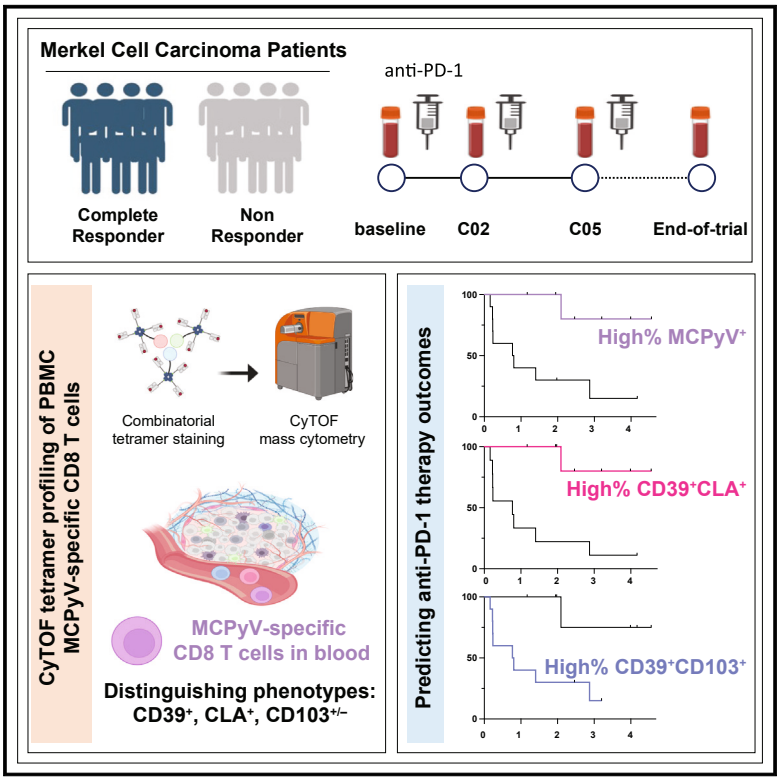
Cell Reports
Medicine

Article

Merkel cell polyomavirus-specific and CD39$^+$CLA$^+$ CD8 T cells as blood-based predictive biomarkers for PD-1 blockade in Merkel cell carcinoma

Graphical abstract



Highlights

- MCPyV-specific CD8 T cells in blood predict MCC patient response to anti-PD-1
- Blood MCPyV-specific CD8 T cells often express CD39, CLA, or CD103
- Bulk CD39$^+$CLA$^+$ CD8 T cells predict response to anti-PD-1
- Bulk CD39$^+$CD103$^+$ CD8 T cells correlate with tumor burden and predict poor outcomes

Authors

Heeju Ryu, Timothy M. Bi, Thomas H. Pulliam, ..., David M. Koelle, Paul Nghiem, Evan W. Newell

Correspondence

enewell@fredhutch.org

In brief

Ryu et al. profile MCC patient peripheral CD8 T cells and find that higher frequencies of MCPyV-specific cells predict response to anti-PD-1 therapy. Based on antigen-specific cell profiles, they identify CD39$^+$CLA$^+$ and CD39$^+$CD103$^+$ frequencies among total CD8 T cells as potential markers of tumor reactivity and as biomarkers for immunotherapy response.

Ryu et al., 2024, Cell Reports Medicine *5*, 101390
February 20, 2024 © 2023 The Author(s).
https://doi.org/10.1016/j.xcrm.2023.101390

CellPress

Cell Reports Medicine

CellPress
OPEN ACCESS

Article

Merkel cell polyomavirus-specific and CD39$^+$CLA$^+$ CD8 T cells as blood-based predictive biomarkers for PD-1 blockade in Merkel cell carcinoma

Heeju Ryu,[1] Timothy M. Bi,[1] Thomas H. Pulliam,[2] Korok Sarkar,[1] Candice D. Church,[2] Nandita Kumar,[1,3] Koshlan Mayer-Blackwell,[1] Saumya Jani,[2,3] Nirasha Ramchurren,[4] Ulla K. Hansen,[5] Sine R. Hadrup,[5] Steven P. Fling,[4] David M. Koelle,[1,3,6,7,8] Paul Nghiem,[2] and Evan W. Newell[1,3,9,*]

[1]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Center, Seattle, WA, USA
[2]Department of Medicine, Division of Dermatology, University of Washington, Seattle, WA, USA
[3]Department of Lab Medicine and Pathology, University of Washington, Seattle, WA, USA
[4]Cancer Immunotherapy Trails Network, Fred Hutchinson Cancer Center, Seattle, WA, USA
[5]Department of Health Technology, Technical University of Denmark, Kongens Lyngby, Denmark
[6]Department of Medicine, Division of Allergy and Infectious Diseases, University of Washington, Seattle, WA, USA
[7]Department of Global Health, University of Washington, Seattle, WA, USA
[8]Benaroya Research Institute, Seattle, WA, USA
[9]Lead contact
*Correspondence: enewell@fredhutch.org
https://doi.org/10.1016/j.xcrm.2023.101390

SUMMARY

Merkel cell carcinoma is a skin cancer often driven by Merkel cell polyomavirus (MCPyV) with high rates of response to anti-PD-1 therapy despite low mutational burden. MCPyV-specific CD8 T cells are implicated in anti-PD-1-associated immune responses and provide a means to directly study tumor-specific T cell responses to treatment. Using mass cytometry and combinatorial tetramer staining, we find that baseline frequencies of blood MCPyV-specific cells correlated with response and survival. Frequencies of these cells decrease markedly during response to therapy. Phenotypes of MCPyV-specific CD8 T cells have distinct expression patterns of CD39, cutaneous lymphocyte-associated antigen (CLA), and CD103. Correspondingly, overall bulk CD39$^+$CLA$^+$ CD8 T cell frequencies in blood correlate with MCPyV-specific cell frequencies and similarly predicted favorable clinical outcomes. Conversely, frequencies of CD39$^+$CD103$^+$ CD8 T cells are associated with tumor burden and worse outcomes. These cell subsets can be useful as biomarkers and to isolate blood-derived tumor-specific T cells.

INTRODUCTION

Merkel cell carcinoma (MCC) is a rare and aggressive form of skin cancer with a high mortality rate. The etiology of MCC has been linked to both Merkel cell polyomavirus (MCPyV) and UV-induced mutagenesis.[1] The persistent expression of two MCPyV oncoprotein isoforms, large T antigen (LTA) and small T antigen (STA), is detectable in 80% of MCC cases in the US, which are categorized as virus positive (VP). The rest of cases where MCPyV cannot be detected are considered virus negative (VN).[2]

Involvement of immune cells in MCC is an important aspect of both disease progression and prognosis, and both VP and VN MCCs are known to be immunogenic.[3,4] Higher intratumoral T cell counts and robust CD8 T cell infiltration within the tumor have been associated with improved survival regardless of the tumor stage at diagnosis.[5] In contrast to many other solid tumors, treatment with the anti-PD-1 antibody has demonstrated remarkable efficacy in both VP and VN advanced MCCs.[6] Clinical studies, including the Cancer Immunotherapy Trials Network-09 (CITN-09)/KEYNOTE-017, have revealed that initial response rates reached up to 58%, with a notable 30% of treatment recipients achieving complete and durable response.[7,8] These promising outcomes provide strong motivation for further investigation into T cell-based biomarkers for MCC.

In particular, MCPyV-specific CD8 T cells play a major role in the immunopathogenesis of MCC. They are found at the site of the tumor and are enriched among tumor-infiltrating lymphocytes.[9] Major histocompatibility complex (MHC) class I-restricted MCPyV oncoprotein processing and presentation by mammalian cells have been shown to lead to CD8-mediated cytotoxicity, and a number of virally derived T cell epitopes have been characterized in the context of MCC.[10,11] Moreover, MCPyV-specific T cells have been identified in both MCC tumors and peripheral blood mononuclear cells (PBMCs) of patients with MCC but are more abundantly found in the former.[12] Even though MCPyV-specific T cells are implicated in immunotherapeutic responses, further qualitative and quantitative analyses of those cells among responders and non-responders are

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

needed to understand why only some VP patients respond to anti-PD-1 checkpoint blockade immunotherapy.

Currently, very few predictive or prognostic biomarkers exist for the diagnosis or treatment of MCC. Prior to the use of checkpoint blockade as a therapy for MCC, it was found that the presence of tumor-infiltrating MCPyV-specific T cells was associated with improved patient survival[9] and that frequencies of peripheral virus-specific cells expressing PD-1 and Tim-3 are generally directly associated with disease burden.[12,13] Consistent with the association between MCPyV-specific CD8 T cell frequencies in the blood and tumor burden, levels of MCPyV-oncoprotein-specific antibodies are also associated with tumor mass. These antibody levels can be used as a prognostic marker for relapse in patients with advanced MCC, particularly those at high risk for recurrent disease.[14] Additionally, intratumoral T cell clonality has been found to be a differentiating characteristic between VP MCCs and VN MCCs, with the former showing higher clonality indices than the latter.[9] However, as of yet, there are no clear blood-based characteristics that can predict which patients are more likely to respond to anti-PD-1 therapy, and the mechanisms of response and resistance remain poorly understood.

Because of the high rates of complete initial responses and the consistent expression of viral antigens, VP MCC provides an ideal model system to track and evaluate T cell responses during anti-PD-1 checkpoint blockade immunotherapy. Using MHC class I tetramers against MCPyV oncoproteins, it is possible to explore shared tumor-specific T cell responses across patients with VP MCC without the requirement of identifying mutation-derived antigens or other tumor-associated antigens, which are often challenging to identify and typically not always shared among patients. Additionally, leveraging the ability to detect tumor-specific T cell responses in treatment-responsive patients with VP MCC may also help inform therapeutic strategies for patients with VN MCC, which are driven by less defined and likely more heterogeneous UV-induced neoantigens and are therefore more difficult to study. Ultimately, this could help identify correlates of anti-PD-1 response and inform strategies to improve efficacy.

In this study, we examine tumor-specific T cell dynamics in the blood of patients with MCC treated with an anti-PD-1 antibody, pembrolizumab. Taking advantage of the persistent expression of MCPyV oncoproteins in VP MCC tumors, we analyze CD8 T cells that target these proteins to directly study and measure tumor-specific T cell responses. We utilized a mass-cytometry-based multiplexed peptide-MHC class I-tetramer staining approach to first assess the frequencies and phenotypes of MCPyV-specific CD8 T cells in the blood of patients with MCC receiving pembrolizumab. Based on the phenotypes of the virus-specific cells, we identified that the frequencies of certain phenotypic subsets within total blood CD8 T cells were predictive of patient response to therapy, and these phenotypic signatures may be useful to indirectly identify tumor- and MCPyV-specific cells. Our findings provide valuable insights into tumor-specific immune responses in MCC and identify a potential biomarker for predicting immunotherapy response and isolating tumor-specific T cells from blood.

RESULTS

Identification and quantification of MCPyV-specific CD8 T cells in blood from patients with MCC

To examine the effect of pembrolizumab treatment on peripheral CD8 T cell responses in the context of MCC, we analyzed longitudinally collected PBMC samples from patients enrolled in the CITN-09 clinical trial (ClinicalTrials.gov: NCT02267603). We carefully selected a patient sample cohort that encompassed diverse clinical responses and viral status (complete response [CR], n = 13; partial response [PR], n = 6; and stable and progressive disease [SD/PD], n = 5; viral status: VP, n = 17; VN, n = 7). To ensure robust data collection and analysis, we considered human leukocyte antigen (HLA)-typing information and the availability of PBMCs for at least baseline and end-of-treatment time points (see method details). Specifically, we analyzed samples obtained prior to treatment (C01), at 3 weeks post-treatment (C02), at 12 weeks post-treatment (C05), and at the end of treatment (EOT) (Table S1).

To detect treatment-associated phenotypic changes in circulating MCPyV-specific T cells, we used mass cytometry (CyTOF) combined with multiplexed MHC class I-tetramer staining (as described previously) to screen 76 MCPyV-associated epitopes and 44 MCC-unrelated viral epitopes (spanning A*01:01, A*02:01, A*03:01, A*11:01, A*24:02, A*68:01, B*07:02, B*08:01, B*15:01, B*35:01, B*37:01, and B*57:01) (Figures S1A and S1B; Table S2).[15–17] Our MCPyV epitope panel allowed for the direct detection of tumor-specific CD8 T cells in VP MCC and included both previously known epitopes for both small and large T oncoproteins as well as epitopes that were computationally predicted to bind to HLA allelic variants in our cohort (see methods details).[11,12] The 35 markers in our phenotyping panel allowed us to detect broad immune compartment changes as well as T cell phenotypic markers to identify cell states such as differentiation, trafficking, activation, and exhaustion (Figure 1A; key resources table).[15] Ultimately, this enabled us to identify a total of 111 T cell populations that were specific for 16 MCPyV T oncoprotein-associated epitopes including 5 novel epitopes (based on epitope predictions) and 3 MCPyV capsid-protein-associated epitopes from 72 samples (n = 24 patient over 2–4 time points), which we could then further phenotypically characterize (Figure 1B). To comprehensively measure the overall changes to antigen-specific T cells, we combined the frequencies of cells detected by tetramers loaded with epitopes of the same antigen specificity within each patient across different time points (Figures 1C–1E).

First, we analyzed the frequencies of peripheral MCPyV-specific T cells and other MCC-unrelated virus-specific T cells at baseline to identify any correlations to response. As expected, VN patients had lower levels of MCPyV-specific CD8 T cells compared to VP patients ($p = 0.0567$ Mann-Whitney test) (Figures S1C and S1D). This is in line with previous reports that quantified MCPyV-specific CD8 T cells among tumor-infiltrating T cells (tumor-infiltrating lymphocytes [TILs]) in patients with VP MCC.[12] In addition, we found that the baseline frequency of circulating MCPyV-specific CD8 T cells summed across all MCPyV-epitope-specific T cell populations detected for each sample was associated with improved response (Figure 1C).

CellPress
OPEN ACCESS

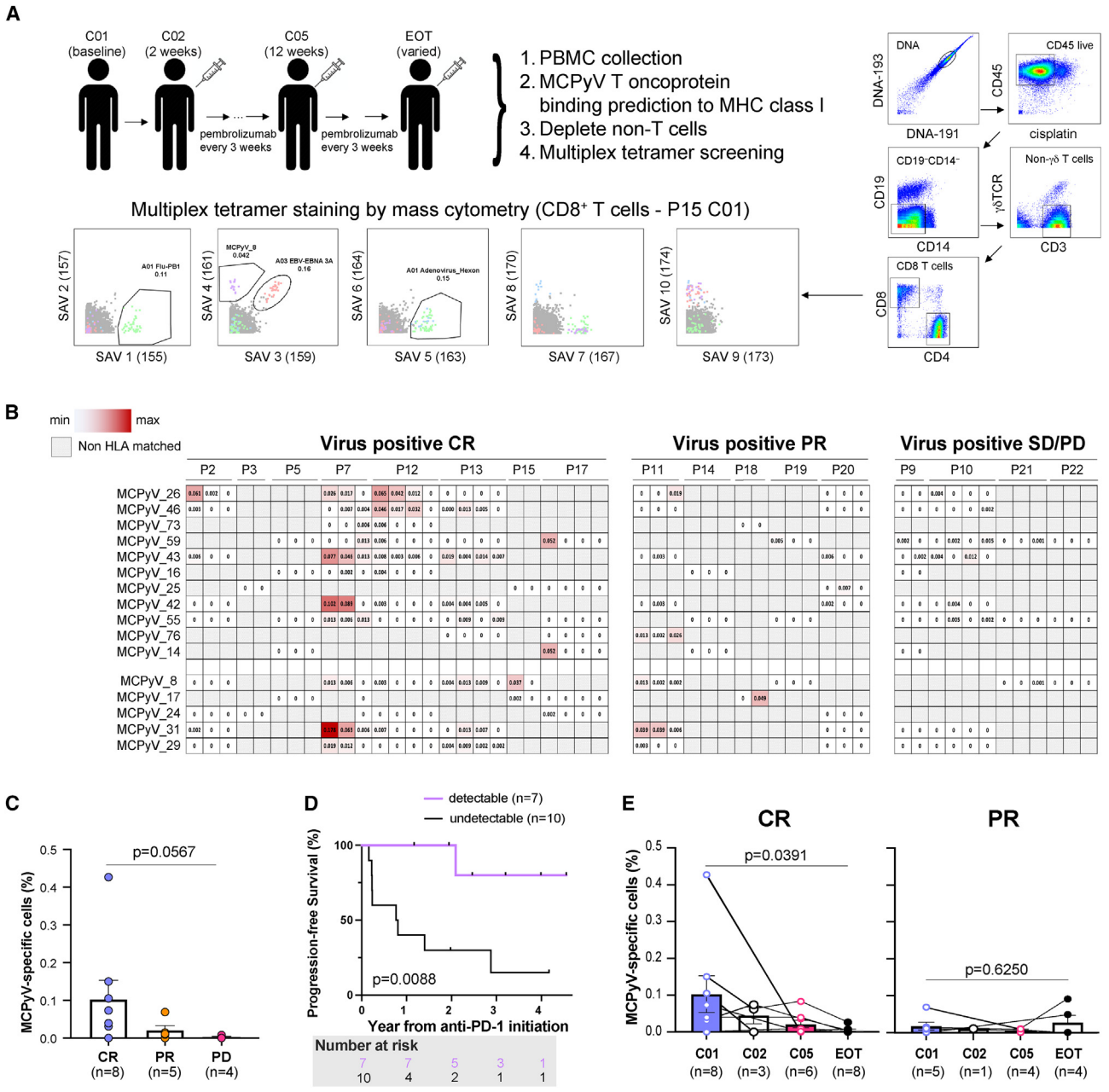


Figure 1. Identification of MCPyV-specific CD8 T cells and correlation with MCC pembrolizumab treatment clinical outcomes

(A) Experimental schematic and representative mass cytometry gating tree and dot plots of CD8 T cells showing the detection of MCPyV-specific and MCC-unrelated epitopes. Representative plots from one patient (P15) and one time point (C01, baseline).

(B) Heatmap of individual MCPyV epitopes detected across all VP patients with complete response (CR; n = 8), partial response (PR; n = 5), stable disease (SD; n = 1), and progressive disease (PD; n = 3).

(C) Integrated frequencies of MCPyV-specific CD8 T cells in PBMCs prior to the treatment. Data are represented as mean ± SEM; Mann-Whitney test.

(D) Kaplan-Meier curve of progression-free survival in patients with detectable (>0.01% of CD8 T cells, violet) or undetectable (<0.01% of CD8 T cells, black) MCPyV-specific CD8 T cells in PBMCs. Detection limit was determined by the highest frequency observed in VN patients. Log-rank test.

(E) Frequencies of MCPyV-specific CD8 T cells in PBMCs over the course of the therapy in patients with CR (left) and PR (right). Data are represented as mean ± SEM. Wilcoxon test. SAV, streptavidin.

Specifically, the mean frequency of MCPyV-specific CD8 T cells in CR patients was 5 times higher than in PR patients (mean = 0.1032 vs. 0.02077, p = 0.2199 Mann-Whitney test) and 38 times higher than in SD/PD patients (mean = 0.002654, p = 0.0727 Mann-Whitney test). Furthermore, patients with detectable (>0.01%) MCPyV-specific CD8 T cells in their blood at baseline

had improved overall survival and progression-free survival compared to patients that did not (Figures 1D and S1E; p = 0.0088 and 0.0405, respectively, log-rank test).

When analyzing on-treatment kinetics in MCPyV-specific T cell frequencies in VP patients, we found that pembrolizumab treatment was associated with a significant decrease in frequencies of MCPyV-specific CD8 T cells in CR patients (p = 0.0391 Wilcoxon) but were only marginally different in PR and SD/PD patients across treatments (Figures 1E and S1F). The frequencies of MCC-unrelated virus-specific cells (e.g., cytomegalovirus [CMV], Epstein-Barr virus [EBV], herpes simplex virus [HSV], and influenza) remained unchanged over the course of the therapy (p = 0.5879 Wilcoxon), suggesting that MCPyV-specific cells in the blood are closely associated with tumor immune responses (Figure S1G). Taken together, these results indicate that tumor-associated MCPyV-specific CD8 T cells are present in the periphery and are closely linked to patient clinical outcomes.

Phenotypic characterization of tumor-specific CD8 T cells in patients with MCC

To further investigate peripheral tumor-associated MCPyV-specific CD8 T cells that correlate with patient clinical response, we performed deep immune phenotyping of these cells and compared their phenotypes to MCC-unrelated virus-specific T cells and the peripheral CD8 compartment as a whole. The bulk peripheral CD8 T cell compartment of patients with MCC is highly heterogeneous and comprised of cells with various distinct phenotypes including naive ($CCR7^+CD45RA^+$), effector memory (EM; $CCR7^-CD45RA^-CD45RO^+$), effector memory re-expressing CD45RA (TEMRA; $CCR7^-CD45RA^+CD45RO^-$), tissue resident-like (RM; $CD103^+$), and activated/exhausted ($CD38^+PD\text{-}1^+HLA\text{-}DR^+CD39^+$). To examine high-dimensional phenotypes of peripheral CD8 T cells, we embedded cell surface markers using UMAP (uniform manifold approximation and projection) and applied the Leiden algorithm based clustering to samples across the entire cohort (Figure 2A; see method details). In parallel, we also quantified the frequencies of each marker expression across all antigen-specific cells identified for each patient time point to corroborate these findings (exemplified in Figure S1H).

When comparing the phenotypes of MCPyV-specific and MCC-unrelated virus-specific CD8 T cells (CMV, HSV, and influenza), we identified unique phenotypic profiles that differentiate these cell populations based on their antigen specificity (Figures 2B and S1I). Our analysis of virus-specific T cells with MCC-unrelated specificities showed that CMV-specific T cells expressed high levels of CD45RA, CD57, and KLRG1, consistent with TEMRA-like phenotypes, and EBV-specific cells were enriched for CXCR5, CD27, and CD45RO (Figure 2C). As previously reported, HSV-specific cells expressed high levels of cutaneous lymphocyte-associated antigen (CLA), a skin-homing marker, and influenza-specific cells expressed high levels of CXCR3 and CCR5 (Figure 2C).[18]

In contrast, we found that MCPyV-oncoprotein-specific CD8 T cells (LTA and STA) expressed various co-stimulatory and inhibitory receptors, such as PD-1 and TIGIT, and were highly enriched for CD39, a marker previously reported to distinguish tumor-specific exhausted CD8 TILs (Figure 2C).[19–22] These cells also highly expressed proliferation, activation, and effector T cell markers, such as HLA-DR, CD38, CD71, and CXCR3. Interestingly, they also highly expressed the skin-homing marker CLA and sporadically expressed a tissue-recirculating marker CD103.[23–25] These observations were even more prominent when we quantified the marker expression of each epitope-specific cell population individually (i.e., when multiple epitope-specific populations for each patient time point were detected) (Figure 2D). Note that in this analysis, we observed some MCPyV-specific cell populations to be highly elevated for CD103 expression, which was not as readily apparent when MCPyV-specific cells were aggregated within samples. We also found that the expression levels of CD39, CLA, and CD103 on MCPyV-specific cells were maintained over the course of therapy (Figure 2C).

Notably, we also found that the phenotypes of CD8 T cells that were specific for the non-oncogenic MCPyV capsid protein were similar to MCPyV-oncoprotein-specific CD8 T cells. However, these cells exhibited lower levels of activation markers such as HLA-DR and CD71 and terminal differentiation markers such as CD45RA and KLRG1. Viral capsid-protein-specific CD8 T cells, like their oncoprotein-specific counterparts, expressed relatively high levels of CD39 in comparison to other bystander T cells, which might be a result of persistent MCPyV capsid expression in the tumor. To test this hypothesis, we analyzed public MCC tumor RNA sequencing datasets and aligned gene sequences of LTA, STA, and major capsid protein VP1 by employing the Basic Local Alignment Search Tool (BLAST).[26–28] Expression of LTA and STA transcripts were detectable in more than 90% of VP patients, with a much lower prevalence of detection in VN patients (Figures S1J and S1K). Interestingly, VP1 expression was identified in 30% of VP patients, whereas none of the VN patients expressed detectable VP1 transcripts (Figure S1L). This finding supports the notion that persistent MCPyV capsid expression may contribute to the observed phenotype of capsid-protein-specific CD8 T cells. Taken together, MCPyV-specific CD8 T cells display a unique phenotype characterized by elevated expression of co-stimulatory and inhibitory marker receptors and enrichment for CD39, CLA, and CD103. The maintenance of these phenotypes throughout the course of therapy suggests their potential as immune biomarkers and therapeutic targets.

$CD39^+CLA^+$ CD8 T cells in blood as predictive biomarkers for favorable clinical outcomes in patients with MCC

Since the phenotypes of MCPyV-specific CD8 T cells remained consistent across time points, we sought to explore if the decline in MCPyV-specific cell frequencies could be attributed to changes in the overall composition of the CD8 T cell compartment, independent of antigen specificities detected by MHC class I tetramers. In line with this, we observed notable alterations in the composition of bulk CD8 T cells, specifically within populations characterized by high expression levels of CD39, CLA, and CD103 (Figure S2A). This was one indication that we might be able to leverage our knowledge of the detailed phenotypes of MCPyV-specific cells to identify high-dimensional

CellPress
OPEN ACCESS

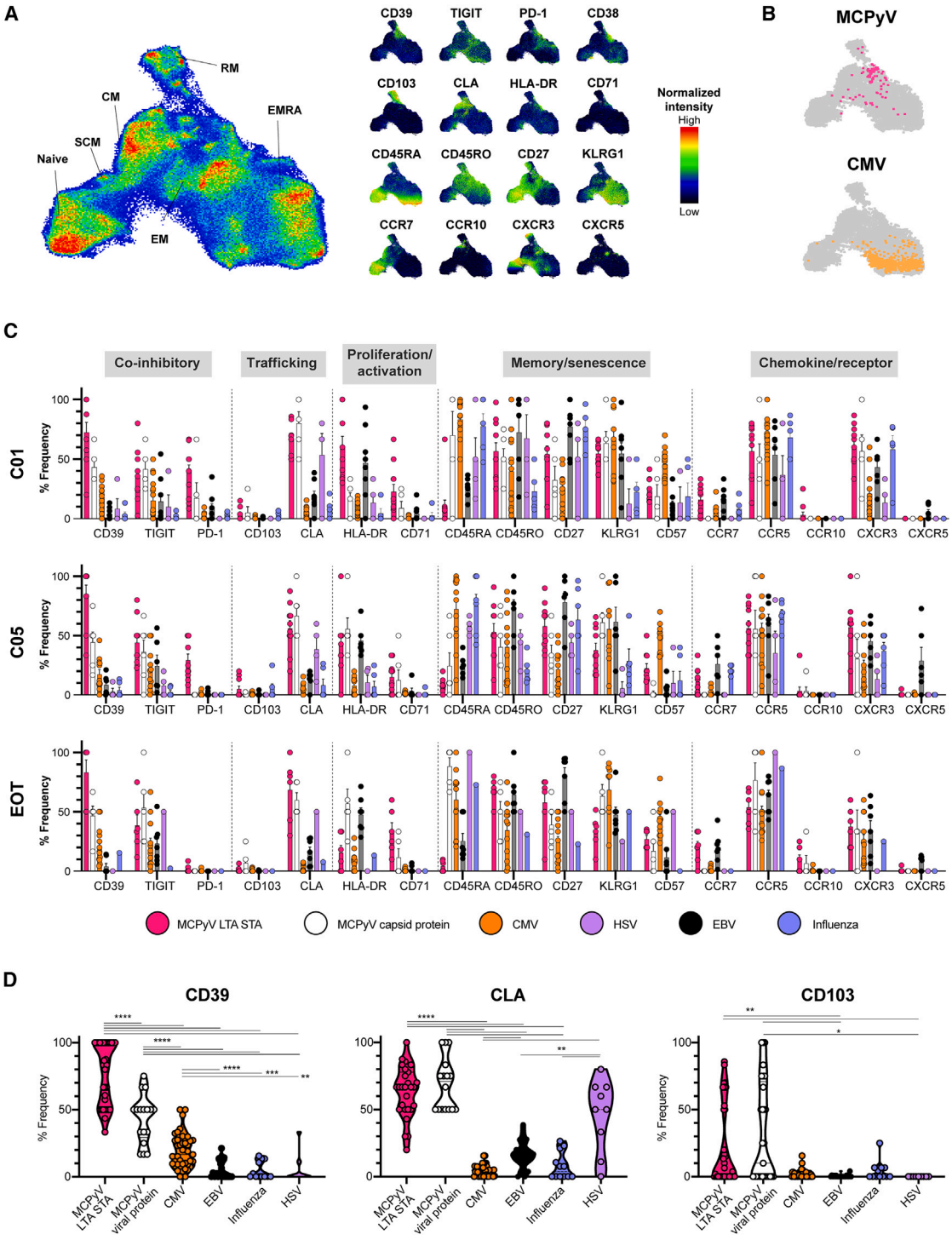


Figure 2. Phenotypic profiling of MCPyV-specific CD8 T cells in patients with MCC

(A) UMAP embedding of CD8 T cells from all patients (left). Normalized expression of selected markers defining CD8 T cell clusters (right).

(B) MCPyV-specific (top) and CMV-specific (bottom) cells projected onto a UMAP embedding. Manually gated individual tetramer$^+$ cells were concatenated.

(C) Expression of markers by indicated tetramer$^+$ cells within CD8 T cells from individual patients over the course of pembrolizumab (top: C01 or baseline, middle: C05, bottom: end of treatment [EOT]). Data are represented as mean ± SEM. LTA, large T antigen; STA, small T antigen; CMV, cytomegalovirus; HSV, herpes simplex virus; EBV, Epstein-Barr virus.

(D) Expression of CD39, CLA, and CD103 by individual tetramer$^+$ cells within CD8 T cells of MCPyV LTA and STA (n = 32), MCPyV viral protein (n = 17), CMV (n = 41), EBV (n = 21), influenza (n = 12), and HSV (n = 9). Mann-Whitney test, *p < 0.05, **p < 0.01, ***p < 0.001, and ****p < 0.0001.

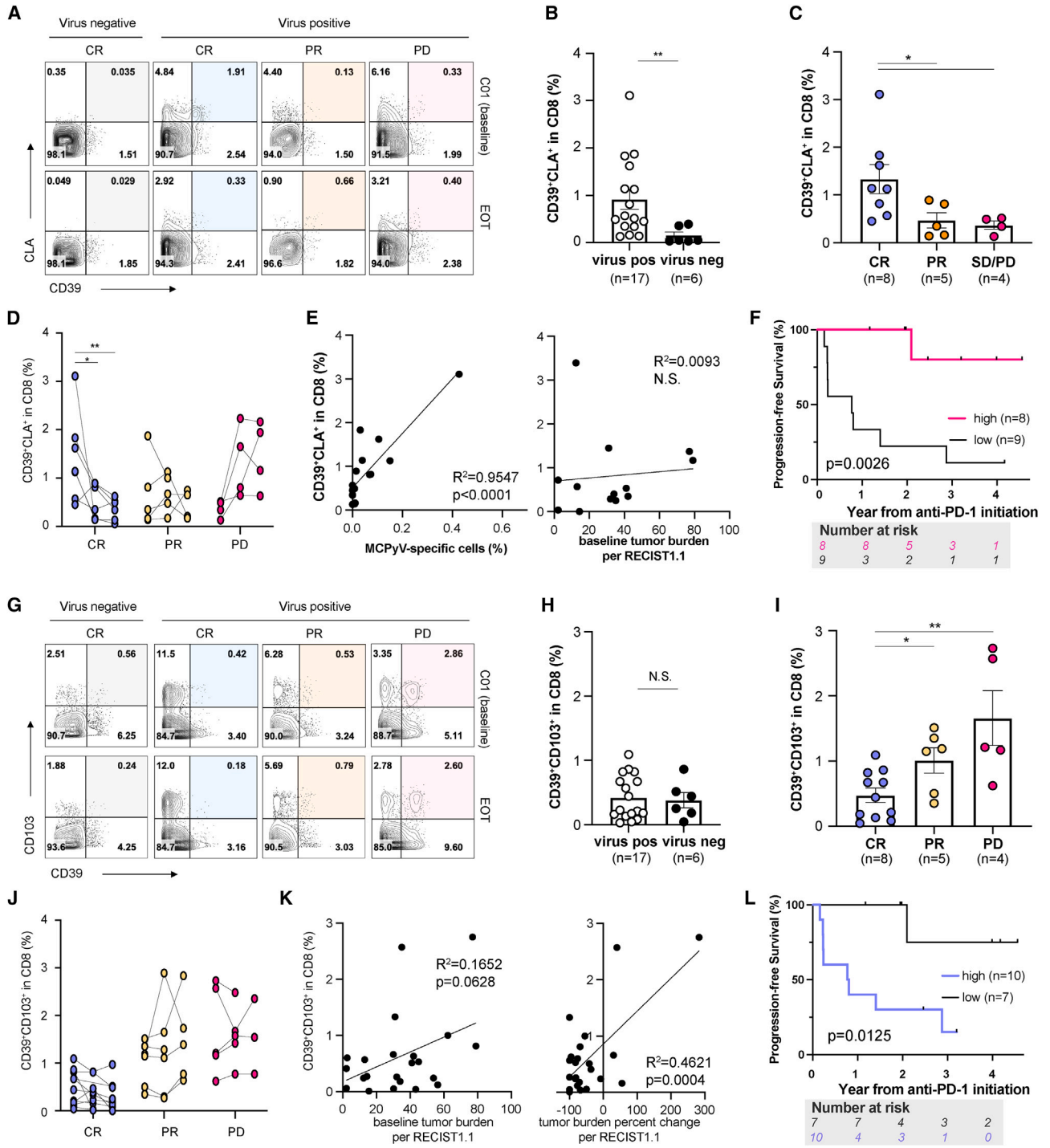


Figure 3. Prognostic potential of CD39, CLA, and CD103 expression among CD8 T cells in the peripheral blood of patients with MCC

(A) Contour plots identifying CD39$^+$CLA$^+$ CD8 T cells by mass cytometry. CD8 T cells were defined and pregated as live/CD45$^+$/CD19$^-$/CD14$^-$/CD3$^+$/γδTCR$^-$/CD4$^-$/CD8$^+$. Representative data.

(B) Frequencies of CD39$^+$CLA$^+$ CD8 T cells prior to treatment in all patients. Data are represented as mean ± SEM; Mann-Whitney test.

(C) Frequencies of CD39$^+$CLA$^+$ CD8 T cells prior to treatment in VP patients. Data are represented as mean ± SEM; Mann-Whitney test.

(D) Frequencies of CD39$^+$CLA$^+$ CD8 T cells over the course of pembrolizumab in VP patients. Data are represented as mean ± SEM; Wilcoxon test.

(E) Linear regression analysis of frequencies of CD39$^+$CLA$^+$ CD8 T cells and frequencies of MCPyV-specific CD8 T cells (left) or baseline tumor burden (right) in VP patients.

(F) Kaplan-Meier curve of progression-free survival in VP patients with high or low CD39$^+$CLA$^+$ CD8 T cells. Log-rank test.

(legend continued on next page)

clusters of bulk CD8 T cells that could also be useful as biomarkers for immunotherapeutic responses.

To identify unbiased correlates of clinical outcome, we conducted clustering analysis (Figures S2B and S2C) followed by regularized regression analysis using LASSO to identify high-dimensionally defined clusters of CD8 T cells with frequencies predictive of patient outcome (see STAR Methods). This analysis identified cluster 18 ($CD39^+CLA^+$) as a positive predictor of the patient's clinical outcome (Figure S2D). It is noteworthy that in CR patients, the frequency of cluster 18 increased shortly after the first post-therapy time point (C02) but declined post-treatment (EOT). However, in PR and SD/PD patients, these cells persisted at the same level or were even elevated (Figure S2E).

Since CD39 has been previously identified as a marker of tumor-specific, terminally exhausted CD8 TILs and is highly expressed by patient MCPyV-specific cells, we initially investigated whether the expression of CD39 in CD8 T cells alone could serve as a predictor of patient response to therapy.[19,20,29] We found that the baseline frequency of $CD39^+$ CD8 T cells in CR patients was associated with baseline tumor burden; however, it did not show any correlation with the clinical outcome or viral status of the patients (Figures S2F and S2G). After anti-PD-1 therapy, the frequencies of $CD39^+$ CD8 T cells decreased in patients who achieved CR but remained the same or decreased in patients who achieved PR and SD/PD (Figure S2H). Additionally, the frequencies of $CD39^+$ CD8 T cells in VP and VN patients correlated with the patients' baseline tumor burden (Figure S2I).

Further informed by the phenotypes of MCPyV-specific CD8 T cells, we hypothesized that CLA and CD103 could be used to better delineate $CD39^+$ CD8 T cells in blood samples, which could improve our ability to identify biomarkers for predicting the clinical prognosis of patients (Figure 3A). We found that the frequencies of $CD39^+CLA^+$ CD8 T cells were higher in VP patients than in VN patients, suggesting a potential association with MCPyV-driven tumors ($p = 0.0032$ Mann-Whitney test; Figure 3B). Similar to MCPyV-specific cells, the frequency of these cells was higher at baseline in patients with VP MCC who achieved a complete response to therapy ($p = 0.0295$ for CR vs. PR; $p = 0.0162$ for CR vs. PD, Mann-Whitney test; Figure 3C) and decreased over time, while the frequency of these cells remained the same or increased after the therapy in patients with PR or SD/PD ($p = 0.0469$ and $p = 0.0096$ Wilcoxon test; Figure 3D). This suggests that the presence of $CD39^+CLA^+$ cells at baseline may indicate a favorable treatment outcome in MCC. The frequencies of these cells were correlated with the frequency of MCPyV-tetramer-positive cells, but not with baseline tumor burden, unlike the total frequency of $CD39^+$ CD8 T cells (Figure 3E). This indicates a potential relationship between the presence of these cells and tumor-specific responses in patients with MCC that are not confounded by baseline tumor burden.

To validate the efficacy of $CD39^+CLA^+$ as a binary classifier that can distinguish between complete response (CR) and non-complete response (non-CR), we employed a receiver operating characteristic curve (ROC) to estimate an optimal threshold for classification (area under the curve [AUC] = 0.9028, $p = 0.0053$ for AUC > 0.5; Figure S3A). Note that the ROC analysis only included VP patients due to the scarcity of $CD39^+CLA^+$ CD8 T cells in VN patients. By utilizing a threshold of 0.81% as determined by ROC analysis, we found that patients with a higher frequency of $CD39^+CLA^+$ CD8 T cells prior to treatment achieved a 75% progression-free survival (PFS) rate at 3 years post-treatment. In contrast, patients that fell below this threshold exhibited a substantially lower PFS rate of 11% (Figure 3F; $p = 0.0026$ log-rank test). Furthermore, patients with high $CD39^+CLA^+$ levels demonstrated an overall survival rate of 100% at the 4 year mark, while only 40% patients below the threshold achieved a similar outcome (Figure S3B, left; $p = 0.0285$ log-rank test). Notably, the significance of this predictor was retained even when VN patients were included in the analysis, despite the threshold being initially determined using VP patients (Figure S3C, left; $p = 0.0026$ log-rank test). These findings underscore the potential utility and robustness of the frequency of $CD39^+CLA^+$ CD8 T cells as a predictive marker for therapeutic response.

$CD39^+CD103^+$ CD8 T cells in blood as a predictive biomarker for poor clinical outcomes in patients with MCC

A similar analysis of $CD39^+CD103^+$ CD8 T cells showed a different association with clinical outcomes. That is, these cells were present in both MCPyV-positive and -negative patients and were found to be higher in patients with MCC with a poor clinical outcome (Figures 3G–3I). The frequencies of these cells slightly decreased ($p = 0.0955$) over the course of pembrolizumab treatment in CR, but not in PR and SD/PD, patients (Figure 3J). Furthermore, frequencies of $CD39^+CD103^+$ among CD8 T cells were correlated with baseline tumor burden and changes in tumor burden (Figure 3K). Overall, this suggests that the presence of $CD39^+CD103^+$ CD8 T cells may associate with a less favorable response to treatment in MCC.

To validate the potential utility of $CD39^+CD103^+$ for binary classification of responsiveness to PD-1 blockade, we determined an optimal threshold of 0.82% through ROC analysis (AUC = 0.9444, $p = 0.002$ for AUC > 0.5; Figure S3A) for VP patients. Patients with a higher frequency of $CD39^+CD103^+$ CD8 T cells prior to treatment achieved a PFS rate of 10% at 3 years,

(G) Contour plots identifying $CD39^+CD103^+$ CD8 T cells by mass cytometry, CD8 T cells were defined and pregated as live/$CD45^+$/$CD19^-$/$CD14^-$/$CD3^+$/$\gamma\delta TCR^-$/$CD4^-$/$CD8^+$. Representative data.

(H) Frequencies of $CD39^+CD103^+$ CD8 T cells prior to treatment in all patients. Data are represented as mean ± SEM; Mann-Whitney test.

(I) Frequencies of $CD39^+CD103^+$ CD8 T cells prior to treatment in VP patients. Data are represented as mean ± SEM; Mann-Whitney test.

(J) Frequencies of $CD39^+CD103^+$ CD8 T cells over the course of pembrolizumab in VP patients. Data are represented as mean ± SEM; Wilcoxon test.

(K) Linear regression analysis of frequencies of $CD39^+CD103^+$ CD8 T cells and frequencies of MCPyV-specific CD8 T cells (left) or baseline tumor burden (right) in VP patients.

(L) Kaplan-Meier curve of progression-free survival in VP patients with high or low $CD39^+CD103^+$ CD8 T cells. Log-rank test. $^*p < 0.05$ and $^{**}p < 0.01$. Each symbol represents an individual patient.

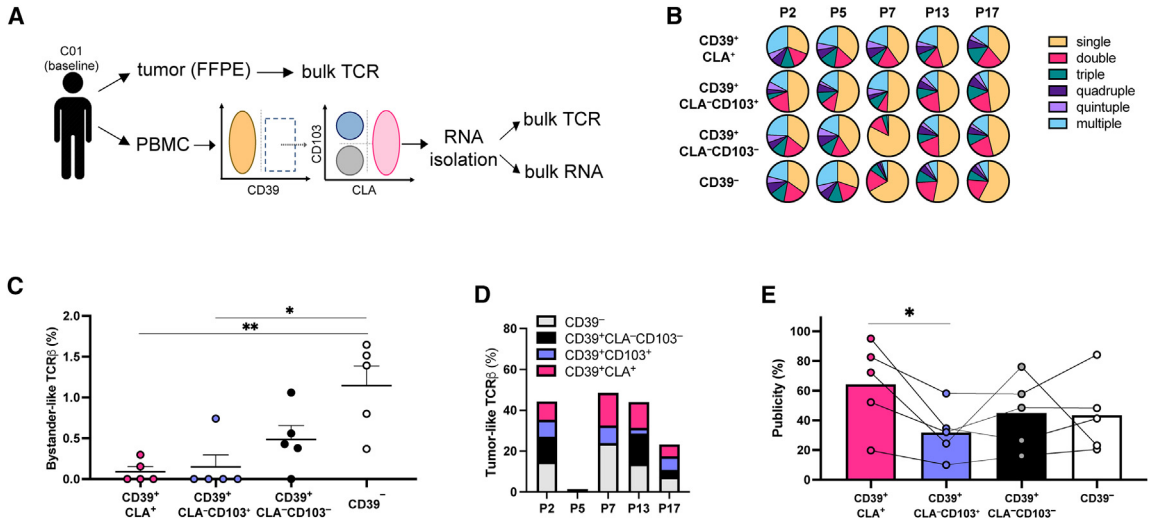


Figure 4. TCR repertoire analysis of circulating CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ CD8 T cells from patients with MCC

(A) Experimental schematic. Bulk TCRβ sequencing was performed on FFPE tumor biopsy samples (n = 5) obtained from patients with a complete response. Patient-matched PBMC samples collected prior to treatment were subjected to flow-based sorting followed by RNA isolation, bulk TCRβ sequencing, and RNA sequencing.

(B) Pie chart illustrating the distribution of unique TCR clonotypes within each population of CD8 T cells.

(C) Frequency of bystander-like TCRβ from each CD8 population determined by clustering and quantifying their connection frequency with VDJ database based on sequence similarity (distance $\leq$ 12). Data are represented as mean ± SEM; Mann-Whitney test.

(D) Frequency of tumor-like TCRβ from each CD8 population determined by clustering and quantifying their connection frequency with patient-matched tumor TCRβ based on sequence similarity (distance $\leq$ 12).

(E) Frequency of publicity determined by clustering and quantifying their connection with other patients based on sequence similarity (distance $\leq$ 12). Data are represented as mean ± SEM, Wilcoxon test. *p < 0.05 and **p < 0.01. Each symbol represents an individual patient.

while patients falling below this threshold demonstrated a significantly higher PFS rate of 70% (Figure 4L; p = 0.0125 log-rank test). This observation also held true when analyzing overall survival (Figure S3B, right; p = 0.0453 log-rank test) and even when including VN patients (Figure S3C, right; p = 0.0048 log-rank test).

Since we discovered that these two phenotypically defined cell subsets predict patients' clinical outcomes in opposite directions, we hypothesized that the ratio of their frequencies would yield improved predictive accuracy. By employing an optimal threshold of 0.695, inferred from ROC analysis of the CD39$^+$CLA$^+$/CD39$^+$CD103$^+$ ratio (AUC = 0.9444, p = 0.0021 for AUC > 0.5), we found that this ratio can indeed predict improved PFS (Figure S3D) and overall survival (Figure S3E) in both VP (left) and VN patients (right). Altogether, these findings suggest that CD39$^+$ cells co-expressing CLA or CD103 in blood could serve as potential blood-based biomarkers of response to immunotherapy in patients with MCC.

Distinct tumor antigen-specificity of CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ CD8 T cells in blood

Having identified two biomarkers with contrasting predictions for patient clinical outcomes, we next sought to assess the antigen specificity of these cells. Among VP patients at the baseline time point, we observed that within the CD39$^+$CLA$^+$ cell populations, approximately 5% of the cells bound to MCPyV tetramers. In contrast, only 0.05% of the CD39$^+$CD103$^+$ cells showed MCPyV tetramer positivity. This indicates that CD39$^+$CLA$^+$ CD8 T cells are significantly enriched for MCPyV reactivity, despite the possibility of false negatives affecting these frequencies (Figure S2J).

As an alternative method to test whether CD39$^+$CLA$^+$ and CD39$^+$CD103$^+$ populations were enriched for tumor reactivity, we used additional PBMC samples and sorted bulk CD8 T cells into four subpopulations: (1) CD39$^-$, (2) CD39$^+$CLA$^-$CD103$^-$, (3) CD39$^+$CLA$^+$CD103$^{+/-}$, and (4) CD39$^+$CLA$^-$CD103$^+$. These subpopulations were sorted from the baseline blood samples of VP CR patients (n = 5), and bulk T cell receptor (TCR) sequencing was performed on the sorted populations to assess the TCR repertoire. Depending on the sample, 5,000–100,000 cells were sorted for each population prior to bulk TCRα and -β sequencing (analysis was focused on TCRβ; see method details). Additionally, we obtained TCRβ clonotypes from patient-matched tumors (n = 5) preserved in formalin-fixed, paraffin-embedded (FFPE) samples (Figure 4A). All sorted populations contained a range of singleton and expanded clones (Figure 4B), and the number of detected clonotypes varied mostly as a function of the number of cells sorted (Figure S4A). However, when we corrected for sampling biases using the true diversity method, which considers both the number of clonotypes present and their relative abundances, the effective number of clonotypes was the lowest in CD39$^+$CLA$^+$ cells, followed by CD39$^+$CLA$^-$CD103$^+$.[30] This observation is consistent with some degree of clonal selection within these populations and indicates a more restricted diversity within CD39$^+$CLA$^+$ cells (Figure S4B). To explore antigen specificity, we compared the TCRβ sequences of each group to TCRs with known specificity in the VDJ database (see

method details).[31] TCRβ sequences recovered from cells with CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ phenotypes were less frequently matched with the database receptors compared to CD39$^-$ and CD39$^+$CLA$^-$CD103$^-$ cells (Figure 4C). This suggests that CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ cells are less likely to represent MCC-unrelated bystander T cells.

Next, to examine potential overlap of similar TCRβ sequences between tumor and blood populations, we computed the Morisita-Horn indexes as well as the pairwise sequence distances (TCRdist) among receptors and grouped highly similar sequences into clusters of *meta*-clonotypes (Figures S4C–S4E).[32,33] The number of clones recovered varied greatly among T cell populations (Figure S4A). Thus, before clustering to identify sequence overlap, we normalized repertoires to the clone count of the smallest repertoires. After downsampling to correct for the variable numbers of clonotypes detected in each of the samples, we found that CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ populations exhibited a similar degree of shared TCRβ sequences with tumors compared to CD39$^-$ and CD39$^+$CLA$^-$CD103$^-$ populations (Figure 4D). Comparable results were obtained using the Morisita-Horn index, a measure of sequence overlap, which corrects for the degree of sampling (Figure S4C). Considering that the CD39$^-$ population—but not the CD39$^+$CLA$^+$ or CD39$^+$CLA$^-$CD103$^+$ population—were found to be enriched for bystander activity in blood (Figure 4C) and in tumors,[19,20,34] we infer that CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ populations play a role in recognizing non-bystander cells that infiltrate tumors, indicating their potential specificity toward tumor recognition.

Lastly, to compare the clonal sharing of the CD39$^+$CLA$^+$ cells (most correlated with MCPyV-specific cells and associated with favorable clinical outcomes) to the CD39$^+$CLA$^-$CD103$^+$ cells (less correlated with MCPyV-specific cell frequencies and associated with worse clinical outcomes), we calculated a publicity score for each sample (see method details) to determine the extent to which the clones within these two populations were shared between individuals. From this, we observed that the TCRs in T cells of CD39$^+$CLA$^-$CD103$^+$ had low publicity compared to TCRs in CD39$^+$CLA$^+$ (Figure 4E). This is consistent with the hypothesis that CD39$^+$CLA$^+$ cells are enriched for shared antigens such MCPyV-derived antigens, whereas CD39$^+$CLA$^-$CD103$^+$ might be more enriched for T cells that recognize a range of private or patient-specific antigens. Additionally, we constructed sequence similarity networks from TCRs derived from both the circulating CD39$^+$CLA$^+$ CD8 T cells and tumor samples to explore shared CDR3 motifs present within and across individuals (Figure S4F). Highly similar clusters of TCRs found across individuals can arise due to convergent selection for shared epitope specificity for shared viral antigens. Altogether, TCR sequence similarity analyses revealing large public clusters of TCRs between blood and tumor samples are consistent with the enrichment of T cells with viral or tumor specificity among CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ CD8 T cell populations.

Differential transcriptional profiles of exhausted T cells categorized by CLA and CD103 expression

Next, to gain insights into the roles of blood-derived T cell populations with opposing predictive functions, we sought to perform a more detailed comparison of the phenotype and gene expression of profiles of CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ CD8 T cells. We started by identifying differentially expressed protein markers in the mass cytometry data between these populations that delineated a response. By examining the ratio of median differential protein expression of these cells (CD39$^+$CLA$^+$/CD39$^+$CLA$^-$CD103$^+$) prior to treatment, we observed that the frequencies of CD71 expression on CD39$^+$CLA$^+$ cells were higher in CR patients compared to PR or SD/PD patients. Conversely, frequencies of TIM3 expression were higher in CD39$^+$CLA$^-$CD103$^+$ cells among patients with PR or SD/PD responses (Figure S2K).

To assess the functional states of these cells more comprehensively, we performed transcriptomic profiling on the same sorted populations described for the TCR sequencing analysis (Figure 4A). As expected, sorted populations with CD39 expression showed higher levels of *ENTPD1*, which encodes CD39 (Figure 5A), and CD39$^+$CLA$^-$CD103$^+$ cells expressed higher levels of *ITGAE*, which encodes the αE chain of CD103. CD39$^+$CLA$^+$ cells expressed slightly higher levels of *SELPLG*, which encodes one of the key components, PSGL-1, of CLA, as well as *ZNF683*, which encodes the tissue-resident T cell transcription factor HOBIT (Figure S5A).[35,36]

Pairwise comparison of defined differentially expressed genes (DEGs) and subset-specific transcriptional signatures was performed and summarized using principal-component analysis (PCA) to highlight a distinct transcriptional program of CD39$^+$CLA$^-$CD103$^+$ cells compared to other subpopulations (Figure 5B). CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ cell subsets shared upregulated genes (DEG^{UP}) with each other but fewer with CD39$^-$ and CD39$^+$CLA$^-$CD103$^-$ subsets (Figure 5C). For example, ∼40% of DEGs from CD39$^+$CLA$^+$ cells were also upregulated in CD39$^+$CLA$^-$CD103$^+$ cells, whereas less than 10% of these genes were shared by CD39$^-$ and CD39$^+$CLA$^-$CD103$^-$ subsets. This suggests that CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ cells have a similar transcriptional program compared to other subsets.

To analyze subset-specific biology, we clustered DEGs and performed Gene Ontology analysis. Among the 5 clusters of DEGs identified, cluster 1, which was most enriched within CD39$^+$ populations, contained canonical exhaustion signatures such as *TOX* and *CXCL13* (Figure 5D). In addition, Gene Ontology (GO) biological pathway analysis revealed lymphocyte activation/differentiation and cell adhesion features in cluster 1 (Figure 5E). Clusters 2 and 3 contained genes that were preferentially upregulated in CD39$^+$CLA$^-$CD103$^+$ and were enriched for genes related to cell differentiation, activation, and apoptosis, which is a signature of T cell dysfunction. Cluster 4 captured genes that were downregulated in CD39$^+$CLA$^-$CD103$^+$ and included stemness/effector signature genes and cytokines. Lastly, cluster 5 contained genes at low levels across all CD39$^+$ cell populations, which included metabolic-related pathways (Figures 5D and 5E).

Overall, when comparing the transcriptional profiles of CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ subpopulations, we found they both exhibited signs of exhaustion, which is in line with CD39 being a marker of terminal exhaustion (Figures S5B and S5C).[29] The exhaustion scores of CD39$^+$CLA$^+$ and

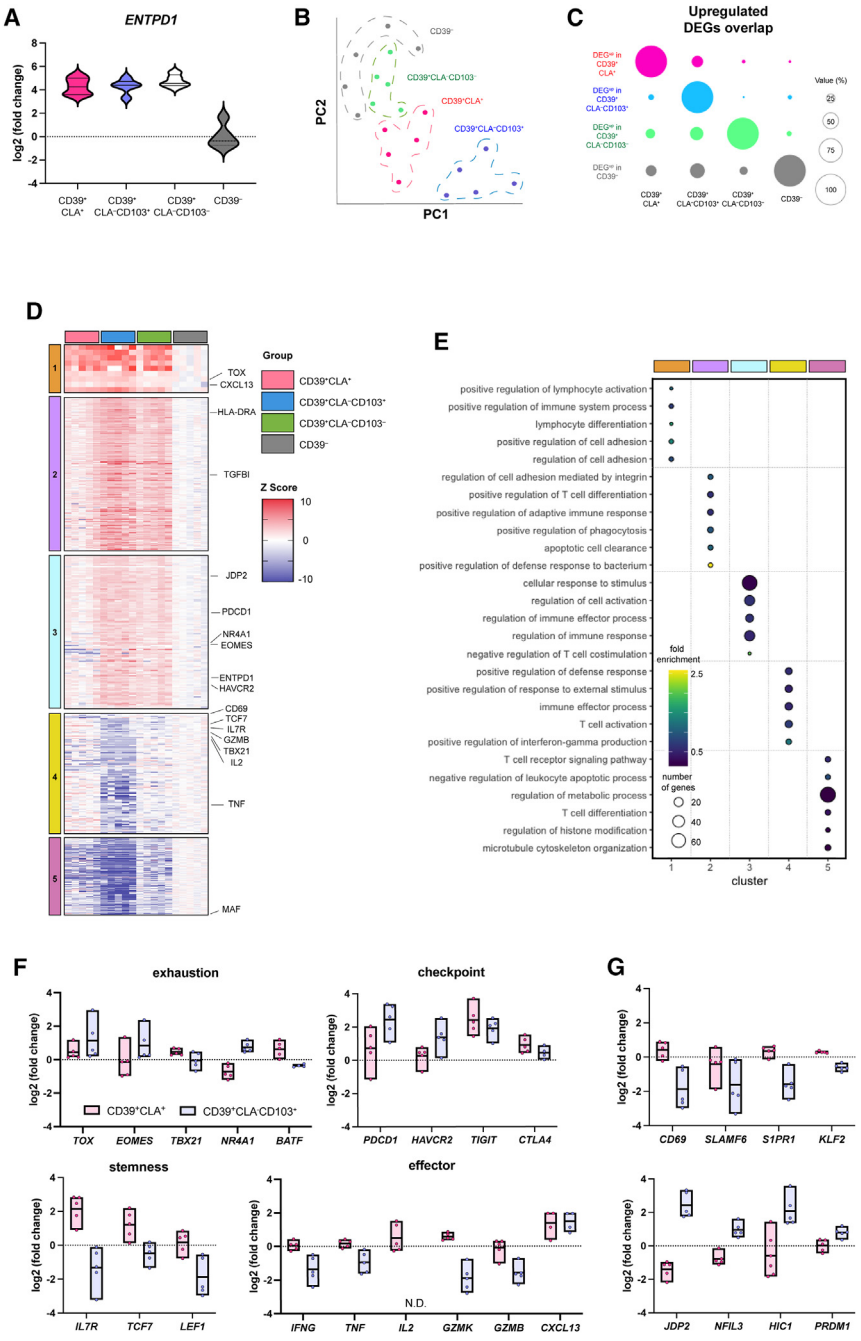


Figure 5. Distinct transcriptional signatures of circulating CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ CD8 T cells

(A) Expression of *ENTPD1* in each CD8 T cell population, normalized to those of CD39$^-$.

(B) PCA projection of sorted CD39$^+$CLA$^+$, CD39$^+$CLA$^-$CD103$^+$, CD39$^+$CLA$^-$CD103$^-$, and CD39$^-$ CD8 T cells from each patient. Each symbol represents an individual patient.

(C) DEG^{UP} overlaps between indicated populations. Bubble size represents the proportion of DEGs^{UP} from each population (y axis) also found to be upregulated in indicated subsets (x axis).

(D) Heatmap illustrating all DEGs (2-fold change, $p < 0.05$, and false discovery rate [FDR] < 0.1) clustered by using Pearson's correlation matrix. Color legend indicates *Z* scores.

(E) Pathway enrichment analysis of significantly enriched pathways by log(q value) for each cluster defined in Figure 4D. PCA, principal-component analysis; DEG, differentially expressed gene.

expressed lower levels of effector cytokine production (*IFNG*, *TNF*, *IL2*, *GZMK*, *GZMB*) and stemness markers (*IL7R*, *TCF7*, *LEF1*). CD39$^+$CLA$^-$CD103$^+$ cells also upregulate the transforming growth factor β (TGF-β) signaling pathway, which leads to the CD39$^+$CLA$^-$CD103$^+$ expression in T cells needed for epithelial homing (Figure S5F).[41] Since CD39$^+$CLA$^-$CD103$^+$ cells exhibited both tissue-resident markers (CD103 and *ZNF683*) and exhaustion markers (*TOX*, *NR4A1*, *LAYN*, *CXCL13*, *HAVCR2*, and *TIGIT*), we hypothesize that CD39$^+$CLA$^-$CD103$^+$ cells are more closely related to recirculating resident T and terminally exhausted cells.[22,36,42–46] A recent study has shown novel transcription factors that distinguish resident and terminally exhausted cells.[47] Expression of *JDP2* and *NFIL3*, which is only found in exhaustion program rather than resident T cells, was significantly higher in CD39$^+$CLA$^-$CD103$^+$ than CD39$^+$CLA$^+$, suggesting that CD39$^+$CLA$^-$CD103$^+$ cells were terminally exhausted rather than originating from typical tissue-resident memory T cells (Figure 5G). Of note, CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ cells were also higher than those of CD39$^-$ and CD39$^+$CLA$^-$CD103$^-$ subsets (Figures 5F and S5D). When contrasting these populations, CD39$^+$CLA$^+$ cells showed higher signatures of stemness and effector function compared to CD39$^+$CLA$^-$CD103$^+$ cells (Figures 5F and S5E).[37–39] Specifically, CD39$^+$CLA$^+$ cells expressed higher levels of *TIGIT*, *CTLA4*, and *TBX21* than CD39$^+$CLA$^-$CD103$^+$ cells, which are often associated with effector function. Additionally, these cells expressed higher levels of *IL7R* and *CCR7*, which are associated with naive or central memory CD8 T cells.[40] Conversely, CD39$^+$CD103$^+$ cells had comparable levels of *CXCL13*, which is known to be associated with tumor reactivity (Figure 5F). Altogether, these results suggest that CD39$^+$CLA$^+$ and CD39$^+$CLA$^-$CD103$^+$ CD8 T cell populations have distinct transcriptional profiles and functional characteristics.

DISCUSSION

Immune checkpoint therapy has shown significant advancements in the treatment of patients with cancer; nevertheless,

the impact of treatment on tumor-specific T cells and the identification of predictive biomarkers for patient response are lacking. In this study, we utilized a mass-cytometry-based multiplexed peptide-MHC class I-tetramer staining approach to analyze MCPyV-oncoprotein-specific T cells longitudinally in the blood of 24 patients treated with PD-1 blockade. Our findings revealed that patients with a higher frequency of baseline MCPyV-specific cells have better clinical outcomes and that the frequency of MCPyV-specific CD8 T cells decreased during pembrolizumab treatment in complete responders. Additionally, we identified two distinct populations of tumor-specific cells that correlated with clinical outcomes: $CD39^+CLA^+$ cells, which positively correlated with MCPyV-specific cell frequencies and the magnitude of clinical outcome, and $CD39^+CD103^+$ CD8 T cells, which positively correlated with the baseline tumor burden and negatively correlated with the magnitude of clinical outcome. These findings suggest that MCPyV-specific cells, $CD39^+CLA^+$ cells, and $CD39^+CD103^+$ cells could serve as blood-based biomarkers for predicting response to immunotherapy and as targets for novel cellular therapeutic strategies in patients with MCC. Here, we discuss the potential reasons behind the correlation of these populations with patient response to therapy and their implications for the further development of therapy.

Previous studies have demonstrated the presence of MCPyV-specific T cells in the tumor microenvironment of patients with MCC, which is associated with improved clinical outcomes.[9] However, studies of tumor-specific T cells in blood are limited due to difficulties in identifying cancer-specific T cells and the rarity of these cells in blood. Herein, we show that frequency of tumor-specific T cells in blood positively correlated to response to anti-PD-1 therapy. Of note, these results are supported by a parallel study that showed that the frequency of MCPyV-specific CD8 T cells in the blood correlated with pathological response in a clinical trial of neoadjuvant anti-PD-1 therapy.[48] We further show that anti-PD-1 therapy reduced the frequency of MCPyV-specific CD8 T cells in patients with complete response over the course of pembrolizumab but had only a marginal change in patients with PR. Phenotyping of these cells showed that although MCPyV-specific cells in the blood were enriched for various protein and gene expression profiles associated with terminal exhaustion, they exhibited less exhaustion than would be expected for tumor-infiltrating T cells and expressed genes associated with a maintained effector function. Consistent with our findings, other studies have shown that MCPyV-specific cells found within tumors are at late stages of exhaustion and do not strongly correlate with patient response to immunotherapy.[48,49] Therefore, prior to initiating PD-1 blockade, MCPyV-specific CD8 T cells in the blood could serve as a non-invasive biomarker for predicting patient outcomes, as has been reported in other cancer types.[50,51]

Our multimodal analysis provided insights into the complex nature of CD8 T cells in patients with MCC, particularly those with tumor-specific phenotypes and distinct co-stimulatory and inhibitory receptor expression patterns. By comparing these cells to bystander CD8 T cells, those recognizing common non-MCC-associated viral antigens (e.g., CMV, EBV, HSV, and influenza), we elucidated specific features unique to MCPyV-specific cells. Notably, MCPyV-specific cells exhibited high expression of the activation/exhaustion marker CD39, the skin-homing receptor CLA, and the tissue-recirculating marker CD103. This observation suggests a potential role for these cells in the local immune response within the skin, which is particularly relevant since MCC is a skin cancer and the initial encounter of T cells with the antigen is in skin-draining lymph nodes. Previous research has linked CLA-expressing tumor-infiltrating T cells with increased total T cell infiltration and highlighted the association of MCC tumors lacking CLA expression with immune evasion and poorer clinical outcomes. Our study found that MCPyV-specific T cells expressed high levels of CLA, and the frequency of $CD39^+CLA^+$ cells in the blood is associated with the frequency of MCPyV-specific cells and is predictive of favorable immunotherapy clinical outcomes.

The finding that the presence of $CD39^+CLA^+$ CD8 T cells in the blood is associated with favorable clinical outcomes suggests that these cells play a crucial role in the immune response against MCC. It is possible that these cells possess a higher degree of tumor reactivity and that they retain their functional capacity, allowing them to effectively target and eliminate tumor cells. The elevated expression of inhibitory receptors, such as PD-1, TIGIT, and CTLA-4, in $CD39^+CLA^+$ cells indicates that their activation and effector function may be regulated by immune checkpoint signaling, aligning with recent findings in patients with oral cancer.[52] Therefore, combination therapies targeting multiple inhibitory receptors could potentially enhance the anti-tumor immune response mediated by $CD39^+CLA^+$ cells.

Interestingly, we identified a population of $CD39^+CD103^+$ CD8 T cells with otherwise similar high-dimensional phenotypes to MCPyV-specific cells and $CD39^+CLA^+$ cells. The presence of these cells in the tumor have been reported in various solid tumors, such as non-small cell lung cancer, melanoma, breast cancer, and ovarian cancer.[19,20,53–56] Notably, these cells have been demonstrated to be enriched for tumor reactivity and are associated with improved survival and increased response to PD-1 blockade. Tracking TCRs of these cells within tumors has revealed dynamic changes in circulating T cells in response to immunotherapy, with a higher frequency of these TCRs in patients with disease progression.[57–59] We extended our analysis beyond cell specificity and quantification, providing a comprehensive characterization of $CD39^+CD103^+$ cells in the blood through CyTOF and bulk RNA sequencing. $CD39^+CD103^+$ cells in the blood exhibited characteristics of terminal exhaustion, including expression of PD-1, lack of cytokine production, and stemness. These cells are associated with baseline tumor burden and a worse clinical outcome. The expression of TIM3, an inhibitory receptor, further supports their exhausted phenotype. Strategies aimed at reinvigorating these exhausted cells, such as dual or triple immune checkpoint blockade, may be necessary to enhance their functionality and improve patient responses to immunotherapy.

Understanding the phenotypic and functional characteristics of tumor-specific T cells and their correlation with patient response to immunotherapy has important implications for the further development of therapy for MCC and other immunogenic cancers. Our findings suggest that measuring the frequency of MCPyV-specific cells, $CD39^+CLA^+$ cells, and $CD39^+CD103^+$ cells in the blood could serve as non-invasive biomarkers to

predict patient responses to immunotherapy. These biomarkers can aid in patient stratification and guide treatment decisions, facilitating personalized therapeutic approaches for MCC. Similar cell populations and CD8 T cell profiles might have comparable utility in other cancer types. Therefore, we think it could be worthwhile to monitor these or similar populations in other contexts. These cell populations could also be useful for rapid identification of patient-specific tumor-reactive T cells and TCRs that could be used for the design of novel cellular therapy strategies.[60]

Based on our findings, we propose that measuring CD8 T cells expressing both CD39 and CLA in the bloodstream prior to treatment can predict the frequency of MCPyV-specific cells and positive clinical response to pembrolizumab. CD39, known for its immunosuppressive role, has been suggested as a marker for T cell activation and exhaustion. Our study revealed that MCPyV-specific cells expressing CD39 and CLA associate with a favorable clinical outcome. CLA, which guides memory T cells back to the skin, has been associated with higher total T cell infiltration and better clinical outcomes in MCC.[61] In contrast, we also identified a $CD39^+CD103^+$ CD8 T cell population that is transcriptionally distinct and inversely correlated with patient response to pembrolizumab. CD103, a resident memory T cell marker that also mediates cell binding to epithelial and endothelial cells, has also been implicated as a marker for tumor reactivity in TILs[19,62,63] and in blood[64] in the context of other tumor types. The association between the frequencies of these cells and baseline tumor burden might partially explain their association with poor clinical outcomes. In addition, perhaps because of their more terminally exhausted profiles, these cells might be less receptive to disinhibition by PD-1 blockade. We believe this aligns well with a previous study showing that peripheral detection of TCR clonotypes corresponding to those of terminally exhausted tumor-infiltrating T cells was also associated with poor clinical outcomes after immunotherapy.[58]

TCR sequencing analysis provided insights into the antigen specificity of these T cell populations. Circulating CD8 T cells shared significant TCRβ sequences with tumor-infiltrating T cells, indicating dynamic movement of T cells between the bloodstream and the tumor. $CD39^+CLA^+$ cells showed a high degree of clonal sharing and were correlated with MCPyV-specific cells, suggesting their tracking of tumor-specific cells. $CD39^+CD103^+$ cells displayed lower clonal sharing among patients and were correlated with baseline tumor burden, suggesting that these cells might also be tumor reactive but possibly specific for private non-viral antigens such as mutation-derived neoantigens or other less ubiquitously expressed tumor-associated antigens. However, this would need to be tested in further studies.

In summary, our study highlights the significance of MCPyV-specific cells, $CD39^+CLA^+$ cells, and $CD39^+CD103^+$ cells as potential blood-based biomarkers for predicting patient responses to immunotherapy and as potential sources of tumor-specific T cells and TCRs in MCC. These subsets of tumor-specific cells exhibit distinct phenotypic and functional characteristics, which may contribute to their correlation with differential responses to immune checkpoint blockade. Understanding the implications of these populations in the context of immunotherapy response and their association with tumor burden and exhaustion status can guide the development of more effective and personalized treatment strategies for patients with MCC as well other cancers. Further investigation is warranted to validate these findings and explore the therapeutic potential of targeting these tumor-specific cell populations in MCC and other cancer types.

Limitations of the study

The study has some limitations that should be considered when interpreting the findings. Firstly, the study was based on a single trial and involved a relatively small number of patients, which may limit the generalizability of the results. An independent study (Pulliam, Jani et al., manuscript submitted) has demonstrated that the baseline MCPyV-specific cell frequencies can predict immunotherapy responses in a separate clinical trial. However, further validation of other predictive biomarkers is still required. Additionally, the study did not assess the functional activity of $CD39^+CLA^+$ or $CD39^+CD103^+$ CD8 T cells derived from blood against MCC tumors. Finally, as this study was cross-sectional in nature, it cannot establish causality or the direction of associations between the variables examined. Further mechanistic studies are required to investigate how $CD39^+CLA^+$ and $CD39^+CD103^+$ cells modulate the immune response to MCPyV by pembrolizumab treatment. These studies should also examine the detailed phenotypes and importance of MCPyV-specific CD4 or other immune subsets affected by anti-PD-1.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Patients and patient samples
- METHOD DETAILS
  - Epitope prediction and multiplexed peptide-MHC-I-tetramer preparation
  - Mass cytometry staining
  - Mass cytometry data analysis
  - Cell sorting and RNA isolation
  - Bulk TCR sequencing and data processing
  - Tumor T cell receptor sequencing
  - Bulk RNAseq and data processing
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrm.2023.101390.

ACKNOWLEDGMENTS

This project was supported by funding from a CCSG pilot award (P30 CA015704 to E.W.N.), a Fred Hutchinson Cancer Center IIRC post-doctoral

fellowship (to H.R.), a post-doctoral training fellowship from the National Research Foundation of Korea (2020R1A6A3A03037852 to H.R.), an NIH R01 grant (CA264646 to E.W.N.), an NIH N01 grant (75N93019C00063 to D.M.K.), a Kelsey Dickson Team Science Courage Research Award (to P.N., C.D.C., and S.J.), Merck & Co., and a NIH P01 grant (P01 CA225517 to P.N.). We thank all patients and participants in the CITN-09 clinical trial. We thank all other members of the Newell and Nghiem labs for their feedback and support for this project. We also thank Suzanne Topalian for critical feedback and contributions to the design of the pembrolizumab trial.

AUTHOR CONTRIBUTIONS

Conceptualization, H.R., T.M.B., T.H.P., S.J., S.R.H., D.M.K., P.N., and E.W.M.; methodology, H.R., T.M.B., T.H.P., and E.W.N.; formal analysis, H.R. and K.M.-B.; investigation, H.R., T.M.B., U.K.H., and K.S.; resources and data curation, C.D.C., N.R., and S.P.F.; writing – original draft, H.R.; writing – review & editing, H.R., T.M.B., T.H.P., K.S., C.D.C., N.K., K.M.-B., S.J., U.K.H., D.M.K., P.N., and E.W.N.; visualization, H.R. and K.M.-B.; supervision, S.R.H., D.M.K., P.N., and E.W.N.; project administration, H.R., T.H.P., C.D.C., N.R., S.P.F., P.N., and E.W.N.; funding acquisition, H.R., D.M.K., P.N., and E.W.N.

DECLARATION OF INTERESTS

E.W.N. is a co-founder, advisor, and shareholder for ImmunoScape Pte. Ltd., a scientific advisory board member and shareholder for Neogene Therapeutics, and a scientific advisory board member for Nanostring Biotechnologies and Trojan Biotechnologies. D.M.K. and P.N. are co-inventors on an institutionally owned patent concerning MCPyV-specific TCRs.

Received: July 3, 2023
Revised: November 29, 2023
Accepted: December 21, 2023
Published: February 9, 2024

REFERENCES

1. Becker, J.C., Stang, A., DeCaprio, J.A., Cerroni, L., Lebbé, C., Veness, M., and Nghiem, P. (2017). Merkel cell carcinoma. Nat. Rev. Dis. Primers *3*, 17077.
2. DeCaprio, J.A. (2021). Molecular Pathogenesis of Merkel Cell Carcinoma. Annu. Rev. Pathol. *16*, 69–91.
3. Bhatia, S., Afanasiev, O., and Nghiem, P. (2011). Immunobiology of Merkel Cell Carcinoma: Implications for Immunotherapy of a Polyomavirus-Associated Cancer. Curr. Oncol. Rep. *13*, 488–497.
4. Sihto, H., Böhling, T., Kavola, H., Koljonen, V., Salmi, M., Jalkanen, S., and Joensuu, H. (2012). Tumor Infiltrating Immune Cells and Outcome of Merkel Cell Carcinoma: A Population-Based Study. Clin. Cancer Res. *18*, 2872–2881.
5. Paulson, K.G., Iyer, J.G., Simonson, W.T., Blom, A., Thibodeau, R.M., Schmidt, M., Pietromonaco, S., Sokil, M., Warton, E.M., Asgari, M.M., and Nghiem, P. (2014). CD8+ Lymphocyte Intratumoral Infiltration as a Stage-Independent Predictor of Merkel Cell Carcinoma Survival: A Population-Based Study. Am. J. Clin. Pathol. *142*, 452–458.
6. Yarchoan, M., Hopkins, A., and Jaffee, E.M. (2017). Tumor Mutational Burden and Response Rate to PD-1 Inhibition. N. Engl. J. Med. *377*, 2500–2501.
7. Nghiem, P.T., Bhatia, S., Lipson, E.J., Kudchadkar, R.R., Miller, N.J., Annamalai, L., Berry, S., Chartash, E.K., Daud, A., Fling, S.P., et al. (2016). PD-1 Blockade with Pembrolizumab in Advanced Merkel-Cell Carcinoma. N. Engl. J. Med. *374*, 2542–2552.
8. Nghiem, P., Bhatia, S., Lipson, E.J., Sharfman, W.H., Kudchadkar, R.R., Brohl, A.S., Friedlander, P.A., Daud, A., Kluger, H.M., Reddy, S.A., et al. (2021). Three-year survival, correlates and salvage therapies in patients receiving first-line pembrolizumab for advanced Merkel cell carcinoma. J. Immunother. Cancer *9*, e002478.
9. Miller, N.J., Church, C.D., Dong, L., Crispin, D., Fitzgibbon, M.P., Lachance, K., Jing, L., Shinohara, M., Gavvovidis, I., Willimsky, G., et al. (2017). Tumor-Infiltrating Merkel Cell Polyomavirus-Specific T Cells Are Diverse and Associated with Improved Patient Survival. Cancer Immunol. Res. *5*, 137–147.
10. Iyer, J.G., Afanasiev, O.K., McClurkan, C., Paulson, K., Nagase, K., Jing, L., Marshak, J.O., Dong, L., Carter, J., Lai, I., et al. (2011). Merkel Cell Polyomavirus-Specific CD8+ and CD4+ T-cell Responses Identified in Merkel Cell Carcinomas and Blood. Clin. Cancer Res. *17*, 6671–6680.
11. Jing, L., Ott, M., Church, C.D., Kulikauskas, R.M., Ibrani, D., Iyer, J.G., Afanasiev, O.K., Colunga, A., Cook, M.M., Xie, H., et al. (2020). Prevalent and Diverse Intratumoral Oncoprotein-Specific CD8+ T Cells within Polyomavirus-Driven Merkel Cell Carcinomas. Cancer Immunol. Res. *8*, 648–659.
12. Miller, N.J., Church, C.D., Fling, S.P., Kulikauskas, R., Ramchurren, N., Shinohara, M.M., Kluger, H.M., Bhatia, S., Lundgren, L., Cheever, M.A., et al. (2018). Merkel cell polyomavirus-specific immune responses in patients with Merkel cell carcinoma receiving anti-PD-1 therapy. J. Immunother. Cancer *6*, 131.
13. Afanasiev, O.K., Yelistratova, L., Miller, N., Nagase, K., Paulson, K., Iyer, J.G., Ibrani, D., Koelle, D.M., and Nghiem, P. (2013). Merkel Polyomavirus-Specific T Cells Fluctuate with Merkel Cell Carcinoma Burden and Express Therapeutically Targetable PD-1 and Tim-3 Exhaustion Markers. Clin. Cancer Res. *19*, 5351–5360.
14. Samimi, M., Molet, L., Fleury, M., Laude, H., Carlotti, A., Gardair, C., Baudin, M., Gouguet, L., Maubec, E., Avenel-Audran, M., et al. (2016). Prognostic value of antibodies to Merkel cell polyomavirus T antigens and VP1 protein in patients with Merkel cell carcinoma. Br. J. Dermatol. *174*, 813–822.
15. Cheng, Y., Gunasegaran, B., Singh, H.D., Dutertre, C.-A., Loh, C.Y., Lim, J.Q., Crawford, J.C., Lee, H.K., Zhang, X., Lee, B., et al. (2021). Non-terminally exhausted tumor-resident memory HBV-specific T cell responses correlate with relapse-free survival in hepatocellular carcinoma. Immunity *54*, 1825–1840.e7.
16. Newell, E.W., Sigal, N., Bendall, S.C., Nolan, G.P., and Davis, M.M. (2012). Cytometry by Time-of-Flight Shows Combinatorial Cytokine Expression and Virus-Specific Cell Niches within a Continuum of CD8$^+$ T Cell Phenotypes. Immunity *36*, 142–152.
17. Newell, E.W., Sigal, N., Nair, N., Kidd, B.A., Greenberg, H.B., and Davis, M.M. (2013). Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. Nat. Biotechnol. *31*, 623–629.
18. Koelle, D.M., Liu, Z., McClurkan, C.M., Topp, M.S., Riddell, S.R., Pamer, E.G., Johnson, A.S., Wald, A., and Corey, L. (2002). Expression of cutaneous lymphocyte-associated antigen by CD8+ T cells specific for a skin-tropic virus. J. Clin. Invest. *110*, 537–548.
19. Duhen, T., Duhen, R., Montler, R., Moses, J., Moudgil, T., de Miranda, N.F., Goodall, C.P., Blair, T.C., Fox, B.A., McDermott, J.E., et al. (2018). Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. Nat. Commun. *9*, 2724.
20. Simoni, Y., Becht, E., Fehlings, M., Loh, C.Y., Koo, S.-L., Teng, K.W.W., Yeong, J.P.S., Nahar, R., Zhang, T., Kared, H., et al. (2018). Bystander CD8+ T cells are abundant and phenotypically distinct in human tumour infiltrates. Nature *557*, 575–579.
21. Hanada, K.-i., Zhao, C., Gil-Hoyos, R., Gartner, J.J., Chow-Parmer, C., Lowery, F.J., Krishna, S., Prickett, T.D., Kivitz, S., Parkhurst, M.R., et al. (2022). A phenotypic signature that identifies neoantigen-reactive T cells in fresh human lung cancers. Cancer Cell *40*, 479–493.e6.
22. Veatch, J.R., Lee, S.M., Shasha, C., Singhi, N., Szeto, J.L., Moshiri, A.S., Kim, T.S., Smythe, K., Kong, P., Fitzgibbon, M., et al. (2022). Neoantigen-specific CD4+ T cells in human melanoma have diverse differentiation

states and correlate with CD8+ T cell, macrophage, and B cell function. Cancer Cell *40*, 393–409.e9.

23. Wijeyesinghe, S., Beura, L.K., Pierson, M.J., Stolley, J.M., Adam, O.A., Ruscher, R., Steinert, E.M., Rosato, P.C., Vezys, V., and Masopust, D. (2021). Expansible residence decentralizes immune homeostasis. Nature *592*, 457–462.
24. Yenyuwadee, S., Sanchez-Trincado Lopez, J.L., Shah, R., Rosato, P.C., and Boussiotis, V.A. (2022). The evolving role of tissue-resident memory T cells in infections and cancer. Sci. Adv. *8*, eabo5871.
25. Peng, C., Huggins, M.A., Wanhainen, K.M., Knutson, T.P., Lu, H., Georgiev, H., Mittelsteadt, K.L., Jarjour, N.N., Wang, H., Hogquist, K.A., et al. (2022). Engagement of the costimulatory molecule ICOS in tissues promotes establishment of $CD8^+$ tissue-resident memory T cells. Immunity *55*, 98–114.e5.
26. Das, B.K., Kannan, A., Velasco, G.J., Kunika, M.D., Lambrecht, N., Nguyen, Q., Zhao, H., Wu, J., and Gao, L. (2023). Single-cell dissection of Merkel cell carcinoma heterogeneity unveils transcriptomic plasticity and therapeutic vulnerabilities. Cell Rep. Med. *4*, 101101.
27. Salmikangas, M., Laaksonen, M., Edgren, H., Salgado, M., Suoranta, A., Mattila, P., Koljonen, V., Böhling, T., and Sihto, H. (2023). Neurocan expression associates with better survival and viral positivity in Merkel cell carcinoma. PLoS One *18*, e0285524.
28. Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T.L., Agarwala, R., and Schäffer, A.A. (2008). Database indexing for production MegaBLAST searches. Bioinformatics *24*, 1757–1764.
29. Gupta, P.K., Godec, J., Wolski, D., Adland, E., Yates, K., Pauken, K.E., Cosgrove, C., Ledderose, C., Junger, W.G., Robson, S.C., et al. (2015). CD39 Expression Identifies Terminally Exhausted CD8+ T Cells. PLoS Pathog. *11*, e1005177.
30. Chiffelle, J., Genolet, R., Perez, M.A., Coukos, G., Zoete, V., and Harari, A. (2020). T-cell repertoire analysis and metrics of diversity and clonality. Curr. Opin. Biotechnol. *65*, 284–295.
31. Goncharov, M., Bagaev, D., Shcherbinin, D., Zvyagin, I., Bolotin, D., Thomas, P.G., Minervina, A.A., Pogorelyy, M.V., Ladell, K., McLaren, J.E., et al. (2022). VDJdb in the pandemic era: a compendium of T cell receptors specific for SARS-CoV-2. Nat. Methods *19*, 1017–1019.
32. Mayer-Blackwell, K., Schattgen, S., Cohen-Lavi, L., Crawford, J.C., Souquette, A., Gaevert, J.A., Hertz, T., Thomas, P.G., Bradley, P., and Fiore-Gartland, A. (2021). TCR meta-clonotypes for biomarker discovery with tcrdist3 enabled identification of public, HLA-restricted clusters of SARS-CoV-2 TCRs. Elife *10*, e68605.
33. Dash, P., Fiore-Gartland, A.J., Hertz, T., Wang, G.C., Sharma, S., Souquette, A., Crawford, J.C., Clemens, E.B., Nguyen, T.H.O., Kedzierska, K., et al. (2017). Quantifiable predictive features define epitope-specific T cell receptor repertoires. Nature *547*, 89–93.
34. Scheper, W., Kelderman, S., Fanchi, L.F., Linnemann, C., Bendle, G., de Rooij, M.A.J., Hirt, C., Mezzadra, R., Slagter, M., Dijkstra, K., et al. (2019). Low and variable tumor reactivity of the intratumoral TCR repertoire in human cancers. Nat. Med. *25*, 89–94.
35. Maverakis, E., Kim, K., Shimoda, M., Gershwin, M.E., Patel, F., Wilken, R., Raychaudhuri, S., Ruhaak, L.R., and Lebrilla, C.B. (2015). Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: A critical review. J. Autoimmun. *57*, 1–13.
36. Parga-Vidal, L., Behr, F.M., Kragten, N.A.M., Nota, B., Wesselink, T.H., Kavazović, I., Covill, L.E., Schuller, M.B.P., Bryceson, Y.T., Wensveen, F.M., et al. (2021). Hobit identifies tissue-resident memory T cell precursors that are regulated by Eomes. Sci. Immunol. *6*, eabg3533.
37. Jansen, C.S., Prokhnevska, N., Master, V.A., Sanda, M.G., Carlisle, J.W., Bilen, M.A., Cardenas, M., Wilkinson, S., Lake, R., Sowalsky, A.G., et al. (2019). An intra-tumoral niche maintains and differentiates stem-like CD8 T cells. Nature *576*, 465–470.
38. Galletti, G., De Simone, G., Mazza, E.M.C., Puccio, S., Mezzanotte, C., Bi, T.M., Davydov, A.N., Metsger, M., Scamardella, E., Alvisi, G., et al. (2020). Two subsets of stem-like CD8+ memory T cell progenitors with distinct fate commitments in humans. Nat. Immunol. *21*, 1552–1562.
39. Apetoh, L., Smyth, M.J., Drake, C.G., Abastado, J.-P., Apte, R.N., Ayyoub, M., Blay, J.-Y., Bonneville, M., Butterfield, L.H., Caignard, A., et al. (2015). Consensus nomenclature for CD8+ T cell phenotypes in cancer. OncoImmunology *4*, e998538.
40. van der Leun, A.M., Thommen, D.S., and Schumacher, T.N. (2020). CD8+ T cell states in human cancer: insights from single-cell analysis. Nat. Rev. Cancer *20*, 218–232.
41. El-Asady, R., Yuan, R., Liu, K., Wang, D., Gress, R.E., Lucas, P.J., Drachenberg, C.B., and Hadley, G.A. (2005). TGF-β–dependent CD103 expression by CD8+ T cells promotes selective destruction of the host intestinal epithelium during graft-versus-host disease. J. Exp. Med. *201*, 1647–1657.
42. Liu, X., Wang, Y., Lu, H., Li, J., Yan, X., Xiao, M., Hao, J., Alekseev, A., Khong, H., Chen, T., et al. (2019). Genome-wide analysis identifies NR4A1 as a key mediator of T cell dysfunction. Nature *567*, 525–529.
43. Chen, J., López-Moyado, I.F., Seo, H., Lio, C.-W.J., Hempleman, L.J., Sekiya, T., Yoshimura, A., Scott-Browne, J.P., and Rao, A. (2019). NR4A transcription factors limit CAR T cell function in solid tumours. Nature *567*, 530–534.
44. Seo, H., Chen, J., González-Avalos, E., Samaniego-Castruita, D., Das, A., Wang, Y.H., López-Moyado, I.F., Georges, R.O., Zhang, W., Onodera, A., et al. (2019). TOX and TOX2 transcription factors cooperate with NR4A transcription factors to impose CD8+ T cell exhaustion. Proc. Natl. Acad. Sci. USA *116*, 12410–12415.
45. Khan, O., Giles, J.R., McDonald, S., Manne, S., Ngiow, S.F., Patel, K.P., Werner, M.T., Huang, A.C., Alexander, K.A., Wu, J.E., et al. (2019). TOX transcriptionally and epigenetically programs CD8+ T cell exhaustion. Nature *571*, 211–218.
46. Scott, A.C., Dündar, F., Zumbo, P., Chandran, S.S., Klebanoff, C.A., Shakiba, M., Trivedi, P., Menocal, L., Appleby, H., Camara, S., et al. (2019). TOX is a critical regulator of tumour-specific T cell differentiation. Nature *571*, 270–274.
47. Chung, H.K., Cong, L., Eduardo, C., Brent, C., Bryan, M., Jun, W., Peixiang, H., Ming, S., Shixin, M., Qiyuan, Y., et al. (2023). Multiomics atlas-assisted discovery of transcription factors enables specific cell state programming. Preprint at bioRxiv, https://doi.org/10.1101/2023.01.03.522354.
48. Pulliam, T., Jani, S., Jing, L., Ryu, H., Jojic, A., Shasha, C., Zhang, J., Kulikauskas, R., Church, C., Garnett-Benson, C., et al. (2023). Circulating cancer-specific CD8 T cell frequency is associated with response to PD-1 blockade in Merkel cell carcinoma. Cell Rep Med *5* (*2*).
49. Lyngaa, R., Pedersen, N.W., Schrama, D., Thrue, C.A., Ibrani, D., Met, Ö., thor Straten, P., Nghiem, P., Becker, J.C., and Hadrup, S.R. (2014). T-cell Responses to Oncogenic Merkel Cell Polyomavirus Proteins Distinguish Patients with Merkel Cell Carcinoma from Healthy Donors. Clin. Cancer Res. *20*, 1768–1778.
50. Yamauchi, T., Hoki, T., Oba, T., Jain, V., Chen, H., Attwood, K., Battaglia, S., George, S., Chatta, G., Puzanov, I., et al. (2021). T-cell CX3CR1 expression as a dynamic blood-based biomarker of response to immune checkpoint inhibitors. Nat. Commun. *12*, 1402.
51. Fehlings, M., Jhunjhunwala, S., Kowanetz, M., O'Gorman, W.E., Hegde, P.S., Sumatoh, H., Lee, B.H., Nardin, A., Becht, E., Flynn, S., et al. (2019). Late-differentiated effector neoantigen-specific CD8+ T cells are enriched in peripheral blood of non-small cell lung carcinoma patients responding to atezolizumab treatment. J. Immunother. Cancer *7*, 249.
52. Luoma, A.M., Suo, S., Wang, Y., Gunasti, L., Porter, C.B.M., Nabilsi, N., Tadros, J., Ferretti, A.P., Liao, S., Gurer, C., et al. (2022). Tissue-resident memory and circulating T cells are early responders to pre-surgical cancer immunotherapy. Cell *185*, 2918–2935.e29.
53. Ganesan, A.-P., Clarke, J., Wood, O., Garrido-Martin, E.M., Chee, S.J., Mellows, T., Samaniego-Castruita, D., Singh, D., Seumois, G., Alzetani, A., et al. (2017). Tissue-resident memory features are linked to the

magnitude of cytotoxic T cell responses in human lung cancer. Nat. Immunol. *18*, 940–950.

54. Edwards, J., Wilmott, J.S., Madore, J., Gide, T.N., Quek, C., Tasker, A., Ferguson, A., Chen, J., Hewavisenti, R., Hersey, P., et al. (2018). CD103+ Tumor-Resident CD8+ T Cells Are Associated with Improved Survival in Immunotherapy-Naïve Melanoma Patients and Expand Significantly During Anti–PD-1 Treatment. Clin. Cancer Res. *24*, 3036–3045.
55. Park, M.H., Kwon, S.Y., Choi, J.E., Gong, G., and Bae, Y.K. (2020). Intratumoral CD103-positive tumour-infiltrating lymphocytes are associated with favourable prognosis in patients with triple-negative breast cancer. Histopathology *77*, 560–569.
56. Webb, J.R., Milne, K., Watson, P., deLeeuw, R.J., and Nelson, B.H. (2014). Tumor-Infiltrating Lymphocytes Expressing the Tissue Resident Memory Marker CD103 Are Associated with Increased Survival in High-Grade Serous Ovarian Cancer. Clin. Cancer Res. *20*, 434–444.
57. Caushi, J.X., Zhang, J., Ji, Z., Vaghasia, A., Zhang, B., Hsiue, E.H.-C., Mog, B.J., Hou, W., Justesen, S., Blosser, R., et al. (2021). Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. Nature *596*, 126–132.
58. Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D.T., Le, P.M., Forman, J., Huang, T., Li, S., Zhang, W., et al. (2021). Phenotype, specificity and avidity of antitumour CD8+ T cells in melanoma. Nature *596*, 119–125.
59. Rahim, M.K., Okholm, T.L.H., Jones, K.B., McCarthy, E.E., Liu, C.C., Yee, J.L., Tamaki, S.J., Marquez, D.M., Tenvooren, I., Wai, K., et al. (2023). Dynamic CD8+ T cell responses to cancer immunotherapy in human regional lymph nodes are disrupted in metastatic lymph nodes. Cell *186*, 1127–1143.e18.
60. Norberg, S.M., and Hinrichs, C.S. (2023). Engineered T cell therapy for viral and non-viral epithelial cancers. Cancer Cell *41*, 58–69.
61. Dowlatshahi, M., Huang, V., Gehad, A.E., Jiang, Y., Calarese, A., Teague, J.E., Dorosario, A.A., Cheng, J., Nghiem, P., Schanbacher, C.F., et al. (2013). Tumor-Specific T Cells in Human Merkel Cell Carcinomas: A Possible Role for Tregs and T-Cell Exhaustion in Reducing T-Cell Responses. J. Invest. Dermatol. *133*, 1879–1889.
62. Banchereau, R., Chitre, A.S., Scherl, A., Wu, T.D., Patil, N.S., de Almeida, P., Kadel Iii, E.E., Madireddi, S., Au-Yeung, A., Takahashi, C., et al. (2021). Intratumoral CD103+ CD8+ T cells predict response to PD-L1 blockade. J. Immunother. Cancer *9*, e002231.
63. Lee, Y.J., Kim, J.Y., Jeon, S.H., Nam, H., Jung, J.H., Jeon, M., Kim, E.-S., Bae, S.J., Ahn, J., Yoo, T.-K., et al. (2022). CD39+ tissue-resident memory CD8+ T cells with a clonal overlap across compartments mediate antitumor immunity in breast cancer. Sci. Immunol. *7*, eabn8390.
64. Wang, Z., Ahmed, S., Labib, M., Wang, H., Wu, L., Bavaghar-Zaeimi, F., Shokri, N., Blanco, S., Karim, S., Czarnecka-Kujawa, K., et al. (2023). Isolation of tumour-reactive lymphocytes from peripheral blood via microfluidic immunomagnetic cell sorting. Nat. Biomed. Eng. *7*, 1188–1203.
65. Wagih, O. (2017). ggseqlogo: a versatile R package for drawing sequence logos. Bioinformatics *33*, 3645–3647.
66. Paulson, K.G., Carter, J.J., Johnson, L.G., Cahill, K.W., Iyer, J.G., Schrama, D., Becker, J.C., Madeleine, M.M., Nghiem, P., and Galloway, D.A. (2010). Antibodies to Merkel Cell Polyomavirus T Antigen Oncoproteins Reflect Tumor Burden in Merkel Cell Carcinoma Patients. Cancer Res. *70*, 8388–8397.
67. Shuda, M., Arora, R., Kwun, H.J., Feng, H., Sarid, R., Fernández-Figueras, M.T., Tolstov, Y., Gjoerup, O., Mansukhani, M.M., Swerdlow, S.H., et al. (2009). Human Merkel cell polyomavirus infection I. MCV T antigen expression in Merkel cell carcinoma, lymphoid tissues and lymphoid tumors. Int. J. Cancer *125*, 1243–1249.
68. Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. *199*, 3360–3368.
69. Cheng, Y., Zhu, Y.O., Becht, E., Aw, P., Chen, J., Poidinger, M., de Sessions, P.F., Hibberd, M.L., Bertoletti, A., Lim, S.G., and Newell, E.W. (2019). Multifactorial heterogeneity of virus-specific T cells and association with the progression of human chronic hepatitis B infection. Sci. Immunol. *4*, eaau6905.
70. Hagberg, A.A., Schult, D.A., and Swart, P. (2008). Exploring Network Structure, Dynamics, and Function Using NetworkX.
71. Britanova, O.V., Shugay, M., Merzlyak, E.M., Staroverov, D.B., Putintseva, E.V., Turchaninova, M.A., Mamedov, I.Z., Pogorelyy, M.V., Bolotin, D.A., Izraelson, M., et al. (2016). Dynamics of Individual T Cell Repertoires: From Cord Blood to Centenarians. J. Immunol. *196*, 5005–5013.
72. Samokhina, M.P., Aleksandr, I., Nazarov, V., immunarch.bot, Rumynskiy, E., gracecodeadventures, Zarodniuk, M., and tsvvas. (2019). Immunarch: an R package for painless bioinformatics analysis of T-cell and B-cell immune repertoires. Zenodo *10*, 5281.
73. Robins, H.S., Campregher, P.V., Srivastava, S.K., Wacher, A., Turtle, C.J., Kahsai, O., Riddell, S.R., Warren, E.H., and Carlson, C.S. (2009). Comprehensive assessment of T-cell receptor β-chain diversity in αβ T cells. Blood *114*, 4099–4107.
74. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.
75. Zhou, G., Soufan, O., Ewald, J., Hancock, R.E.W., Basu, N., and Xia, J. (2019). NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. *47*, W234–W241.
76. Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L.-P., and Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci. *31*, 8–22.
77. Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A Greedy Algorithm for Aligning DNA Sequences. J. Comput. Biol. *7*, 203–214.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
7-AAD	BD Biosciences	Cat# 559925, RRID:AB_2869266
APC/Cyanine7 anti-human CD14 Antibody (clone: M5E2)	Biolegend	Cat# 301820, RRID:AB_493695
APC/Cyanine7 anti-human CD19 Antibody (clone: HIB19)	Biolegend	Cat# 302218, RRID:AB_314247
BUV395 Mouse Anti-Human CD3 (clone: UCHT1)	BD Biosciences	Cat# 563546, RRID:AB_2744387
BUV805 CD4 Mouse anti-Human (clone: RPA-T4)	BD Biosciences	Cat# 742000, RRID:AB_2871299
BUV737 CD8 Mouse anti-Human (clone: RPA-T8)	BD Biosciences	Cat# 612755, RRID:AB_2870086
Brilliant Violet 711 anti-human CD39 Antibody (clone: A1)	Biolegend	Cat# 328228, RRID:AB_2632894
CD103 (Integrin alpha E) Monoclonal Antibody (B-Ly7), APC	eBioscience	Cat# 17-1038-42, RRID:AB_10670631
PE anti-human/mouse Cutaneous Lymphocyte Antigen (CLA) Antibody (clone: HECA-452)	Biolegend	Cat# 321312, RRID:AB_2565589
89 - CD45 (clone: HI30)	Fluidigm	Cat# 3089003, RRID:AB_2661851
110 - CD4 (clone: RPA-T4)	Biolegend	Cat# 300502, RRID:AB_314070
111 - CD8a (clone: RPA-T8)	Biolegend	Cat# 301002, RRID:AB_314120
112 - CD14 (clone: TuK4)	Thermofisher	Cat# MHCD1400, RRID:AB_10371749
113 - CD19 (clone: HIB19)	Biolegend	Cat# 302202, RRID:AB_314232
114 - CD56 (clone: REA196)	Miltenyi Biotec	Cat# 130-108-016, RRID:AB_2658728
115 - CD57 (clone: HNK-1)	Biolegend	Cat# 359602, RRID:AB_2562403
116 - CD3 (clone: UCHT1)	Biolegend	Cat# 300402, RRID:AB_314056
141 - CLA (HECA-452)	Biolegend	Cat# 321302, RRID:AB_492894
142 - HLA-DR (clone: L243)	Biolegend	Cat# 307602, RRID:AB_314680
143 - ITB7 (clone: FIB504)	BD Biosciences	Cat# 555943, RRID:AB_396240
144 - TIGIT (clone: MBSA-43)	Thermofisher	Cat# 16-9500-82, RRID:AB_10718831
145 - Granzyme K (clone: GM6C3)	Santa cruz	Cat# SC-56125, RRID:AB_2263772
TCR gamma/delta Monoclonal Antibody (clone: 5A6.E9)	Thermofisher	Cat# MHGD04, RRID:AB_1502165
146 - anti-PE (clone: PE001)	Biolegend	Cat# 408102, RRID:AB_2168924
147 - CCR10 (clone: 314305)	R&D	Cat# MAB3478, RRID:AB_2275692
148 - CD45RO (clone: UCHL1)	Biolegend	Cat# 304202, RRID:AB_314418
149 - CD161 (clone: HP-3G10)	Biolegend	Cat# 339902, RRID:AB_1501090
150 - KLRG1 (clone: 13F2F12)	Thermofisher	Cat# 16-9488-85, RRID:AB_2637116
151 - CD27 (clone: O323)	Biolegend	Cat# 302802, RRID:AB_314294
152 - ICOS (clone: C398.4A)	Thermofisher	Cat# 14-9949-82, RRID:AB_468637
153 - CD103 (clone: B-Ly7)	Thermofisher	Cat# 14-1038-80, RRID:AB_467411
154 - Granzyme B (clone: 2C5/F5)	BD Biosciences	Cat# 550558, RRID:AB_393751
156 - CD25 (clone: M-A251)	Biolegend	Cat# 356102, RRID:AB_2561751
158 - CXCR3 (clone: G025H7)	Biolegend	Cat# 353702, RRID:AB_10983073
160 - PD-1 (clone: eBioJ105)	Thermofisher	Cat# 14-2799-80, RRID:AB_763476
162 - TIM3 (clone: F38-2E2)	Biolegend	Cat# 345002, RRID:AB_2116574
165 - CXCR5 (clone: RF8B2)	BD Biosciences	Cat# 552032, RRID:AB_394324
168 - CCR7 (clone: 150503)	R&D	Cat# MAB197, RRID:AB_2072803
169 - CD45RA (clone: HI100)	Biolegend	Cat# 304102, RRID:AB_314406
171 - CCR5 (clone: HEK/1/85a)	BioRAD	Cat# MCA2175GA, RRID:AB_324205
172 - CD39 (clone: A1)	Biolegend	Cat# 328202, RRID:AB_940438

(*Continued on next page*)

CellPress
OPEN ACCESS

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
175 - Perforin (clone: B-D48)	Fluidigm	Cat# 3175004B, RRID:AB_2895147
176 - CD38 (clone: HIT2)	Biolegend	Cat# 303502, RRID:AB_314354
209 - CD16 (clone: 3G8)	Fluidigm	Cat# 3209002B, RRID:AB_2756431
Chemicals, peptides, and recombinant proteins		
Merkel cell polyomavirus peptides	Mimotopes	Custom
Control viral peptides	Mimotopes	Custom
UV-cleavable MHC class I monomers	Fred Hutchinson Cancer Center	Custom
Streptavidin	Fred Hutchinson Cancer Center	Custom
Critical commercial assays		
NucleoSpin® RNA Plus XS	Takara Bio	740990
SMARTer® Human TCR a/b Profiling Kit v2	Takara Bio	634479
Illumina Truseq stranded mRNA library prep kit	Illumina	20020594
Deposited data		
Bulk TCR sequencing fastq	This paper	Zenodo: https://doi.org/10.5281/zenodo.8102089
Bulk RNA sequencing BAM	This paper	Zenodo: https://doi.org/10.5281/zenodo.8104398
Software and algorithms		
FlowJo	BD Biosciences	V9.9.4 and V10.8.0
R	R Foundation for Statistical Computing	V4.2.1
GraphPad Prism	GraphPad	V9
tcrdist3	Mayer-Blackwell et al.[32]	https://github.com/kmayerb/tcrdist3
ggseqlogo	Wagih et al.[65]	https://github.com/omarwagih/ggseqlogo

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and regents should be directed to and will be fulfilled by the lead contact, Evan W. Newell (enewell@fredhutch.org).

Materials availability

All unique and stable reagents generated in this study are available to the scientific community upon request and following execution of materials transfer agreement by contacting the lead contact.

Data and code availability

Bulk TCR sequencing, bulk RNA sequencing data discussed in this study have been deposited on Zenodo (Zenodo: https://doi.org/10.5281/zenodo.8102089, https://doi.org/10.5281/zenodo.8104398). No original code was generated in this study. Further data that support the finding of this study are available from the corresponding authors upon reasonable request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Patients and patient samples

All patients enrolled in this study provided written informed consent. Patients received pembrolizumab intravenously every 3 weeks at a dose of 2 mg/kg, for a maximum period of 2 years with radiologic assessment every 9 weeks. Investigators reported clinical responses based on CT scans per RECIST 1.1, as follows: complete response (CR), partial response (PR), stable disease (SD) or progressive disease (PD) based on imaging collected from time of enrollment to 08/01/2016. An initial response was confirmed by a serial CT scan showing the same result to be considered a confirmed response. Blood samples were drawn for correlative laboratory analyses. Peripheral blood mononuclear cells (PBMC) were cryopreserved after routine Ficoll preparation by a specimen processing facility at the Cancer Immunotherapy Trials Network. Tumor viral status was defined by expression of Large T-antigen within the tumor or by production of antibodies to small T-antigen as both are restricted to patients with MCPyV-positive tumors, as previously described.[66,67] All patients were HLA class I genotyped to determine eligibility for CD8 T cell specific MCPyV-tetramer screening

(Bloodworks Northwest, Seattle, WA). No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment. The samples were provided by Cancer Immunotherapy Trails Network (trial registration: ClinicalTrials.gov NCT02267603) and the analysis was performed according to the IRB file/approval number IR File #10686 (Table S1).

METHOD DETAILS

Epitope prediction and multiplexed peptide-MHC-I-tetramer preparation

Binding of epitopes of MCPyV LTA and STA proteins to HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*24:02, HLA-B*07:02 was predicted using the NetMHC4.0 algorithm.[68] Candidate epitopes (8–11 amino acids) with a predicted affinity <500 nM were selected and synthesized (Mimotopes). For MCPyV HLA-A*68:01, HLA-B*08:01, HLA-B*15:01, HLA-B*35:01, HLA-B*37:01, B*57:01, MHC-I easYmers were purchased from immuAware.[11] 34 common viral epitopes from HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*24:02, HLA-B*07:02, HLA-B*08:01, HLA-B*35:01 were chosen based on the previous literatures (Table S2).[11,15,20,69]

For multiplex MHC-I-tetramer staining, each tetramer was labeled with a combination of three metal-labelled streptavidin (SAV).[17] Using ten different metal-labeled SAV, 120 possible combinations were generated. Each specific combination was associated with a different peptide and HLA allele. Each metal-labeled SAV (50 μg/mL) was mixed for each combination using an automated pipetting device (TECAN). In a 96 well plate, peptide (1 mM) was added to corresponding HLA monomer (100 μg/mL), with a different peptide in each well. The plate was exposed to UV (365 nm) for 10 min for peptide exchange and left overnight at 4 °C. For the tetramerization, each peptide–MHC-I complex–metal-labelled SAV combination (50 μg/mL) was added in three steps according to the coding scheme. Then, tetramerized peptide-MHC-I complexes were incubated with free biotin for 10 min (10 μM) (Sigma). All different tetramers were combined and concentrated using a 50-kDa Amicon filter (Millipore). Before staining, the tetramer cocktail was filtered using a 0.1-μm filter (Millipore).

Mass cytometry staining

Purified antibodies lacking carrier proteins were purchased according to the provided in the Key resources table. Antibody conjugation was performed according to the protocol provided by Fluidigm. SAV was expressed, purified and heavy-metal conjugated as previously described.[16] HLA-A*01:01, HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-A*24:02 and HLA-B*07:02 monomer was refolded with appropriate UV-cleavable peptide and biotinylated as described. HLA-A*68:01, HLA-B*08:01, HLA-B*15:01, HLA-B*35:01, HLA-B*37:01, B*57:01 were purchased from immunAware. Tetramerization was performed with metal-labeled SAV as previously described. PBMCs were thawed and enriched for T cells by Pan T cell isolation kit (removing cells expressing CD14, CD15, CD16, CD19, CD34, CD36, CD56, CD123, and CD235a) using MACS LS columns (Miltenyi). Cells were then stained for cisplatin (5 mM) in PBS for 5 min on ice, washed and incubated for 1h at room temperature with tetramer cocktails in FACS buffer (PBS, 0.5% BSA, 0.02% sodium azide). Cells were then incubated with primary and secondary surface antibodies cocktails for 20min each on ice and fixed overnight in PFA 2%. The next day, cells were washed and permeabilized using permeabilization buffer (eBioscience) and stained with intracellular antibodies for 30 min at room temperature followed by DNA staining for 30 min at room temperature. Cells were washed three times with Cell Acquisition Solution (Fluidigm) and run on CyTOF (Helios, Fluidigm).

Mass cytometry data analysis

After CyTOF acquisition, which was performed as previously described,[20] any zero values were randomized using a uniform distribution of values between zero and −1 using an R script (as was the default operation of previous CyTOF software). Note also that all other integer values measured by the mass cytometer are randomized in a similar fashion by default. The signal of each parameter was then normalized based on the EQ beads (Fluidigm) as previously described. Cells were manually de-barcoded using FlowJo.

In line with the previously described method, we initially utilized an R script to identify tetramer positive cells.[69] In brief, we determined the thresholds for identifying tetramer positive events in CD8 T cells by manually gating out the SAV-metal channel in FlowJo. Based on the thresholds (Threshold X = Tx and Threshold Y = Ty) of every SAV-metal channel, the safety factors then automatically identify the tetramer positive population of CD8 T cells using the pre-set geometric criteria (Y/X Slope = k, X/Y Slope = k, and Width = w). The deconvolution algorithm excluded any tetramer positive cells that have less, or more than three SAV-metals coding. All events which were identified as tetramer positive cells underwent manual cross-checking of their three SAV-metal codes and HLA-allelic variants and then considered as validated antigen-specific CD8 T cells.

Samples were then used for UMAP analysis similar to that previously described using customized R scripts based on the 'flowCore' and 'uwot' R packages. In R, all data were transformed using the logicleTransform function (flowCore package) using parameters: w = 0.25, t = 16409, m = 4.5, a = 0 to roughly match scaling historically used in FlowJo. For heatmaps, median intensity corresponds to a logical data scale using the formula previously described. The colors in the heatmap represent the measured means intensity value of a given marker in each sample. A seven-color scale was used with black–blue indicating low expression values, green–yellow indicating intermediately expressed markers, and orange-red representing highly expressed markers. Samples were then used for UMAP analysis similar to that previously described[20] using customized R scripts based on the 'flowCore'

CellPress
OPEN ACCESS

and 'uwot' R packages. In R, all data were transformed using the logicleTransform function (flowCore package) using parameters: $w = 0.25$, $t = 16409$, $m = 4.5$, $a = 0$ to roughly match scaling historically used in FlowJo.

Cell sorting and RNA isolation

PBMCs were thawed and stained with Human TruStain FcX (Biolegend) for 10 min at room temperature followed by 20-min staining on ice with fluorescent antibody cocktails (Key resources table). CD8 T cells were isolated by FACS sorting (BD Bioscience) and subjected to RNA isolation (Takara). RNA purity was determined by assaying 1 μL of total RNA extract on a NanoDrop8000 spectrophotometer. Total RNA integrity was checked using an Agilent Technologies 2100 Bioanalyzer with an RNA Integrity Number (RIN) value.

Bulk TCR sequencing and data processing

RNA samples isolated from sorted CD8 T cell populations were subjected to TCR sequencing library preparation using SMARTer Human TCR a/b Profiling Kit v2 (Takara Bio USA). Briefly, after isolating mRNA from the sorted cells (NucleoSpin RNA Plus XS, Takara Bio USA), cDNA was synthesized and amplified following manufacturer's instructions. Quality of final libraries was assessed using Agilent 2200 TapeStation with High Sensitivity D5000 ScreenTape, quantified using a Qubit Fluorometer (ThermoFisher) and KAPA library quantification kit (Roche). The resulting libraries were sequenced on Illumina Miseq system (paired-end 300 bp). Raw reads of sequence data were aligned and assembled using Cogent NGS Immune Profiler Software (Takara Bio USA).

We computed TCR distances between all recovered TCRβ receptors using tcrdist3 software.[32,33] Retaining all distances $\leq$ 12 tcr distance units (corresponding roughly to an edit distance of 1–2), we visualized a similarity network using the Python package networkX.[70] We extracted the connected components within the sequence similarity network, and – for each connected component – we sampled a background set of TCRs from umbilical cord blood[71] with the same V- and J-gene usage to permit generation of background subtracted CDR3 motifs, which emphasize conserved non-germline encoded residues. We exported the aligned CDR3 motifs and used a custom R script (www.github.com/kmayerb/cdr3art/) to visualize each cluster CDR3 motif and gene usage using R packages ggseqlogo[65] (version 0.1.0) and ggsankey (https://github.com/davidsjoberg/ggsankey, version 0.0.9). True diversity and Morisita-Horn index were calculated using R package immunarch.[72] We used R 4.1.2.

Tumor T cell receptor sequencing

Pre-treatment formalin-fixed paraffin-embedded (FFPE) tumor biopsy material (20–25 μm thick molecular curls or material scraped from pre-cut slides) were submitted to Adaptive Biotechnologies for genomic DNA extraction of tissue, TCRβ sequencing and normalization as previously described.[73]

Bulk RNAseq and data processing

mRNA sequencing libraries were prepared according to the manufacturer's instructions (Illumina Truseq stranded mRNA library prep kit). mRNA was purified and fragmented from total RNA (1ug) using poly-T oligo-attached magnetic beads using two rounds of purification. Cleaved RNA Fragments primed with random hexamers were reverse transcribed into first strand cDNA using reverse transcriptase, random primers, dUTP in place of dTTP. (The incorporation of dUTP quenches the second strand during amplification, because the polymerase does not incorporate past this nucleotide.) These cDNA fragments then had the addition of a single 'A' base and subsequent ligation of the adapter. The products were purified and enriched with PCR to create the final strand specific cDNA library. The quality of the amplified libraries was verified by capillary electrophoresis (Bioanalyzer, Agilent). After qPCR using SYBR Green PCR Master Mix (Applied Biosystems), we combined libraries that index tagged in equimolar amounts in the pool and performed Nextseq 2000 sequencing system (Illumina) with 2 × 100bp read length. FASTQ files for each sample were mapped to the reference genome (human hg19) by STAR aligner.[74] The aligned results were further analyzed by NetworkAnalyst 3.0 to normalize data and report differentially expressed genes.[75] Genes with 2-fold change, p value < 0.05 and FDR<0.1 were selected for ontology analysis.[76] Data are visualized by R 4.1.2.

For LTA, STA, and viral capsid protein expression on tumor analysis, gene sequences of LTA, STA and VP1 were retrieved from the NCBI database (Gene ID: 10987417; 10987419; 10987416, respectively) and bulk RNA sequencing data of MCC cell lines and tumors were obtained from the SRA database.[26,27] To align and compare the expression patterns, we employed the MegaBLAST algorithm of the BLASTn suite from NCBI to identify highly similar sequences (Match/Mismatch Scores: 1, −2; Gap Costs: linear; suitable for 95% conserved) as previously reported.[28,77]

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistical tests were conducted as per the specifications mentioned in each figure. To determine the optimal threshold for classifiers, ROC analysis was employed, which was represented as the distance from the linear line. Pearson (r) or Spearman (ρ) correlation was used to determine linear concordance, and a two-sided t-test was used to see if the result was significantly different from zero. All statistical analysis was carried out using R 4.1.2 or Prism 9.