour samples, supporting previous work where CLDN3 protein was overexpressed in prostate cancer [15] as well as other epithelial tumours [16]. CLDN3 has also been examined as a potential prostate cancer biomarker, where it was able to discriminate malignant from normal prostate tissue better than PSA [17]. *CLDN3* is widely expressed in the body and its levels in serum are more variable than PSA levels, making it unlikely to be useful as a predictive or prognostic marker.

Whilst the analysis of individual genes provides insights into specific pathways of early oncogenesis, unsupervised and unbiased hierarchical clustering showed distinct clustering of samples, which provides evidence that distinct gene expression profiles are present in epithelial cells from early localized prostate cancer in which intra-patient noise has been subtracted. Cancer samples that contained ERG signalling with (patient 3) and without (patients 1 and 2) detectable *TMPRSS2-ERG* rearrangement clustered together, suggesting a similar biological profile; however, the finding that ERG upregulation occurred in the absence of TMPRSS2-ERG rearrangement in some patients raises the question that TMPRSS2-ERG rearrangement might not necessarily be the only mechanism for inducing this profile.

Our finding of significant downregulation of gene expression in malignant cells suggests that silencing of key tumour-suppressor pathways may be early events in prostate carcinogenesis. Our inability to detect mutations in these pathways implies that the underlying mechanisms of gene downregulation may be epigenetic; a hypothesis that we could not confirm with the available coding sequence information.

Despite the overall downregulation of individual genes, GSEA revealed typical cancer-associated pathways (MYC, PI3K/AKT/mTOR pathways) to be over-represented in malignant prostate cancer epithelial cells, indicating that the cells have established signalling pathways supporting their oncogenesis.

The method in which cells were prepared in this study is very different from that used in other genome-wide studies in that it used highly enriched epithelial cells that express the cell surface marker EpCAM. This allowed a much more focused interrogation of the molecular events occurring in malignant prostate cells and may be a reason for the identification of a relatively small number of differentially expressed genes.

Multidisciplinary meetings usually concentrate on the clinical aspects of care [18]. Careful study of patient samples in ethics-approved protocols can lead to findings that might be of clinical relevance, and the multidisciplinary meeting may be a useful forum to remind clinicians of the value of correlation of translational research with clinical outcomes. Careful study of patient samples can lead to clinically relevant insights even with small numbers of samples. Large-scale studies looking for population-level changes may dilute the significance of specific changes in individual patient samples, especially if analyses are performed using unpurified samples without appropriate matched controls. This report shows that: early oncogenic events can be detected and characterized at the level of individual patients; ERG pathway signalling in prostate cancer is likely to be biologically relevant in some patients but not required in others; upregulation of ERG signalling does not necessarily require early *TMPRSS2-ETS* family fusion events; *ANGPTL4* and *CLDN3* upregulation may be common and occurs early; and downregulation of pathways is much more frequent in early prostate cancer than upregulation.

Acknowledgements

We gratefully acknowledge the support of Associate Professor David Powell and Dr Kirill Tsyganov from Monash University Bioinformatics Platform, and of Genevieve Whitty, the staff of Austin Health Department of Anatomical Pathology, and the Austin Health Urology Unit for collection and processing of samples. Funding for this work was provided in part by the Eastern Health Foundation. The FACSAria III was purchased with equipment grant support from the National Health and Medical Research Foundation (NHMRC) through Monash University. IDD is supported by an NHMRC Practitioner Fellowship (APP1102604).

Conflicts of Interest

Dr Sengupta reports other from Mundipharma Australia, other from Janssen Australia, other from Ipsen Australia, other from MSD Australia, other from Eastern Melbourne Primary health network, outside the submitted work, personal fees from Mundipharma Australia, personal fees from Janssen Australia, personal fees from Ipsen Australia, personal fees from MSD Australia, personal fees from Eastern Melbourne Primary health network, outside the submitted work. Dr Davis reports grants from National Health and Medical Research Council, grants from Monash University, grants from Eastern Health Foundation, during the conduct of the study. The remaining authors have no conflicts of interest.

References

1 **Australian Institute of Health and Welfare**. Cancer in Australia 2014: actual incidence data from 1982 to 2011 and mortality data from 1982 to 2012 with projections to 2014. *Asia Pac J Clin Oncol* 2015; 11: 208–20

2 **Chen RC**, **Rumble RB**, **Loblaw DA et al.** Active surveillance for the management of localized prostate cancer (Cancer Care Ontario Guideline): American Society of Clinical Oncology Clinical Practice Guideline Endorsement. *J Clin Oncol* 2016; 34: 2182–90

3 **Boorjian SA**, **Karnes RJ**, **Rangel LJ**, **Bergstralh EJ**, **Blute ML**. Mayo Clinic validation of the D'amico risk group classification for predicting survival following radical prostatectomy. *J Urol* 2008; 179: 1354–60

4 **Epstein JI**, **Egevad L**, **Amin MB et al.** The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: definition of Grading Patterns and Proposal for a New Grading System. *Am J Surg Pathol* 2016; 40: 244–52

5 **Cancer Genome Atlas Research Network**. The molecular taxonomy of primary prostate cancer. *Cell* 2015;163:1011–25.

6 **Mosquera JM**, **Perner S**, **Genega EM et al.** Characterization of TMPRSS2-ERG fusion high-grade prostatic intraepithelial neoplasia and potential clinical implications. *Clin Cancer Res* 2008; 14: 3380–5

7 **Niranjan B**, **Lawrence MG**, **Papargiris MM et al.** Primary culture and propagation of human prostate epithelial cells. *Methods Mol Biol* 2013; 945: 365–82

8 **McCarthy DJ**, **Chen Y**, **Smyth GK**. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 2012; 40: 4288–97

9 **Subramanian A**, **Tamayo P**, **Mootha VK et al.** Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 2005; 102: 15545–50

10 **Boormans JL**, **Korsten H**, **Ziel-van der Made AJ et al.** Identification of TDRD1 as a direct target gene of ERG in primary prostate cancer. *Int J Cancer* 2013; 133: 335–45

11 **Xiao L**, **Lanz RB**, **Frolov A et al.** The germ cell gene TDRD1 as an ERG target gene and a novel prostate cancer biomarker. *Prostate* 2016; 76: 1271–84

12 **Le Jan S**, **Amy C**, **Cazes A et al.** Angiopoietin-like 4 is a proangiogenic factor produced during ischemia and in conventional renal cell carcinoma. *Am J Pathol* 2003; 162: 1521–8

13 **Ifon ET**, **Pang AL**, **Johnson W et al.** U94 alters FN1 and ANGPTL4 gene expression and inhibits tumorigenesis of prostate cancer cell line PC3. *Cancer Cell Int* 2005; 5: 19

14 **Feng S**, **Agoulnik IU**, **Truong A et al.** Suppression of relaxin receptor RXFP1 decreases prostate cancer growth and metastasis. *Endocr Relat Cancer* 2010; 17: 1021–33

15 **Bartholow TL**, **Chandran UR**, **Becich MJ**, **Parwani AV**. Immunohistochemical profiles of claudin-3 in primary and metastatic prostatic adenocarcinoma. *Diagn Pathol* 2011; 6: 12

16 **Hewitt KJ**, **Agarwal R**, **Morin PJ**. The claudin gene family: expression in normal and neoplastic tissues. *BMC Cancer* 2006; 6: 186

17 **Amaro A**, **Esposito AI**, **Gallina A et al.** Validation of proposed prostate cancer biomarkers with gene expression data: a long road to travel. *Cancer Metastasis Rev* 2014; 33: 657–71

18 **Kinnear N**, **Smith R**, **Hennessey DB**, **Bolton D**, **Sengupta S**. Implementation rates of uro-oncology multidisciplinary meeting decisions. *BJU Int* 2017; 120(Suppl 3): 15–20

Correspondence: Pavel Sluka, Eastern Health Clinical School, Level 2, 5 Arnold Street, Box Hill, Vic. 3128, Australia.

e-mail: pavel.sluka@monash.edu

Abbreviations: EpCAM, epithelial cell adhesion molecule; RNA-Seq, RNA sequencing; FCS, fetal calf serum; GSEA, gene set enrichment analysis.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Independent validation of gene expression.
Table S1. Real-time PCR primers.