

ARTICLE



Molecular Diagnostics

A DNA copy number alteration classifier as a prognostic tool for prostate cancer patients

Walead Ebrahimizadeh^{1,8}, Karl-Philippe Guérard¹, Shaghayegh Rouzbeh¹, Eleonora Scarlata¹, Fadi Brimo², Palak G. Patel³, Tamara Jamaspishvili^{3,4}, Lucie Hamel¹, Armen G. Aprikian¹, Anna Y. Lee⁵, David M. Berman³, John M. S. Bartlett^{5,6,7}, Simone Chevalier¹ and Jacques Lapointe¹

© The Author(s), under exclusive licence to Springer Nature Limited 2023

BACKGROUND: Distinguishing between true indolent and potentially life-threatening prostate cancer is challenging in tumours displaying clinicopathologic features associated with low or intermediate risk of relapse. Several somatic DNA copy number alterations (CNAs) have been identified as potential prognostic biomarkers, but the standard cytogenetic method to assess them has a limited multiplexing capability.

METHODS: Multiplex ligation-dependent probe amplification (MLPA) targeting 14 genes was optimised to survey 448 tumours of patients with low or intermediate risk (Grade Group 1–3, Gleason score ≤ 7) who underwent radical prostatectomy. A 6-gene CNA classifier was developed using random survival forest and Cox proportional hazard modelling to predict biochemical recurrence.

RESULTS: The classifier score was significantly associated with biochemical recurrence after adjusting for standard clinicopathologic variables and the known prognostic index CAPRA-S score with a hazard ratio of 2.17 and 1.80, respectively ($n = 406$, $P < 0.01$). The prognostic value of this classifier was externally validated in published CNA data from three radical prostatectomy cohorts and one radiation therapy pre-treatment biopsy cohort.

CONCLUSION: The 6-gene CNA classifier generated by a single MLPA assay compatible with the small quantities of DNA extracted from formalin-fixed paraffin-embedded (FFPE) tissue specimens has the potential to improve the clinical management of patients with low or intermediate risk disease.

British Journal of Cancer (2023) 128:2165–2174; <https://doi.org/10.1038/s41416-023-02236-8>

INTRODUCTION

Prostate cancer (PCa) affected 1.4 million men worldwide and caused 375,000 deaths in 2020 [1]. While recent data indicate that the incidence of high risk PCa is rising, likely due to the discontinuation of routine prostate-specific antigen (PSA) screening [2], most prostate tumours are still diagnosed at the localised stage, with Grade Group (GG) 1–3—Gleason score (GS) ≤ 7 —and a low-intermediate risk of treatment failure, as indicated by biochemical recurrence (BCR) after radical prostatectomy (RP) or radiotherapy. Out of concerns about the overtreatment of indolent tumours, active surveillance (AS) is now proposed to some low and favourable intermediate risk patients as an alternative to immediate definitive treatment [3]. While clinicopathologic parameters usually identify patients at high risk of tumour recurrence, distinguishing the most suitable patients for AS from those who would require immediate curative treatments and, perhaps, may benefit from adjuvant therapies is challenging in tumours displaying clinicopathologic features associated with low or intermediate risk of relapse. Thus,

there is an unmet need for clinically applicable molecular biomarkers that would aid the decision-making process particularly in the post-widespread PSA screening era.

DNA copy number alterations (CNAs) targeting oncogenes (gain) and tumour suppressors (deletion) drive PCa tumorigenesis and increase in frequency as the disease progresses [4]. Several CNAs such as *PTEN* deletion have been identified as potential prognostic biomarkers independent of clinicopathologic parameters [5, 6]. Combinations of specific CNAs, as well as global CNA burden (percentage of tumour genome affected by CNA), have shown to further improve risk stratification [5, 7–9]. However, fluorescence in situ hybridisation (FISH), the gold standard method to assess CNAs, has limited multiplexing capability and genome-wide approaches remain expensive and difficult to implement outside academic centres. There is therefore a need to optimise the selection of CNAs for prognostic purposes and their assessment method that would ensure the integration of CNA classifiers into routine clinical settings.

¹Department of Surgery, Division of Urology, McGill University and the Research Institute of the McGill University Health Centre (RI MUHC), Montreal, QC, Canada. ²Department of Pathology, McGill University and the Research Institute of the McGill University Health Centre (RI MUHC), Montreal, QC, Canada. ³Department of Pathology & Molecular Medicine, Queen's University, Kingston, ON, Canada. ⁴Department of Pathology & Laboratory Medicine, SUNY Upstate Medical University, Syracuse, NY, USA. ⁵Ontario Institute for Cancer Research, Toronto, ON, Canada. ⁶Laboratory Medicine & Pathobiology, University of Toronto, Toronto, ON, Canada. ⁷Cancer Research UK Edinburgh Centre, Institute of Genetics and Cancer, The University of Edinburgh, Edinburgh, UK. ⁸Present address: Current affiliation: IMV Inc., Dartmouth, Canada. ✉email: jacques.lapointe@mcgill.ca

Multiplex ligation-dependent probe amplification (MLPA) is a PCR-based method of CNA quantification [10] allowing simultaneous assessment of up to 50 different loci, compatible with small quantities of formalin-fixed paraffin-embedded (FFPE) extracted DNA (50 ng). We recently developed and optimised an MLPA-based assay for profiling 10 CNAs [11] either known to be associated with PCa clinical outcome or with the potential to improve risk stratification. The objective of the current study was to assess CNA profile of PCa samples using MLPA and test whether a combination of CNAs was associated with poor clinical outcome. Here, we describe the expansion and application of the MLPA assay in an RP cohort of patients with low or intermediate risk disease GG 1–3 ($GS \leq 7$). We developed and validated a CNA classifier, which predicts BCR independently of clinicopathologic features and, as a result, improved overall risk stratification.

MATERIALS AND METHODS

Patients and samples

The study was conducted in accordance with the REMARK guidelines [12]. The RP cohort consisted of formalin-fixed paraffin-embedded (FFPE) RP specimens with clinical correlates that were collected between 1996 and 2013 at McGill and Queen's University Health Centres. After slide review by pathologists (FB and DMB) to identify the cancer area and assign the final Gleason grade according to the International Society of Urological Pathology/World Health Organization recommendations [13], 448 cases GG 1–3 ($GS \leq 7$) with no prior treatment were included in the study approved by the Research Ethics Board of McGill University Health Centre (Quebec, Canada, BDM-10-115) amended to include samples from Kingston General Hospital collected with the approval of the Queen's University Research Ethics Board (Ontario, Canada, PATH-144-14). The sample size was considered adequate based on our previous biomarker studies [5, 14]. Biopsy slides were available for 286 of these cases and were reviewed for Gleason grading by pathologists (FB and DMB). Two tumour areas representing the highest (A sample) and lowest (B sample) Gleason grade patterns were cored from the RP tissue blocks followed by DNA extraction as previously reported [11]. In GG 1 (3 + 3) cases, two samples were also cored and assigned to A or B randomly. DNA was successfully extracted from both A and B samples, from sample A only, and from sample B only in 317, 76 and 55 cases, respectively. DNA from an independent set of 13 PCa samples described previously [11] was used to test the expanded MLPA assay before profiling the cohort.

Expansion of the MLPA assay

Further literature search and data mining of our previous PCa CNA profile data [4] to identify potential biomarkers has led to the design of four additional MLPA probes targeting *RWDD3*, *PDZD2*, *GTF2H2* and *WRN*, as well as the new reference gene, *DHRS4L2* as reported previously [11]. These probes were added to the original 10-gene PCa probe mix [11]. The expanded MLPA probe set (14 target genes plus 10 reference loci) used for this study is shown in Supplementary Table S1. The sensitivity and specificity of the additional CNA targeting probes were assessed by comparing the MLPA results with fluorescence in situ hybridisation (FISH) done on a tissue microarray (TMA) representing regions adjacent to those used for DNA extraction [11]. The following BAC clones: RP11–335D10 (1p21.3), RP11–437P15 (5p13.3), RP11–195E2 (5q13.2) and RP11–363L24 (8p12) were labelled with spectrum Orange dUTP (Enzo Life Science) to probe *RWDD3*, *PDZD2*, *GTF2H2* and *WRN*, respectively. Spectrum green labelled chromosome 1 subtelomere probe (Cytocell), RP11–530D2 BAC (5p12), and CEP8 (Abbot Molecular) were used as control reference probes where appropriate. The FISH procedure, including probe labelling and data analysis, was done as previously reported [11]. Images were acquired using an Olympus IX-81 inverted microscope at 96X magnification and Image-Pro Plus 7.0 software (Media Cybernetics).

MLPA analysis

The MLPA analysis was performed blindly to the clinical outcome. Normal control DNA consisted of healthy female genome (Promega) along with DNA extracted from FFPE kidney tissue (McGill University Health Centre) and breast lymph node (Ontario Institute for Cancer Research). The PC-3 cell line (ATCC Cat# CRL-1435) genome served as a positive control and a no-DNA reaction served as a negative control. MLPA was performed on all

DNA samples in duplicate for a total of 22 batches with 50 ng of DNA per reaction as described before [11] according to the manufacturer (MRC Holland) guidelines using MLPA One-Tube general protocol and EK5-FAM kit. MLPA data were processed with Coffalyser software (Version 140721.1958) and the P.I.N.P.2 protocol as previously reported [11]. Quality control (QC) failure of a MLPA reaction was defined when a standard deviation (SD) of more than 0.1 was computed in more than four probes by the Coffalyser software or in the absence of a fluorescent signal in the reaction. Intra-sample normalisation consisted of computing the median value of the test probe divided by each of the reference probes, while inter-sample normalisation consisted of computing the average value of the intra-sample normalised probe signal of the test sample divided by the signal of the three reference DNA samples separately. If the 95% confidence interval of the probe for a test sample was above or below the interval for a reference sample, a gain or deletion call respectively was made, otherwise a normal copy number was assigned to that probe. CNA in a gene was thus defined when both duplicate reactions showed the same CNA in at least one of the probes targeting the gene and relative to at least two of the three reference samples. Based on our previous report showing a suboptimal performance of the *PDPK1* Exon 14 targeting probe [11], only the intron 10 probe was considered for calling a CNA for this gene. MLPA calls for gain of *MYC*, *PDPK1* and *PDZD2* as well as deletion of the remaining 11 targeted genes were considered as CNAs in the analyses. Unless otherwise specified, the patient tumour CNA profile combines, for each locus, the CNAs status of both A and B tumour's samples successfully profiled by MLPA.

Statistical analysis

The primary endpoint of this study was the prediction of BCR, defined as two post-prostatectomy PSA measurements of more than 0.2 ng/ml for RP cohorts, or a PSA level increase of more than 2 ng/ml above the post-radiation nadir value for the radiation therapy cohort. Random survival forest (R package "randomForestSRC" v2.9.3 [15]) was used to perform variable selection (feature reduction) among the 14 genes assessed by MLPA, where the CNA status of each gene defined a binary variable. For patients in our RP cohort, the binary CNA data and survival data (time to BCR with censoring) were used to build a random survival forest, where 1000 trees showing a stable out-of-the-bag error rate were generated using the log-rank score to split nodes. Genes with positive variable importance (VIMP) values and 95% confidence intervals above zero after bootstrapping (100 times) indicated that the predictive power of the forest depended on these variables and were selected, whereas those with zero and negative values contributed nothing (zero) or even generated noise (negative) to the prediction model and were discarded.

Cox proportional hazard regression (R package "survival" v3.2–7) was used to build the CNA classifier model from variables previously selected by the random survival forest and used the Wald test to evaluate univariate and multivariate hazard ratios. The predictive power of the CNA classifier alone or in combination with other clinicopathologic features or risk assessment scores was assessed by calculating Harrell C-index (R package "dynpred" v0.1.2). The optimal Cox CNA classifier cut-off score to stratify patients into low and high risk group of BCR was determined by maximally selected rank statistics with the adaptive method (R Shiny application "Evaluate Cutpoints" [16]). Recurrence-free survival analyses were performed using Kaplan–Meier method and the log-rank test (R package "survcomp" v1.38.0). To validate our findings, we used CNA profiles of GG 1–3 ($GS \leq 7$) and local disease ($\leq pT3$) from three RP datasets: MSKCC [17], Cambridge [18], CPC-GENE [19] and one radiation therapy pre-treatment biopsy dataset: Toronto [8]. In all statistical tests, *P* values of less than 0.05 were considered as statistically significant.

RESULTS

DNA copy number analysis

The adequate performance of the new probes added to the original 10-gene prostate MLPA probe mix [11] was confirmed in 13 independent clinical PCa samples previously used for the development of the assay [11] and for which the gold standard CNA status was determined by FISH (Supplementary Figs. S1 and S2). The expanded 14-gene MLPA assay was then applied to 448 RP cases represented by samples A and/or B taken from the highest and the lowest Gleason pattern areas whenever applicable, respectively. In 97% of cases ($n = 433$), both duplicate MLPA

Table 1. Clinicopathologic features of radical prostatectomy cases from the MLPA cohort.

Clinicopathologic variables	Category	n (%)
Total number of cases	<i>n</i>	412
Patients with A samples		362 (88%)
Patients with B samples		343 (83%)
Patients with A and B samples		293 (71%)
Age (years)	Median	60
	Min–max	43–77
Pathological stage (T-stage)	pT2	245 (59%)
	pT3A	143 (35%)
	pT3B	24 (6%)
Surgical margin status	Positive	126 (31%)
Follow-up (months)	Median (min–max)	69 (1–226)
Biochemical recurrence (BCR)	Positive	79 (19%)
Grade Group at surgery, Gleason score indicated	1 (3 + 3)	112 (27%)
	2 (3 + 4)	209 (51%)
	3 (4 + 3)	91 (22%)
	<i>n</i> ^a	286 (69%)
Grade Group at biopsy, Gleason score indicated	1 (3 + 3)	155 (54%)
	2 (3 + 4)	106 (37%)
	3 (4 + 3)	20 (7%)
	≥4 (≥4 + 4)	5 (2%)
Preoperative prostate-specific antigen (ng/ml)	<i>n</i> ^a	406 (99%)
	Mean (±SD)	7.81 (±5.26)
	PSA ≤ 10	334 (82%)
	PSA > 10	72 (18%)
Clinical stage (C-stage)	<i>n</i> ^a	252 (61%)
	cT1	170 (67%)
	cT2	81 (32%)
	cT3A	0 (0%)
	cT3B	1 (<1%)

^aValues not available for all 412 cases (*n* noted for each variable).

reactions of A or B samples passed the QC allowing CNA assessment. Ratio charts of successful and failed MLPA reactions are shown as examples in Supplementary Fig. S3. Clinicopathologic data were retrieved for 412 of these cases which were considered for subsequent analyses and referred to as the MLPA cohort (Table 1). CNAs were detected in 75% of cases, of which the majority harboured 1–3 CNAs (Fig. 1a). The most frequent CNAs were on chromosome 8p with deletions of *NKX3-1* (8p21.2) and *WRN* (8p12) detected in 35% and 30% of the cases, respectively (Fig. 1b). Overall, there was no statistical difference (Kruskal–Wallis test, $P = 0.847$) in the CNA frequency of the assessed 14 genes between the MLPA cohort and the previously reported CNA profile generated by array-CGH or DNA sequencing of GG 1–3 ($GS \leq 7$) cohorts used for validation of the current study findings [8, 17–19].

Developing a CNA classifier

Random survival forest was applied on the MLPA data to identify the best predictive CNAs for BCR and reduce the number of variables in the classifier. Six CNAs were identified as contributing to BCR prediction (positive variable importance) among the 14 CNAs (Fig. 1c) and included in a Cox proportional hazard regression to build the final

6-gene classifier model. The score value was calculated according to the Cox regression coefficients and the CNA status (0 vs 1): $0.83 \cdot RWDD3 + 0.37 \cdot WRN + 0.41 \cdot PTEN + 0.45 \cdot TP53 + 0.23 \cdot MYC + 0.92 \cdot PDPK1 - 0.32$ (centring constant), which was able to predict BCR with a univariate hazard ratio (HR) of 2.76 (Table 2A). Using an optimised cut-off score of ≥ 0.09 for this model, patients with high risk of early BCR can be distinguished from lower risk patients in Kaplan–Meier analysis (Fig. 1d, $P < 0.0001$). As expected, the low risk group included a higher proportion of GG 1 (3 + 3) (32% vs 16%) and a lower proportion of GG 3 (4 + 3) (18% vs 31%) than the high risk group, respectively (Supplementary Fig. S4, $P = 0.0005$). The prediction of BCR by the model in Cox regression was superior to a single CNA (*PTEN* deletion) already known as a prognostic marker (Harrell C-index, 0.64 vs 0.57). In multivariate analysis after adjusting for standard clinicopathologic features (Table 2A) or the CAPRA-S (cancer of the prostate risk assessment post-surgical) [20] score risk group (Table 2B), the 6-gene model retained its significance (Wald test, $P < 0.01$) with a hazard ratio of 2.17 and 1.80, respectively. Moreover, the CNA model improved the BCR prediction of the standard clinicopathologic variables and CAPRA-S score risk group, as shown by the increased Harrell C-index of the combinations (Fig. 1e).

From the perspective of patient management, the impact of sampling performance on the CNA classifier was investigated in cases with BCR information and a CNA call available for both samples A and B representing, when applicable, the highest and the lowest Gleason grade pattern, respectively ($n = 293$). For this comparative analysis, the classifier was applied to the CNA data generated from the combination of both sample A and B as done so far in this study, sample A alone, sample B alone, and from a randomly selected sample A or B. Harrell C-index results indicate that the CNA classifier performed better when both samples were considered (Fig. 1f) and thus this method continued to be applied for the remaining outcome analyses of this study. There were no significant differences in the 6-gene CNA calls made from sample A and B except for two genes (*MYC*, *WRN*) which were more likely to be called from the A sample (highest grade) in the GG 2–3 ($GS = 7$) cases (McNeimar's test, $P < 0.05$, Supplementary Table S2).

CNA classifier validation in RP cohorts

The 6-gene model performance was then assessed on published CNA and clinical data of GG 1–3 ($GS \leq 7$) cases from three RP cohorts: MSKCC [17], Cambridge [18] and CPC-GENE [19]. In all of them, the model was able to stratify patients into low and high risk of BCR as shown on Kaplan–Meier curves (Fig. 2a–c, left panels, $P < 0.05$). The prognostic value of the model remained significant in multivariate analyses (Table 2 C–G) with standard clinicopathologic variables and the CAPRA-S score risk group (surgical margins and CAPRA-S data not available for CPC-GENE cohort). In all instances, the 6-gene model improved the prognostic value (C-index) of the standard markers and the CAPRA-S (Fig. 2a–c, right panels) in the post-treatment assessment.

To explore the potential value of the CNA model at the diagnostic stage, multivariate analyses were performed with the biopsy GG, clinical stage (cT) and preoperative PSA levels as well as with three well-known prognostic indexes: CAPRA (cancer of the prostate risk assessment) [21], D'Amico [22], and NCCN (national comprehensive cancer network) [23] in MLPA cases for which these pre-treatment data were available. As shown in Table 3A–D, the 6-gene model remained a significant prognostic marker of BCR with the standard pre-treatment variables and the three risk classifiers. The contribution of the CNA model to the prognostication was translated into an improved C-index when combined with the clinicopathologic variables and the three classifiers (Fig. 3a). Similarly, the CNA model was a significant prognostic marker in multivariate analyses (Supplementary Table S3A–I) and increased the C-index of standard pre-treatment variables as well as of D'Amico and NCCN classifiers in MSKCC, Cambridge and CPC-GENE datasets (Fig. 3b–d).

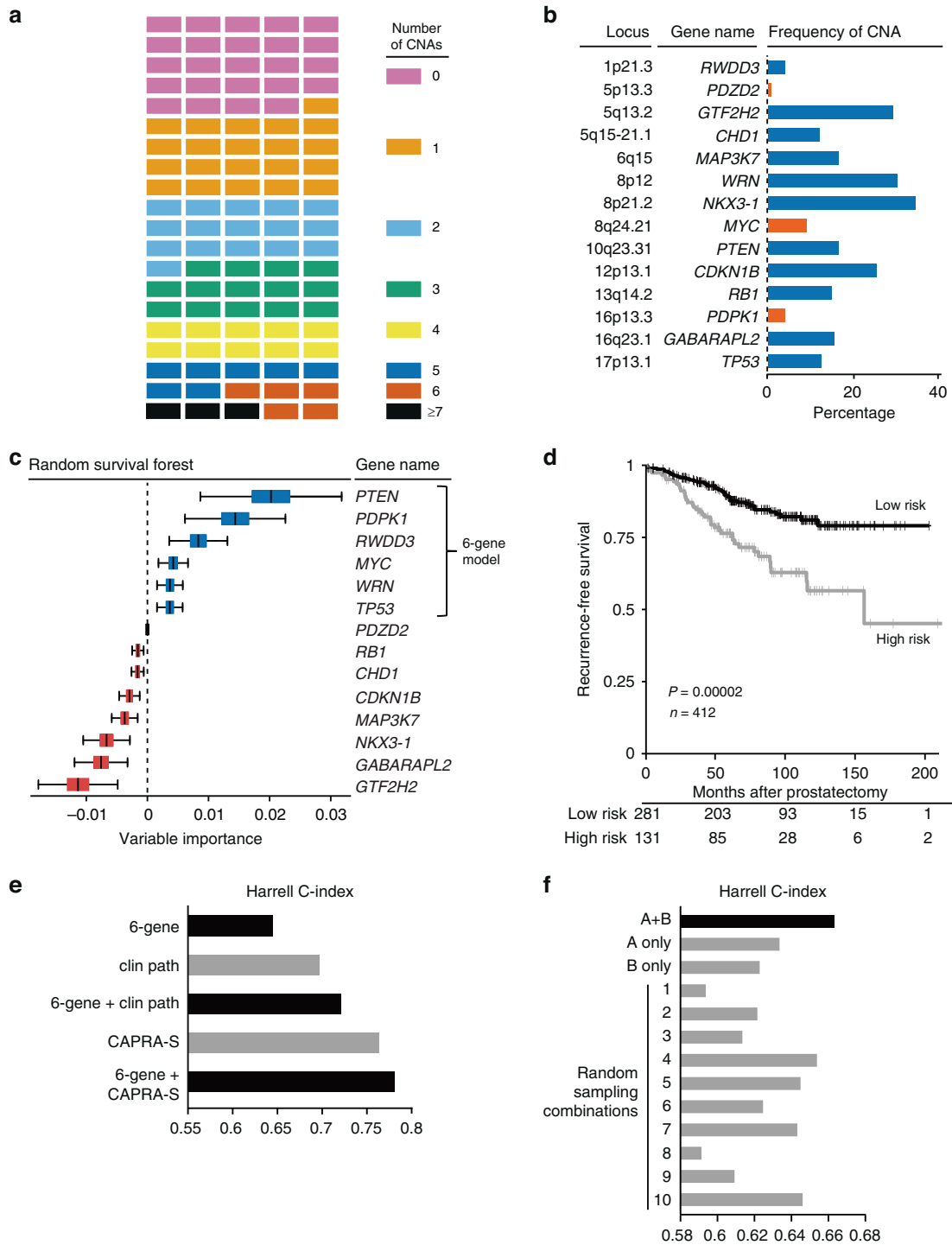


Fig. 1 Copy number alteration (CNA) profiles of the MLPA cohort and generation of a 6-gene CNA model to predict biochemical recurrence (BCR). **a** Waffle chart showing the distribution of MLPA cohort radical prostatectomy cases ($n = 412$) according to the number of CNAs detected by the 14-gene multiplex ligation-dependent probe amplification (MLPA) assay. One rectangle represents 1%. **b** Frequency of individual CNA with chromosomal loci and targeted genes shown in red for gain and blue for deletion. **c** Selection of CNAs to predict BCR using random survival forest analysis. Boxplot shows positive (blue) and negative (red) variable importance (VIMP) values. Means, quartiles and 95% confidence intervals are indicated. **d** Kaplan–Meier BCR survival analysis of the MLPA cohort stratified according to the 6-gene Cox model score (low <0.09 vs. high ≥ 0.09), log-rank test, P value and the number of patients are indicated. **e** Harrell C-index for BCR from 6-gene model applied alone or in combination with the clinicopathologic variables (clin path: preoperative PSA, pathological GG 3 (4 + 3) vs. 1–2 ($\leq 3 + 4$), pathological stage T3 vs. $\leq T2$, and surgical margins) or the CAPRA-S score risk group, $n = 406$ patients. **f** Harrell C-index for BCR based on the 6-gene model when the CNA status for each case ($n = 293$) was assigned by considering both A and B samples (A + B), only A, only B, and A or B sample randomly selected (ten sampling combinations).

Table 2. Univariate and multivariate Cox proportional hazard analysis for 6-gene model adjusted for clinicopathologic parameters and CAPRA-S risk score.

Variables	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% confidence interval)	P value	Hazard ratio (95% confidence interval)	P value
<i>MLPA dataset (n = 406)</i>				
(A) Standard clinicopathologic parameters				
6-gene model ^a	2.76 (1.87–4.08)	<0.0001	2.17 (1.44–3.27)	0.0002
Preoperative prostate-specific antigen ^a	1.06 (1.03–1.09)	0.0001	1.04 (1.01–1.07)	0.02
GG 3 (4 + 3) vs. GG 1–2 (≤3 + 4)	2.27 (1.44–3.60)	0.0005	1.39 (0.86–2.25)	0.2
pT-stage (T3 vs ≤T2)	3.29 (2.06–5.26)	<0.0001	2.22 (1.33–3.71)	0.002
Surgical margin (positive vs negative)	2.13 (1.36–3.34)	0.001	1.49 (0.92–2.41)	0.1
(B) CAPRA-S risk				
6-gene model ^a	2.76 (1.87–4.08)	<0.0001	1.80 (1.19–2.71)	0.005
CAPRA-S risk				
Low (0–2)	–	–	–	–
Intermediate (3–5)	3.45 (1.69–7.04)	0.0007	3.34 (1.64–6.83)	0.001
High (≥6)	18.56 (9.19–37.49)	<0.0001	15.57 (7.60–31.91)	<0.0001
<i>MSKCC dataset (n = 127)</i>				
(C) Standard clinicopathologic parameters				
6-gene model ^a	2.82 (1.69–4.72)	<0.0001	3.33 (1.89–5.89)	<0.0001
Preoperative prostate-specific antigen ^a	1.01 (1.00–1.01)	<0.0001	1.01 (1.00–1.01)	<0.0001
GG 3 (4 + 3) vs. GG 1–2 (≤3 + 4)	3.58 (1.67–7.67)	0.001	3.57 (1.52–8.40)	0.004
pT-stage (T3 vs ≤T2)	1.58 (0.77–3.22)	0.2	1.10 (0.51–2.40)	0.8
Surgical margin (positive vs negative)	2.71 (1.31–5.59)	0.007	2.73 (1.29–5.81)	0.009
(D) CAPRA-S risk				
6-gene model ^a	2.82 (1.69–4.72)	<0.0001	2.03 (1.16–3.54)	0.01
CAPRA-S risk				
Low (0–2)	–	–	–	–
Intermediate (3–5)	7.10 (2.85–17.72)	<0.0001	6.11 (2.41–15.47)	0.0001
High (≥6)	6.00 (1.83–19.70)	0.003	4.19 (1.20–14.68)	0.03
<i>Cambridge dataset (n = 104)</i>				
(E) Standard clinicopathologic parameters				
6-gene model ^a	3.96 (1.69–9.30)	0.002	2.96 (1.18–7.41)	0.02
Preoperative prostate-specific antigen ^a	1.04 (0.93–1.17)	0.5	1.06 (0.94–1.20)	0.3
GG 3 (4 + 3) vs. GG 1–2 (≤3 + 4)	5.23 (1.92–14.28)	0.001	4.94 (1.56–15.59)	0.006
pT-stage (T3 vs ≤T2)	1.23 (0.40–3.81)	0.7	0.76 (0.24–2.45)	0.6
Surgical margin (positive vs negative)	1.90 (0.69–5.25)	0.2	1.71 (0.61–4.76)	0.3
(F) CAPRA-S risk				
6-gene model ^a	3.96 (1.69–9.30)	0.002	4.01 (1.57–10.21)	0.004
CAPRA-S risk				
Low (0–2)	–	–	–	–
Intermediate (3–5)	1.44 (0.43–4.79)	0.6	0.94 (0.26–3.42)	0.9
High (≥6)	3.08 (0.77–12.35)	0.1	2.18 (0.52–9.11)	0.3
<i>CPC-GENE dataset (n = 135)</i>				
(G) Standard clinicopathologic parameters				
6-gene model ^a	2.09 (1.18–3.69)	0.01	2.11 (1.06–4.20)	0.03
Preoperative prostate-specific antigen ^a	1.03 (0.97–1.11)	0.4	1.00 (0.92–1.09)	1.00
GG 3 (4 + 3) vs. GG 1–2 (≤3 + 4)	2.19 (0.99–4.83)	0.05	1.63 (0.73–3.67)	0.2
pT-stage (T3 vs ≤T2)	4.42 (1.84–10.65)	0.0009	3.94 (1.61–9.61)	0.003

^aAnalysed as a continuous variable; GG: Grade Group, Gleason score indicated; pT-stage: pathological stage. P value Wald test.

CNA-classifier validation in a radiation therapy biopsy cohort

Moving a step closer to a diagnostic application, the 6-gene model was finally applied to a published radiation therapy cohort of GG 1–3 (GS ≤ 7) cases with the clinicopathologic parameters and CNA profiles generated from pre-treatment biopsies (Toronto cohort [8]).

The Kaplan–Meier analysis shows that these patients can be stratified in low and high risk group for BCR by the CNA model (Fig. 3e, left panel). The 6-gene model was also a significant variable in multivariate analyses with standard parameters and with both D'Amico and NCCN classifiers (Table 3E–G). All combinations of the

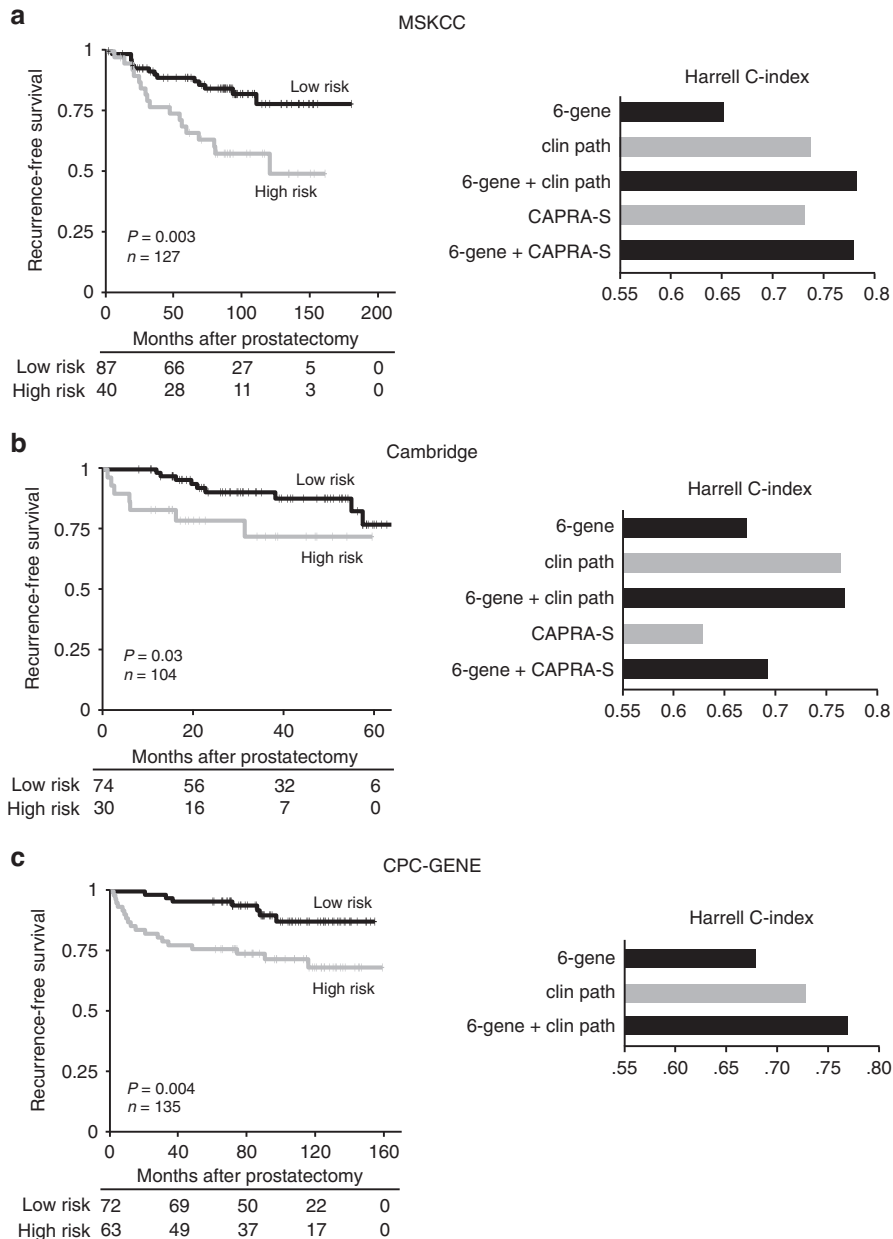


Fig. 2 Prognostic value of the 6-gene CNA model in published radical prostatectomy datasets. a–c Left panels, Kaplan–Meier survival analyses of patients from (a) MSKCC, (b) Cambridge and (c) CPC-GENE cohorts stratified with the 6-gene model score (low <0.09 vs. high ≥ 0.09), log-rank test, P value and the number of patients, as indicated. a–c Right panels, Harrell C-index for BCR based on the 6-gene model applied alone or in combination with the clinicopathologic variables (clin path: preoperative PSA, pathological GG 3 (4 + 3) vs. 1–2 ($\leq 3 + 4$), pathological stage T3 vs. $\leq T2$, and surgical margins) or the CAPRA-S score risk group. Surgical margins and CAPRA-S score are not available for the CPC-GENE cohort.

CNA model with the standard variables or the two classifiers resulted in an increased C-index supporting its potential benefit at the diagnostic stage (Fig. 3e, right panel).

DISCUSSION

MLPA provides a fast, easy and cost-effective approach for the assessment of candidate CNAs in small quantities of DNA extracted from archived FFPE samples. Here, we used an expanded MLPA probe mix to generate the profiles of 433 low or intermediate risk ($GS \leq 7$) prostate tumours for 14 genes in cytobands with recurrent alterations. The performance of the assay was comparable to what we previously reported with the first version of the MLPA probe mix with $>90\%$ of the tumours

sampled twice with good quality data [11]. The expansion of the probe mix allowed the assessment of four additional CNAs, including two that were selected by random survival forest for our BCR classifier 1p21.3 (*RWDD3*) and 8p12 (*WRN*) deletion. Deletion of 1p21.3 was previously detected in advanced prostate tumours [4, 24] and the expression of *RSUME*, encoded by its putative driver *RWDD3*, was reduced in neuroendocrine pancreatic cancer [25]. *RSUME* enhances protein sumoylation, a post-translational modification, that can increase PTEN stability and tumour suppressor function [25, 26]. 8p12 (*WRN*) deletion has been associated with BCR [27] and an aggressive molecular subtype of PCa [4]. *WRN* is a RecQ-like helicase involved in the maintenance of genomic stability (review in [28]) and germline variants have been detected in castrate-resistant prostate cancer (CRPC)

Table 3. Univariate and multivariate Cox proportional hazard analysis for 6-gene model adjusted for pre-surgical or pre-radiation clinicopathologic parameters.

Variables	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% confidence interval)	P value	Hazard ratio (95% confidence interval)	P value
<i>MLPA dataset subcohort</i>				
(A) Standard clinicopathologic parameters (<i>n</i> = 249)				
6-gene model ^a	2.23 (1.31–3.79)	0.003	2.17 (1.25–3.76)	0.006
Preoperative prostate-specific antigen ^a	1.05 (1.01–1.10)	0.01	1.05 (1.00–1.09)	0.04
Biopsy GG ≥3 (≥4 + 3) vs. GG 1–2 (≤3 + 4)	1.93 (0.59–6.28)	0.3	2.07 (0.62–6.87)	0.2
C-stage (T2 vs T1)	1.08 (0.56–2.09)	0.8	0.98 (0.50–1.92)	1.0
(B) D'Amico (<i>n</i> = 235)				
6-gene model ^a	2.87 (1.62–5.11)	0.0003	2.73 (1.53–4.88)	0.0007
D'Amico risk				
Low	–	–	–	–
Intermediate	1.67 (0.79–3.54)	0.2	1.58 (0.75–3.34)	0.2
High	3.60 (1.42–9.16)	0.007	3.23 (1.26–8.24)	0.01
(C) NCCN (<i>n</i> = 254)				
6-gene model ^a	2.25 (1.35–3.74)	0.002	2.08 (1.24–3.49)	0.005
NCCN risk				
Low	–	–	–	–
Intermediate	1.95 (0.97–3.91)	0.06	1.78 (0.88–3.58)	0.1
High	4.22 (1.66–10.75)	0.003	3.82 (1.49–9.76)	0.005
(D) CAPRA risk (<i>n</i> = 238)				
6-gene model ^a	2.08 (1.18–3.69)	0.01	1.92 (1.06–3.49)	0.03
CAPRA risk				
Low (0–2)	–	–	–	–
Intermediate (3–5)	2.42 (1.14–5.13)	0.02	2.27 (1.07–4.82)	0.03
High (≥6)	7.05 (2.40–20.71)	0.0004	6.39 (2.16–18.86)	0.001
<i>Toronto dataset (n = 126)</i>				
(E) Standard clinicopathologic parameters				
6-gene model ^a	3.82 (2.20–6.65)	<0.0001	3.20 (1.82–5.65)	<0.0001
Preoperative prostate-specific antigen ^a	1.13 (1.06–1.21)	0.0003	1.10 (1.03–1.18)	0.004
Biopsy GG 3 (4 + 3) vs. GG 1–2 (≤3 + 4)	1.01 (0.67–1.52)	1.0	1.25 (0.66–2.38)	0.5
C-stage (T2 vs T1)	0.97 (0.55–1.70)	0.9	0.99 (0.56–1.76)	1.0
(F) D'Amico				
6-gene model ^a	3.82 (2.20–6.65)	<0.0001	3.79 (2.16–6.65)	<0.0001
D'Amico risk				
Low	–	–	–	–
Intermediate	0.84 (0.37–1.86)	0.7	0.99 (0.44–2.21)	1.0
High	2.07 (0.43–10.08)	0.4	1.73 (0.35–8.43)	0.5
(G) NCCN				
6-gene model ^a	3.82 (2.20–6.65)	<0.0001	3.83 (2.19–6.68)	<0.0001
NCCN intermediate vs. low risk ^b	0.86 (0.39–1.92)	0.7	1.01 (0.45–2.25)	1.0

^aAnalysed as a continuous variable; GG: Grade Group, Gleason score indicated; C-stage: clinical stage.

P value: Wald test.

^bNo NCCN high risk case in the dataset.

patients [29]. The association of *RWDD3* and *WRN* deletions with poor clinical outcome in our study provides further rationale for future investigation of their potential involvement in the pathobiology of PCa.

In addition to *RWDD3* and *WRN* deletion, the 6-gene CNA classifier included *PTEN* and *TP53* deletion as well *MYC* and *PDPK1* gain consistent with the biological function of these genes and the reported ability of these CNAs to predict BCR [6, 14, 30–32].

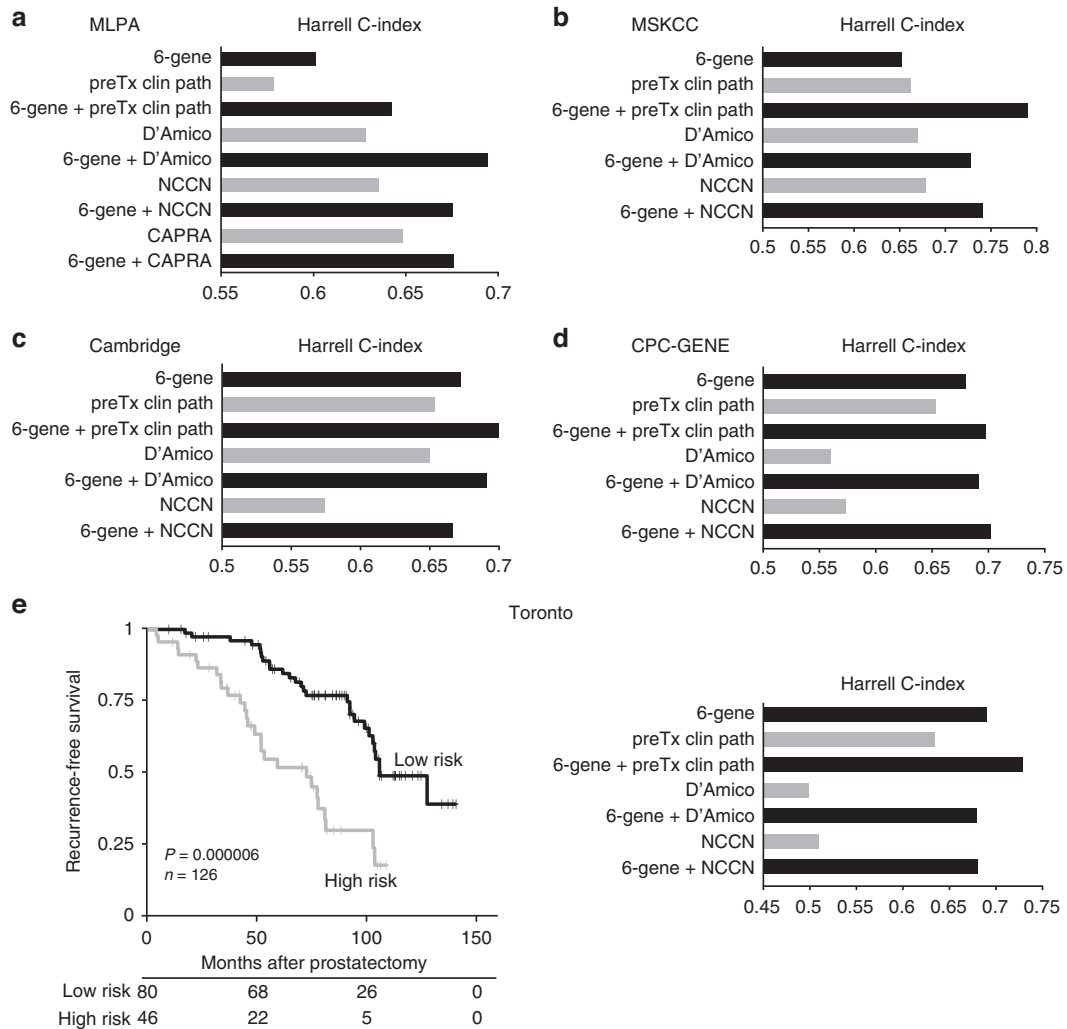


Fig. 3 Six-gene CNA model improved prognostication of pre-treatment clinicopathologic variables in MLPA, published radical prostatectomy, and pre-radiation biopsy datasets. a–d Harrell C-index for BCR based on the 6-gene model applied alone or in combination with pre-treatment clinicopathologic variables (preTx clin path: preoperative PSA, biopsy GG ≥ 3 ($\geq 4 + 3$) vs. 1–2 ($\leq 3 + 4$) and clinical stage T2 \geq vs. T1), D'Amico, NCCN, and CAPRA scores from (a) MLPA, (b) MSKCC, (c) Cambridge and (d) CPC-GENE cohorts. e Left panel, Kaplan–Meier survival analysis of patients from the Toronto pre-radiation biopsy cohort stratified based on the 6-gene model score (low < 0.09 vs. high ≥ 0.09), log-rank test, P value and the number of patients are indicated. Right panel, Harrell C-index for BCR based on the 6-gene model applied alone or in combination with pre-treatment clinicopathologic variables (preTx clin path: pre-radiation PSA, biopsy GG 3 (4 + 3) vs. 1–2 ($\leq 3 + 4$) and clinical stage T2 \geq vs. T1), D'Amico and NCCN scores.

Some of the CNAs included in our classifier such as *PTEN* deletion and *MYC* gain were reportedly enriched in tumour displaying Gleason pattern 4 with cribriform and intraductal gland architecture [33]. In the current cohort, the different histological Gleason patterns 4 or the presence of intraductal carcinoma were not recorded at the time of histological characterisation. Given that those histological patterns predict poor outcome after RP [34, 35], it would be interesting to assess, in future studies, their association with the 6-gene classifier score.

We previously showed that combining *PTEN* deletion and *PDPK1* gain improved the prognostic value of these two individual CNAs assessed by FISH [5]. With the multiplex capability of MLPA, the current study showed that a combination of 6 CNAs was optimal and superior to a single CNA (*PTEN* deletion) to predict BCR. The results also suggest that the classifier may help determine which patients are more at risk of recurrence after RP by providing additional prognostic information to the standard clinicopathologic variables, including the CAPRA-5 score. Recent data indicate that adjuvant radiation therapy might be advantageous to patients with adverse pathology at RP [36]; the 6 CNAs

may potentially identify additional patients who would benefit from such adjuvant treatments. The classifier might also serve to orient therapeutic decisions at diagnosis given that it remained a significant prognostic factor after adjusting for the clinicopathologic variables obtained before treatment, even in the validation cohort for which the CNAs were surveyed in biopsy specimens before radiation therapy. Further validation is warranted on prospective cohorts using the MLPA assay on biopsy specimens.

An important advantage of this study was that each RP case was sampled from two distinct tumour areas which were separately processed for DNA extraction and CNA profiling. The double sampling strategy ensured that a CNA profile can still be derived in instances of inadequate DNA quality/quantity or failed MLPA reaction of one of the samples. Our data showed that either the A or B sample provided prognostic information suggesting that the CNA classifier may work even in suboptimal sampling conditions associated with prostate biopsies. Sampling two tumour areas nevertheless provided a slightly better discrimination power as shown by its C-index, which might be explained, at least in part, by CNAs that were not detected in one sample but were in the

other, a phenomenon that probably occurred in the other datasets analysed in this project where only one area was sampled. In support of this hypothesis, *WRN* deletion and *MYC* gain were detected more often in sample A (highest grade) than B (lowest grade). These results are in agreement with previous findings that have highlighted the intra-tumoral heterogeneity of PCa [37]. In contrast, no significant difference in CNA detection between A and B samples was observed in our previous MLPA report, which may be explained by the limited sample size ($n = 20$ cases) representing mostly higher grade tumours (GG 3 and up, $GS \geq 4 + 3$) [11] than those in the current study.

Most of the PCa prognostic classifier reported are derived from gene expression, but there are CNAs-based signatures that have been developed recently [38–41], including one based on MLPA assay [39]. While there is a certain level of overlap between the regions of CNA assessed by the 6-gene classifier and those signatures, their compositions vary. The 6-gene classifier was derived and validated on GG 1–3 ($GS \leq 7$) cases, while the studies mentioned above were based on different cohorts that included $GG \geq 4$ ($GS \geq 8$) cases, which may account for some of the differences observed. Given that the MLPA assay can assess more than 6 regions of CNA (14 in the current report), it would be possible, in futures studies, to evaluate additional promising CNAs to the 6-gene classifier by replacing probes that underperformed as a prognostic variable in the random survival forest analysis.

In the context of molecular pathology, a seamless addition of the 6-gene CNA classifier to the clinicopathologic assessment workflow could be envisioned. The MLPA assay requires only 50 ng of DNA, which can be extracted from a standard FFPE biopsy specimen without any additional procedure for the patient. The histopathologic diagnosis is not compromised and allows the selection of tumour cell areas optimal for DNA extraction. The assay costs less than ten dollars in consumable supplies and can be implemented in any laboratory that performs PCR and capillary electrophoresis. Future studies would aim to specify its application with the clinicopathologic classifiers and determine in which subgroups of patients it could be most useful.

In conclusion, a CNA classifier improving risk prediction was developed and validated in 492 additional patients from four PCa datasets treated by RP or biopsied prior to radiation therapy. The MLPA-based assay is robust and can be easily implemented worldwide in the clinical setting to improve the management of patients with localised PCa.

DATA AVAILABILITY

Data are available from the corresponding author upon reasonable request.

REFERENCES

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209–49.
- Fletcher SA, von Landenberg N, Cole AP, Gild P, Choueiri TK, Lipsitz SR, et al. Contemporary national trends in prostate cancer risk profile at diagnosis. *Prostate Cancer Prostatic Dis*. 2020;23:81–7.
- Eastham JA, Auffenberg GB, Barocas DA, Chou R, Crispino T, Davis JW, et al. Clinically localized prostate cancer: AUA/ASTRO guideline, part i: introduction, risk assessment, staging, and risk-based management. *J Urol*. 2022;208:10–8.
- Lapointe J, Li C, Giacomini CP, Salari K, Huang S, Wang P, et al. Genomic profiling reveals alternative genetic pathways of prostate tumorigenesis. *Cancer Res*. 2007;67:8504–10.
- Bramhecha YM, Rouzbeh S, Guerard KP, Scarlata E, Brimo F, Chevalier S, et al. The combination of PTEN deletion and 16p13.3 gain in prostate cancer provides additional prognostic information in patients treated with radical prostatectomy. *Mod Pathol*. 2019;32:128–38.
- Yoshimoto M, Cunha IW, Coudry RA, Fonseca FP, Torres CH, Soares FA, et al. FISH analysis of 107 prostate cancers shows that PTEN genomic deletion is associated with poor clinical outcome. *Br J Cancer*. 2007;97:678–85.

- Hieronimus H, Schultz N, Gopalan A, Carver BS, Chang MT, Xiao Y, et al. Copy number alteration burden predicts prostate cancer relapse. *Proc Natl Acad Sci USA*. 2014;111:11139–44.
- Lalonde E, Ishkanian AS, Sykes J, Fraser M, Ross-Adams H, Erho N, et al. Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: a retrospective cohort study. *Lancet Oncol*. 2014;15:1521–32.
- Zafarana G, Ishkanian AS, Malloff CA, Locke JA, Sykes J, Thoms J, et al. Copy number alterations of c-MYC and PTEN are prognostic factors for relapse after prostate cancer radiotherapy. *Cancer*. 2012;118:4053–62.
- Stuppia L, Antonucci I, Palka G, Gatta V. Use of the MLPA assay in the molecular diagnosis of gene copy number alterations in human genetic diseases. *Int J Mol Sci*. 2012;13:3245–76.
- Ebrahimizadeh W, Guerard KP, Rouzbeh S, Bramhecha YM, Scarlata E, Brimo F, et al. Design and development of a fully synthetic multiplex ligation-dependent probe amplification-based probe mix for detection of copy number alterations in prostate cancer formalin-fixed, paraffin-embedded tissue samples. *J Mol Diagn*. 2020;22:1246–63.
- McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, et al. Reporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer*. 2005;93:387–91.
- Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO classification of tumours of the urinary system and male genital organs-part B: prostate and bladder tumours. *Eur Urol*. 2016;70:106–19.
- Bramhecha YM, Guerard KP, Rouzbeh S, Scarlata E, Brimo F, Chevalier S, et al. Genomic gain of 16p13.3 in prostate cancer predicts poor clinical outcome after surgical intervention. *Mol Cancer Res*. 2018;16:115–23.
- Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. 2008;2:841–60.
- Ogluszka M, Orzechowska M, Jedroszka D, Witas P, Bednarek AK. Evaluate cut-points: adaptable continuous data distribution system for determining survival in Kaplan-Meier estimator. *Comput Methods Prog Biomed*. 2019;177:133–9.
- Taylor BS, Schultz N, Hieronimus H, Gopalan A, Xiao Y, Carver BS, et al. Integrative genomic profiling of human prostate cancer. *Cancer Cell*. 2010;18:11–22.
- Ross-Adams H, Lamb A, Dunning M, Halim S, Lindberg J, Massie C, et al. Integration of copy number and transcriptomics provides risk stratification in prostate cancer: a discovery and validation cohort study. *EBioMedicine*. 2015;2:1133–44.
- Espiritu SMG, Liu LY, Rubanova Y, Bhandari V, Holgersen EM, Szyca LM, et al. The evolutionary landscape of localized prostate cancers drives clinical aggression. *Cell*. 2018;173:1003–13.e15.
- Punnen S, Freedland SJ, Presti JC, Aronson WJ, Terris MK, Kane CJ, et al. Multi-institutional validation of the CAPRA-S score to predict disease recurrence and mortality after radical prostatectomy. *Eur Urol*. 2014;65:1171–7.
- Cooperberg MR, Pasta DJ, Elkin EP, Litwin MS, Latini DM, Du Chane J, et al. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. *J Urol*. 2005;173:1938–42.
- D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. *J Am Med Assoc*. 1998;280:969–74.
- Mohler JL, Antonarakis ES, Armstrong AJ, D'Amico AV, Davis BJ, Dorff T, et al. Prostate cancer, version 2.2019, NCCN clinical practice guidelines in oncology. *J Natl Compr Cancer Netw*. 2019;17:479–505.
- Sun J, Liu W, Adams TS, Sun J, Li X, Turner AR, et al. DNA copy number alterations in prostate cancers: a combined analysis of published CGH studies. *Prostate*. 2007;67:692–700.
- Wu Y, Tedesco L, Lucia K, Schlitter AM, Garcia JM, Esposito I, et al. RSUME is implicated in tumorigenesis and metastasis of pancreatic neuroendocrine tumors. *Oncotarget*. 2016;7:57878–93.
- Huang J, Yan J, Zhang J, Zhu S, Wang Y, Shi T, et al. SUMO1 modification of PTEN regulates tumorigenesis by controlling its association with the plasma membrane. *Nat Commun*. 2012;3:911.
- Fukasawa S, Kino M, Kobayashi M, Suzuki H, Komiya A, Imamoto T, et al. Genetic changes in pT2 and pT3 prostate cancer detected by comparative genomic hybridization. *Prostate Cancer Prostatic Dis*. 2008;11:303–10.
- van Wietmarschen N, Nathan WJ, Nussenzweig A. The WRN helicase: resolving a new target in microsatellite unstable cancers. *Curr Opin Genet Dev*. 2021;71:34–8.
- Hart SN, Ellingson MS, Schahl K, Vedell PT, Carlson RE, Sinnwell JP, et al. Determining the frequency of pathogenic germline variants from exome sequencing in patients with castrate-resistant prostate cancer. *BMJ Open*. 2016;6:e010332.
- Choucair KA, Guerard KP, Ejdelman J, Chevalier S, Yoshimoto M, Scarlata E, et al. The 16p13.3 (PDPK1) genomic gain in prostate cancer: a potential role in disease progression. *Transl Oncol*. 2012;5:453–60.

31. Fromont G, Godet J, Peyret A, Irani J, Celhay O, Rozet F, et al. 8q24 amplification is associated with Myc expression and prostate cancer progression and is an independent predictor of recurrence after radical prostatectomy. *Hum Pathol.* 2013;44:1617–23.
32. Kluth M, Harasimowicz S, Burkhardt L, Grupp K, Krohn A, Prien K, et al. Clinical significance of different types of p53 gene alteration in surgically treated prostate cancer. *Int J Cancer.* 2014;135:1369–80.
33. Bottcher R, Kweldam CF, Livingstone J, Lalonde E, Yamaguchi TN, Huang V, et al. Cribriform and intraductal prostate cancer are associated with increased genomic instability and distinct genomic alterations. *BMC Cancer.* 2018;18:8.
34. Trudel D, Downes MR, Sykes J, Kron KJ, Trachtenberg J, van der Kwast TH. Prognostic impact of intraductal carcinoma and large cribriform carcinoma architecture after prostatectomy in a contemporary cohort. *Eur J Cancer.* 2014;50:1610–6.
35. Kweldam CF, Wildhagen MF, Steyerberg EW, Bangma CH, van der Kwast TH, van Leenders GJ. Cribriform growth is highly predictive for postoperative metastasis and disease-specific death in Gleason score 7 prostate cancer. *Mod Pathol.* 2015;28:457–64.
36. Tilki D, Chen MH, Wu J, Huland H, Graefen M, Wiegel T, et al. Adjuvant versus early salvage radiation therapy for men at high risk for recurrence following radical prostatectomy for prostate cancer and the risk of death. *J Clin Oncol.* 2021;39:2284–93.
37. Boutros PC, Fraser M, Harding NJ, de Borja R, Trudel D, Lalonde E, et al. Spatial genomic heterogeneity within localized, multifocal prostate cancer. *Nat Genet.* 2015;47:736–45.
38. Chen Y, Sadasivan SM, She R, Datta I, Taneja K, Chitale D, et al. Breast and prostate cancers harbor common somatic copy number alterations that consistently differ by race and are associated with survival. *BMC Med Genomics.* 2020;13:116.
39. Liu W, Hou J, Petkewicz J, Na R, Wang CH, Sun J, et al. Feasibility and performance of a novel probe panel to detect somatic DNA copy number alterations in clinical specimens for predicting prostate cancer progression. *Prostate.* 2020;80:1253–62.
40. Wang S, Li H, Song M, Tao Z, Wu T, He Z, et al. Copy number signature analysis tool and its application in prostate cancer reveals distinct mutational processes and clinical outcomes. *PLoS Genet.* 2021;17:e1009557.
41. Wang X, Grasso CS, Jordahl KM, Kolb S, Nyame YA, Wright JL, et al. Copy number alterations are associated with metastatic-lethal progression in prostate cancer. *Prostate Cancer Prostatic Dis.* 2020;23:494–506.

ACKNOWLEDGEMENTS

Authors thank all the patients who consented to provide prostate tissues for research.

AUTHOR CONTRIBUTIONS

Conception and design: WE, KPG and JL. Development of methodology: WE, KPG and JL. Acquisition of the data: WE, KPG, SR, ES, FB, PGP, TJ, LH, AGA, AYL, DMB, JMSB, SC and JL. Analysis and interpretation of the data: WE, KPG, AYL and JL. Writing, review, and/or revision of the manuscript: WE, KPG, SR, ES, FB, PGP, TJ, LH, AGA, AYL, DMB,

JMSB, SC and JL. Administrative, technical, or material support: ES, LH, TJ and AGA. Study supervision: JL.

FUNDING

This study was funded by the Movember Foundation through Prostate Cancer Canada (grant #T2014.01) and carried out as a part of personalised risk stratification for patients with early prostate cancer (PRONTO). WE received a studentship and doctoral training award from the Research Institute of McGill University Health Centre and Fonds de la Recherche du Québec-Santé (FRQS), respectively.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study was approved by the Research Ethics Board of McGill University Health Centre (Quebec, Canada, BDM-10-115) and by the Queen's University Research Ethics Board (Ontario, Canada, PATH-144-14). Informed consent was obtained from participants and the study was performed in accordance with the Declaration of Helsinki.

CONSENT FOR PUBLICATION

Not applicable.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41416-023-02236-8>.

Correspondence and requests for materials should be addressed to Jacques Lapointe.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.