PBRM1 loss-of-function mutations and response to immune checkpoint blockade in clear cell renal cell carcinoma

AUTHORS: Jake Conway¹,², Amaro Taylor-Weiner², David Braun¹, Ziad Bakouny¹, Toni K. Choueiri¹, Eliezer M. Van Allen¹

AFFILIATIONS:
1. Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA
2. Biomedical Informatics and Genomics Program, Harvard Medical School, Boston, MA

ABSTRACT
The biological role of PBRM1 loss-of-function (LOF) mutations mediating response to immune checkpoint blockade (ICB) in clear cell renal cell carcinoma (ccRCC) remains an area of active investigation. A recent study (hereafter, ‘the MSK study’) evaluated this hypothesis in a clinically heterogeneous single institution retrospective ccRCC cohort using targeted sequencing and did not find a univariable statistically significant association for PBRM1 LOF and time-to-treatment failure (TTF), but did report a significant association between continuous tumor mutational burden (TMB) and ICB TTF. We re-analyzed this cohort to (i) match the methodology from previous studies and (ii) account for differences in cohort development, statistical approaches, mutation classifications, and outcome measurements. Univariable analysis of the PBRM1 LOF mutation hypothesis, when performed analogously to prior studies, demonstrated a modest statistically significant association with TTF on ICB (p = 0.047; HR = 0.7, 95% CI = 0.49 - 0.99). By contrast, when using the different methodology presented in the MSK study, statistical modeling demonstrated that there was insufficient power to detect this association given the altered classification scheme and modest effect size. In addition, after appropriately normalizing TMB to account for anomalous outliers, TMB was not significantly associated with ICB response in ccRCC. Thus, this study provides further support for the biological association between PBRM1 LOF and favorable clinical outcomes on ICB in specific ccRCC clinical contexts. However, we again strongly caution against interpreting that this PBRM1 biological association, likely one of many modest mediators of tumor-immune-stromal interactions, achieves broadly generalizable predictive biomarker status for ICB in diverse clinical contexts. The PBRM1 LOF biological association has only been shown to be modestly predictive in the post-VEGF TKI ICB-therapy ccRCC patient population. Given emerging data about the biological role of PBRM1 in contributing to interwoven immune and angiogenesis programs (among many others), and the rapidly shifting therapeutic combination strategies invoking these processes in ccRCC, focused biological and predictive biomarker analyses in specific clinical settings using appropriately annotated clinical trial cohorts are necessary.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
MAIN TEXT
Two clear cell renal cell carcinoma (ccRCC) clinical trial cohorts evaluating single-agent immune checkpoint blockade (ICB) in a VEGF receptor tyrosine kinase inhibitor (VEGF TKI)-refractory patient population identified a statistically significant association between PBRM1 loss-of-function (LOF) mutations and ICB response (1-3), whereas two trial cohorts in the first-line treatment-naive metastatic setting with different ICB-based therapeutic combinations did not observe a significant association (4, 5). Hakimi et al (hereafter, ‘the MSK study’) investigated associations of mutations in PBRM1 and other PBAF complex members with response to immune checkpoint blockade (ICB) in a clinically heterogeneous single institution retrospective cohort of ccRCC, and generally across cancer types (6). They did not observe a significant association for PBRM1 LOF in ccRCC in univariable analysis, nor in multivariable analysis incorporating two factors they found significantly associated with ICB response in ccRCC: tumor mutational burden (TMB) and treatment class. These observations raise key questions for the field about the biological and biomarker relevance of PBRM1 mutations in ccRCC and cancer immunotherapy, and also more generally interpretation of candidate molecular mediators of ICB response from biological and clinical perspectives.

To investigate the MSK study in the context of prior reports, we first assessed the univariable outcomes in ccRCC reported in this study using the processed data provided to us by the corresponding author. Given the desire to evaluate clinical outcomes that directly link to potential biological relevance of these findings, we first examined time-to-treatment failure (TTF) in this cohort (as a proxy for progression free survival [PFS]). Information on patient selection or technical quality control from sequencing data was not reported or made available for sensitivity analyses to identify potential biases. Beyond these general issues, the MSK study differed from prior studies in two key ways: 1) They chose different criteria for LOF mutations (excluding splice site mutations as LOF variants for this tumor suppressor, even though this research group and others generally considered this variant class as LOF in prior genomic or immunohistochemistry studies (7-12)); and 2) They split the control arm into two groups (PBRM1 non-LOF that included splice site mutated patients, and PBRM1 wild-type). As a result, power was reduced to detect a significant association between PBRM1 LOF and ICB response by including putative LOF variants in a separated non-LOF control group and by lowering the sample size of each group.

When we re-evaluated this dataset using the previously established definition of PBRM1 LOF mutations (1-3) to directly assess the LOF hypothesis, we observed a statistically significant association between PBRM1 LOF and improved TTF with ICB (p = 0.047; HR = 0.7, 95% CI = 0.49 - 0.99; Fig. 1A; Methods). We noted that there were minimal differences between the statistically significant HR in our standardized analysis and the original study’s non-statistically significant HR estimate (HR = 0.7 vs 0.73), and a small difference in p-values (p = 0.047 vs p = 0.11 in our analysis and the original MSK study, respectively). Since the methodological choices in the MSK study decreased statistical power to detect the same effect size for this hypothesis, we then simulated the minimum cohort sizes necessary to observe a significant association between PBRM1 LOF mutations and TTF, given the study population, effect size and the statistical results reported in the originating study (Methods). This simulation indicated that, with the cohort size, composition, and originating analytical setup, there was 39% power to detect a statistically significant difference (Fig. 1B), and the subgroup analyses split by line of therapy were similarly underpowered (Fig. 2A-B; Methods). This observation highlights the challenges of binary interpretation of p-values, specifically differentiating between statistically significant results (which are highly dependent on sample sizes) and either biologically or clinically significant results (13). After adjusting the PBRM1 classifications to match multiple prior studies (1, 3), their univariable results from a single institution, retrospective ccRCC cohort therefore provides further support for
the significant but modest effect size association between _PBRM1_ LOF mutations and ICB response.

Regarding the overall survival (OS) univariable analyses, the reported OS and time-to-treatment failure (TTF) findings for _PBRM1_ in ccRCC (n = 189 patients) yielded opposing hazard ratios, and the median OS for this advanced ccRCC cohort was 68.2 months. This median OS for the patients in this study, which was a mix of first-line and >= 2 line metastatic ccRCC patients, far exceeded the median OS of any known metastatic ccRCC clinical trial cohort in either first-line or >= second line treatment settings (e.g. one modern clinical trial of first-line treatment-naïve ccRCC patients treated with ICBs reported a median OS of 47 months (14), and a second trial in >= 2 line setting reported median OS of 25 months (15)). In the MSK study, OS was defined from ccRCC diagnosis at any stage, rather than the standard definition of time from start of ICB therapy for metastatic ccRCC (A. Hakimi, personal communication, September 10, 2020), rendering the reported OS challenging to interpret. This issue illustrates a general issue in retrospective single institution molecular analyses, whereby unaccountable biases (e.g. whether drug class impacted censoring, timing of biopsy used for sequencing relative to treatment exposures, among many others) may influence interpretation of results in unknowable ways that can be circumvented through analysis of homogeneously collected tumor samples from well-defined cohorts within specific clinical settings, e.g. clinical trial cohorts (preferably randomized clinical trials).

While the original MSK study identified no association with _PBRM1_ LOF, they reported a significant univariable association between continuous TMB and ICB response in ccRCC. This observation would have major clinical implications for ccRCC if validated, but contrasts findings from prior studies in ccRCC (1, 2, 4, 5, 8, 16). However, this TMB analysis had two key issues. First, their reported significant effect of TMB as a continuous predictor is driven primarily by irregularities in the TMB data and the outcomes of a handful of outliers, rather than a linear relationship to TTF. Specifically, ccRCC patients in this study had a higher median TMB than previously reported in metastatic ccRCC (3.9 mut/Mb) (2, 4), 7 patients (3.7%) had TMB = 0 mutations/Mb (noting unknown technical quality control, and expecting ccRCC to have > 0 mutations with expanded sequencing) and 3 patients (1.6%) had TMB > 10 mutations/Mb (Fig. 3A). TMB inflation is expected since the authors are reporting TMB derived from a targeted driver gene panel; however, since there were no harmonization efforts included in the study, it is impossible to distinguish between the effect of TMB versus the number of driver events (one of which is _PBRM1_). To demonstrate this, we compared univariable models including TMB or log-transformed TMB (Fig. 3B). Log transformation is routinely applied in such TMB analyses to bring outliers closer to the rest of the data distribution, reducing the effect of outliers on the estimated model coefficients (16-18). After log-transforming TMB, there was no significant association with TTF (log-transformed TMB: Cox PH, \( p = 0.16 \), HR = 0.85, 95% CI = 0.68 - 1.07; Methods). Further, because the observed effect of TMB may actually be driven by having a higher number of _PBRM1_ mutations, we re-assessed the relationship between TMB and TTF in samples without _PBRM1_ LOF mutations. When categorizing TMB by tertile and completely removing the effect of _PBRM1_ LOF mutations on TTF, there was no difference in the TTF between samples in the highest and lowest TMB tertiles (\( p = 0.52 \), log-rank test; Methods; Fig. 3C).

The TMB analysis issue was further compounded by including it as a covariate in the subsequent multivariable models presented (Supplementary Fig. 1). Specifically, there was no prior evidence that TMB is significantly associated with ICB response in ccRCC; however, TMB is colinear with driver genes in a small gene panel enriched for cancer genes (e.g. _PBRM1_). Multicollinearity reduces precision of estimated coefficients, which weakens the statistical power of a regression model, and thus one is unable to fully trust the p-values to identify independent variables that are statistically significant. The MSK study then introduce other confounders into the model by
combining single agent and dual anti-PD-1/CTLA4 immunotherapies, even though the clinical impact of these drugs in ccRCC may be distinct (19). After removing TMB from their multivariable model, PBRM1 LOF remains non-statistically significant, although with consistent trends (Cox PH, p = 0.15, HR = 0.75, 95% CI = 0.51 - 1.10). Therefore, statistical issues in TMB analysis and cohort heterogeneity confounded the interpretation of the TTF multivariable model presented in this study. After correction, the multivariable analysis remained underpowered, and while not statistically significant for PBRM1 LOF, it yielded an association with improved TTF that was consistent with the univariable analysis.

Broadly, we conclude that ICB-associated molecular analyses should explicitly distinguish between two goals: 1) Discovering associations warranting further translational and biological investigation, and 2) Developing predictive biomarkers for immediate clinical use in specific clinical settings. This study provides further evidence in support of our prior findings specifically regarding PBRM1 LOF associating with single-agent ICB response, which has biological plausibility given preclinical studies of PBRM1 loss and tumor-immune interactions (1, 22-24) that has prompted further translational investigation. We submit that this study, given the cohort size, p-values close to the threshold used for statistical significance in the setting of modest effect size, technical and statistical anomalies, varying patient populations, and different outcome measurements, does not warrant dismissing a potential biological role of PBRM1 LOF in ccRCC tumor immunology contexts.

Distinct from direct evaluation of the association of PBRM1 LOF with benefit from ICI monotherapy, further functional and immunological characterization of other mutations and mutation classes (i.e. missense) in chromatin regulator genes like PBRM1 using saturation mutagenesis assays may provide more granularity regarding their biological roles. Multiple studies have demonstrated distinct and often subtle functional effects for different missense mutations in different PBRM1 bromodomains and other mSWI/SNF members, which is further supported by transcriptional differences between LOF, non-LOF (missense or in-frame), and wild-type classes and the absence of specific bromodomain mutation clusters (rather a general increased mutation rate in PBRM1 exons) (Supplementary Fig. 2) (25-29). As with other tumor suppressors (30, 31), saturation assays could further inform diverse biological effects for PBRM1 missense mutations in different exonic bromodomains, and potentially their differential impact on tumor-immune interactions, that may not be discriminated by in silico or immunohistochemistry techniques (which does not always correlate with genetic loss-of-function) (12, 32).

Most importantly, as we emphasized in prior immunogenomic correlative studies (1, 16, 33, 34), we again strongly caution against overinterpretation of this biological association (among many others) as simultaneously establishing predictive biomarker status in diverse ccRCC clinical contexts or other diseases. For instance, PBRM1 has also been associated with many other tumor-intrinsic processes (e.g. hypoxia, TP53 interactions) (25, 35), as well as to VEGF TKI monotherapy response in some clinical settings but not others (4, 5, 36). Since VEGF TKIs and ICB combinations are now a standard of care, and these therapies have numerous complex effects on the already intricate tumor and immune microenvironment interactions, it is unlikely that PBRM1 LOF (or any single molecular feature, compared to composite transcriptional programs that capture the biological complexity of this context) has strong clinical biomarker relevance in that combination therapy context. Dedicated translational analyses from randomized clinical trials representing specific clinical contexts are necessary to elucidate these factors.

Clinical and molecular data linked to functional studies can invoke new biology but often do not translate to predictive biomarkers. These biological observations may only capture one component of the complex tumor-immune-stromal processes that are perturbed by ICBs and are
very unlikely to be driven exclusively by a single molecular event. For the ICB predictive biomarker field, TMB best illustrates this challenge, as it has a biological rationale linked to neoantigens and clinico-genomic associations in specific settings, but has no relevance as a predictive biomarker in ccRCC and may have limited performance as a singleton predictive biomarker in diseases where TMB is established as a genomic correlate of response after considering other molecular and clinical variables (37).

METHODS
Survival Analysis
To determine if there are significant differences between the survival curves of two or more groups we used the log-rank test from the survival R package. To determine the hazard ratios and significance of covariates on survival outcomes we performed Cox proportional hazard analysis using the survival R package. To generate the TMB tertile groups for survival analysis, we determined the TMB tertiles from the entire cohort, followed by filtering to only non-LOF and WT PBRM1 patients.

Power analysis (simulations)
To perform simulations for power analysis two assumptions were made: (1) The proportion of PBRM1 mutation statuses (e.g. LOF, non-LOF, and WT) in the authors cohort (n = 189) approximates the relative proportion of these events in ccRCC patient populations. Specifically, the frequency of PBRM1 LOF mutations in ccRCC patient populations is 32%. (2) The hazard ratios from performing Cox proportional hazard analysis on the originating cohort are the postulated hazard ratios for PBRM1 LOF mutations in this cancer type and therapeutic settings (e.g. overall, first-line, second-line). To determine the power to detect a significant hazards ratio for PBRM1 LOF mutations we used the lifelines python package.

Log transformation of TMB
A total of 7 samples had a TMB of 0. To account for these samples during log transformation, we added a constant 0.5 to the TMB values. For reference, the lowest non-zero TMB observed in the cohort was 0.9 mutations per megabase.

Code availability
All data (as provided by the corresponding author of the initiating study) and code written to perform the analyses presented herein is available at: https://github.com/vanallenlab/PBRM1_MSK_RCC_analysis

REFERENCES


Figure 1: Effect of PBRM1 classification criteria from previously published studies and univariable Cox proportional hazards power simulations using the originating study criteria

a) When classifying PBRM1 alterations as LOF using the same criteria as previous studies (i.e. including splice-site mutations as LOF), and keeping the non-LOF patients as a single control arm like previous studies, there is a significant improvement in TTF for PBRM1 LOF patients (Cox PH, p = 0.047, HR = 0.7 [0.49 - 0.99]).

b) Simulations showing the power to determine that PBRM1 LOF mutations have a significantly different hazard by the number of patients included in the cohort. The proportion of PBRM1 LOF patients used in the power calculations was 32%, the same as the originating study. The power calculations were performed under the assumptions that the true hazard ratio was the lower bound of the 95% confidence interval (green), the hazards ratio estimate (gold), and the upper bound of the 95% confidence interval (brown) from the univariable Cox proportional hazards analysis in the originating study.
Figure 2: Univariable Cox proportional hazards power simulations by line of therapy using the originating study criteria
Simulations showing the power to determine that PBRM1 LOF mutations have a significantly different hazard in the a) front-line setting and b) second-line setting by the number of patients included in the cohort. The proportion of PBRM1 LOF patients used in the power calculations was 32%, the same as the overall PBRM1 LOF frequency originating study. The power calculations were performed under the assumptions that the true hazard ratio was the lower bound of the 95% confidence interval (green), the hazards ratio estimate (gold), and the upper bound of the 95% confidence interval (brown) from univariable Cox proportional hazards analysis on patients from a) front-line and b) second-line settings, respectively.
Figure 3: The TMB association in originating study is driven by outliers and TMB is not associated with ICB response in ccRCC

a) The distribution of TMB in the originating study cohort (n = 189). A total of 7 samples (3.7%) have a TMB of 0 and 3 samples (1.6%) have a TMB > 10 mutations/Mb (occurs at a rate of < 1% in ccRCC).

b) The distribution of log-transformed TMB from the originating study cohort (n = 189). Log-transformation reduces the effect of outliers on the data, and after log-transformation TMB is no longer associated with ICB response (Cox PH, p = 0.16, HR = 0.85, 95% CI = 0.68 - 1.07).

c) TTF survival curves between the lowest and highest TMB tertiles filtered down to non-LOF and WT PBRM1 samples.
Supplementary Figure 1: TMB is nonlinear in the originating study data
The martingale residuals for both a) TMB and b) log-normalized TMB are nonlinear, which violates the Cox proportional hazards model assumption for continuous variables, and suggests that the TMB covariate is not properly fit to the data.
Supplementary Figure 2: PBRM1 expression differences by PBRM1 mutation status

There is no difference in PBRM1 expression between tumors that are PBRM1 WT or have non-LOF mutations (e.g. missense and in-frame indels) in PBRM1. Conversely, tumors with putative LOF mutations (including splice-site) in PBRM1 have significantly lower PBRM1 expression compared to PBRM1 WT tumors ($p = 2.5 \times 10^{-8}$) and tumors with non-LOF mutations ($p = 0.0022$) in PBRM1.