

Original Research

Mutational signatures and their association with survival and gene expression in urological carcinomas

Peeter Karihtala^{a,*}, Outi Kilpivaara^{b,c,d,†}, Katja Porvari^{e,†}

^a Department of Oncology, Helsinki University Hospital Comprehensive Cancer Center and University of Helsinki, Helsinki FI-00290, Finland

^b Applied Tumor Genomics Research Program, Faculty of Medicine, University of Helsinki, Helsinki FI-00014, Finland

^c Department of Medical and Clinical Genetics, Medicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland

^d HUSLAB Laboratory of Genetics, HUS Diagnostic Center, Helsinki University Hospital, Helsinki FI-00014, Finland

^e Department of Pathology, Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu FI-90220, Finland



ARTICLE INFO

Keywords:

Bladder carcinoma

COSMIC

Mutations

Renal cell carcinoma cancer

Prognosis

Signature

Urothelial cancer

ABSTRACT

Different sources of mutagenesis cause consistently identifiable patterns of mutations and mutational signatures that mirror the various carcinogenetic processes. We used publicly available data from the Cancer Genome Atlas to evaluate the associations between the activity of the mutational signatures and various survival endpoints in six types of urological cancers after adjusting for established prognostic factors. The predictive power of the signatures was evaluated with dynamic area under curve models. In addition, links between mutational signature activities and differences in gene expression patterns were analysed. APOBEC-related signature SBS2 was associated with improved overall survival (OS) and disease-specific survival (DSS) in bladder carcinomas in the multivariate analysis, while clock-like signature SBS1 predicted shortened DSS and progression-free interval (PFI) in clear cell renal cell carcinomas (ccRCC). In papillary renal cell carcinomas (pRCC), SBS45 was a predictor of improved outcomes, and APOBEC-related SBS13 was a predictor of worse outcomes. Gene expression analyses revealed various enriched pathways between the low- and high-signature groups. Interestingly, in both the ccRCC and pRCC cohorts, the genes of several members of the melanoma antigen (MAGE) family were highly upregulated in the signatures, which predicted poor outcomes, and downregulated in signatures, which were associated with improved survival. To summarize, SBS signatures provide substantial prognostic value compared with just the traditional prognostic factors in certain cancer types. APOBEC-related SBS2 and SBS13 seem to provide robust prognostic information for particular urological cancers, maybe driven by the expression of specific groups of genes, including the MAGE gene family.

1. Introduction

Urological cancers are among the most frequently occurring cancers, and they have a significant impact on total cancer mortality worldwide [1]. Their survival rates vary from the excellent long-term disease-specific survival (DSS) rates observed in testicular cancers and local prostate cancers to just a few years of median survival in cases of metastatic urinary bladder and renal cell carcinomas (RCC), despite the introduction of immuno-oncological treatments during the last few years [2,3]. Virtually all urological carcinomas lack any established prognostic factors in addition to the TNM stage and histological subtype. More precise and reproducible prognostic factors would thus be needed to optimize their surgical and oncological treatments, and surveillance.

Mutational signatures are characteristic patterns of mutations that reflect different carcinogenetic pathways [4]. The incidence of these unique mutational patterns in cancer genomes can give insights into, for example, past exposure to carcinogens and defects in DNA repair mechanisms. Single-base substitution (SBS) mutational signatures consist of 53 distinct signatures in the latest version of the Catalogue of Somatic Mutations in Cancer (COSMIC) database [5]. SBS signatures may tell, for example, about the activation of AID/APOBEC cytidine deaminases (signatures SBS2 and SBS13), defects in specific DNA proofreading mechanisms (e.g. SBS10), exposure to specific chemotherapies (e.g. SBS17), or they may be secondary to smoking (e.g. SBS4) [5,6]. The aetiology of some signatures is still unknown, although knowledge is rapidly progressing [5].

* Corresponding author.

E-mail address: peeter.karihtala@hus.fi (P. Karihtala).

† These authors contributed equally to this work.

The association between COSMIC mutational signatures and outcomes in patients with urological cancers has not yet been studied. We used publicly available data from the Cancer Genome Atlas (TCGA) database to evaluate whether these signatures would be able to show differences in survival endpoints in bladder transitional cell carcinoma (BTCC), clear cell RCC (ccRCC), papillary RCC (pRCC), renal chromophob carcinoma, prostate adenocarcinoma, and testicular germ cell tumours (TGCT). These studies were complemented by gene expression analysis to gain insight into the potential molecular mechanisms connected to mutational signatures and survival.

2. Materials and methods

2.1. Data

Mutational signature activity data published in [5] were accessed from The International Cancer Genome Consortium (ICGC) data portal: https://dcc.icgc.org/releases/PCAWG/mutational_signatures/Signatures_in_Samples/SP_Signatures_in_Samples. The baseline characteristics for the cohorts used in this study are shown in the Supplementary Tables 1-6. The data comprised whole-genome sequenced tumours from the Pan-Cancer Analysis of Whole-Genomes (PCAWG) consortium and whole-exome sequenced tumours from TCGA. Here we concentrated to TCGA data, where only SBS and ID signatures were available for the samples. Open-access RNA-seq harmonized (aligned to GRCh38 reference genome) gene expression data (HTSeq generated counts) from TCGA were downloaded from the Genomic Data Commons (GDC) portal [8] using the R package *TCGAbiolinks* v2.16.4 [9].

2.2. Univariate survival analysis

The association between mutational signatures and overall survival (OS) was first tested in a univariate approach utilizing the R packages *survival* (v3.2-7) [10] and *survminer* (v0.4.8) [11]. Only TCGA primary tumour samples and patients with available vital status and survival/follow-up time were included. The association between signature activity and OS in each cancer type was analysed according to the following criteria: at least twenty samples had both signature and survival data, there were at least five death events among the patients, and there were at least five samples with nonzero signature activity. The association with survival between low-activity and high-activity tumours was then tested using the log-rank test for the given signature. Low-activity tumours were defined as those with a median or lower activity of the signature within the cancer type, and high-activity tumours were defined as those with above-median activity. For the association between the gene expression of the MAGE family genes and OS, log₂-transformed RSEM counts were used, and the subjects were dichotomized (high and low expression) based on the mean gene expression values. Only primary tumour samples were considered. For each analysis, a Kaplan–Meier curve was plotted using the function *ggsurvplot* from the *survminer* package.

2.3. Multivariate survival analysis

Multivariate survival analysis was performed for BTCC, ccRCC, pRCC, RCC, prostate adenocarcinoma, and TGCT using TCGA data and the *survival* and *survminer* packages. For each cancer type, a Cox proportional hazard regression with multiple variables was fitted, considering all signatures or selected MAGE expressions with nonzero activity in at least 5% of patients within the cancer type and selected clinical variables. For each cancer type, a separate regression model was fitted for each of the following endpoints: OS, DSS, PFI and DFI. The function *cox.zph* was used to test the proportional hazard assumptions and plot the Schoenfeld residuals for each variable and the combined model. The function *coxph* was used to run the Cox regression. Forest plots illustrating the hazard ratios of each variable were generated using the

function *ggforest* from the package *survminer* v0.4.8 [11]. In cases where no uncensored patients (i.e. patients with a qualifying progression event) with a valid value of a clinical variable were available, these variables were either left out, or patients with particular values of that variable were removed. Cross-tables were produced using the R package *gtsummary* [12].

2.4. Time-dependent area under curve (AUC)

Time-dependent AUC was calculated for each Cox proportional hazard regression with at least one significant COSMIC mutational signature variable. For each model the incident/dynamic AUC was estimated as proposed by Song and Zhou [13] using the *survAUC* R package across the entire range of endpoint timepoints with one-year intervals. The summary AUC measure (integrated AUC; iAUC) reported is the integral of AUC on the entire time range.

2.5. MAGE gene expressions as predictors of overall survival

Open-access RNA-seq harmonized (aligned to GRCh38 reference genome) gene expression data (HTSeq generated counts) of the The Cancer Genome Atlas (TCGA) Kidney Renal Clear Cell Carcinoma: TCGA-KIRC (ccRCC) and Cervical Kidney renal papillary cell carcinoma: TCGA-KIRP (pRCC) were downloaded from the Genomic Data Commons (GDC) using R package *TCGAbiolinks*, v. 2.16.4. Gene expression data were quantile normalized using the *preprocessCore* R package. For ccRCC the following genes from the MAGE family were considered: *MAGEB2*, *MAGEC2*, *MAGEA4*, and *MAGEA6* while for pRCC *MAGEA3*, *MAGEA11*, *MAGEA12*, *MAGEC1*, and *MAGEC2*. Patients were dichotomized based on median of each MAGE gene expression with patients higher than median classified as High for MAGE.

2.6. Differential expression analysis

The HTseq count matrix was filtered to include only subjects with available mutational signature information. Additionally, data from only primary tumour samples were considered. Duplicated samples (different aliquots of the same patient) were also excluded. Samples (subjects) were dichotomized into low and high mutational signature groups based on the population median mutational signature activity. Finally, the count matrix was filtered to exclude genes with fewer than 10 reads in total across all samples. Data normalization and differential expression analysis were performed using the R package *DESeq2* [14] by contrasting gene expression in the high versus the low group for each cancer type and mutational signature combination. Differentially expressed genes were tested on the null hypothesis that the log₂-fold change between groups was equal to 0 (*lfc0*). Significant genes were identified as those with an adjusted p-value of < 0.05 and an absolute log₂-fold change of > 1.

2.7. Gene ontology enrichment

Overrepresentation tests of differentially expressed genes in gene ontology, biological processes, molecular function, and cellular component terms were performed using the R package *clusterProfiler* [15,16]. Hypergeometric test p-values were corrected for multiple testing using the Benjamini–Hochberg multiple testing adjustment. Terms with an adjusted p-value of < 0.05 were considered significantly enriched. Enriched terms were further visualized using the R package *clusterProfiler*.

3. Results

3.1. COSMIC mutational signatures and survival

First, we studied whether specific mutational signatures were

Table 1

Summary of Cox proportional hazard (CoxPH) analyses from TCGA data. The clinical variables used in the models included age, gender (if applicable), tissue, or organ of origin, and primary diagnosis. American Joint Committee on Cancer (AJCC) staging was also included, except for prostate adenocarcinoma, where this variable was replaced with two Gleason score groups: (1) a Gleason score of < 7 and 7 (3 + 4) and (2) a Gleason score of > 7 and 7 (4 + 3).

Cancer type	Endpoint	Signatures eligible for the analysis	Excluded variables / variable levels	Variables with rejected proportional hazards assumptions	Signatures with p-value < 0.05 from the CoxPH test
Bladder transitional cell carcinoma	OS	SBS1, SBS2, SBS3, SBS5, SBS13	AJCC pathologic: stage I samples, Primary diagnosis: Carcinoma, Papillary adenocarcinoma and Squamous cell carcinoma samples, Tissue: Bladder neck and Ureteric orifice samples		SBS2 (Improved survival)
	DSS	SBS1, SBS2, SBS3, SBS5, SBS13	Same as above		SBS2 (Improved survival)
	DFI	SBS1, SBS2, SBS3, SBS5, SBS13	Tissue or organ of origin and AJCC pathologic stage and signature SBS3 not included in the model	SBS2, SBS5, SBS13	
	PFI	SBS1, SBS2, SBS3, SBS5, SBS13	AJCC pathologic: stage I samples, Primary diagnosis: Carcinoma, Papillary adenocarcinoma and Squamous cell carcinoma samples, Tissue: Bladder neck and Ureteric orifice samples	SBS1	SBS2 (Improved survival)
Clear cell renal cell carcinoma	OS	SBS1, SBS5, SBS40, SBS45, SBS52	Tissue or organ of origin and Primary diagnosis not included in the model	Age	
	DSS	SBS1, SBS5, SBS40, SBS45, SBS52	Same as above	SBS40	SBS1 (Decreased survival)
	DFI	SBS1, SBS5, SBS40, SBS45, SBS52	Tissue or organ of origin, Primary diagnosis and SBS45 not included in the model, AJCC pathologic stage IV samples		
	PFI	SBS1, SBS5, SBS40, SBS45, SBS52	Tissue or organ of origin and Primary diagnosis not included in the model	SBS45	SBS1 (Decreased survival)
Renal papillary cell carcinoma	OS	SBS1, SBS2, SBS5, SBS13, SBS45	Same as above	AJCC pathologic stage, gender, SBS1	SBS45 (Improved survival)
	DSS	SBS1, SBS2, SBS5, SBS13, SBS45	Same as above	AJCC pathologic stage	SBS45 (Improved survival)
	DFI	SBS1, SBS2, SBS5, SBS13, SBS45	Tissue or organ of origin and Primary diagnosis not included in the model, AJCC pathologic stage IV samples	AJCC pathologic stage	SBS13 (Decreased survival)
	PFI	SBS1, SBS2, SBS5, SBS13, SBS45	Tissue or organ of origin and Primary diagnosis not included in the model	AJCC pathologic stage, SBS1, SBS5	SBS13 (Decreased survival)
Renal chromophobic carcinoma	OS	SBS1, SBS5, SBS17a, SBS46	Tissue or organ of origin and Primary diagnosis not included in the model, AJCC pathologic stage and signatures SBS5 and SBS46 also as model did not converge		
	DSS	SBS1, SBS5, SBS17a, SBS46	Tissue or organ of origin and Primary diagnosis not included in the model, AJCC pathologic stage and signatures SBS5, SBS17a and SBS46 also as model did not converge		
	DFI	SBS1, SBS5, SBS17a, SBS46	Tissue or organ of origin and Primary diagnosis not included in the model, AJCC pathologic stage and all signatures as model did not converge		
	PFI	SBS1, SBS5, SBS17a, SBS46	Tissue or organ of origin and Primary diagnosis not included in the model, AJCC pathologic stage and signatures SBS5, SBS17a and SBS46 also as model did not converge		
Prostate adenocarcinoma	OS	SBS1, SBS5, SBS10b, SBS15, SBS40, SBS45	Tissue or organ of origin not included; samples with primary diagnosis 'Adenocarcinoma with mixed subtypes' or 'Mucinous adenocarcinoma', signatures SBS10b and SBS15 due to inflated coefficients		SBS40, SBS45 (Decreased survival)
	DSS	SBS1, SBS5, SBS10b, SBS15, SBS40, SBS45	Tissue or organ of origin not included; samples with primary diagnosis 'Adenocarcinoma with mixed subtypes' or 'Mucinous adenocarcinoma', Gleason group and signatures SBS10b and SBS15 due to inflated coefficients		SBS5, SBS40 (Decreased survival)
	DFI	SBS1, SBS5, SBS10b, SBS15, SBS40, SBS45	Tissue or organ of origin not included; samples with primary diagnosis 'Adenocarcinoma with mixed subtypes' or 'Mucinous adenocarcinoma'		
	PFI	SBS1, SBS5, SBS10b, SBS15, SBS40, SBS45	Same as above		
Testicular germ cell tumours	OS	SBS1, SBS2, SBS5, SBS7a, SBS13, SBS15, SBS19, SBS24, SBS42, SBS44	Tissue or organ of origin not included, 'AJCC pathologic stage' subtypes combined into 3 stages, 'Primary diagnosis' categories with few samples combined into 'Other' group, signatures SBS2, SBS7a, SBS13, SBS15, SBS19, SBS24 and SBS44 due to inflated coefficients		
	DSS	SBS1, SBS2, SBS5, SBS7a, SBS13, SBS15, SBS19, SBS24, SBS42, SBS44	Tissue or organ of origin not included, 'AJCC pathologic stage' and 'Primary diagnosis' and signatures SBS1, SBS2, SBS7a, SBS13, SBS15, SBS19, SBS24 and SBS44 due to inflated coefficients		
	DFI	SBS1, SBS2, SBS5, SBS7a, SBS13, SBS15, SBS19, SBS24, SBS42, SBS44	Tissue or organ of origin not included, 'AJCC pathologic stage' due to inflated coefficients	SBS44	
	PFI	SBS1, SBS2, SBS5, SBS7a, SBS13, SBS15, SBS19, SBS24, SBS42, SBS44		Primary diagnosis, SBS2	

OS = overall survival; DSS = disease-specific survival; DFI = disease-free interval; PFI = progression-free interval; SBS = single-base substitution.

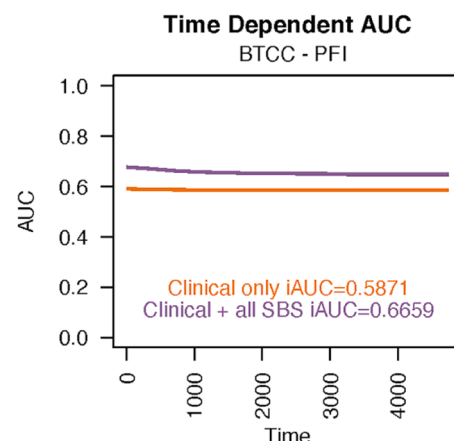
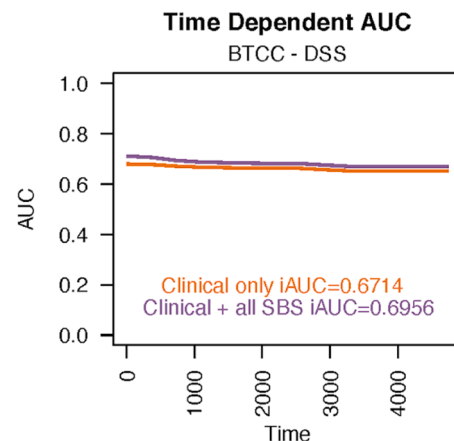
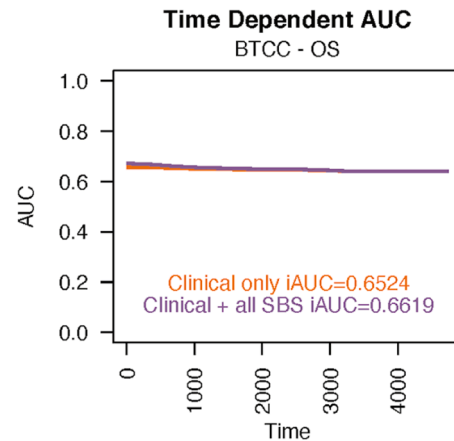
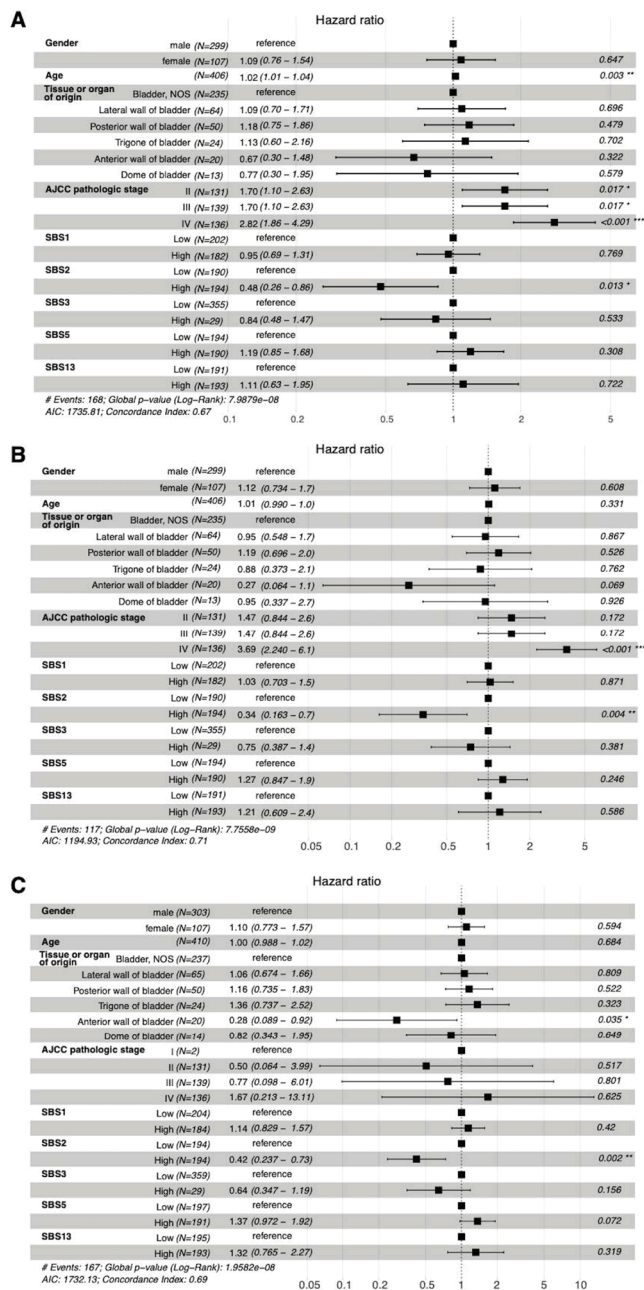


Fig. 1. Cox proportional hazards regression analysis results for bladder transitional cell carcinoma. In these models, nonmetastatic disease and high SBS2 prevalence were associated with better OS (A) and DSS (B). A high number of SBS2 signatures also predicted a better PFI (C). AIC = Akaike Information Criterion; AJCC = American Joint Committee on Cancer; NOS = not otherwise specified; SBS = single-base substitution.

associated with survival endpoints (overall survival (OS), DSS, disease-free interval (DFI), or progression-free interval (PFI) in the six studied cancer types from the TCGA cohort. We performed Cox proportional hazard analyses, which are reported in more detail in Table 1.

The high prevalence of SBS2 signatures in BTCC was associated in multivariate analysis with improved PFI (Hazard Ratio (HR) 0.42, 95% Confidence Interval (CI) 0.24–0.73), DSS (HR 0.34, 95% CI 0.16–0.70), and OS (HR 0.48, 95% CI 0.26–0.86; see Figure 1). The iAUC values in the Cox model were 0.6659 for PFI, 0.6956 for DSS and 0.6619 for OS (Figure 2). Adding mutational signatures to the model improved especially the prognostic value of PFI when compared to using only the traditional prognostic factors. The prognostic value of the SBS2

Fig. 2. The time-dependent integrated areas under the curves (iAUC) values in bladder transitional cell carcinoma for progression-free interval (A), for disease-specific survival (B) and for overall survival (C). SBS = single-base substitution.

signature also remained separate in the stage I–III and stage IV cohorts (Figure 3).

In ccRCC, the high prevalence of the SBS1 signature was associated with decreased DFI (HR 7.69, 95% CI 1.45–40.80) and PFI (HR 1.90, 95% CI 1.19–3.00; see Figure 4). The iAUC values were 0.8139 and 0.7773, respectively (Figure 5). There was a notable increase in the prognostic power of the DFI endpoint, when mutational signatures were added to the model apart from clinical prognostic factors.

Again, in pRCC, both SBS13 and SBS45 signatures were significantly associated with survival endpoints after adjusting for traditional

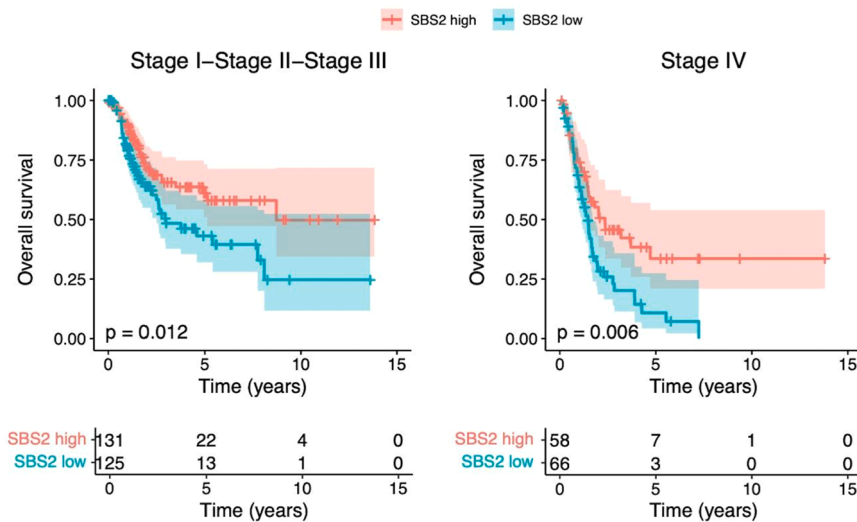


Fig. 3. Kaplan-Meier curves showing that the prognostic value of SBS2 in terms of OS was observed separately in both stage I-III (left) and stage IV (right) bladder transitional cell carcinomas. The coloured areas represent 95% confidence intervals.

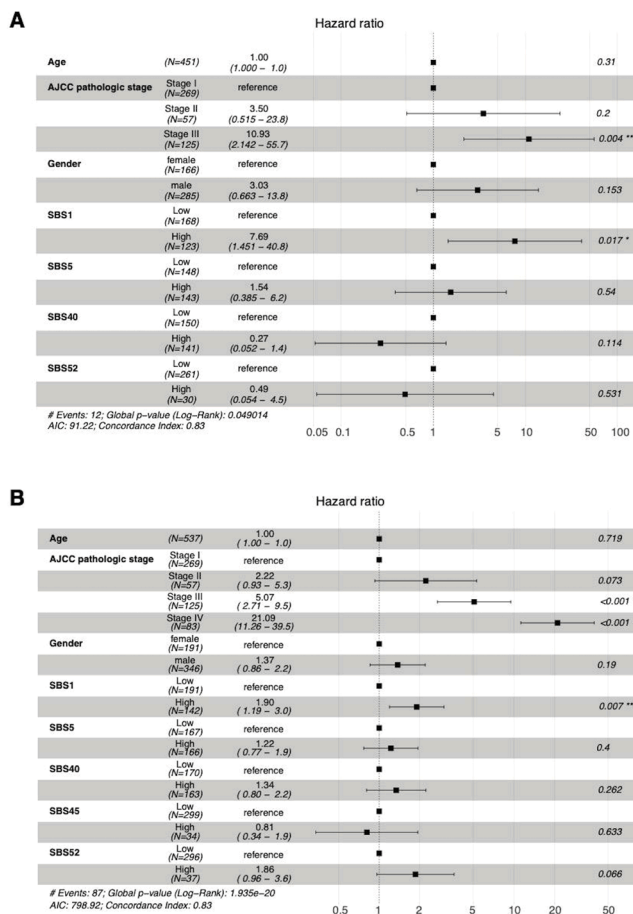


Fig. 4. In patients with ccRCC, the high activity of SBS1 was associated with a worse DFI (A) and PFI (B). Stage was still the most predominant prognostic factor in both cases. AJCC = American Joint Committee on Cancer; SBS = single-base substitution.

prognostic factors (Figure 6). A high number of SBS13 signatures was associated with worse DFI (HR 4.00, 95% CI 1.20–13.30) and PFI (HR 3.15, 95% CI 1.30–7.60), while the presence of SBS45 was linked to improvements in the clinically more significant endpoints DSS (HR 0.27;

95% CI 0.083–0.85) and OS (HR 0.39; 95% CI 0.154–0.97). The predictive power analyzed for the Cox model with ROC showed iAUC value of 0.7367 for DFI, 0.7625 for PFI, 0.9259 for DSS and 0.7824 for OS, with substantial increase in the iAUC value especially with the OS endpoint (Figure 7).

SBS5, SBS40, and SBS45 were found to be significant predictors of DSS and OS in prostate adenocarcinoma but with very wide CIs (Figure 8). Specifically, the high prevalence of SBS40 predicted both decreased DSS (HR 98.73; 95% CI 3.47–2810.71) and OS (HR 8.42; 95% CI 1.243–57.10). The SBS5 signature was associated with shorter DSS (HR 24.27; 95% CI 1.75–335.59) and SBS45 with shorter OS (HR 17.02; 95% CI 2.606–111.20). Adding mutational signatures to the traditional clinical prognostic factors yielded a lot more accurate predictive model, with all SBS iAUC values of 0.9104 for DSS and 0.8347 for OS (Figure 9).

No significant associations between survival and COSMIC mutational signatures were found in renal chromophobe carcinoma, or TGCT after adjusting for traditional prognostic factors.

3.2. Mutational signatures and gene expression analysis

Next, we identified putative links between mutational signature activity (low activity vs. high activity) and gene expression patterns based on TCGA data. To elucidate the potential drivers of the observed survival differences at the gene expression level, these analyses were conducted only for cancer types and signatures, in which statistically significant survival differences were observed in multivariate analyses.

Among the BTCC cohort, between SBS-low and SBS-high subjects, 550 genes were differentially expressed (absolute log2-fold change of > 1 and an adjusted p-value of < 0.05) (Supplementary Figure 1). The most downregulated gene in the SBS2-high group was delta-like 1 homologue, with 9.65-fold (95% CI 1.38; p = 1.9*10⁻⁸) expression in the SBS2-low group in patients with BTCC (Table 2). In the BTCC cohort, somatostatin was among the most downregulated genes in the SBS2 group.

In the ccRCC group, 455 genes were differentially expressed between SBS1-low and SBS1-high subjects. Of these genes, 343 were upregulated in the SBS1-high group, and 90 were immunoglobulin-related genes. The 22 most upregulated genes in the SBS1-high group included four members of the melanoma antigen (MAGE) family: *MAGE-B2* (10.1-fold expression, 95% CI 1.54, p = 0.00018), *MAGE-C2* (7.52-fold expression, 95% CI 1.56, p = 0.00078), *MAGE-A4* (6.81-fold expression, 95% CI 1.75, p = 0.018), and *MAGE-A6* (4.38-fold expression, 95% CI 1.64, p = 0.048). The high expressions of *MAGE-B2*, *MAGE-C2*, *MAGE-A4*, and

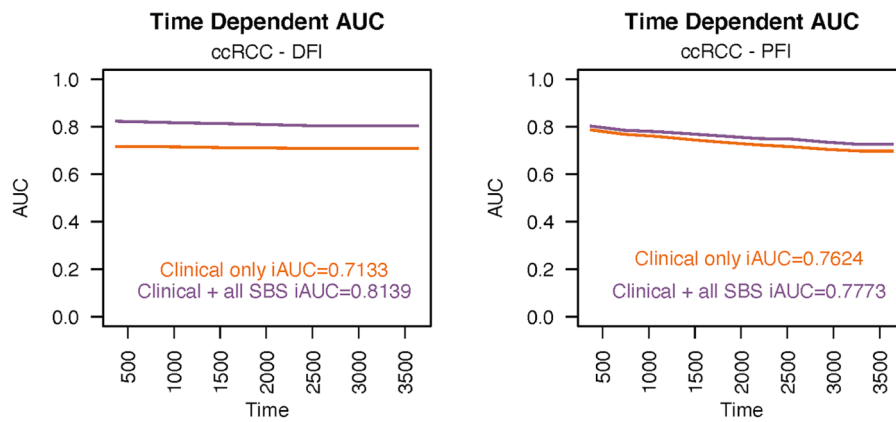


Fig. 5. The time-dependent integrated areas under the curves show (iAUC) values in ccRCC for progression-free interval (A) and for disease-free interval (B). SBS = single-base substitution.

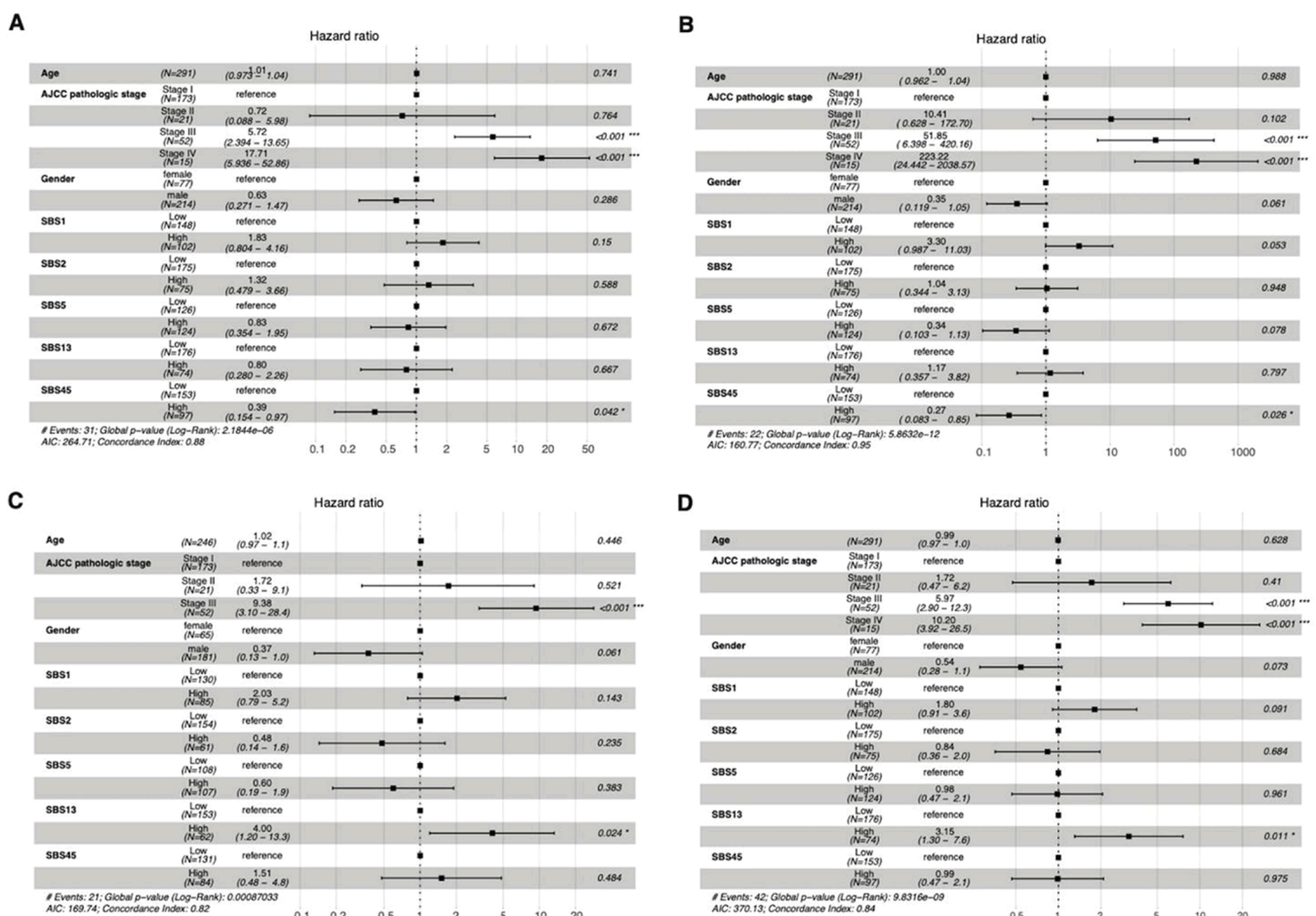


Fig. 6. In pRCC patients, a high number of SBS45 signatures predicted longer OS (A) and DSS (B). Again, a high number of SBS13 signatures were associated with worse DFI (C) and PFI (D). AIC = Akaike information criterion; AJCC = American Joint Committee on Cancer; SBS = single-base substitution.

MAGE-A6 were all associated with poor OS (Figure 10). We also studied if *MAGE-B2*, *MAGE-C2*, *MAGE-A4* or *MAGE-A6* gene expression could provide prognostic power over the traditional prognostic factors in the ccRCC cohort, but the additional prognostic value of the gene expression was very limited (Figure 11A).

As another potentially interesting gene, encoding uromodulin (UMOD, Tamm-Horsfall protein), was 4.69-fold upregulated (95% CI 1.31, $p = 4.7 \times 10^{-6}$) in SBS1-high subjects in the ccRCC cohort. The five most enriched biological processes in SBS1-high patients were all related

to immunoglobulins and immune responses (Supplementary Figure 2).

High SBS45 levels, which were associated with improved DSS and OS in pRCC patients, were associated with 337 differentially expressed genes, of which 312 were downregulated in high SBS45-high patients compared to SBS45-low pRCC patients. *MAGE-A3* was the most downregulated gene in the SBS45-high group (27.6-fold change, 95% CI 1.90; $p = 0.00011$) in the pRCC cohort, and the 10 most downregulated genes also included *MAGE* family members *MAGEC2* and *MAGEC1*, with 12.1-fold change (95% CI 2.08, $p = 0.041$) and 7.98-fold (95% CI 1.84, $p =$

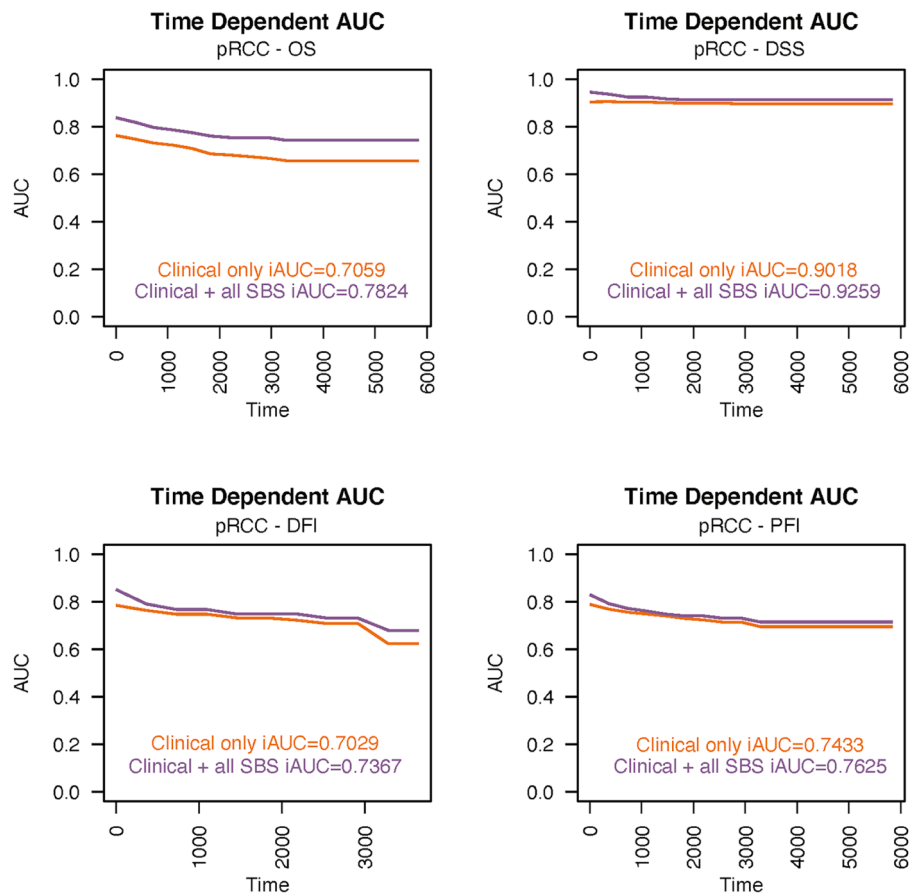


Fig. 7. The integrated areas under the curves (iAUC) values in the pRCC cohort demonstrate the predictive power over time for disease-free interval (A), progression-free interval (B), disease-specific survival (C) and overall survival (D). SBS = single-base substitution.

0.041) downregulations, respectively. Kallikrein-1 was the second most upregulated gene in the SBS45-high group in the pRCC cohort, with a 5.00-fold increase (95% CI 1.32, $p = 3.9 \times 10^{-6}$) compared to the SBS45-low group in the same cohort. The most enriched biological processes within the differentially expressed genes were associated predominantly with processes such as epidermis development, keratinocyte differentiation, and keratinization (Supplementary Figure 3). High SBS45 prevalence was also connected with several genes related to ion transport in patients with pRCC (Figure 12A).

As described above, SBS13 was associated with worse outcomes for pRCC patients. Upon closer investigation of different gene expressions between high and low SBS13 phenotype individuals, 842 genes were differentially expressed between the SBS13 groups, and 744 of them were upregulated. For example, various enzyme activities belong to the most enriched gene ontology molecular function terms describing the differentially expressed genes between SBS13 low and high patient groups (Figure 12B). In contrast with the list of the most upregulated genes in the SBS45-high pRCC cohort, there was a significant upregulation of several MAGE family members in the SBS13-high pRCC cohort: *MAGE-C2* (69.7-fold upregulation, 95% CI 2.154, $p = 5.1 \times 10^{-6}$), *MAGE-C1* (27.7-fold upregulation, 95% CI 1.93, $p = 4.4 \times 10^{-5}$), *MAGE-A11* (12.2-fold upregulation, 95% CI 2.03, $p = 0.013$), *MAGE-A3* (11.3-fold upregulation, 95% CI 1.97, $p = 0.011$) and *MAGE-A12* (10.3-fold upregulation, 95% CI 1.70, $p = 0.00088$) in the SBS13-high group. The high expressions of *MAGE-A3*, *MAGE-C1*, and *MAGE-C2* were all associated with poor OS (Figure 13). When studied in the Cox model, the expression of *MAGE-A3*, *MAGE-A11*, *MAGE-A12*, *MAGE-C1*, and *MAGE-C2* genes was a better predictor than just the traditional prognostic factors, and this model almost reached the prognostic level of mutational signatures and traditional prognostic factors combined

(Figure 11B). Again, kallikrein-related peptidase 3, kallikrein 1, and kallikrein-related peptidase 2 were among the most downregulated genes in the SBS13-high group. In particular, the processes linked to mitosis and nuclear segregation were upregulated in the patients with the worst prognoses, i.e. those with a high SBS13 prevalence (Supplementary Figure 4). Among those, *BUB1* and *survivin (BIRC5)* upregulation were identified as genes of specific interest.

In patients with prostate adenocarcinoma, the SBS45 signature was associated with the enrichment of several pathways related to skeletal muscle development and regulation (Supplementary Figure 5). Again, between SBS40-high and SBS40-low patients, differentially expressed genes especially included those having functions in sperm and spermatid capacitation and function and in antibacterial and antimicrobial humoral responses (Supplementary Figure 6). Instead, SBS5 was associated with several regulators of blood pressure and vasoconstriction (Supplementary Figure 7). Mutational signatures were compared with clinical and pathological parameters, and these results are shown in Supplementary Tables 7–12.

3.3. Mutational burden

As expected, age-related mutational signatures SBS1 and SBS5, in particular, correlated with the total number of mutations (Supplementary Figure 8). The correlation between the individual signatures and the mutational burden was most frequently observed in the BTCC cohort, where SBS1, SBS2, SBS5, and SBS13 were associated with the total mutation count.

Mutation rate of MAGE family genes across urological cancers is shown in Supplementary Figure 9. Highest mutation rates were detected in the BTCC cohort, while mutations in MAGE genes were infrequent in

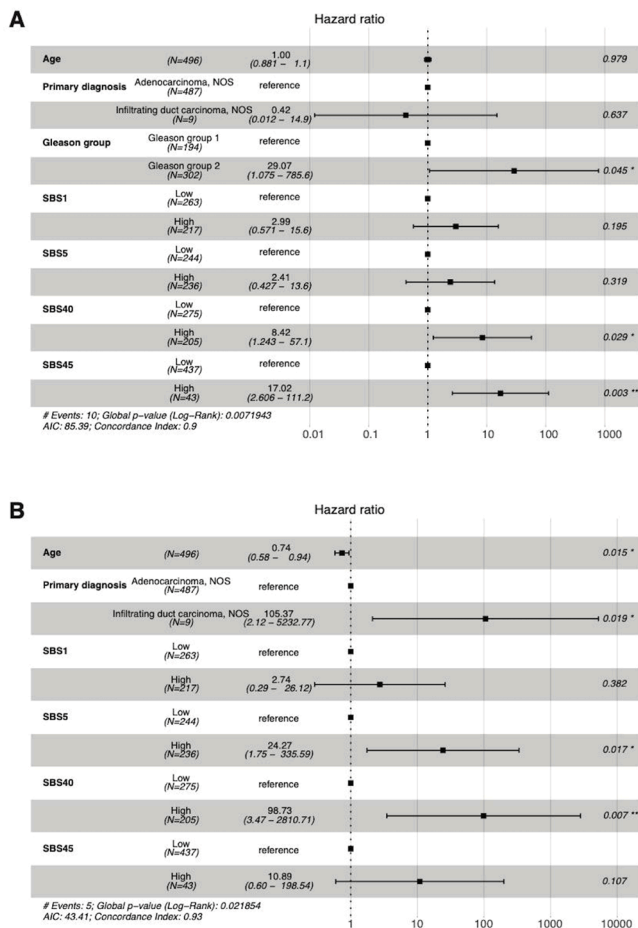


Fig. 8. The high numbers of SBS40 and SBS45 signatures were predictors of poor OS (A) and SBS5 and SBS45 of DSS (B) in patients with prostate adenocarcinoma in Cox multivariate analysis. AIC = Akaike Information Criterion; AJCC = American Joint Committee on Cancer; NOS = not otherwise specified; SBS = single-base substitution.

other cohorts.

4. Discussion

This is the first study to assess the prognostic significance of COSMIC mutational signatures in urological cancers. One of the main findings was the association of a high number of SBS2 signatures with improved

PFI, DSS, and OS in BTCC. In particular, PFI, and DSS have been considered reliable endpoints in this TCGA cohort [7]. The effect of SBS2 on DSS was of a similar magnitude as the traditionally with the most powerful predictor of outcome and stage in the multivariate analysis (HR 0.34, 95% CI 0.16–0.70). Notably, SBS2 activity was not associated with the traditional prognostic factors, and a survival benefit was observed in stage I–III and stage IV BTCCs. The SBS2 signature is formed due to APOBEC deamination of cytosine to uracil, and the APOBEC family of enzymes is a frequent source of hypermutation in most urothelial carcinomas [17]. While uracil subsequently enters the replication process uncorrected and pairs with A, this yields C-to-T mutations, which are characteristic of SBS2 [5,18,19]. Indeed, SBS2 is one of the most well-defined mutational signatures, and it is often accompanied by SBS13. Logically, with the hypermutator role of APOBEC, we observed a strict correlation between the number of SBS2 signatures and the total mutational burden in BTCC.

Recent data suggest that the APOBEC-driven subgroups of high SBS2 and SBS13 would form the largest subgroup of metastatic urothelial carcinomas and that SBS2 and SBS13 would be the major features of radiotherapy-associated BTCC [20,21]. APOBEC-related mutations also seem to drive the aggressiveness of non-muscle invasive BTCC and are associated with tobacco smoking, being an early, a critical event in bladder carcinogenesis [22–25]. Patients with APOBEC-related mutations have compared OS rates compared to patients with APOBEC-low BTCC [22]. The APOBEC-A-high SBS5 phenotype was linked to the absence of recurrences in a series of 62 high-grade, T1 non-muscle-invasive bladder cancers [26]. Non-COSMIC-based APOBEC mutational signatures have also been linked to improved outcomes in patients with advanced urothelial cancers, with most patients having BTCC [27]. To summarize, our results of a favourable prognosis in patients with a high number of SBS2 signatures are in line with the previously reported results regarding the behaviours of less aggressive cancers in APOBEC-high BTCC. Our results complete the previous reports by using COSMIC database, reporting the prognostic value from all stages, and using also additional endpoints than used in the previous studies. In gene expression analysis, we observed notable upregulation of, for example, delta-like noncanonical Notch ligand 1-gene expression in the SBS2-low group (worsened prognosis) in BTCC patients. This gene has been previously linked to a poor prognosis in a variety of different cancers [28], but has remained unstudied in bladder cancers *in vivo*.

In the ccRCC cohort, the high prevalence of the SBS1 signature provided additional prognostic information over staging in terms of decreased DFI and PFI. In particular, PFI has been considered a reliable endpoint in ccRCCs in TCGA data, with data containing 162 progression events in 362 evaluable patients [7]. SBS1 signatures are produced after the deamination of 5-methylcytosine to thymine, generating G:T mismatches in double-stranded DNA and further C-to-T substitutions [5].

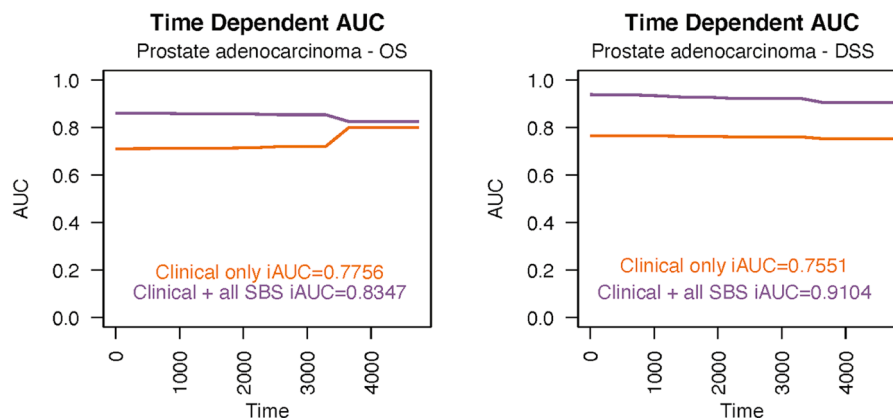


Fig. 9. The integrated areas under the curves (iAUC) values in the prostate adenocarcinoma cohort demonstrate the predictive power over time for disease-specific survival (A) and overall survival (B). SBS = single-base substitution.

Table 2

Examples of the differences in gene expression between the low- and high-signature cohorts. Renal chromophobic carcinoma and TCGT are not included in the table since there were no associations with survival outcomes in these cohorts in the multivariate analysis.

Cancer	Signature	Differentially expressed genes between low-signature and high-signature groups	Examples of differentially expressed genes between low-signature and high-signature groups (fold-change, adjusted p-value)	Examples of enriched GO terms
Bladder transitional cell carcinoma	SBS2 (association with increased survival)	550	Delta-like noncanonical Notch ligand 1 (9.65-fold downregulation, $p = 1.9 \times 10^{-9}$) Somatostatin (7.70-fold downregulation, $p = 8.4 \times 10^{-8}$)	Humoral immune response, regulation of metal ion transport
Clear cell renal cell carcinoma	SBS1 (association with decreased survival)	455	Melanoma antigen-B2 (10.1-fold upregulation, $p = 1.8 \times 10^{-5}$) Melanoma antigen-C2 (7.52-fold upregulation, $p = 0.00078$) Melanoma antigen A4 (6.82-fold upregulation, $p = 0.019$) Uromodulin (4.68-fold upregulation, $p = 4.7 \times 10^{-5}$) Melanoma antigen A6 (4.38-fold upregulation, $p = 0.048$)	Antigen binding, immunoglobulin genes, complement, T-cell/B-cell response related pathway upregulated
Papillary renal cell carcinoma	SBS13 (association with decreased survival)	842	Melanoma antigen-C2 (69.7-fold upregulation, $p = 3.2 \times 10^{-8}$) Melanoma antigen-C1 (27.7-fold upregulation, $p = 4.4 \times 10^{-5}$) Melanoma antigen-A11 (12.2-fold upregulation, $p = 0.013$) Melanoma antigen-A3 (11.3-fold upregulation, $p = 0.011$) Melanoma antigen-A12 (10.3-fold upregulation, $p = 0.000088$) BUB1 (2.05-fold upregulation, $p = 7.8 \times 10^{-7}$) Survivin (2.2-fold upregulation, $p = 3.1 \times 10^{-7}$) Kallikrein-related peptidase 3 (9.45-fold downregulation, $p = 7.9 \times 10^{-8}$) Kallikrein 1 (7.69-fold downregulation, $p = 1 \times 10^{-10}$) Kallikrein-related peptidase 2 (5.36-fold downregulation, $p = 0.00017$)	Mitosis, cell division, nuclear segregation
Papillary renal cell carcinoma	SBS45 (association with increased survival)	337	Melanoma antigen-A3 (27.6-fold downregulation, $p = 0.00011$) Melanoma antigen-C2 (12.1-fold downregulation, $p = 0.041$) Melanoma antigen-C1 (7.98-fold downregulation, $p = 0.042$) Kallikrein-1 (5.00-fold upregulation, $p = 3.9 \times 10^{-6}$)	Several pathways related to ion channel activity downregulated
Prostate adenocarcinoma	SBS5 (association with decreased survival)	144	Cytochrome c oxidase subunit VIa polypeptide 1 (COX6A1) (3.03-fold downregulation, $p = 0.002$) Cytochrome c oxidase subunit 7B2 (5.23-fold upregulation, $p = 0.01$)	Related to vascular system
	SBS40 (association with decreased survival)	56	Prostate and testis expressed 1 (5.77-fold upregulation, $p = 2.7 \times 10^{-10}$)	Sperm function and spermatid development
	SBS45 (association with decreased survival)	108	Myosin light chain 3 (3.83-fold upregulation, $p = 0.00034$) Actin alpha 1, skeletal muscle (2.96-fold upregulation, $p = 0.046$) Troponin C1, slow skeletal, and cardiac type (2.62-fold upregulation, $p = 0.016$) Troponin I1, slow skeletal type (2.37-fold upregulation, $p = 0.042$)	Several pathways related skeletal muscles

GO = gene ontology; SBS = single-base substitution.

The burden of SBS1 is associated with age in virtually all cancer types examined [29], but we still do not observe an association with OS in the ccRCC cohort.

The mutational load of ccRCCs is among the most common of all cancer types, and ccRCCs are consequently among the most immunologically active carcinomas. In our gene expression analysis, immune-related pathways were activated in SBS1-high (poor prognosis) subjects in the ccRCC cohort. Among these, 90 immunoglobulin genes were upregulated in patients with a high SBS1 signature. Several of these immunoglobulin genes have been associated with survival in ccRCC [30]. In line with our results, immune cell infiltration and enhanced

immunological response are established adverse prognostic factors in ccRCC [31–33]. There is no previous literature on COSMIC signatures and prognosis in ccRCCs, and neither has the relationship between SBS1 and immune cell activation been previously evaluated in RCC. In primary and metastatic cutaneous melanomas, SBS1 activity was associated with lower activity of cytotoxic T and NK cells [34].

Several MAGE genes were upregulated in SBS1-high tumours in the ccRCC cohort. MAGE genes are highly conserved in all eukaryotes, and they play a crucial role in adaptation against environmental stress [35]. For example, *MAGE-B2*, which was the most upregulated gene in the SBS1-high ccRCC cohort, promotes stress tolerance, giving a growth

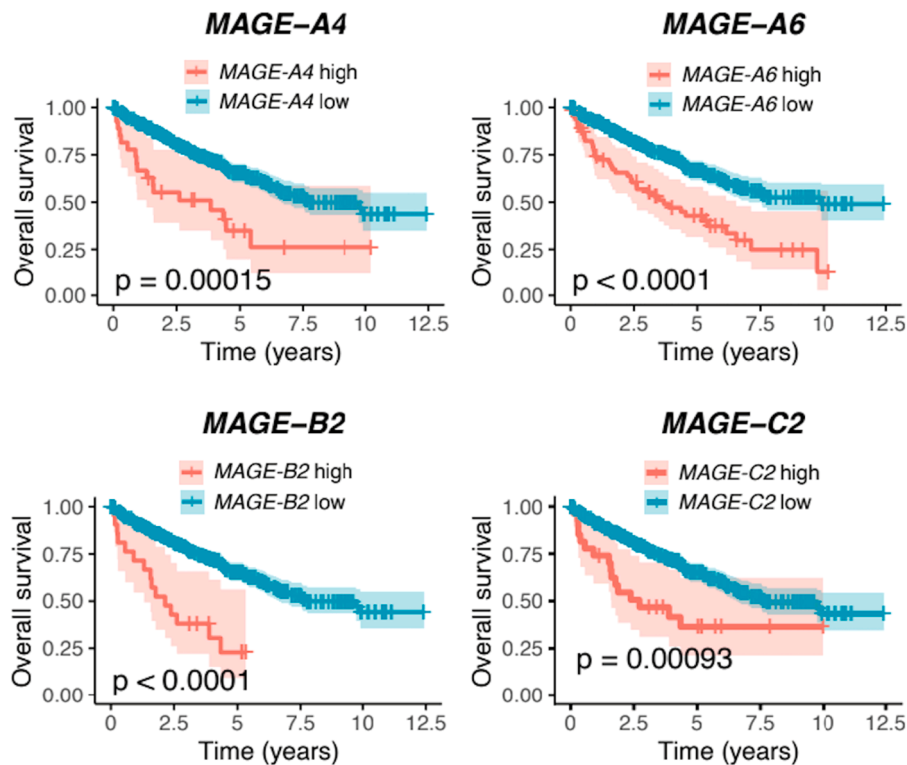


Fig. 10. The high expressions of the *MAGE-B2*, *MAGE-C2*, *MAGE-A4*, and *MAGE-A6* genes were associated with worse overall survival in patients with clear cell renal cell carcinoma (ccRCC) in univariate analysis. These genes were significantly upregulated in patients with a high SBS1 signature (poor prognosis signature) in the ccRCC cohort.

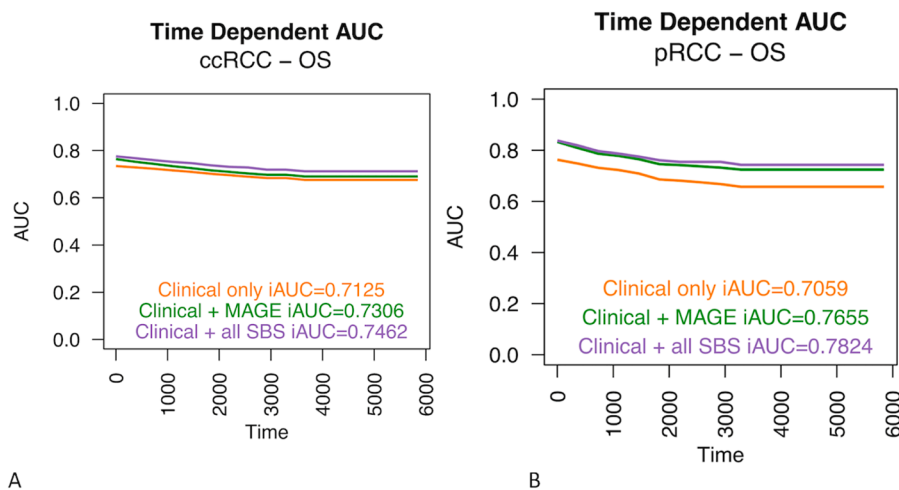


Fig. 11. The expression of genes *MAGE-B2*, *MAGE-C2*, *MAGE-A4* or *MAGE-A6* provided only minor additional prognostic value over the traditional prognostic (clinical) values in the ccRCC cohort (A). In pRCC, the expression of *MAGE-A3*, *MAGE-A11*, *MAGE-A12*, *MAGE-C1*, and *MAGE-C2* genes was a better predictor when compared only to the traditional prognostic factors. This model also almost reached the prognostic level of mutational signatures and traditional prognostic factors combined (B).

advantage to cancer cells, and its high expression has been implicated in tumour growth and progression in various cancers *in vitro* but has not been studied previously in RCCs [35–38]. *MAGE-C2*, with the observed 7.4-fold upregulation in SBS1 ccRCCs, is a cancer cell-specific regulator of TRIM28, and its high expression is associated with poor survival in various malignancies, including bladder and prostate carcinomas; however, the role of *MAGE-C2* has not yet been evaluated in RCCs [39–42]. Other highly induced MAGE genes in the SBS1-high group in the ccRCC cohort included *MAGE-A4* and *MAGE-A6*. In particular, *MAGE-A6* seems to have significant pro-oncogenic roles [43,44], whereas the role of *MAGE-A4* is less well understood. The high RNA levels of *MAGE-B2*, *MAGE-A4*, and *MAGE-A6* were related to a dismal prognosis. A single potentially interesting gene, *UMOD*, was > 4.5-fold

upregulated in SBS1-high subjects in the patients with ccRCC. For a small set of samples, *UMOD* RNA expression has been reported as one of the most upregulated transcripts in ccRCCs compared to benign renal tissue [45]. Based on our ccRCC results, it is possible that the dismal prognosis for SBS1-high patients may be driven by the upregulation of the immune system and MAGE family genes.

Of all renal cancers, pRCC accounts for approximately 15% and is genetically and morphologically distinct from other RCCs [46]. With the exception of staging, reliable prognostic factors are virtually absent in pRCC. According to our analyses, a high number of SBS13 signatures was associated with worse DFI and PFI in the TCGA pRCC cohort. SBS1 and SBS13 have also been attributed to the activity of AID/APOBEC family of cytidine deaminases [5,47]. However, while high APOBEC

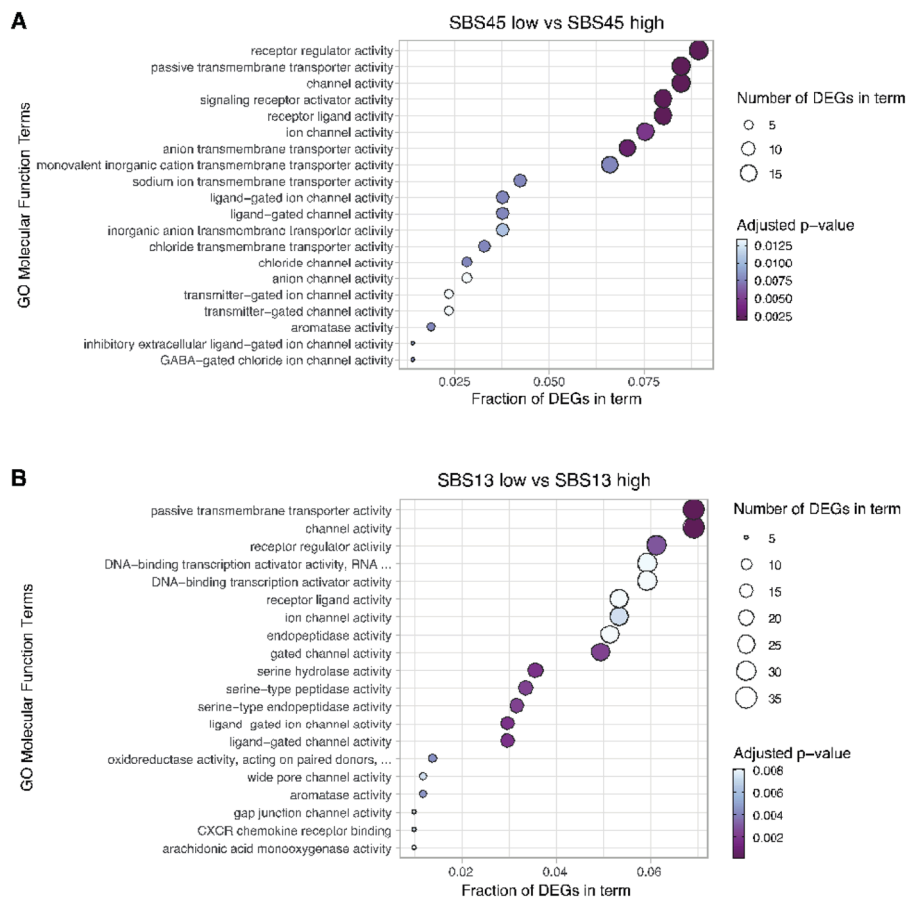


Fig. 12. Top twenty significantly enriched (adjusted p-value of < 0.05) gene ontology molecular function terms in differentially expressed genes between SBS45-low and SBS45-high (A) and SBS13-low and SBS13-high (B) patients with papillary renal cell carcinoma.

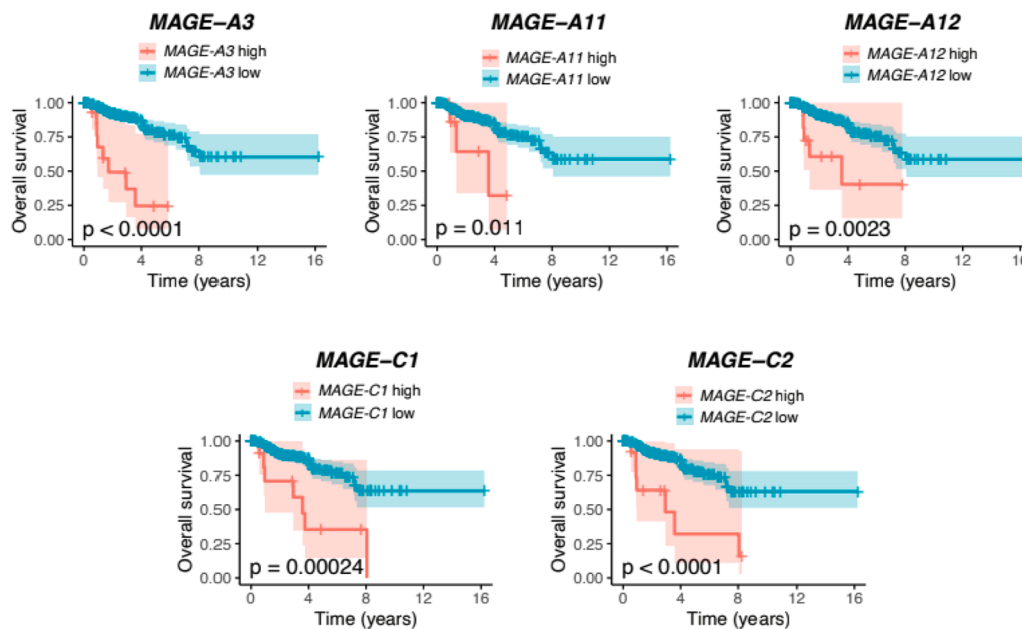


Fig. 13. The high expressions of the *MAGE-A3*, *MAGE-A11*, *MAGE-A12*, *MAGE-C1*, and *MAGE-C2* genes were associated with worse overall survival in patients with papillary renal cell carcinoma (pRCC) in univariate analysis. These genes were significantly upregulated in patients with a high SBS13 signature (poor prognosis signature) in the pRCC cohort.

activity has been constantly linked to improved outcomes in BTCC, there are no previous data on the role of APOBEC in pRCC. In line with our results, high SBS13 activity is associated with a higher risk of post-resection recurrence in non-small-cell lung cancers [48] and with radioiodine treatment resistance in papillary thyroid carcinomas [49]. We found that SBS13 activity was associated with the upregulation of 744 genes in pRCC, which were mainly linked to mitosis, cell division, and nuclear segregation. Notably, several MAGE family members were upregulated in pRCC in the SBS13 high cohort. The most upregulated of all genes in this cohort was *MAGE-C2*, which was upregulated 70-fold in the SBS13 high group in the pRCC cohort. In addition, *MAGE-C1*, *MAGE-A11*, *MAGE-A3*, and *MAGE-A12* were among the most upregulated genes, with at least a 10-fold increase in the SBS13 high group.

In urological cancers, *MAGE-A11* is best known for its proto-oncogenic role in prostate cancer [50]. Although MAGE family members in RCCs have rarely been studied, there is evidence that a specific single-nucleotide polymorphism of *MAGE-A11* provides prognostic information on OS in RCCs in materials of mixed histologies [51]. The *MAGE-A3* gene shares a 96% sequence similarity with more well-characterized prooncogenic *MAGE-A6* [52]. *MAGE-C1* is highly expressed in several cancers and has prognostic relevance in multiple myeloma and breast cancers [53,54]. The possible biological relationship between APOBEC enzymes and MAGE family members or kallikreins remains unstudied. However, it is known that the expression of *MAGE-A* genes is regulated by DNA methylation [55], which could be modified by APOBEC enzymes.

The high prevalence of SBS45 signatures was associated with improved OS and DSS in pRCCs, and in the time-dependent SBS45-including model the iAUC value was as high as 0.93 for DSS. Adding mutational signatures to the model with only traditional clinical variables increased especially the predictive power of OS analysis in the pRCC cohort. Among the SBS45-high patients, *MAGE-A3*, *MAGE-C1*, and *MAGE-C2* were among the most downregulated genes, and kallikrein 1 was highly upregulated in the SBS45-high patients. The function of these genes, and the respective proteins with tumour-suppressive functions, are well-characterised in prostate carcinogenesis, for example, but no previous literature exists on their role in RCCs [56,57]. The high RNA expression of *MAGE-A3*, *MAGE-C1*, and *MAGE-C2* individually associated with dismal survival rates in pRCC. According to the ROC analysis, gene expression of the five selected *MAGE* genes almost reached the prognostic significance of mutational signatures and traditional prognostic factors combined and exceeded substantially the prognostic power of that of just the traditional prognostic factors. In clinical practice, using gene expression analysis could be more feasible prognostic tool than mutational signatures requiring whole-exome sequencing (WES). It is worth noting that the SBS45 signature may be a sequencing artefact due to the use of 8-oxoguanine during sequencing [58].

In patients with prostate cancer, SBS5, SBS40, and SBS45 were associated with worse OS and DSS. There were very wide CIs in these analyses due to the limited number of events in this disease, which has a generally favourable prognosis. In addition, only PFI and DFI have been considered reliable endpoints in prostate cancer, while OS and DSS should be used with caution [7]. Thus, this study could not reveal any usable prognostic factors for use in patients with prostate cancer. We propose that the prognostic value of mutational signatures should be evaluated in the future, preferably in local high-risk or metastasized prostate cancers.

Despite urological TCGA cohorts containing a substantial number of patients with comprehensive clinical parameters, the major limitation of this study was the use of only a single dataset. WES is currently performed only to a minority of cancer patients, which limits the immediate implementation of the WES-based prognostic tools. As discussed above and by [7], the sample sizes in different TCGA cohorts vary considerably, and the usability of different survival endpoints is diverse. As TCGA contains only exome data, not all mutations in the human genome are covered. This is also a likely reason for our final analyses, including

only SBSs, but not the other types of mutational signatures. Furthermore, mutations occurring at noncoding regulatory regions, including promoters, distal genomic enhancers, and silencer elements, might play crucial roles in gene expression levels. As the current study was not able to evaluate if the observed changes in gene expressions were drivers or passengers, this topic requires functional studies in the future. No significant associations between survival and COSMIC mutational signatures were found in chromophobe RCCs or in TGCT after adjusting for the traditional prognostic factors. This is due, at least partly, to the small number of events in both datasets.

Conclusions

Several COSMIC signatures seem to provide prognostic information in urological cancers, and a significant proportion of the prognostic signatures are related to the activation of APOBEC enzymes. Based on the iAUC values, adding mutational signatures to Cox models seem to provide best prognostic information in pRCC with the iAUC values up to 0.93. The prognostic relevance was worst in BTCC, although mutational signatures are able to provide additional prognostic information also for BTCC patients over traditional prognostic factors. We presented a novel finding of the prominent upregulation of the genes belonging to the MAGE family in the groups where the signature predicted worse survival (SBS1 in ccRCC and SBS13 in pRCC) and the respective downregulation of MAGE genes in signature group with improved survival (SBS45 in pRCC). With using only selected MAGE genes we were also able to provide prognostic information near to that of mutational signatures and clinical variables combined. This suggests that MAGE genes could be key drivers of the current results and deserve further functional studies.

Funding

No reportable funding was used to conduct this study.

Author contributions

PK drafted the manuscript. All authors participated in the planning the study, evaluating the results and writing the final versions of the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The datasets used and analyzed are available on reasonable request.

Acknowledgments

We thank Genevia Technologies for their help in statistical analyses and in composing the figures and tables. Open access funded by Helsinki University Library.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.neo.2023.100933](https://doi.org/10.1016/j.neo.2023.100933).

References

- [1] H Sung, J Ferlay, RL Siegel, M Laversanne, I Soerjomataram, A Jemal, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, *CA Cancer J. Clin.* 71 (2021) 209–249.
- [2] RJ Motzer, NM Tannir, DF McDermott, M Burotto, TK Choueiri, Conditional survival and 5-year follow-up in CheckMate 214: first-line nivolumab + ipilimumab (N+I) versus sunitinib (S) in advanced renal cell carcinoma (aRCC), *Ann. Oncol.* 32 (2021) S678.
- [3] AT Lenis, PM Lec, K Chamie, MD. Mshs, Bladder cancer: a review, *JAMA* 324 (19) (2020 Nov 17) 1980–1991.
- [4] LB Alexandrov, MR Stratton, The patterns of somatic mutations hidden in cancer genomes, *Curr. Opin. Genet. Dev.* 24 (2014) 52–60.
- [5] LB Alexandrov, J Kim, NJ Haradhvala, MN Huang, AW Tian Ng, Y Wu, The repertoire of mutational signatures in human cancer, *Nature* 578 (2020) 94–101.
- [6] G Koh, A Degasperi, X Zou, S Momen, S. Nik-Zainal, Mutational signatures: emerging concepts, caveats and clinical applications, *Nat. Rev. Cancer* 21 (2021) 619–637.
- [7] J Liu, T Lichtenberg, KA Hoadley, LM Poisson, AJ Lazar, AD Cherniack, AJ Kovatich, CC Benz, DA Levine, AV Lee, L Omberg, DM Wolf, CD Shriver, V Thorsson, Cancer Genome Atlas Research Network, H. Hu, An integrated TCGA pan-cancer clinical data resource to drive high quality survival outcome analytics, *Cell* 173 (2) (2018) 400–416, e11.
- [8] RL Grossman, AP Heath, V Ferretti, HE Varmus, DR Lowy, WA Kibbe, LM. Staudt, Toward a shared vision for cancer genomic data, *N. Engl. J. Med.* 375 (12) (2016 Sep 22) 1109–1112.
- [9] A Colaprico, TC Silva, C Olsen, L Garofano, C Cava, D Garolini, TS Sabedot, TM Malta, SM Pagnotta, I Castiglioni, M Ceccarelli, G Bontempi, H. Noushmehr, TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data, *Nucleic. Acids. Res.* 44 (8) (2016 May 5) e71.
- [10] Therneau, T. A package for survival analysis. In R Package Version 3.2–7. 2020. <https://CRAN.R-project.org/package=survival>.
- [11] Kassambara, A, Kosinski, M, Biecek, P. *Survminer: drawing survival curves using 'ggplot2'*. R Package Version 0.4.8. 2020. <https://CRAN.R-project.org/package=survminer>.
- [12] DD Sjoberg, K Whiting, M Curry, JA Lavery, J. Larmarange, Reproducible summary tables with the *gtsummary* package, *R. J.* 2021; 13 (2021) 570–580.
- [13] X Song, X-H. Zhou, A semiparametric approach for the covariate specific ROC curve with survival outcome, *Stat. Sin.* 18 (2008) 947–965.
- [14] MI Love, W Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, *Genome Biol.* 15 (2014) 550.
- [15] G Yu, L-G Wang, Y Han, Q-Y. He, clusterProfiler: an R package for comparing biological themes among gene clusters, *OMICS* 16 (2012) 284–287.
- [16] T Wu, E Hu, S Xu, M Chen, P Guo, Z Dai, T Feng, L Zhou, W Tang, L Zhan, Fu x, S Liu, X Bo, G Yu, clusterProfiler 4.0: a universal enrichment tool for interpreting omics data, *Innovation (Camb)* 2 (3) (2021), 100141, <https://doi.org/10.1016/j.xinn.2021.100141>.
- [17] AP Glaser, D Fantini, A Shilatifard, EM Schaeffer, JJ. Meeks, The evolving genomic landscape of urothelial carcinoma, *Nat. Rev. Urol.* 14 (4) (2017 Feb 7) 215–229.
- [18] S Nik-Zainal, H Davies, J Staaf, M Ramakrishna, D Glodzik, X Zou, I Martincorena, LB Alexandrov, S Martin, DC Wedge, P Van Loo, YS Ju, M Smid, AB Brinkman, S Morganello, MR Aure, OC Lingjærde, A Langerød, M Ringnér, SM Ahn, S Boyault, JE Brock, A Broeks, A Butler, C Desmedt, L Dirix, S Dronov, A Fatima, JA Foekens, M Gerstung, GK Hooijer, SJ Jang, DR Jones, HY Kim, TA King, S Krishnamurthy, HJ Lee, JY Lee, Y Li, S McLaren, A Menzies, V Mustonen, S O'Meara, I Pauptert, X Pivot, CA Purdie, K Raine, K Ramakrishnan, FG Rodríguez-González, G Romieu, AM Sieuwerts, PT Simpson, R Shepherd, L Stebbings, OA Stefansson, J Teague, S Tommasi, I Treilleux, GG Van den Eynden, P Vermeulen, A Vincent-Salomon, L Yates, C Caldas, L van't Veer, A Tutt, S Knappskog, BK Tan, J Jonkers, Å Borg, NT Ueno, C Sotiriou, A Viari, PA Futreal, PJ Campbell, PN Span, S Van Laere, SR Lakhani, JE Eyfjord, AM Thompson, E Birney, HG Stunnenberg, MJ van de Vijver, JW Martens, AL Børresen-Dale, AL Richardson, G Kong, G Thomas, MR. Stratton, Landscape of somatic mutations in 560 breast cancer whole-genome sequences, *Nature* 2 (534) (2016) 47–54.
- [19] K Chan, SA Roberts, LJ Klimczak, JF Sterling, N Saini, EP Malc, J Kim, DJ Kwiatkowski, DC Fargo, PA Mieczkowski, G Getz, DA. Gordenin, An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers, *Nat. Genet.* 47 (9) (2015 Sep) 1067–1072.
- [20] JA Nakamura-González, M Rijnders, J van Riet, MS van der Heijden, J Voortman, E Cuppen, N Mehra, S van Wilpe, SF Oosting, LL Rijstenberg, HM Westgeest, EC Zwarthoff, R de Wit, AAM van der Veldt, HJG van de Werken, MPJ Lolkema, JL. Boormans, Comprehensive molecular characterization reveals genomic and transcriptomic subtypes of metastatic urothelial carcinoma, *Eur. Urol.* 81 (4) (2022 Apr) 331–336.
- [21] M Mossanen, FLF Carvalho, V Muralidhar, MA Preston, B Reardon, JR Conway, C Curran, D Freeman, S Sha, G Sonpavde, M Hirsch, AS Kibel, EM Van Allen, KW. Mouw, Genomic features of muscle-invasive bladder cancer arising after prostate radiotherapy, *Eur. Urol.* 81 (5) (2022 May) 466–473.
- [22] AG Robertson, J Kim, H Al-Ahmadie, J Bellmunt, G Guo, AD. Cherniack, Comprehensive molecular characterization of muscle-invasive bladder cancer, *Cell* 171 (2017) 540–556.
- [23] J Kim, A Akbani, CJ Creighton, SP Lerner, JN Weinstein, G. Getz, Invasive bladder cancer: genomic insights and therapeutic promise, *Clin. Cancer Res.* 21 (2015) 4514–4524.
- [24] SV Lindskog, F Prip, P Lamy, A Taber, CS Groeneveld, K Birkenkamp-Demtröder, JB Jensen, T Strandgaard, I Nordentoft, E Christensen, M Sokac, NJ Birkbak, L Maretty, GG Hermann, AC Petersen, V Weyerer, MO Grimm, M Horstmann, G Sjødahl, M Höglund, T Steiniche, K Mogensen, A de Reyniès, R Nawroth, B Jordan, X Lin, D Dragicic, DG Ward, A Goel, CD Hurst, JD Raman, JI Warrick, U Segersten, D Sikic, KEM van Kessel, T Maurer, JJ Meeks, DJ DeGraff, RT Bryan, MA Knowles, T Simic, A Hartmann, EC Zwarthoff, PU Malmström, N Malats, FX Real, L. Dyrskjøet, An integrated multi-omics analysis identifies prognostic molecular subtypes of non-muscle-invasive bladder cancer, *Nat. Commun.* 12 (1) (2021 Apr 16) 2301.
- [25] S Koutros, N Rao, LE Moore, ML Nickerson, D Lee, B Zhu, LA Pardo, D Baris, M Schwenn, A Johnson, K Jones, M Garcia-Closas, L Prokunina-Olsson, DT Silverman, N Rothman, M. Dean, Targeted deep sequencing of bladder tumors reveals novel associations between cancer gene mutations and mutational signatures with major risk factors, *Clin. Cancer Res.* 27 (13) (2021 Jul 1) 3725–3733.
- [26] J Bellmunt, J Kim, B Reardon, J Perera-Bel, A Orsola, A Rodriguez-Vida, SA Wankowicz, M Bowden, JA Barletta, J Morote, I de Torres, N Juanpere, J Lloreta-Trull, S Hernandez, KW Mouw, ME Taplin, P Cejas, HW Long, EM Van Allen, G Getz, DJ Kwiatkowski, Genomic predictors of good outcome, recurrence, or progression in high-grade T1 non-muscle-invasive bladder cancer, *Cancer Res.* 80 (20) (2020 Oct 15) 4476–4486.
- [27] D Natesan, L Zhang, HJ Martell, T Jindal, P Devine, B Stohr, C Espinosa-Mendez, J Grenet, J Van Ziffle, N Joseph, S Umetsu, C Onodera, M Turski, E Chan, A Desai, R Aggarwal, A Wong, S Porten, J Chou, T Friedlander, L Fong, EJ Small, A Sweet-Cordero, VS. Koshkin, APOBEC mutational signature and tumor mutational burden as predictors of clinical outcomes and treatment response in patients with advanced urothelial cancer, *Front. Oncol.* 12 (2022 Mar 7), 816706.
- [28] JFH Pittaway, C Lipsos, K Mariniello, L. Guasti, The role of delta-like non-canonical Notch ligand 1 (DLK1) in cancer, *Endocr. Relat. Cancer* 28 (12) (2021 Oct 15) R271–R287.
- [29] LB Alexandrov, PH Jones, DC Wedge, JE Sale, PJ Campbell, S Nik-Zainal, MR. Stratton, Clock-like mutational processes in human somatic cells, *Nat. Genet.* 47 (2015) 1402–1407.
- [30] ZN Xia, XY Wang, LC Cai, WG Jian, C. Zhang, IGLL5 is correlated with tumour-infiltrating immune cells in clear cell renal cell carcinoma, *FEBS Open Bio* 11 (3) (2021 Mar) 898–910.
- [31] X Gao, J Yang, Y Chen, Identification of a four immune-related gene signature based on an immunogenomic landscape analysis of clear cell renal cell carcinoma, *J. Cell. Physiol.* 235 (12) (2020) 9834–9850.
- [32] CM Diaz-Montero, BI Rini, JH. Finke, The immunology of renal cell carcinoma, *Nat. Rev. Nephrol.* 16 (12) (2020) 721–735.
- [33] H Shen, J Liu, S Chen, X Ma, Y Yang, J Li, W Wang, X Wang, L. Xie, Prognostic value of tumor-associated macrophages in clear cell renal cell carcinoma: a systematic review and meta-analysis, *Front. Oncol.* 11 (2021 Apr 26), 657318.
- [34] C Roufas, I Georgakopoulos-Soares, A. Zaravinos, Distinct genomic features across cytolytic subgroups in skin melanoma, *Cancer Immunol. Immunother.* 70 (11) (2021 Nov) 3137–3154.
- [35] Z Bai, Y Zhao, X Yang, L Wang, X Yin, Y Chen, J. Lu, A novel prognostic ferroptosis-related long noncoding RNA signature in clear cell renal cell carcinoma, *J. Oncol.* 2022 (2022 Feb 22), 6304824.
- [36] LY Peche, MF Ladelfa, MF Toledo, M Mano, JE Laiseca, C Schneider, M Monte, Human MAGEB2 protein expression enhances E2F transcriptional activity, cell proliferation, and resistance to Ribotoxic stress, *J. Biol. Chem.* 290 (49) (2015 Dec 4) 29652–29662.
- [37] KM Pattani, E Soudry, CA Glazer, MF Ochs, H Wang, J Schussel, W Sun, P Hennessy, W Mydlarz, M Loyo, S Demokan, IM Smith, JA. Califano, MAGEB2 is activated by promoter demethylation in head and neck squamous cell carcinoma, *PLoS One* 7 (9) (2012) e45534.
- [38] R Yamada, A Takahashi, T Torigoe, R Morita, Y Tamura, T Tsukahara, T Kanaseki, T Kubo, K Watarai, T Kondo, Y Hirohashi, N. Sato, Preferential expression of cancer/testis genes in cancer stem-like cells: proposal of a novel sub-category, cancer/testis/stem gene, *Tissue Antigens* 81 (6) (2013 Jun) 428–434.
- [39] L von Boehmer, L Keller, A Mortezaei, M Provenzano, G Sais, T Hermanns, T Sulser, AA Jungbluth, LJ Old, G Kristiansen, M van den Broek, H Moch, A Knuth, PJ. Wild, MAGE-C2/CT10 protein expression is an independent predictor of recurrence in prostate cancer, *PLoS One* 6 (7) (2011) e21366.
- [40] A Hodgson, AA Jungbluth, N Katani, B Xu, MR. Downes, Evaluation of cancer testis antigen (CT10, PRAME) and MHC I expression in high-grade urothelial carcinoma of the bladder, *Virchows. Arch.* 476 (4) (2020 Apr) 535–542.
- [41] Q Zhao, WT Xu, T. Shalier, Pilot study on MAGE-C2 as a potential biomarker for triple-negative breast cancer, *Dis. Markers* 2016 (2016), 2325987.
- [42] X Chen, L Wang, J Liu, L Huang, L Yang, Q Gao, X Shi, J Li, F Li, Z Zhang, S Zhao, B Zhang, P Van der Bruggen, Y. Zhang, Expression and prognostic relevance of MAGE-A3 and MAGE-C2 in non-small cell lung cancer, *Oncol. Lett.* 13 (3) (2017 Mar) 1609–1618.
- [43] CT Pineda, S Ramanathan, K Fon Tacer, JL Weon, MB Potts, YH Ou, MA White, PR Potts, Degradation of AMPK by a cancer-specific ubiquitin ligase, *Cell* 160 (25679763) (2015) 715–728.
- [44] Yang, MAGE-A, MAGE-B, and MAGE-C proteins form complexes with KAP1 and suppress p53-dependent apoptosis in mage-positive cell lines, *Cancer Res.* 67 (17942928) (2007) 9954–9962.
- [45] OS Eikrem, P Strauss, C Beisland, A Scherer, L Landolt, A Flatberg, S Leh, V Beisvag, T Skogstrand, K Hjellev, A Shresta, HP. Marti, Development and confirmation of potential gene classifiers of human clear cell renal cell carcinoma using next-generation RNA sequencing, *Scand. J. Urol.* 50 (6) (2016 Dec) 452–462.

- [46] N Mendhiratta, P Muraki, AE Sisk Jr., B. Shuch, Papillary renal cell carcinoma: review, *Urol. Oncol.* 39 (6) (2021 Jun) 327–337.
- [47] PJ Stephens, PS Tarpey, H Davies, P Van Loo, C Greenman, DC Wedge, S Nik-Zainal, S Martin, I Varela, GR Bignell, LR Yates, E Papaemmanuil, D Beare, A Butler, A Cheverton, J Gamble, J Hinton, M Jia, A Jayakumar, D Jones, C Latimer, KW Lau, S McLaren, DJ McBride, A Menzies, L Mudie, K Raine, R Rad, MS Chapman, J Teague, D Easton, A Langerød, Oslo Breast Cancer Consortium (OSBREAC), MT Lee, CY Shen, BT Tee, BW Huimin, A Broeks, AC Vargas, G Turashvili, J Martens, A Fatima, P Miron, SF Chin, G Thomas, S Boyault, O Mariani, SR Lakhani, M van de Vijver, L van 't Veer, J Foekens, C Desmedt, C Sotiriou, A Tutt, C Caldas, JS Reis-Filho, SA Aparicio, AV Salomon, AL Børresen-Dale, AL Richardson, PJ Campbell, PA Futreal, MR Stratton, The landscape of cancer genes and mutational processes in breast cancer, *Nature* 486 (2012) 400–404.
- [48] R Caso, F Sanchez-Vega, KS Tan, B Mastrogiacomo, J Zhou, GD Jones, B Nguyen, N Schultz, JG Connolly, WS Brandt, MJ Bott, G Rocco, D Molena, JM Isbell, Y Liu, MW Mayo, PS Adusumilli, WD Travis, DR. Jones, The underlying tumor genomics of predominant histologic subtypes in lung adenocarcinoma, *J. Thorac. Oncol.* 15 (12) (2020 Dec) 1844–1856.
- [49] S Siraj, T Masoodi, AK Siraj, S Azam, Z Qadri, SK Parvathareddy, R Bu, KS Siddiqui, SS Al-Sobhi, M AlDawish, KS Al-Kuraya, APOBEC SBS13 mutational signature—a novel predictor of radioactive iodine refractory papillary thyroid carcinoma, *Cancers (Basel)* 14 (6) (2022 Mar 21) 1584.
- [50] YM. Willett, Evolution of melanoma antigen-A11 (MAGEA11) during primate phylogeny, *J. Mol. Evol.* 86 (29574604) (2018) 240–253.
- [51] S Su, Q Gu, A Xu, S Shen, S Liu, C Zhang, C Miao, C Qin, B Liu, Z. Wang, Genetic variations in MAGE-A11 predict the risk and survival of renal cell cancer, *J. Cancer* 10 (20) (2019 Aug 27) 4860–4865.
- [52] CT Pineda, PR. Potts, Oncogenic MAGEA-TRIM28 ubiquitin ligase downregulates autophagy by ubiquitinating and degrading AMPK in cancer, *Autophagy* 11 (5) (2015) 844–846.
- [53] S Hou, M Sang, L Zhao, R Hou, B. Shan, The expression of MAGE-C1 and MAGE-C2 in breast cancer and their clinical significance, *Am. J. Surg.* 211 (1) (2016 Jan) 142–151.
- [54] C Pabst, J Zustin, F Jacobsen, T Luetkens, N Kröger, G Schilling, C Bokemeyer, G Sauter, D Atanackovic, A. Marx, Expression and prognostic relevance of MAGE-C1/CT7 and MAGE-C2/CT10 in osteolytic lesions of patients with multiple myeloma, *Exp. Mol. Pathol.* 89 (2) (2010 Oct) 175–181.
- [55] A Colemon, TM Harris, S. Ramanathan, DNA hypomethylation drives changes in MAGE-A gene expression resulting in alteration of proliferative status of cells, *Genes Environ.* 42 (2020) 24–36.
- [56] S Hassan, T Blick, EW Thompson, ED. Williams, Diversity of epithelial-mesenchymal phenotypes in circulating tumour cells from prostate cancer patient-derived xenograft models, *Cancers (Basel)* 13 (11) (2021 Jun 1) 2750.
- [57] Z Zhang, H Wu, H Zhou, Y Gu, Y Bai, S Yu, R An, J. Qi, Identification of potential key genes and high-frequency mutant genes in prostate cancer by using RNA-Seq data, *Oncol. Lett.* 15 (4) (2018 Apr) 4550–4556.
- [58] M Costello, TJ Pugh, TJ Fennell, C Stewart, L Lichtenstein, JC Meldrim, JL Fostel, DC Friedrich, D Perrin, D Dionne, S Kim, SB Gabriel, ES Lander, S Fisher, G. Getz, Discovery and characterization of artifactual mutations in deep coverage targeted capture sequencing data due to oxidative DNA damage during sample preparation, *Nucleic. Acids. Res.* 41 (6) (2013 Apr 1) e6.