Supplemental Online Content

Cloughesy TF, Petrecca K, Walbert T, et al. Effect of vocimagene amiretrorepvec in combination with flucytosine vs standard of care on survival following tumor resection in patients with recurrent high-grade glioma: a randomized clinical trial. *JAMA Oncol*. Published online October 29, 2020. doi:10.1001/jamaoncol.2020.3161

eMethods and eMaterials
eReferences
eFigure 1. Kaplan-Meier overall survival curve with Toca FC cycle numbers.
eFigure 2. Kaplan-Meier overall survival curve for pre-planned subpopulations.
eFigure 3. Kaplan-Meier overall survival curve for subpopulations within the second recurrence.
eFigure 4. Workflow for whole exome sequencing and RNA sequencing data analyses.
eFigure 5. Quality assessment of phase 2 and phase 3 portions of Toca 5 tumor RNA sequencing.
eFigure 6. Assessment of Toca 5 RNA sequencing data quality.
eFigure 7. Boxplot comparing intra-tumor Pearson correlations in mRNA expression versus inter-tumor Pearson correlations in mRNA expression.
eFigure 8. mRNA expression profiles of Toca 5 tumors.
eFigure 9. Molecular classification of Toca 5 tumor samples.
eFigure 10. Relationships between patients' tumor molecular classification and survival.
eFigure 11. Sequencing coverage metrics across targeted regions for normal and tumor samples.
eFigure 12. Summary of DNA sequencing analyses of Toca 5 tumors.
eFigure 13. Comparison of genetic profiles from 1st and 2nd recurrence tumors.
eFigure 14. Comparison of genetic profiles from 2nd recurrence patients in the two treatment arms.
eFigure 15. Relationships between tumor DNA alterations and survival.
eFigure 16. Tumor copy number variants inferred from exome sequencing data.
eFigure 17. CNV subtypes for IDH1-wildtype GBM patients.
eFigure 18. Mutational signatures in hypermutated tumors.
eFigure 19. Number of predicted tumor neoantigens correlates with tumor mutational burden.
eFigure 20. Relationships between tumor purity and molecular subtype.
eFigure 21. Pairwise correlations among genetic alterations, RNA subtypes and histology.
eFigure 22. Relative and absolute abundance of leukocyte populations in Toca 5 tumors.
eFigure 23. mRNA expression differences between IDH1 mutant tumors and IDH1 wild-type tumors.

eFigure 24. Comparison between treatment-arms of immune cell population levels in tumors at time of surgery as measured by iSort.
eFigure 25. Stratification of patients based on tumor immune cell levels inferred by iSort.
eFigure 26. Differences in immune cell composition between IDH1 mutant tumors and IDH1 wildtype tumors.
eFigure 27. Differences in immune cell and neoantigen composition between AA tumors and GBM tumors.
eFigure 28. Differences in immune cell composition between first and second recurrence tumors.
eFigure 29. Differences in peripheral immune cell composition between IDH1 widltype and IDH1 mutant patients.
eFigure 30. Differences in peripheral immune cell composition between AA and GBM patients.
eFigure 31. Differences in peripheral immune cell composition between patients at first and second recurrence.
eFigure 32. Baseline peripheral immune balance between the control arm and the Toca 511/FC arm
eTable 1. Baseline demographics and neuro-oncology history for Phase 2 and Phase 3 patients.
eTable 2. Toca FC and SOC cycle numbers and schedule for patients.
eTable 3. Secondary endpoints for Toca 5 randomized clinical trial.
eTable 4. Baseline patient characteristics for patients in second recurrence.
eTable 5. Toca 511 viral RNA and DNA signal by cycle.
eTable 6. Proposed etiology for SBS Mutational signatures, from COSMIC.

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods and eMaterials

Results and Discussion

1.1. RNA sequencing of Toca 5 tumor biopsies

Total RNA was isolated from frozen tumor biopsies at Tocagen. RNA from one to six frozen biopsies per tumor was profiled in order to account for intra-tumor heterogeneity. Illumina sequencing libraries were prepared and sequenced by Cofactor Genomics in a CLIA-like setting. Sequencing results were obtained from 957 samples, representing 335 patients (175 on Toca 511/FC arm, and 160 in standard of care arm). Single-end 75-base reads were aligned to annotated protein coding genes (GRCh38.87) and the human genome using STAR (eFigure 3).[1] Nineteen million reads mapped, on average, to the human genome (interquartile range: 16-21 x 10^6), with 33% of reads, on average, mapping to protein coding exons (interquartile range: 28-39%). Transcript per million (TPM) expression values were generated using the pseudoalignment program Kallisto. [2]

1.2. Quality Control Assessment of Toca 5 RNA sequencing

Variation in RNA quality, RNA isolation and sequencing library preparation can have profound effects on RNA sequencing results. Unless accounted for, and if necessary, batch-corrected or compromised samples removed, these technical factors confound subsequent analyses. In order to identify potential confounding technical issues and affected samples we collected a number of quality-related parameters and asked how these parameters tracked with variation in RNA sequencing results. These parameters include initial RNA quantity and quality, measured by RNA integrity number (RIN),[3] as well as sequencing/mapping parameters including RNA strand bias, 5' to 3' bias, total number of mapped reads (reads aligned), the fraction of reads mapping to protein coding exons (exon_frac), introns (intron_frac), intergenic regions (intergenic_frac), ribosomal RNAs (rRNA_frac), and the mitochondrial genome (chrM_frac) (eDataset 1). Following initial quality assessment of the phase 2 RNA sequencing results, we increased the RIN threshold for generating RNA sequencing libraries from two to four and added an additional DNA removal step in the RNA isolation procedure. As such the RIN values for phase 3 samples (median 8.0) is higher than phase 2 samples (median 6.4), as illustrated by the violin plots in eFigure 4A. The distributions of other quality parameters are comparable between phase 2 and phase 3 samples (eFigure 4B-D). As shown in eFigure 4E, RNA sequencing results from phase 2 (red) and phase 3 (blue) were indistinguishable based on a scatterplot of the first two principle components (PCs) of the RNA sequencing dataset.

In order to relate quality metrics to variation in mRNA expression, we calculated the first ten PCs of the RNA expression dataset. The first PC accounts for 9% of variation in expression, and the first ten in sum only account for 37% of variation in expression, highlighting the multidimensionality of the dataset (eFigure 5B). eFigure 5A shows a heatmap representation of the pairwise Spearman rank correlations between quality metrics and the first 10 PCs. PC1 negatively correlated with exon_frac (r= -0.74) and RIN (r= -0.56), and positively correlated with intergenic_frac (r=0.66), highlighting the influence of RNA quality on RNA sequencing results. PC3 negatively tracked with intron_frac (r=-0.79). No other pairwise correlations between quality metrics and PC 1-10 are greater than 0.5. Thus, while quality metrics track with the first and third PCs, their overall contribution to variation in mRNA expression is small.

We next identified and removed outlier samples based on quality metrics. A scatterplot of PC1 and PC3 highlighted a number of outlier samples, a disproportionate fraction of which had exon_frac < 0.20 and/or RIN < 4 (red) (eFigure 5C). Using these criteria as cut-offs, 838 of 957 samples were retained for further analyses.

1.3. Toca 5 HGG tumor mRNA expression programs

The first hypothesis we tested was whether mRNA expression profiles of samples taken from the same tumor biopsy were invariably more similar to each other than to tumor samples taken from other patients. In order to test this hypothesis, the pairwise Pearson correlations between all samples' normalized mRNA expression was calculated (TPM values for transcript isoforms were summed, filtered to remove mRNAs with TPM < 3, $\log_2$+1 transformed, and each mRNA was mean centered across samples), and the intra-sample correlations were plotted vs inter-sample correlations (eFigure 6). The median intra-tumor Pearson correlation was 0.58 (interquartile range of 0.34-0.74) compared to a median inter-tumor correlation of -0.01 (interquartile range of -0.17-0.16). Thus, samples from the same biopsy tended to have a higher correlation in RNA expression, but with many exceptions (TPM values in eDataset 2).

As a first pass analysis, the 2794 mRNAs with the greatest variation in expression (top 25% of all mRNAs with average TPM $\geq$ 3) were retrieved and their relative expression was used as a metric of similarity for hierarchical clustering (eFigure 7). Many of the mRNAs in this set fell into one of seven distinct clusters (color bars to left of heatmap); the five largest clusters include mRNAs with common functional themes. For instance, many of the mRNAs in the cluster marked with orange bar encode proteins involved in "cell cycle" functions (83 of 396, p <1e-16 hypergeometric distribution), likely highlighting tumor samples with a higher fraction of proliferating cells. The cluster marked by the turquoise bar includes a disproportionate number of mRNAs encoding proteins linked to "immune response" (153 of 599, p <1e-67), including "innate immune response" (87, p < 1e-37). The salmon cluster includes many mRNAs encoding proteins involved in neuronal functions, including "synaptic signaling" (95 of 975, p<1e-37) and "neurogenesis" (143, p < 1e-14). We have previously shown that these mRNAs track with non-enhancing regions of the tumor and may be preferentially expressed in samples with a higher fraction of non-neoplastic cells.[4]

1.4. Molecular classification of Toca 5 tumor samples

HGG tumors' mRNA expression profiles tend to segregate into three to six groups, coinciding with specific genomic alterations and potential cell-type of origin of the tumors.[5-8] The Cancer Genome Atlas (TCGA) originally proposed a four-class grouping based on clustering analyses (herein referred to as "clustering") called "classical" -- astrocytes, "mesenchymal" – microglia, "proneural" – oligodendrocytes and "neural" – neurons. They are often found at roughly equal proportions in large scale profiling studies.[5,7] Subsequently, TCGA struck out the "neural" group, suggesting it was largely driven by expression of non-neoplastic neuronal cells. They moved to a single sample gene set enrichment method (herein referred to as "TCGA") to classify samples into the three groups with an associated p-value obtained from resampling.[8] Recently, Neftel, et al. used an integrative approach spanning single-cell RNA-sequencing of 26 tumors, bulk genetic and expression analysis of 401 specimens and defined six types of cellular states of GBM tumor cells (herein referred to as "sc-Seq"). The four main classifications are, mesenchymal-like (MES-like), neural-progenitor-like (NPC-like), Oligodendrocyte-progenitor-like (OPC-like), and astrocyte-like (AC-like). The MES-like samples are further subdivided into hypoxia-independent (MES1) and hypoxia-dependent (MES2) and NPC-like was further subdivided into OPC-related NPC1 and neuronal NPC2.[6]

We classified Toca 5 samples using the three methods and color-coded each sample's classification above the overview heatmap (eFigure 7) (eDataset 3). Classical/AC samples are black, mesenchymal/MES samples are red, proneural/NPC1 samples are blue and neural/NPC2 samples are green. Classical/AC and neural/NPC1 are enriched in samples that have relatively high expression of "cell cycle" mRNAs; mesenchymal/MES are enriched in samples that have relatively high expression of "immune response" mRNAs; neural/NPC2 are enriched in samples that have relatively high expression of neuron-related mRNAs. Consistent with published results, samples from AA tumors were predominantly proneural/NPC1 (eFigure 8 A-C). There was a similar distribution of subtypes in 1st and 2nd recurrence patients (eFigure 8 D-F). Among second recurrence only tumors, there was a similar distribution of subtypes in standard of care and Toca 511 and Toca FC treatment arms (eFigure 8 G-I).

We next asked if a tumor's molecular subtype tracked with treatment-related clinical outcomes. We performed univariate Cox proportional hazards regression comparing survival times of patients classified to each molecular subtype (highest quality RNA sample used for each patient/tumor). As shown in the forest plot in eFigure 9, molecular subtype did not track with treatment-related survival.

1.5. *IDH1* mutant tumors have a distinct expression profile

We performed a supervised analysis to identify mRNAs whose expression tracked with IDH1 R132H/S/C mutations. We identified hundreds of such mRNAs using Sleuth (eFigure 22).[9] A disproportionate number of the mRNAs whose relative expression was higher in IDH1 mutant tumors encode proteins involved in "translation initiation" and "CNS development". Conversely, mRNAs whose relative expression was higher in *IDH1* wild-type tumors encode an abundance of proteins involved "innate immune response", "regulation of cell shape" and "cell adhesion".

1.6. Leukocyte populations in Toca 5 tumors

Tumor immune cells contribute to bulk RNA expression profiles and play important roles in patient outcomes and response to therapy. We collaborated with Cibermed to de-convolute immune cell composition in Toca 5 tumors using iSort, the gold-standard method for enumeration of cell subsets from whole tissue RNA expression profiles.[10-12]

iSort takes raw RNA sequencing data as input and outputs estimates of the relative and absolute abundance of 23 immune cell types (eDataset 4 and eDataset 5). According to iSort, monocyte-derived cells are the major immune cell population in Toca 5 tumors regardless of molecular subtype (eFigure 21A), including M2 macrophages (28% of all 23 immune cells on average), monocytes (12%), MDSCs (7%), M0 macrophages (7%) and M1 macrophages (1%). In sum the monocyte-derived cells account for ~55% of the 23 immune cells on average. Among T-cell populations, resting memory CD4+ cells were most abundant (7% on average), followed by CD4+ naïve (4%) and CD8+ cells (3%). The relative abundance of immune cells did not track with RNA subtype (determined by clustering method). However, on an absolute scale, there were more noticeable differences in the estimated total percentage of these immune cells in tumors according to molecular subtype (eFigure 21B); for example, immune cells were estimated to account for 17% of all cells in mesenchymal tumor samples on average, 11% in neural, 10% in classical and 10% in proneural. The higher abundance of monocyte-derived immune cell types is consistent with molecular profiles of mesenchymal tumors. Tumor immune cell levels were indistinguishable between the two treatment arms (eFigure 23). We observed differences in the relative levels of seven immune cell types comparing tumors from AA patients and GBM patients (eFigure 25). Similar to what we observed comparing IDH1 wild-type tumors and IDH1 mutant tumors (Figure 3), GBM tumors tended to have higher levels of M0 macrophages and follicular helper T-cells than AA tumors, whereas AA tumors tended to have higher levels of naïve B cells, memory-resting CD4 T-cells, monocytes, activated dendritic cells and resting NK cells. We also observed differences in the relative levels of five immune cell populations in second recurrence tumors compared to first recurrence tumors (eFigure 27).

The levels of specific immune cell types in patients' tumors could, in principle, can be used to identify patients most likely to benefit from Toca 511 and Toca FC therapy in subsequent trials. In order to test this hypothesis, we utilized Lasso regression to build a predictive model based on the relative iSort dataset using only patients on the Toca 511 and Toca FC arm. This analysis led to an "optimized" model utilizing CD4 memory activated T-cells (weight = -25.6), M1 macrophages (-5.0), memory B-cells (-1.9), plasma cells (-1.0), monocytes (-0.51), activated dendritic cells (4.3), activated mast cells (3.9) and M2 macrophages (0.62). The optimized threshold to stratify the Toca 511 and Toca FC patients into two groups resulted in 48 patients in the "low" group and 114 patients in the "high" group. We then compared the four survival curves after splitting patients by treatment and the iSort classification (eFigure 24). Compared to each of the other three groups patients classified as Low/Toca5 lived significantly longer with a median OS of 19 months (CI 17.2-32); median OS for Low/SoC was 11.6 months (CI 8.3-28.1); median OS for High/Toca5 was 10.1 months (CI 8.4-11.6); median OS for High/SoC was 12.8 months (CI 10.7-15.4) (eFigure 24B). While the survival curves that were significantly difference, the performance of the signature as a potential predictive marker were modest, with a sensitivity of 0.47, a specificity of 0.58 and an AUC of 0.56 (eFigure 24C). Nonetheless, this work provides an interesting lead towards a potential Toca 511 and FC predictive immune marker.

1.7. Exome sequencing of Toca 5 tumor biopsies and paired blood samples

Paired tumor-whole blood exome sequencing on 148 Phase 2 and 179 Phase 3 Toca 5 patient samples was performed by Siemens utilizing the SureSelect V7 exome capture kit (Agilent). The analysis workflow to map reads to the genome and call somatic variants was developed on the DNAnexus platform and described in M&M. Briefly, paired-end sequencing reads were mapped to the human genome (hg19) with BWA-MEM, refined/realigned with Picardtools, duplicates removed, variants called and then filtered (eFigure 3). We primarily used a single variant caller, Sentieon, which mirrors the GATK workflow, but we also used results from MuSE to filter lower quality single-nucleotide variants (SNVs). Filtered variants were annotated with SNPeff (eDataset 6).

In targeted regions, tumor samples have a mean coverage of 133x, and on average 94% of targeted bases have at least 30x coverage and 85% of targeted bases have at least 50x coverage. Blood samples have a mean coverage of 70x, and on average 82% of targeted bases have at least 30x coverage and 60% of targeted bases have at least 50x coverage (eFigure 10 and eDataset 7).

Given that hundreds of samples were retrieved and processed in parallel, it's likely some sample mix-ups occurred. We therefore performed pairwise correlation analyses of germline mutations to validate tumor-normal sample pairings. This analysis identified 33 sample mix-ups, 18 of which we were able to unambiguously resolve.

1.8. Identification of driver mutations

GBM tumors frequently have mutations and/or copy number variants (CNVs) in *EGFR* and/or p53, Rb, and receptor tyrosine kinase (RTK)/Ras/phosphoinositide 3-kinase (PI3K) signaling pathways.[5,13] According to our variant calling criteria, 70% of Toca 5 tumors had CNVs or mutations in the p53 pathway (*MDM2*, *MDM4*, *TP53*, *CDKN2A*), 66% had alterations in the Rb pathway (*CDK4*, *CDK6*, *CCND2*, *CDKN2A*/*B* and *RB1*), 27% had alterations in the PI3K pathway (*PIK3CA*, *PIK3R1*, *PIK3C2G*, *PIK3CG*, *PIK3CB*, *PIK3C2B*, *PIK3C2A*, *PIK3R2*), 70% had alterations in either the PI3K pathway or *PTEN*, and 73% had alterations in the RTK pathway (*EGFR*, *PDGFRA*, *MET*, *FGFR1*/*2*/*3*). RTK alterations are largely accounted for by *EGFR* mutations or CNVs (65% of tumors). These include 62% of tumors with focal amplifications of *EGFR*, 25% of tumors with mutations in *EGFR* and 49% with whole chromosome 7 amplifications (eFigure 16).

Driver mutations in *IDH1*, usually R132H, are common in lower grade gliomas and AA (~75%),[14,15] but less frequent in GBM (<10%).[5] *IDH1* mutation is a positive prognostic marker for survival; *IDH1* mutant tumors should be viewed as categorically different from *IDH1* wild-type tumors. Thus it is important to evaluate which patients' tumors are *IDH1* mutant. Of 323 patients evaluated for *IDH1* mutations, 36 patients were identified with R132 mutations; 32 had the *IDH1*-R132H mutation, two had an *IDH1*-R132G mutation, and two patients had an *IDH1*-R132C mutation. Twenty-one of the 36 tumors with *IDH1* mutations were classified as AA. Of 34 *IDH1* mutants that had blood samples available for sequencing and had somatic variant calling performed, 33 patients also had mutations in *TP53.*

Telomerase reverse transcriptase (*TERT*) promoter mutations are present in ~55% of GBM tumors and are associated with a poor prognosis, possibly due to their inverse correlation with *IDH1* mutations.[16] In our cohort, at least 19% of patients had *TERT* promoter mutations (C228T, C250T), although coverage in this region was insufficient to determine the presence of *TERT* promoter mutations in most samples: at either given locus, 48-60% of samples were covered at <=1x. *TERT* promoter mutations were observed exclusively in *IDH1*-R132H/G/S wildtype patients.

We failed to identify a putative GBM driver mutation or copy number variant (CNV) in 37 tumors (11%), even when expanding the list of putative GBM-associated gene targets. Tumors with no known driver had lower tumor purity than other samples (p <1e-4), and tumors with RNA-seq data available were classified almost exclusively as neural (36%) or mesenchymal (44%) subtype, or a mix of neural and mesenchymal (17%); a single sample was classified as neural/classical. It was previously determined that even in samples with no other known drivers, small CNVs at 8p11.22 affecting the gene *ADAM3A* were common in tumors from our phase 1 study (manuscript in preparation). A similar pattern was observed in Toca 5 patients: losses were observed in 64% of tumors, gains in 34% of tumors, and only seven tumors did not have a CNV altering *ADAM3A*. Among patients with no known driver mutations or CNVs, 16 had a gain of *ADAM3A* (43%) and 21 had a loss (57%). Various ADAM family proteins have been implicated in cancer: *ADAM3A* deletions are frequently observed in gliomas, and relatively low expression levels of *ADAM3A* and *ADAM33* correlate with poor prognosis in other cancers.[17-20]

We next asked how driver mutations tracked with various clinical features. The most abundant driver mutations had similar frequencies between first and second recurrence tumors (eFigure 12) and between treatment arms (eFigure 13). We performed univariate cox regression analyses to determine if driver mutations tracked with differences in treatment-related outcomes. Patients' whose tumors had IDH1, ATRX and/or TP53 mutations (which often co-occur) tended to live longer on the Toca 511/FC arm (eFigure 14). Similar trends were seen for patients with PTEN or EGFR mutations.

1.9. DNA copy number changes in tumor samples

Whole exome sequencing data was used to identify tumor CNVs. We observed well-known GBM-associated CNVs, such as chromosome 7 amplifications, chromosome 10 deletions, *EGFR* focal amplification and *NF1* focal deletion (eFigure 11, eFigure 15, and eDataset 8).

Certain commonly observed tumor CNVs have prognostic value. *IDH1*-wildtype GBM patients can be stratified into three distinct prognostic subtypes based on the presence or absence of whole chromosome 1 gain, whole chromosome 19 gain, and co-amplification of *CDK4* and *MDM2*: W1 (worst survival), W2 (intermediate survival), and W3 (best survival).[21] As an example, W2, the most common subtype in GBM, is characterized by no whole chromosome 1 gain, no whole chromosome 19 gain, and no co-amplification of *CDK4* and *MDM2*. Clinical trials often disproportionately enroll patients with the W3 subtype, complicating comparisons to historical results.[22] Toca 5 patients were biased towards the W1 and W2 subtypes, while W3 subtype patients were underrepresented compared to the expected distribution (eFigure 16). Specifically, in two GBM patient cohorts, TCGA (n=256) and German Glioma Network (GGN) (n=243), 34% and 32% of patients belonged to the W3 copy number subtype, respectively.[22] In Toca 5, 24% of *IDH1*-wildtype tumors were classified as W3 (eFigure 16). CNV subtype was biased for *IDH1* mutant tumors in second recurrence (eFigure 12), but balanced between the two treatment arms (eFigure 13).

We did not observe significant treatment-related survival differences for patients whose tumors had EGFR amplification, CDk4 amplification, PDGFRA amplification or PTEN deletion, nor for patients whose tumors had W1, W2 or W3 CNV subtype (eFigure 14).

1.10. Tumor mutation burden and sources of hypermutation

Previous reports on mutation profiles of primary and recurrent HGG tumors suggest (i) most primary and recurrent tumors have relatively low tumor mutation burden compared to other tumor types and (ii) a small subset of patients exhibit “TMZ-induced hypermutation” phenotype at recurrence, whereby mutation of one or more DNA damage genes leads to a dramatic increase in the number of mutations following TMZ treatment.[23,24] Two hundred ninety-four of 318 tumors (93%) had fewer than 400 high-confidence SNVs and indels (~10/Mb), ranging from 2 to 389, median = 137 (eFigure 11). Thirteen tumors exhibited the hypermutation phenotype, which we defined as having over 1000 high-confidence SNVs or indels.

Mutations can arise from many different processes, each of which leaves a “mutational signature” in the form of biases in the specific point mutation as well as the trinucleotide context surrounding a mutation.[25] We were particularly interested in the mechanisms that gave rise to so-called hypermutant tumors. We filtered the dataset for high-confidence SNVs and used SigProfiler to deconvolute mutation signatures (eTable 6).[26] Ten of the 13 hypermutated tumors had a signature indicating that the mutations arose from an alkylating agent such as TMZ (eFigure 17); these tumors are characterized by an excess of C to T and G to A mutations. Only two of these ten patients had *IDH1* R132H/G/C mutations, which are associated with TMZ-induced hypermutation. The eight *IDH1*-wildtype patients exhibiting a TMZ hypermutation signature each had several mutations in the p53 and AKT-mTOR signaling pathways (including genes TP53, PIK3CA, PTEN, PIK3R1, PIK3C2G, PIK3CG, PIK3CB, PIK3C2B, PIK3C2A, PIK3R2, MDM2, MDM1, PDGFRA, MET, and FGFR3) and DNA repair pathways (including genes MSH6, MSH2, MLH1, PMS2), which are commonly affected in hypermutated recurrent gliomas;[27,28] the average number of mutations in these pathways combined were 4.25 mutations per patient, compared to 0.5 mutations per patient in all other IDH1-wildtype patients. A single patient with the greatest tumor mutation burden (16,530 high-confidence SNVs or indels) had a unique signature indicating concurrent polymerase epsilon mutation and defective DNA mismatch repair. This patient’s tumor was the only sample with missense mutations in two of four subunits of Pol ε (*POLE* and *POLE2*). Two hypermutated tumors had a strong SBS 07b mutation signature that is observed nearly exclusively in skin cancers and are thought to arise from exposure to ultraviolet light. However, the specific trinucleotide mutation distribution differed from specific skin cancer / UV light signatures, and appears more similar to a signature characterizing either an alkylating agent or prior treatment with ionizing radiation, as determined from previous studies,[29,30] but with a greater number of CC>TC and CC>TT mutations. A signature with a high number of CC>TC and CC>TT mutations has been associated with treatment with platinum-based drugs cisplatin or carboplatin,[31] but neither of these two patients had received prior treatment using carboplatin. Furthermore, both patients had *IDH1*-R132H mutations. Therefore, these patients might be exhibiting a combination of different signatures, including TMZ-induced hypermutation that was not accurately resolved.

Thus, consistent with previous reports, Toca 5 tumors generally displayed relatively low mutation levels, with a small subset of tumors displaying hypermutation phenotype, in most cases likely linked to TMZ treatment. We did not observe treatment related differences in survival for patients with hypermutated tumors (eFigure 14).

1.11. HLA alleles

The human leukocyte antigen (HLA) gene complex on human chromosome 6 is one of the most polymorphic regions in the genome and contributes to inter-individual variation in immune responses. Class I (HLA-A, -B, and -C) and class II (HLA-DR, -DP, and -DQ) HLA gene products function by presenting foreign antigens to T cells to trigger immune responses.[32] Patient HLA class I genotype influences response to immunotherapy in various cancer types, and loss of heterozygosity at HLA-I loci in tumors is linked to decreased survival, especially among patients whose tumors had a low tumor mutation burden.[33] In addition, certain HLA alleles are associated with favorable prognosis in GBM.[34,35] Specifically, previous studies have found correlations between OS or GBM occurrence and HLA alleles Cw*01, A*32 (specifically A*3201), B*55, the HLA-B62 supertype (including HLA-B*15:01), B*13, B*07, and B*07-Cw*07.[33-35]

We performed HLA class I and II genotyping from the whole exome sequencing data. Of 327 patients tested, 66 patients (20%) were homozygous for at least one HLA class I locus (HLA-A, -B or -C), and 261 patients (80%) were heterozygous for all three loci. 164 patients (50%) had the HLA-B44 supertype, and 55 patients (17%) had the B*07-Cw*07 haplotype. In previous studies, patients with B44 subtype have been shown to have extended survival in melanoma cohorts of patients treated with immune checkpoint blockade,[33] and the B*07-Cw*07 haplotype, a common haplotype in Caucasian populations, was positively associated with the occurrence of GBM.[35] No HLA supertype was associated with survival.

1.12. Driver mutations are common sources of tumor neoantigens

Tumor-specific antigens, referred to as neoantigens, represent novel peptide sequences encoded by somatic mutations in the cancer genome. Neoantigens may be recognized as foreign by T-cells, enabling an immune response against the tumor. Across a patient population, those whose tumors have a relatively higher number of potential neoantigens are often more likely to respond to immunotherapy, such as checkpoint inhibitors, in various cancers.[36-38] We performed *in silico* neoantigen discovery from RNA and DNA sequencing data. Putative tumor-specific neoantigens derived from expressed, nonsynonymous missense or frameshift mutations are prioritized based on predicted expression and binding affinity. We also predicted putative antigens derived from the Toca 511 viral genome, which provides another source of potential neoantigens.

The number of predicted neoantigens derived from mutated peptides binding to MHC class I or II alleles were positively correlated with tumor mutational burden (eFigure 18). Neoantigens in non-hypermutated samples most frequently arose from mutations in the following genes, when predicting binding to MHC class II alleles: *EGFR* (38 patients), *TP53* (31 patients), *PTEN* (20 patients), *IDH1* (17 patients), *PIK3CA* (6 patients); and when predicting binding to MHC class I alleles: *TP53* (11 patients), *EGFR* (11 patients), *PTEN* (11 patients), *PIK3CA* (7 patients), *TTN* (6 patients), *PTPRZ1* (5 patients) and *IDH1* (5 patients).

We did not observe treatment related differences in survival for patients with relatively high levels of MHCI ($\geq$ 6) or MHCII ($\geq$ 14) neoantigens (eFigure 20).

1.13. Assessment of tumor purity

The tumor microenvironment is a complex milieu including non-neoplastic cells such as fibroblasts, infiltrating immune cells, endothelial cells and epithelial cells.[39] The fraction of cancer cells in a tumor, called tumor purity, is often associated with clinical outcomes.[40] We estimated tumor purity from whole exome sequencing data. Tumors classified as neural or mesenchymal subtype had significantly lower tumor purity than proneural or classical subtype tumors (eFigure 19A), and overall, there was a positive correlation between the total number of mutations called and tumor purity (R^2=0.22, eFigure 19B). Tumors with no identified driver mutations had lower tumor purity estimates. There was a negative correlation between estimated tumor purity and the absolute abundance levels of nearly every immune cell type, and total abundance, as determined by iSort (eFigure 19C & D), indicating that tumor purity is in large part determined by the abundance of immune infiltrates in Toca 5 tumor samples.

We split patients into four groups based on their tumor purity (quartiles) and compared survival curves between treatment arms. There was no clear relationship between tumor purity and treatment-related survival (eFigure 14).

1.14. Interrelationships among genetic alterations, RNA subtypes and histology

There are well-documented inter-relationships among tumor histology, genetic alterations and RNA subtypes in HGG.[5] We explored these relationships among Toca 5 tumors using pairwise Jaccard correlations as a metric of similarity for hierarchical clustering (eFigure 20). The correlations, represented as a heatmap, affirm many previously documented observations. For instance, AA histology tracks with mutations in IDH1, TP53 and ATRX and is enriched for tumors with a proneural RNA subtype. EGFR mutations and amplification track with the classical/AC RNA subtype, W3 CNV subtype, high tumor purity and PTEN deletion. Mesenchymal RNA subtype tracks with W2 CNV subtype. Neural/NPC2 RNA subtype is associated with low tumor purity and is not associated with specific genetic alterations.

2. METHODS AND MATERIALS

2.1. RNA sequencing

Tumor pieces were separated from bulk tumor on dry ice and total RNA was subsequently isolated using Maxwell 16 Total RNA Purification Kit (Promega AS1050). Phase 3 RNA samples were also further purified using Zymo RNA Clean and Concentrator-5 (R1014). RNA integrity was determined using the Tapestation 4200 (Agilent) following manufacturer's recommendations.

Total RNA was shipped to Cofactor Genomics (http://cofactorgenomics.com, St. Louis, MO) to perform library construction and sequencing. Total RNA sample quantity and quality was determined by Qubit 4 Fluorometer (Invitrogen, Carlsbad, CA) and Agilent RNA 6000 Nano assay (Agilent Technologies, Sanata Clara, CA) respectively. Total RNA library construction was performed using TruSeq Stranded Total RNA Library Prep Gold Kit (Illumina, San Diego, CA) according to manufacturer's instructions: Briefly, rRNA-probes were hybridized to total RNA in order to remove nuclear-encoded and mitochondrial contaminating ribosomal RNA from the sample. The resulting ribo-depleted RNA was then fragmented. First-strand cDNA synthesis was performed using reverse transcriptase and random primers in the presence of Actinomycin D, followed by second-strand cDNA synthesis with DNA polymerase I and RNase H. Double-stranded cDNA was end-repaired and A-tailed for subsequent adaptor ligation. Indexed adaptors were ligated to the A-tailed cDNA. Enrichment by PCR was performed to generate the final cDNA sequencing library. Libraries were quantified using Qubit 4 Fluorometer (Invitrogen, Carlsbad, CA) and quality assessed using Agilent D1000 assay (Agilent Technologies, Sanata Clara, CA). Libraries were sequenced as single-end 75 base pair reads on an Illumina NextSeq500 following the manufacturer's protocols.

2.2. RNA sequencing analysis workflow

Quantification of RNA expression levels was performed with Kallisto 0.46.0 against "protein coding" transcripts in Ensembl GRCh38.87 using the following command line options: --single -b 100 -l 220 -s 60 --rf-stranded.[2] Normalized RNA expression values in the form of TPMs were combined into a dataframe (eDataset 2). For gene-centric analyses TPMs for transcripts corresponding to a single ENSEMBL gene_id were summed. TPM data were filtered to remove unreliable measurements (average TPM < 3), $\log_2$+1 transformed, then mean centered across each gene/transcript.

Sequencing reads were also aligned to the human ENSEMBL reference RNA database and then the human genome using STAR,[1] on DNAnexus (Palo Alto, CA, USA) platform (version 1.2.0: --quantMode TranscriptomeSAM --outSAMunmapped Within --outSAMattributes NH HI AS NM MD --outSAMstrandField intronMotif --outSAMtype BAM SortedByCoordinate --quantMode TranscriptomeSAM GeneCounts).

Sequencing metrics were obtained using several methods. Qualimap2 was used to obtain basic alignment and coverage statistics (rnaseq -a proportional).[41] We calculated the fraction of reads mapping to the mitochondrial genome (chrM_frac) from the Samtools (version 1.1) idxstats file. We estimated the fraction of reads mapping to ribosomal RNAs (rRNA_frac) by mapping the first 1e6 reads in each fastq to a reference genome containing human 5S, 5.8S, 18S and 28S sequences using BWA mem under default conditions. Quality metrics for each sample are available in eDataset 1.

We used Sleuth v0.30.0 identify mRNAs whose expression tracks with *IDH1* mutation status.[9] Sleuth was run in "gene mode", such that p-values for transcripts mapping to a single gene were aggregated.

Identification of functional themes associated with specific sets of mRNAs was performed using the hypergeometric distribution using mRNAs with reliable measurements as background. P-values were adjusted with Benjamini-Hochberg procedure. Gene sets were obtained from MSigDB: "c5.bp.v6.2.symbols.gmt".

2.3. Tumor molecular classification

Four-class gene lists were provided by TCGA GBM consortium leader Roel Veerhak in August 2014. We performed hierarchical clustering using the union of the four gene lists (Pearson correlation, average linkage) and assigned samples using the top nodes in the dendrogram. An R script for assigning samples to TCGA's newer 3-class system was obtained from Wang et al.[8] We used a p-value cut-off of 0.2. The six-class classification method was based on Neftel, et al.[6]

2.4. Deconvolution of leukocyte populations in tumors using iSort

Tumor infiltrating leukocytes were enumerated from bulk tumor RNA sequencing using iSort^{TM} (CiberMed, Inc.), a gene expression deconvolution framework based on CIBERSORT technology that currently supports the detection of twenty-three functionally-defined human immune subsets in bulk tumor expression data, and allows for the "relative" quantification of leukocyte composition in relation to total immune content and "absolute" quantification of leukocytes in relation to total tumor cellularity.[11,12]

2.5. Identification of a potential predictive iSort signature

For these analyses we used relative iSort values from each patient's highest quality RNA sample (RIN + exon_frac). The R package "glmnet" was used for Cox Lasso regression.[42] We identified the optimal binary threshold based on finding the maximal log-rank statistic ("maxstat" package in R). Performance metrics were determined using R packages "survivalROC" and "riskRegression".

2.6. Exome sequencing

Genomic DNA was extracted from tumor or blood using Qiagen Blood and Tissue DNA Purification Kit (Qiagen, 69506) and sent to Siemens Healthineers (Berkeley, CA) for library preparation and sequencing.

Isolated genomic DNA was assessed for quantity and quality using the Agilent 2200 TapeStation with Genomic DNA ScreenTape (Agilent, 5067-5365). DNA was sheared using the Covaris M220 Focused-Ultrasonicator to yield fragments in the range of 150 to 200 bp, which was confirmed using the D1000 ScreenTape (Agilent, 5067-5582). Following end repair, A-tailing, and ligation with sequencing adapters, the libraries were amplified by PCR and purified using AMPure XP paramagnetic particles (Beckman Coulter, A63881).

SureSelectXT Human All Exon V6+Cosmic (Agilent, 5190-9307) and SureSelectXT2 Reagent Kits (Agilent, G9621B) were utilized for enrichment. After purification using Dynabeads Streptavidin T1 (Thermo Fisher, 65602), the libraries were amplified and enriched a second time.

Libraries were then pooled at ratios to provide sufficient reads for both tumor (75% of targeted bases were covered at 50X) and blood (75% of targeted bases were covered at 30X for blood) samples. PhiX DNA (Illumina, FC-110-3001) was added before pair-end sequencing using Illumina NextSeq 500 sequencers and NextSeq® 500/550 High Output v2 (150 cycles) flow cells (Illumina, FC-404-2002).

2.7. Exome sequencing analysis workflow

Sequencing data were processed on the DNAnexus platform. Reads were aligned to the human genome (hg19) with BWA-MEM (version 0.7.12) using default parameters.[43,44] Duplicates were removed from sorted bam files (Samtools version 1.1)[45] with PicardTools MarkDuplicates (version 2.0.1). Point mutations and indels were identified with MuSE (1.0rc) and Sentieon (201611) using default settings.[46,47] VCF files were filtered against variants in dbSNP138 (after removing variants with SAO flags of 2 or 3), Exome Sequencing Project v. 6500, and 1000 Genomes Project phase 1 indels, using VCFtools (version 0.1.15).[48] VCF files were then filtered to retain only "PASS" variants and annotated using SnpEff (version 4.3.2).[49]

We primarily utilized a single variant caller, Sentieon, which mirrors the GATK workflow, but also considered information from variant calls made by MuSE to filter out lower quality SNVs. The union of "moderate" and "high" impact mutations (called by SnpEff) from both variant callers, MuSE and Sentieon, was used to calculate the total number of mutations and identify driver mutations. Custom python scripts were used to parse annotated variant output, determine the trinucleotide context of each mutation, calculate information about driver mutations, and generate plots. Samtools (version 1.1) mpileup command was used to identify *TERT* promoter mutations.[45] A sample was considered to contain a *TERT* promoter mutation C228 (chr5:1295228) or C250 (chr5:1295250) if the alternate allele was present in at least one read. SigProfiler was used to deconvolute mutational signatures based on the trinucleotide context of mutations in a tumor sample.[26] A diagram of the workflow is shown in eFigure 3.

Qualimap2 (version 1.1.0) was used to obtain basic alignment and coverage statistics.[41] Samples were flagged and re-sequenced if the variants called were unexpectedly high in number and were mostly annotated germline variants, indicating a possible sample mix-up. After re-sequencing these samples, tumor-normal pairings were confirmed by calculating the correlation between raw FASTQ files using NGSCheckMate (version 1.0).[50]

2.8. Identification of copy number variants

Exome data was analyzed for CNVs using CNVkit (version 0.9.1).[51] Copy number subtypes were defined by four genetic loci: gain of whole chromosome 1 (gChr1), gain of whole chromosome 19 (gChr19), and co-amplification of CDK4/MDM2 (caCDK4/MDM2); subtypes were called according to the algorithm described in Cimino *et al.* 2018,[22] as follows: W1 = [No gChr1 + No gChr19 + caCDK4/MDM2] or [gChr1 + caCDK4/MDM2]; W2 = [No gChr1 + No gChr19 + No caCDK4/MDM2]; W3 = [No gChr1 + gChr19] or [gChr1 + No caCDK4/MDM2]. The copy number variant calling threshold was determined by calculating the threshold that would call a known 1p/q19 codeletion, which was identified in a pathology report, in one patient from a phase I study. This threshold was a $\log_2$-scale copy number ratio of 0.14. Copy number subtypes in TCGA dataset were obtained from Cimino *et al.* 2018.[22] Heatmap plots were generated with the 'heatmap' command in CNVkit with segment files.

2.9. Estimating tumor purity

Tumor purity estimates were obtained using R packages BubbleTree (version 2.10.0) and SciClone (version 1.1.0) (Bioconductor version 3.7).[52,53] For SciClone, the highest cluster means under 0.5 were considered, and if one of the estimates was close to the estimated obtained using BubbleTree, that was considered to be the correct tumor purity estimate. If estimates between the two tools did not match, the entire range of possible tumor purity estimates was considered. The ranges for tumor purity indicated in eFigure 19 are the range of possible tumor purity estimates obtained by SciClone and BubbleTree. The mid-range was used for calculating correlation with other variables and for plotting figures. Calculating statistical significance for tumor purity versus expression-based molecular subtype was done using a one-way ANOVA, and pairwise t-tests between categories.

2.10. Determining HLA allele zygosity

Raw reads from blood and tumor sample whole exome sequencing data were merged using Samtools version 1.1.[45] HLA typing was performed on merged FASTQ files using HLA-HD version 1.2.0.1[54] using the IPD-IMGT/HLA database version 3.36.0.[55] The output was parsed using custom python scripts, and whether a patient was heterozygous or homozygous for a HLA class I allele was determined based on the 4-digit type.

2.11. Neoantigen calling workflow

We identified putative neoantigens by implementing the *in silico* workflow (strategy IV) described in Karasaki et al. 2017.[56] Briefly, proteins were filtered for those that are expressed in RNA-seq data (average TPM ≥ 1) and also showing expression of the mutant peptide (threshold for mutant allele frequency ≥ 0.04). Next, K-mers surrounding each mutation were generated using custom python scripts, and the binding strength between each K-mer and a patient's HLA alleles was calculated using NetMHCpan version 3.0 for HLA class I alleles and NetMHCIIpan version 4.0 for HLA class II alleles.[57,58] Mutant peptides with a binding score ≥ 0.5 were considered strong binders for MHC class I alleles, and peptides with an affinity $IC_{50} \leq 500$ nM were considered strong binders for MHC class II alleles.

2.12. Flow Cytometry immunophenotyping

Peripheral blood mononuclear cells (PBMCs) were isolated from venous blood within 48 hours of collection using Vacutainer CPT™ Cell Preparation Tube with Sodium Heparin (BD Biosciences). PBMCs were cryopreserved in RPMI 1640 (Cellegro) with 12.5% human serum albumin (Gemini bio-products) and 10% DMSO (Sigma), and then stored in vapor phase liquid nitrogen until needed. Samples were de-identified during processing by assigning lab ID numbers, and then analyzed in random order by technicians blind to sample identities. Thawed PBMCs were rested overnight (16-20 hours) at 37oC with 5% CO2 with a density of approximately 1E+06 cells per mL in 25cm2 flasks (Corning) in AIMV media (Gibco) containing 10% human serum, (Sigma) and ≥1000U/mL IL-2. After resting, 1E+05 to 5E+05 PBMCs were transferred to each well of 96-well U-bottom polypropylene plates (Costar). To prepare fluorescence minus one (FMO) controls, multiple PBMC samples were pooled from each batch. Single color controls were prepared using UltraComp eBeads (Invitrogen) and ArC Amine Reactive Compensation beads (Invitrogen) according to manufacturer's instructions. In addition, a healthy PBMC control sample was run with every batch to track variability in results. Samples were stained with LIVE/DEAD Fixable Green Dead Cell Stain kit (Invitrogen), Fc blocked using Human TruStain FcX™ (Biolegend), surface stained using fluorochrome-conjugated antibodies (Biolegend, eBioscience and BD Biosciences), fixed and permeabilized using Foxp3/Transcription Factor Staining Buffer Set staining kit (eBioscience), and then stained intracellularly using fluorochrome-conjugated antibodies (Biolegend, eBioscience and BD Biosciences). Data were acquired using a BD Biosciences LSRFortessaTM X-20 flow cytometer equipped with high throughput sampler (HTS) and Fluidics Flow Supply System (FFSS). At least 10,000 events per sample replicate was acquired using BD FACSDiva software version 8.0.1. (BD Biosciences), and compensation was performed using single color controls prior to exporting FCS files. Cell population gating was carried out using FlowJo™ v10.5.3 (Tree Star Inc.) using FMO controls and unstained samples to determine cutoffs for positive and negative events. Statistical analyses and graphical visualizations were performed using R version 3.4.1 or later [59], with data converted from wide format to long format using the reshape and reshape2 packages[60,61] in order to facilitate plotting using ggplot2[62] .

3. DATASETS

eDataset 1. RNA sequencing quality metrics.

eDataset 2. RNA sequencing TPM values.

eDataset 3. Summary of patients RNA and DNA sequencing analyses.

eDataset 4. Absolute iSort results.

eDataset 5. Relative iSort results.

eDataset 6. High-confidence protein impacting mutations.

eDataset 7. Exome sequencing coverage summary.

eDataset 8. CNV values for a subset of loci.

4. LIST OF ABBREVIATIONS

AA	anaplastic astrocytoma
CNV	copy number variant
GATK	Genome Analysis Tool Kit
GBM	glioblastoma
HGG	high-grade glioma
NGS	next-generation sequencing
OS	overall survival
SNV	single-nucleotide variant
ssGSEA	single sample Gene Set Enrichment Analysis
TCGA	The Cancer Genome Atlas
Toca 511	vocimagene amiretrorepvec
Toca FC	extended-release 5-fluorocytosine
TPM	transcripts per million

5. REFERENCES

1. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.
2. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology.* 2016;34:525.
3. Schroeder A, Mueller O, Stocker S, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology.* 2006;7:3-3.
4. Cloughesy TF, Landolfi J, Hogan DJ, et al. Phase 1 trial of vocimagene amiretrorepvec and 5-fluorocytosine for recurrent high-grade glioma. *Science translational medicine.* 2016;8(341):341ra375.
5. Brennan CW, Verhaak RGW, McKenna A, et al. The Somatic Genomic Landscape of Glioblastoma. *Cell.* 2013;155(2):462-477.
6. Neftel C, Laffy J, Filbin MG, et al. An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell.* 2019;178(4):835-849.e821.
7. Verhaak RGW, Hoadley KA, Purdom E, et al. An integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR and NF1. *Cancer cell.* 2010;17(1):98.
8. Wang Q, Hu B, Hu X, et al. Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer cell.* 2018;33(1):152.
9. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature methods.* 2017;14(7):687-690.
10. Gentles AJ, Newman AM, Liu CL, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nature medicine.* 2015;21(8):938-945.
11. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nature methods.* 2015;12(5):453-457.
12. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol.* 2019;37(7):773-782.
13. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature.* 2008;455(7216):1061-1068.
14. Killela PJ, Pirozzi CJ, Reitman ZJ, et al. The genetic landscape of anaplastic astrocytoma. *Oncotarget.* 2013;5(6):1452-1457.
15. Wang J, Bettegowda C. Genomic discoveries in adult astrocytoma. *Current opinion in genetics & development.* 2015;30:17-24.
16. Nonoguchi N, Ohta T, Oh J-E, Kim Y-H, Kleihues P, Ohgaki H. TERT promoter mutations in primary and secondary glioblastomas. *Acta Neuropathologica.* 2013;126(6):931-937.
17. Barrow J, Adamowicz-Brice M, Cartmill M, et al. Homozygous loss of ADAM3A revealed by genome-wide analysis of pediatric high-grade glioma and diffuse intrinsic pontine gliomas. *Neuro-Oncology.* 2011;13(2):212-222.
18. Mochizuki S, Okada Y. ADAMs in cancer cell proliferation and progression. *Cancer Science.* 2007;98(5):621-628.
19. Manica GCM, Ribeiro CF, Oliveira MASd, et al. Down regulation of <i>ADAM33</i> as a Predictive Biomarker of Aggressive Breast Cancer. *Scientific Reports.* 2017;7:44414.

20. Wang Y, Zhang X, Wu YL. Homozygous deletion of ADAM3A revealed by genome-wide analysis of early-stage NSCLC in China showed to be correlated with poor prognosis. *Journal of Clinical Oncology.* 2011;29(15_suppl):e21177-e21177.
21. Cimino PJ, Zager M, McFerrin L, et al. Multidimensional scaling of diffuse gliomas: application to the 2016 World Health Organization classification system with prognostically relevant molecular subtype discovery. *Acta Neuropathol Commun.* 2017;5(1):39.
22. Cimino PJ, McFerrin L, Wirsching HG, et al. Copy number profiling across glioblastoma populations has implications for clinical trial design. *Neuro Oncol.* 2018;20(10):1368-1373.
23. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Medicine.* 2017;9(1):34.
24. Kim J, Lee I-H, Cho HJ, et al. Spatiotemporal Evolution of the Primary Glioblastoma Genome. *Cancer Cell.* 2015;28(3):318-328.
25. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415-421.
26. Alexandrov LB, Kim J, Haradhvala NJ, et al. The Repertoire of Mutational Signatures in Human Cancer. *bioRxiv.* 2019:322859.
27. Johnson A, Severson E, Gay L, et al. Comprehensive Genomic Profiling of 282 Pediatric Low- and High-Grade Gliomas Reveals Genomic Drivers, Tumor Mutational Burden, and Hypermutation Signatures. *The Oncologist.* 2017;22(12):1478-1490.
28. Johnson BE, Mazor T, Hong C, et al. Mutational Analysis Reveals the Origin and Therapy-driven Evolution of Recurrent Glioma. *Science (New York, NY).* 2014;343(6167):189.
29. Davidson PR, Sherborne AL, Taylor B, Nakamura AO, Nakamura JL. A pooled mutational analysis identifies ionizing radiation-associated mutational signatures conserved between mouse and human malignancies. *Scientific reports.* 2017;7(1):7645-7645.
30. Sherborne AL, Davidson PR, Yu K, Nakamura AO, Rashid M, Nakamura JL. Mutational Analysis of Ionizing Radiation Induced Neoplasms. *Cell Rep.* 2015;12(11):1915-1926.
31. Pich O, Muiños F, Lolkema MP, Steeghs N, Gonzalez-Perez A, Lopez-Bigas N. The mutational footprints of cancer therapies. *bioRxiv.* 2019:683268.
32. Shiina T, Hosomichi K, Inoko H, Kulski JK. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of Human Genetics.* 2009;54(1):15-39.
33. Chowell D, Morris LGT, Grigg CM, et al. Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science (New York, NY).* 2018;359(6375):582-587.
34. Song W, Ruder AM, Hu L, et al. Genetic Epidemiology of Glioblastoma Multiforme: Confirmatory and New Findings from Analyses of Human Leukocyte Antigen Alleles and Motifs. *PLoS ONE.* 2009;4(9).
35. Tang J, Shao W, Dorak MT, et al. Positive and Negative Associations of Human Leukocyte Antigen Variants with the Onset and Prognosis of Adult Glioblastoma Multiforme. *Cancer Epidemiology and Prevention Biomarkers.* 2005;14(8):2040-2044.

36. Johanns TM, Bowman-Kirigin JA, Liu C, Dunn GP. Targeting Neoantigens in Glioblastoma: An Overview of Cancer Immunogenomics and Translational Implications. *Neurosurgery.* 2017;64(CN_suppl_1):165-176.
37. McGranahan N, Furness AJS, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science.* 2016;351(6280):1463.
38. Snyder A, Makarov V, Merghoub T, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *The New England journal of medicine.* 2014;371(23):2189-2199.
39. Joyce JA, Pollard JW. Microenvironmental regulation of metastasis. *Nature Reviews Cancer.* 2009;9(4):239-252.
40. Aran D, Sirota M, Butte AJ. Systematic pan-cancer analysis of tumour purity. *Nature Communications.* 2015;6.
41. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics.* 2016;32(2):292-294.
42. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of statistical software.* 2010;33(1):1-22.
43. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:13033997 [q-bio].* 2013.
44. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England).* 2009;25(14):1754-1760.
45. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078-2079.
46. Fan Y, Xi L, Hughes DST, et al. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biology.* 2016;17(1):178.
47. Weber JA, Aldana R, Gallagher BD, Edwards JS. *Sentieon DNA pipeline for variant detection - Software-only solution, over 20× faster than GATK 3.3 with identical results.* PeerJ Inc.; 2016/01/22/ 2016. e1672v2.
48. Danecek P, Auton A, Abecasis G, et al. The variant call format and VCFtools. *Bioinformatics.* 2011;27(15):2156-2158.
49. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w(1118); iso-2; iso-3. *Fly.* 2012;6(2):80-92.
50. Lee S, Lee S, Ouellette S, Park W-Y, Lee EA, Park PJ. NGSCheckMate: software for validating sample identity in next-generation sequencing studies within and across data types. *Nucleic Acids Research.* 2017;45(11):e103.
51. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol.* 2016;12(4):e1004873.
52. Zhu W, Kuziora M, Creasy T, et al. BubbleTree: an intuitive visualization to elucidate tumoral aneuploidy and clonality using next generation sequencing data. *Nucleic Acids Research.* 2016;44(4):e38.
53. Miller CA, White BS, Dees ND, et al. SciClone: Inferring Clonal Architecture and Tracking the Spatial and Temporal Patterns of Tumor Evolution. *PLOS Computational Biology.* 2014;10(8):e1003665.

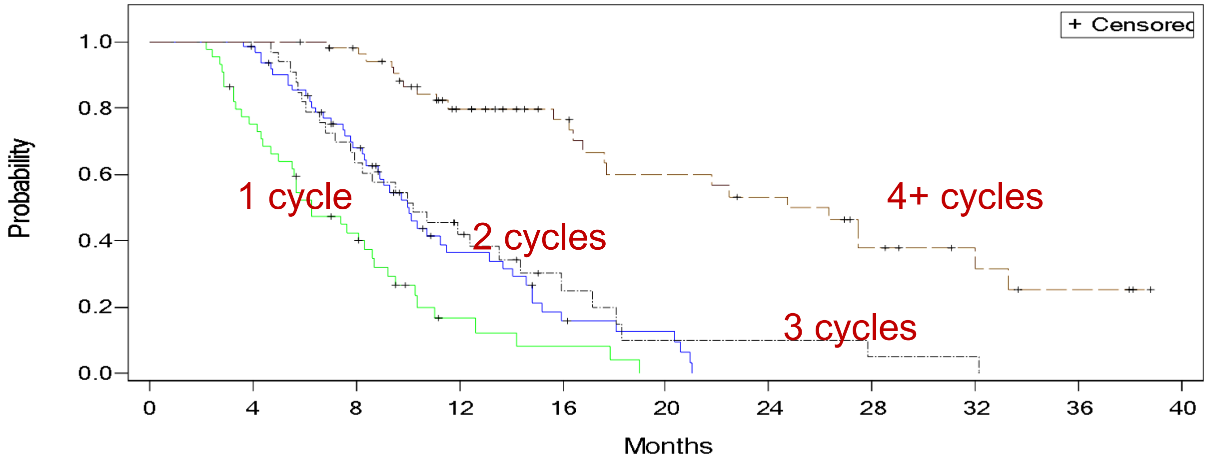
54. Kawaguchi S, Higasa K, Shimizu M, Yamada R, Matsuda F. HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data. *Human Mutation.* 2017;38(7):788-797.
55. Lefranc M-P, Giudicelli V, Ginestoux C, et al. IMGT, the international ImMunoGeneTics information system. *Nucleic acids research.* 2009;37(Database issue):D1006-D1012.
56. Karasaki T, Nagayama K, Kuwano H, et al. Prediction and prioritization of neoantigens: integration of RNA sequencing data with whole-exome sequencing. *Cancer science.* 2017;108(2):170-177.
57. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *Journal of immunology (Baltimore, Md : 1950).* 2017;199(9):3360-3368.
58. Karosiene E, Rasmussen M, Blicher T, Lund O, Buus S, Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. *Immunogenetics.* 2013;65(10):711-724.
59. RC T. R Foundation for Statistical Computing. *Vienna, Austria.* 2019.
60. Zhang Z. Reshaping and aggregating data: an introduction to reshape package. *Ann Transl Med.* 2016;4(4):78.
61. Wickham H. Reshaping Data with the reshape Package. *Journal of statistical software.* 2007;21(12):1-20.
62. Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer-Verlag; 2016.

6. SUPPLEMENTARY ONLINE FIGURES AND TABLES

eFigure 1. Kaplan-Meier overall survival curve with Toca FC cycle numbers.
eFigure 2. Kaplan-Meier overall survival curve for pre-planned subpopulations.
eFigure 3. Kaplan-Meier overall survival curve for subpopulations within the second recurrence.
eFigure 4. Workflow for whole exome sequencing and RNA sequencing data analyses.
eFigure 5. Quality assessment of phase 2 and phase 3 portions of Toca 5 tumor RNA sequencing.
eFigure 6. Assessment of Toca 5 RNA sequencing data quality.
eFigure 7. Boxplot comparing intra-tumor Pearson correlations in mRNA expression versus inter-tumor Pearson correlations in mRNA expression.
eFigure 8. mRNA expression profiles of Toca 5 tumors.
eFigure 9. Molecular classification of Toca 5 tumor samples.
eFigure 10. Relationships between patients' tumor molecular classification and survival.
eFigure 11. Sequencing coverage metrics across targeted regions for normal and tumor samples.
eFigure 12. Summary of DNA sequencing analyses of Toca 5 tumors.
eFigure 13. Comparison of genetic profiles from 1st and 2nd recurrence tumors.
eFigure 14. Comparison of genetic profiles from 2nd recurrence patients in the two treatment arms.
eFigure 15. Relationships between tumor DNA alterations and survival.
eFigure 16. Tumor copy number variants inferred from exome sequencing data.
eFigure 17. CNV subtypes for IDH1-wildtype GBM patients.
eFigure 18. Mutational signatures in hypermutated tumors.
eFigure 19. Number of predicted tumor neoantigens correlates with tumor mutational burden.
eFigure 20. Relationships between tumor purity and molecular subtype.
eFigure 21. Pairwise correlations among genetic alterations, RNA subtypes and histology.
eFigure 22. Relative and absolute abundance of leukocyte populations in Toca 5 tumors.
eFigure 23. mRNA expression differences between IDH1 mutant tumors and IDH1 wild-type tumors.
eFigure 24. Comparison between treatment-arms of immune cell population levels in tumors at time of surgery as measured by iSort.
eFigure 25. Stratification of patients based on tumor immune cell levels inferred by iSort.
eFigure 26. Differences in immune cell composition between IDH1 mutant tumors and IDH1 wildtype tumors.
eFigure 27. Differences in immune cell and neoantigen composition between AA tumors and GBM tumors.
eFigure 28. Differences in immune cell composition between first and second recurrence tumors.
eFigure 29. Differences in peripheral immune cell composition between IDH1 widltype and IDH1 mutant patients.
eFigure 30. Differences in peripheral immune cell composition between AA and GBM patients.
eFigure 31. Differences in peripheral immune cell composition between patients at first and second recurrence.
eFigure 32. Baseline peripheral immune balance between the control arm and the Toca 511/FC arm

eTable 1. Baseline demographics and neuro-oncology history for Phase 2 and Phase 3 patients.
eTable 2. Toca FC and SOC cycle numbers and schedule for patients.

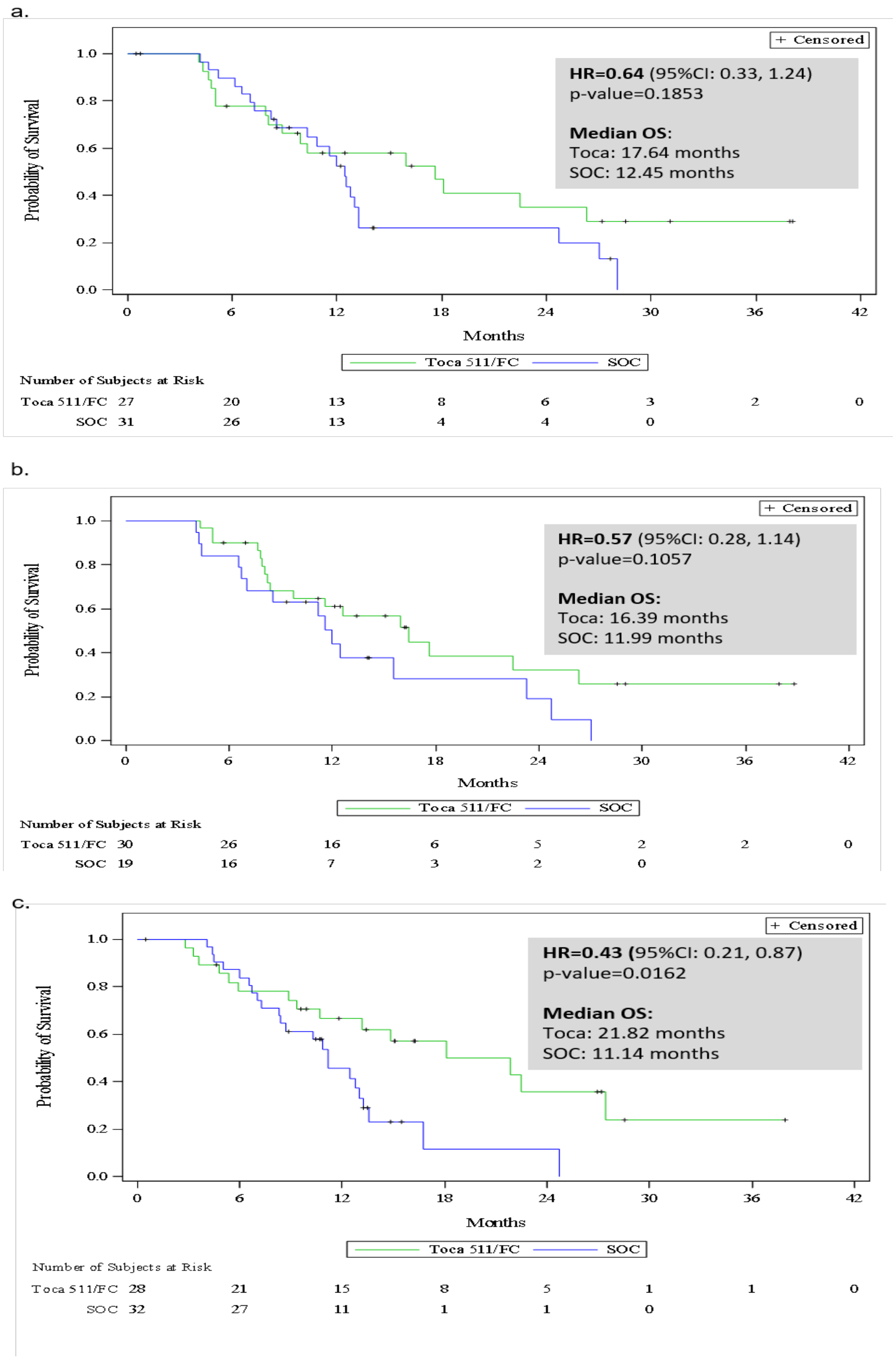
eTable 3. Secondary endpoints for Toca 5 randomized clinical trial.
eTable 4. Baseline patient characteristics for patients in second recurrence.
eTable 5. Toca 511 viral RNA and DNA signal by cycle.
eTable 6. Proposed etiology for SBS Mutational signatures, from COSMIC.



Number of Subjects at Ri

Toca-Cyc1	45	33	17	4	2	0				
Toca-Cyc2	63	61	38	15	6	4	0			
Toca-Cyc3	33	33	21	12	5	2	2	1	1	0
Toca-Cyc4+	56	56	51	33	24	18	15	9	5	3	0

eFigure 1. Kaplan-Meier overall survival curve with Toca FC cycle numbers.

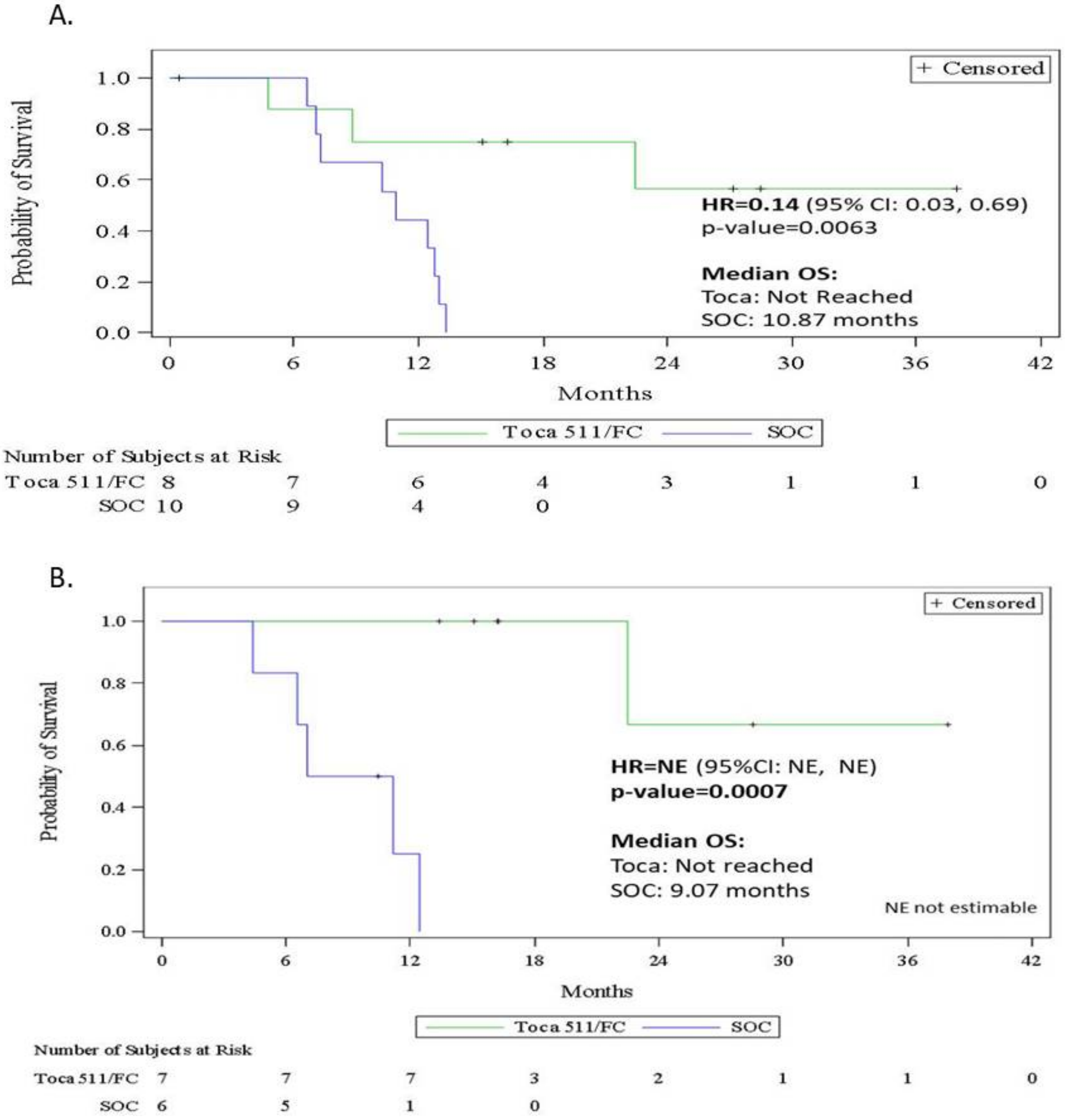


eFigure 2. Kaplan-Meier overall survival curve for pre-planned subpopulations

A. IDH1 mutant subpopulation

B. AA subpopulation

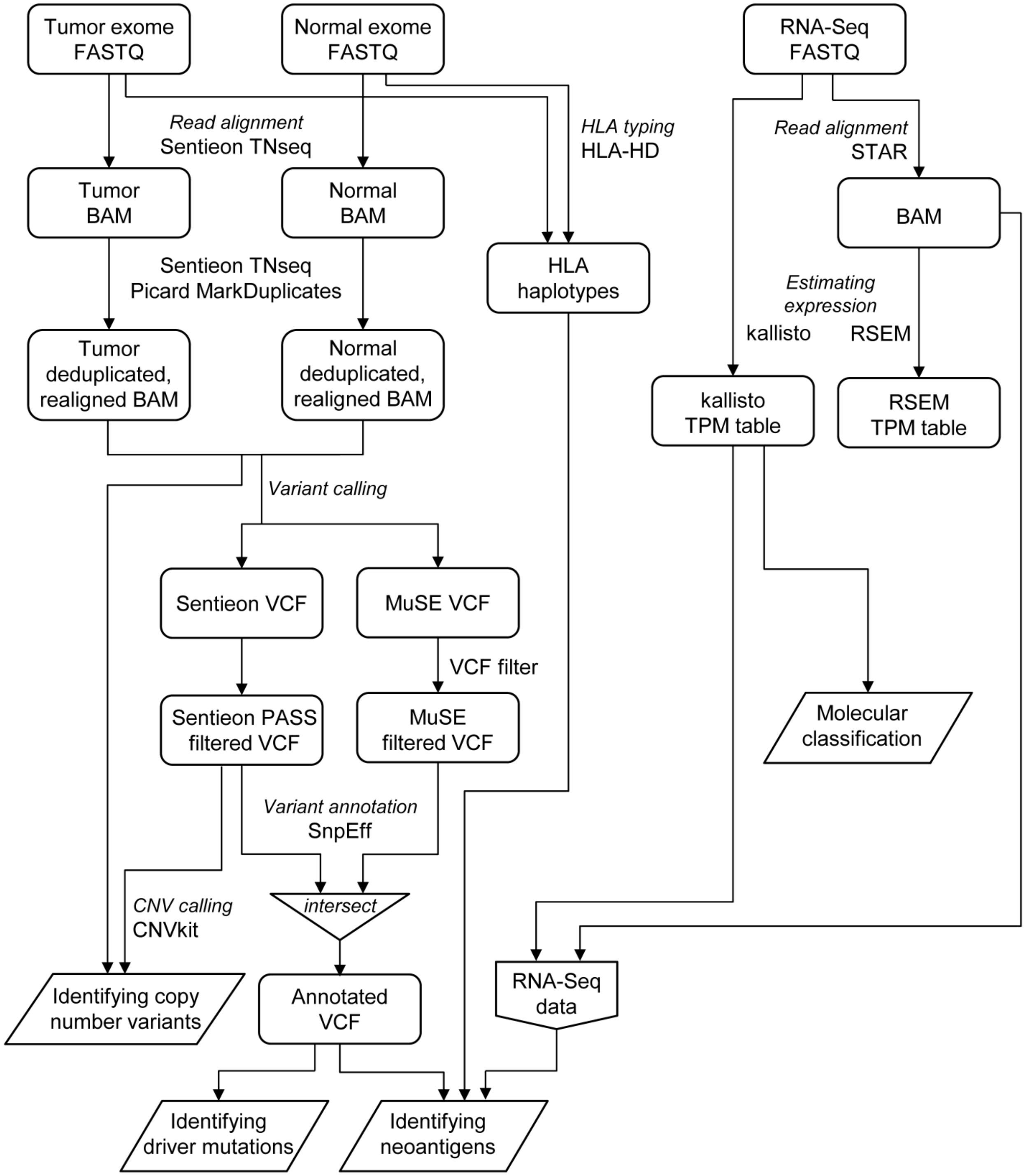
C. Second recurrence



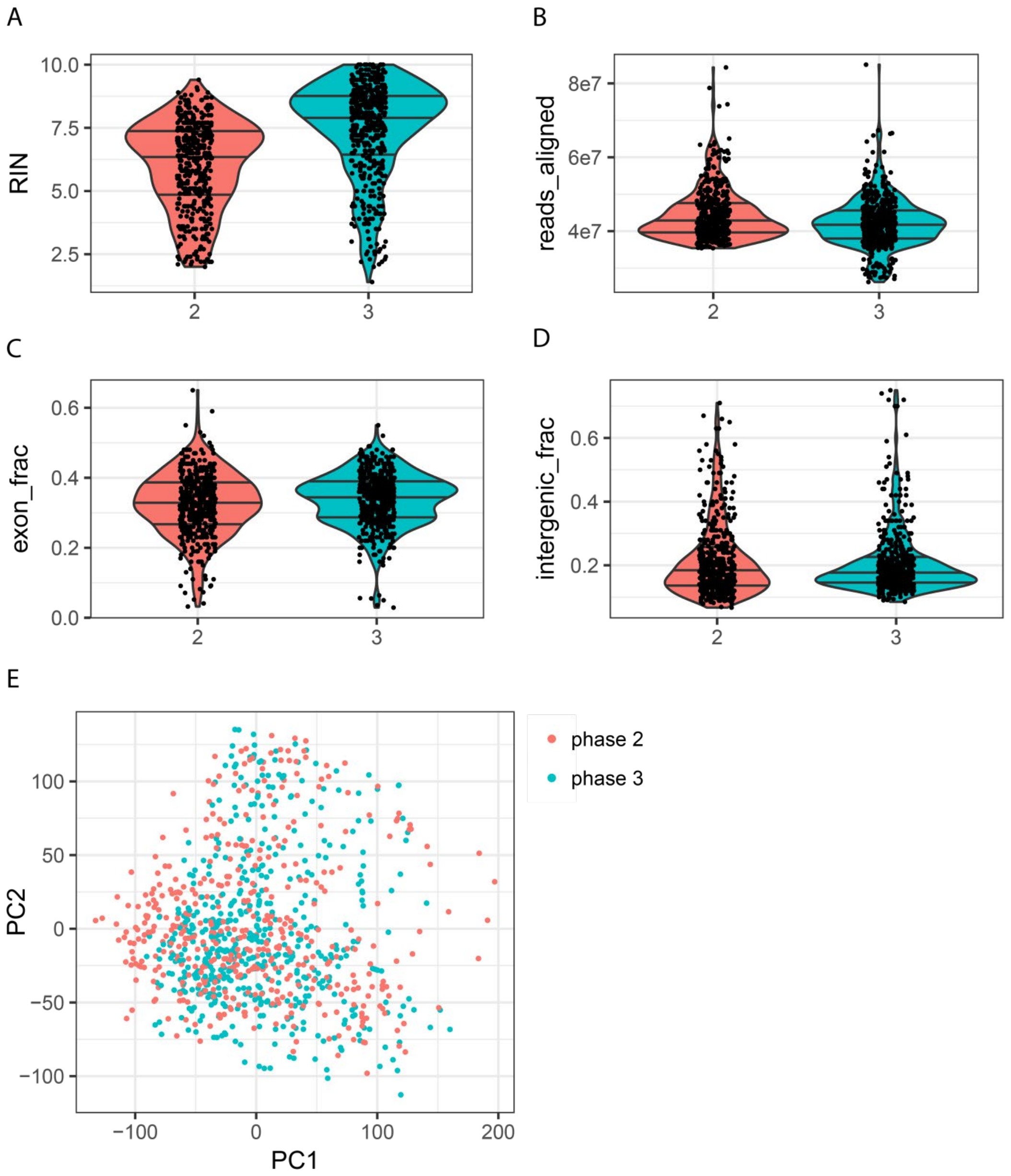
eFigure 3. Kaplan-Meier overall survival curve for subpopulations within the second recurrence.

A. IDH1 mutant subpopulation

B. AA subpopulation



eFigure 4. Workflow for whole exome sequencing and RNA sequencing data analyses.



eFigure 5. Quality assessment of phase 2 and phase 3 portions of Toca 5 tumor RNA sequencing.

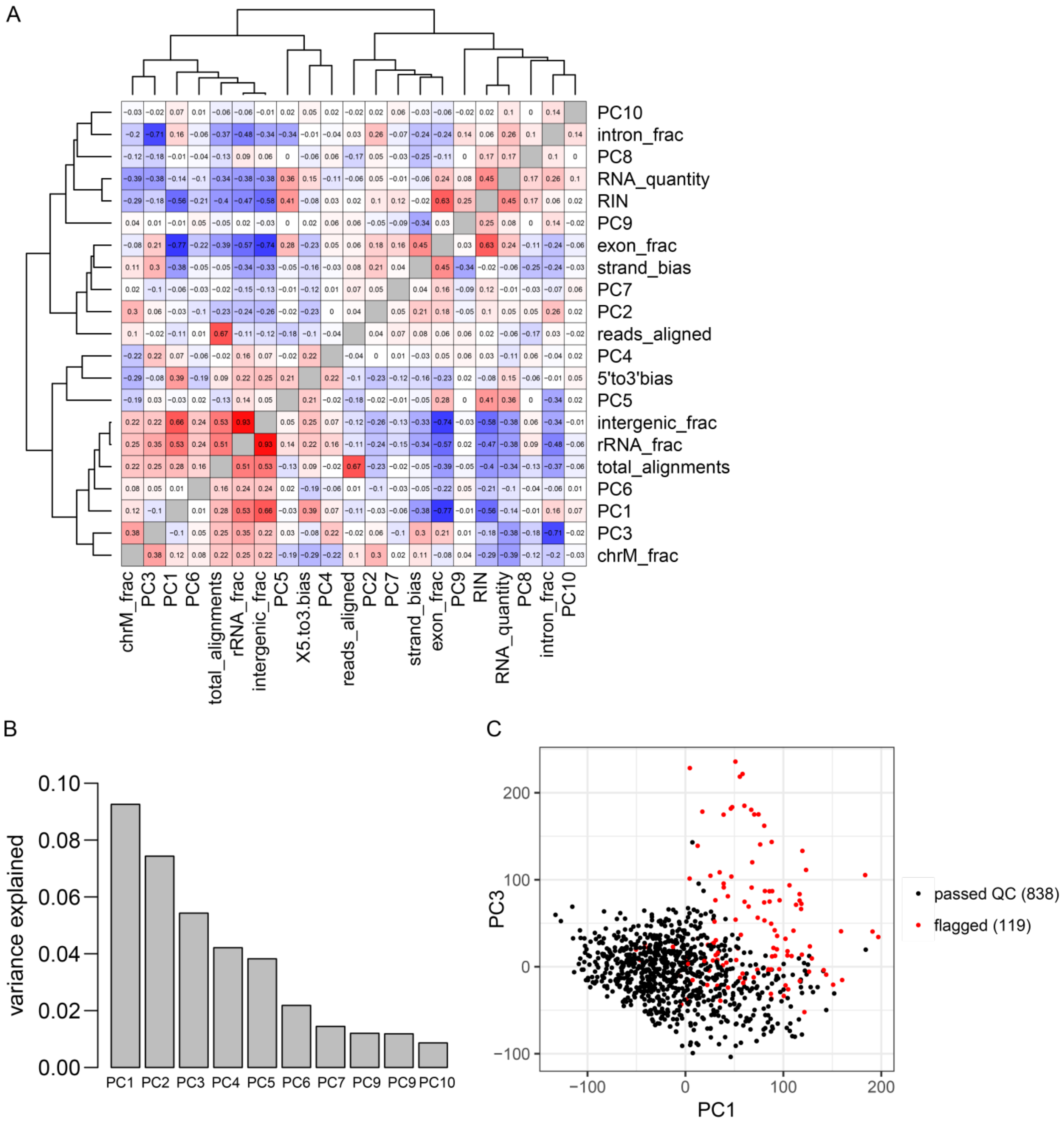
(A) Violin plots of RIN values from phase 2 (red) and phase 3 (blue) tumor samples.

(B) Violin plots of the number of reads that mapped to reference mRNA set and/or human genome from phase 2 and phase 3 tumor samples.

(C) Violin plots of the fraction of mapped reads that mapped to annotated exons from phase 2 and phase 3 tumor samples.

(D) Violin plots of the fraction of mapped reads that mapped to “intergenic regions” from phase 2 and phase 3 tumor samples.

(E) Scatterplot of the first two principal components of the RNA sequencing dataset. Samples are color-coded by phase.

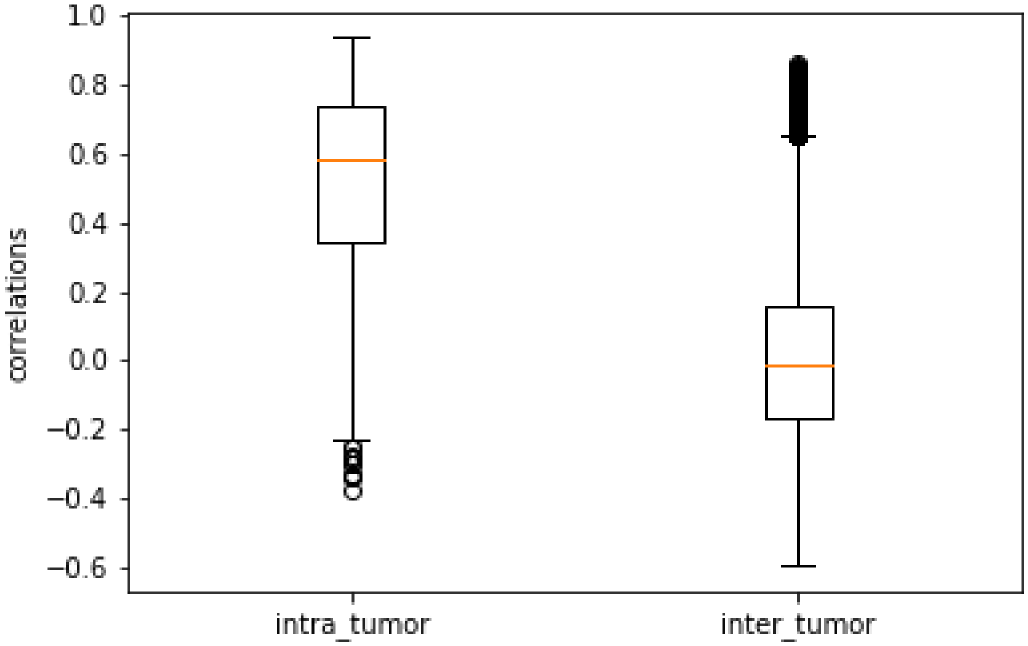


eFigure 6. Assessment of Toca 5 RNA sequencing data quality.

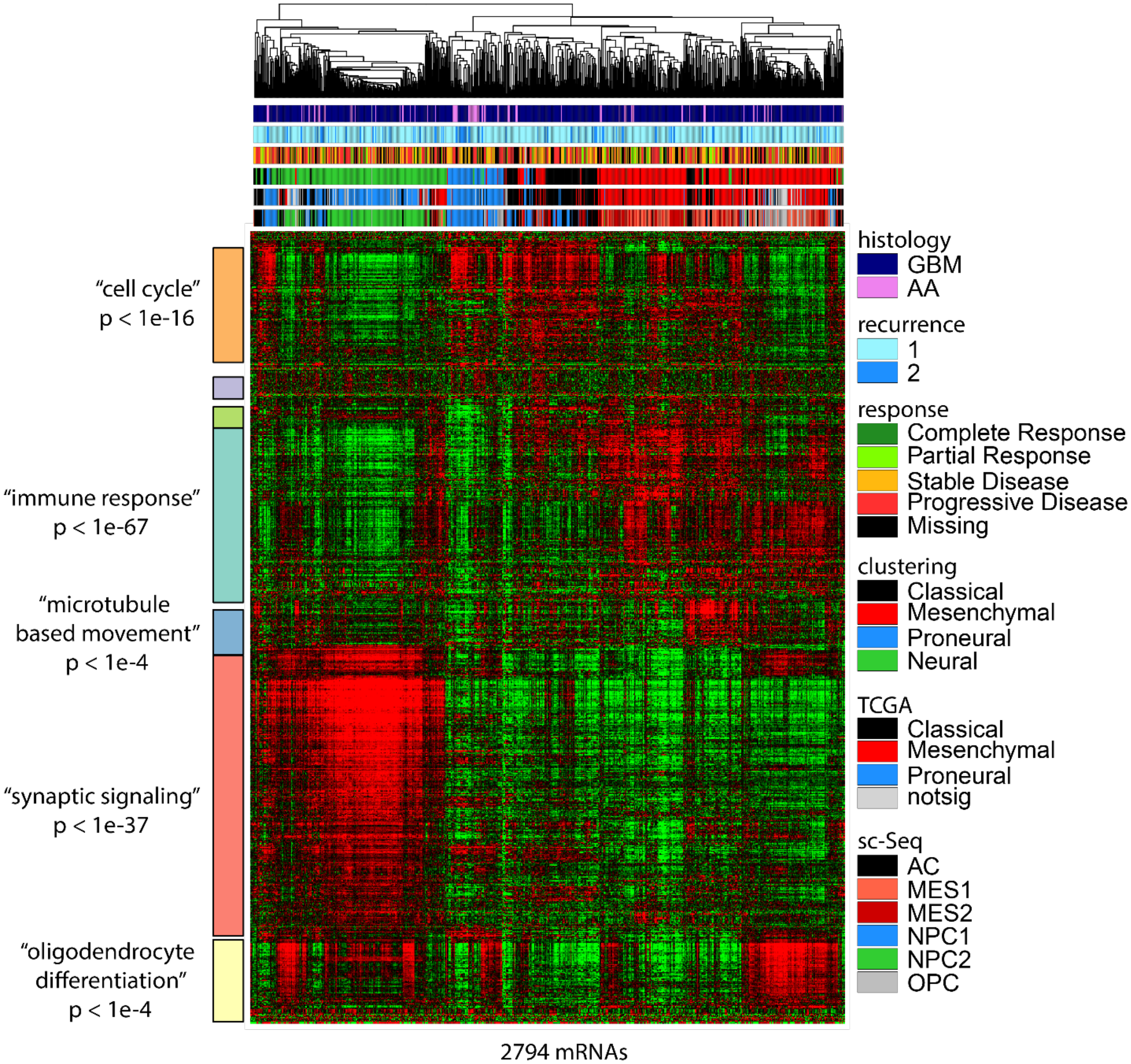
(A) Heatmap representation of pairwise Spearman rank correlations between various RNA quality and sequencing metrics and the first ten principal components of the RNA expression dataset.

(B) Barplot showing the variance explained by first ten principle components of Toca 5 RNA sequencing dataset.

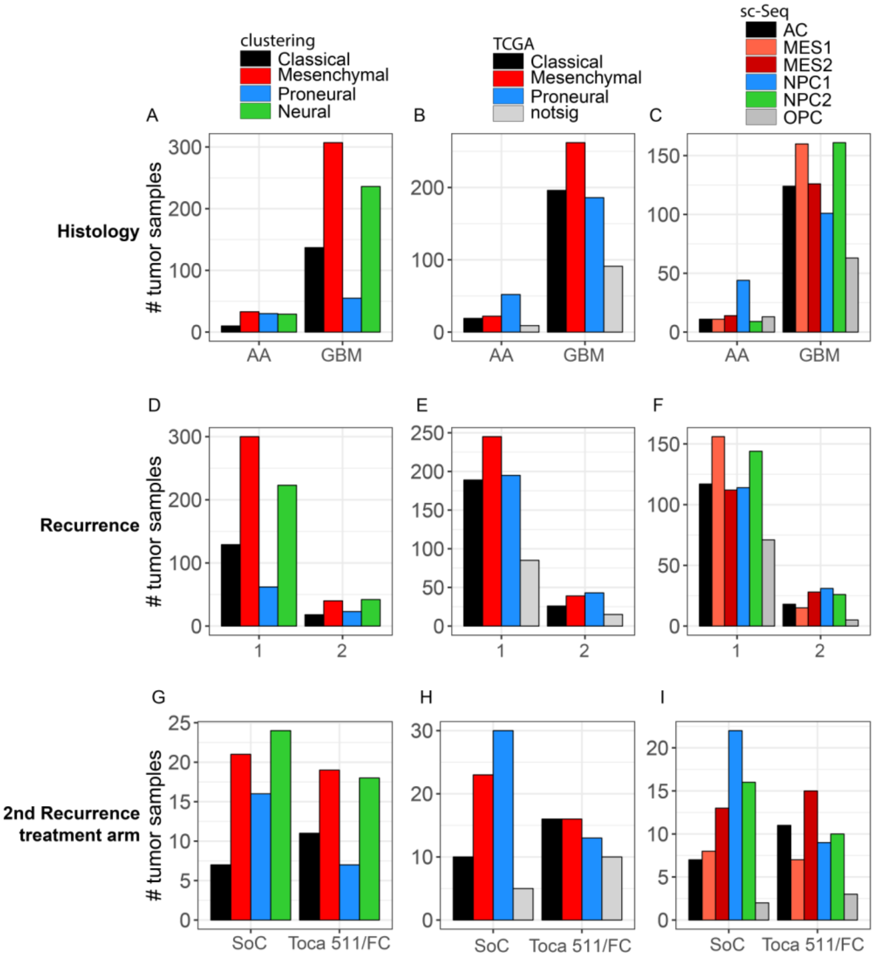
(C) Scatterplot between first and third principle components in RNA expression dataset highlighting that outlier samples tend to have anomalously low RNA quality (RIN < 4) and/or low fraction of reads that map to annotated protein coding exons (exon_frac < 0.2). Thus samples that met these criteria were flagged (red) and not used for subsequent analyses.



eFigure 7. Boxplot comparing intra-tumor Pearson correlations in mRNA expression versus inter-tumor Pearson correlations in mRNA expression.



eFigure 8. mRNA expression profiles of Toca 5 tumors.
Unsupervised hierarchical clustering of mRNA expression across 838 Toca 5 tumor samples from 319 biopsies. Many of the 2794 mRNAs in this set (selected for standard deviation in top quartile) fell into one of seven distinct clusters (marked by colored bars to the left of the heatmap) that included mRNAs with common functional themes, as indicated. Above the heatmap we included the cognate patient's histology, recurrence and treatment response. Additionally, we show molecular classifications of the tumor samples using three different methods (4-class system based on clustering, 3-class system from TCGA using ssGSEA and 6-class system based on single cell RNA sequencing).



eFigure 9. Molecular classification of Toca 5 tumor samples.

(A) Bar plot showing the number of AA and GBM samples in each molecular subtype as defined by clustering.

(B) Bar plot showing the number of AA and GBM samples in each molecular subtype as defined by 3-class TCGA method.

(C) Bar plot showing the number of AA and GBM samples in each molecular subtype as defined by single-cell RNA sequencing.

(D) Bar plot showing the number of 1st and 2nd recurrence samples in each molecular subtype as defined by clustering.

(E) Bar plot showing the number of 1st and 2nd recurrence samples in each molecular subtype as defined by 3-class TCGA method.

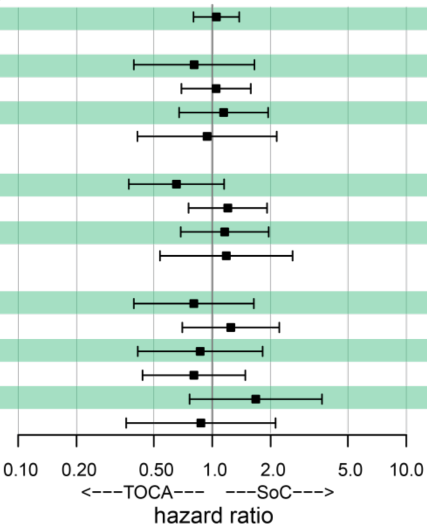
(F) Bar plot showing the number of 1st and 2nd recurrence samples in each molecular subtype as defined by single-cell RNA sequencing.

(G) Bar plot showing the number of standard of care and Toca 511 and Toca FC treatment arm samples from 2nd recurrence tumors in each molecular subtype as defined by clustering.

(H) Bar plot showing the number of standard of care and Toca 511 and Toca FC treatment arm samples from 2nd recurrence tumors in each molecular subtype as defined by 3-class TCGA method.

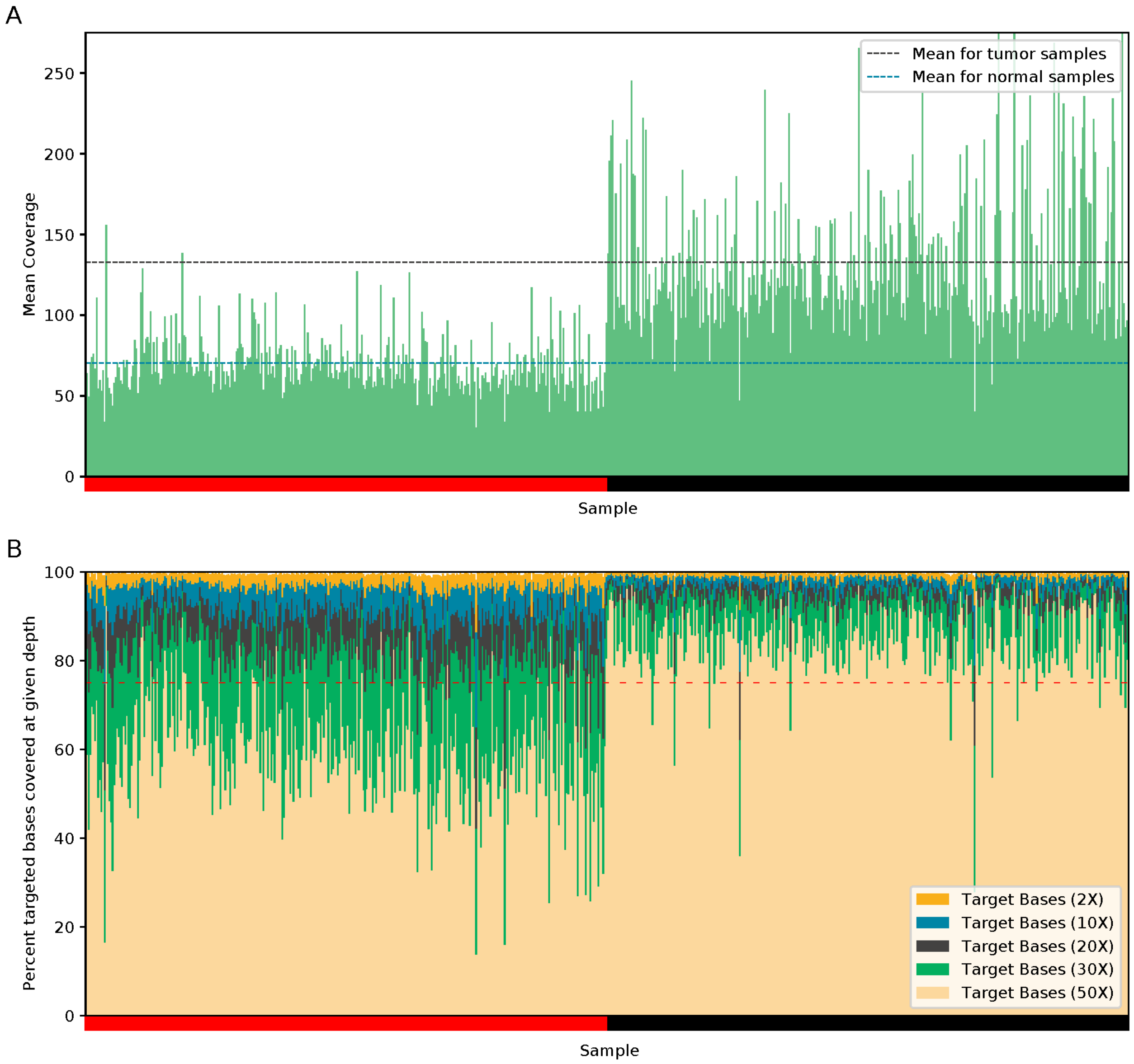
(I) Bar plot showing the number of standard of care and Toca 511 and Toca FC treatment arm samples from 2nd recurrence tumors in each molecular subtype as defined by single-cell RNA sequencing.

group	subgroup	mOS Toca	mOS SoC	Events/# Toca	Events/# SoC	HR	p−value
overall		11.5	12.5	116/169	95/148	1.05	0.7
clustering	Classical	18.3	15.7	13/26	21/31	0.81	0.6
	Mesenchymal	10.7	10.7	56/76	39/58	1.05	0.8
	Neural	13.7	14.2	35/51	23/43	1.14	0.6
	Proneural	8.9	12.5	12/16	12/16	0.94	0.9
TCGA	Classical	11.6	11.6	25/37	27/40	0.65	0.1
	Mesenchymal	13.7	13.6	42/63	32/52	1.2	0.4
	Proneural	9.9	12.5	34/48	25/37	1.16	0.6
	not significant	11.3	13.2	15/21	11/19	1.18	0.7
sc-Seq	AC	15.9	12.8	14/21	21/32	0.8	0.5
	MES1	10.4	13.2	25/35	23/36	1.25	0.5
	MES2	11.5	11.8	20/30	12/22	0.87	0.7
	NPC1	9.9	12.4	21/32	21/26	0.8	0.5
	NPC2	11.1	14.2	17/27	11/24	1.68	0.2
	OPC	13.5	9.8	19/24	7/8	0.87	0.8



eFigure 10. Relationships between patients' tumor molecular classification and survival.

For this analysis the tumor sample from each biopsy with the highest RNA quality (RIN + fraction of reads mapping to exons) was used. The forrest plots show the hazard ratio between the Toca 511/FC arm and standard of care along with 95% confidence intervals.

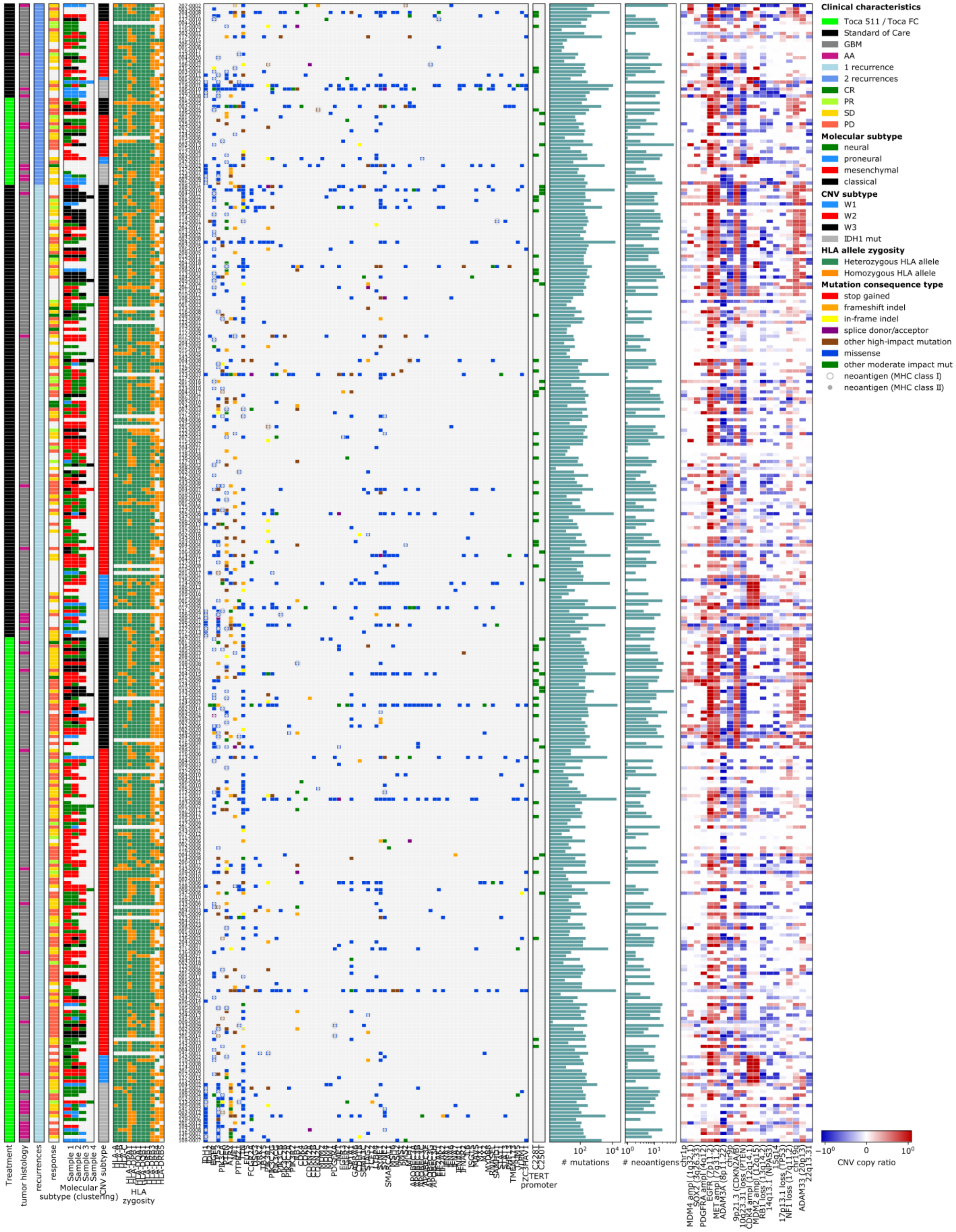


eFigure 11. Sequencing coverage metrics across targeted regions for normal and tumor samples.

The target region comprises 66MB as defined by the Agilent SureSelect V6+ kit used for target capture.

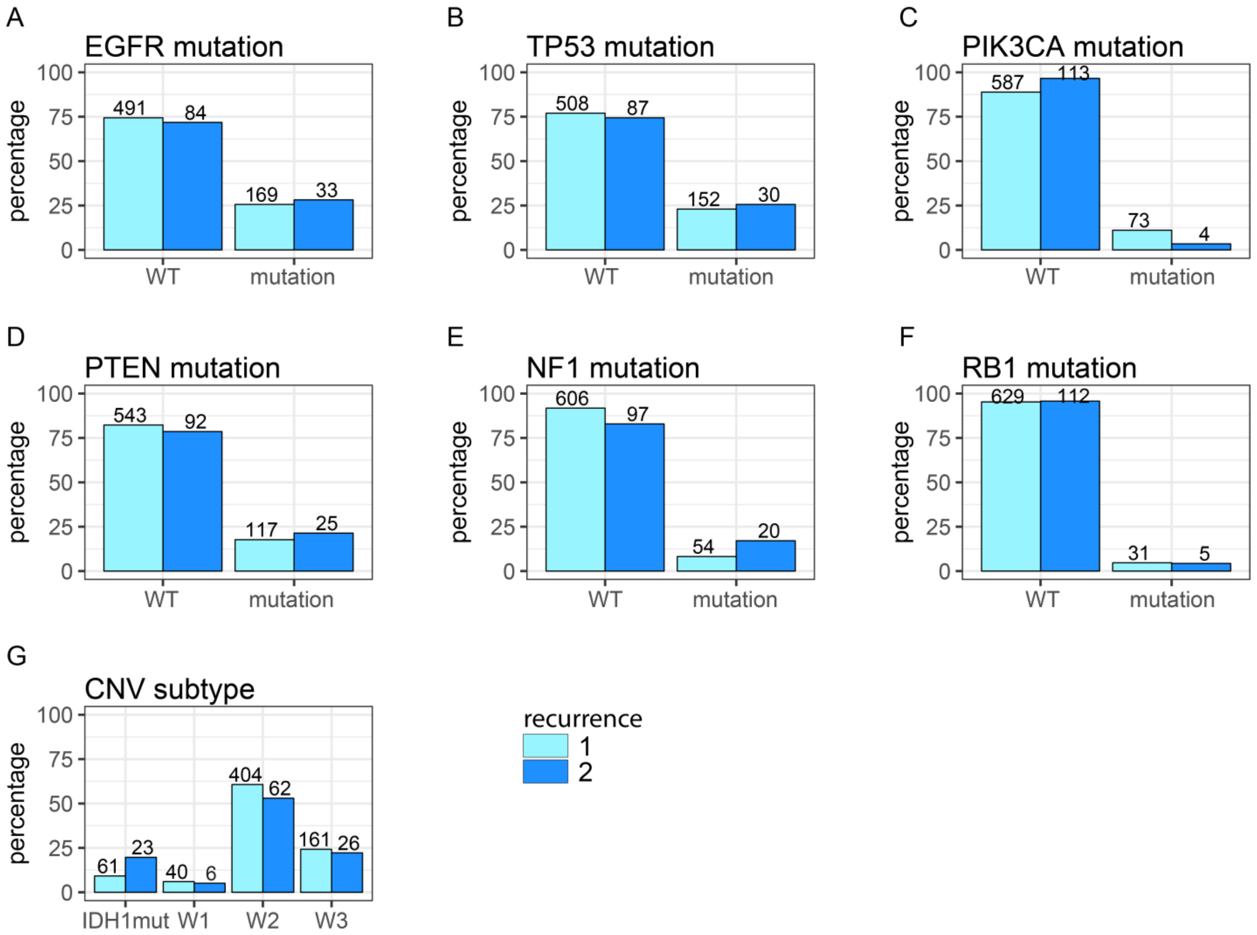
(A) Mean depth of coverage across targeted regions. Samples on the left half are from DNA isolation from blood (normal) and samples on the right half are from DNA isolated from tumor.

(B) Percentage of targeted bases covered at a given depth. The red dashed line indicates the target coverage threshold. For blood samples (left side of plot), the goal was at least 75% of targeted bases covered at 30x, and for tumor samples, at least 75% of targeted bases covered at 50x.


Clinical characteristics
Toca 511 / Toca FC
Standard of Care
GBM
AA
1 recurrence
2 recurrences
CR
PR
SD
PD
Molecular subtype
neural
proneural
mesenchymal
classical
CNV subtype
W1
W2
W3
IDH1 mut
HLA allele zygosity
Heterozygous HLA allele
Homozygous HLA allele
Mutation consequence type
stop gained
frameshift indel
in-frame indel
splice donor/acceptor
other high-impact mutation
missense
other moderate impact mut
neoantigen (MHC class I)
neoantigen (MHC class II)
Treatment
tumor histology
recurrences
response
Sample 1
Sample 2
Sample 3
Sample 4
Molecular subtype (clustering)
CNV subtype
HLA-A
HLA-B
HLA-C
HLA zygosity
IDH1
TP53
PIK3CA
ATRX
PTEN
SMARCA4
ZC3HAV1
TERT promoter
C228T
C250T
10^2
10^4
mutations
10^0
10^1
neoantigens
chr1p
MDM4 ampl (1q32.1)
SOX2 (3q26.33)
PDGFRA ampl (4q12)
EGFR (7p11.2)
MET ampl (7q31.2)
ADAM3A (8p11.22)
chr9p
9p21.3 (CDKN2A/B)
10q23.31 loss (PTEN)
CDK4 ampl (12q14.1)
MDM2 ampl (12q15)
RB1 loss (13q14.2)
14q13.1 (NPAS3)
15q14
17p13.1 loss (TP53)
NF1 loss (17q11.2)
chr19q
ADAM33 (20p13)
22q13.31
-10^0
0
10^0
CNV copy ratio

eFigure 12. Summary of DNA sequencing analyses of Toca 5 tumors.

(left to right) Patients are ordered by number of recurrences and CNV subtype. The first four columns show clinical features, including eligibility for Phase III trial (Toca 5 eligible – green), tumor grade as determined by a clinical site pathologist (grade IV – grey, grade III – black), number of recurrences, response (CR – green, PR – yellow-green, SR – yellow, PD – orange). The next 6 columns summarize molecular subtype calculated from RNA sequencing results: molecular subtype (mesenchymal – red, classical – black, neural – green, proneural – blue). Molecular subtype was calculated according to three different methods, and different colors in the two columns for each method indicate that the classifications differed between samples from the same patient. The next column indicates CNV subtype, which was determined from exome data. The next column shows whether the patient was homozygous or heterozygous for an HLA class I or class II allele (heterozygous – green, homozygous – orange, indeterminate – yellow, data not available – light grey). The next block of columns summarizes mutations of key glioma-associated genes as determined by exome sequencing. Mutations are shown if a variant was called in either tumor sample by either variant caller. In case of multiple annotations, the most deleterious predicted consequence is indicated. The next two columns indicate whether there were *TERT* promoter mutations (C228T or C250T) present, if coverage was sufficient. The first bar plot shows the total number of high confidence mutations, and specifically, the union of moderate and high impact variants called from exome data by either variant caller, and in case of multiple samples, values indicate the mean of multiple samples. The second bar plots shows the number of predicted neoantigens based on binding of mutated peptides to HLA class II molecules. The last bar plot indicates the range of tumor purity estimates obtained by two different tools for each sample from a given patient. The heat map shows copy number variants commonly amplified or lost in GBM. Shades of red (gains) and blue (losses) indicate log2 ratios and are shown on a log scale. If a chromosomal region contains a gene associated with GBM, the gene name is given in parentheses. Abbreviations: OS, overall survival; CNV, copy number variant; mut, mutation; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; HLA, human leukocyte antigen.



eFigure 13. Comparison of genetic profiles from 1st and 2nd recurrence tumors.

(A) Barplot showing the percentage of patients in 1st (lighter blue) or 2nd (darker blue) recurrence with EGFR mutation (right). The numbers above the barplots are the number of patients represented in the bar. There is not a significantly higher fraction of patients with EGFR mutations in 2nd recurrence compared to 1st recurrence (Fisher Exact test).

(B) Same as (A) except for TP53 mutations. No significant difference.

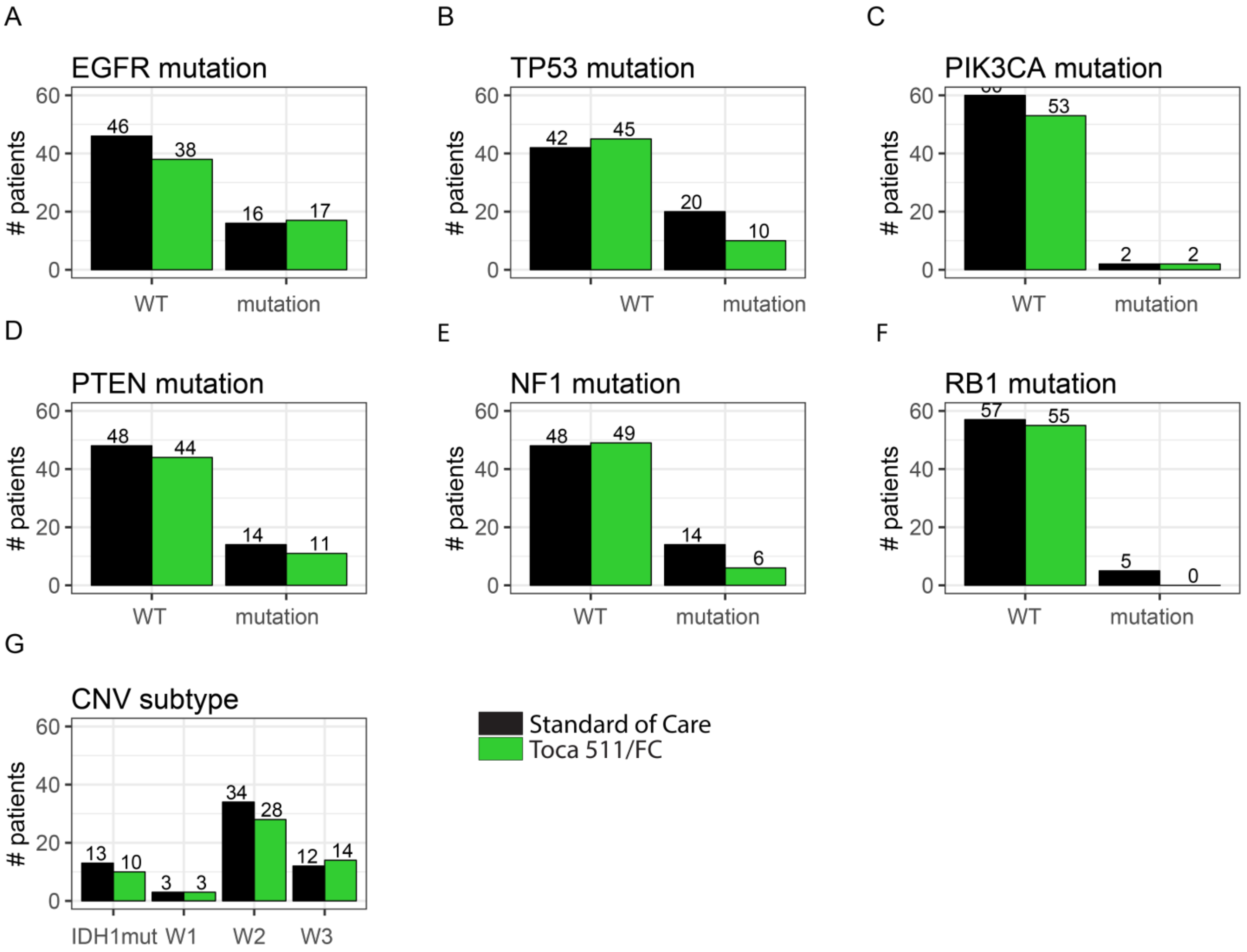
(C) Same as (A) except for PIK3CA mutations. P = 0.007.

(D) Same as (A) except for PTEN mutations. No significant difference.

(E) Same as (A) except for NF1 mutations. P = 0.005

(F) Same as (A) except for RB1 mutations. No significant difference.

(G) Same as (A) except for CNV subtype. P = 0.02.



eFigure 14. Comparison of genetic profiles from 2nd recurrence patients in the two treatment arms.

(A) Barplot showing the number of 2nd recurrence patients in standard of care (black) or Toca 511/FC (green) arms with EGFR mutations (right). No significant difference.

(B) Same as (A) except for TP53 mutations. No significant difference.

(C) Same as (A) except for PIK3CA mutations. No significant difference.

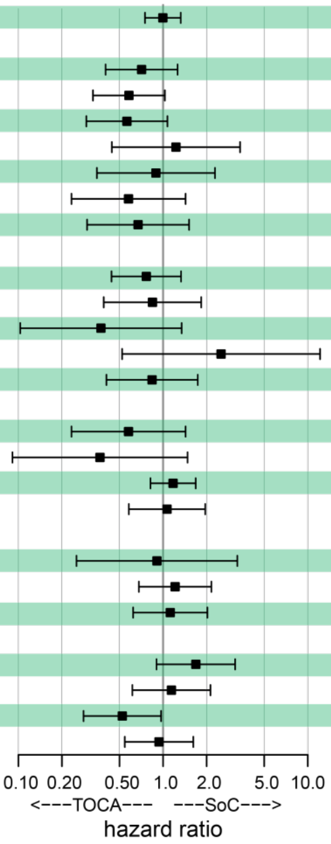
(D) Same as (A) except for PTEN mutations. No significant difference.

(E) Same as (A) except for NF1 mutations. No significant difference.

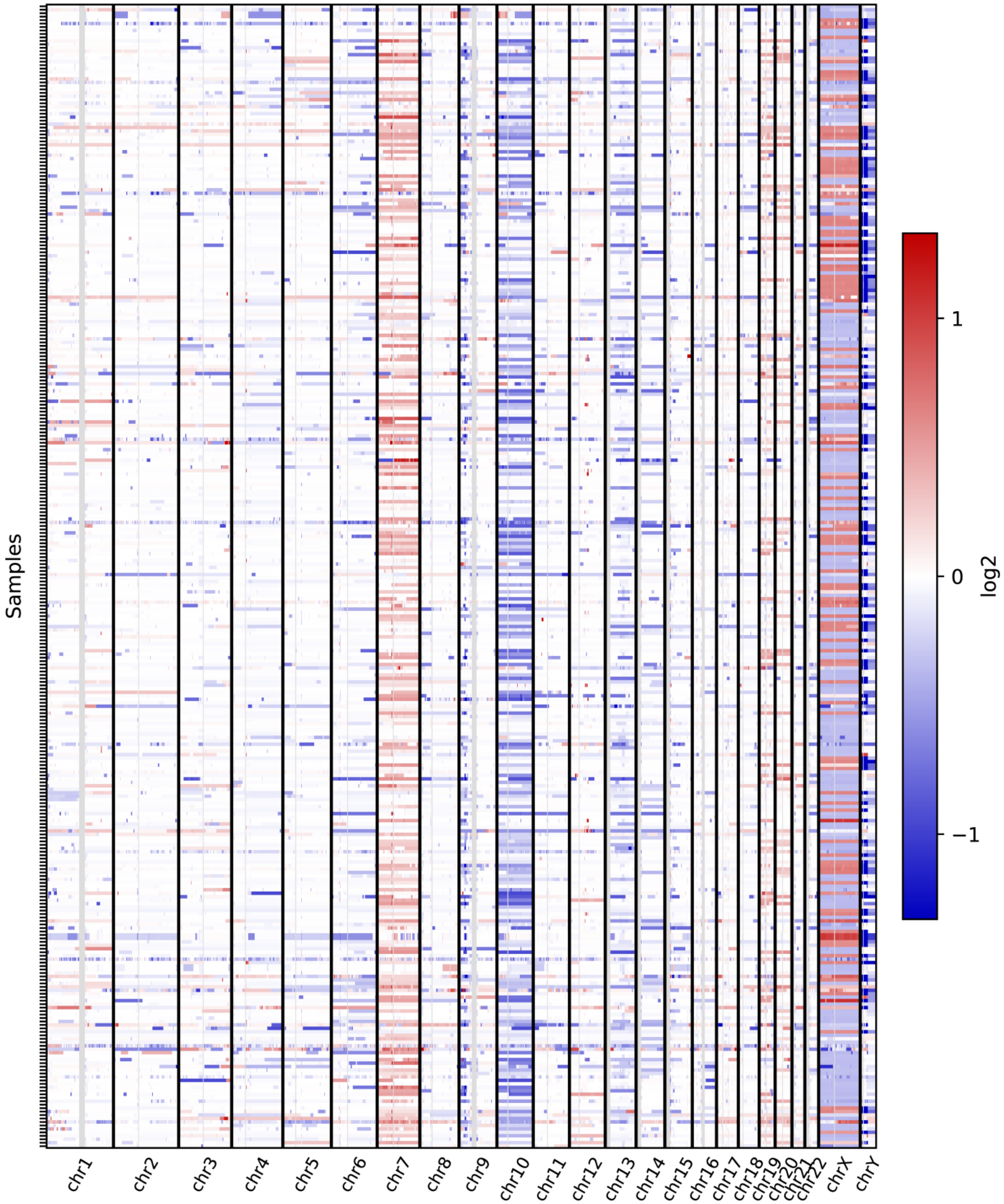
(F) Same as (A) except for RB1 mutations. No significant difference.

(G) Same as (A) except for CNV subtype. No significant difference.

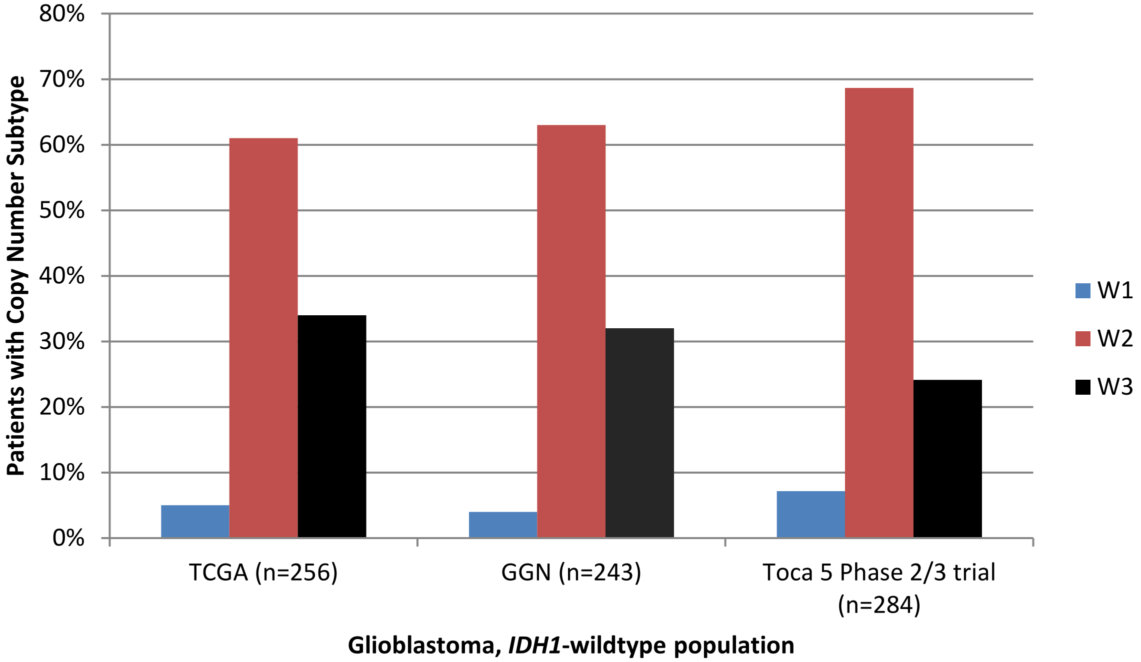
group	subgroup	mOS Toca	mOS SoC	Events/# Toca	Events/# SoC	HR	p−value
overall		11.6	12.4	107/157	86/133	1	1
mutated gene	EGFR	10.3	9.1	26/36	24/33	0.71	0.2
	TP53	14.1	10.7	24/38	25/32	0.58	0.06
	PTEN	10.1	8.9	22/30	22/27	0.56	0.08
	NF1	9.4	12.5	6/9	10/18	1.23	0.7
	PIK3CA	17.6	12.2	7/11	13/16	0.89	0.8
	IDH1	17.6	12.5	10/19	9/11	0.58	0.2
	ATRX	15.2	12.3	13/17	12/14	0.67	0.3
gene gain or loss	EGFR amp > 2−fold	13.2	14.4	30/46	24/40	0.77	0.3
	EGFR amp > 20−fold	9.4	10.7	15/22	12/15	0.84	0.7
	CDK4 amp > 2−fold	21.6	12.5	6/14	6/7	0.37	0.1
	PDGFRA amp > 2−fold	6.4	10.7	8/11	2/4	2.52	0.2
	PTEN deletion > 1.5 fold	10.2	10.2	18/27	13/19	0.84	0.6
CNV class	IDH1 mutant	17.6	12.5	10/19	9/11	0.58	0.2
	W1	21.6	10.6	3/10	6/8	0.37	0.1
	W2	10.7	11.6	68/93	53/85	1.17	0.4
	W3	10.3	15.9	26/35	18/29	1.07	0.8
mutation burden	hypermutant	9.9	12.4	5/7	6/7	0.91	0.9
	MHCI neoantigens ≥ 6	8.6	10.1	24/32	25/30	1.21	0.5
	MHCII neoantigens ≥ 14	7.8	11.1	24/32	22/29	1.12	0.7
tumor purity	1st quartile	10.7	15.7	25/36	20/37	1.68	0.1
	2nd quartile	13.5	12.8	28/44	16/28	1.14	0.7
	3rd quartile	14.1	10.7	23/35	26/37	0.52	0.04
	4th quartile	10.2	11.8	31/42	23/30	0.94	0.8



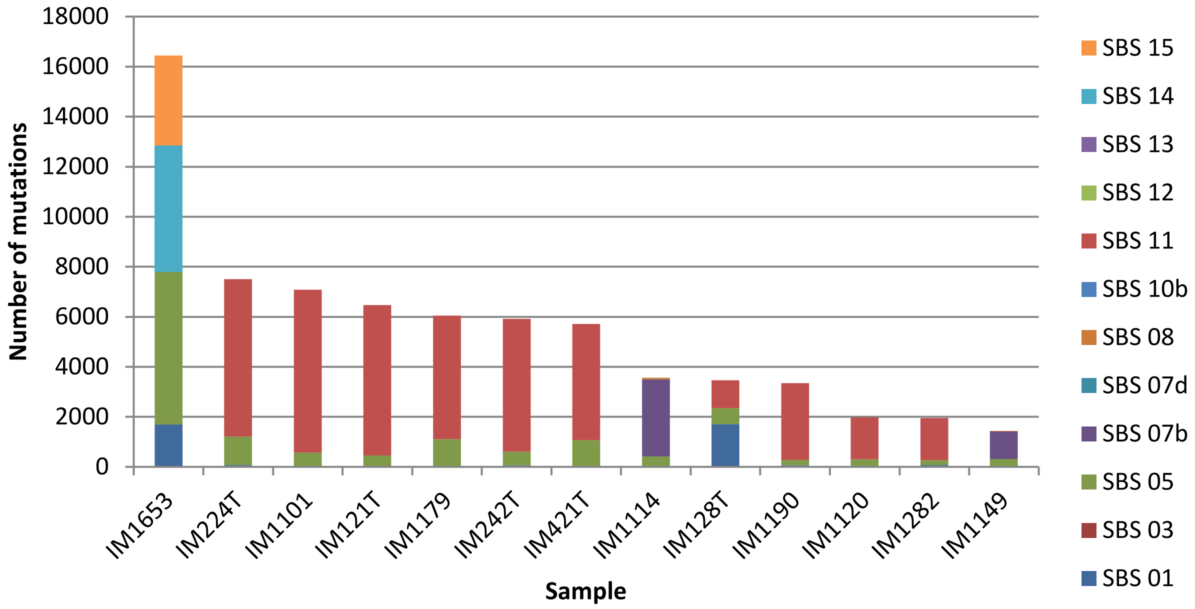
eFigure 15. Relationships between tumor DNA alterations and survival.
For the mutation burden group, hypermutant tumor is defined as having > 1000 high-confidence mutations; for tumors with ≥ 6 MHC1 neoantigens or ≥ 14 MHCII neoantigens, hypermutant tumors were excluded. The forrest plots show the hazard ratio between the Toca 511/FC arm and standard of care along with 95% confidence intervals.



eFigure 16. Tumor copy number variants inferred from exome sequencing data.
Red and blue lines indicate gain and loss events, respectively.

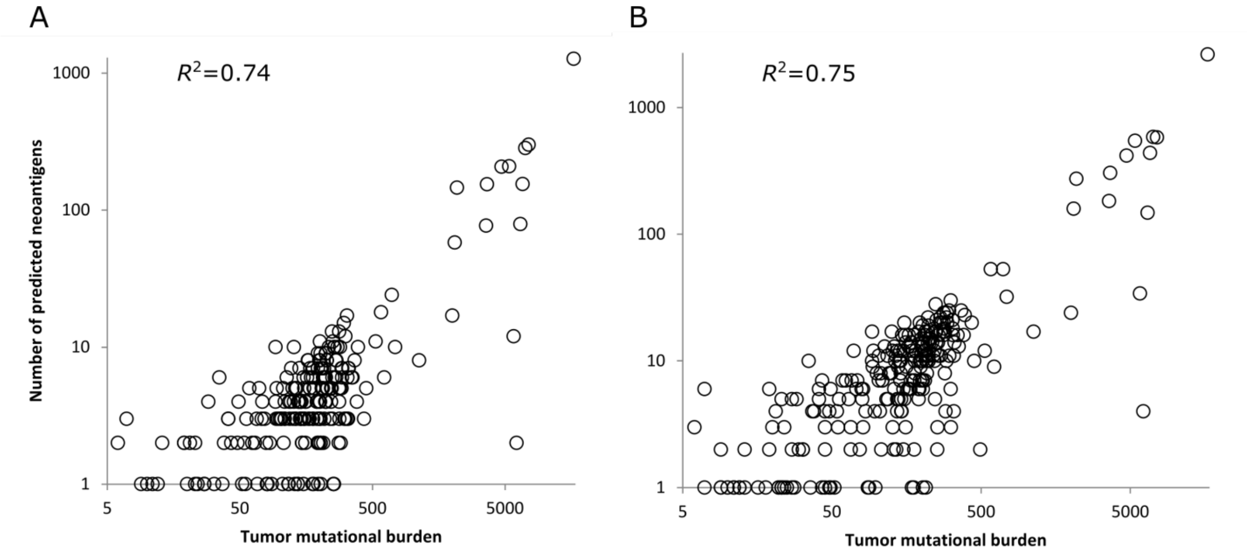


eFigure 17. CNV subtypes for IDH1-wildtype GBM patients.
Barplots show the distribution of CNV subtypes in IDH1-wildtype Toca 5 patients compared to those of two published GBM cohorts from The Cancer Genome Atlas (TCGA) and German Glioma Network (GGN). The W3 subtype (best survival) is underrepresented in Toca 5 relative to the other two cohorts.



eFigure 18. Mutational signatures in hypermutated tumors.

Barplots showing the number of high-confidence mutations color-coded by the predicted source, as determined by SigProfiler. Tumors with >1000 mutations are shown. The proposed etiology for each SBS signature, from COSMIC, is defined in eTable 6. SBS 11 exhibits a mutational pattern resembling that of alkylating agents. Patient histories indicate an association between previous treatment with the alkylating agent temozolomide and SBS11 mutations.

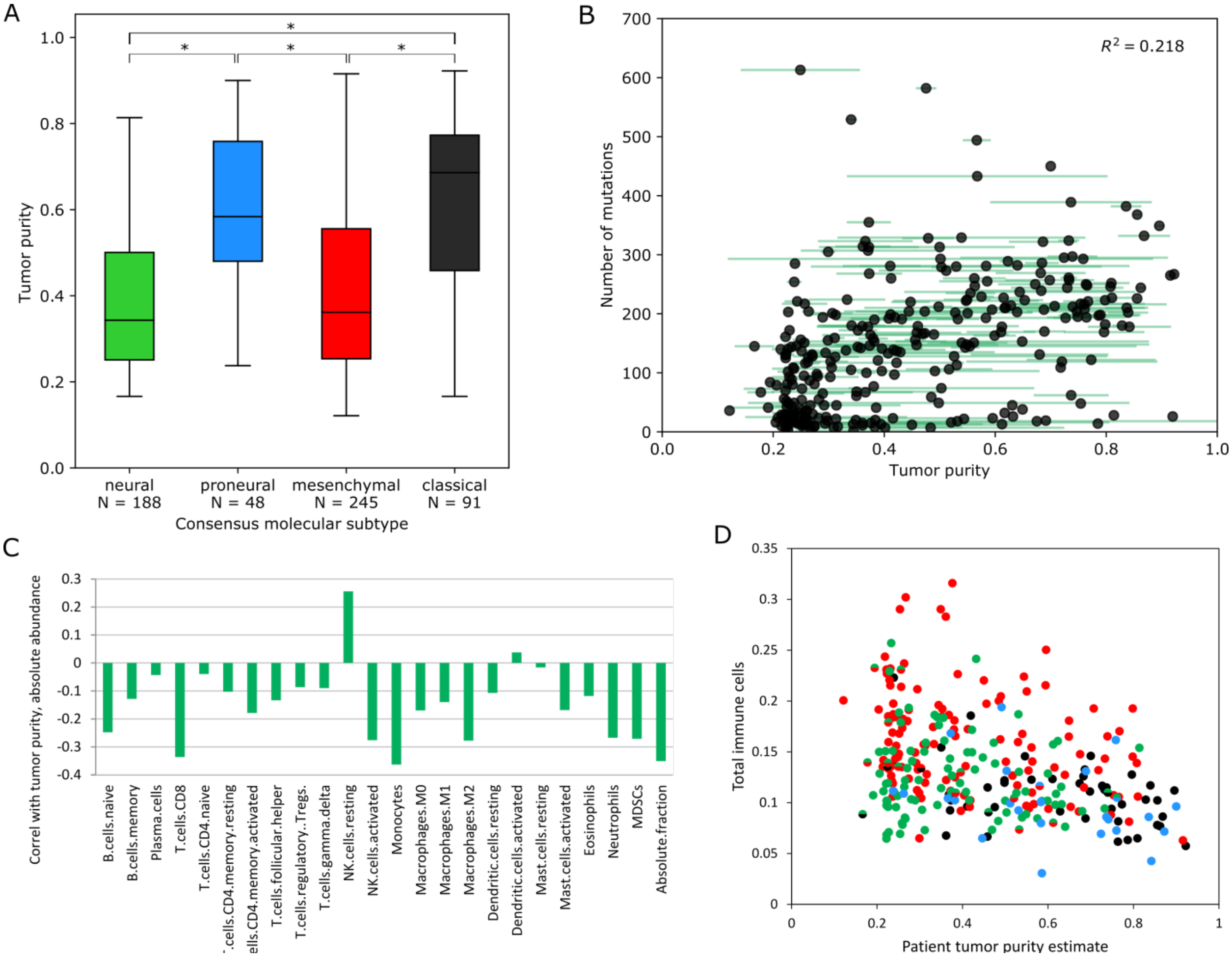


eFigure 19. Number of predicted tumor neoantigens correlates with tumor mutational burden.

The number of neoantigens predicted to bind to a patient's HLA molecules correlates with the number of high-confidence SNVs and indels called from exome data.

(A) Neoantigens predicted to bind to MHC class I molecules only.

(B) Neoantigens predicted to bind to MHC class II molecules only.



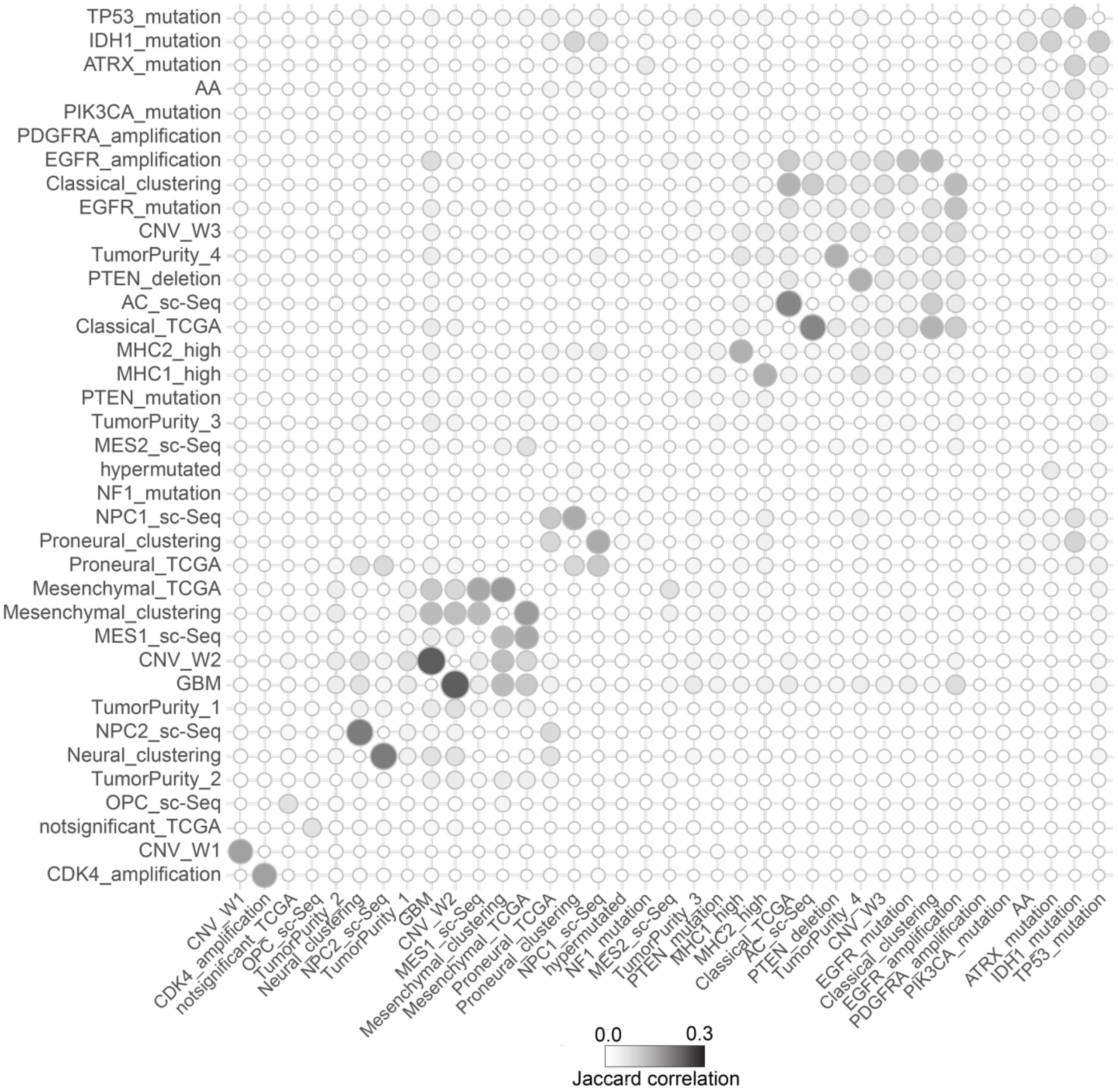
eFigure 20. Relationships between tumor purity and molecular subtype.

(A) Box plot showing tumor purity estimates from whole-exome sequencing data categorized according to consensus molecular subtype based on RNA sequencing. Asterisks indicate p-value < 1e-8.

(B) Scatterplot showing the correlation between tumor purity estimates (mid-range) and the mean number of somatic mutations. Green horizontal bars indicate the range of possible tumor purity estimates for a given patient, as calculated by two different tools.

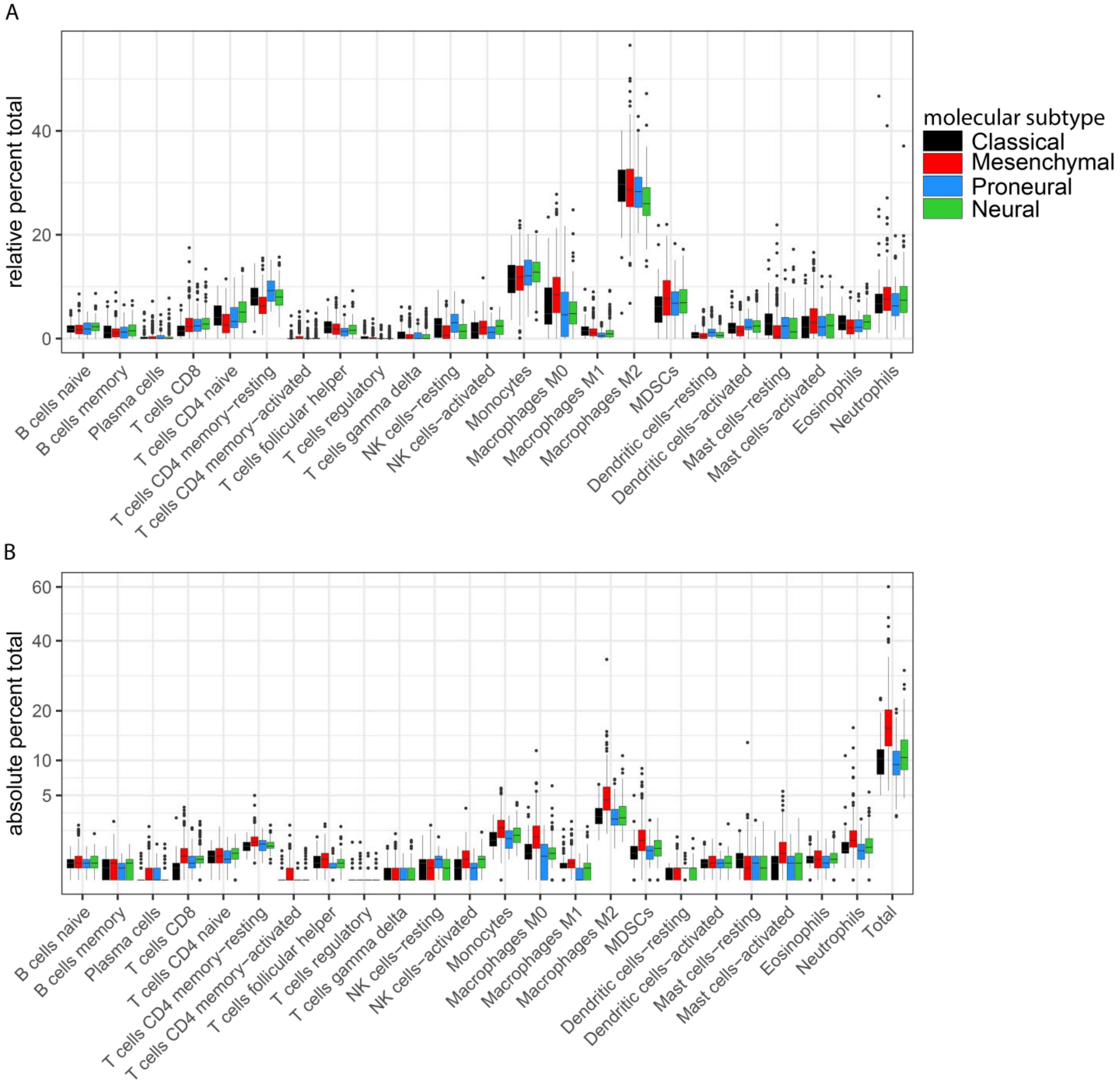
(C) Barplots showing correlations between tumor purity (mid-range) and absolute levels of various immune cell types in tumor samples as measured by iSort.

(D) Scatterplot showing details of the correlation between tumor purity and total absolute levels of immune cells that was shown in (C). Each point represents a single patient, points are colored according to the consensus molecular subtype; split colors indicate a tie between two molecular subtypes. Colors indicate the following molecular cell types: green, neural; blue, proneural; red, mesenchymal; black, classical.



eFigure 21. Pairwise correlations among genetic alterations, RNA subtypes and histology.

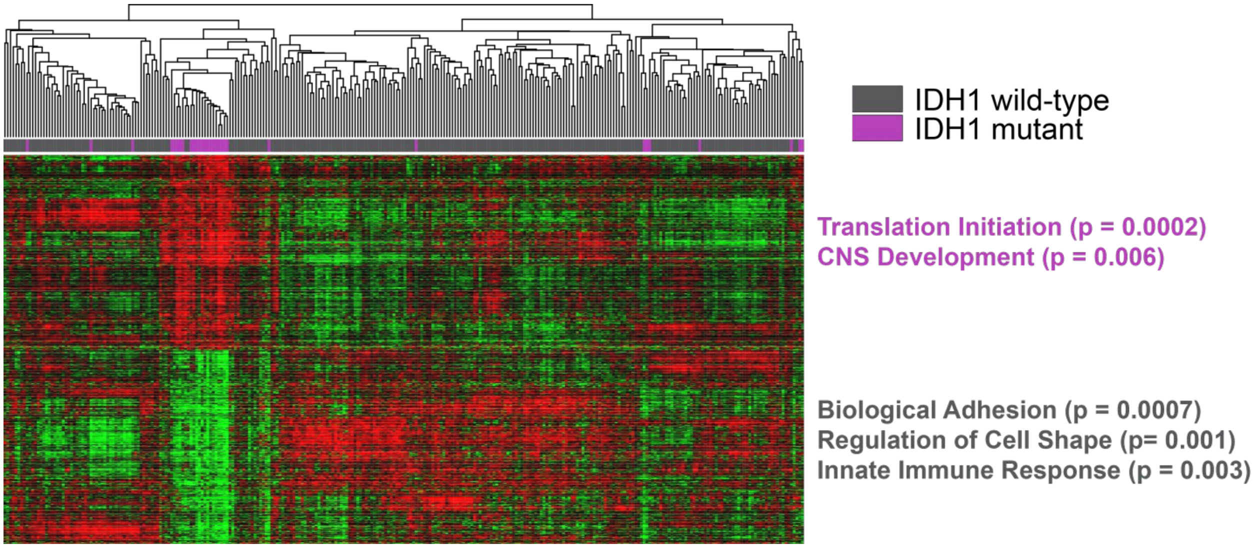
Genetic alterations and RNA subtypes are defined in eFigure 10 and eFigure 7, respectively.



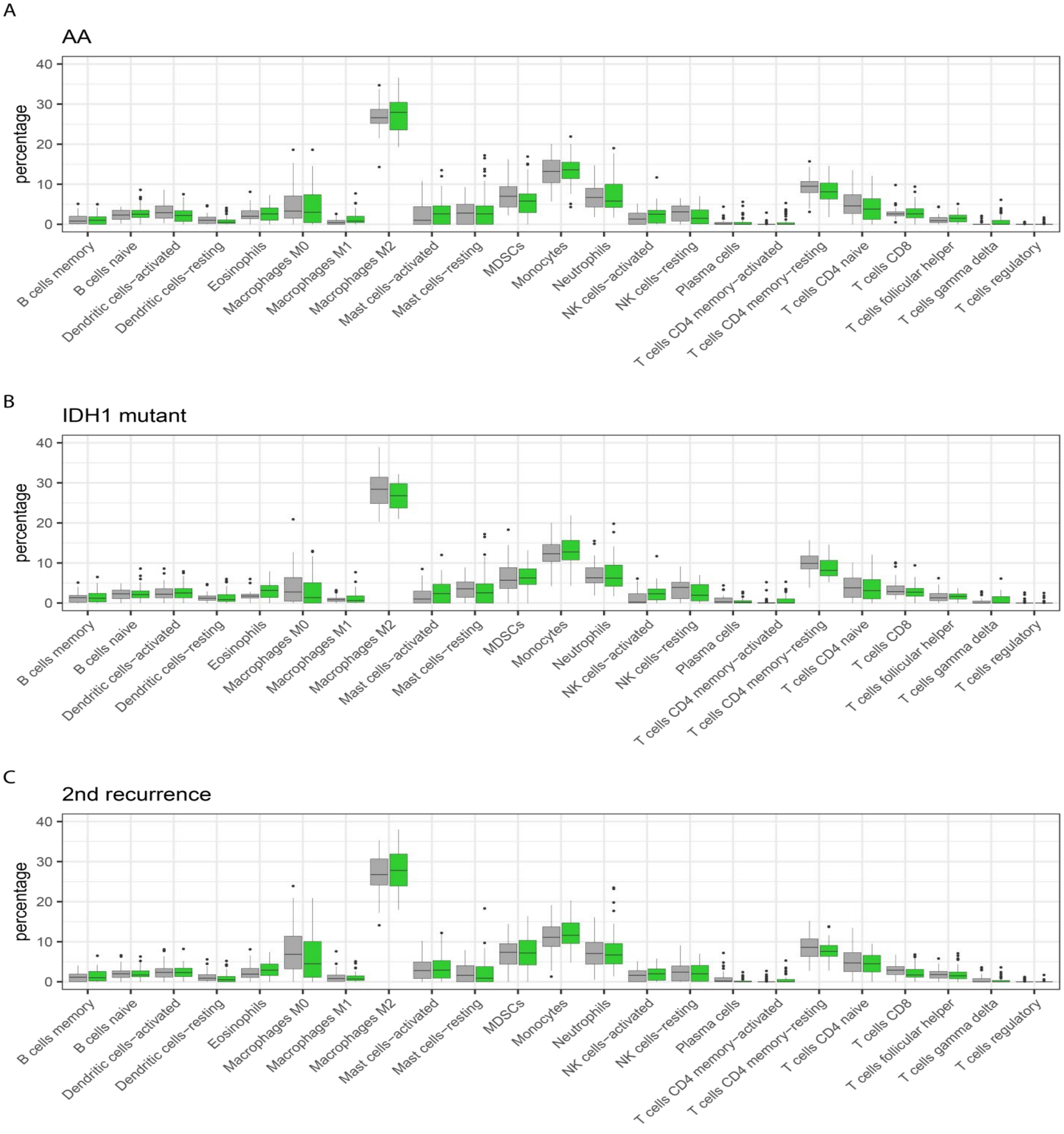
eFigure 22. Relative and absolute abundance of leukocyte populations in Toca 5 tumors.

(A) The relative levels of 23 leukocyte populations as measured by iSort in Toca 5 tumors grouped by molecular subtype (clustering).

(B) The estimated absolute levels (% of all cells in tumor sample) of 23 leukocyte populations as measured by iSort in Toca 5 tumors grouped by molecular subtype (clustering).



eFigure 23. mRNA expression differences between IDH1 mutant tumors and IDH1 wild-type tumors.
For this analysis the tumor sample from each biopsy with the highest RNA quality (RIN + fraction of reads mapping to exons) was used.

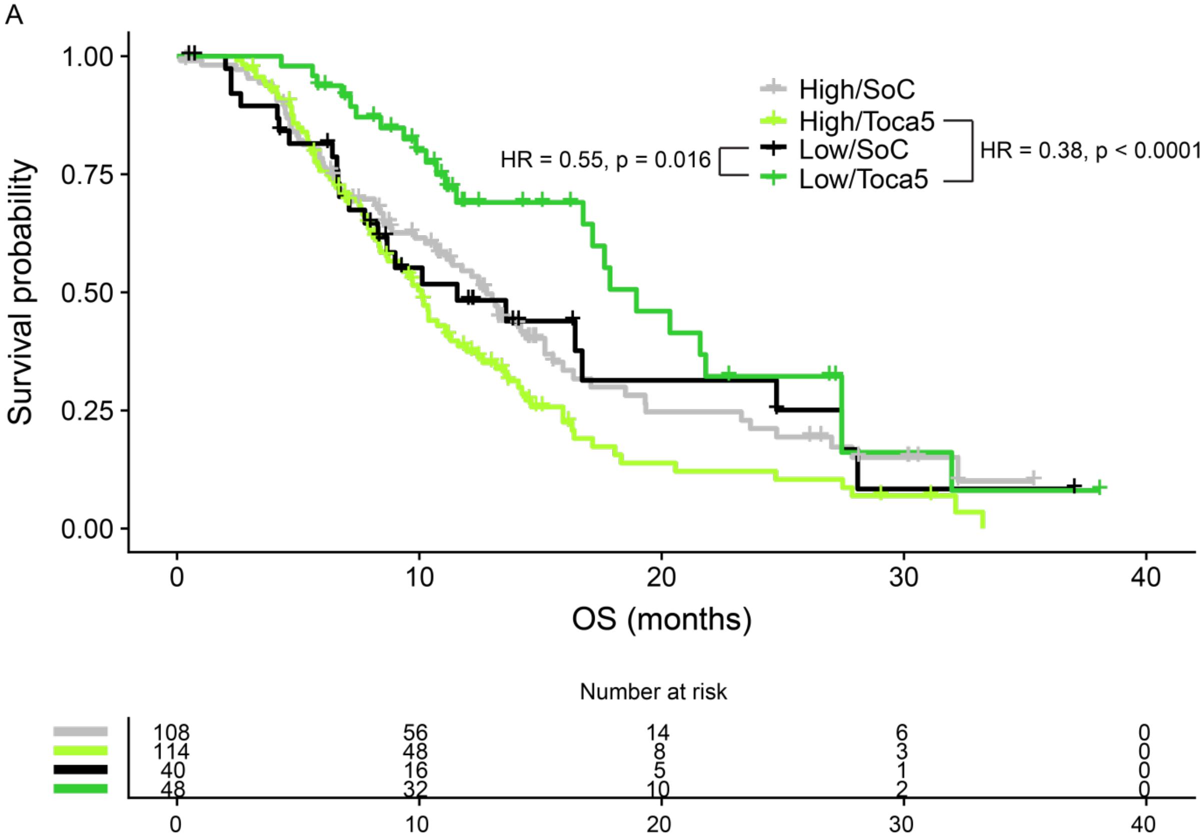


eFigure 24. Comparison between treatment-arms of immune cell population levels in tumors at time of surgery as measured by iSort.

(A) The relative percent total of each of 23 immune cell populations in AA tumors from patients on standard of care (grey) and Toca 511 and Toca FC (green) are shown as boxplots.

(B) The relative percent total of each of 23 immune cell populations in IDH1 mutant tumors from patients on standard of care (grey) and Toca 511 and Toca FC (green) are shown as boxplots.

(C) The relative percent total of each of 23 immune cell populations in second recurrence tumors from patients on standard of care (grey) and Toca 511 and Toca FC (green) are shown as boxplots.



B

strata	n	events	mOS	95% LCL	95% UCL
High/SoC	108	72	12.8	10.7	15.4
High/Toca5	114	99	10.1	8.4	11.6
Low/SoC	40	24	11.6	8.3	28.1
Low/Toca5	48	24	19	17.2	32

C

Performance Metrics 9-month survival

Sensitivity	0.48 (0.41, 0.54)
Specificity	0.58 (0.48, 0.68)
AUC	0.525

Performance Metrics 16-month survival

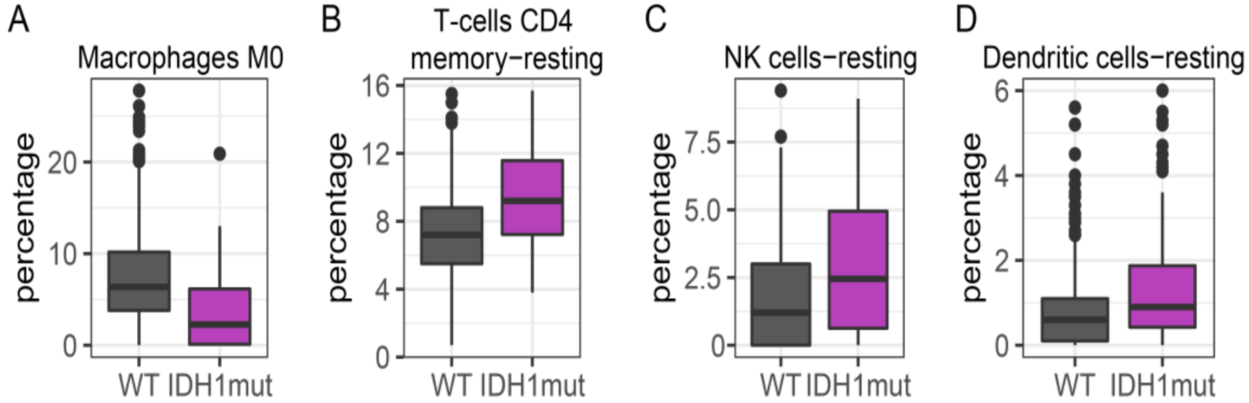
Sensitivity	0.85 (0.80, 0.89)
Specificity	0.22 (0.14, 0.31)
AUC	0.55

eFigure 25. Stratification of patients based on tumor immune cell levels inferred by iSort.

(A) Kaplan-Meier curves of patients stratified by "High" and "Low" immune signature score and by treatment arm.

(B) Summary of the survival curves.

(C) Summary of the predictive value of the immune signature stratification.



eFigure 26. Differences in immune cell composition between IDH1 mutant tumors and IDH1 wildtype tumors.

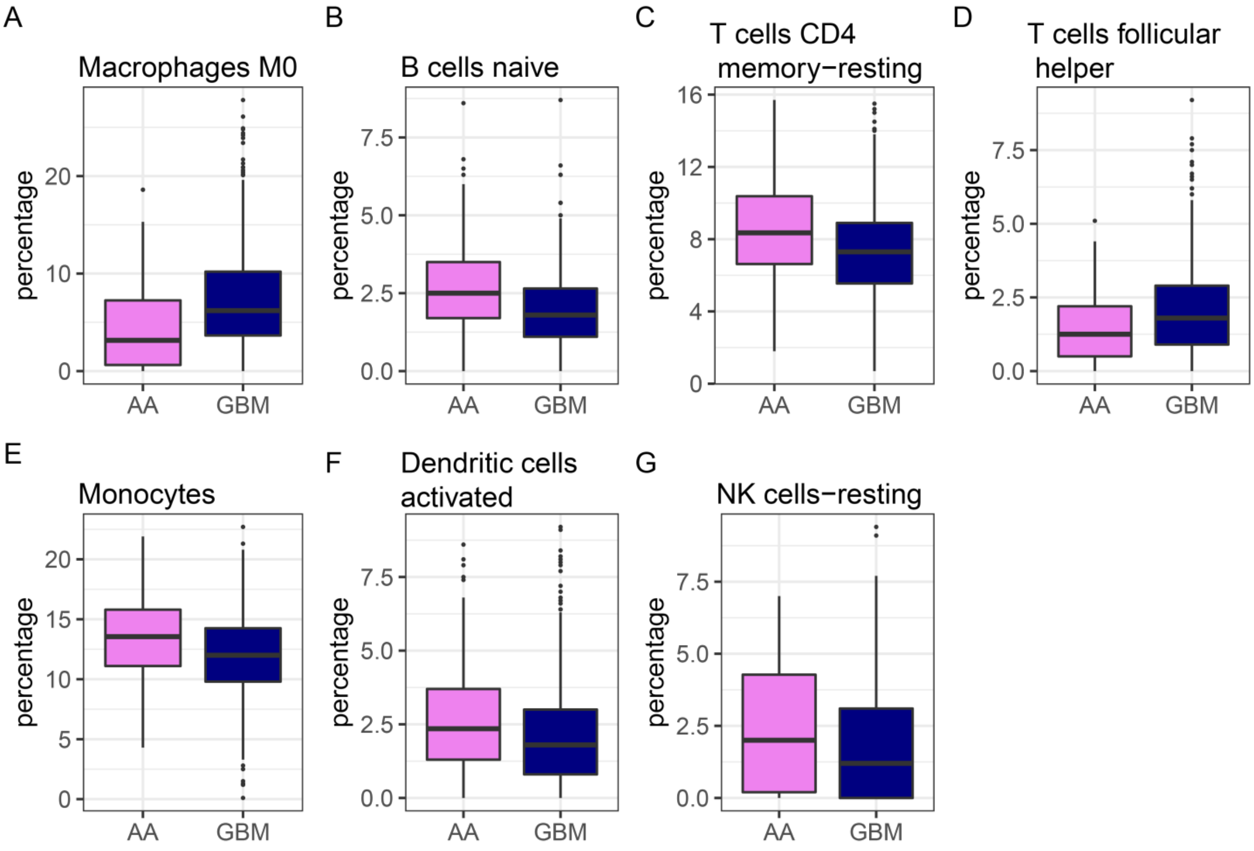
The relative levels of 23 immune cells were determined by iSort analyses of tumor RNA sequencing data.

(A) IDH1 mutant tumors (purple) have significantly lower levels of M0 macrophages than IDH1 wild-type tumors (dark gray) ($P < 1e-10$, Mann-Whitney test).

(B) IDH1 mutant tumors have significantly higher levels of CD4 memory-resting T-Cells than IDH1 wild-type tumors ($P < 1e-9$).

(C) IDH1 mutant tumors have significantly higher levels of resting NK cells than IDH1 wild-type tumors ($P < 1e-4$).

(D) IDH1 mutant tumors have significantly higher levels of resting dendritic cells than IDH1 wild-type tumors ($P < 1e-4$).



eFigure 27. Differences in immune cell and neoantigen composition between AA tumors and GBM tumors.

The relative levels of 23 immune cells were determined by iSort analyses of tumor RNA sequencing data.

(A) AA tumors (violet) have significantly lower levels of M0 macrophages than GBM tumors (navy blue) ($p < 1e-6$, Mann-Whitney test).

(B) AA tumors have significantly higher levels of naïve B cells than GBM tumors ($p < 1e-4$).

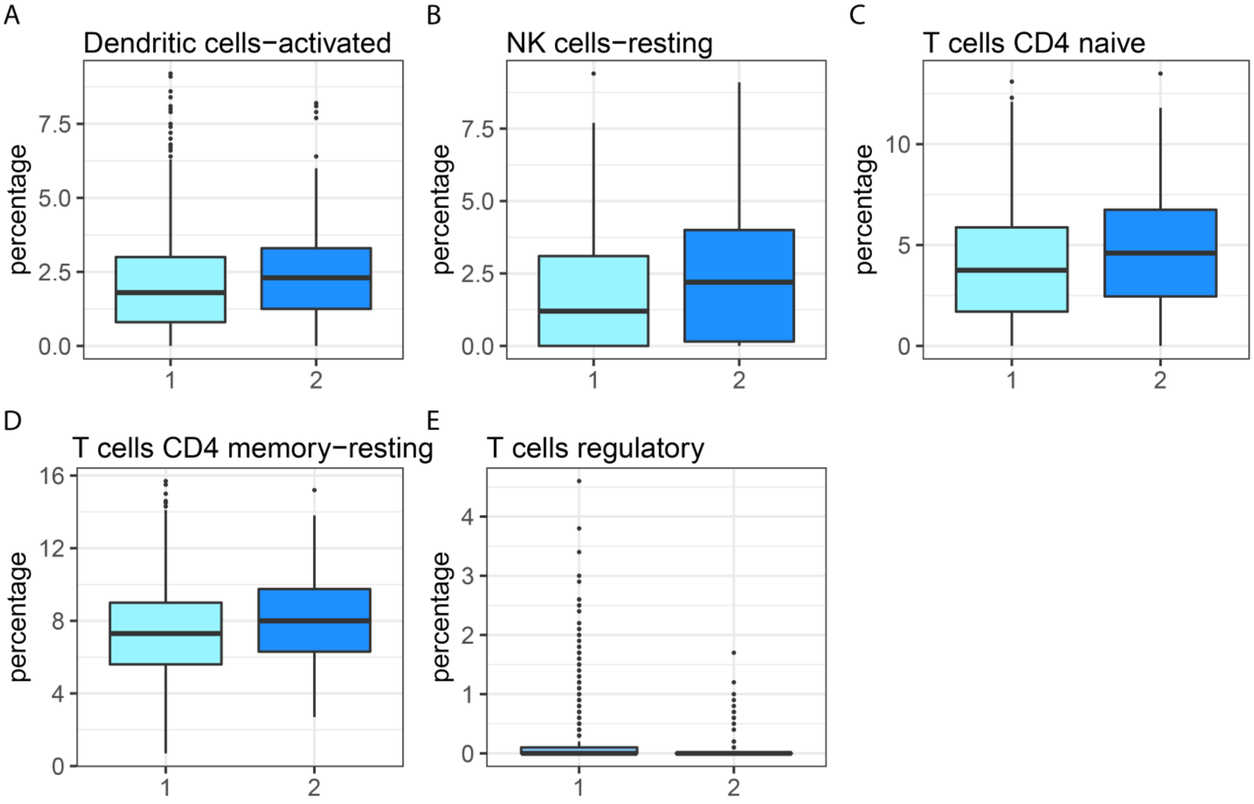
(C) AA tumors have significantly higher levels of resting CD4 memory-resting T-cells than GBM tumors ($p < 1e-4$).

(D) AA tumors have significantly lower levels of follicular help T-cells than GBM tumors ($p < 1e-4$).

(E) AA tumors have significantly higher levels of monocytes than GBM tumors ($p < 1e-3$).

(F) AA tumors have significantly higher levels of activated dendritic cells than GBM tumors ($p = 0.007$).

(G) AA tumors have significantly higher levels of resting NK-cells GBM tumors ($p = 0.01$).



eFigure 28. Differences in immune cell composition between first and second recurrence tumors.

The relative levels of 23 immune cells were determined by iSort analyses of tumor RNA sequencing data.

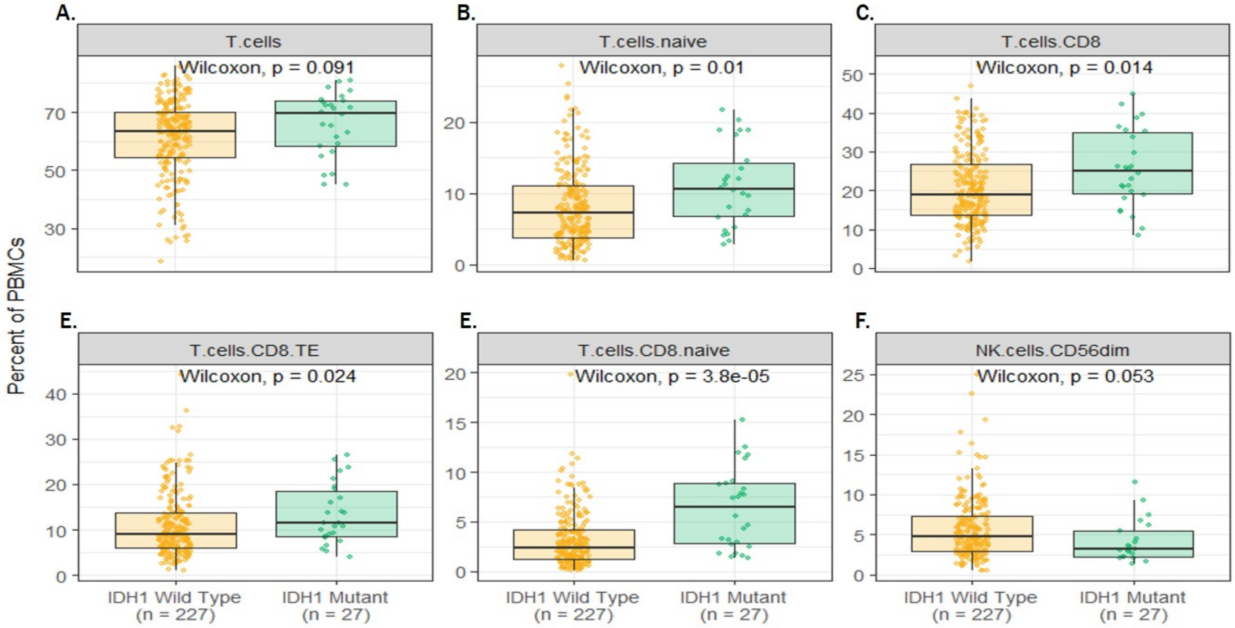
(A) Second recurrence tumors (darker blue) have significantly higher levels of activated dendritic cells than first recurrence tumors (lighter blue) (p = 0.001, Mann-Whitney test).

(B) Second recurrence tumors have significantly higher levels of resting NK cells than first recurrence tumors (p = 0.003).

(C) Second recurrence tumors have significantly higher levels of naive CD4 T-cells than first recurrence tumors (p = 0.003).

(D) Second recurrence tumors have significantly higher levels of follicular memory-resting CD4 T-cells than first recurrence tumors (p = 0.005).

(E) Second recurrence tumors have significantly lower levels of regulatory T-cells than first recurrence tumors (p = 0.007).



eFigure 29. Differences in peripheral immune cell composition between IDH1 widltype and IDH1 mutant patients.

The relative levels of peripheral immune cells were determined by FLOW analyses from patients' peripheral blood mononuclear cells.

(A) T cell population

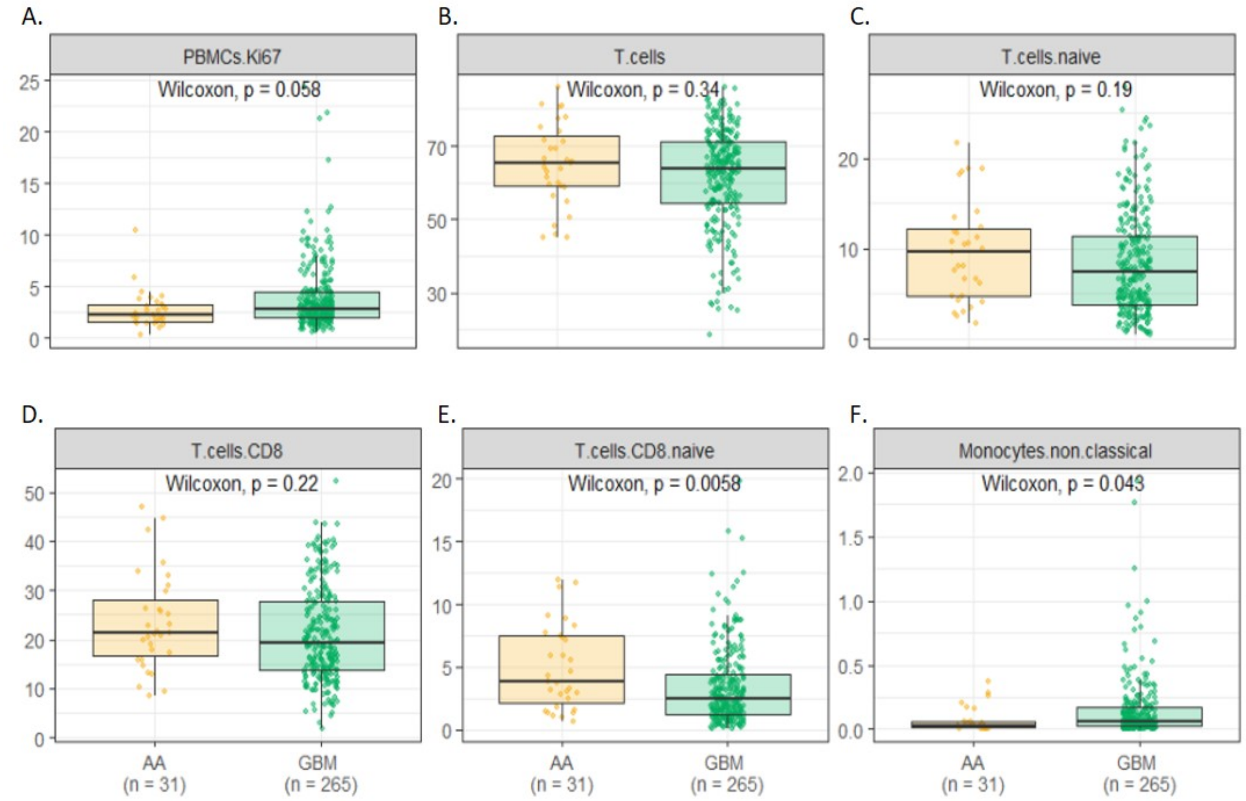
(B) Naïve T cell population

(C) CD8 T cell population

(D) Effector CD8 T cell population

(E) Naïve CD8 T cell population

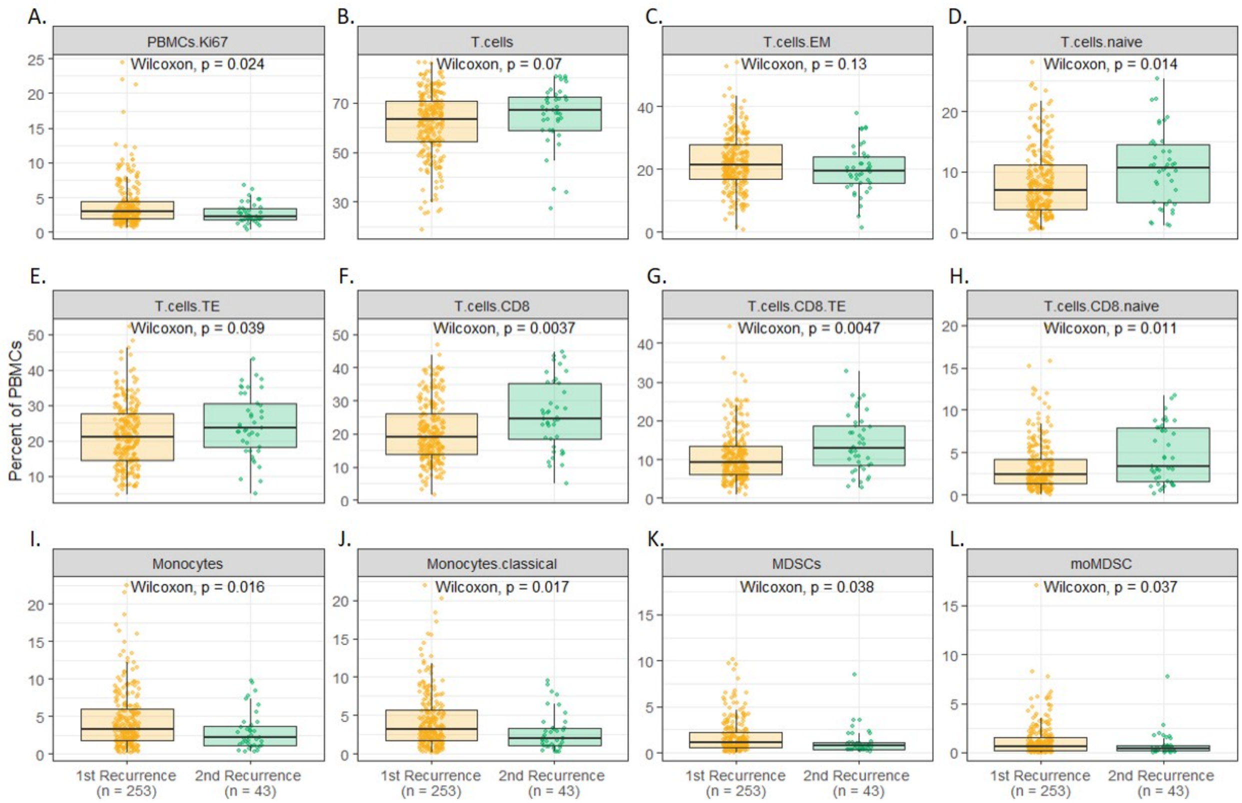
(F) CD56dim NK cell population



eFigure 30. Differences in peripheral immune cell composition between AA and GBM patients.

The relative levels of peripheral immune cells were determined by FLOW analyses from patients' peripheral blood mononuclear cells.

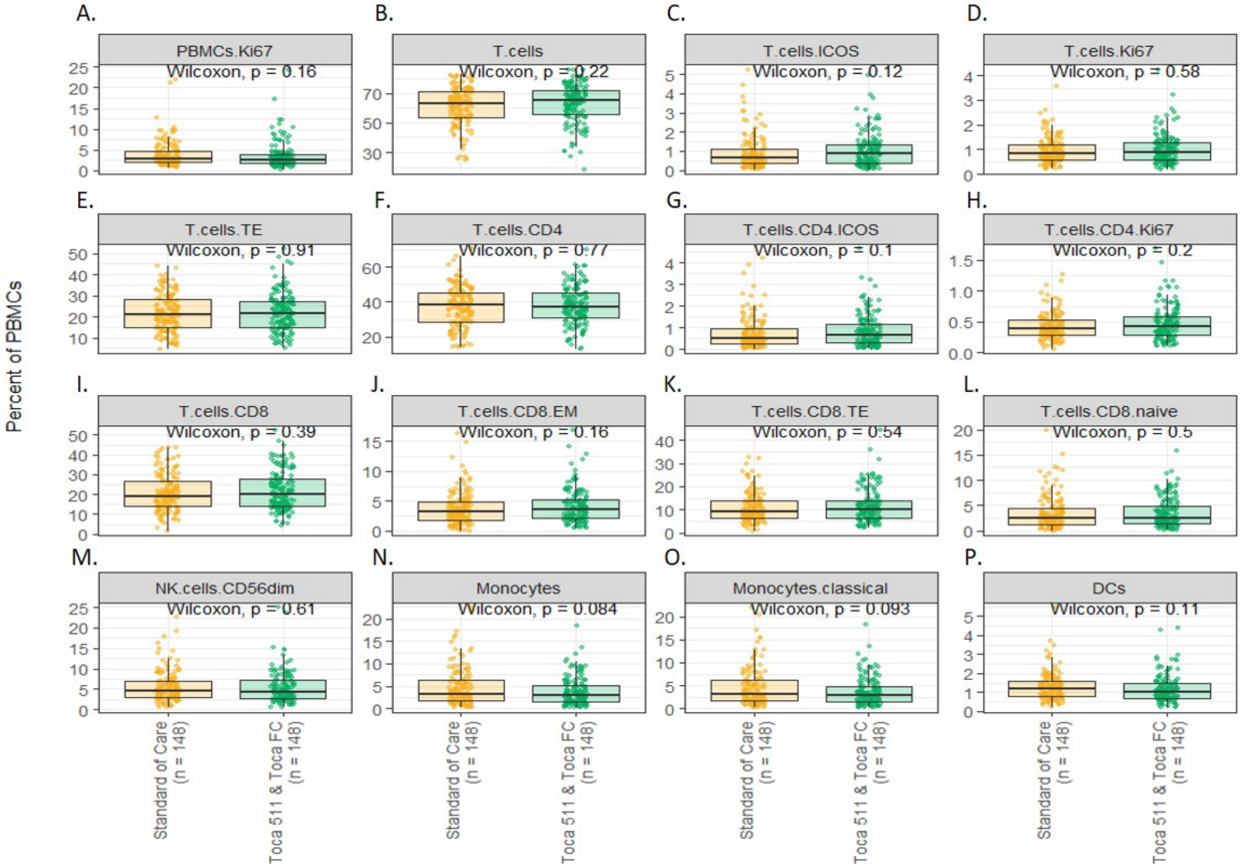
(A) Overall PBMC population
(B) T cell population
(C) Naïve T cell population
(D) CD8 T cell population
(E) Naïve CD8 T cell population
(F) Non-classical monocyte population



eFigure 31. Differences in peripheral immune cell composition between patients at first and second recurrence.

The relative levels of peripheral immune cells were determined by FLOW analyses from patients' peripheral blood mononuclear cells.

(A) Overall PBMC population
(B) T cell population
(C) Effector memory T cell population
(D) Naïve T cell population
(E) Effector T cell population
(F) CD8 T cell population
(G) Effector CD8 T cell population
(H) Naïve CD8 T cell population
(I) Monocyte population
(J) Classical monocyte population
(K) Myeloid derived suppressor cell population
(L) M0 Myeloid derived suppressor cell population



eFigure 32. Baseline peripheral immune balance between the control arm and the Toca 511/FC arm

The relative levels of peripheral immune cells were determined by FLOW analyses from patients' peripheral blood mononuclear cells.

(A) Ki67+ PBMC population
(B) T cell population
(C) ICOS+ T cell population
(D) Ki67+ T cell population
(E) Effector T cell population
(F) CD4 T cell population
(G) ICOS+ CD4 T cell population
(H) Ki67+ CD4 T cell population
(I) CD8 T cell population
(J) Effector memory CD8 T cell population
(K) Effector CD8 T cell population
(L) Naïve CD8 T cell population
(M) CD56dim NK cell population
(N) Monocyte population
(O) Classical monocyte population
(P) Dendritic cell population

eTable 1. Baseline demographics and neuro-oncology history for Phase 2 and Phase 3 patients.

	Phase 2 (N=187)			Phase 3 (N=216)		
	Toca 511 and Toca FC (N=93)	Standard of Care (N=94)	Total (N=187)	Toca 511 and Toca FC (N=108)	Standard of Care (N=108)	Total (N=216)
Age (Years) Median	55.0	56.0	56.0	58.5	56.0	58.0
<65	73(78.5%)	75 (79.8%)	148 (79.1%)	82 (75.9%)	89 (82.4%)	171 (79.2%)
≥65	20 (21.5%)	19 (20.2%)	39 (20.9%)	26 (24.1%)	19 (17.6%)	45 (20.8%)
Sex						
Male	61 (65.5%)	57 (60.6%)	118 (63.1%)	64 (59.3%)	70 (64.8%)	134 (62.0%)
Female	32 (34.4%)	37 (39.4%)	69 (36.9%)	44 (40.7%)	38 (35.2%)	82 (38.0%)
Race						
White	80 (86.0%)	83 (88.3%)	163 (87.2%)	94 (87.0%)	90 (83.3%)	184 (85.2%)
KPS Category						
70-80	28 (30.1%)	30 (31.9%)	58 (31.0%)	36 (33.3%)	36 (33.3%)	72 (33.3%)
90-100	65 (69.9%)	64 (68.1%)	129 (69.0%)	72 (66.7%)	72 (66.7%)	144 (66.7%)
Tumor Histology at Study Entry from Investigator						
Glioblastoma	72 (77.4%)	78 (83.0%)	150 (80,2%)	86 (79.6%)	98 (90.7%)	184 (85.2%)
Anaplastic Astrocytoma	17 (18.3%)	14 (14.9%)	31 (16.6%)	21 (19.4%)	10 (9.3%)	31 (14.4%)
IDH1 Mutation Status						
Present (Mutant)	15 (16.1%)	17 (18.1%)	32 (17.1%)	12 (11.1%)	14 (13.0%)	26 (12.0%)
Absent (Wildtype)	78 (83.9%)	77 (81.9%)	155 (82.9%)	96 (88.9%)	94 (87.0%)	190 (88.0%)

	Phase 2 (N=187)			Phase 3 (N=216)		
	Toca 511 and Toca FC (N=93)	Standard of Care (N=94)	Total (N=187)	Toca 511 and Toca FC (N=108)	Standard of Care (N=108)	Total (N=216)
Number of Recurrence						
One	78 (83.9%)	78 (83%)	156 (83.4%)	95 (88.0%)	92 (85.2%)	187 (86.6%)
Two	15 (16.1%)	16 (17.0%)	31 (16.6%)	13 (12.0%)	16 (14.8%)	29 (13.4%)

eTable 2. Toca FC and SOC cycle numbers and schedule for patients.

	Toca 511 and Toca FC Arm	Standard of Care Arm			
Number of cycles treatment	Toca FC	Metronomic TMZ	TMZ	Lomustine	Bevacizumab
n	197	18	31	66	26
Mean (SD)	3.4 (3.36)	4.7 (3.80)	4.4 (3.64)	2.3 (1.68)	6.2 (5.49)
Median	2	3	3	2	4
Min, Max	1, 21	1, 11	1, 15	1, 7	1, 24
Drug dosing schedule	Every 6 weeks	Every day	Every 4 weeks	Every 6 weeks	Every 2 weeks
Cycle length	6 weeks	4 weeks	4 weeks	6 weeks	4 weeks

eTable 3. Secondary endpoints for Toca 5 randomized clinical trial.

	Toca 511 and Toca FC (N=201)	Standard of Care (N=202)	HR (95% CI)	P-value
Durable response rate (%)	5 (2.5%)	9 (4.5%)	NA	0.4156
Durable clinical benefit rate (%)	5 (2.5%)	9 (4.5%)	NA	0.4156
Duration of durable response (months)	11.17	6.93	0.59 (0.14, 2.56)	0.5161
OS-12 (%)	45.6	51.4	NA	0.2848

eTable 4. Baseline patient characteristics for patients in second recurrence.

	Toca 511 and Toca FC (N=28)	Standard of Care (N=32)	Total (N=60)
KPS Category			
70-80	8 (28.6%)	8 (25.0%)	16 (26.7%)
90-100	20 (71.4%)	24 (75.0%)	44 (73.3%)
IDH1 Mutation Status			
Present (Mutant)	8 (28.6%)	10 (31.3%)	18 (30.0%)
Absent (Wildtype)	20 (71.4%)	22 (68.8%)	42 (70.0%)
1p/19q Co-Deletion Status			
Present (Co-Deletion)	0 (0%)	0 (0%)	0 (0%)
Absent (No Co-Deletion)	28 (100%)	32 (100%)	60 (100%)
Tumor Histology at Study Entry			
Anaplastic Astrocytoma	7 (11.7%)	6 (10.0%)	13 (21.7%)
Glioblastoma	21 (35.0%)	26 (43.3%)	47 (78.3%)
Total	28 (46.7%)	32 (53.3%)	60 (100%)

eTable 5. Toca 511 viral RNA and DNA signal by cycle.

	Viral RNA Signal (RT-PCR)	Viral DNA Signal (qPCR)
Post-Surgery	134/189 (70.9%)	1/184 (0.5%)
Cycle 1 Day 1	18/184 (9.8%)	8/178 (4.5%)
Cycle 2 Day 1	1/146 (0.7%)	1/139 (0.7%)
Cycle 4 Day 1	0/24 (0%)	0/19 (0%)
Cycle 8 Day 1	0/11 (0%)	0/10 (0%)
Cycle 12 Day 1	0/2 (0%)	0/2 (0%)
Cycle 16 Day 1	0/24 (0%)	0/4 (0%)
End of Treatment[1]	1/108 (0.9%)	3/103 (2.9%)

[1] End of treatment samples were either from Cycle 1 Day 1 or Cycle 2 Day 1 time points.

eTable 6. Proposed etiology for SBS Mutational signatures, from COSMIC.

Signature	Proposed etiology
SBS 01	An endogenous mutational process initiated by spontaneous or enzymatic deamination of 5-methylcytosine to thymine which generates G:T mismatches in double stranded DNA. Failure to detect and remove these mismatches prior to DNA replication results in fixation of the T substitution for C.
SBS 03	Defective homologous recombination-based DNA damage repair which manifests predominantly as small indels and genome rearrangements due to abnormal double strand break repair but also in the form of this base substitution signature.
SBS 05	Unknown. SBS5 mutational burden is increased in bladder cancer samples with ERCC2 mutations and in many cancer types due to tobacco smoking.
SBS 07b	SBS7a/SBS7b/SBS7c/SBS7d are found in cancers of the skin from sun exposed areas and are likely to be due to exposure to ultraviolet light. SBS7b may possibly be the consequence of just one of the two major known UV photoproducts, cyclobutane pyrimidine dimers or 6-4 photoproducts. However, there is no evidence for this hypothesis and it is unclear which of these photoproducts may be responsible for SBS7b.
SBS 07d	SBS7a/SBS7b/SBS7c/SBS7d are found in cancers of the skin from sun exposed areas and are likely to be due to exposure to ultraviolet light. SBS7d is possibly the consequence of translesion DNA synthesis by error-prone polymerases with greater propensity to insert G, rather than A, opposite UV light induced thymidine and cytidine photodimers.
SBS 08	Unknown.
SBS 10b	Polymerase epsilon exonuclease domain mutations.
SBS 11	SBS11 exhibits a mutational pattern resembling that of alkylating agents. Patient histories indicate an association between previous treatment with the alkylating agent temozolomide and SBS11 mutations.
SBS 12	Unknown.
SBS 13	SBS13 is closely associated with SBS2. SBS13 is also associated with DBS11, which is characterised predominantly by CC>TT doublet base substitutions as well as other CC>NN doublet base substitutions.
SBS 14	Concurrent polymerase epsilon mutation and defective DNA mismatch repair.
SBS 15	Defective DNA mismatch repair.