

Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma

Junfei Zhao^{1,2,13}, Andrew X. Chen^{1,13}, Robyn D. Gartrell³, Andrew M. Silverman³, Luis Aparicio^{1,2}, Tim Chu^{1,2}, Darius Bordbar³, David Shan³, Jorge Samanamud⁴, Aayushi Mahajan⁴, Ioan Filip¹, Rose Orenbuch¹, Morgan Goetz⁵, Jonathan T. Yamaguchi⁶, Michael Cloney⁶, Craig Horbinski^{6,7}, Rimas V. Lukas⁸, Jeffrey Raizer⁸, Ali I. Rae⁹, Jinzhou Yuan², Peter Canoll¹⁰, Jeffrey N. Bruce⁴, Yvonne M. Saenger¹¹, Peter Sims², Fabio M. Iwamoto^{12*}, Adam M. Sonabend^{6*} and Raul Rabadan^{1,2*}

Immune checkpoint inhibitors have been successful across several tumor types; however, their efficacy has been uncommon and unpredictable in glioblastomas (GBM), where <10% of patients show long-term responses. To understand the molecular determinants of immunotherapeutic response in GBM, we longitudinally profiled 66 patients, including 17 long-term responders, during standard therapy and after treatment with PD-1 inhibitors (nivolumab or pembrolizumab). Genomic and transcriptomic analysis revealed a significant enrichment of *PTEN* mutations associated with immunosuppressive expression signatures in non-responders, and an enrichment of *MAPK* pathway alterations (*PTPN11*, *BRAF*) in responders. Responsive tumors were also associated with branched patterns of evolution from the elimination of neoepitopes as well as with differences in T cell clonal diversity and tumor micro-environment profiles. Our study shows that clinical response to anti-PD-1 immunotherapy in GBM is associated with specific molecular alterations, immune expression signatures, and immune infiltration that reflect the tumor's clonal evolution during treatment.

Glioblastoma is the most common primary brain malignancy in adults. The current standard of care for newly diagnosed glioblastoma has limited efficacy, with a median overall survival of approximately 16–20 months¹. There is still no effective treatment for progressive or relapsed GBM, which invariably occurs in most patients. In the last decade, immunotherapy with checkpoint inhibitors has shown remarkable success in treating a variety of tumors, including advanced melanoma², non-small-cell lung cancer³, and Hodgkin's lymphoma⁴, among others. There has been considerable interest in utilizing immunotherapy in GBM, but a recent clinical trial of programmed cell death 1 (PD-1) immune checkpoint inhibitors in recurrent glioblastoma showed that only a small subset of patients (8%) demonstrated objective responses⁵. The mechanistic basis for the variation in response patterns remains to be explained.

Improved response to anti-PD-1 therapy has been found to be associated with higher mutational burdens in tumors across mul-

iple cancer types^{6,7} as well as with levels of T cell infiltration in the tumor microenvironment⁸. However, compared with melanomas or non-small-cell lung cancer, GBM harbors a lower burden of somatic mutations⁹ and a more immunosuppressive tumor microenvironment. One mechanism of immunosuppression is T cell exhaustion and apoptosis via PD-1 ligands (PD-L1/2) expressed by tumor cells: on binding PD-1 on the surface of cytotoxic T cells, the T cells become incapable of eliciting effective anti-tumor responses. PD-1 inhibitor therapy impairs this immune checkpoint and enhances the anti-tumor immune response⁸. Given the variable and unpredictable response of patients with GBM to PD-1 inhibitor therapies, we have extensively profiled 66 patients across a variety of timepoints, collecting DNA, RNA, tissue imaging, and clinical data (Fig. 1a). We sought to evaluate the genomic and stromal features associated with clinical outcomes and to gain insight into the underlying mechanisms of immunotherapy response.

We compiled a retrospective series of 66 adult patients with GBM who were treated with PD-1 inhibitors (pembrolizumab or nivolumab) on recurrence. Baseline patient characteristics and available data modalities of our cohort can be found in Supplementary Table 1 and Extended Data Fig. 1a.

Patients were classified as responders if they met at least one of the following two criteria:

- (1) Tissue sampled during surgery after PD-1 inhibitor therapy grossly showed only an inflammatory response and very few to no tumor cells (as associated with pseudo-progression).
- (2) Tumor volumes, as seen from magnetic resonance imaging (MRI), were either stable or shrinking continually over at least 6 months.

In Fig. 1b, we show brain MRIs of two patients treated with nivolumab with their corresponding relative timelines. Patient NU 7 showed progression after 2 months of nivolumab as measured by the RANO criteria¹⁰. Meanwhile, patient NU 11 showed stable disease without progression after 17 months.

Demographic and clinical characteristics, including response pattern, age at treatment initiation, gender, and choice of PD-1

¹Department of Systems Biology, Columbia University, New York, NY, USA. ²Department of Biomedical Informatics, Columbia University, New York, NY, USA. ³Department of Pediatrics, Pediatric Hematology/Oncology/SCT, Columbia University Irving Medical Center, New York, NY, USA. ⁴Department of Neurosurgery, Columbia University, New York, NY, USA. ⁵Department of Biomedical Engineering, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. ⁶Department of Neurological Surgery, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ⁷Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ⁸Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA. ⁹Department of Neurological Surgery, Oregon Health & Sciences University, Portland, OR, USA. ¹⁰Department of Pathology and Cell Biology, Columbia University, New York, NY, USA. ¹¹Department of Medicine, Hematology/Oncology, Columbia University Irving Medical Center, New York, NY, USA. ¹²Department of Neurology, College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, USA. ¹³These authors contributed equally: Junfei Zhao, Andrew X. Chen. *e-mail: fi2146@cumc.columbia.edu; Adam.Sonabend@nm.org; rr2579@cumc.columbia.edu

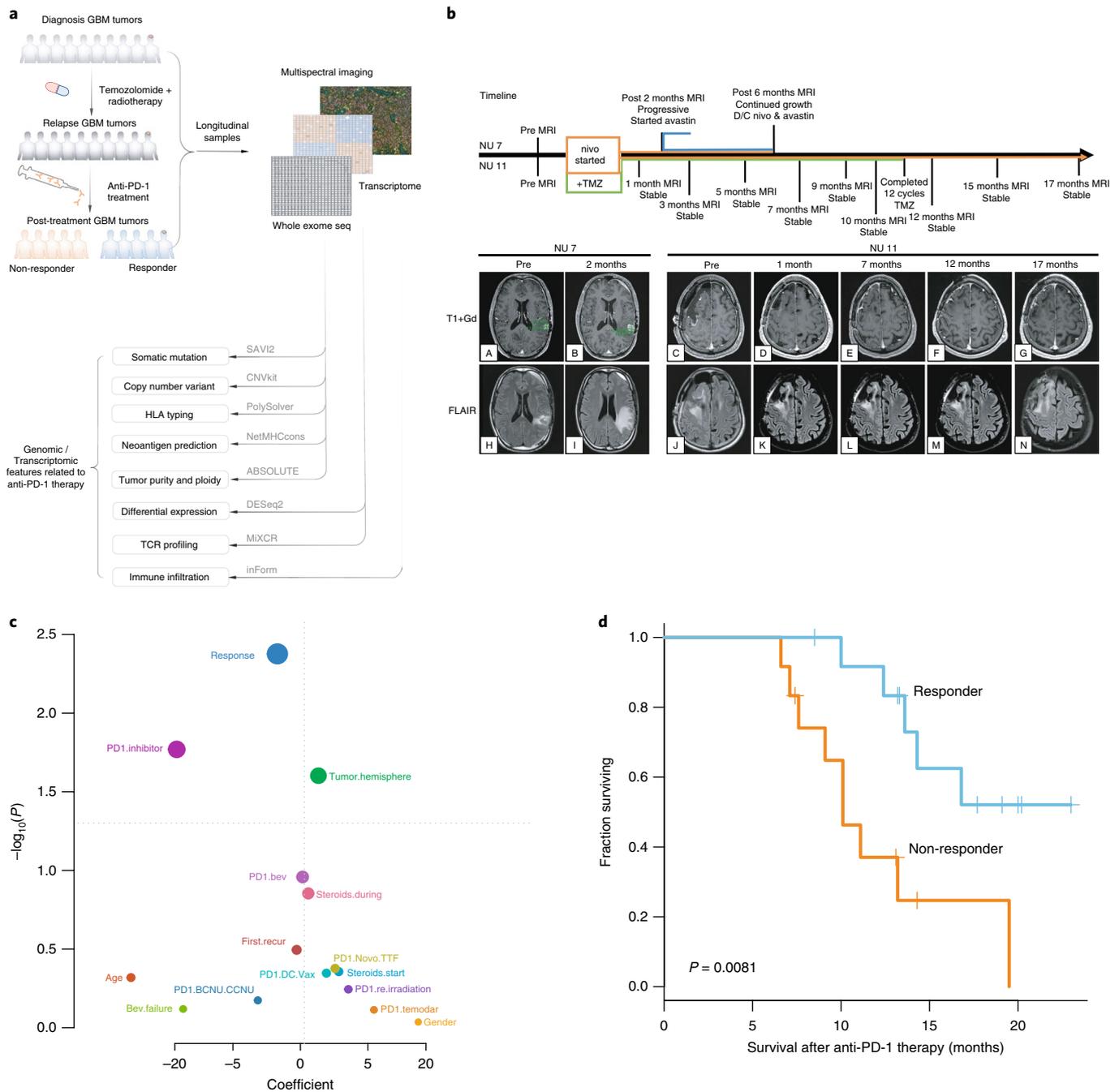


Fig. 1 | Analysis pipeline and clinical characteristics of the cohort. a, Sample collection and computational workflow. **b**, Brain MRIs of two patients treated with nivolumab, one of whom showed disease progression following 2 months of treatment (left, NU 7) while the other showed stable disease without progression after 17 months of treatment (right, NU 11). **c**, Univariate survival analysis revealed that only response to anti-PD-1 therapy is significantly correlated with overall survival of the patients ($n = 25$); P value, two-sided log-rank test. **d**, Kaplan-Meier curve comparing overall survival of patients who responded to anti-PD-1 therapy ($n = 13$) with those that did not respond ($n = 12$); P value, two-sided log-rank test.

inhibitor, were evaluated in a univariate survival analysis (Fig. 1c). Since criterion 2 of our definition of response contains a temporal component, to avoid potential confounding, we excluded from this analysis patients whose survival was less than 6 months. Response to the PD-1 inhibitor was found to be most significantly associated with overall survival as measured from the initiation of immunotherapy: patients who showed a responsive pattern to anti-PD-1 immunotherapy had a median survival of 14.3 months compared with the 10.1 months of non-responsive patients ($P = 0.0081$, log-rank test) (Fig. 1d). Without excluding patients, this effect was even

stronger: responders had a median survival of 15.5 months, whereas it was 5.7 months for non-responders ($P = 2.2 \times 10^{-5}$, log-rank test) (Extended Data Fig. 1b,d). Similarly, survival as measured from initial diagnosis was also increased in responders ($P = 1.6 \times 10^{-3}$, log-rank test; Extended Data Fig. 1c). However, there was no significant difference in the time span between initial diagnosis and the start of anti-PD-1 treatment between the two groups ($P = 0.96$, Wilcoxon rank-sum test).

We analyzed 58 whole exomes and 38 transcriptomes from longitudinal tumor-matched blood normal samples for 17 patients,

and also incorporated the results from a cancer gene panel of 39 patients (Fig. 2a).

We identified a median of 47 non-synonymous somatic mutations in the 33 tumors, with a range from 14 to 83, typical for GBM¹¹ (Supplementary Table 2). Contrary to previous observations in other tumor types^{6,7,12}, we did not find more non-synonymous single nucleotide variants (nsSNVs) in the responsive compared with the non-responsive baseline tumors (Extended Data Fig. 2). In fact, we observed a non-significant trend in the opposite direction; on the basis of the pre-treatment samples from the first surgery for each patient, non-responders had a median nsSNV count of 40, whereas it was 26 for responders ($P=0.11$, Wilcoxon rank-sum test). A statistically non-significant trend was also observed between response and aneuploidy ($P=0.88$, t -test; Extended Data Fig. 2)¹³. Similarly, human leukocyte antigen class I (HLA-I) neo-epitope load predictions yielded similar patterns for the two groups (median of 45 in non-responders and 35 in responders, $P=0.41$, Wilcoxon rank-sum test). Although a recent study has shown that zygosity at HLA-I genes influences survival of advanced melanoma and non-small-cell lung cancer patients treated with immunotherapies¹⁴, we did not find any significant association between HLA zygosity with immunotherapy response or survival. Additionally, there was no significant difference of tumor purity between these two groups (median of 0.41 in non-responders and 0.38 in responders, $P=0.19$, Wilcoxon rank-sum test).

We then sought to identify mutations (nsSNVs and indels) that were significantly enriched in either responsive or non-responsive tumors. In total, we identified 11 *IDH1* R132G/H mutated tumors, of which four were found in responders and seven in non-responders. Focusing on the remaining 45 *IDH1* wild type tumors, we found 23 *PTEN* mutations among the 32 non-responders, but only three mutations among the 13 responders (Fig. 2b,c). Within the cohort, *PTEN* was significantly more frequently mutated in the non-responsive tumors than in the responsive ones (Fisher $P=0.0063$, odds ratio=8.5, false discovery rate (FDR)-corrected $P<0.05$; Fig. 2b, right). Considering that the background *PTEN* mutation rate is around 33% (154 of 458 tumors in *IDH1* wild type glioblastomas from The Cancer Genome Atlas (TCGA)¹⁵), *PTEN* mutations were also more enriched in non-responders than expected (Fisher $P=0.0018$, odds ratio=3.3, FDR-corrected $P<0.05$; Fig. 2b, left; see Methods).

Notably, existing studies in melanoma have shown that *PTEN* loss in tumor cells increases the expression of immunosuppressive cytokines, resulting in decreased T cell infiltration in tumors and inhibited autophagy, which decreases T cell-mediated cell death¹⁶. Meanwhile, a study in glioblastoma has shown that tumor-specific T cells lysed *PTEN* wild type glioma cells more efficiently than those expressing mutant *PTEN*¹⁷. By utilizing single-sample gene-set enrichment analysis (ssGSEA) to calculate the enrichment score of the PI3K-AKT pathway for each tumor in our cohort, we also observed significantly higher PI3K-AKT pathway activity among *PTEN* mutant non-responsive tumors (t -test $P=0.049$; Extended Data Fig. 3a). However, we did not find any difference in *CD274* (which encodes PD-L1) RNA expression between responsive and non-responsive tumors (t -test $P=0.374$; Extended Data Fig. 3b).

We also found four mutations in the MAPK pathway components (including *BRAF* and *PTPN11*) among the 13 responders, and only one among the 32 non-responders (Fig. 2b). Considering the rarity of MAPK pathway mutations among *IDH1* wild-type glioblastoma (mutation rate 7.8%, 36 of 458 tumors from TCGA), MAPK pathway genes were significantly more frequently mutated in the responsive tumors than expected (Fisher $P=0.018$, odds ratio=5.1, FDR-corrected $P<0.05$). Similarly, MAPK pathway mutations are also significantly enriched in responders within our cohort (Fisher $P=0.019$, odds ratio=12.8, FDR-corrected $P<0.05$). Given the high prevalence of *BRAF* mutations in melanoma and the dramatic

success of immunotherapy in treating advanced melanoma, this finding may have relevant implications for the MAP kinase pathway and immune response¹⁸. Concordantly, the MAPK pathway was recently implicated in the modulation of T cell recognition of melanoma cells in a genome-wide CRISPR screen analysis¹⁹.

Recent studies in cancer immuno-editing have shown that the immune system selects for tumor variants with reduced immunogenicity, a phenomenon that could contribute to the immune-evasive features of gliomas²⁰. The pattern of initial response and later relapse among the responders in our cohort led us to investigate the evolution of tumors undergoing anti-PD-1 immunotherapy. The number of mutations exclusive to, or in common with, each sample was used to construct evolutionary trees for five patients (two non-responders and three responders) for whom we had both pre- and post-immunotherapy tumor samples (Fig. 2d).

We found that the tumors from non-responders and responders exhibited different patterns of evolution. The higher fraction of mutations exclusive to post-immunotherapy tumors, compared with the pre-anti-PD-1 treatment case, in the two non-responders (patients 20 and 53) suggests that they followed the classical linear model of tumor evolution. In contrast, the tumors from two responders (patients 55 and 71) were more similar to the branched model, with clonal alterations in the pre-anti-PD-1 dominant clone not present after therapy, suggesting that specific alterations and evolutionary patterns are associated with treatment (Fig. 2e). In the case of patient 55, we found three missense mutations (*MYPN* R409H, *UBQLN3* R159W, *CYP27B1* G194E) that were present before anti-PD-1 therapy (recurrent 1), but not after (recurrent 2). Interestingly, one of these mutations (*CYP27B1* G194E) is highly expressed (reads per kilobase per million mapped reads (RPKM) > 5) and predicted to result in immunogenic neoantigens. Similarly, for two other responders (patients 71 and 101), we found a missense mutation (*FNIP1* T409M and *TCF12* A605S, respectively) missing after immunotherapy, which is also highly expressed (RPKM > 5) and predicted to generate a neoantigen (Fig. 2f).

We also tracked the evolution of lymphocytes within the tumor by identifying T cell antigen receptor (TCR) and immunoglobulin RNA sequences. Across seven patients, we assessed the total number of relevant reads, and the clonal diversity via Shannon entropy, an information theory measure of randomness (Extended Data Fig. 4). We found that non-responders had a greater increase in clonal diversity among T cells compared with responders (Fig. 3a; $P=0.024$, exact Mann-Whitney U test). Likewise, the same effect was seen in the clonality of immunoglobulin reads, suggesting a similar response in B cells (Extended Data Fig. 5; $P=0.048$, exact Mann-Whitney U test).

Expression subtyping into proneural, mesenchymal, and classical subtypes²¹ did not result in any association with response (Extended Data Fig. 6). From a differential enrichment analysis across a total of 9,292 gene sets from MSigDB, we observed that before PD-1 inhibitor treatment, gene sets related to regulatory T (T_{reg}) cells were enriched among the top-ranked gene sets ($P=0.037$; Fig. 3b). Additionally, enrichment analysis showed that genes upregulated in T_{reg} cells were significantly more active in non-responsive tumors (Extended Data Fig. 7). FOXP3-expressing T_{reg} cells, which suppress aberrant immune response against self-antigens, have been shown to be negatively associated with clinical response to adaptive immunotherapy in human cancers²². Interestingly, following immunotherapy, gene sets related to immunosuppression were more active in responsive tumors, including FOXP3 and STAT3 signatures, as well as an immune evasion signature previously reported in renal cell carcinoma²³ (Fig. 3b).

Immunohistochemistry imaging of five tumors after anti-PD-1 treatment (two responders and three non-responders) did not identify CD4⁺FOXP3⁺ T_{reg} cells, suggesting that the aforementioned FOXP3 expression signature originated from another cell type. To investigate the origin of this immune signature, we

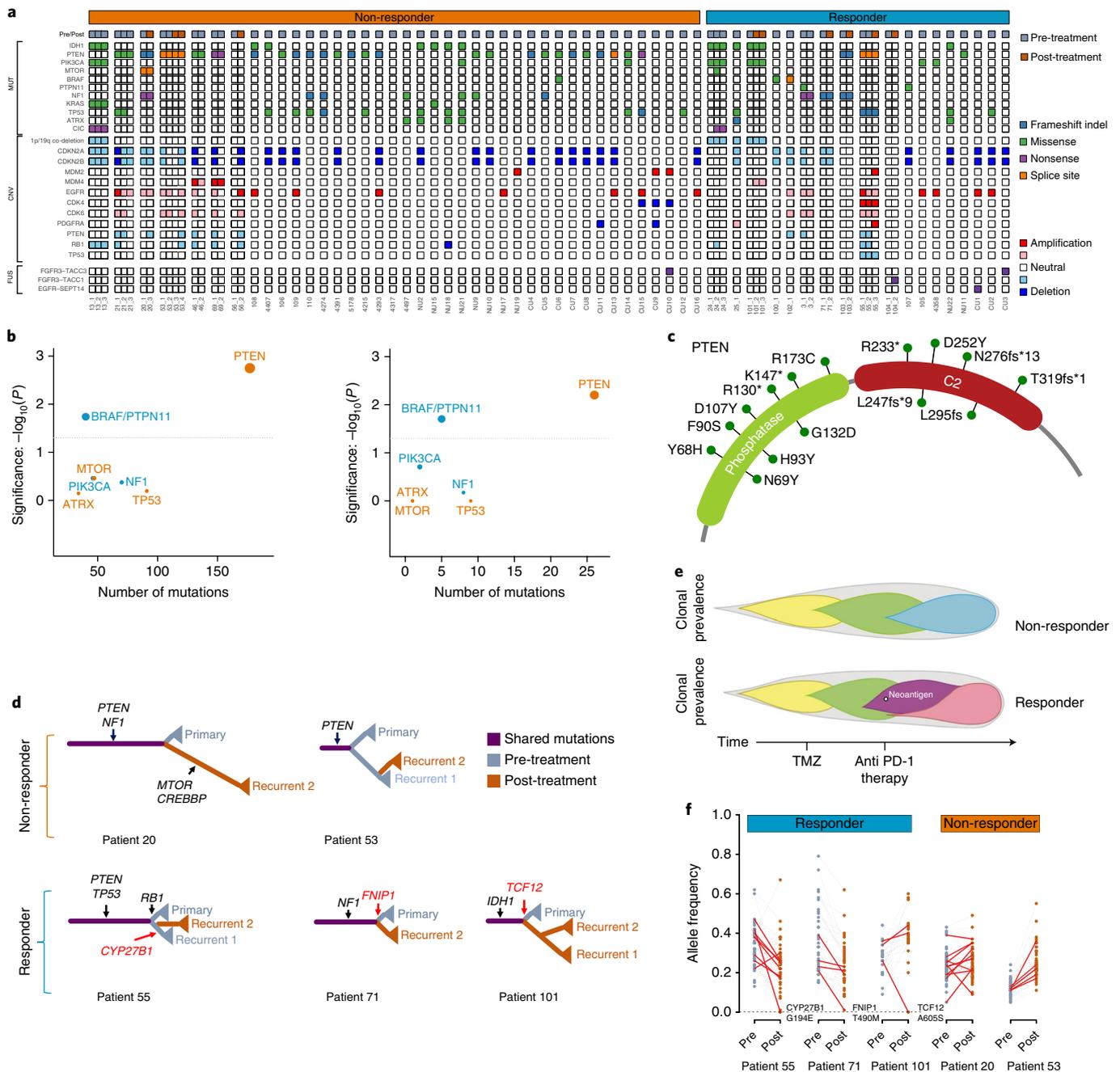


Fig. 2 | Mutational landscape, genomic correlates of response, and tumor evolution under anti-PD-1 therapy. **a**, Clinical and genetic profiles of the cohort. **b**, Enrichment of *BRAF/PTPN11* and *PTEN* mutations in tumors from responders and non-responders, respectively, compared with the TCGA-GBM background (left, $n=503$ patients) and within the cohort (right, $n=45$ patients); two-tailed Fisher's exact test, see Methods. **c**, Locations of identified mutations within the PTEN protein. **d**, Evolutionary trees of five patients (two non-responders and three responders) evaluated by whole-exome sequencing. Selected driver mutations are labeled in black. The variants that were eliminated after anti-PD-1 therapy and predicted to generate neoantigens are labeled in red. **e**, Different tumor evolution models characterize non-responders and responders. The upper panel represents non-responders following a linear pattern of evolution. The lower panel represents responders following a branching pattern of evolution, with the elimination of a clone possessing a neoantigen after anti-PD-1 therapy. **f**, Variant allele frequency of protein coding mutations before and after immunotherapy. Predicted expressed neoantigens are depicted in red.

studied the transcriptional profiles of 9,000 cells from three GBMs, including a *PTEN*-mutated tumor²⁴. Cells associated with the signature were enriched in a *PTEN*-mutated tumor ($P < 1 \times 10^{-16}$, Kolmogorov–Smirnov test; Extended Data Fig. 8), consistent with associations found in TCGA *PTEN*-mutated samples ($P < 1 \times 10^{-16}$, Kolmogorov–Smirnov test; Extended Data Fig. 8). Using topological data analysis²⁵, we identified three major cellular populations:

microglia, actively proliferating tumor cells (Ki-67⁺), and tumor cells with migrational markers (CD44⁺). Of these groups, the immunosuppressive signature was most associated with the CD44⁺ tumor subpopulation of the *PTEN*-mutated case ($P < 1 \times 10^{-16}$, t -test; Fig. 3c and Extended Data Fig. 9).

The observation that these immune signatures are upregulated in non-responsive tumors suggests that *PTEN* may play a role in

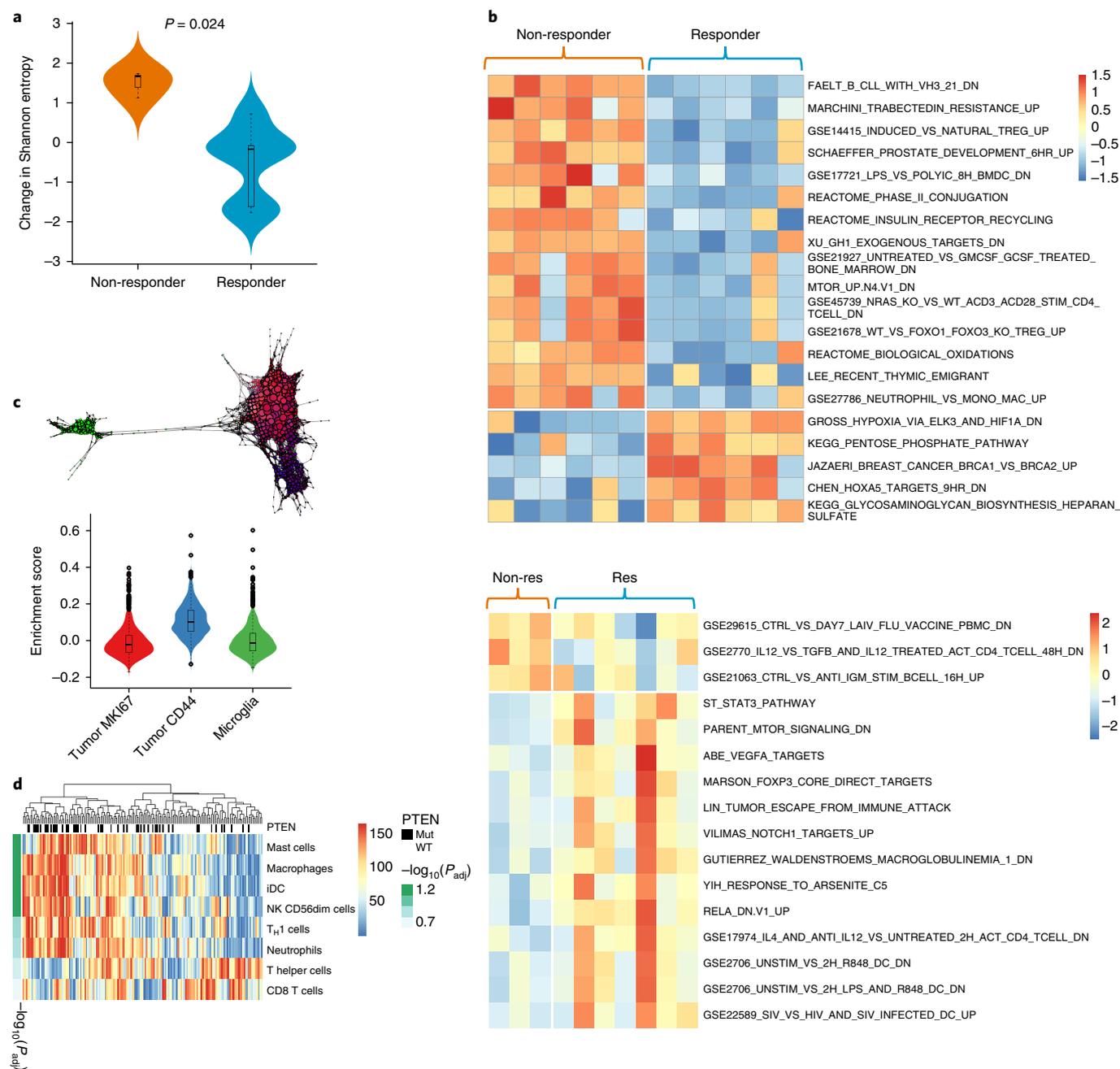


Fig. 3 | Transcriptomic signatures related to response to anti-PD-1 therapy. **a**, T cell clonal diversity before and after immunotherapy was assessed by identifying TCR RNA sequences within the tumor. Non-responders had a greater increase in Shannon entropy among T cells compared with responders ($P=0.024$, two-sided exact Mann-Whitney U test; $n=16$ independent timepoints from seven patients). **b**, Heatmap showing the top gene sets differentially enriched in responders versus non-responders before (top) and after (bottom) immunotherapy. **c**, Single-cell RNA-seq identifies a cluster of CD44-expressing tumor cells that are enriched in displaying an immunosuppressive profile ($n=4,000$ cells). **d**, Heatmap showing the associations between *PTEN* mutations and immune cell enrichment in TCGA (two-sided Wilcoxon rank-sum test; $n=167$ samples). All boxplots show the median, interquartile range, and whiskers (1.5 times interquartile range); violin plots represent sample distributions via kernel density estimation.

the formation of the tumor immune microenvironment. To further explore the potential immunological impact of *PTEN* mutations, we analyzed RNA-seq data from 172 samples from TCGA. We found that *PTEN* mutations are significantly correlated with the aforementioned *FOXP3*-related transcriptional signature and with lower tumor purity ($P=0.028$, Wilcoxon rank test; Extended Data Fig. 10). Then, ssGSEA was employed to measure the per sample infiltration levels of 24 immune cell types²⁶. Consistent with a previous report²⁷, correlation analysis revealed that *PTEN* mutations are

significantly associated with higher levels of macrophages, microglia, and neutrophils in the tumor microenvironment ($P<0.05$, Wilcoxon rank-sum test, FDR-corrected $P<0.1$; Fig. 3d). As the predominant immune cells infiltrating gliomas, tumor-associated macrophages have been shown to release a wide array of growth factors and cytokines that can facilitate tumor proliferation, survival, and migration²⁸.

Next, we investigated whether *PTEN* mutations were associated with the structure of the tumor microenvironment, which

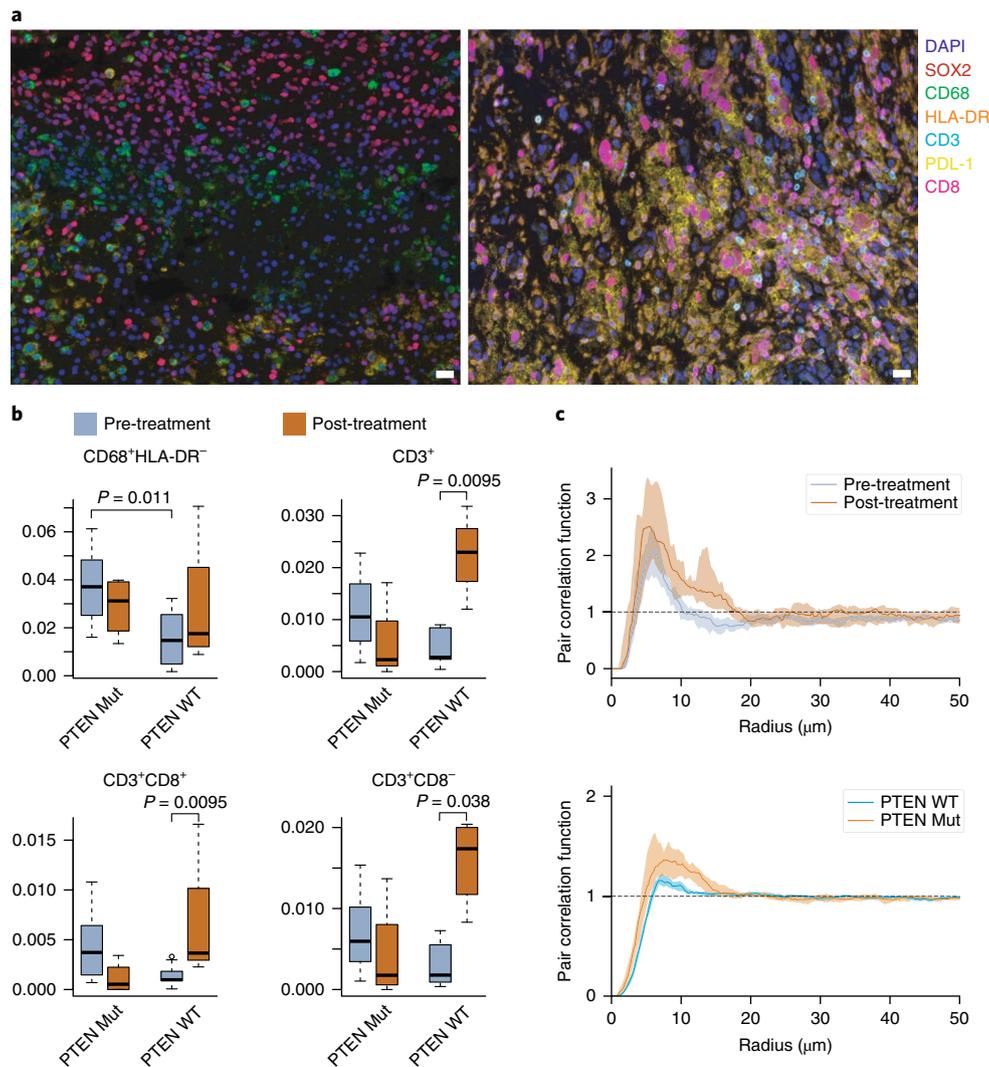


Fig. 4 | Tumor microenvironment profiling through qmIF. **a**, Representative multispectral images from pre-treatment samples showing DAPI (nuclei, blue), SOX2 (tumor, red), CD68 (microglia/macrophages, green), HLA-DR (activation marker, orange), CD3 (T cells, cyan), PD-L1 (immune suppression, yellow), and CD8 (cytotoxic T lymphocytes, magenta) in a non-responder (left) and a responder (right). The white bars represent 10 μm . A total of $n=337$ images were acquired. **b**, Cellular proportions for identified cell types are shown before and after immunotherapy, as a fraction of the total cell count ($n=17$ patients; P values, two-sided Wilcoxon rank-sum test). Boxplots show the median, interquartile range, and whiskers (1.5 times interquartile range). **c**, Pair correlation functions compare the degree of clustering of cells as a function of radius, for macrophages in *PTEN* wild-type patients (above, $n=126$ images) and for tumor cells before immunotherapy (below, $n=204$ images). Lines represent the pointwise median across samples; shaded regions represent 95% confidence intervals.

we explored with quantitative multiplex immunofluorescence (qmIF)^{29,30}. We stained and analyzed formalin-fixed paraffin-embedded (FFPE) specimens from 17 patients with matched pre- and post-anti-PD-1 treatment samples (7 non-responders, 10 responders; Fig. 4a). We found that *PTEN*-mutated tumors tended to have higher overall levels of CD68⁺ macrophage infiltration, although the difference did not reach statistical significance. Given that macrophages have heterogeneous roles, upon further categorization we observed in *PTEN*-mutated tumors a significantly higher density of CD68⁺HLA-DR⁻ macrophages ($P=0.011$, Wilcoxon rank-sum test; Fig. 4b), a subpopulation that indicates poor survival in melanoma²⁹. Finally, after immunotherapy, the density of CD3⁺ T cells in *PTEN* wild-type samples significantly increased compared with pre-treatment samples ($P=0.0095$, Wilcoxon rank-sum test; Fig. 4b), while the *PTEN*-mutated samples did not show this change. The same pattern was also observed in both CD3⁺CD8⁻ ($P=0.0095$,

Wilcoxon rank-sum test; Fig. 4b) and CD3⁺CD8⁺ T cells ($P=0.038$, Wilcoxon rank-sum test; Fig. 4b).

To assess the degree of clustering between cell types, we applied a technique from spatial statistics, the pair correlation function. We found that before immunotherapy, tumor cells clustered more strongly with each other in *PTEN*-mutated cases compared with *PTEN* wild type ($P=2.4 \times 10^{-4}$, Wilcoxon rank-sum test; Fig. 4c). Furthermore, in *PTEN* wild-type cases, macrophages became more strongly clustered with each other following treatment ($P=0.0012$, Wilcoxon rank-sum test; Fig. 4c); however, this effect was reversed in *PTEN* mutants ($P=0.032$).

In summary, we have confirmed that GBM patients who were responsive to anti-PD-1 immunotherapy (as evaluated through radiology and pathology) had significantly better overall survival after treatment. In our cohort, tumors from non-responders were significantly enriched for *PTEN* mutations; furthermore, RNA-seq

analysis indicated that these *PTEN* mutations may induce a distinct immunosuppressive microenvironment. Single-cell RNA profiling revealed that the source of this signature originates not from T_{reg} cells, but rather from tumor cells overexpressing CD44, a marker associated with cellular mobility and GBM aggressiveness³¹. Immunohistochemistry analysis confirmed the lack of increase of T cell infiltration in *PTEN* mutant tumors, concurring with transcriptomic signatures seen in pretreated cases from TCGA. We also identified differences in spatial structure of the tumor microenvironment that were associated with *PTEN* status; the increased clustering of tumor cells in *PTEN* mutants may impede immune infiltration. Similar results were observed in melanoma¹⁶ and in uterine leiomyosarcoma³², where *PTEN* loss was associated with reduced immune infiltration and resistance to anti-PD-1 therapy. Furthermore, the AKT–mTOR pathway downstream of *PTEN* has been implicated in both PD-L1 expression³³ and immune evasion in cancers³⁴. These in turn may determine the response pattern of patients with GBM to anti-PD-1 immunotherapy.

There is good evidence in the literature that alterations in the MAPK signaling pathway are implicated in the development of an unfavorable cancer immune phenotype³⁵. However, preclinical evidence suggests that MAPK pathway inhibition can dramatically increase the efficacy of immunotherapy³⁶. The observation in our cohort that *BRAF/PTPN11* mutations are enriched in tumors responsive to anti-PD-1 therapy supports the rationale for combining checkpoint inhibitors with MAPK-targeted therapy in multiple cancers^{35,37}. Furthermore, these genomic phenomena may be interconnected, as a study in prostate cancer has proposed a *PTEN-PTPN11* axis that is responsible for immunosuppression³⁸.

Another finding in our study is the distinct evolutionary patterns of responding and non-responding tumors under immunotherapy. Although we did not find higher somatic mutation and neoepitope loads in responding tumors, our analysis of their evolution provides strong evidence that the immune system plays an important role in the negative selection of clones containing immunogenic neoepitopes and thus promotes tumors in escaping immune surveillance. This also confirms previous observations from gastrointestinal cancers where low mutational loads did not preclude tumor infiltration by mutation-reactive, class I- and class II-restricted T cells³⁹. Additionally, non-responders were found to have a greater increase in T cell diversity following immunotherapy, suggesting that there was a failure of selective recruitment of lymphocytes into the tumor microenvironment. Supporting the role of tumor evolution in shaping the microenvironment, gene sets associated with immunosuppression were more active in non-responders before immunotherapy, but were more active in responders following treatment. These findings are consistent with the notion that tumors of non-responders possessed primary resistance to immunotherapy, whereas responders demonstrate a gradual acquisition of resistance following successful selection pressure⁴⁰.

In conclusion, our study identified multiple genomic features related to response to anti-PD-1 therapy in patients with GBM and depicted distinct evolutionary patterns of GBM under immunotherapy. While overall PD-1 inhibitors do not provide a survival benefit for patients with GBM, our study showed that a sub-group of patients might benefit from this therapy, suggesting a molecular, personalized approach for refining patient selection for immunotherapy. Although this approach requires further validation, it might provide a means for the effective application of therapy for glioblastoma.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at <https://doi.org/10.1038/s41591-019-0349-y>.

Received: 23 August 2018; Accepted: 8 January 2019;

Published online: 11 February 2019

References

- Gilbert, M. R. et al. A randomized trial of bevacizumab for newly diagnosed glioblastoma. *N. Engl. J. Med.* **370**, 699–708 (2014).
- Wolchok, J. D. et al. Nivolumab plus ipilimumab in advanced melanoma. *N. Engl. J. Med.* **369**, 122–133 (2013).
- Garon, E. B. et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N. Engl. J. Med.* **372**, 2018–2028 (2015).
- Ansell, S. M. et al. PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015).
- Filley, A. C., Henriquez, M. & Dey, M. Recurrent glioma clinical trial, CheckMate-143: the game is not over yet. *Oncotarget* **8**, 91779 (2017).
- Le, D. T. et al. PD-1 blockade in tumors with mismatch-repair deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).
- Rizvi, N. A. et al. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).
- Tumeh, P. C. et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568 (2014).
- Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. *Nature* **500**, 415 (2013).
- Wen, P. Y. et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. *J. Clin. Oncol.* **28**, 1963–1972 (2010).
- Wang, J. et al. Clonal evolution of glioblastoma under therapy. *Nat. Genet.* **48**, 768 (2016).
- Hugo, W. et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell* **165**, 35–44 (2016).
- Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* **355**, eaaf8399 (2017).
- Chowell, D. et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* **359**, 582–587 (2018).
- Brennan, C. W. et al. The somatic genomic landscape of glioblastoma. *Cell* **155**, 462–477 (2013).
- Peng, W. et al. Loss of PTEN promotes resistance to T cell-mediated immunotherapy. *Cancer Discov.* **6**, 202–216 (2016).
- Parsa, A. T. et al. Loss of tumor suppressor PTEN function increases B7-H1 expression and immunoresistance in glioma. *Nat. Med.* **13**, 84 (2007).
- Dong, C., Davis, R. J. & Flavell, R. A. MAP kinases in the immune response. *Annu. Rev. Immunol.* **20**, 55–72 (2002).
- Pan, D. et al. A major chromatin regulator determines resistance of tumor cells to T cell-mediated killing. *Science* **359**, 770–775 (2018).
- Arrieta, V. A. et al. The possibility of cancer immune editing in gliomas. A critical review. *Oncoimmunology* **7**, e1445458 (2018).
- Wang, Q. et al. Tumor evolution of glioma-intrinsic gene expression subtypes associates with immunological changes in the microenvironment. *Cancer Cell* **32**, 42–56. e46 (2017).
- Yao, X. et al. Levels of peripheral CD4⁺FoxP3⁺ regulatory T cells are negatively associated with clinical response to adoptive immunotherapy of human cancer. *Blood* **119**, 5688–5696 (2012).
- Lin, K.-Y. et al. Ectopic expression of vascular cell adhesion molecule-1 as a new mechanism for tumor immune evasion. *Cancer Res.* **67**, 1832–1841 (2007).
- Yuan, J. et al. Single-cell transcriptome analysis of lineage diversity and microenvironment in high-grade glioma. *Genome Med.* **10**, 57 (2018).
- Rizvi, A. H. et al. Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nat. Biotechnol.* **35**, 551–560 (2017).
- Senbabaoglu, Y. et al. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol.* **17**, 231 (2016).
- Zhang, C. et al. Tumor purity as an underlying key factor in glioma. *Clin. Cancer Res.* **23**, 6279–6291 (2017).
- Hussain, S. F. et al. The role of human glioma-infiltrating microglia/macrophages in mediating antitumor immune responses. *Neuro-oncology* **8**, 261–279 (2006).
- Gartrell, R. D. et al. Quantitative analysis of immune infiltrates in primary melanoma. *Cancer Immunol. Res.* **6**, 481–493 (2018).
- Stack, E. C., Wang, C., Roman, K. A. & Hoyt, C. C. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of Tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods* **70**, 46–58 (2014).
- Mooney, K. L. et al. The role of CD44 in glioblastoma multiforme. *J. Clin. Neurosci.* **34**, 1–5 (2016).

32. George, S. et al. Loss of PTEN is associated with resistance to anti-PD-1 checkpoint blockade therapy in metastatic uterine leiomyosarcoma. *Immunity* **46**, 197–204 (2017).
33. Lastwika, K. J. et al. Control of PD-L1 expression by oncogenic activation of the AKT–mTOR pathway in non-small cell lung cancer. *Cancer Res.* **76**, 227–238 (2016).
34. Noh, K. H. et al. Activation of Akt as a mechanism for tumor immune evasion. *Mol. Ther.* **17**, 439–447 (2009).
35. Bedognetti, D., Roelands, J., Decock, J., Wang, E. & Hendrickx, W. The MAPK hypothesis: immune-regulatory effects of MAPK-pathway genetic dysregulations and implications for breast cancer immunotherapy. *Emerg. Top. Life Sci.* **1**, 429–445 (2017).
36. Ebert, P. J. et al. MAP kinase inhibition promotes T cell and anti-tumor activity in combination with PD-L1 checkpoint blockade. *Immunity* **44**, 609–621 (2016).
37. Deken, M. A. et al. Targeting the MAPK and PI3K pathways in combination with PD1 blockade in melanoma. *Oncoimmunology* **5**, e1238557 (2016).
38. Toso, A. et al. Enhancing chemotherapy efficacy in Pten-deficient prostate tumors by activating the senescence-associated antitumor immunity. *Cell Rep.* **9**, 75–89 (2014).
39. Tran, E. et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* **350**, 1387–1390 (2015).
40. Sharma, P., Hu-Lieskovan, S., Wargo, J. A. & Ribas, A. Primary, adaptive, and acquired resistance to cancer immunotherapy. *Cell* **168**, 707–723 (2017).

Acknowledgements

This work has been funded by NIH grants R01 CA185486 (R.R.), R01 CA179044 (R.R.), U54 CA193313, U54 209997 (R.R.), R01 NS103473 (P.C., J.N.B., P.S.), NSF/SU2C/V-Foundation Ideas Lab Multidisciplinary Team (PHY-1545805) (R.R.), 2018 Stand Up To Cancer Phillip A. Sharp Innovation in Collaboration Awards (R.R.) and Keep Punching Foundation (F.M.I.). Funding support from Northwestern 5DP5OD021356-04 (A.M. Sonabend), P50CA221747 SPORE for Translational Approaches to Brain Cancer

(A.M.Sonabend, R.V.L., C.H.). Developmental funds from The Robert H Lurie NCI Cancer Center Support Grant no. P30CA060553 (A.M. Sonabend). A.X.C. is funded by the Medical Scientist Training Program (T32GM007367). R.D.G. is funded by CUIMC CTSA as TL1 Precision Medicine Fellow (1TL1TR001875-01) and Swim Across America.

Author contributions

J.Z. performed the majority of experiments and analyses. R.R., A.M. Sonabend, J.Z., and A.X.C. wrote the manuscript. A.I.R., J.T.Y., M.C., R.V.L., and J.R. compiled the clinical data for analysis. J.N.B., J.S., A.M., P.C., and C.H. procured and reviewed the tumor specimens for sequencing. J.Y. and P.S. provided the single-cell transcriptomic data. R.D.G., A.M. Silverman, D.B., D.S., and Y.M.S. provided the single-cell immunofluorescence. L.A. performed single-cell transcriptomic data analysis. I.F. and R.O. performed HLA genotyping. M.G. performed survival analysis. T.C. conducted figure design. R.R., A.M. Sonabend, and F.M.I. designed and supervised the entire project.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41591-019-0349-y>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41591-019-0349-y>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to F.M.I., A.M.S. or R.R.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2019

Methods

Patient selection. We collected a series of adult patients with GBM (treated with pembrolizumab or nivolumab on recurrence) from two institutions: Northwestern University ($n=20$) and Columbia University ($n=46$) with proper IRB approval at each institution. Informed consent was obtained from all patients, and this study complied with all relevant ethical regulations. All patients were treated with the standard therapy of temozolomide and radiation⁴¹ before the administration of PD-1 inhibitors. We excluded patients for whom there were no pre-immunotherapy specimens (either at diagnosis or after standard therapy recurrence) available. The baseline patient characteristics for our cohort are found in Supplementary Table 1, and the distribution of available data modalities across the cohort is found in Extended Data Fig. 1.

Sequencing and mapping. On average, 100-fold exome-wide target coverage was achieved for all of the sequenced tumor samples, and 60-fold for matched blood normal samples was achieved. High-quality reads for these samples were mapped by BWA⁴² to the hg19 human genome assembly with default parameters. All mapped reads were then marked for duplicates by Picard to eliminate potential duplications.

Somatic mutations. To identify somatic mutations from whole-exome sequencing data for samples from patients with GBM, we applied the variant-calling software SAVI2 (statistical algorithm for variant frequency identification⁴³), which is based on an empirical Bayesian method. Specifically, we first generated a list of variant candidates by successively eliminating positions without variant reads, positions with low sequencing depth, positions that were biased for one strand, and positions that contained only low-quality reads. Then, the numbers of high-quality reads for forward-strand reference alleles, reverse-strand reference alleles, forward-strand non-reference alleles, and reverse-strand non-reference alleles were calculated at the remaining candidate positions to build the prior and the posterior distribution of the mutation allele fraction. Finally, somatic mutations were identified on the basis of the posterior distribution of differences in the mutation allele fraction between normal and tumor samples. SAVI2 was able to assess mutations by simultaneously considering multiple tumor samples, as well as their corresponding RNA samples, if available. Only variants with a mutant allele frequency of 5% or greater were included for further analysis.

Analysis of mutation frequencies. Our pipeline is capable of detecting intragenic exon deletions, whereas these alterations are not reported in TCGA. When comparing the rate of *PTEN* mutations between our cohort and TCGA, for the sake of egalitarianism, we have excluded the three calls of these intragenic exon deletions from our cohort.

Analysis of copy number changes. CNVkit⁴⁴ was used to detect copy number changes from whole-exome sequencing data.

Tumor purity estimation. In our cohort, we calculated tumor purity with ABSOLUTE⁴⁵. For TCGA, pre-computed values from ESTIMATE⁴⁶ were used.

Gene fusion detection. ChimeraScan⁴⁷ was used to generate the starting set of gene fusion candidates. To reduce the false positive rate and nominate potential driver events, we applied the Pegasus annotation and prediction pipeline. We reconstructed the entire fusion sequence on the basis of breakpoint coordinates and assigned a driver score to each candidate fusion via a machine learning model trained largely on GBM data⁴⁸.

Gene expression analysis. Paired-end transcriptome reads were processed using STAR⁴⁹ aligner on the basis of the Ensembl GRCh37 human genome assembly with default parameters. Normalized gene expression values were calculated by featureCounts⁵⁰ as RPKM. ssGSEA was performed using the R package GSVA⁵¹. We compared the transcriptomic profiles of the two tumor groups using ssGSEA on the basis of five collections of annotated gene sets from the Molecular Signature Database v6.0 (C2 curated gene sets, C4 computational gene sets, C6 oncogenic gene sets, and C7 immunologic gene sets)⁵². Then, differentially enriched gene sets between the responders and non-responders were defined by an effect size of GSVA score differences being greater than 0.8 and a *t*-test *P* value of less than 0.01.

HLA typing and neoantigen prediction. To test if HLA zygosity affects immunotherapy response or survival, we determined HLA genotypes of 17 patients for whom we had normal blood whole-exome sequencing data and verified using tumor RNA. HLA typing for each sample was performed on blood DNA using the POLYSOLVER algorithm⁵³. We furthermore validated the calls with tumor RNA, which achieved perfect concordance on samples where both were available. We used the pVAC-Seq⁵⁴ pipeline with the NetMHCcons⁵⁵ binding strength predictor to identify neoantigens. NetMHCcons integrates three state-of-the-art methods, NetMHC, NetMHCpan, and PickPocket, to give the most accurate predictions with consideration of the patient's HLA type. As required, we used the variant effect predictor from Ensembl to annotate variants for downstream processing by pVAC-Seq. For each single-residue missense alteration, major histocompatibility complex

binding affinity was predicted for all the wild type and mutant peptides of 8, 9, 10, and 11 amino acids in length. The mutant peptide with the strongest binding affinity was kept for further analysis.

Single-cell data analysis. Single-cell transcriptional profiles were obtained from 9,000 cells over three samples⁵⁴. GSEA was used to assess enrichment of transcriptomic signatures among the samples. Topological representations of cellular expression were constructed with Mapper (Ayasdi Inc.), outputting a network where nodes represent sets of cells with similar characteristics. RGB values were computed for each node in proportion to its composition of Ki-67⁺ tumor cells, microglia, and CD44⁺ tumor cells, respectively.

Tumor purity estimation and cellular fraction. ABSOLUTE was used to infer tumor purity and ploidy for each whole-exome sequencing sample by integrating mutational allele frequencies and copy number calls.

Lymphocyte clonality analysis. TCR and immunoglobulin RNA sequences were processed via MiXCR⁵⁶. This was performed for a total of seven patients for whom we had pre- and post-immunotherapy RNA-seq data of sufficient quality. Of these patients (two non-responders and five responders), one from each response criteria had two samples post-therapy (patients 53 and 101). Clonal diversity was calculated through Shannon entropy.

Quantitative multiplex immunofluorescence (qmIF) analysis. FFPE tumor samples were collected for each sample, and H&E slides were reviewed by a neuropathologist (P.C.) to confirm presence of tumor. Opal multiplex staining was performed on FFPE immunoblot slides for CD3 (T cells), CD8 (cytotoxic T lymphocytes), CD68 (microglia/macrophages), HLA-DR (activation marker), PD-L1 (immunosuppression marker), and SOX2 (tumor marker)^{45,46}. Images were acquired using Vectra (PerkinElmer) for whole-slide scanning, and multispectral images were acquired for all areas with at least 99% tissue, using inForm software (PerkinElmer) to unmix and remove autofluorescence. Multispectral images were analyzed using inForm software and R to evaluate density of immune phenotypes within the tumor microenvironment.

Spatial analysis. Phenotyped immunofluorescence data were processed into pair correlation functions (PCFs) using the spatstat R package⁵⁷. Inhomogeneous PCFs were calculated up to a radius of 50 μm for tumor and CD68⁺ cells, provided that there were a minimum of 20 cells of that type in the sample. The isotropic edge correction and a normalization power of 2 were used. The area under the curve for each PCF was used as a summary statistic for quantifying clustering, and plots represent the point-wise median PCF across samples with 95% confidence intervals obtained via bootstrapping.

Statistical analysis. All statistical analyses were conducted in R and Python. In all boxplots, the center lines represent the median, lower and upper box limits are, respectively, the first and third quartiles, and whiskers represent the maximal values up to 1.5 times the interquartile range. All values extending beyond this range are considered fliers/outliers. Violin plots use the Gaussian kernel to estimate densities. The two-sided Wilcoxon rank-sum (Mann–Whitney) test was generally used to compare two populations nonparametrically, unless we had prior knowledge that distributions were normal, in which case the two-sided *t*-test was used. The *P* values were adjusted for multiple comparisons using the Benjamini–Hochberg (FDR) procedure, and statistical significance was assessed at an adjusted *P* value threshold of 0.05.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Code availability

All the custom code will be made available upon request.

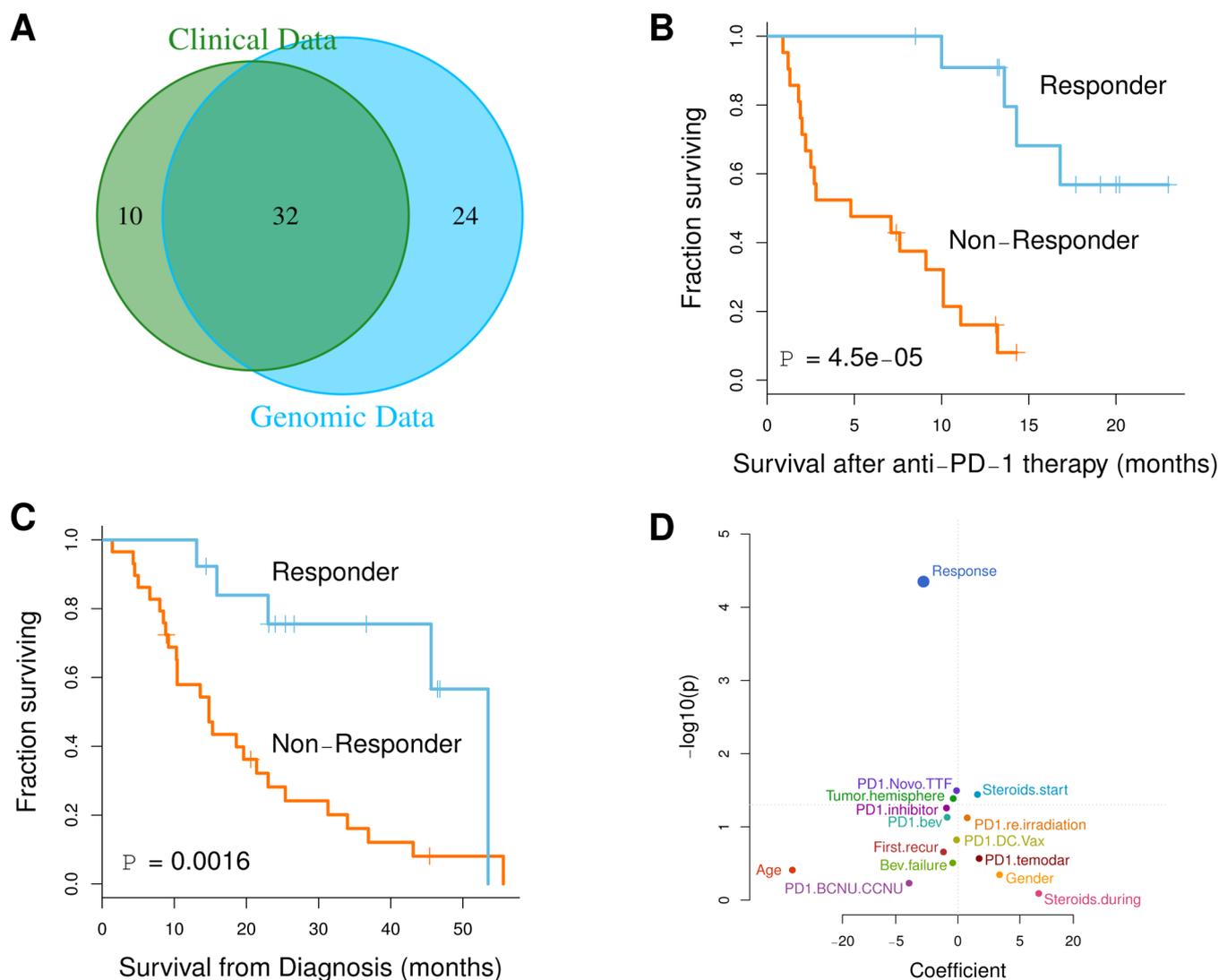
Data availability

All the sequencing data have been deposited in SRA [PRJNA482620](https://www.ncbi.nlm.nih.gov/sra/PRJNA482620). Processed data and basic association analyses will be made available upon request.

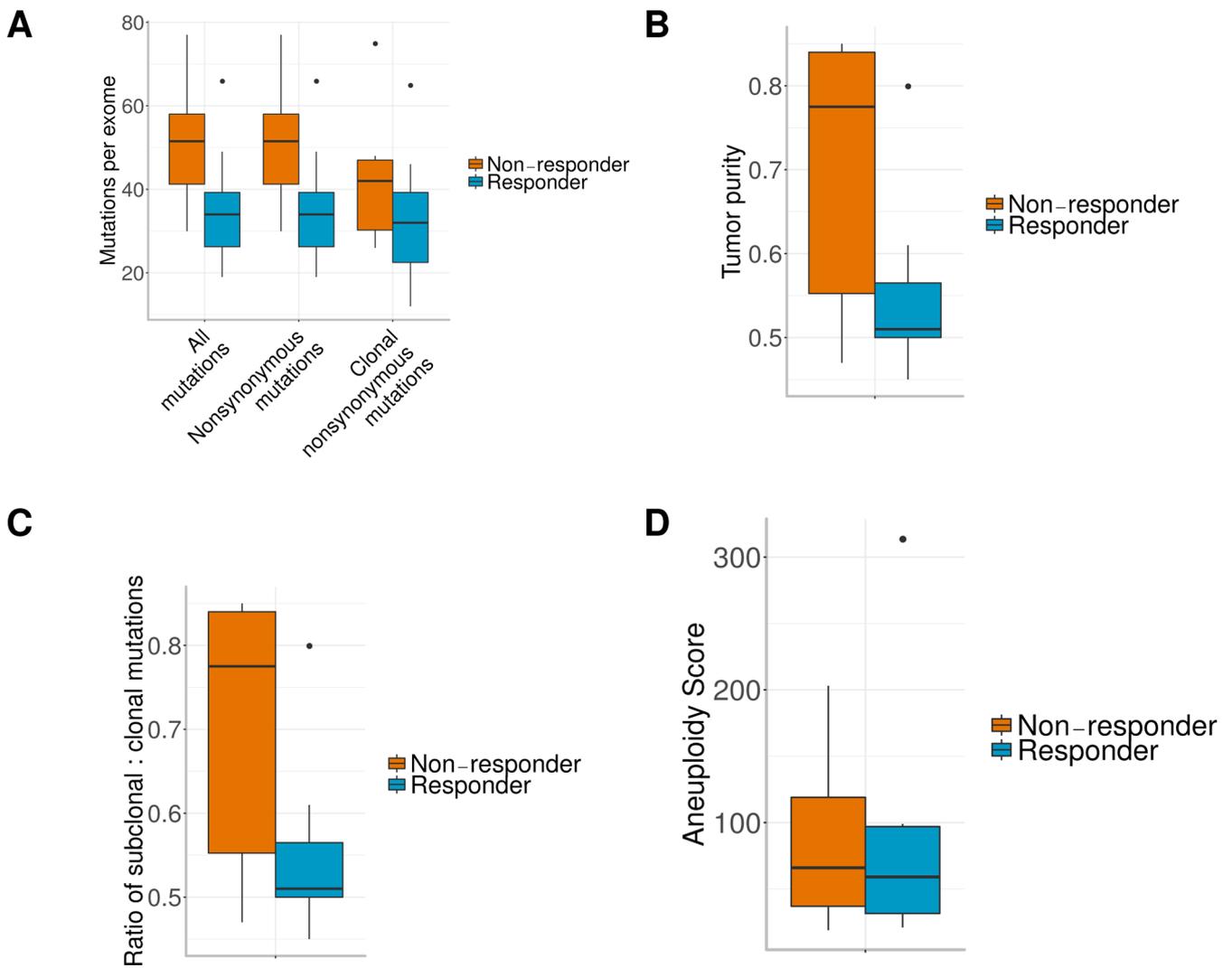
References

- Stupp, R. et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* **352**, 987–996 (2005).
- Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
- Trifonov, V., Pasqualucci, L., Tiacci, E., Falini, B. & Rabadan, R. SAVI: a statistical algorithm for variant frequency identification. *BMC Syst. Biol.* **7**, S2 (2013).
- Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).

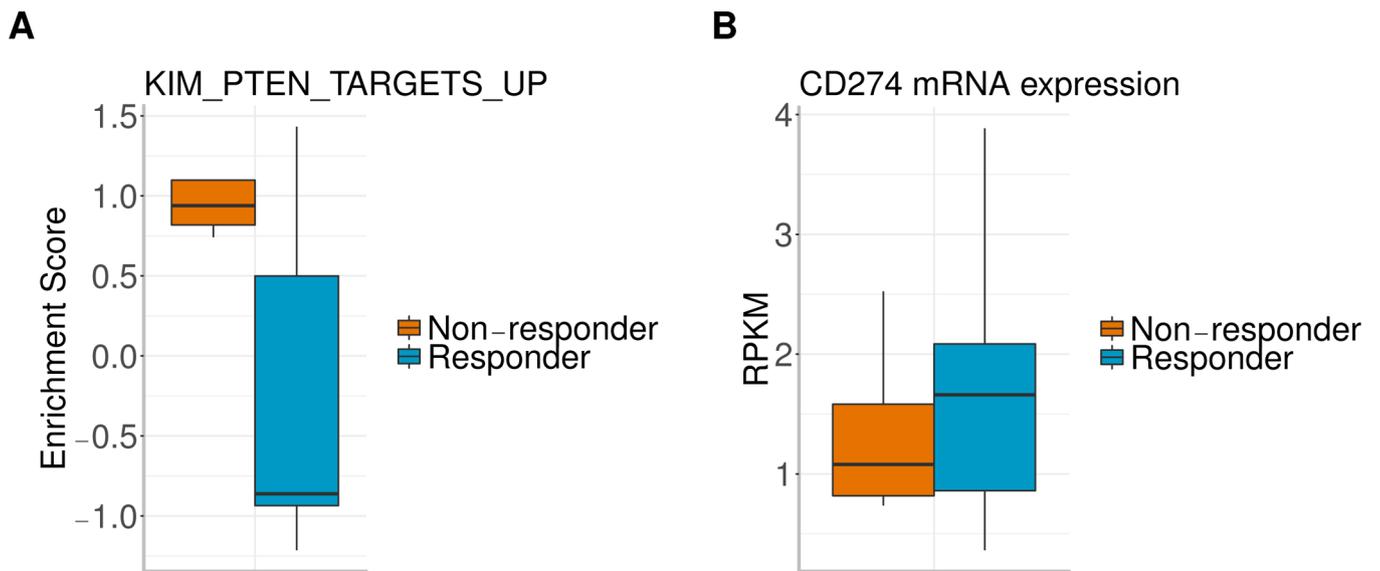
45. Carter, S. L. et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat. Biotechnol.* **30**, 413 (2012).
46. Yoshihara, K. et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* **4**, 2612 (2013).
47. Iyer, M. K., Chinnaiyan, A. M. & Maher, C. A. ChimeraScan: a tool for identifying chimeric transcription in sequencing data. *Bioinformatics* **27**, 2903–2904 (2011).
48. Abate, F. et al. Pegasus: a comprehensive annotation and prediction tool for detection of driver gene fusions in cancer. *BMC Syst. Biol.* **8**, 97 (2014).
49. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
50. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2013).
51. Hänzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013).
52. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
53. Shukla, S. A. et al. Comprehensive analysis of cancer-associated somatic mutations in class I HLA genes. *Nat. Biotechnol.* **33**, 1152 (2015).
54. Hundal, J. et al. pVAC-Seq: a genome-guided in silico approach to identifying tumor neoantigens. *Genome Med.* **8**, 11 (2016).
55. Karosiene, E., Lundegaard, C., Lund, O. & Nielsen, M. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* **64**, 177–186 (2012).
56. Bolotin, D. A. et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat. Methods* **12**, 380 (2015).
57. Baddeley, A, Rubak, E. & Turner, R. *Spatial Point Patterns: Methodology and Applications with R* (Chapman and Hall/CRC Press, Boca Raton, FL, USA, 2015).



Extended Data Fig. 1 | Additional clinical characteristics of the cohort. a, Venn diagram of the data modalities available across the 66-patient cohort. **b,c**, Kaplan-Meier curve comparing post-treatment survival (**b**) and overall survival from diagnosis (**c**) of patients who responded to anti-PD-1 therapy ($n=13$) with those that did not respond ($n=29$; $P=4.5 \times 10^{-5}$ (**b**), $P=0.0045$ (**c**), two-sided log-rank test), assessed across the entire cohort. **d**, Univariate survival analysis reveals that response to anti-PD-1 therapy is still most correlated with post-treatment survival of the patients when assessed across the entire cohort ($n=42$, 13 responders, 29 non-responders; P value, two-sided log-rank test).



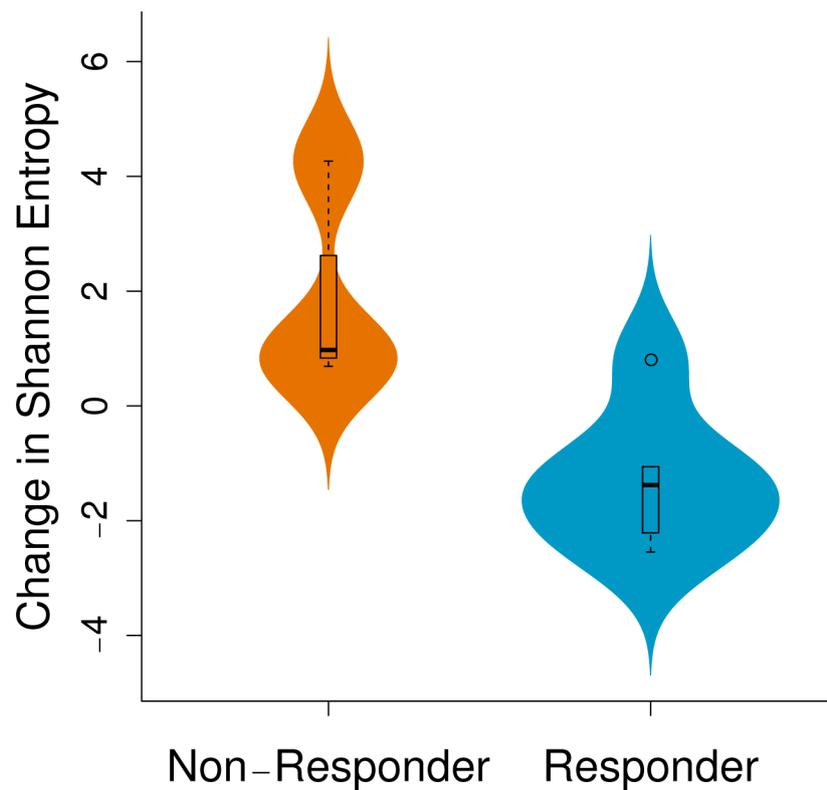
Extended Data Fig. 2 | Additional analysis of genomic correlates of response to anti-PD-1 immunotherapy. a, Mutation burden by response group ($n=17$ patients). **b**, Tumor purity, as estimated by ABSOLUTE, by response group. **c**, Ratio of subclonal to clonal mutations, as estimated by ABSOLUTE, by response group. **d**, Aneuploidy score analysis of non-responders versus responders. Boxplots show the median, interquartile range, and whiskers (1.5 times interquartile range).



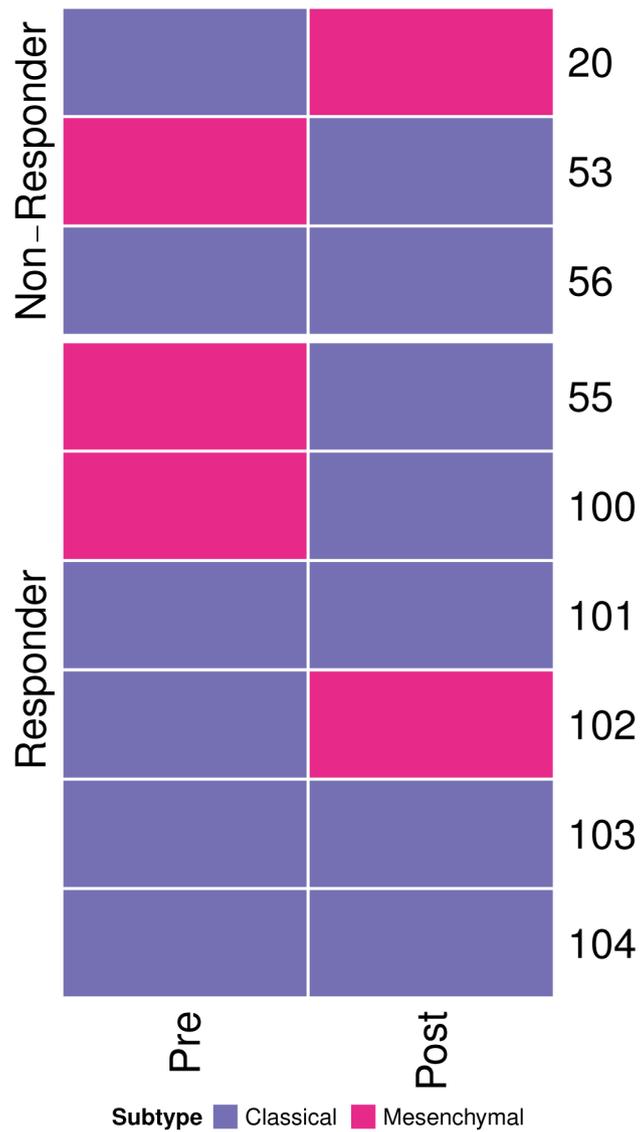
Extended Data Fig. 3 | Additional analysis of transcriptomic correlates of response to anti-PD-1 immunotherapy. a, GSEA enrichment score of gene-set KIM_PTEN_TARGETS_UP for non-responders versus responders ($n=12$ patients). The boxplot shows the median, interquartile range, and whiskers (1.5 times interquartile range). **b**, Boxplot of *CD274* (encoding PD-L1) messenger RNA expression in responders versus non-responders ($n=12$ patients). The boxplot shows the median, interquartile range, and whiskers (1.5 times interquartile range).



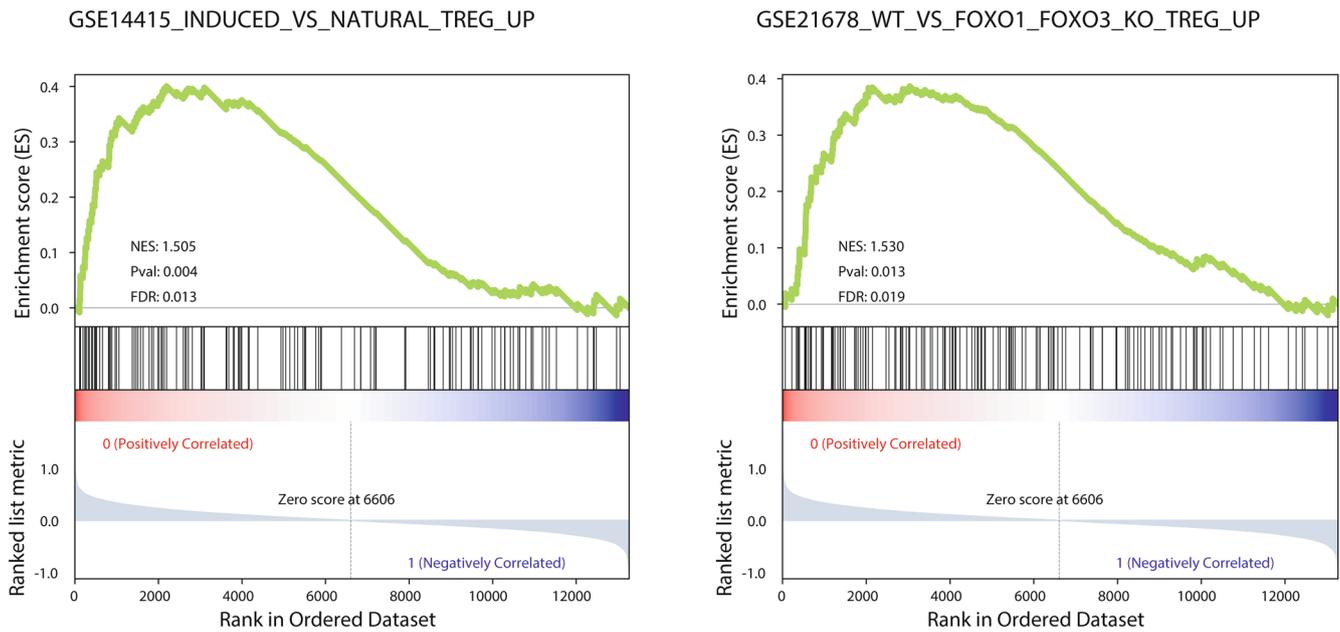
Extended Data Fig. 4 | Clonal diversity of lymphocytes before and after immunotherapy. Within seven patients with longitudinal information on TCR and immunoglobulin (Ig) RNA expression, MiXCR was used to group reads into T cell (**a**) and B cell clones (**b**). Each color on a bar represents the fractional presence of a different clone, with the total clonal read count, n , listed above.



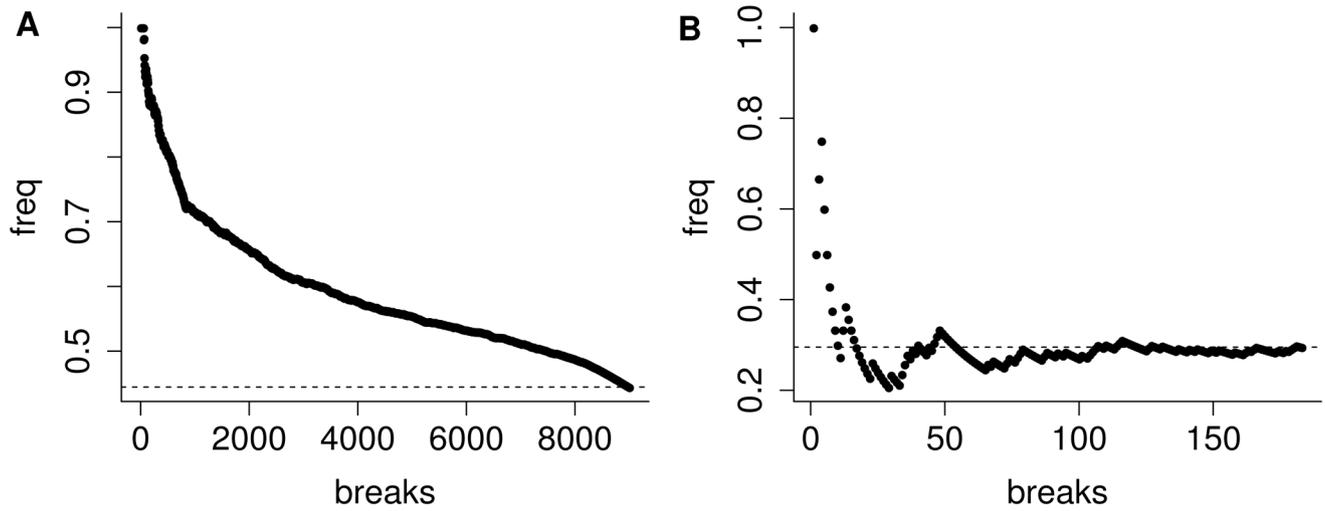
Extended Data Figure 5 | Non-responders demonstrate a greater increase in clonal diversity of B cells following immunotherapy. B cell clonal diversity before and after immunotherapy was assessed by identifying immunoglobulin RNA sequences within the tumor. Non-responders had a greater increase in Shannon entropy among B cells compared with responders ($P=0.048$, two-sided exact Mann-Whitney U test; $n=16$ independent timepoints from seven patients). The boxplot shows the median, interquartile range, and whiskers (1.5 times interquartile range); the violin plot represents sample distributions via kernel density estimation.



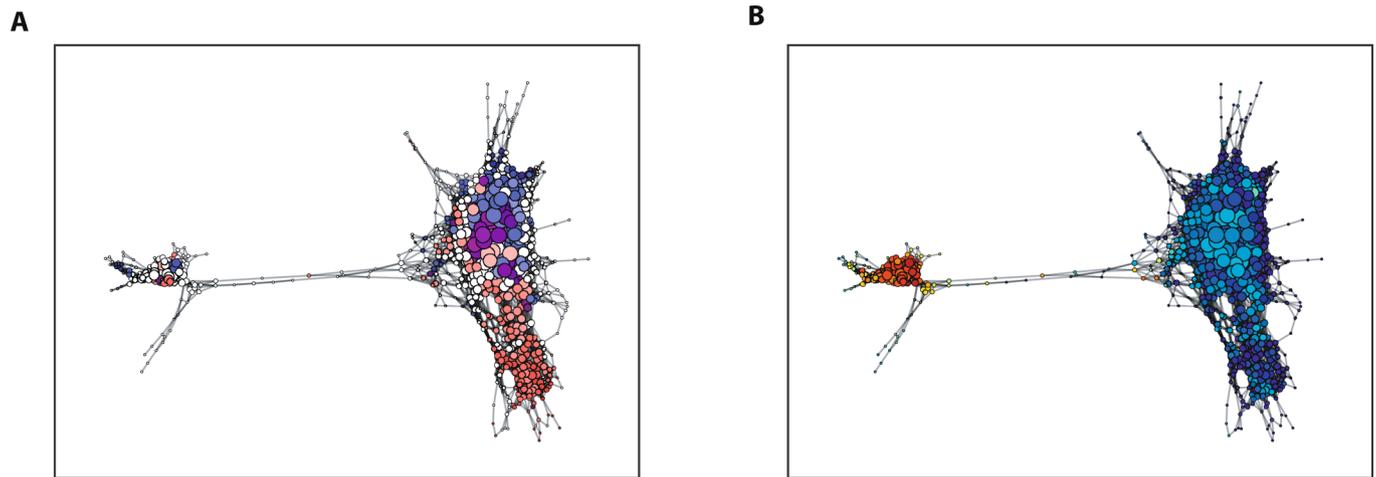
Extended Data Figure 6 | Tumor subtype. Expression subtyping of tumors from nine patients (pre- and post-treatment) into proneural, mesenchymal, and classical subtypes.



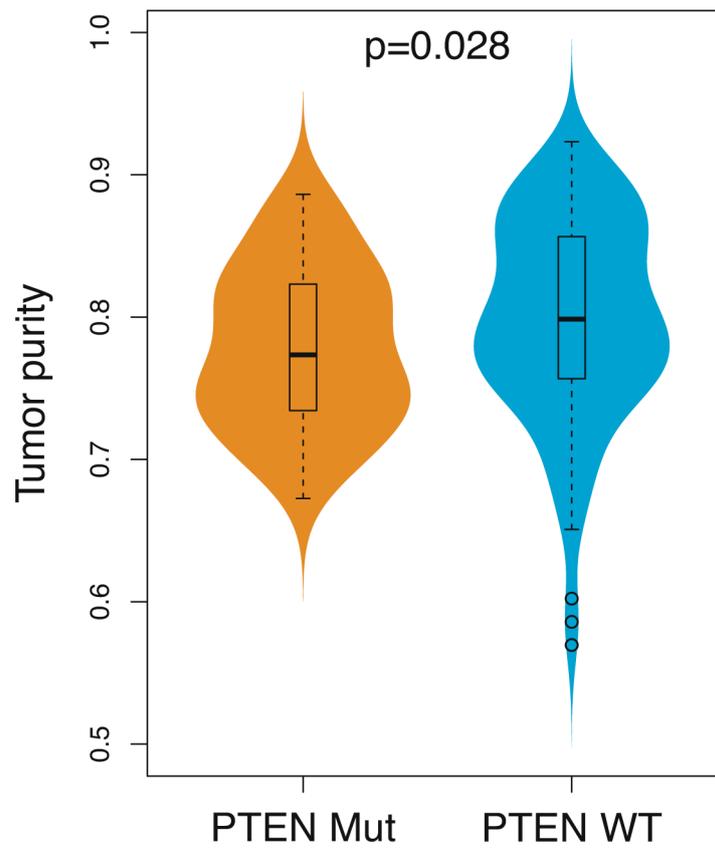
Extended Data Figure 7 | GSEA analysis. GSEA enrichment plots ($n=12$ patients; six responders versus six non-responders) of two T_{reg} -cell-related gene sets; $P=0.004$ (left), $P=0.013$ (right), two-sided Kolmogorov-Smirnov test.



Extended Data Figure 8 | Enrichment of T_{reg} cell signature. **a.** Cells associated with the T_{reg} cell signature were enriched in a *PTEN*-mutated tumor. **b.** Tumors associated with the T_{reg} cell signature were enriched in *PTEN*-mutated samples.



Extended Data Figure 9 | Single-cell RNA-seq data analysis. Topological data analysis of single-cell RNA-seq data ($n=4,000$ cells) from a *PTEN*-mutated tumor, demonstrating clusters of cells with high expression of *CD44* (A, in red) and of microglial signatures (B, in red).



Extended Data Figure 10 | Tumor purity analysis. *PTEN*-mutated GBM tumors have significantly lower tumor purity compared with *PTEN* wild-type tumors ($n=172$, two-sided Wilcoxon rank-sum test). The boxplot shows the median, interquartile range, and whiskers (1.5 times interquartile range); the violin plot represents sample distributions via kernel density estimation.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated
- Clearly defined error bars
State explicitly what error bars represent (e.g. SD, SE, CI)

Our web collection on [statistics for biologists](#) may be useful.

Software and code

Policy information about [availability of computer code](#)

Data collection

inForm version 2.4

Data analysis

R version 3.4.3
 Python version 2.7.15
 bwa version 7.15
 Picard version 2.9.2
 GATK version 3.8.0
 STAR version 2.5.3a
 featurecounts version 1.5.2
 ChimeraScan version 0.4.5a
 ABSOLUTE version 1.4
 PolySolver version 1.0
 pVACseq version 1.1
 MiXCR version 2.1.10

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All the sequencing data have been deposited in SRA PRJNA482620. Processed data and basic association analyses will be made available upon request.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size was based on the number of available, qualified tumor samples for this study. Nonparametric statistical analysis yielded significant p values which support our conclusion.
Data exclusions	No data were excluded in our analysis.
Replication	Experimental replications were not attempted. All the computational work in our manuscript are reproducible.
Randomization	The samples in our cohort are compiled retrospectively. Randomization is not relevant to our computational analysis.
Blinding	Patients are not recruited actively during this study. Blinding is not relevant to our computational analysis.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Unique biological materials
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants

Methods

n/a	Involvement in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	<ol style="list-style-type: none"> 1. CD68: BioGenex #AM416-5M, Dilution 1:RTU, Clone KP1, Species: ms 2. SOX2: Abcam #ab92494, Dilution 1:50000, Clone EPR3131, Species: rb 3. CD8: Leica #PA0183, Dilution: 1:RTU, Clone 4B11, Species: ms 4. PDL1: Abcam #ab205921, Dilution 1:650, Clone 28-8, Species: rb 5. CD3: Leica #NCL-L-CD3-565, Dilution 1:150, Clone LN10, Species: ms 6. HLA-DR: Abcam #ab166777, Dilution 1:200, Clone LN-3, Species: ms
-----------------	--

Validation	<p>All the antibodies have been validated in manufacturer's documentation.</p> <ol style="list-style-type: none"> 1. CD68: http://store.biogenex.com/us/anti-cd68-clone-kp1-4753.html 2. SOX2: https://www.abcam.com/sox2-antibody-epr3131-ab92494.html 3. CD8: https://www.leicabiosystems.com/ihc-ish-fish/immunohistochemistry-ihc-antibodies-novocastra-reagents/primary-antibodies/products/cd8/ 4. PDL1: https://www.abcam.com/pd-l1-antibody-28-8-ab205921.html
------------	--

5. CD3: <https://www.leicabiosystems.com/ihc-ish-fish/immunohistochemistry-ihc-antibodies-novocastra-reagents/primary-antibodies/products/cd3/>
6. HLA-DR: <https://www.abcam.com/hla-dr-antibody-ln-3-ab166777.html>

Human research participants

Policy information about [studies involving human research participants](#)

Population characteristics

Patient metadata including age, gender, survival, dose, etc. are included in the Supplementary Table 1.

Recruitment

The samples are compiled retrospectively. No active recruitment were conducted during this study. All the relevant covariants are tested for any possible confounding effect.