# PHYLOGENETIC TREE CONSTRUCTION AND "TRUNCAL LOSS" ANALYSIS REVEAL HIDDEN ASSOCIATIONS BETWEEN LOSS OF PROTEIN EXPRESSION IN SWI/SNF COMPLEX COMPONENTS AND TUMOR STAGE AND SURVIVAL IN CLEAR CELL RENAL CELL CARCINOMA (CCRCC)

Wei Jiang[1], Essel Dulaimi[2], Theodore Parsons[1], Karthik Devarajan[2], Qiong Wang[1], Raymond O'Neill[1], Charalambos C. Solomides[1], Stephen C. Peiper[1], Robert Uzzo[2], Joseph R. Testa[2], and Haifeng Yang[1]

[1]Thomas Jefferson University, Philadelphia, PA, United States; [2]Fox Chase Cancer Center/Temple University, Philadelphia, PA, United States

## BACKGROUND

Polybromo-1 (PBRM1), a targeting subunit of the SWI/SNF chromatin remodeling complex, is mutated at a rate of ~40% in clear cell Renal Cell Carcinoma (ccRCC), second only to VHL. Whether its mutation is correlated with tumor stage is controversial. As other components of the SWI/SNF complex were also reported to be mutated in ccRCC, we aim to examine the protein expression patterns of PBRM1, ARID1A, BRG1, and BRM in ccRCC, and to investigate possible association between their loss of expression and tumor stage, as well as survival. We also included a histone modifier, SETD2, which is recently discovered to be mutated in ~15% of ccRCC.

## DESIGN

160 ccRCC, with 40 per tumor stage (1-4), diagnosed at Fox Chase Cancer Center, were used to generate tissue microarray (TMA). Four 1x1 mm$^2$ foci (dots) from different regions of each tumor were selected to represent tumor heterogeneity. Standard immunohistochemistry (IHC) was performed on the TMA slides at Thomas Jefferson University Hospital, and was scored by two pathologists (W. J. and T. P.), with the clinical and staging details blinded. Loss of expression was defined as 0-5% of staining within tumor nuclei in any 1 mm$^2$ focus. Discrepancies in scores were resolved by re-review by the two pathologists, and consensus was reached in such cases. Clinical data were also collected and overall survival (OS) and recurrence-free survival (RFS) were calculated. Appropriate statistical analyses were performed (see details in RESULTS).

## RESULTS

49/160 (31%), 81/160 (51%), 23/160 (14%), 24/160 (15%), and 61/160 (38%) tumors show loss of PBRM1, ARID1A, SETD2, BRG1, and BRM expressions, respectively. For individual protein, very high percentage of tumors show loss of expression in only a fraction of the four foci, displaying heterogeneity. Striking co-loss patterns among different proteins are also observed.

"Truncal Loss" for individual protein was defined as the most ubiquitous loss of expression in the foci from the same tumor, and "The Only Truncal Loss" if there was no co-loss with other proteins. "Truncal Loss", like truncal genetic mutation, is most likely an early event in tumorigenesis, therefore, for each individual tumor, phylogenetic tree of protein expression can be constructed (see examples in Figure 1).

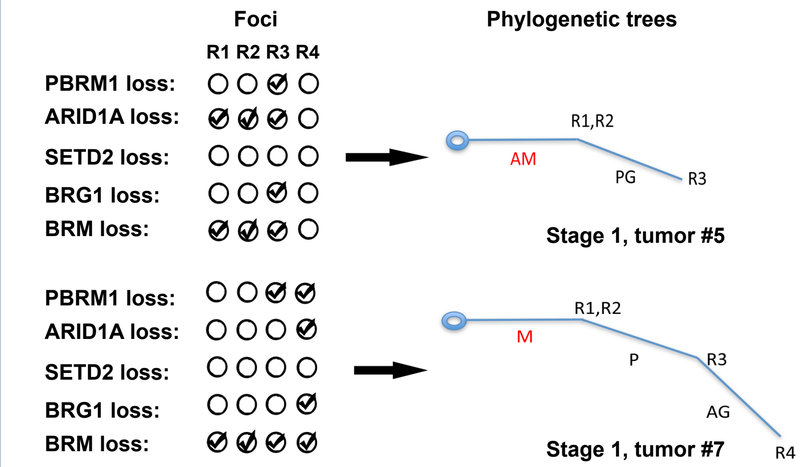


**Figure 1.** Example of construction of phylogenetic tree. On the left are the schematic representations of loss of expression of each protein on the foci of stage 1 tumors 5 and 7. On the right are the phylogenetic trees that were built upon the assumption that the more prevalent changes occurred earlier during tumor evolution. R-Region, A-ARID1A, M-BRM, G-BRG1, P-PBRM1.

Statistic analysis reveals that certain protein expression pattern is strongly associated with tumor stage (see Table 1).

**Table 1.** *p* values showing association between protein expression pattern and tumor stage (Fisher's exact test).

| | The Only Truncal Loss | Truncal Loss | Any Loss in Tumor |
|---|---|---|---|
| PBRM1 | 0.0002 | <0.0001 | 0.003 |
| BRM | 0.001 | 0.012 | 0.05 |
| ARID1A | 0.028 | 0.2 | 0.21 |
| BRG1 | 1 | 1 | 0.12 |
| SETD2 | 0.06 | 0.032 | 0.07 |

Univariate analyses for overall survival (OS) and recurrence-free survival (RFS) were performed using the Cox proportional hazards (PH) model. The following variables showed significant association with RFS: SETD2.Any loss ($p$=0.018, RR=0.50), BRM.Any loss ($p$=0.015, RR=1.84); BRG1.Truncal loss ($p$=0.031, RR=1.85), and BRM.Truncal loss ($p$=0.00028, RR=1.79). For OS, SETD2. Any loss ($p$=0.013, RR=0.50), PBRM1.Truncal loss ($p$=0.004, RR=0.6); BRG1.Truncal loss ($p$=0.017, RR=1.97), and BRM.Truncal loss ($p$=0.002, RR=1.60). If RR > 1 then it suggests that the risk of death increases with the value of the corresponding variable (as it goes from 0 to 1 or from 0 to 2). A similar interpretation holds when RR < 1. Variables with $p$ <= 0.1 were used in multivariable analyses using the Cox PH model (Table 2 and 3).

**Table 2.** Multivariate analysis of protein expression pattern with recurrence-free survival (RFS).

| | RR | *p* | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|
| SETD2.Any loss | 0.272 | 0.034 | 0.082 | 0.904 |
| BRG1.Any loss | 1.627 | 0.38 | 0.549 | 4.822 |
| BRM.Any loss | 1.571 | 0.244 | 0.735 | 3.358 |
| SETD2.Truncal loss | 1.133 | 0.695 | 0.607 | 2.112 |
| BRG1.Truncal loss | 1.351 | 0.429 | 0.641 | 2.851 |
| BRM.Truncal loss | 1.324 | 0.218 | 0.847 | 2.069 |

**Table 3.** Multivariate analysis of individual protein expression pattern with overall survival (OS).

| | RR | *p* | Lower 95% CI | Upper 95% CI |
|---|---|---|---|---|
| ARID1A.Any loss | 0.226 | 0.001 | 0.093 | 0.549 |
| SETD2.Any loss | 0.299 | 0.0003 | 0.154 | 0.578 |
| BRM.Any loss | 3.588 | 0.0016 | 1.621 | 7.94 |
| PBRM1.Truncal loss | 0.425 | 2.45E-05 | 0.286 | 0.632 |
| BRG1.Truncal loss | 2.553 | 0.0027 | 1.384 | 4.709 |
| BRM.Truncal loss | 0.987 | 0.948 | 0.657 | 1.481 |

## CONCLUSIONS

1. Using phylogenetic tree construction and "Truncal Loss" analysis, we identified statistically significant associations between loss of protein expression and tumor stage in SWI/SNF complex components in ccRCC, while the commonly used "Any Loss in Tumor" analysis either failed to do so or did so weakly.

2. We also made the novel findings that "Truncal Loss" of BRG1 or "Any Loss in Tumor" of BRM is significantly associated with better prognosis for OS in multivariate analysis. PBRM1's association with clinical outcome has been controversial, and our result is consistent with the view that PBRM1 loss is a bad prognosis factor. Consistent with previous reports, "Any Loss in Tumor" of ARID1A or SETD2 is significantly associated with worse prognosis for OS.

## ACKNOWLEDGEMENT

This work was partially supported by an NIH grant awarded to HFY.