

Supplementary Appendix to:

Neoantigen Landscape in an Ultramutated Glioblastoma Arising in a Patient with Germline *POLE* Deficiency Treated with Checkpoint Blockade Immunotherapy

Tanner M. Johanns, MD, PhD^{1,2,*}, Christopher A. Miller, PhD^{3,*}, Ian G. Dorward, MD⁴, George Anstas, MD¹, Christina Tsien⁵, Edward Chang, MD⁶, Arie Perry, MD⁷, Ravindra Uppaluri, MD, PhD⁸, Cole Ferguson, MD, PhD⁹, Robert Schmidt, MD, PhD⁹, Sonika Dahiya, MD⁹, Elaine R. Mardis, PhD^{3,**}, Gavin P. Dunn, MD, PhD^{2,4,10,**}

¹Division of Oncology, Department of Medicine, Washington University School of Medicine, St. Louis, Missouri

²Center for Human Immunology and Immunotherapy Programs, Washington University School of Medicine, St. Louis, Missouri

³The McDonnell Genome Institute, Washington University, St. Louis, MO, USA, 63108

⁴Department of Neurological Surgery, Washington University School of Medicine, St. Louis, Missouri

⁵Department of Radiation Oncology, Washington University School of Medicine, St. Louis, Missouri

⁶Department of Neurological Surgery, University of California, San Francisco, San Francisco, United States

⁷Department of Pathology, University of California, San Francisco, CA.

⁸Department of Otolaryngology, Washington University School of Medicine, St. Louis, Missouri

⁹Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri

¹⁰The Alvin J. Siteman Cancer Center at Barnes-Jewish Hospital and Washington University School of Medicine, St. Louis, Missouri

*These authors contributed equally to this manuscript.

**To whom correspondence should be addressed

A. Supplementary Methods

A.1 Library Construction and sequencing

Libraries were captured using xGen Exome Research Panel v1.0 from IDT. Capture was performed following manufacturer's recommendations. KAPA qPCR was used to quantify the libraries and determine the appropriate concentration to produce optimal recommended cluster density on a HiSeq2500 v4 (PE125bp), sequencing run. All sequencing runs were completed according the manufacturer's recommendations (Illumina Inc, San Diego, CA).

cDNA was generated using the TruSeq Stranded Total RNA with Ribo-Zero Gold kit according to the manufacturer's recommendations (Illumina Inc, San Diego, CA). Libraries were captured using xGen Exome Research Panel v1.0 from IDT. Capture was performed following manufacturer's recommendations (IDT Technologies, Coralville, IA). KAPA qPCR (KAPA Biosystems, Boston, MA) was used to quantify the libraries and determine the appropriate concentration to produce optimal recommended cluster density on a HiSeq2500 v4 (PE125bp), sequencing run. All sequencing runs were completed according the manufacturer's recommendations (Illumina Inc, San Diego, CA).

Sequence coverage of the target space was as follows:

Sample	Sequence Coverage
Normal	86.60x
Primary	208.02x
C7-T2 Metastasis	226.74x
T7-8 Metastasis	266.50x

A.2 Somatic variant analysis

Sequence data was aligned to reference sequence build GRCh37-lite-build37 using bwa¹ version 0.5.9 (params: -t 4 -q 5::), then merged and deduplicated using picard version 1.46 (<https://broadinstitute.github.io/picard/>).

SNVs were detected using the union of four callers: 1) samtools² version r982 (params: -A -B) intersected with Somatic Sniper³ version 1.0.4 (params: -F vcf -G -L -q 1 -Q 15) and processed through false-positive filter v1 (params: --bam-readcount- version 0.4 --bam-readcount-min-base-quality 15 --min-mapping-quality 40 --min-somatic-score 40), 2) VarScan⁴ version 2.3.6 filtered by varscan-high-confidence filter version v1 and processed through false-positive filter v1 (params: --bam-readcount-version 0.4 --bam-readcount-min-base-quality 15), 3) Strelka⁵ version 1.0.11 (params: isSkipDepthFilters = 1), and 4) mutect⁶ v1.1.4

Indels were detected using the union of 4 callers: 1) GATK⁷ somatic-indel version 5336) pindel version 0.5 filtered with pindel⁸ false- positive and VAF filters (params: --variant-freq-cutoff=0.08), 3) VarScan⁴ version 2.3.6 filtered by varscan-high-confidence- indel version v1 and 4) Strelka⁵ version 1.0.11 (params: isSkipDepthFilters = 1).

SNVs and Indels were further filtered by removing artifacts found in a panel of 905 normal

exomes, removing sites that exceeded 0.1% frequency in the 1000 genomes or NHLBI exome sequencing projects, and then using a bayesian classifier (<https://github.com/genome/genome/blob/master/lib/perl/Genome/Model/Tools/Validation/IdentifyOutliers.pm>) and retaining variants classified as somatic with a binomial log-likelihood of at least 5.

Copy number aberrations were detected using VarScan⁴ v2.3.6, segmented with the DNACopy⁹ package for R (params: `undo.splits="sdundo"`, `undo.SD=4`, `min.width=2`). Segments were retaining if they had 50 supporting probes, then adjacent segments that differed by less than 0.2 were merged.

A.3 RNAseq analysis

RNA-seq data was aligned to reference sequence build GRCh37-lite-build37 with Tophat v2.0.8¹⁰ (denovo mode, params: `--library-type fr-firststrand --bowtie-version=2.1.0`) and expression levels were calculated with Cufflinks¹¹ v2.1.1 (params: `--max-bundle-length 10000000 --max-bundle-frags 10000000`). Allelic expression was determined for somatic variant calls using bam-readcount (<https://github.com/genome/bam-readcount>)

Gene fusions were inferred jointly from RNA and DNA using INTEGRATE version 0.2 with default parameters.¹² Subtyping was performed using the KNN Method in GlioVis (<http://gliovis.bioinfo.cnio.es/>), with the following results:

Sample	Subtype	Probability Classical	Probability Mesenchymal	Probability Proneural
Primary	Classical	0.62	0.23	0.15
C7-T2 Metastasis	Mesenchymal	0.06	0.94	0
T7-8 Metastasis	Mesenchymal	0.06	0.94	0

Note that this subtyping uses a more recent 3-group classification from Verhaak, et al., rather than the older 4-type classification¹³.

A.4 Subclonal inference

Subclonal inference was performed using copy-number neutral SNVs and sciClone¹⁴ version 1.1. (params: `minimumDepth=300`, `maximumClusters=15`)

A.5 Neoepitope prediction

Both the wild type and mutant peptide sequences for each missense mutations from each tumor were created, then MHC class I binding predictions were generated using NetMHC v3.4, as well as 5 other algorithms from the Immune Epitope Database and Analysis resource (IEDB, iedb.org): `netmhcons`, `netmhcpn`, `pickpocket`, `smm`, and `smmpmbe`. Predictions were retained if the best score had an $ic50 < 500$ and better binding of the mutant peptide than the wild type (fold-change > 1). Results were then filtered to require expression of the mutant allele (FPKM > 1 and at least one variant-supporting read in the RNA). Variants above 25% VAF in all three tumors were considered founding clone variants.

A.6 Mutational Signatures

Mutational signatures were inferred using the `deconstructSigs` package for R, version 1.6.0 with default parameters (`signature.cutoff = 0.06`)¹⁵.

B. Supplemental Figures

Figure S1: Mutation Spectra of each subclonal population

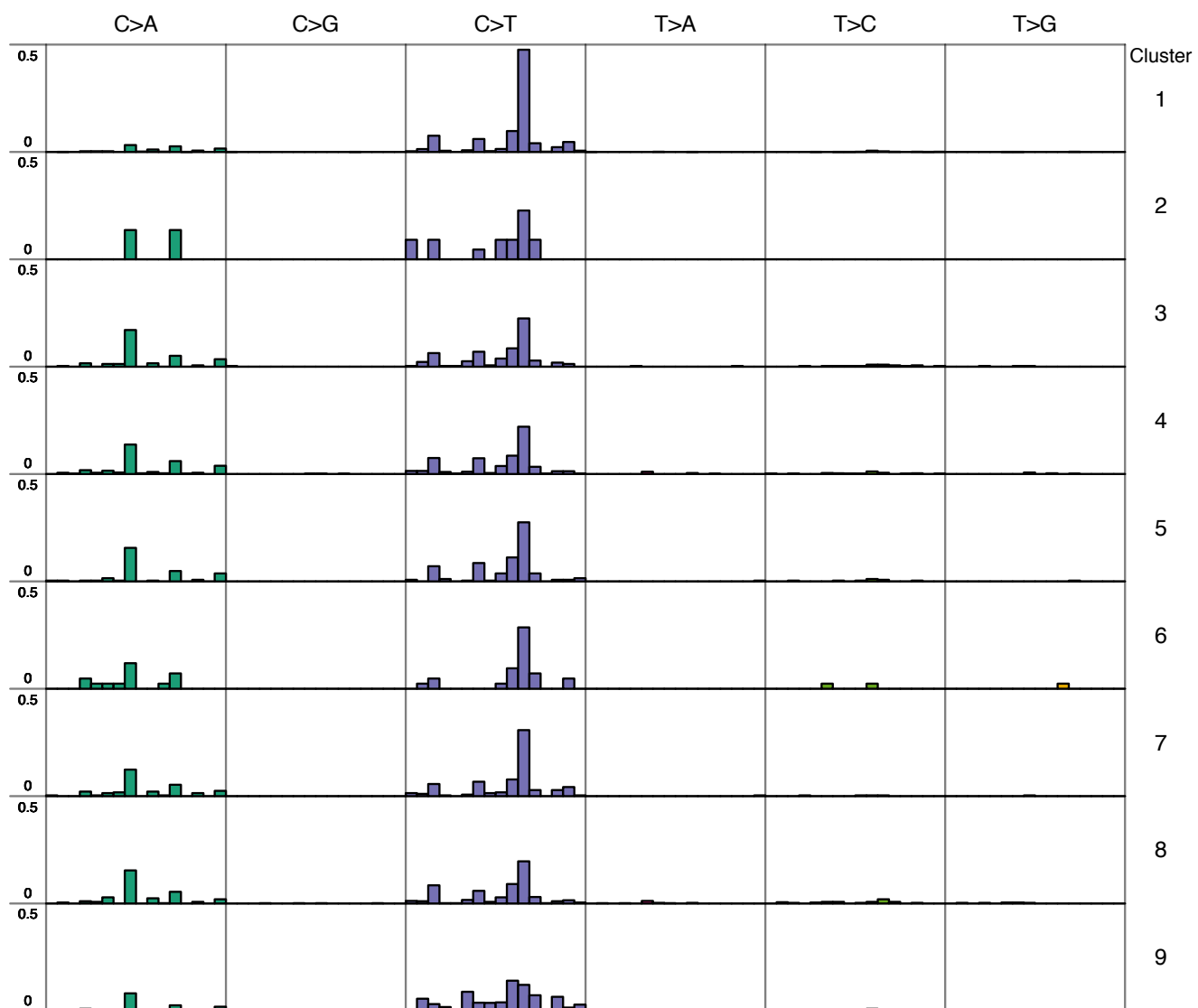


Figure S2:

Somatic Copy Number calls. Top: Primary, Middle: C7-T2 Metastasis, Bottom: T7-8 Metastasis

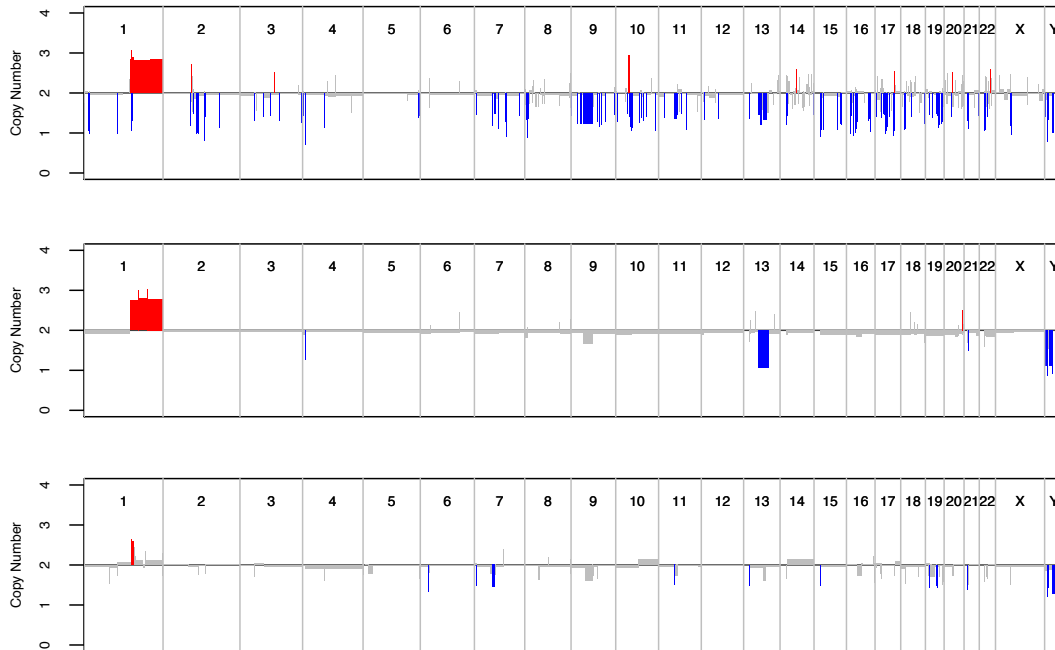
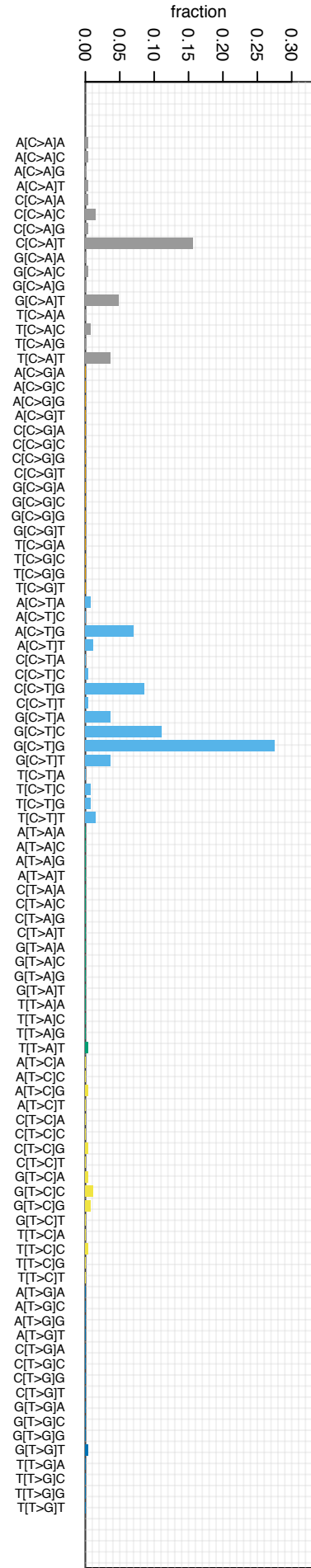


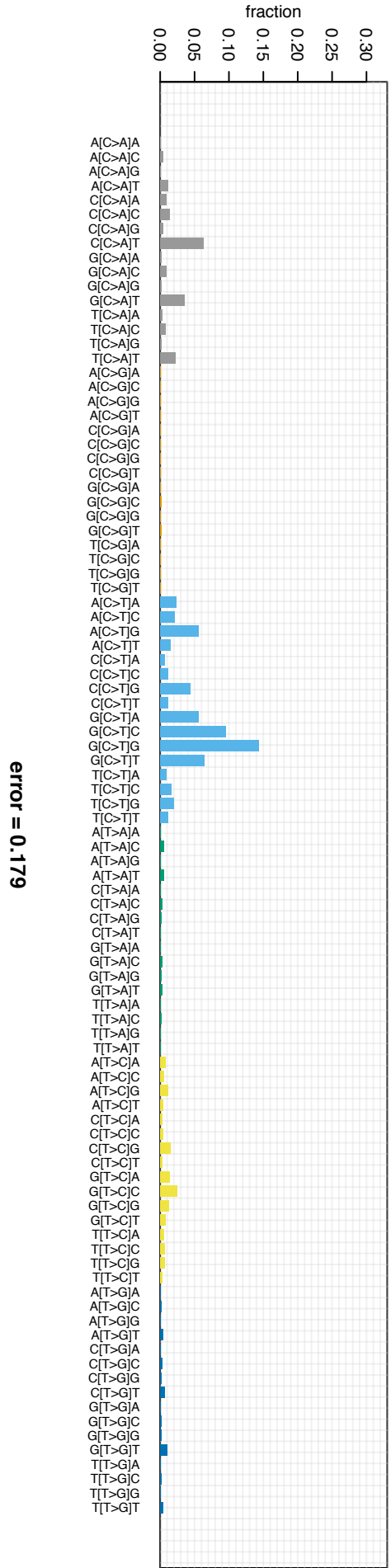
Figure S3:

Mutation signatures derived for each subclone independently

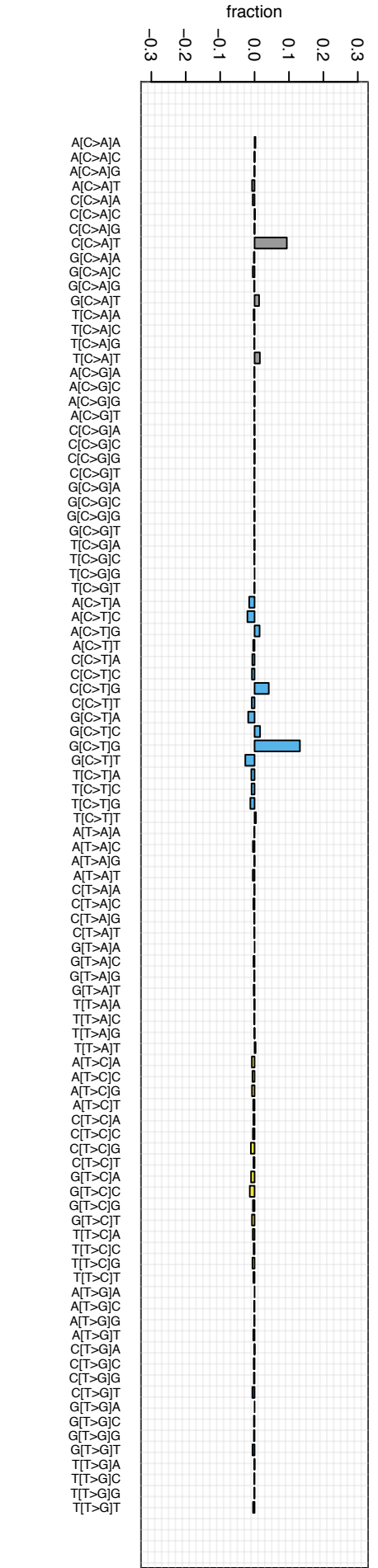
GBM27 -- Cluster 5



Signature:6 : 0.334 & Signature:14 : 0.274 & Signature:15 : 0.393

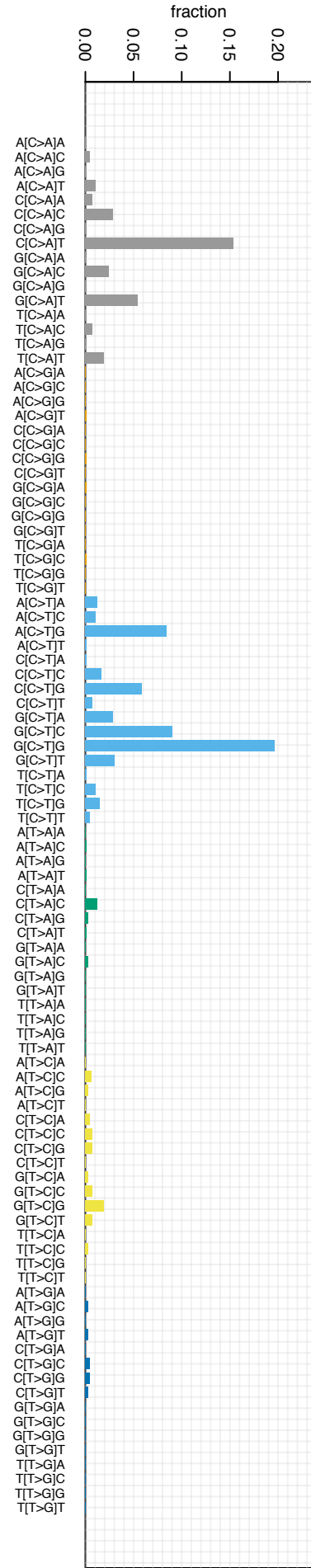


error = 0.179

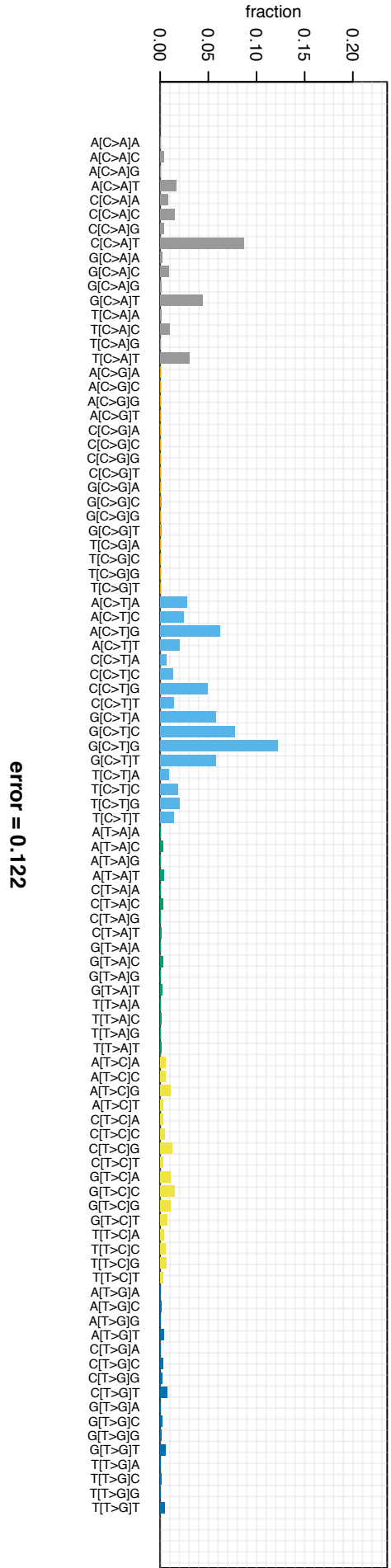


C>A
G>G
T>C
T>A
G>T
C>G
T>G

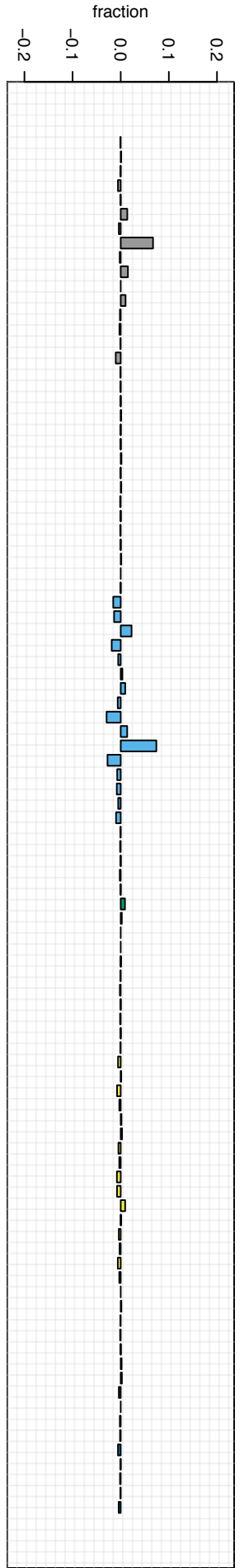
GBM27 -- Cluster 8



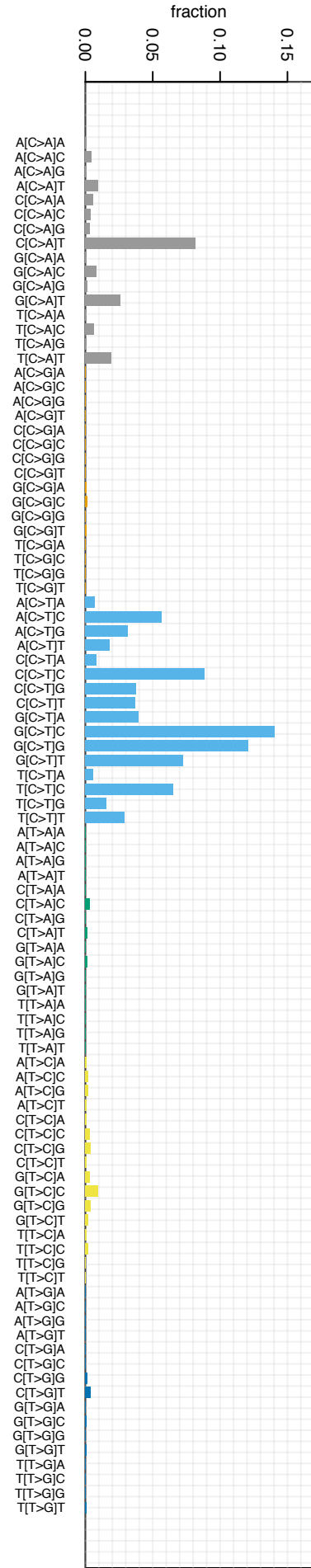
Signature:6 : 0.363 & Signature:14 : 0.503 & Signature:15 : 0.134



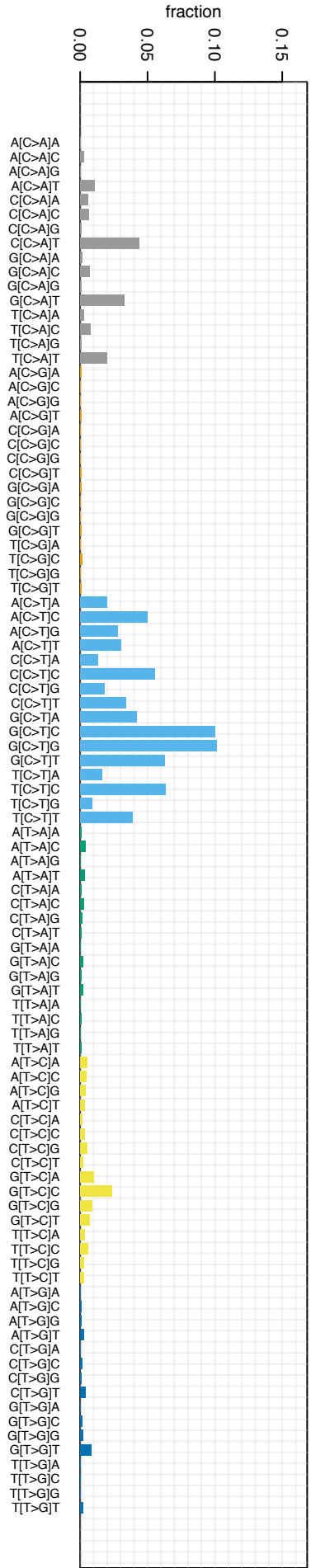
error = 0.122



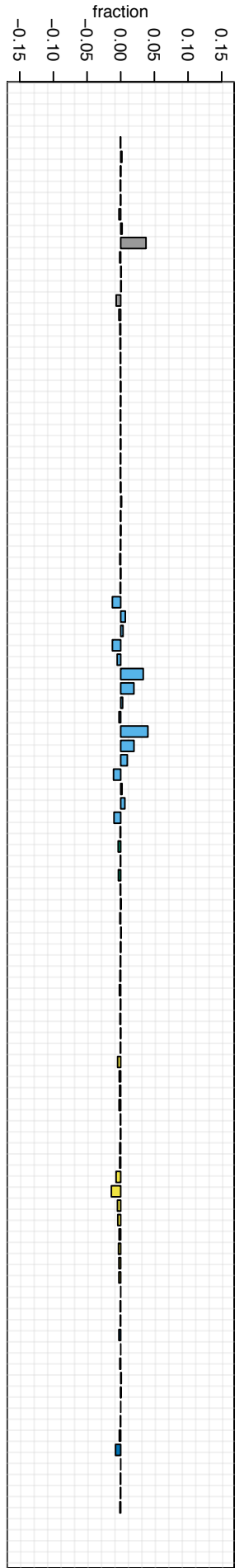
C>A
C>G
G>T
T>A
T>C
T>G



Signature.11 : 0.311 & Signature.14 : 0.275 & Signature.15 : 0.392

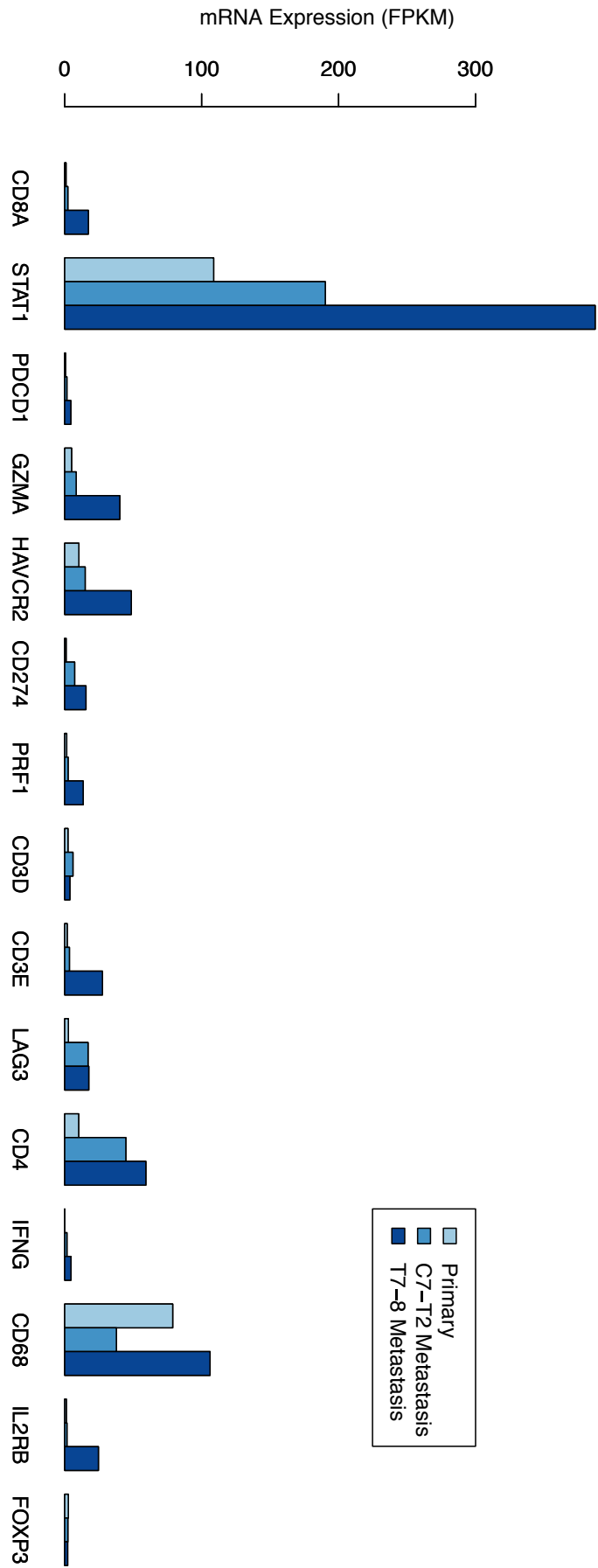


error = 0.079



C>A
G>T
T>A
T>C
T>G

Figure S4: Expression of immune-related genes



C. Supplemental Tables

Table S1:

Somatic SNVs and Indels with transcript annotation, population frequencies, and readcounts appended. Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable1_variants.xls

Table S2:

Segmented somatic copy number calls. Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable2_copynumber.xls

Table S3:

Gene Fusions identified with INTEGRATE. Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable3_fusions.xls

Table S4:

Gene Expression from RNA-sequencing in each of the three tumors. Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable4_expression.xls

Table S5:

DNA repair pathway mutations found in the founding clone (Cluster 1). Gene list was derived from http://sciencepark.mdanderson.org/labs/wood/dna_repair_genes.html, and annotated with data from ClinVar (www.ncbi.nlm.nih.gov/clinvar/), COSMIC (<http://cancer.sanger.ac.uk/cosmic>), the Foundation Medicine report provided to this patient, and SIFT/PolyPhen predictions from ExAC (<http://exac.broadinstitute.org/>). Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable5_repairGenes.xls

Table S6:

Clonal clustering assignments generated from sciClone. Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable6_sciclone.xls

Table S7:

Predicted neoantigens generated as described above. This table contains all peptides with predicted ic50 less than 500 and fold-change greater than 1. The list is provided before filtering for expression or presence in the founding clone. Available in the Supplementary Data, or at:

http://genome.wustl.edu/pub/supplemental/Johanns_Miller_2016/suppTable7_neoepitopes.xls

D. References

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma Oxf Engl* 2009;25(14):1754–60.
2. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl* 2009;25(16):2078–9.
3. Larson DE, Harris CC, Chen K, et al. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinforma Oxf Engl* 2012;28(3):311–7.
4. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22(3):568–76.
5. Saunders CT, Wong WSW, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinforma Oxf Engl* 2012;28(14):1811–7.
6. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013;31(3):213–9.
7. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20(9):1297–303.
8. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinforma Oxf Engl* 2009;25(21):2865–71.
9. Olshen AB, Venkatraman ES, Lucito R, Wigler M. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostat Oxf Engl* 2004;5(4):557–72.
10. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009;25(9):1105–11.
11. Trapnell C, Roberts A, Goff L, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 2012;7(3):562–78.
12. Zhang J, White NM, Schmidt HK, et al. INTEGRATE: Gene fusion discovery using whole genome and transcriptome data. *Genome Res* 2015;gr.186114.114.
13. Wang Q, Hu X, Muller F, et al. Tumor evolution of glioma intrinsic gene expression subtype associates with immunological changes in the microenvironment. *bioRxiv* 2016;52076.
14. Miller CA, White BS, Dees ND, et al. SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS Comput Biol* 2014;10(8):e1003665.
15. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. deconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol* 2016;17:31.