Received: 5 April 2022 | Revised: 23 June 2022 | Accepted: 8 July 2022

DOI: 10.1111/cts.13372

ARTICLE

Liquid biopsy-based targeted gene screening highlights tumor cell subtypes in patients with advanced prostate cancer

Seta Derderian[1,2] | Quentin Vesval[1,3] | Michel D. Wissing[1,4] | Lucie Hamel[1] | Nathalie Côté[1] | Marie Vanhuyse[4] | Cristiano Ferrario[5] | Franck Bladou[6] | Armen Aprikian[1,2,4] | Simone Chevalier[1,2,4,7]

[1]Urologic Oncology Research Group, Cancer Research Program, Research Institute (RI)-McGill University Health Center (MUHC), Montreal, Canada

[2]Department of Surgery (Urology Division), MUHC and McGill University, Montreal, Canada

[3]Department of Urology, Centre Hospitalier Régional et Universitaire (CHRU) de Rennes, Rennes, France

[4]Department of Oncology, MUHC and McGill University, Montreal, Canada

[5]Department of Oncology, Jewish General Hospital (JGH) and McGill University, Montreal, Canada

[6]Department of Urology, Centre Hospitalier Universitaire de Bordeaux, Bordeaux, France

[7]Department of Medicine, McGill University, Montreal, Canada

Correspondence
Simone Chevalier, Research Institute, McGill University Health Centre, 1001 Blvd Décarie, E M2.2210, Montréal, Quebec, Canada, H4A 3J1.
Email: simone.chevalier@mcgill.ca

Funding information
Prostate Cancer Canada/Movember; MUHC Prostate Cancer Research funds

Abstract

Prostate cancer (PCa) clinical heterogeneity underscores tumor heterogeneity, which may be best defined by cell subtypes. To test if cell subtypes contributing to progression can be assessed noninvasively, we investigated whether 14 genes representing luminal, neuroendocrine, and stem cells are detectable in whole blood RNA of patients with advanced PCa. For each gene, reverse transcription quantitative polymerase chain reaction assays were first validated using RNA from PCa cell lines, and their traceability in blood was assessed in cell spiking experiments. These were next tested in blood RNA of 40 advanced PCa cases and 40 healthy controls. Expression in controls, which was low or negative, was used to define stringent thresholds for gene overexpression in patients to account for normal variation in white blood cells. Thirty-five of 40 patients overexpressed at least one gene. Patients with more genes overexpressed had a higher risk of death (hazard ratio 1.42, range 1.12–1.77). Progression on androgen receptor inhibitors was associated with overexpression of stem (odds ratio [OR] 7.74, range 1.68–35.61) and neuroendocrine (OR 13.10, range 1.24–142.34) genes, while luminal genes were associated with taxanes (OR 2.7, range 1.07–6.82). Analyses in PCa transcriptomic datasets revealed that this gene panel was most prominent in metastases of advanced disease, with diversity among patients. Collectively, these findings support the contribution of the prostate cell subtypes to disease progression. Cell-subtype specific genes are traceable in blood RNA of patients with advanced PCa and are associated with clinically relevant end points. This opens the door to minimally invasive liquid biopsies for better management of this deadly disease.

Seta Derderian, Quentin Vesval, and Michel D. Wissing contributed equally to the content of this manuscript.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Clinical and Translational Science* published by Wiley Periodicals LLC on behalf of American Society for Clinical Pharmacology and Therapeutics.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

Prostate cancer (PCa) is clinically heterogeneous and difficult to treat, especially in an advanced stage. However, all patients are treated similarly. The one-treatment-fits-all approach could be revisited by a better understanding of the tumor PCa cell subtypes present in blood.

WHAT QUESTION DID THIS STUDY ADDRESS?

Can circulating genes represent the three PCa cell subtypes depicted in tumors?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

The genes selected to represent PCa cell subtypes contribute to progression as transcriptomic datasets showed their predominant overexpression in metastases. Their detection in the blood of patients with PCa with advanced disease is not related to white blood cells, but to their clinical features and treatments.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

This approach can be applied in the clinic to identify overexpressed prostate cell subtypes' genes in the blood of patients with advanced disease, to select optimal treatments and impact on progression and death from PCa, and to assess overexpression longitudinally to offer alternative therapies earlier.

INTRODUCTION

Despite recent therapeutic improvements, prostate cancer (PCa) remains among the leading causes of cancer mortality worldwide, with virtually all deaths occurring at the metastatic castration-resistant (mCRPC) stage.[1] Tumor cells surviving androgen-deprivation therapy (ADT) continue to grow despite inhibition of steroid synthesis and androgen receptor (AR) signaling.[2] Resistance may involve the emergence of constitutively active AR variants (AR-Vs) and/or cell plasticity, a concept based on the relationship among reserve stem, neuroendocrine, and luminal cells in glandular acini, which remain relevant in cancer.[2,3] PCa is initially defined as well-differentiated adenocarcinomas with predominantly AR-positive luminal-like cells. Neuroendocrine differentiation frequently emerges under ADT and AR inhibitors (ARIs), giving rise to AR-independence/androgen-insensitivity.[4,5] ADT also favors cells expressing AR-Vs, which are linked to epithelial-mesenchymal transition and stemness.[6,7] Accordingly, cell plasticity and trans-differentiation resulting from treatment failure support the emergence of cancer cells displaying neuroendocrine- and stem-like properties, alongside altered luminal-like cells.

Identifying biomarkers of intra-tumoral cell diversity is critical for offering and developing optimal therapies. A major limitation of testing markers in advanced disease remains the invasiveness of biopsies of metastases. This can be overcome by liquid biopsies, particularly blood which contains extracellular vesicles (EVs), circulating tumor cells (CTCs), and cell-free nucleic acids.[8,9] CTCs can be isolated by various methods, including antigen-based positive-selection or separation based on their physical properties.[10,11] However, this often leads to an underestimation of CTCs due to loss of certain subpopulations, as highlighted in PCa.[10–15] Whole blood RNA has been used as an attractive alternative to define gene expression profiles related to progression for diverse cancers, including PCa.[16–21]

In this pilot study, we assessed the (over)expression of 14 genes representing prostate cell subtypes in whole blood RNA of 40 patients with advanced PCa and 40 healthy controls. We observed a variety of gene patterns in blood samples, also validated in PCa transcriptomic datasets. Circulating genes were associated with disease stage, treatment, and survival. Thus, this minimally invasive approach may potentially predict treatment resistance and outcome and refine therapeutic options.

MATERIALS AND METHODS

PCa cell lines

Human PCa cell lines LNCaP (ATCC CRL-1740), DU145 (ATCC HTB-81), and PC-3 (ATCC CRL-1435) were from American Type Culture Collection (ATCC, Manassas, VA). NCI-H660 (ATCC CRL-5813) and 22Rv1 (ATCC CRL-2505) were generous gifts from Dr. A. Zoubeidi (Prostate Centre, Vancouver, BC, Canada) and Dr. M. Tremblay (McGill Goodman Cancer Institute, Montreal, QC, Canada), respectively. Cells were cultured according to

ATCC recommendations for no more than passage 30, with routine mycoplasma testing (MycoFluor Mycoplasma kit; Thermo Scientific, Waltham, MA).

Patients and healthy volunteers

This study (MP-37-2017-3158) was approved by the Ethics Review Board of the McGill University Health Center and accepted by the Ethic Review Board of the Jewish General Hospital. Participants voluntarily signed an informed consent. Blood was drawn from 40 patients with CRPC at follow-up visits with oncologists (authors M.V. and C.F.), and 40 healthy controls: 16 men below 50 years old, 15 men older than 50 years, and nine women recruited among students, personnel, and patients with no prostate-related disease or cancer. A pseudonymized database was generated from patients' medical records.

Blood processing and RNA extraction

Peripheral blood was collected in PAXgene tubes (Qiagen; Germantown, MD), kept at room temperature for 2 h for cell lysis, frozen overnight at −20°C, and long-term storage at −80°C. RNA was extracted using PAXGene RNA Extraction kits and QIAcube following Qiagen protocols. RNA integrity (RIN) and concentration were determined by BioAnalyzer 2100 (Agilent, Millcreek, ON, Canada) and Nanodrop ND-1000 (Thermo Scientific), respectively (RIN average 8.8 for patients and controls). RNA was stored at −80°C.

Reverse transcription quantitative polymerase chain reaction

Total RNA (1 μg) was reverse transcribed using SuperScript First Strand Synthesis kit (Thermo Scientific) with dNTPs (10 mM), oligo$(dT)_{12-18}$ (0.5 μg/μl), and water (total volume of 20 μl). The cDNA was stored at −20°C. Primers were obtained from OriGene (Rockville, MD), according to best designed sequences, except *ARV7* for which primers were based on literature (Table S1).[22] Specificity was determined using Primer-BLAST.[23]

Each quantitative polymerase chain reaction (qPCR) reaction contained 10 μl of SsoAdvanced SYBRGreen Master Mix (Bio-Rad, Mississauga, ON, Canada), 0.5 μl each of reverse and forward primers (10 μM), 4 μl of diluted cDNA (1:4) and 5 μl of water. Reactions were run in triplicate in 96-well plates, using CFX96 Touch (Bio-Rad). The program was 30 s initiation at 95°C, 15 s denaturation at 95°C, 15 s annealing, and elongation at 60°C (58°C for *ARV7*) for 40 cycles. Amplicon length was verified by running qPCR reactions on agarose gels.

The CFX Manager software (version 3.1) was used to quantify gene expression. Relative normalized gene expression was determined by the $2^{-\Delta Ct}$ method. Three reference genes were chosen based on previous literature and interindividual stability verified using GeNorm.[24,25] RNA from PCa cell lines mixed in equal proportions was used as positive controls and inter-run calibrator for all genes in each plate. Water was used as a negative control. Examples of qPCR amplification are shown (Figure S1).

Statistical analyses

Overexpression was defined as 2.58 SDs above the mean expression of each gene in blood of volunteers, ensuring with 99.5% certainty that expression in controls was lower than this threshold.

To detect whether patient characteristics or treatments were associated with overexpression of individual genes, all characteristics were dichotomized, and differences evaluated using Chi-square tests in two by two tables. PSA levels were compared to *KLK3* expression using the Wilcoxon rank-sum test, due to the non-normal distribution of PSA levels (evaluated using a Shapiro–Wilk test). To adjust for repeated measurements within individuals, multilevel, mixed effects logistic regression was used to identify variables associated with overexpression of luminal, neuroendocrine, and stem cell genes. Survival analyses for disease progression and death were performed using the Cox proportional hazards model. If no event had occurred, patients were censored at the last date they were known to be alive and/or without disease progression, either clinical (worsening of symptoms: fractures, and pain), biological (PSA rising in two subsequent measurements), or radiological (new lesion or increased size of existing lesions). Results with $p < 0.05$ were considered significant. All statistical analyses were performed using Stata version 15.0.

Datasets and bioinformatic analyses

Cell line reverse transcription qPCR (RT-qPCR) results were compared to RNA-sequencing from the Cancer Cell Line Encyclopedia Expression Atlas.[26] Gene expression data of white blood cells (WBCs) was obtained from the consensus dataset from the Human Protein Atlas (v20. proteinatlas.org).[27] Gene expression microarrays were accessed through Gene Expression Omnibus for the Stanford (GSE3933), Cambridge (GSE70770), and Memorial Sloan Kettering Cancer Center (MSKCC; GSE21032) datasets.

RNA-sequencing results were authorized and accessed through dbGaP for SU2C (phs000915.v2.p2) and GDC for The Cancer Genome Atlas (TCGA; phs000178.v11.p8). For the Stanford dataset, preprocessed data were normalized as initially reported.[28] For the MSKCC and Cambridge datasets, preprocessed data were normalized using the LIMMA Bioconductor package in R.[29] For the SU2C and TCGA datasets, fastq and BAM files were processed using the Canadian Centre for Computational Genomics GenPipes RNA-seq pipeline, followed by count normalization using the DESeq2 and LIMMA Bioconductor packages in R.[29–31] Overexpression was defined as 2.58 SDs over the expression in benign samples from the same dataset, except for SU2C which was compared to TCGA benign samples after normalization and batch correction. Fisher's exact test was used to compare gene overexpression by categories of cases for each dataset. Cochrane-Armitage test for trends was used to analyze gene overexpression by Gleason score (GS). Kaplan–Meier survival analysis and uni- and multivariable Cox analyses were performed for MSKCC and TCGA datasets.

RESULTS

Prostate cell-subtype genes are differentially expressed in PCa cell lines and are traceable in blood

Cell subtype-specific genes were chosen based on a literature review encompassing prostate luminal, neuroendocrine (neuroendocrine differentiation), and stem cells. The luminal genes were: *KLK3* (PSA), *ARV7*, *FOLH1* (prostate-specific membrane antigen/PSMA), and *AR*.[32] For the neuroendocrine subtype, *SYP* (synaptophysin), *ENO2* (neuron specific enolase), *VEGFA* (vascular endothelial growth factor-A), and *NCAM1* (neural cell adhesion molecule/CD56) were chosen.[4,33] For stem cells, *ALDH1A1* (aldehyde dehydrogenase), *CD44*, *EZH2* (enhancer of zeste homolog 2), *NANOG*, *POU5F1* (octamer-binding transcription factor/OCT4), and *SOX2* (SRY/sex determining region Y-box 2) were included.[34]

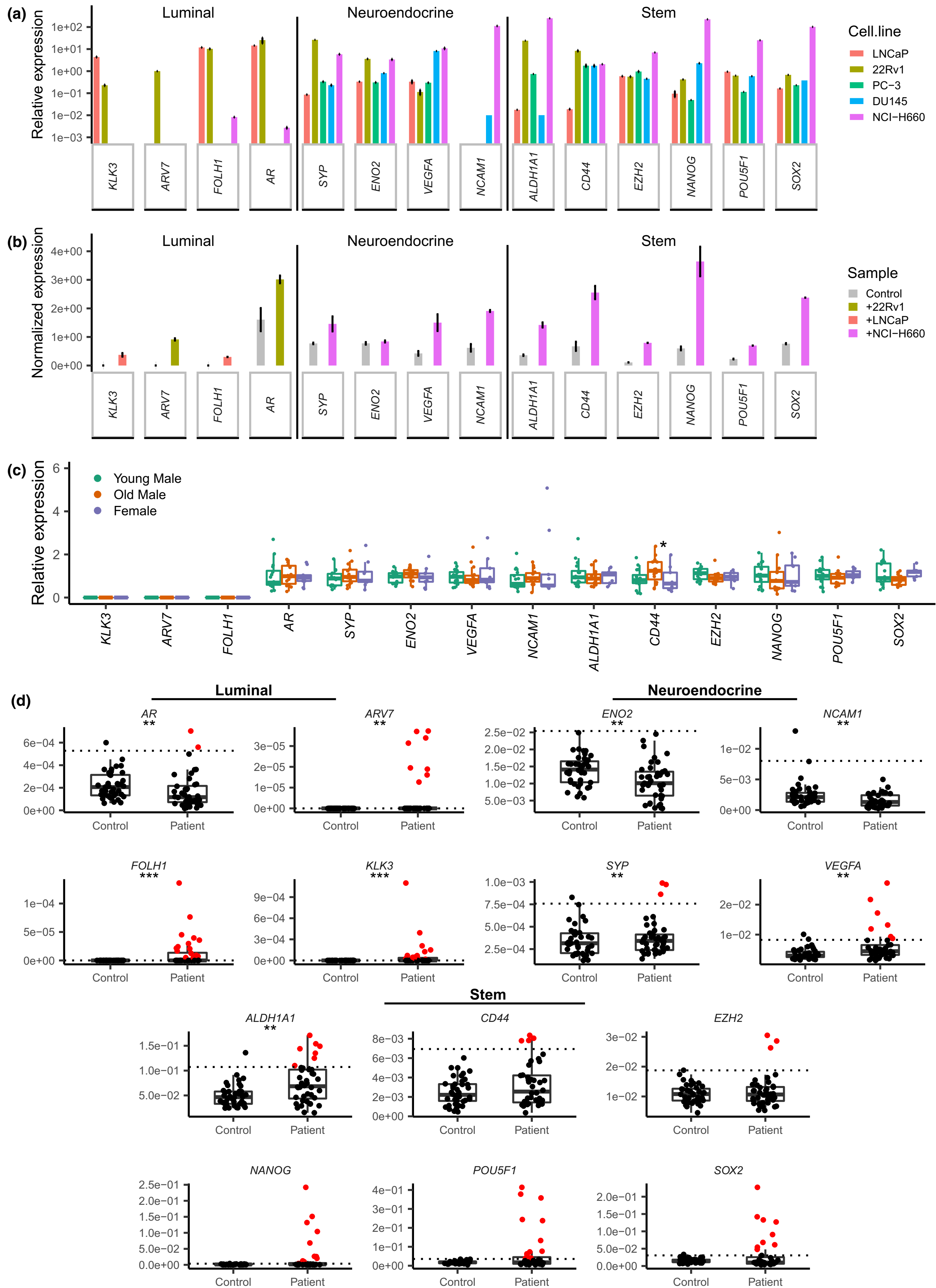
These genes were first assayed in five PCa cell lines (Figure 1a). Luminal genes were expressed in 22Rv1 and LNCaP (except 22Rv1-specific *ARV7*) and negative in the AR-negative PC-3 and DU145, whereas NCI-H660 showed low *AR* and *FOLH1* expression. Among neuroendocrine genes, *NCAM1* was mainly expressed in NCI-H660 with low expression in DU145, whereas *SYP*, *ENO2*, and *VEGFA* were detected in all cell lines, notably in the neuroendocrine/NCI-H660 model. Similarly, stem cell genes were highest in NCI-H660, but also detected in other cell lines. These results are in line with CCLE RNA-sequencing data (Figure S2). Altogether, each PCa cell line displays some but not all features of pure cell subtypes.

To mimic the situation in patients, fresh control blood was spiked with 10–50 PCa cells (Figure 1b). *KLK3*, *FOLH1*, and *ARV7* were detected upon spiking with LNCaP or 22Rv1 cells but not in controls, whereas other genes showed baseline expression in controls which increased upon spiking with 22RV1 (for *AR*) or NCI-H660 cells (for neuroendocrine and stem genes). Higher gene expression upon spiking blood with few PCa cells suggested that cell-subtype genes may be traceable in blood.

Patient cohort and features

Forty patients with advanced PCa were enrolled in this pilot study; baseline characteristics are displayed in Table S2. Patients were intermediate or high-risk at

FIGURE 1 Genes representing prostatic cell subtypes are differentially expressed in prostate cancer (PCa) cell lines and are traceable in blood. Gene expression assays were optimized in PCa cell lines and tested in control blood spiked with PCa cells before testing in 40 patients and 40 controls. (a) RNA was extracted from five human PCa cell lines (LNCaP, 22Rv1, PC-3, DU145, and NCI-H660) and reverse transcribed to test our panel of genes by reverse transcription quantitative polymerase chain reaction. Results were expressed for each gene as mean normalized relative quantities on a logarithmic scale. (b) Blood drawn from a healthy male in PAXgene tubes (2.5 ml) was spiked with LNCaP, 22Rv1, or NCI-H660 cells. For LNCaP and 22Rv1, 10 cells were added to the blood (equivalent of 4 circulating tumor cells per ml of blood). NCI-H660 cells grow as clumps in suspension, therefore we estimated that ~50 cells were added to the control blood. Genes were tested in control versus spiked blood. Results are presented on a linear scale. (c) Genes were tested in control blood and compared by age and sex. Relative normalized gene expression is shown in box plots for each gene in male controls by age (<50 vs. >50 years old) and female controls (25–70 years old; no difference by age). Wilcoxon rank-sum test is significant ($p < 0.05$) for *CD44* between young versus older men (*). No significant differences were detected by sex. Similar analysis of gene expression with age of patients also showed no correlation (data not shown). (d) Box plots of results in controls and patients for luminal, neuroendocrine, and stem cell genes. Red dots for patients represent overexpression, as defined by values greater than means of controls +2.58 SD (Table S3). Black dots in patients represent expression below cutoff values for each gene (dashed line). Some genes (*KLK3*, *ARV7*, and *FOLH1*) are not detected in controls and in subsets of patients. Wilcoxon rank-sum test p values comparing patients and controls are denoted by * for $p < 0.05$; ** for $p < 0.01$; *** for $p < 0.001$ (under gene names).


(a)
Relative expression
Luminal
Neuroendocrine
Stem
Cell.line
LNCaP
22Rv1
PC-3
DU145
NCI-H660
KLK3
ARV7
FOLH1
AR
SYP
ENO2
VEGFA
NCAM1
ALDH1A1
CD44
EZH2
NANOG
POU5F1
SOX2
(b)
Normalized expression
Sample
Control
+22Rv1
+LNCaP
+NCI-H660
(c)
Relative expression
Young Male
Old Male
Female
(d)
Luminal
Neuroendocrine
Stem
Control
Patient

diagnosis, with 52.5% having metastatic disease. At enrollment, all but one patient were diagnosed with mCRPC, for a median duration of 19 months. More than half of the patients were progressing and 12.5% were on end-of-life palliative care. Ninety percent had received systemic therapy for mCRPC, mainly ARIs and taxanes. Median follow-up was 7 months, during which 12 patients died.

Cell-subtype genes are not related to age or WBCs but to clinical features

Our gene panel was tested in the blood of 40 patients and 40 controls. Analysis of expression in controls showed no differences between men and women (Figure 1c). Comparisons by age showed increased *CD44* expression in older versus younger men, with no difference for other genes, despite an increased age-related risk of benign prostatic hyperplasia. Accordingly, stringent thresholds were defined for gene overexpression in patients at the 99.5% confidence interval of expression in 40 controls (Table S3), except for *CD44* whose threshold was set using the 15 older men.

Results of patients and controls are shown in box plots (Figure 1d). The accuracy, sensitivity, and specificity of identifying patients versus controls by overexpression of at least one gene were 88.8%, 87.5%, and 90%, respectively. As expected from Figure 1c, there was no signal for *KLK3*, *FOLH1*, and *ARV7* in controls and basal expression for other genes. The mean expression in patients was significantly higher than controls for six of 14 genes. Nonetheless, 12 of 14 genes were overexpressed in subsets of cases. *ENO2* and *NCAM1* were not overexpressed in any patients.

As baseline levels of 11 genes in controls suggest expression in WBCs, we investigated these genes in RNA-sequencing data from normal WBC lineages (Figure 2a). Most genes were negative or minimally expressed, except for *CD44* found at varying levels in all lineages, *ALDH1A1* in monocytes, and *ENO2* and *EZH2* in subsets of lymphocytes.

Patients' WBCs (Figure S3A) were next analyzed in relation to gene expression in blood. Most genes were not significantly associated to proportions of different WBC lineages, except *SYP* with neutrophils and *CD44* with monocytes (Figure 2b). However, when genes showing any correlation with WBCs were represented in scatter plots (Figure S3B), not all patients with high neutrophil or monocyte counts overexpressed *SYP* or *CD44*, respectively. These findings further strengthen the use of stringent thresholds when assessing overexpression in patients' blood. In contrast, 9/14 genes were related to patients' clinical features (Figure 2c), such as PSA, visceral metastases, Eastern Cooperative Oncology Group (ECOG) score, line of therapy, progression, surgery, and past or current mCRPC therapies. For instance, *KLK3* expression was associated with higher PSA levels (Figure 2d) as well as disease progression and death in survival analyses (Table 1). Overexpression of *AR*, *VEGFA*, or *CD44* was also associated with increased risk of death. Altogether, cell-subtype genes are overexpressed in liquid biopsies of patients with advanced PCa and gene patterns are associated with clinical features.

Cell-subtype genes in liquid biopsies support phenotypic diversity

The diversity of gene expression in blood among patients is illustrated in Figure 3a by a heatmap. Thirty-five of 40 patients (87.5%) overexpressed at least one gene. Luminal genes were the most prevalent (62.5%), followed by stem (45%) and neuroendocrine (15%) genes. Overall, 26 unique patterns were identified, with up to 11 genes in P09. No patient overexpressed all markers of a category, supporting diversity within subtypes. Correlations between genes showed positive trends among luminal genes (Figure 3b). For neuroendocrine genes, positive correlations were found among *VEGFA*, *SYP*, and *NCAM1*, whereas *ENO2* displayed a distinct pattern. Among stem cell genes, positive correlations were observed among *SOX2*, *NANOG*,

FIGURE 2 Cell-subtype genes are not related to WBCs but to clinical features of patients with prostate cancer (PCa). In (a) are bioinformatic analyses of the 14 genes in normal WBC lineages in consensus data from the Human Protein Atlas (v20.proteinatlas.org), as described in the Methods section. Normalized expression from 0 (white) to 70 (red) is shown. (b) Spearman correlation matrix between gene expression in the 40 patients with mCRPC and their WBC counts at the time of blood draw. Blue represents a negative association, red a positive association. The p values are entered when correlations were significant ($p < 0.05$), and at least >0.4 or <−0.4. (c) Heatmap depicting associations between patients and treatment characteristics and overexpression of individual genes. Blue represents a negative association, and red a positive correlation. The darker the color, the stronger the association between gene overexpression and patients' characteristics. The p values are entered when the patient's characteristic was significantly associated with overexpression of that gene ($p < 0.05$). (d) Association between the detection of the *KLK3* transcript in the blood of patients and PSA levels (logarithmic scale) at time of blood draw. The p value was calculated using a Wilcoxon rank-sum test. AR, androgen receptor; ECOG, Eastern Cooperative Oncology Group; mCRPC, metastatic castration-resistant; T-reg, regulatory T cells; WBC, white blood cell.

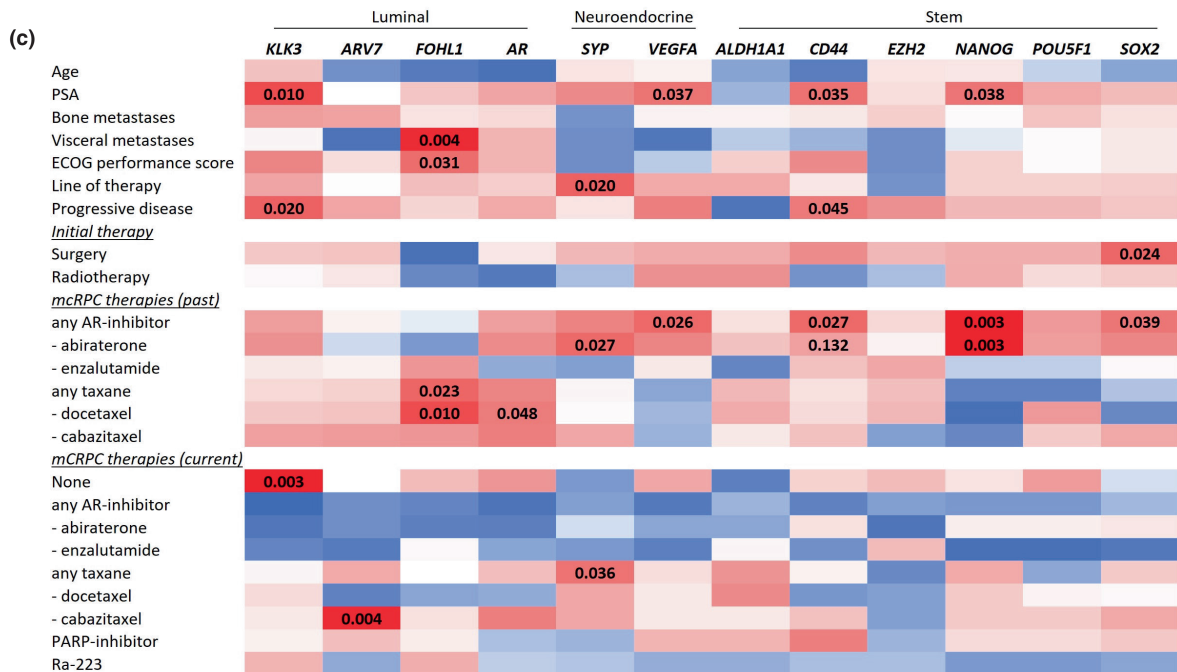
(a)

Lineage		*KLK3*	*ARV7*	*FOLH1*	*AR*	*SYP*	*ENO2*	*VEGFA*	*NCAM1*	*ALDH1A1*	*CD44*	*EZH2*	*NANOG*	*POU5F1*	*SOX2*
Granulocyte	Eosinophil	0	.	0.3	0	0.4	0.1	1.9	0	0.5	6	1.4	0.5	0.5	0
	Basophil	0	.	0.1	0.2	1.7	1.3	0	0	0	19.1	0.5	3.2	1.4	0
	Neutrophil	0	.	0.2	0.3	5.3	6.2	0.5	0.4	1.1	16.4	0.2	2.3	1.8	0
Monocyte	Intermediate	0	.	0.6	0	0.7	0	0	0	5.8	45.7	1.5	0.4	0.7	0
	Non-classical	0	.	0.5	0	0.4	0.2	0	0	0.4	59.3	3.9	0.5	0.5	0
	Classical	0	.	0.8	0	0.8	0	1.2	0.1	34.2	65.2	2.2	0.7	1	0
Lymphocyte	Memory B-cell	0	.	0.6	0	0.7	5.6	0	0	0	22.7	1.5	0.6	1	0
	Naive B-cell	0	.	0.9	0	1.3	3.3	0	0	0.3	16.2	0.8	1.1	3	0
	Naive CD4 T-cell	0	.	0.7	0	0.5	7.1	0	0	0	34.5	1.9	0.7	0.9	0
	Memory CD4 T-cell	0	.	1.2	0	0.5	10.8	0	0.1	0	39.8	3.4	0.5	0.4	0
	Memory CD8 T-cell	0	.	0.5	0	0.7	9.6	0	0.2	0	34.9	2.9	0.4	0.4	0
	MAIT T-cell	0	.	0.2	0	0.4	9.7	0	0.6	0	37	1.3	0.4	0.5	0.3
	Naive CD8 T-cell	0	.	0.3	0	0.5	7.4	0	0.8	0	28.7	1.2	0.8	0.6	0
	T-reg	0	.	0.4	0	0.3	5.5	0	0.9	0	53.1	10.3	0.3	0.3	0
	GdT-cell	0	.	0.9	0	0.2	8.2	0	2	0	26.2	1.7	0.6	0.8	0

(b)

Lineage		*KLK3*	*ARV7*	*FOLH1*	*AR*	*SYP*	*ENO2*	*VEGFA*	*NCAM1*	*ALDH1A1*	*CD44*	*EZH2*	*NANOG*	*POU5F1*	*SOX2*
Granulocyte	% Eosinophil														
	% Neutrophil					0.015						0.045			
Monocyte	% Monocyte										0.040				
Lymphocyte	% Lymphocyte					0.001		0.029			0.030			0.045	0.030

(c)



(d)

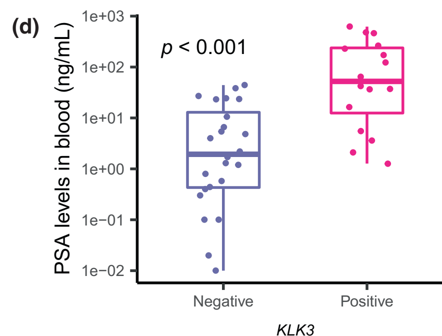


TABLE 1 Overexpression of circulating genes is associated with disease progression and patient death

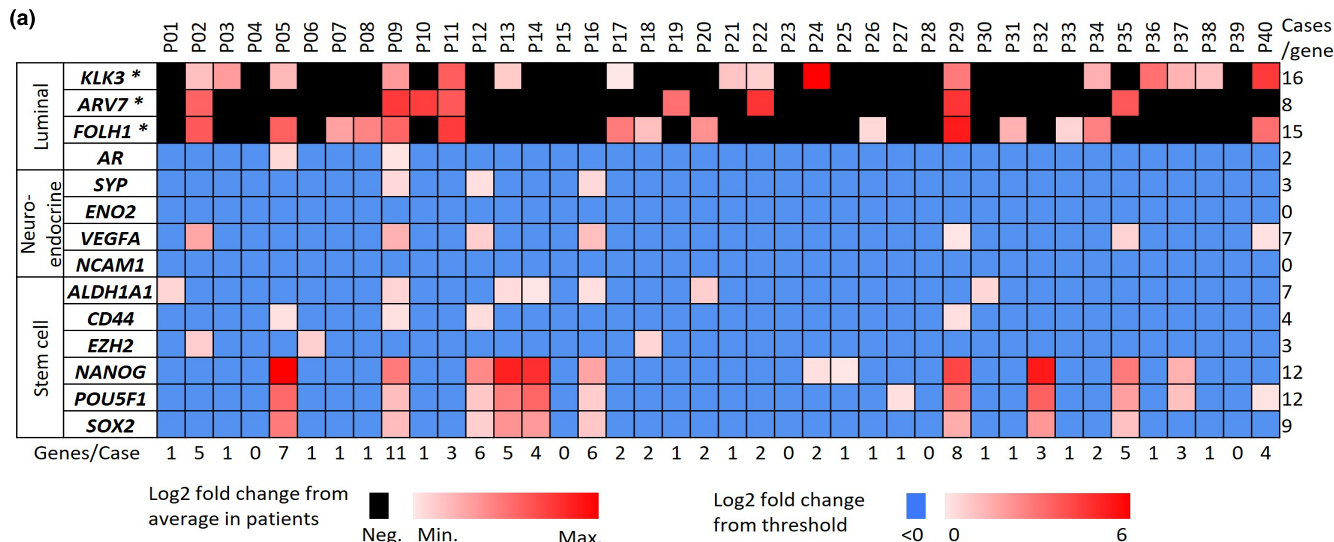
Genes and cell subtypes	Disease progression	Death
KLK3	**2.37 (1.09–5.13), *p* = 0.029**	**3.95 (1.06–14.75), *p* = 0.041**
ARV7	1.56 (0.66–3.70)	1.98 (0.60–6.50)
FOLH1	1.17 (0.55–2.50)	1.99 (0.62–6.37)
AR	2.00 (0.46–8.57)	**60.89 (5.39–688.62), *p* = 0.001**
SYP	1.53 (0.46–5.09)	2.15 (0.26–17.97)
ENO2	N/A	N/A
VEGFA	1.96 (0.82–4.67), *p* = 0.130	**4.44 (1.11–17.80), *p* = 0.035**
ALDH1A1	0.62 (0.21–1.81)	0.85 (0.18–4.07)
CD44	2.12 (0.71–6.29), *p* = 0.177	**7.59 (1.68–34.38), *p* = 0.009**
EZH2	2.06 (0.61–6.98)	1.33 (0.16–11.07)
NANOG	1.74 (0.81–3.73), *p* = 0.156	4.26 (0.95–19.16), *p* = 0.059
POU5F1	1.38 (0.64–2.99)	4.08 (0.90–18.36), *p* = 0.128
SOX2	1.47 (0.64–3.33)	3.22 (0.71–14.52)
Number of genes	1.13 (0.99–1.29), *p* = 0.071	**1.43 (1.13–1.79), *p* = 0.002**
Number of genes excluding *CD44* and *SOX2*	1.16 (0.98–1.38), *p* = 0.078	**1.60 (1.19–2.15), *p* = 0.002**
Number of luminal genes	1.40 (0.99–1.98), *p* = 0.057	**2.16 (1.23–3.80), *p* = 0.008**
Number of neuroendocrine genes	1.41 (0.83–2.40)	2.20 (0.93–5.19), *p* = 0.073
Number of stem cell genes	1.12 (0.91–1.37)	1.41 (0.97–2.04), *p*-0.070
Any luminal gene	1.48 (0.65–3.37)	4.27 (0.54–33.90), *p* = 0.169
Any neuroendocrine gene	1.96 (0.82–4.67), *p* = 0.130	**4.44 (1.11–17.80), *p* = 0.035**
Any stem cell gene	1.13 (0.91–1.40)	**1.50 (1.00–2.25), *p* = 0.048**
Any luminal + any neuroendocrine gene	2.19 (0.80–5.97), *p* = 0.127	**6.81 (1.69–27.52), *p* = 0.007**
Any luminal + any stem cell gene	1.81 (0.83–3.98), *p* = 0.139	**9.62 (1.93–47.86), *p* = 0.006**
Any neuroendocrine + any stem cell gene	1.96 (0.82–4.67), *p* = 0.130	**4.44 (1.11–17.80), *p* = 0.035**
Overexpression of at least one gene in all three groups	2.19 (0.80–5.97), *p* = 0.127	**6.81 (1.69–27.52), *p* = 0.007**
No gene	0.48 (0.11–2.04)	N/A

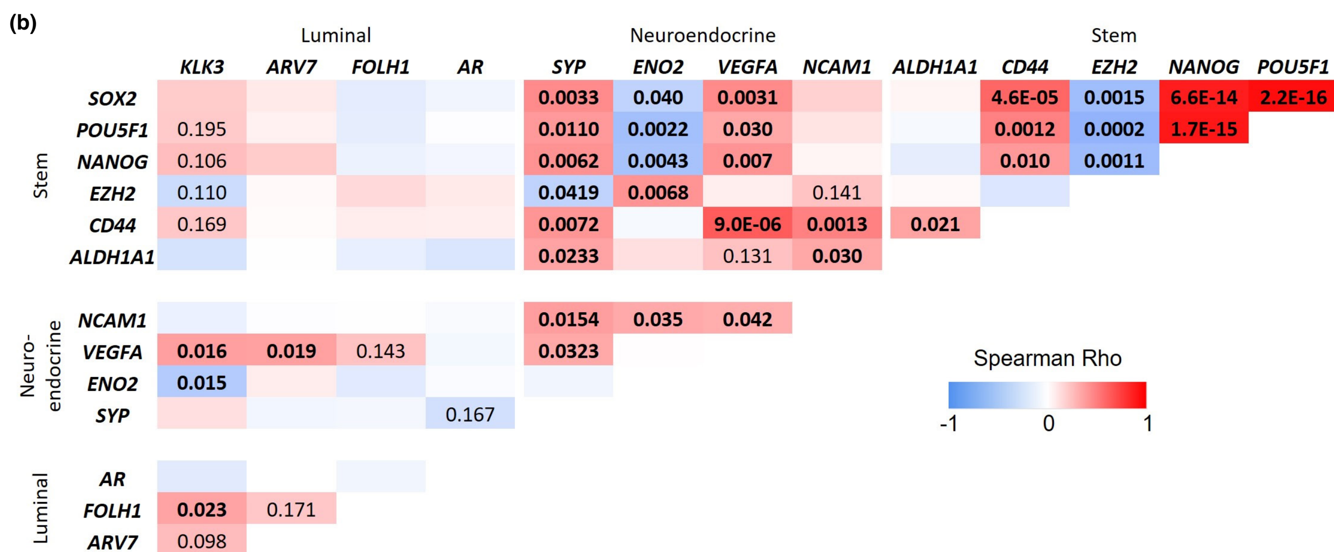
Note: Hazards ratios and 95% confidence intervals for disease progression and death in patients overexpressing the selected circulating genes. Hazards ratios and respective 95% confidence intervals were calculated using Cox proportional hazard models. Numbers are in bold when $p < 0.05$ (p values were added when <0.20).

Abbreviation: N/A, not applicable.

POU5F1 (co-overexpressed in 22.5% of cases; Figure 3a) and *CD44*, whereas *EZH2* was negatively correlated. Interestingly, neuroendocrine and stem cell genes were positively correlated (Figure 3b).

Classifying patients by cell subtypes and combinations thereof highlighted phenotypic diversity in eight subcategories (Figure 3c). In survival analyses (Table 1), patients overexpressing a higher number of genes, regardless of





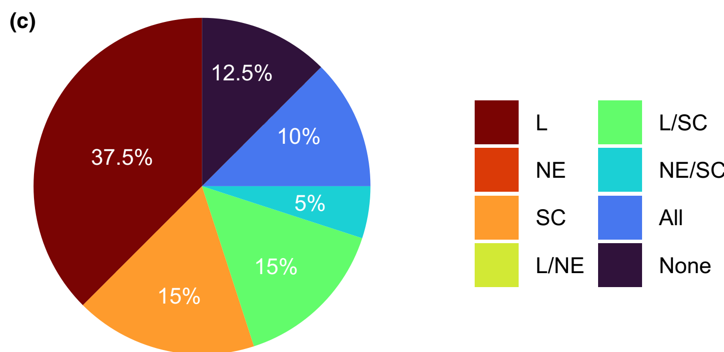


FIGURE 3 Cell-subtype genes in liquid biopsies support phenotypic diversity. (a) Heatmap of gene expression in 40 patients, with black squares representing no signal, blue squares for baseline expression, and white to red squares for overexpression. For *KLK3*, *FOLH1*, and *ARV7* which are completely negative in control samples, results are presented as log2 fold change from the average expression in positive patients. For the remaining genes, results are presented as log2 fold change from the overexpression thresholds calculated from 40 controls, except for *CD44*, whose threshold was established from older male controls. (b) Spearman correlation matrix between the 14 genes in patients with metastatic castration-resistant (mCRPC). Blue and red represent negative and positive associations, respectively. The *p* values are entered when the correlation was significant ($p < 0.05$). (c) Pie chart representing phenotypic diversity by cell subtypes (L: luminal, NE: neuroendocrine, and SC: stem cell) and combinations thereof, as indicated on the right.

TABLE 2 Clinical relevance of circulating genes overexpressed in patients with advanced prostate cancer

Characteristics	Luminal	Neuroendocrine	Stem
Age[a]	0.68 (0.26–1.78)	1.46 (0.21–10.38)	0.88 (0.20–3.80)
PSA[a]	**2.54 (1.03–6.29), $p = 0.044$**	**12.0 (1.10–132.1), $p = 0.042$**	4.00 (0.92–17.44), $p = 0.065$
Bone metastases	3.71 (0.85–16.24), $p = 0.082$	0.91 (0.07–11.54)	1.87 (0.25–13.87)
Visceral metastases	2.01 (0.61–6.62)	N/A	0.88 (0.12–6.18)
ECOG performance score[b]	**3.64 (1.21–10.97), $p = 0.022$**	0.49 (0.03–8.54)	2.22 (0.35–14.21)
Lines of therapy	1.48 (0.97–2.27), $p = 0.071$	2.02 (0.85–4.83), $p = 0.111$	1.66 (0.84–3.31), $p = 0.147$
Progressive disease	**2.99 (1.21–7.39), $p = 0.018$**	6.15 (0.53–70.90), $p = 0.146$	3.00 (0.68–13.32), $p = 0.149$
Initial therapy			
Surgery	1.09 (0.40–2.97)	3.82 (0.54–26.84), $p = 0.178$	**5.20 (1.25–21.69), $p = 0.024$**
Radiotherapy	0.43 (0.09–2.18)	0.83 (0.04–17.15)	1.16 (0.13–10.30)
Non-curative therapy	1.29 (0.49–3.40)	0.27 (0.03–2.20)	**0.19 (0.04–0.79), $p = 0.023$**
Past mCRPC therapy			
Any AR inhibitor	1.64 (0.64–4.21)	**13.3 (1.24–142.3), $p = 0.032$**	**7.77 (1.81–33.38), $p = 0.006$**
Abiraterone	1.46 (0.56–3.81)	**7.88 (1.16–53.55), $p = 0.035$**	**6.14 (1.44–26.10), $p = 0.014$**
Enzalutamide	2.05 (0.46–9.12)	1.18 (0.05–27.66)	1.09 (0.10–12.43)
Any taxane	**2.7 (1.07–6.82), $p = 0.035$**	0.61 (0.08–5.03)	0.96 (0.21–4.39)
Docetaxel	**3.19 (1.27–7.99), $p = 0.013$**	0.71 (0.09–5.88)	0.69 (0.15–3.24)
Cabazitaxel	**5.25 (1.33–20.75), $p = 0.018$**	2.83 (0.19–42.33)	2.87 (0.29–28.53)
Current mCRPC therapy			
None	**4.09 (1.15–14.57), $p = 0.030$**	2.76 (0.19–39.90)	2.50 (0.30–20.58)
Any AR inhibitor	**0.34 (0.14–0.83), $p = 0.018$**	0.31 (0.04–2.51)	0.45 (0.10–2.00)
Abiraterone	0.41 (0.16–1.13), $p = 0.086$	0.61 (0.08–5.03)	0.76 (0.16–3.57)
Enzalutamide	0.45 (0.09–2.20)	N/A	0.24 (0.02–2.77)
Any taxane	1.46 (0.45–4.68)	3.33 (0.44–25.02)	2.66 (0.47–15.16)
Docetaxel	0.36 (0.06–2.27)	2.83 (0.19–42.33)	1.97 (0.19–20.40)
Cabazitaxel	3.89 (0.94–16.10) $p = 0.060$	2.83 (0.19–42.33)	2.87 (0.29–28.54)
PARP inhibitor	1.9 (0.24–15.28)	3.62 (0.09–33.22)	6.23 (0.30–128.61)
Radium-223	3.58 (0.23–56.71)	N/A	N/A

Abbreviations: AR, androgen receptor; ECOG, Eastern Cooperative Oncology Group; mCRPC, metastatic castration-resistant; N/A, not applicable; PSA, prostate specific antigen.

[a]Patients were stratified below and above the median.

[b]Patients were stratified into two groups: ECOG 0–1 and ECOG 2–3. Multilevel, mixed effects logistic regression models were used to calculate odds ratios and their respective 95% confidence intervals, in order to adjust for repeated measurements within individuals. Bold values are significant ($p < 0.05$), and p values <0.2 are indicated. Odds ratios (ORs) with 95% confidence intervals (in parentheses) identifying variables associated with luminal, neuroendocrine, and stem cell markers.

cell subtype, had an increased risk of death. This was also true when excluding *CD44* and *SOX2* which were co-overexpressed with *NANOG* and *POU5F1* in several patients. The total number of luminal genes, overexpression of neuroendocrine or stem cell genes, or overexpression of genes from more than one subtype also influenced overall survival.

Further cell-subtype analyses (Table 2) showed that luminal genes were significantly associated with higher PSA (also true for neuroendocrine), higher ECOG score, and progressive disease at inclusion. Thus, analyses of genes individually (Table 1) or by subtype showed similar results. Patients currently on ARIs were less likely to overexpress luminal genes, whereas patients not on mCRPC therapies were more likely to overexpress these genes. Moreover, patients who previously progressed on ARIs overexpress neuroendocrine or stem cell genes, whereas those who progressed on taxanes overexpressed luminal genes. Stem cell genes were also more common in patients who received surgery as a curative therapy, and less common in those who received noncurative therapies at diagnosis. These findings illustrate the wide diversity of gene patterns in patient's blood, with all cell subtypes being associated with treatments.

Prostate cell-subtype genes are overexpressed in advanced disease and metastases

To relate our findings to phenotypic changes at different stages of PCa, we analyzed our gene panel in five PCa transcriptomic datasets (TCGA, Stanford, Cambridge, MSKCC, and SU2C). Results presented as heatmaps (Figures S4A–E) show that these genes were not commonly overexpressed in primary tumors of hormone-naïve radical prostatectomy (RP) patients but increased with GS and progression, such as in lymph node (LN) metastases collected at RP, primary tumors from hormone-treated cases (TURP/trans-urethral resection of prostate), and in various metastases.

Looking at the number of genes overexpressed overall in primary tumors (Figure 4a–d), we observed that in hormone-naïve RP samples the number of cases overexpressing at least two genes overall increased significantly with GS in all studies, except the MSKCC cohort ($p = 0.061$; Figure 4d). This was also true for the number of luminal or stem cell genes for most studies. Further analyses in TCGA and MSKCC cohorts (Figure S5) showed trends toward predicting earlier biochemical recurrence for the total number of genes overexpressed and for certain subtypes, but significance was lost in multivariable analyses.

Regarding more advanced disease, the likelihood of overexpressing at least two genes overall was significantly higher than in RP samples (Figure 4b–d). LN metastases at RP overexpressed more stem genes than primary tumors (Figure 4b), whereas metastases in the MSKCC cohort showed more cases with luminal and stem genes (Figure 4d). Hormone-treated TURP resembled high GS cases of the hormone-naïve cohort, but with a significant increase in neuroendocrine genes (Figure 4c). Samples with the most genes overexpressed were metastases from mCRPC cases (Figure 4e). Compared to LNs and bones, liver metastases overexpressed more genes overall, more neuroendocrine and stem cell genes, and a trend toward fewer luminal genes. Altogether, an increasing proportion of cases overexpress these genes with further disease progression, whereas gene patterns remain heterogeneous among patients.

This diversity and evolution of gene patterns by cell subtype with progression is illustrated in Figure 5. Of interest, the pattern of circulating genes in the 40 patients with advanced disease in our pilot study resembled the heterogeneity of advanced disease cases, mostly metastatic. The number of cell subtypes represented in each cohort also increased with more aggressive disease and in liver metastases compared to LNs or bones, supporting diversified cell populations in the same tumor. These findings reinforce the relevance of the studied genes in the blood of patients with advanced PCa.

DISCUSSION

This proof-of-concept study demonstrates for the first time the clinical relevance of testing PCa cell-subtype genes in whole blood of patients with advanced disease, setting the stage for studies on larger cohorts to trace circulating gene patterns longitudinally.

Several studies on cancer-related transcripts in blood have focused on CTCs. Despite reports on limitations in isolating such heterogeneous cells, most authors use positive selection-based isolation, often by epithelial cell adhesion molecule or CKs, which are not present on all CTCs, or size-based methods which may miss small CTCs.[9–11,15] Working in blood RNA avoids selecting only specific CTC subsets, while also detecting cell-free RNAs in EVs.[35] Cancer-specific genes were identified in whole blood.[16–21,36–41] The sensitivity of RT-qPCR assays comparing gene expression directly in the blood of patients with PCa versus CTCs was also reported to be similar.[18,42] The central limitation remains the genes expressed in WBCs, making gene selection a key step for success. Tracing immune signatures in parallel may provide useful information but was beyond the scope of this study.[43]

To establish the reliability and specificity of our assays, we first tested each gene in cell lines and spiked control blood. Luminal genes were specific to luminal-like cells, whereas neuroendocrine and stem cell genes were highest in the neuroendocrine NCI-H660 cells. In spiking experiments, luminal genes were negative in control blood, except for *AR* as previously reported.[36,44] All neuroendocrine and stem genes were detected at baseline levels. Similar expression patterns were also observed using probe-based TaqMan assays (not shown), implying that baseline expression was not due to nonspecific amplification. Overall, expression of all genes increased upon spiking with a low number of PCa cells, suggesting their potential detection in patients' blood. *CHGA*, *VIM*, *EPCAM*, *PSCA*, *CD133*, and *ARV12* were considered but excluded due to issues of primer specificity and limitation in RNA quantity from banked blood.

The varying proportions of different WBC lineages in blood of patients showed mostly no or low correlation, except for *CD44* and *SYP*. However, not all patients with high levels of the relevant WBCs showed these genes overexpressed. *CD44* and *SYP* were always overexpressed with other genes of the same cell subtype. These observations highlight the importance of including blood RNA from healthy donors to account for variations in WBCs and

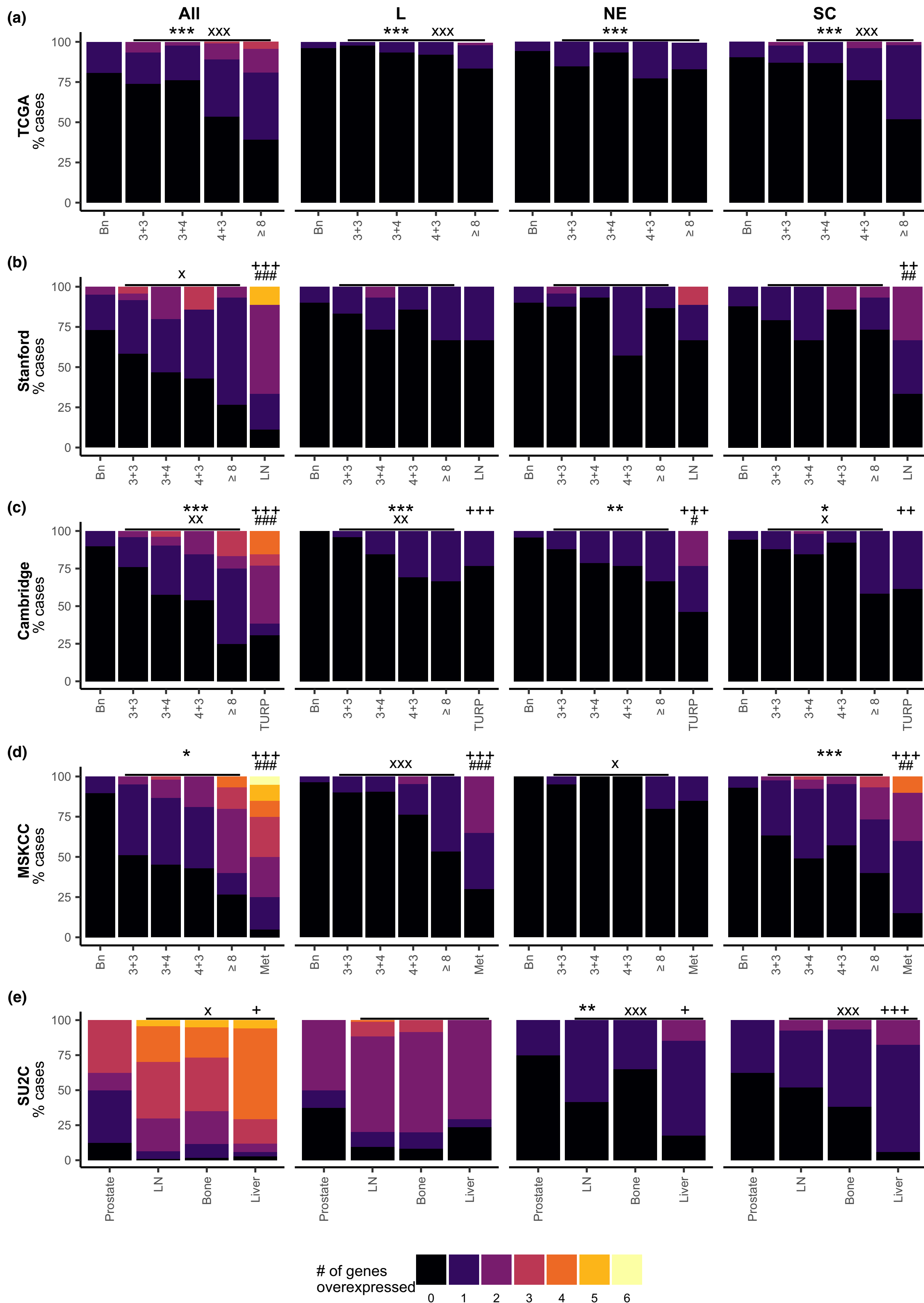

(a)
All
L
NE
SC
TCGA
% cases
(b)
Stanford
% cases
(c)
Cambridge
% cases
(d)
MSKCC
% cases
(e)
SU2C
% cases
Bn
3+3
3+4
4+3
≥ 8
LN
TURP
Met
Prostate
LN
Bone
Liver
# of genes
overexpressed
0
1
2
3
4
5
6

FIGURE 4 Prostate cell-subtype specific genes are overexpressed in PCa tissues of patients at different stages of disease. Stacked bar graphs representing the number of genes overexpressed (color-coded) overall or by cell subtype in different categories of patients from the (a) TCGA, (b) Stanford, (c) Cambridge, (d) MSKCC, and (e) SU2C datasets. Overexpression was defined as +2.58 SDs over the mean expression in benign samples from the same cohort, as indicated in the Methods section. For the SU2C dataset, gene expression was compared to benign samples from the TCGA dataset. (a–e) Denoted above each graph are the Fisher's exact test results comparing the number of genes overexpressed by categories of cases (0 vs. ≥1 gene overexpressed for individual subtypes, or 0–1 versus ≥2 genes overexpressed overall, up to 6 genes). (a–d) Significant differences are shown between benign versus RP (denoted by *), benign versus advanced cases from the same cohort (LN, TURP, metastases; denoted by +), or RP versus advanced cases (#), as well as the Cochrane-Armitage test for trends comparing RP cases by Gleason score (×). (e) In the SU2C dataset, Fisher's exact test results denote difference between LNs versus bones (*), LNs versus liver (×), and bones versus liver metastases (+). (a–e) The level of significance is denoted as: * +/ # / × for $p<0.05$; ** for $p<0.01$; *** for $p<0.001$. Bn, benign; L, luminal; LN, lymph node; Met, metastasis; MSKCC, Memorial Sloan Kettering Cancer Center; NE, neuroendocrine; PCa, prostate cancer; RP, radical prostatectomy; SC, stem cell; TCGA, The Cancer Genome Atlas; TURP, primary tumors removed by transurethral resection from hormone-treated patients.

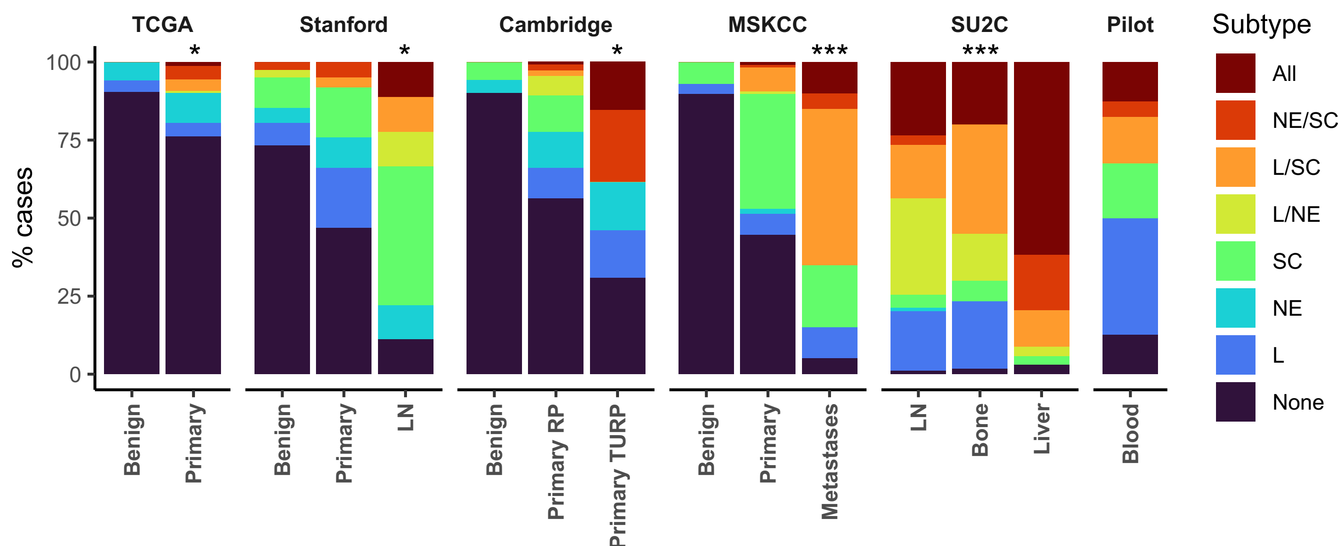


FIGURE 5 Cell subtype patterns in PCa tissues and metastases evolve with disease progression. Stacked bar graphs of cell-subtypes (L: luminal, NE: neuroendocrine, SC: stem cell) and combinations thereof (color-coded on the right) for different categories of patients in the TCGA, Stanford, Cambridge, MSKCC, and SU2C datasets. The number of cell subtypes represented in different categories of patients are compared by Fisher's exact test. Significance levels are represented by * for $p<0.05$; ** for $p<0.01$; *** for $p<0.001$ for benign versus primary (TCGA), primary versus advanced (LN/TURP/Metastases in Stanford/Cambridge/MSKCC), and between metastatic sites (SU2C). The results in blood RNA from our pilot study were added on the right for comparison purposes, mostly resembling metastases and differing from primary and benign prostate tissues. LN, lymph node; MSKCC, Memorial Sloan Kettering Cancer Center; PCa, prostate cancer; RP, radical prostatectomy; TCGA, The Cancer Genome Atlas; TURP, primary tumors removed by transurethral resection from hormone-treated patients.

accurately identify overexpression related to the patient's cancer by establishing stringent thresholds. For cases where overexpression is close to the threshold, we would recommend testing a second blood sample to assess clinical value.

We observed that 87.5% of patients with CRPC (35/40) overexpressed one or more of the target genes in their blood. Cases with no genes overexpressed may be explained by a low number of CTCs and cancer-derived EVs due to less advanced disease or successful therapeutic elimination of cancer cells. RNA quality or quantity was not an issue, being similar across samples. Expanding our gene panel may lead to significant findings in these patients as well. Overall, patients showed a wide phenotypic diversity in line with the clinical heterogeneity of advanced cases and reported for circulating genes in breast, colorectal, lung, and prostate cancers.[16–21,37–39]

Among luminal genes, *KLK3* expression was significantly associated with high levels of PSA in blood, suggesting that PSA reflects tumor burden, whereas circulating *KLK3* would originate from CTCs or EVs. Of note, *KLK3* expression in 16 of 40 (40%) of mCRPC cases is not paralleled by *AR* overexpression. Testing *AR* amplification in circulating tumor DNA, expression of other *ARVs*, or additional AR target genes may be alternatives to assessing AR reactivation. *ARV7* was found in eight of 40 (20%) of cases, in line with reported values ranging from 11–68%.[16–18,36,45] This is clinically meaningful because taxanes may be more effective than ARIs in such patients.[18,46,47] Other *AR*

variants would be interesting to add, given their clinical relevance.[18,36]

Among stem cell genes, *NANOG*, *POU5F1*, and *SOX2* were often, but not always, co-overexpressed. These genes have been studied in whole blood of small cell lung cancer, where *SOX2* was identified as a promising prognostic marker.[39] In our PCa series, circulating *NANOG* and *POU5F1* were more often overexpressed than *SOX2*.

Neuroendocrine genes were infrequent and always found with other subtypes, consistent with the rarity of pure neuroendocrine prostate tumors.[4] *VEGFA* is particularly interesting as a neuro-product promoting metastasis by enhancing PCa cell motility and being targetable in advanced disease.[33,48] *ENO2* was not overexpressed. *NCAM1* was not overexpressed in patients' blood, possibly due to the diversity of splice variants.[49] Nevertheless, these neuroendocrine genes deserve attention as CD56+ (*NCAM1*) and SYP+ CTCs were reported in patients with PCa.[14,50]

Although patterns of individual genes are interesting, they are each overexpressed in a minority of patients, making it difficult to interpret gene-by-gene analysis results. For this reason, we consider analyses by cell subtype more relevant and informative. The contribution of different cell subtypes to tumor development and metastatic progression is important in the context of therapeutic resistance. This was corroborated in our series, despite the low number of patients. Patients treated with ARIs expressed fewer circulating luminal genes, whereas those previously treated with ARIs overexpressed several stem and neuroendocrine genes. This is in line with hormonal therapies targeting differentiated AR-positive luminal cells and promoting neuroendocrine differentiation and stemness features.[4–7] Moreover, the number of genes and cell subtypes represented were associated with an increased risk of death. This suggests less advanced disease or better disease control in patients expressing fewer genes or only luminal genes, whereas patients with more cell subtypes may be more difficult to treat with single agent therapies. Further studies are needed to evaluate whether this gene panel is of prognostic utility in patients with mCRPC.

Bioinformatic analyses in PCa transcriptomic datasets validate our choice of genes and support our clinical findings. Results were consistent among RP cohorts, showing few cases overexpressing the selected genes. They were found more often in advanced cases, especially in mCRPC. Moreover, cell-subtype patterns differed between categories of cases, as well as by metastatic site. Furthermore, the circulating gene patterns of our 40 patients with advanced PCa resemble genes patterns in metastases. Altogether, these findings establish that genes representing epithelial cell subtypes contribute to progression, emphasizing the increasing importance of tumor cell diversity in patients with PCa.

The main limitation of this pilot study remains that patients were at various stages of treatments, although they were all CRPC. Comparing circulating gene patterns with the corresponding transcriptomic data of primary tumors may be informative, and more so if adjacent LN metastases are harvested. Unfortunately, these advanced cases were not necessarily treated at our institutions early in their trajectory and no tissue samples were available for this study. Nonetheless, the gene patterns defined in blood are related to patients' characteristics, previous treatments, and treatment response at time of blood draw, despite the low number of cases tested. This justifies further validation in larger, independent cohorts before translation in clinical practice. We anticipate conducting a study with serial blood sampling to demonstrate whether assessing prostate cell subtypes through liquid biopsies will predict therapeutic response and clinical outcome.

CONCLUSION

This pilot study supports the premise that prostate cell-subtype genes are traceable in the blood of patients with advanced disease. The diversity among patients is relevant to the clinical heterogeneity of PCa. Our findings revealed associations of cell subtype genes in patients' blood with therapeutic regimens and clinical outcome. Bioinformatic analyses in transcriptomic data from over 1000 patients confirmed that our selection of genes and overexpression patterns in the blood of 35 of 40 patients are best represented in metastases, thereby reinforcing the clinical relevance of our findings. Further studies, particularly longitudinal assessments, will determine whether this approach can guide individualized treatments.

AUTHORS CONTRIBUTIONS

S.C., S.D., Q.V., M.W., and A.A. wrote the manuscript. S.C. designed the research. S.D., Q.V., and L.H. performed the research. S.D., Q.V., M.W., M.V., C.F., and F.B. analyzed the data. N.C. contributed reagents/analytical tools.

ACKNOWLEDGEMENTS

The authors thank the patients and volunteers who generously accepted to donate blood for this proof-of-concept study. They acknowledge Dr. Nadia Boufaied and Mr. Karl-Philippe Guérard for helpful discussions in optimizing assays. Initial attempts to test circulating genes in patients' liquid biopsies were possible through a Movember Award via Prostate Cancer Canada to S.C. This research was supported by funds from the Urology division (A.A.) and Urology Research fellowship (from F.B. and A.A.) to Q.V., CIHR fellowship to M.W. and Experimental Surgery and Prostate Cancer studentships to S.D. Part of the bioinformatic analyses were enabled through Calcul Québec (www.calculquebec.ca) and

Compute Canada (www.computecanada.ca). The authors acknowledge Stanford, MSKCC, and Cambridge team members for making PCa datasets publicly available and SU2C and TCGA teams for granting access to their datasets. Our findings are in part based upon data generated by TCGA, managed by the NCI and NHGRI. Information about TCGA can be found at www.cancergenome.nih.gov. SU2C is a program of the Entertainment Industry Foundation administered by the American Association for Cancer Research.

FUNDING INFORMATION

MUHC Prostate Cancer Research funds.

CONFLICT OF INTEREST

The authors declared no competing interests for this work.

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209-249. doi:10.3322/caac.21660
2. Crona DJ, Whang YE. Androgen receptor-dependent and -independent mechanisms involved in prostate cancer therapy resistance. *Cancer*. 2017;9(6):67. doi:10.3390/cancers9060067
3. Ojo D, Lin X, Wong N, Gu Y, Tang D. Prostate cancer stem-like cells contribute to the development of castration-resistant prostate cancer. *Cancer*. 2015;7(4):2290-2308. doi:10.3390/cancers7040890
4. Li Q, Zhang CS, Zhang Y. Molecular aspects of prostate cancer with neuroendocrine differentiation. *Chin J Cancer Res*. 2016;28(1):122-129. doi:10.3978/j.issn.1000-9604.2016.01.02
5. Patel GK, Chugh N, Tripathi M. Neuroendocrine differentiation of prostate cancer-an intriguing example of tumor evolution at play. *Cancer*. 2019;11(10):1405. doi:10.3390/cancers11101405
6. Ware KE, Garcia-Blanco MA, Armstrong AJ, Dehm SM. Biologic and clinical significance of androgen receptor variants in castration resistant prostate cancer. *Endocr Relat Cancer*. 2014;21(4):T87-T103. doi:10.1530/ERC-13-0470
7. Kong D, Sethi S, Li Y, et al. Androgen receptor splice variants contribute to prostate cancer aggressiveness through induction of EMT and expression of stem cell marker genes. *Prostate*. 2015;75(2):161-174. doi:10.1002/pros.22901
8. Underwood JJ, Quadri RS, Kalva SP, et al. Liquid biopsy for cancer: review and implications for the radiologist. *Radiology*. 2019;294(1):5-17. doi:10.1148/radiol.2019182584
9. Morrison GJ, Goldkorn A. Development and application of liquid biopsies in metastatic prostate cancer. *Curr Oncol Rep*. 2018;20(4):35. doi:10.1007/s11912-018-0683-0
10. Gabriel MT, Calleja LR, Chalopin A, Ory B, Heymann D. Circulating tumor cells: a review of non-EpCAM-based approaches for cell enrichment and isolation. *Clin Chem*. 2016;62(4):571-581. doi:10.1373/clinchem.2015.249706
11. Thiele JA, Bethel K, Králíčková M, Kuhn P. Circulating tumor cells: fluid surrogates of solid tumors. *Ann Rev Pathol: Mechan Dis*. 2017;12(1):419-447. doi:10.1146/annurev-pathol-052016-100256
12. Lampignano R, Schneck H, Neumann M, Fehm T, Neubauer H. Enrichment, isolation and molecular characterization of epcam-negative circulating tumor cells. In: Magbanua MJM, Park JW, eds. *Isolation and Molecular Characterization of Circulating Tumor Cells*. Springer International Publishing; 2017:181-203. doi:10.1007/978-3-319-55947-6_10
13. Xu L, Mao X, Imrali A, et al. Optimization and evaluation of a novel size based circulating tumor cell isolation system. *PLOS ONE*. 2015;10(9):e0138032. doi:10.1371/journal.pone.0138032
14. Beltran H, Jendrisak A, Landers M, et al. The initial detection and partial characterization of circulating tumor cells in neuroendocrine prostate cancer. *Clin Cancer Res*. 2016;22(6):1510-1519. doi:10.1158/1078-0432.CCR-15-0137
15. McDaniel AS, Ferraldeschi R, Krupa R, et al. Phenotypic diversity of circulating tumour cells in patients with metastatic castration-resistant prostate cancer. *BJU Int*. 2017;120(5B):E30-E44. doi:10.1111/bju.13631
16. Todenhöfer T, Azad A, Stewart C, et al. AR-V7 transcripts in whole blood RNA of patients with metastatic castration resistant prostate cancer correlate with response to abiraterone acetate. *J Urol*. 2017;197(1):135-142. doi:10.1016/j.juro.2016.06.094
17. To SQ, Kwan EM, Fettke HC, et al. Expression of androgen receptor splice variant 7 or 9 in whole blood does not predict response to androgen-axis-targeting agents in metastatic castration-resistant prostate cancer. *Eur Urol*. 2018;73(6):818-821. doi:10.1016/j.eururo.2018.01.007
18. Liu X, Ledet E, Li D, et al. A whole blood assay for AR-V7 and AR(v567es) in patients with prostate cancer. *J Urol*. 2016;196(6):1758-1763. doi:10.1016/j.juro.2016.06.095
19. Yates DR, Rouprêt M, Drouin SJ, et al. Quantitative RT-PCR analysis of PSA and prostate-specific membrane antigen mRNA to detect circulating tumor cells improves recurrence-free survival nomogram prediction after radical prostatectomy. *Prostate*. 2012;72(12):1382-1388. doi:10.1002/pros.22488
20. Pipinikas CP, Carter ND, Corbishley CM, Fenske CD. HIF-1α mRNA gene expression levels in improved diagnosis of early stages of prostate cancer. *Biomarkers*. 2008;13(7–8):680-691. doi:10.1080/13547500802591992
21. Danila DC, Anand A, Schultz N, et al. Analytic and clinical validation of a prostate cancer–enhanced messenger RNA detection assay in whole blood as a prognostic biomarker for survival. *Eur Urol*. 2014;65(6):1191-1197. doi:10.1016/j.eururo.2013.07.006
22. Liu LL, Xie N, Sun S, Plymate S, Mostaghel E, Dong X. Mechanisms of the androgen receptor splicing in prostate cancer cells. *Oncogene*. 2014;33(24):3140-3150. doi:10.1038/onc.2013.284
23. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinform*. 2012;13:134. doi:10.1186/1471-2105-13-134
24. Falkenberg VR, Whistler T, Murray J'R, Unger ER, Rajeevan MS. Identification of Phosphoglycerate Kinase 1 (PGK1) as a reference gene for quantitative gene expression measurements in human blood RNA. *BMC Res Notes*. 2011;4:324. doi:10.1186/1756-0500-4-324
25. Vandesompele J, de Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;3(7):research0034.1. doi:10.1186/gb-2002-3-7-research0034

26. Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*. 2012;483(7391):603-607. doi:10.1038/nature11003
27. Uhlen M, Zhang C, Lee S, et al. A pathology atlas of the human cancer transcriptome. *Science*. 2017;357(6352):eaan2507. doi:10.1126/science.aan2507
28. Lapointe J, Li C, Higgins JP, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A*. 2004;101(3):811-816. doi:10.1073/pnas.0304146101
29. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47. doi:10.1093/nar/gkv007
30. Bourgey M, Dali R, Eveleigh R, et al. GenPipes: an open-source framework for distributed and scalable genomic analyses. *GigaScience*. 2019;8(6):giz037. doi:10.1093/gigascience/giz037
31. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8
32. Matusik RJ, Jin RJ, Sun Q, et al. Prostate epithelial cell fate. *Differentiation*. 2008;76(6):682-698. doi:10.1111/j.1432-0436.2008.00276.x
33. Chevalier S, Defoy I, Lacoste J, et al. Vascular endothelial growth factor and signaling in the prostate: more than angiogenesis. *Mol Cell Endocrinol*. 2002;189(1–2):169-179. doi:10.1016/s0303-7207(01)00728-6
34. Zhang K, Zhou S, Wang L, et al. Current stem cell biomarkers and their functional mechanisms in prostate cancer. *Int J Mol Sci*. 2016;17(7). doi:10.3390/ijms17071163
35. Mugoni V, Ciani Y, Nardella C, Demichelis F. Circulating RNAs in prostate cancer patients. *Cancer Lett*. 2022;524:57-69. doi:10.1016/j.canlet.2021.10.011
36. Stuopelyte K, Sabaliauskaite R, Bakavicius A, et al. Analysis of AR-FL and AR-V1 in whole blood of patients with castration resistant prostate cancer as a tool for predicting response to abiraterone acetate. *J Urol*. 2020;204(1):71-78. doi:10.1097/JU.0000000000000803
37. Mohammadi P, Saidijam M, Kaki A, Etemadi K, Shabab N, Yadegarazari R. A pilot study of CK19, CK20 and GCC mRNA in the peripheral blood as a colorectal cancer biomarker panel. *Int J Mol Cell Med*. 2016;5(1):30-36.
38. Andergassen U, Hofmann S, Kölbl A, et al. Detection of tumor cell-specific mRNA in the peripheral blood of patients with breast cancer—evaluation of several markers with real-time reverse transcription-PCR. *Int J Mol Sci*. 2013;14(1):1093-1104. doi:10.3390/ijms14011093
39. Sodja E, Rijavec M, Koren A, Sadikov A, Korošec P, Cufer T. The prognostic value of whole blood SOX2, NANOG and OCT4 mRNA expression in advanced small-cell lung cancer. *Radiol Oncol*. 2016;50(2):188-196. doi:10.1515/raon-2015-0027
40. Modlin IM, Kidd M, Malczewska A, et al. The NETest: the clinical utility of multigene blood analysis in the diagnosis and management of neuroendocrine tumors. *Endocrinol Metab Clin North Am*. 2018;47(3):485-504. doi:10.1016/j.ecl.2018.05.002
41. Kwan EM, Fettke H, Docanto MM, et al. Prognostic utility of a whole-blood androgen receptor-based gene signature in metastatic castration-resistant prostate cancer. *Eur Urol Focus*. 2021;7(1):63-70. doi:10.1016/j.euf.2019.04.020
42. Danila DC, Samoila A, Patel C, et al. Clinical validity of detecting circulating tumor cells by adnatest assay compared with direct detection of tumor mRNA in stabilized whole blood, as a biomarker predicting overall survival for metastatic castration-resistant prostate cancer patients. *Cancer J (Sudbury, Mass)*. 2016;22(5):315-320. doi:10.1097/PPO.0000000000000220
43. Ross RW, Galsky MD, Scher HI, et al. A whole-blood RNA transcript-based prognostic model in men with castration-resistant prostate cancer: a prospective study. *Lancet Oncol*. 2012;13(11):1105-1113. doi:10.1016/S1470-2045(12)70263-2
44. Marín-Aguilera M, Jiménez N, Reig Ò, et al. Androgen receptor and its splicing variant 7 expression in peripheral blood mononuclear cells and in circulating tumor cells in metastatic castration-resistant prostate cancer. *Cell*. 2020;9(1):203. doi:10.3390/cells9010203
45. Nakazawa M, Lu C, Chen Y, et al. Serial blood-based analysis of AR-V7 in men with advanced prostate cancer. *Annal Oncol*. 2015;26(9):1859-1865. doi:10.1093/annonc/mdv282
46. Antonarakis ES, Lu C, Luber B, et al. Androgen receptor splice variant 7 and efficacy of taxane chemotherapy in patients with metastatic castration-resistant prostate cancer. *JAMA Oncol*. 2015;1(5):582-591. doi:10.1001/jamaoncol.2015.1341
47. Markowski MC, Silberstein JL, Eshleman JR, Eisenberger MA, Luo J, Antonarakis ES. Clinical utility of CLIA-grade AR-V7 testing in patients with metastatic castration-resistant prostate cancer. *JCO Precis Oncol*. 2017;2017:PO.17.00127. doi:10.1200/PO.17.00127
48. Formica V, Roselli M. Issues and promises of bevacizumab in prostate cancer treatment AU – Cereda, Vittore. *Exp Opin Biol Ther*. 2018;18(6):707-717. doi:10.1080/14712598.2018.1479737
49. Hale JS, Li M, Lathia JD. The malignant social network. *Cell Adh Migr*. 2012;6(4):346-355. doi:10.4161/cam.21294
50. Pal SK, He M, Chen L, et al. Synaptophysin expression on circulating tumor cells in patients with castration resistant prostate cancer undergoing treatment with abiraterone acetate or enzalutamide. *Urol Oncol*. 2018;36(4):162.e1-162.e6. doi:10.1016/j.urolonc.2017.12.006

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Derderian S, Vesval Q, Wissing MD, et al. Liquid biopsy-based targeted gene screening highlights tumor cell subtypes in patients with advanced prostate cancer. *Clin Transl Sci*. 2022;15:2597-2612. doi:10.1111/cts.13372