

Endogenous retrovirus expression is associated with response to immune checkpoint blockade in clear cell renal cell carcinoma

Anshuman Panda,^{1,2} Aguirre A. de Cubas,³ Mark Stein,^{1,4} Gregory Riedlinger,¹ Joshua Kra,¹ Tina Mayer,¹ Christof C. Smith,⁵ Benjamin G. Vincent,⁵ Jonathan S. Serody,⁵ Kathryn E. Beckermann,^{3,6} Shridar Ganesan,^{1,4} Gyan Bhanot,^{1,2,7} and W. Kimryn Rathmell^{3,6}

¹Rutgers Cancer Institute of New Jersey, New Brunswick, New Jersey, USA. ²Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey, USA. ³Vanderbilt-Ingram Cancer Center, Nashville, Tennessee, USA. ⁴Rutgers Robert Wood Johnson Medical School, New Brunswick, New Jersey, USA. ⁵Department of Medicine, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina, USA. ⁶Department of Medicine, Division of Hematology and Oncology, Vanderbilt University Medical Center, Nashville, Tennessee, USA. ⁷Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, New Jersey, USA.

Although a subset of clear cell renal cell carcinoma (ccRCC) patients respond to immune checkpoint blockade (ICB), predictors of response remain uncertain. We investigated whether abnormal expression of endogenous retroviruses (ERVs) in tumors is associated with local immune checkpoint activation (ICA) and response to ICB. Twenty potentially immunogenic ERVs (π ERVs) were identified in ccRCC in The Cancer Genome Atlas data set, and tumors were stratified into 3 groups based on their expression levels. π ERV-high ccRCC tumors showed increased immune infiltration, checkpoint pathway upregulation, and higher CD8⁺ T cell fraction in infiltrating leukocytes compared with π ERV-low ccRCC tumors. Similar results were observed in ER⁺/HER2⁻ breast, colon, and head and neck squamous cell cancers. ERV expression correlated with expression of genes associated with histone methylation and chromatin regulation, and π ERV-high ccRCC was enriched in *BAP1* mutant tumors. *ERV3-2* expression correlated with ICA in 11 solid cancers, including the 4 named above. In a small retrospective cohort of 24 metastatic ccRCC patients treated with single-agent PD-1/PD-L1 blockade, *ERV3-2* expression in tumors was significantly higher in responders compared with nonresponders. Thus, abnormal expression of π ERVs is associated with ICA in several solid cancers, including ccRCC, and *ERV3-2* expression is associated with response to ICB in ccRCC.

Introduction

Immune checkpoint blockade (ICB) leads to durable objective responses in several cancer types (1). A high mutation burden, from exposure to exogenous carcinogens (2, 3) or intrinsic defects in DNA repair and replication (4, 5), predicts response to ICB in some cancer types (6), presumably because of somatic neoantigens. Further, expression of certain exogenous viruses in tumors, such as Epstein-Barr virus in gastric cancer (7) and NK/T cell lymphoma (8) and Merkel cell polyomavirus in Merkel cell cancer (9, 10), is also associated with response to ICB, presumably because of viral antigens. However, some cancers, such as clear cell renal cell carcinoma (ccRCC), have clinically significant and durable responses to ICB (11), despite low mutation burden and absence of known exogenous viral infection.

A recent study (12) reported that loss of the chromatin-modifying gene *PBRM1* correlates with response to ICB in pretreated ccRCC patients. Intriguingly, in multiple cohorts of ccRCC patients, tumors with *PBRM1* loss have lower levels of the CD8⁺ T cell marker (*CD8A*), IFN- γ (*IFNG*), and immune checkpoint genes compared with tumors with intact *PBRM1* (12), although the mechanism underlying these correlations is unknown. Recent studies also show, that ccRCCs, while having low overall mutation burden, are enriched in frameshift mutations, which may be more immunogenic (13). However, the relationship

Authorship note: AP and AAC contributed equally to this work.

Conflict of interest: The authors have declared that no conflict of interest exists.

Submitted: April 6, 2018

Accepted: July 5, 2018

Published: August 23, 2018

Reference information:

JCI Insight. 2018;3(16):e121522.

<https://doi.org/10.1172/jci.insight.121522>

insight.121522.

between levels of frameshift mutations and response to ICB remains unclear. Emerging data also suggest a role for the metabolic environment in balancing (14) or suppressing (15) antitumor immunity in ccRCC. In summary, the mechanisms of response to ICB in ccRCC are currently unknown.

A substantial fraction of the human genome contains endogenous retroviruses (ERVs) (16), the expression of which is normally silenced in most somatic tissues. However, 66 ERVs are known to be transcribed in humans (17), and their expression has been reported in multiple cancers (18). Some recently integrated (19) and well-preserved (19) ERVs, such as those in the ERVK family, are known to retain a functional gag gene (20) and an open-reading frame in *pol* and *env* genes (20), and their simian equivalents are known to induce immune response in Indian rhesus macaques (21). Thus, abnormal expression of some potentially immunogenic ERVs (π ERVs) in tumors may elicit an antitumor immune response spearheaded by CD8⁺ T cells. Tumors may progress by blocking this immune response through upregulation of immune checkpoint pathways, making them sensitive to ICB.

In this study, using RNA-sequencing (RNA-seq) data and previously published (18) ERV expression data of (mostly primary) tumors ($n = 472$ for ccRCC, $n = 4,438$ for 20 other cancers) from The Cancer Genome Atlas (TCGA), and a cohort of metastatic ccRCC patients treated with single-agent PD-1/PD-L1 blockade ($n = 24$) at two institutions, we evaluated the possibility that expression of π ERVs in tumors induces local immune checkpoint activation (ICA) in a subset of tumors and associates with responsiveness to ICB.

Results

π ERVs are abundant in 4 solid cancers from TCGA. To identify π ERVs (see the Methods for details), we evaluated 21 solid cancers from TCGA for correlation between expression levels (18) of 66 transcribed ERVs (17) and RNA-seq-based evidence of local ICA. As shown in Figure 1A, ICA criteria included markers of immune activation, namely overall immune infiltration (“ImmuneScore” from ESTIMATE, ref. 22) and expression of the cytotoxic T cell marker *CD8A*, and markers of checkpoint pathway upregulation, namely expression of genes in the PD-1, CTLA-4, and BTLA/HVEM pathways.

Results of this analysis (summarized in Figure 1B) showed that π ERVs are abundant in 4 cancers, namely ccRCC (from TCGA KIRC database, ER⁺/HER2⁻ breast cancer (ER⁺ HER2⁻), colon cancer (COAD), and head and neck squamous cell cancer (HNSC). In these cancers, expression of 9–20 π ERVs correlated with ICA, while expression of only 0 to 2 π ERVs correlated with ICA in other cancers (Figure 1B). π ERVs were most abundant in ccRCC (KIRC), where expression of 20 π ERVs correlated with ICA, most of which (18 of 20) were members of the ERVK family (Figure 1C). While most ERVs were identified as π ERVs in 0 or 1 cancer type, and several were identified across multiple diseases, *ERV3-2* was identified as a π ERV in 11 different solid cancers (Figure 1D), including the 4 cancers named above.

Expression of π ERVs defines subtypes with differential ICA in ccRCC. In TCGA ccRCC (KIRC) cohort, the 20 π ERVs were mostly coexpressed. Hierarchical clustering of tumors by percentile expression of these 20 π ERVs identified 3 distinct subtypes corresponding to high, intermediate, and low expression of π ERVs (Figure 2A). Since loss-of-function mutations in chromatin regulatory genes (including *PBRM1*, *SETD2*, and *BAP1*) are frequently observed in ccRCC (23), and a recent study (12) reported an association between *PBRM1* loss and response to non-first-line ICB in pretreated ccRCC, we compared the frequency of mutation of these genes in the 3 π ERV expression-based subtypes in KIRC. As shown in Figure 2B, although there was no significant enrichment of *VHL*, *PBRM1*, or *SETD2* mutations in the ccRCC tumors with a high expression of π ERVs (π ERV-high ccRCC tumors), there was a statistically significant enrichment of *BAP1* mutations in the π ERV-high tumors compared with ccRCC tumors with a low expression of π ERVs (π ERV-low tumors) (15.2% vs. 6.8%, odds ratio 2.44 [95% CI: 1.15–5.16], $P = 0.028$).

π ERV-high tumors also had significantly higher immune infiltration (Figure 2C), a significantly higher fraction of CD8⁺ T cells among tumor-infiltrating leukocytes (Figure 2C), and significantly higher mRNA expression of the cytotoxic T cell marker *CD8A* (Figure 2D) compared with π ERV-low tumors, indicating immune activation in the π ERV-high tumors. In addition, follicular helper T cells, $\gamma\delta$ T cells, activated NK cells, resting dendritic cells, and plasma cells constituted a significantly higher fraction of tumor-infiltrating leukocytes in the π ERV-high tumors compared with that in π ERV-low tumors (Figure 2C). Furthermore, M1 macrophages were more abundant in π ERV-high tumors, whereas M2 macrophages were more abundant in π ERV-low tumors (Figure 2C).

π ERV-high tumors also had significantly higher mRNA expression of several checkpoint genes (*PD-1*, *PD-L1*, *CTLA-4*, *CD80*, *BTLA*, *HVEM*, *LAG3*) compared with π ERV-low tumors (Figure 2D),

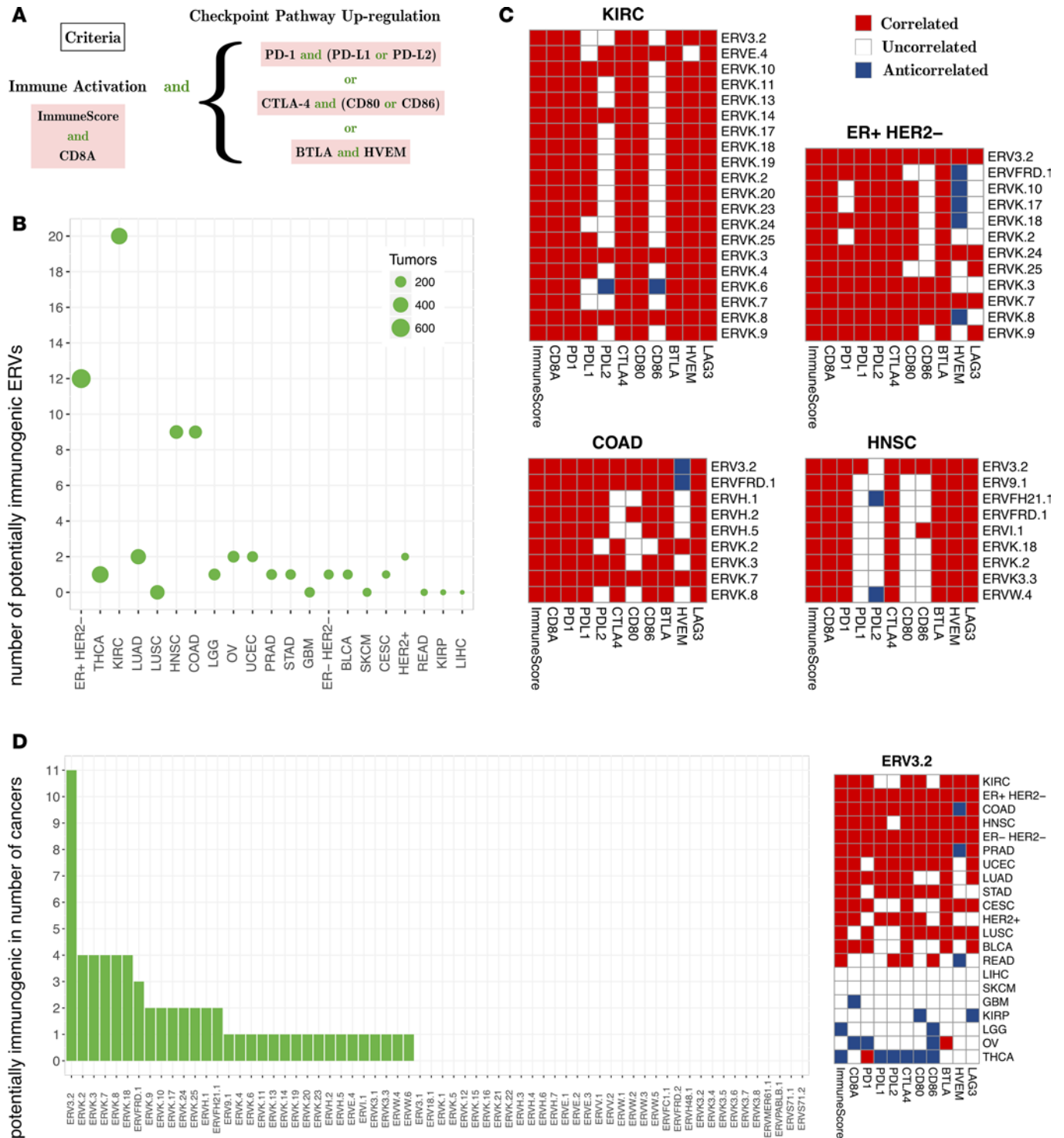


Figure 1. Potentially immunogenic ERVs are abundant in 4 solid cancers from TCGA. (A) Immune checkpoint activation criteria used to identify potentially immunogenic ERVs (π ERVs) in each solid cancer type. (B) The number of π ERVs in each solid cancer type identified 4 cancer types with an unusually high number of π ERVs. (C) Correlation (Spearman) between expression of π ERVs (rows) and levels of immunological variables (columns) in the 4 cancer types. (D) ERV3-2 was identified as a π ERV in 11 different cancer types. Cancer type acronyms are standard TCGA abbreviations (<https://tcga-data.nci.nih.gov/>).

indicating checkpoint pathway upregulation in π ERV-high tumors. Interestingly, expression levels of most π ERVs correlated with expression levels of *PD-L1* but not *PD-L2* and *CD80* but not *CD86* (Figure 1C). Consistently, in contrast to *PD-L1* and *CD80*, *PD-L2* and *CD86* were not differentially expressed in π ERV-high versus π ERV-low tumors (Figure 2D).

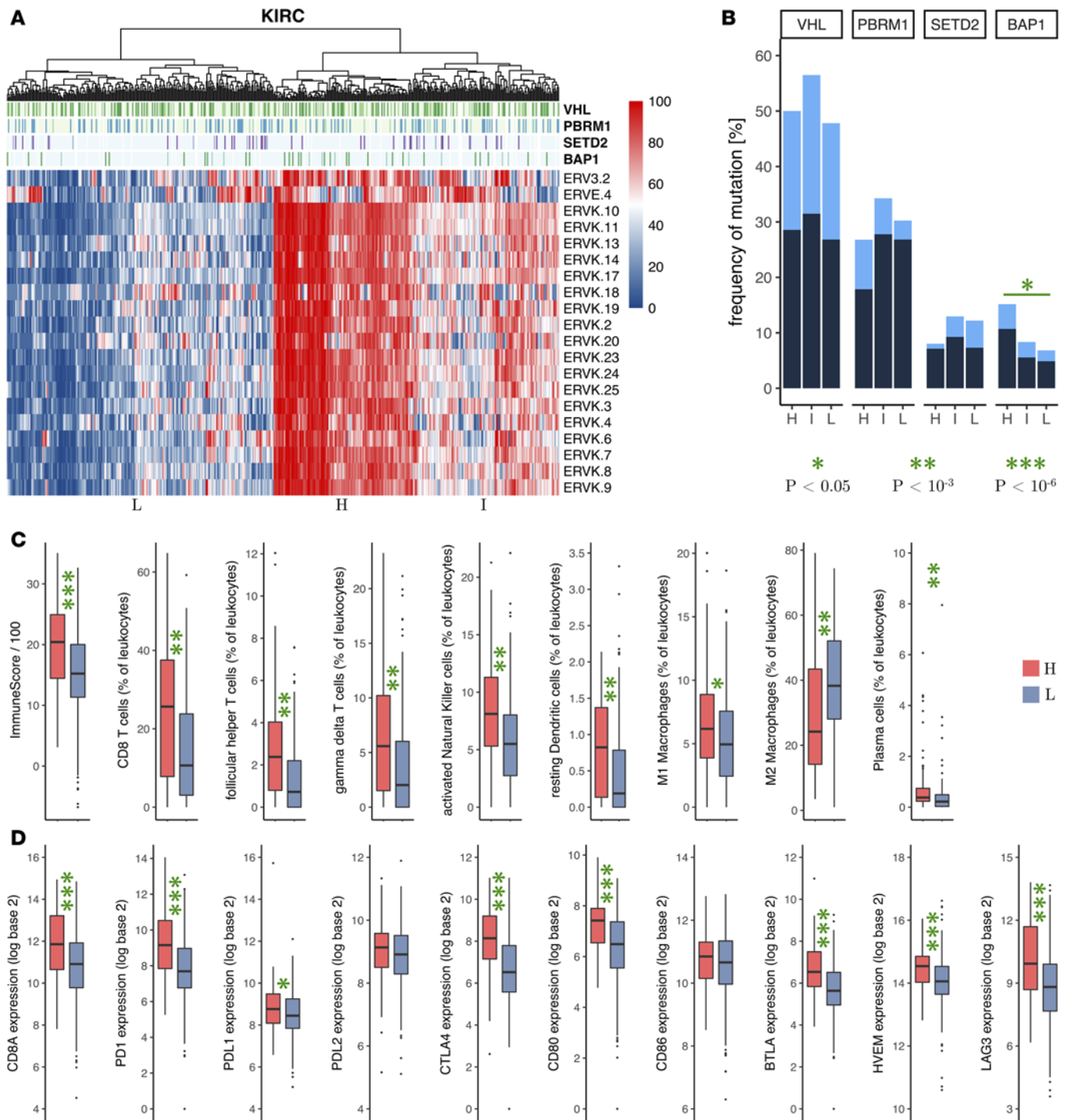


Figure 2. Expression of π ERVs defines subtypes with differential immune checkpoint activation in ccRCC (KIRC). (A) Hierarchical clustering of tumors from TCGA (columns) by expression (percentile) of π ERVs (rows) stratifies tumors into 3 subtypes (high [H], intermediate [I], and low [L]). (B) Frequency of *VHL*, *PBRM1*, *SETD2*, and *BAP1* mutations (dark, truncating mutations; light, other nonsynonymous mutations) in the 3 subtypes. Comparison of (C) overall immune infiltration in tumors ("ImmuneScore") and fractional composition of tumor-infiltrating leukocytes and (D) mRNA expression of *CD8A* (cytotoxic T cell marker) and immune checkpoint genes between π ERV-high and π ERV-low subtypes. Number of samples: (C) ImmuneScore (119 H, 228 L), all other categories (90 H, 134 L), (D) 119 H and 228 L. *P* values reported in bar plots and box plots are from Fisher's exact test and Wilcoxon rank-sum test, respectively (all 2 sided). **P* < 0.05, ***P* < 10⁻³, ****P* < 10⁻⁶.

Table 1. Patient characteristics and ERV3-2 levels (for two different primers) in the validation cohort

Sample	Sex	Age at initiation (yr)	Agent	Best response	Progression-free survival	ERV3-2_1	ERV3-2_2
VICC_1	M	61	Nivolumab	Partial response	>29 mo	37.89	121.65
VICC_2	M	54	Atezolizumab	Partial response	28 mo	360.33	287.72
VICC_3	M	48	Nivolumab	Partial response	>24 mo	77.88	95.92
CINJ_1	M	69	Pembrolizumab	Partial response	>20 mo	281.20	125.63
CINJ_2	M	66	Nivolumab	Partial response	19 mo	48.05	38.81
CINJ_3	M	65	Nivolumab	Partial response	17 mo	3654.11	
CINJ_4	F	64	Nivolumab	Partial response	>15 mo	568.56	168.51
VICC_4	F	73	Nivolumab	Partial response	13 mo	20.12	43.31
VICC_5	M	72	Nivolumab	Partial response	11 mo	765.63	319.27
VICC_6	F	63	Nivolumab	Partial response	9 mo	45.01	57.73
VICC_7	M	63	Nivolumab	Partial response	8 mo	81.90	26.81
VICC_8	M	67	Nivolumab	Partial response	7 mo	235.26	155.56
CINJ_5	F	57	Nivolumab	Partial response	6 mo	93.36	19.73
VICC_9	M	79	Nivolumab	Progressive disease	18 wk	61.98	22.70
CINJ_6	F	63	Nivolumab	Progressive disease	15 wk	65.91	7.10
CINJ_7	M	62	Nivolumab	Progressive disease	13 wk	44.36	21.05
VICC_10	M	46	Nivolumab	Progressive disease	12 wk	45.64	23.36
VICC_11	M	49	Nivolumab	Progressive disease	12 wk	10.66	7.40
CINJ_8	M	61	Nivolumab	Progressive disease	11 wk	90.55	78.14
CINJ_9	F	85	Nivolumab	Progressive disease	11 wk	7.30	26.37
VICC_12	M	54	Nivolumab	Progressive disease	10 wk	32.05	25.05
VICC_13	F	63	Nivolumab	Progressive disease	8 wk	18.64	8.55
VICC_14	M	60	Nivolumab	Progressive disease	8 wk	12.82	28.90
VICC_15	M	82	Nivolumab	Progressive disease	8 wk	3.05	

Expression of ERV3-2 predicts response to ICB in ccRCC. ccRCC (KIRC) showed the strongest evidence of π ERV-associated ICA among the 21 solid cancers (Figure 1B), and *ERV3-2* showed the most consistent correlation with ICA among the 66 transcribed ERVs (17) (Figure 1D). Consequently, we evaluated the expression of *ERV3-2* in ccRCC tumors as a predictor of response to ICB.

The RNA level of *ERV3-2* was measured by real-time quantitative PCRs (RT-qPCRs) in tumors of 24 metastatic ccRCC patients treated with single-agent PD-1/PD-L1 antibody at the Vanderbilt-Ingram Cancer Center and the Rutgers Cancer Institute of New Jersey (Table 1). The cohort consisted of 13 patients with partial response, with progression-free survival of 6 months or longer, and 11 nonresponders who had immediate progressive disease. The RNA level of *ERV3-2* was quantified using 2 different primers (see the Methods for details), referred to here as ERV3-2_1 and ERV3-2_2 (1 responder and 1 nonresponder failed in the case of ERV3-2_2). RNA levels of *ERV3-2*, as measured by either primer, were significantly higher ($P = 0.002$ for ERV3-2_1, $P = 0.0008$ for ERV3-2_2) in tumors from responders compared with tumors from nonresponders and were an excellent predictor of response to ICB (area under the receiver operating characteristic [ROC] curve: 0.86 for ERV3-2_1, 0.90 for ERV3-2_2) in the preliminary analysis, based on this collection of samples (Figure 3A).

As a consistency check, patients were classified into ERV3-2⁺ and ERV3-2⁻ groups, based on whether their tumors had higher or lower expression of *ERV3-2* compared with the optimal cutoff inferred from the ROC curves (marked by green arrows in Figure 3A). As shown in Figure 3B, ERV3-2⁺ patients had significantly higher objective response rates compared with ERV3-2⁻ patients for both ERV3-2_1 (90.0% vs. 28.6%, odds ratio 22.5 [95% CI: 2.1–240.5], $P = 0.004$) and ERV3-2_2 (90.9% vs. 18.2%, odds ratio 45.0 [95% CI: 3.5–584.3], $P = 0.002$). Consistently, ERV3-2⁺ patients had longer progression-free survival (Figure 3B) compared with ERV3-2⁻ patients for both ERV3-2_1 (hazard ratio 0.53 [95% CI: 0.27–1.02], $P = 0.05$) and ERV3-2_2 (hazard ratio 0.15 [95% CI: 0.05–0.44], $P = 0.00003$).

These results suggest that ccRCC tumors with sufficiently high expression of *ERV3-2* may be enriched in the pool of tumors sensitive to ICB. This is particularly interesting because patients with π ERV-high and π ERV-intermediate tumors have significantly shorter overall survival (hazard ratio 1.44 [95% CI: 1.06–1.97],

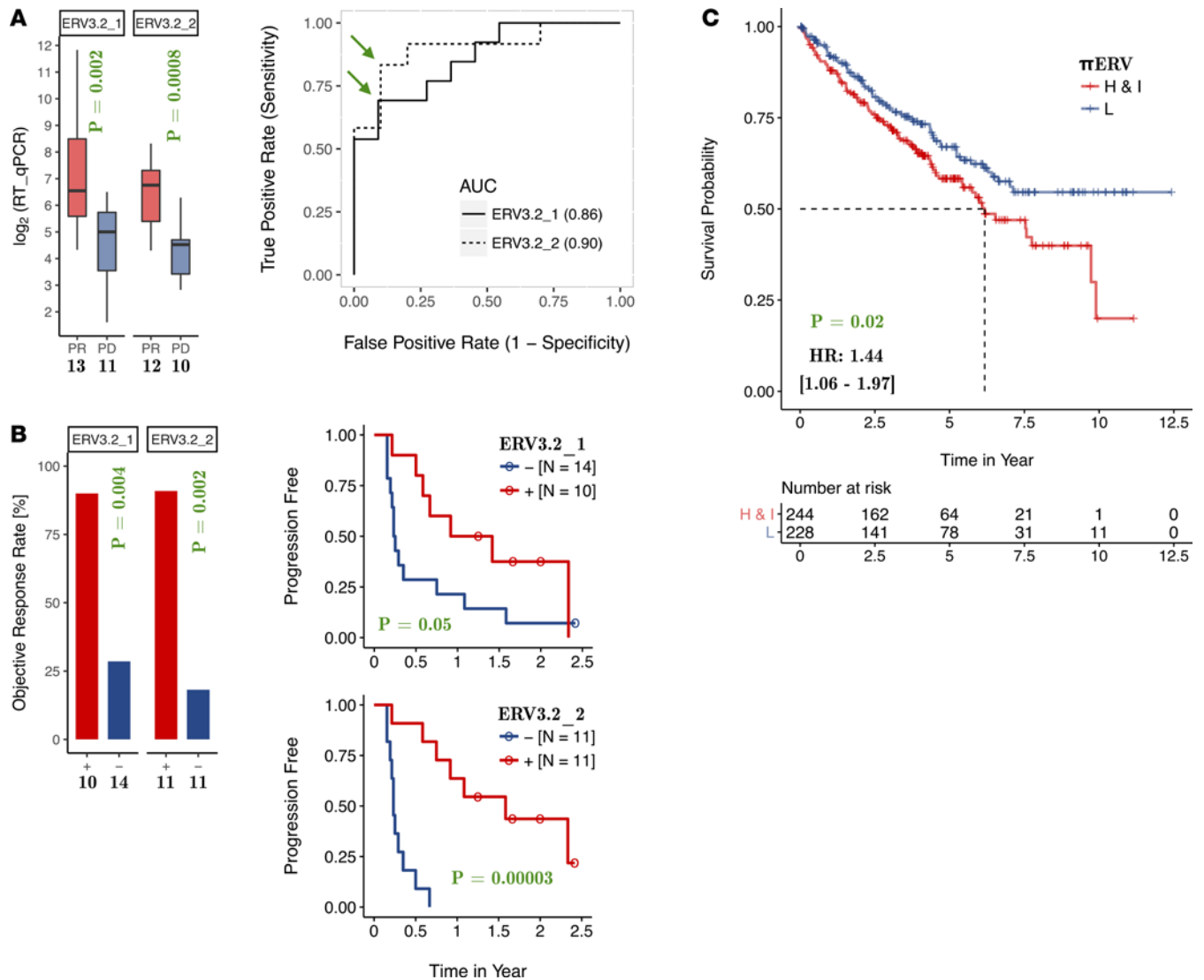


Figure 3. RNA expression of ERV3-2 predicts the response to immune checkpoint blockade in ccRCC. (A) Expression of ERV3-2 is significantly higher in tumors from responders compared with tumors from nonresponders and is an excellent predictor of response to immune checkpoint blockade for both primers. Green arrows mark the optimal cutoffs that were subsequently used to stratify patients into ERV3-2⁺ or ERV3-2⁻ groups for consistency checks. (B) The ERV3-2⁻ group has significantly higher objective response rates and longer progression-free survival compared with the ERV3-2⁺ group for both primers. (C) In contrast, π ERV-high/intermediate subtypes have shorter overall survival under standard therapy compared with π ERV-low subtype in TCGA ccRCC (KIRC) cohort. The number of samples is specified in each panel. P values reported in bar plots, box plots, and Kaplan-Meier plots are from Fisher's exact test, Wilcoxon rank-sum test, and log-rank tests, respectively (all 2 sided).

$P = 0.02$) compared with patients with π ERV-low tumors under standard therapy for ccRCC (Figure 3C). This is consistent with the enrichment of *BAP1* mutations in the π ERV-high subtype of ccRCC (Figure 2B), as *BAP1* mutations are associated with poor prognosis in ccRCC (24).

Expression of π ERVs defines subtypes with differential ICA in ER⁺ HER2⁻, COAD, and HNSC. Like KIRC, π ERV expression-based subtypes were also observed in ER⁺ HER2⁻, COAD, and HNSC (Supplemental Figure 1A, Supplemental Figure 2A, and Supplemental Figure 3A; supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.121522DS1>) in TCGA data. Thus, a clearly identifiable subset of tumors displays broad transcriptional activation of π ERVs in these 4 cancers.

Similar to KIRC, the π ERV-high tumors had significantly higher immune infiltration, a significantly higher fraction of CD8⁺ T cells in infiltrating leukocytes, and significantly higher mRNA expression of the cytotoxic T cell marker *CD8A* compared with π ERV-low tumors in these 3 cancers (Supplemental Figure 1, B–D; Supplemental Figure 2, B–D; and Supplemental Figure 3, B–D), indicating immune activation in the π ERV-high subtype in all 4 cancers. Furthermore, M1 macrophages were more abundant in the π ERV-high

Table 2. Top biological processes from enrichment analysis of genes with expression levels that follow the trends π ERV-high > π ERV-intermediate > π ERV-low > adjacent normal or π ERV-high < π ERV-intermediate < π ERV-low < adjacent normal in ccRCC

GO: biological processes				
No.	ID	Name	FDR	Trend
1	GO:0046649	Lymphocyte activation	1.85E-13	H > I > L > adjn
2	GO:0070489	T cell aggregation	2.36E-13	H > I > L > adjn
3	GO:0042110	T cell activation	2.36E-13	H > I > L > adjn
4	GO:0045321	Leukocyte activation	2.36E-13	H > I > L > adjn
5	GO:0071593	Lymphocyte aggregation	2.36E-13	H > I > L > adjn
6	GO:0070486	Leukocyte aggregation	4.00E-13	H > I > L > adjn
7	GO:0007159	Leukocyte cell-cell adhesion	8.72E-12	H > I > L > adjn
8	GO:0050863	Regulation of T cell activation	3.86E-11	H > I > L > adjn
9	GO:0051249	Regulation of lymphocyte activation	8.17E-11	H > I > L > adjn
10	GO:0001775	Cell activation	9.66E-11	H > I > L > adjn
11	GO:1903037	Regulation of leukocyte cell-cell adhesion	1.43E-10	H > I > L > adjn
12	GO:0002694	Regulation of leukocyte activation	2.02E-10	H > I > L > adjn
1	GO:0045333	Cellular respiration	1.62E-14	H < I < L < adjn
2	GO:0007005	Mitochondrion organization	2.06E-13	H < I < L < adjn
3	GO:1902600	Hydrogen ion transmembrane transport	2.06E-13	H < I < L < adjn
4	GO:0055114	Oxidation-reduction process	2.58E-13	H < I < L < adjn
5	GO:0006119	Oxidative phosphorylation	6.94E-12	H < I < L < adjn
6	GO:0015980	Energy derivation by oxidation of organic compounds	8.09E-12	H < I < L < adjn

H, high; I, intermediate; L, low; GO, gene ontology. Significant values are indicated by > (significantly higher) and < (significantly lower) symbols.

subtype of ER⁺ HER2⁻ and COAD (but not HNSC), whereas M2 or M0 macrophages were more abundant in the π ERV-low subtype of all 3 cancers (Supplemental Figure 1C, Supplemental Figure 2C, and Supplemental Figure 3C). Additionally, activated memory CD4⁺ T cells in COAD and HNSC and follicular helper T cells in HNSC comprised a significantly higher fraction of infiltrating leukocytes and regulatory T cells in COAD comprised a significantly lower fraction of T cells in the π ERV-high tumors compared with the π ERV-low tumors (Supplemental Figure 1C, Supplemental Figure 2C, and Supplemental Figure 3C).

As in KIRC, π ERV-high tumors had significantly higher mRNA expression of checkpoint genes in the PD-1 and CTLA-4 pathways compared with π ERV-low tumors in ER⁺ HER2⁻ and COAD (Supplemental Figure 1D and Supplemental Figure 2D), indicating checkpoint pathway upregulation in the π ERV-high tumors in these cancers. Interestingly, unlike that in ER⁺ HER2⁻ and COAD, expression of most π ERVs in HNSC did not correlate with the expression of the ligands of *PD-1* and *CTLA-4* (Figure 1C). Consistently, *PD-L1* and *PD-L2* were not differentially expressed in π ERV-high versus π ERV-low tumors in HNSC (Supplemental Figure 3D). Instead, the BTLA/HVEM pathway and *LAG3* were upregulated in π ERV-high tumors in HNSC (Supplemental Figure 3D).

π ERV-high tumors in these cancers were also enriched in tumors with known potential predictors of ICA, namely in tumors with APOBEC mutagenesis (25) in ER⁺ HER2⁻, microsatellite instability-high (MSI-H) tumors in COAD, and HPV⁺ tumors in HNSC (Supplemental Figure 1E, Supplemental Figure 2E, and Supplemental Figure 3E). Although hypermutation due to APOBEC and MSI-H etiologies is known to be associated with ICA in ER⁺ HER2⁻ and COAD, respectively (6), a detailed analysis (Supplemental Figure 4) demonstrated that π ERV-high tumors showed evidence of immune activation and checkpoint pathway upregulation compared with π ERV-low tumors in ER⁺ HER2⁻, both with and without APOBEC mutagenesis (25), as well as in both MSI-H and MSI-L/MSS in COAD, and in both HPV⁺ and HPV⁻ in HNSC. These results demonstrate that π ERV expression associates with ICA independently of APOBEC mutagenesis (25) status in ER⁺ HER2⁻, MSI-H status in COAD, and HPV status in HNSC.

Transcriptomic correlates of ERV expression point to chromatin alterations. For further insight, we compared the gene expression profiles of the 3 π ERV subtypes of ccRCC tumors, and tumor-adjacent normal tissues in TCGA data set. Expression levels of 1,048 genes followed the following trend: π ERV-high > π ERV-intermediate > π ERV-low > adjacent normal; whereas 1,103 genes followed the following trend: π ERV-high < π ERV-intermediate < π ERV-low < adjacent normal (where > and < represent statistically

Table 3. Top pathways from enrichment analysis of genes with expression levels that follow the trends π ERV-high > π ERV-intermediate > π ERV-low > adjacent normal or π ERV-high < π ERV-intermediate < π ERV-low < adjacent normal in ccRCC

No.	Source	Pathways Name	FDR	Trend
1	PID	TCR signaling in naive CD8 ⁺ T cells	4.26E-06	H > I > L > adjn
2	PID	TCR signaling in naive CD4 ⁺ T cells	6.02E-06	H > I > L > adjn
3	KEGG	T cell receptor signaling pathway	2.82E-04	H > I > L > adjn
1	REACTOME	The citric acid (TCA) cycle and respiratory electron transport	9.22E-26	H < I < L < adjn
2	REACTOME	Respiratory electron transport, ATP synthesis by chemiosmotic coupling, and heat production by uncoupling proteins	7.82E-21	H < I < L < adjn
3	KEGG	Oxidative phosphorylation	5.83E-19	H < I < L < adjn
4	KEGG	Metabolic pathways	1.12E-17	H < I < L < adjn
5	REACTOME	Respiratory electron transport	2.49E-16	H < I < L < adjn

H, high; I, intermediate; L, low. Significant values are indicated by > (significantly higher) and < (significantly lower) symbols.

significant at $P < 0.05$ in 2-sided Wilcoxon rank-sum test, higher and lower, respectively). Enrichment analysis (Tables 2 and 3) of these two sets of genes using the ToppGene suite (26) showed that the former set was enriched in immune activation genes, whereas the latter set was enriched in genes associated with mitochondrial respiration, which often plays a role in determining disease behavior.

To identify the potential cause of ERV expression, we looked for genes with expression that was correlated with overall ERV expression in each cancer type (see the Methods for details). Six hundred fifty-seven genes satisfied this criterion in KIRC, ER⁺ HER2⁻, and COAD. Enrichment analysis of these genes using ToppGene (26) showed that “methyl” (methyltransferase and methylation) and “histone” pathways were significantly associated with overall ERV expression (Table 4). This suggests that epigenetic alterations, specifically those involving modification of histone methylation–based control of gene expression, may be a functional mechanism of ERV expression in these cancers.

Discussion

These results demonstrate that elevated expression of a set of π ERVs is associated with gene expression-based evidence of ICA in several solid cancers and particularly strikingly in ccRCC. This finding suggests that broad transcriptional activation of π ERVs, present in a subset of tumors, may induce antitumor immune response and subsequent upregulation of immune checkpoint pathways, making such tumors sensitive to ICB. Expression of 7 members of the ERVK family has previously been reported to be correlated with the expression of cytotoxins *GZMA* and *PRFI* in ccRCC (18), and we found an association between *ERV3-2* expression and response to ICB in a small cohort of 24 metastatic ccRCC patients. Although π ERV-high tumors in ccRCC have poor prognosis under standard therapy, such tumors may have significantly improved outcomes with ICB.

In addition to CD8⁺ T cells, M1 macrophages were more abundant and M2 macrophages were less abundant in π ERV-high ccRCC tumors compared with π ERV-low tumors in TCGA. Since M1 macrophages have tumor-inhibitory properties and M2 macrophages may be tumor promoting (27), this suggests that π ERV-high tumors have a relatively favorable immune microenvironment. Consistently, such enrichment of M1 polarization in macrophages has been previously reported in both mutation burden–associated (5, 6) and exogenous virus–associated (7) immunogenicity in other cancers.

Overall ERV expression in tumors correlated with the expression of genes involved in histone methylation and chromatin regulation in multiple cancers, including ccRCC. This suggests that ERV expression may be induced by some underlying dysfunction in chromatin organization, which plays a key role in the normal epigenetic silencing of ERVs. Of note, there was an enrichment of *BAP1* mutations in the π ERV-high subtype of KIRC, suggesting that *BAP1* dysfunction may lead to some chromatin abnormalities resulting in π ERV expression. *BAP1* is a deubiquitinase and is known to functionally associate with chromatin regulating complexes (28), and its role in regulation of π ERV expression warrants further investigation. A recent study (12) reported an association between loss of *PBRM1* and response to non-first-line ICB in pretreated ccRCC. Although we observed no enrichment of *PBRM1* mutations

Table 4. Top GO terms from enrichment analysis of genes with expression levels that correlate with overall ERV expression in ccRCC, ER⁺/HER2⁻ breast cancer, and colon cancer

GO: biological processes			
No.	ID	Name	FDR
1	GO:0018024	Histone-lysine N-methyltransferase activity	4.80E-06
2	GO:0042054	Histone methyltransferase activity	4.17E-05
3	GO:0001071	Nucleic acid-binding transcription factor activity	4.17E-05
4	GO:0003700	Transcription factor activity, sequence-specific DNA binding	4.17E-05
5	GO:0003723	RNA binding	4.17E-05
6	GO:0016279	Protein-lysine N-methyltransferase activity	1.82E-03
7	GO:0016278	Lysine N-methyltransferase activity	1.82E-03
8	GO:0042800	Histone methyltransferase activity (H3-K4 specific)	1.82E-03
9	GO:0033038	Bitter taste receptor activity	5.89E-03
10	GO:0008170	N-methyltransferase activity	5.89E-03
11	GO:0008276	Protein methyltransferase activity	5.89E-03
GO: molecular functions			
1	GO:0006397	mRNA processing	5.44E-06
2	GO:0008380	RNA splicing	5.44E-06
3	GO:0018022	Peptidyl-lysine methylation	1.78E-05
4	GO:0034968	Histone lysine methylation	1.78E-05
5	GO:0016571	Histone methylation	1.78E-05
6	GO:0006479	Protein methylation	2.93E-05
7	GO:0008213	Protein alkylation	2.93E-05
8	GO:0016071	mRNA metabolic process	1.84E-04
9	GO:0031124	mRNA 3'-end processing	1.84E-04
10	GO:0016570	Histone modification	2.02E-04
11	GO:0016569	Covalent chromatin modification	3.56E-04

GO, gene ontology.

in the π ERV-high subtype of KIRC, our results also suggest a link between changes in chromatin and an immune phenotype. It is important to consider that these chromatin modifiers exist with obligated loss of heterozygosity in ccRCC, due to the characteristic chromosome 3p deletion, such that extrapolation to other tumor types, such as papillary RCC or bladder cancers harboring these mutations, may be complex. Prior studies have also shown that ccRCC tumors are relatively enriched in potentially immunogenic frameshift mutations caused by small insertion/deletions (13). The relative association among frameshift mutation burden, presence of *PBRM1* mutations, π ERV expression, and clinical response to ICB needs to be explored in larger prospective studies.

In addition to ccRCC, overexpression of π ERVs was also associated with ICA in ER⁺ HER2⁻, COAD, and HNSC. π ERV overexpression overlapped somewhat with known genomic features potentially associated with ICA in these cancers, such as APOBEC mutagenesis (25) in ER⁺ HER2⁻, microsatellite instability in COAD, and HPV infection in HNSC, raising the possibility that π ERV expression may be biologically related to these genomic features. However, a closer inspection revealed that π ERV overexpression was associated with ICA independently of these features. Moreover, expression of *ERV3-2* in particular was correlated with ICA in 11 solid cancers, including the 4 named above. Thus, like hypermutation or exogenous viral expression in tumors, π ERV expression in tumors is another major mechanism of ICA operative in multiple solid cancers. Our results suggest that overexpression of π ERVs, especially *ERV3-2*, should be investigated further as a feature of immune response in ccRCC and other solid cancers.

In summary, our results suggest that a set of π ERVs is broadly overexpressed in a subset of tumors and is associated with ICA in ccRCC. Analysis of larger cohorts and different cancers, development of clinical-grade assays of π ERV expression, and prospective clinical trials are needed to validate π ERV expression as a predictor of response to ICB in ccRCC and other solid cancers.

Methods

Processing of TCGA RNA-seq data. RNA-seqV2-scaled estimates were obtained from Broad Genome Data Analysis Centers (GDAC) (<http://gdac.broadinstitute.org>) and TCGA data portal (<https://tcga-data.nci.nih.gov/docs/publications/tcga/>). The data were median-adjusted so that the median-scaled estimate was unity in each sample and was then used as input for the ESTIMATE (22) and CIBERSORT (29) algorithms to quantify the level of immune infiltration in tumor (“ImmuneScore”) and the composition of infiltrating leukocytes. Only unambiguous ($P < 0.05$) CIBERSORT outputs were retained. For the remainder of the analysis, the median-adjusted data (x) were log transformed to $y = \log_2(1 + 1,023 \times x)$, so that the lowest possible expression was 0 units and the median expression was 10 units in each sample.

Source of the remaining TCGA data. ERV expression data for a large subset of TCGA cohort were downloaded from a recent study (18) that quantified normalized expression levels of 66 transcribed ERVs (17) by direct remapping from the raw RNA-seq data. Only tumors for which both mRNA and ERV expression data were available were included in the analyses ($n = 472$ for ccRCC, $n = 4,438$ for 20 other cancers). *ERBB2* focal copy number data and *ESR1* mRNA expression data obtained from Broad GDAC were used to classify breast cancer into clinical subtypes (ER⁺/HER2⁻, ER⁻/HER2⁻, HER2⁺), and each subtype was analyzed separately. Clinical and mutation data were downloaded from cBioPortal (<http://www.cbioportal.org>), APOBEC enrichment data were compiled from P-MACD (25) files from Broad GDAC, and microsatellite status and HPV status were compiled from auxiliary files from TCGA Data Portal.

Search for π ERVs. An ERV was considered potentially immunogenic if its expression was correlated with both immune activation and checkpoint pathway upregulation (Figure 1A). The immune activation criterion was considered satisfied if the expression of an ERV correlated with immune infiltration (“ImmuneScore” from ESTIMATE, ref. 22) and mRNA expression of *CD8A* (marker of CD8⁺ T cell infiltration). The checkpoint pathway upregulation criterion was considered satisfied if the expression of an ERV correlated with the PD-1 pathway (i.e., *PD-1* and at least one of its ligands) or the CTLA-4 pathway (i.e., *CTLA-4* and at least one of its ligands) or the BTLA/HVEM pathway. In each cancer type, we tested whether any of the 66 transcribed ERVs (17) satisfied both immune activation and checkpoint pathway upregulation criteria.

Gene expression-based enrichment analyses. By defining overall ERV expression as the total fraction of RNA-seq reads that mapped to the 66 transcribed ERVs (17), genes with expression that correlated (Spearman Rho > 0 , $P < 0.05$) with overall ERV expression in each cancer type were identified. Six hundred fifty-seven genes satisfied this criterion in KIRC, ER⁺ HER2⁻, and COAD. An enrichment analysis of these genes was performed using the ToppGene (26) suite to identify enriched gene ontology terms.

Patient samples in the validation cohort. Formalin-fixed, paraffin-embedded (FFPE) metastatic ccRCC samples from patients treated with single-agent PD-1/PD-L1 blockade were collected at the Vanderbilt-Ingram Cancer Center (using an IRB-exempt waiver of consent, approved by the Vanderbilt University Medical Center Human Research Protection Program) and the Rutgers Cancer Institute of New Jersey (with approval from the Rutgers Institutional Review Board). All tissues were evaluated by a pathologist using hematoxylin and eosin staining, and only the samples containing $\geq 70\%$ tumor cells were included in this study. The cohort included 13 patients who experienced partial response, with progression-free survival of at least 6 months, and 11 patients who demonstrated immediate progressive disease.

Quantification of ERV3-2 expression in the validation cohort. Total RNA isolation was performed using the RNeasy FFPE Kit (Qiagen). DNase treatment was performed during RNA isolation using RNase-free DNase I (Qiagen). RNA quality and concentration were assessed using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). First-strand cDNA synthesis was performed using 250 ng total RNA, random hexamers, and the SuperScript IV Reverse Transcriptase Kit (Life Technologies). RT-qPCRs were performed on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad) using the SYBR green master mix reagent (Life Technologies). Primer sequences for *ERV3-2* and *HPRT1* (fwd: GACACTGGCAAACAATGCAGAC, rev: TGGCTTATATCCAACACTTCGTGG) were designed using PrimerBank (30). *ERV3-2* was quantified using two different primers: ERV3-2_1 (fwd: 5'-CAAGAGGCGGCATAGAAGCAA-3', rev: 5'-GGAGAGTAGCTTGGGGTTTCA-3') and ERV3-2_2 (fwd: 5'-AGCCATTTACAAAGAAAGGGGAC-3', rev: 5'-CTATGCCGCCTCTGTCTGAT-3'). All analyses were performed in triplicate, and relative RNA levels were determined using *HPRT1* as an endogenous internal control. A HeLa control RNA sample was included for interplate calibration. *ERV3-2* expression level was calculated using the $\Delta\Delta C_t$ method.

Statistics. *P* values reported in bar plots, box plots, and Kaplan-Meier plots are from Fisher's exact test, Wilcoxon rank-sum test, and log-rank tests, respectively. Statistical significance was assessed at $P < 0.05$ in 2-sided tests. False discovery rates used the Benjamini-Hochberg procedure, and hazard ratios are from univariate cox regression. Two variables were considered correlated if they had Spearman Rho > 0 and $P < 0.05$, anticorrelated if they had Spearman Rho < 0 with $P < 0.05$, and uncorrelated if $P \geq 0.05$.

Study approval. Work in this study was performed with the approval of the Rutgers Institutional Review Board. Informed consent was deemed unnecessary by the Institutional Review Board of the Vanderbilt University Medical Center Human Research Protection Program.

Author contributions

The study was designed by AP, AADC, SG, GB, and WKR. Data collection was performed by AP and AADC. Samples and clinical data were analyzed by KEB, MS, GR, CCS, BGV, SG, and WKR. The literature search was performed by AP, AADC, SG, GB, and WKR. Figures were created and data analysis was performed by AP and AADC. Data interpretation was performed by AP, AADC, SG, GB, and WKR. Writing was performed by AP, AADC, KEB, SG, GB, and WKR.

Acknowledgments

The study was supported by NIH grants R01-CA198482 (to WKR), K24-CA172355 (to WKR), T32-CA009582-28 (to AADC), and P30-CA072720 (to AP, MS, GR, SG, and GB); a Department of Defense grant CA160728P1 (to WKR); the American Association for Cancer Research — Kure It Research Grant for Immunotherapy in Kidney Cancer (to WKR); a Cancer Research Institute Irvington post-doctoral fellowship (to KEB); a New Jersey Commission on Cancer Research post-doctoral fellowship (to AP); an anonymous gift for precision medicine (to SG); and the Val Skinner Foundation (to SG). Some of this work was performed at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-1607611 (to AP, AADC, and GB).

Address correspondence to: Shridar Ganesan, 195 Little Albany Street, New Brunswick, New Jersey 08903, USA. Phone: 732.235.5211; Email: ganesash@cinj.rutgers.edu. Or to: Gyan Bhanot, 136 Frelinghuysen Road, Piscataway, New Jersey 08854, USA. Phone: 848.391.7508; Email: gyanbhanot@gmail.com. Or to: W. Kimryn Rathmell, 2220 Pierce Avenue, Nashville, Tennessee 37232, USA. Phone: 615.936.3320; Email: kimryn.rathmell@vanderbilt.edu.

- Kim JW, Eder JP. Prospects for targeting PD-1 and PD-L1 in various tumor types. *Oncology (Williston Park, NY)*. 2014;28 Suppl 3:15–28.
- Snyder A, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med*. 2014;371(23):2189–2199.
- Rizvi NA, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348(6230):124–128.
- Le DT, et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 2015;372(26):2509–2520.
- Mehnert JM, et al. Immune activation and response to pembrolizumab in POLE-mutant endometrial cancer. *J Clin Invest*. 2016;126(6):2334–2340.
- Panda A, et al. Identifying a clinically applicable mutational burden threshold as a potential biomarker of response to immune checkpoint therapy in solid tumors. *JCO Precis Oncol*. 2017;2017.
- Panda A, et al. Immune activation and benefit From avelumab in EBV-positive gastric cancer. *J Natl Cancer Inst*. 2018;110(3):316–320.
- Kwong YL, et al. PD1 blockade with pembrolizumab is highly effective in relapsed or refractory NK/T-cell lymphoma failing l-asparaginase. *Blood*. 2017;129(17):2437–2442.
- Nghiem PT, et al. PD-1 blockade with pembrolizumab in advanced merkel-cell carcinoma. *N Engl J Med*. 2016;374(26):2542–2552.
- Kaufman HL, et al. Avelumab in patients with chemotherapy-refractory metastatic Merkel cell carcinoma: a multicentre, single-group, open-label, phase 2 trial. *Lancet Oncol*. 2016;17(10):1374–1385.
- Motzer RJ, et al. Nivolumab versus everolimus in advanced renal-cell carcinoma. *N Engl J Med*. 2015;373(19):1803–1813.
- Miao D, et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. *Science*. 2018;359(6377):801–806.
- Turajlic S, et al. Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. *Lancet Oncol*. 2017;18(8):1009–1021.
- Ascierto ML, et al. The intratumoral balance between metabolic and immunologic gene expression is associated with anti-PD-1 response in patients with renal cell carcinoma. *Cancer Immunol Res*. 2016;4(9):726–733.
- Siska PJ, et al. Mitochondrial dysregulation and glycolytic insufficiency functionally impair CD8 T cells infiltrating human renal cell carcinoma. *JCI Insight*. 2017;2(12):93411.

16. Lander ES, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860–921.
17. Mayer J, Blomberg J, Seal RL. A revised nomenclature for transcribed human endogenous retroviral loci. *Mob DNA*. 2011;2(1):7.
18. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1-2):48–61.
19. Hohn O, Hanke K, Bannert N. HERV-K(HML-2), the best preserved family of HERVs: endogenization, expression, and implications in health and disease. *Front Oncol*. 2013;3:246.
20. Löwer R, et al. Identification of human endogenous retroviruses with complex mRNA expression and particle formation. *Proc Natl Acad Sci USA*. 1993;90(10):4480–4484.
21. Wu HL, et al. Identification and spontaneous immune targeting of an endogenous retrovirus K envelope protein in the Indian rhesus macaque model of human disease. *Retrovirology*. 2016;13:6.
22. Yoshihara K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
23. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013;499(7456):43–49.
24. Kapur P, et al. Effects on survival of BAP1 and PBRM1 mutations in sporadic clear-cell renal-cell carcinoma: a retrospective analysis with independent validation. *Lancet Oncol*. 2013;14(2):159–167.
25. Roberts SA, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat Genet*. 2013;45(9):970–976.
26. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009;37(Web Server issue):W305–W311.
27. Yuan A, et al. Opposite effects of M1 and M2 macrophage subtypes on lung cancer progression. *Sci Rep*. 2015;5:14273.
28. LaFave LM, et al. Loss of BAP1 function leads to EZH2-dependent transformation. *Nat Med*. 2015;21(11):1344–1349.
29. Newman AM, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–457.
30. Wang X, Spandidos A, Wang H, Seed B. PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. *Nucleic Acids Res*. 2012;40(Database issue):D1144–D1149.