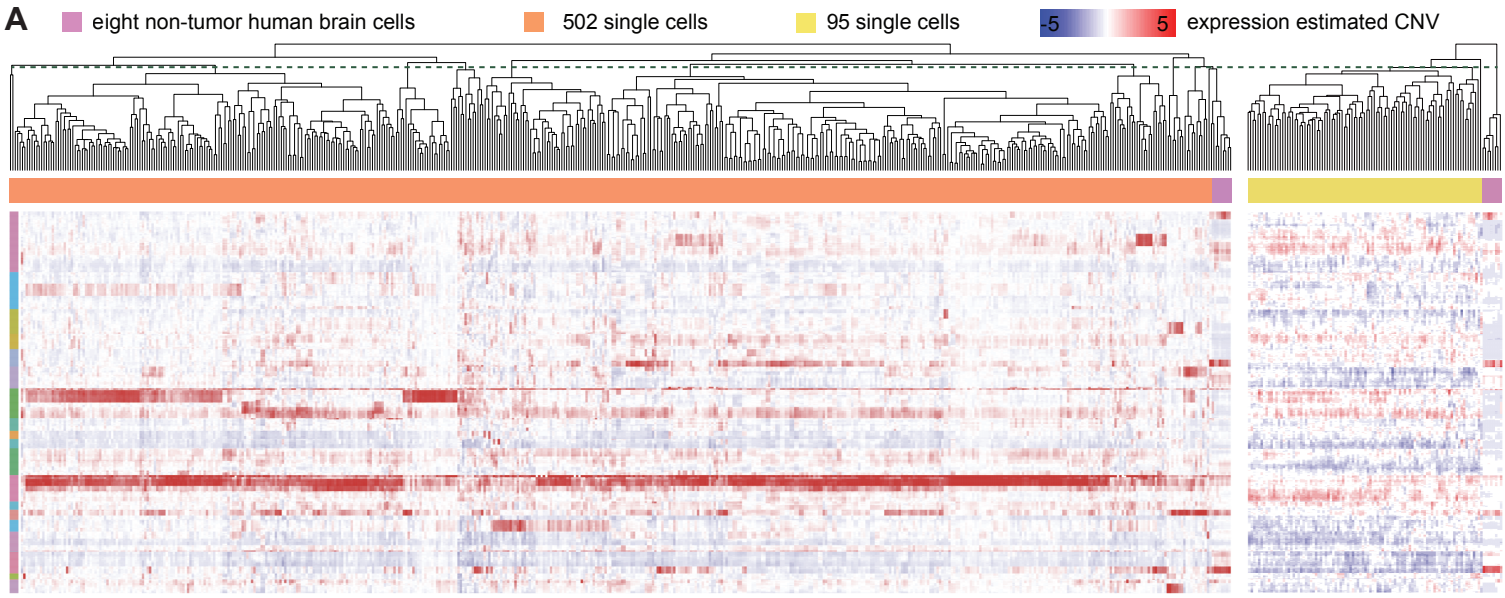
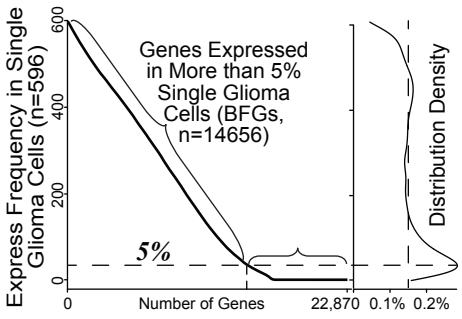
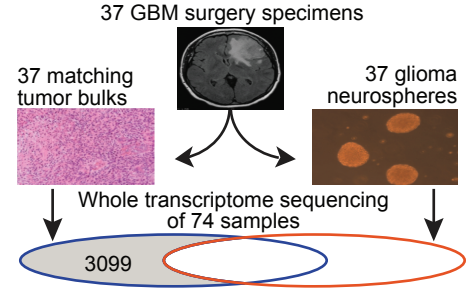
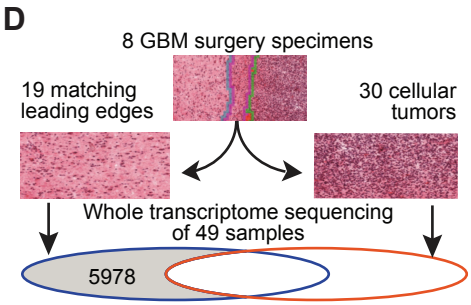
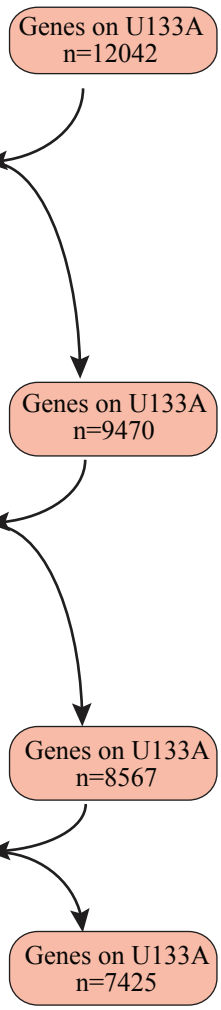
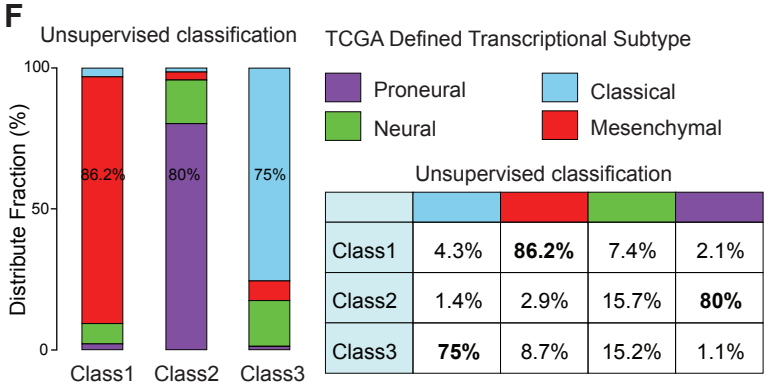






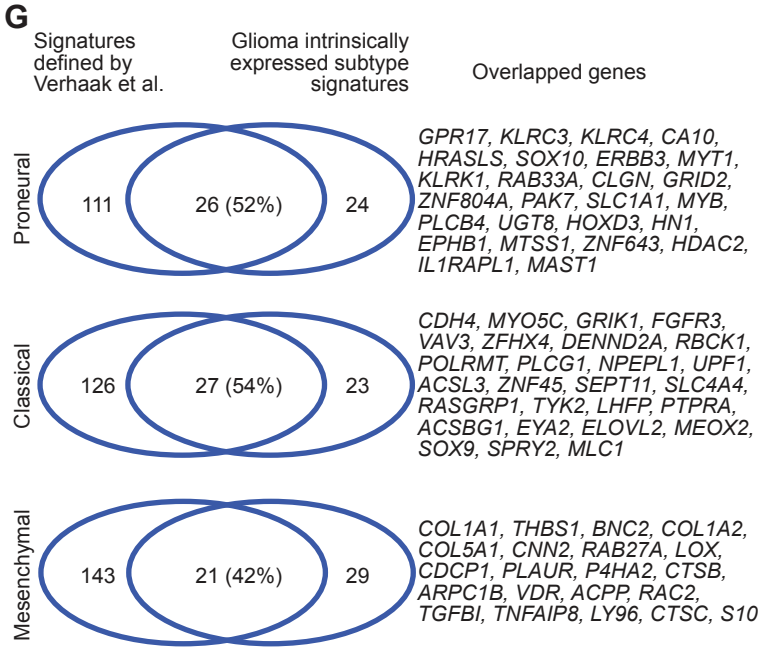






Unsupervised classification

Class1	4.3%	**86.2%**	7.4%	2.1%
Class2	1.4%	2.9%	15.7%	**80%**
Class3	**75%**	8.7%	15.2%	1.1%



H RNA-seq Profiling

Microarray Profiling	PN	CL	MES	
PN	***39***	0	1	40
CL	2	***53***	3	58
MES	2	2	***42***	46
	43	55	46	144

Concordance=**93%**

I Measured by KPKM/TPM

Measured by RPKM	PN	CL	MES	
PN	***13***	0	0	13
CL	0	***30***	0	30
MES	0	0	***13***	13
	13	30	13	46

Concordance=**100%**

J 369 IDH-WT GBMs Separately

Within 533 GBMs	PN	CL	MES	
87	0	0	87	
5	***144***	0	149	
2	4	***127***	133	
94	148	127	369	

Concordance=**97%**

K Batches Combined

Batches Separately	PN	CL	MES	
79	0	5	84	
5	***105***	0	110	
3	0	***93***	96	
87	105	98	290	

Concordance=**93%**

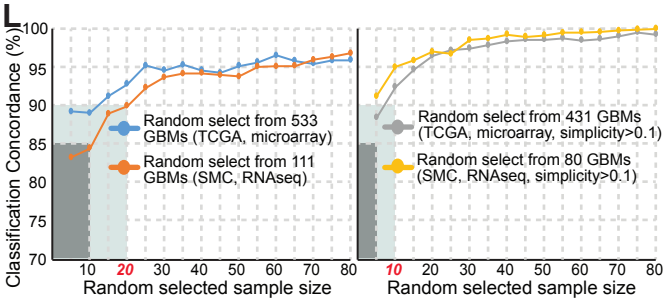


Figure S1. Related to Figure 1.
(**A**) Unsupervised clustering of gene expression established CNV profiles of 502/95 single cells and eight non-tumor cells. Hierarchical clustering was performed. Two batches of single cells were analyzed separately to avoid batch effects. (**B**) Genes expressed in fewer than 5% of a large cohort of single glioma cells were removed. The right panel shows the distribution density of gene express frequency. An unusual occurrence of a distribution peak at 5%. (**C**) Genes significantly higher expressed in GBM tumor bulks pair wisely compared to matching neurospheres (paired Student t-test, FDR adjusted p value<0.01). (**D**) Genes significantly higher expressed in peripheral vs. matching cellular tumors (Student T-test, FDR adjusted p value<0.01). (**E**) The number of genes passing each filtering step and represented on the Affymetrix U133A platform. (**F**) Comparison between IDH-WT GBM specific classification and TCGA defined GBM subtypes. 256 samples were identified as core samples with positive silhouette width core samples. 94, 70 and 92 samples were unsupervised classified class1, class2 and class3, respectively. The previous four transcriptional subtypes of these 256 samples were determined by TCGA Research Network (Brennan et al., 2013). (**G**) Comparison between TCGA GBM transcriptional subtype signatures and developed glioma cells intrinsically expressed subtype signatures. (**H, I**)Classification concordance between RNA-seq and microarray platforms (**H**), and between different gene expression metrics (**I**). (**J**-**L**) Evaluating the interference of embedded non-IDH-WT on classifying IDH-WT GBM samples (**J**), batch effects (**K**) and sample size influence (**L**).

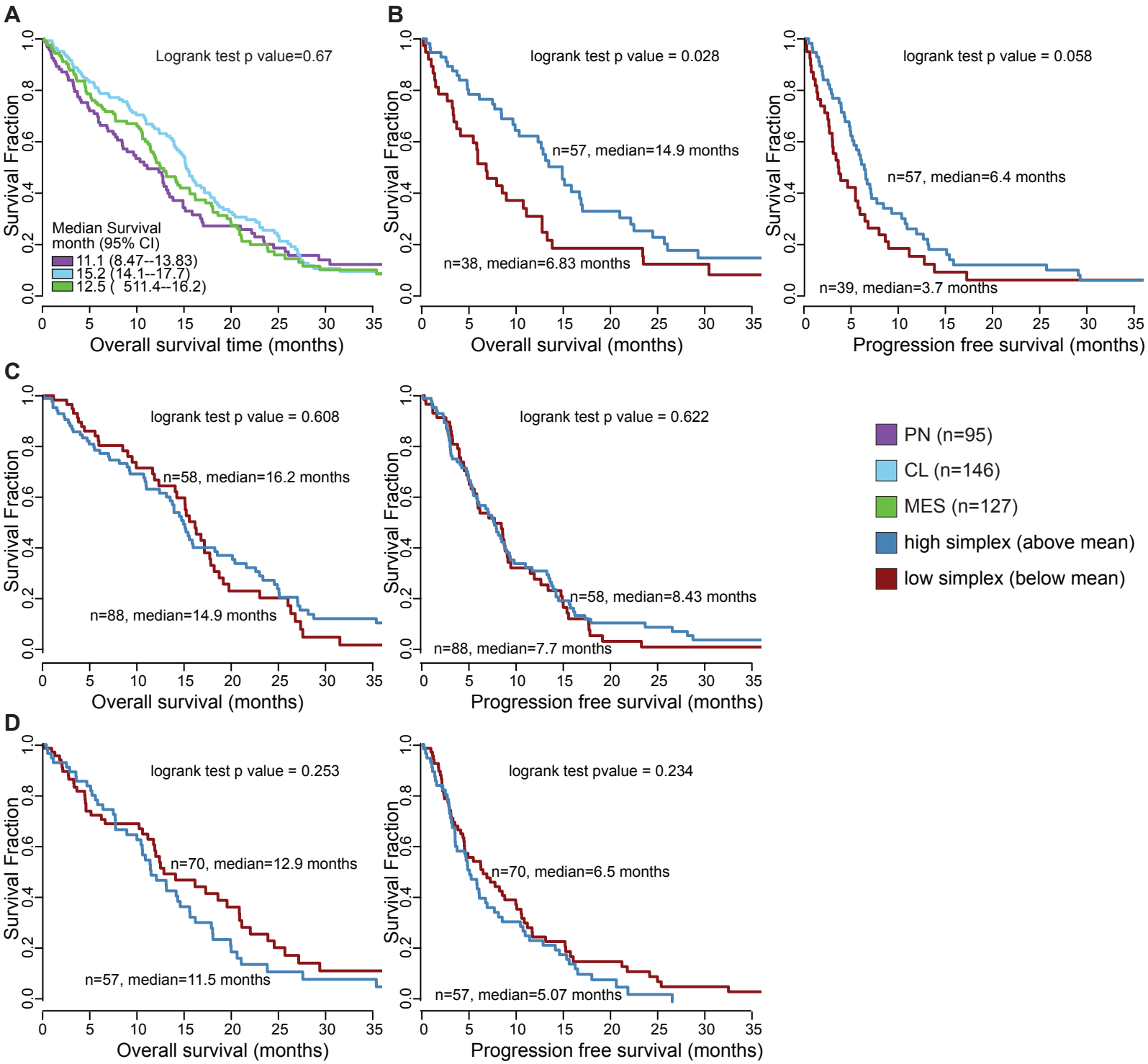
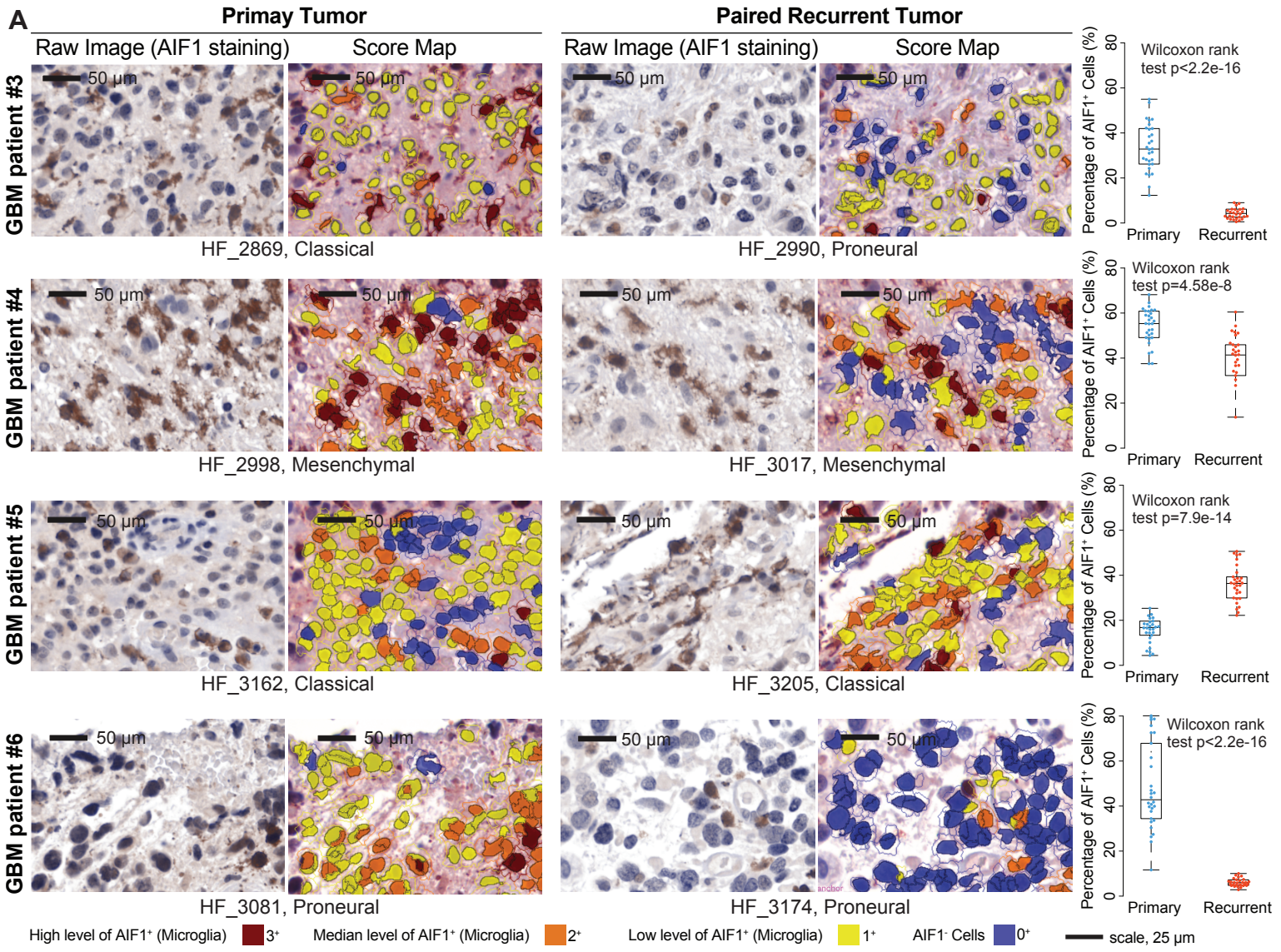
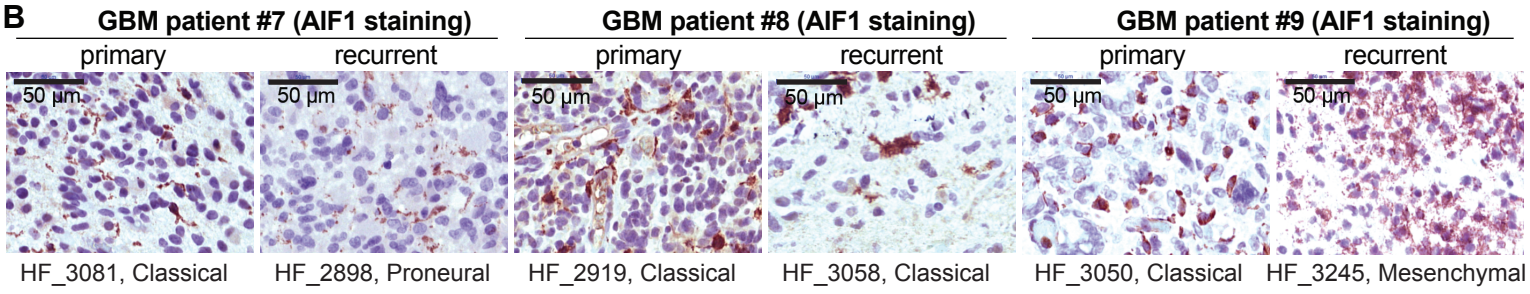


Figure S2. Related to Figure 2.
(**A**) Patient survival differences between transcriptional subtypes. (Note: one of the IDH-WT GBMs in TCGA cohort doesn't have survival data.) (**B-D**) Overall and event free survival analysis comparison between samples with high and low simplicity scores in proneural (**B**), classical(**C**), mesenchymal (**D**).


A
Primay Tumor
Paired Recurrent Tumor
Raw Image (AIF1 staining)
Score Map
GBM patient #3
GBM patient #4
GBM patient #5
GBM patient #6
50 µm
HF_2869, Classical
HF_2990, Proneural
HF_2998, Mesenchymal
HF_3017, Mesenchymal
HF_3162, Classical
HF_3205, Classical
HF_3081, Proneural
HF_3174, Proneural
Percentage of AIF1⁺ Cells (%)
Wilcoxon rank test p<2.2e-16
Wilcoxon rank test p=4.58e-8
Wilcoxon rank test p=7.9e-14
Wilcoxon rank test p<2.2e-16
Primary
Recurrent
High level of AIF1⁺ (Microglia) 3⁺
Median level of AIF1⁺ (Microglia) 2⁺
Low level of AIF1⁺ (Microglia) 1⁺
AIF1⁻ Cells 0⁺
scale, 25 µm


B
GBM patient #7 (AIF1 staining)
GBM patient #8 (AIF1 staining)
GBM patient #9 (AIF1 staining)
primary
recurrent
50 µm
HF_3081, Classical
HF_2898, Proneural
HF_2919, Classical
HF_3058, Classical
HF_3050, Classical
HF_3245, Mesenchymal

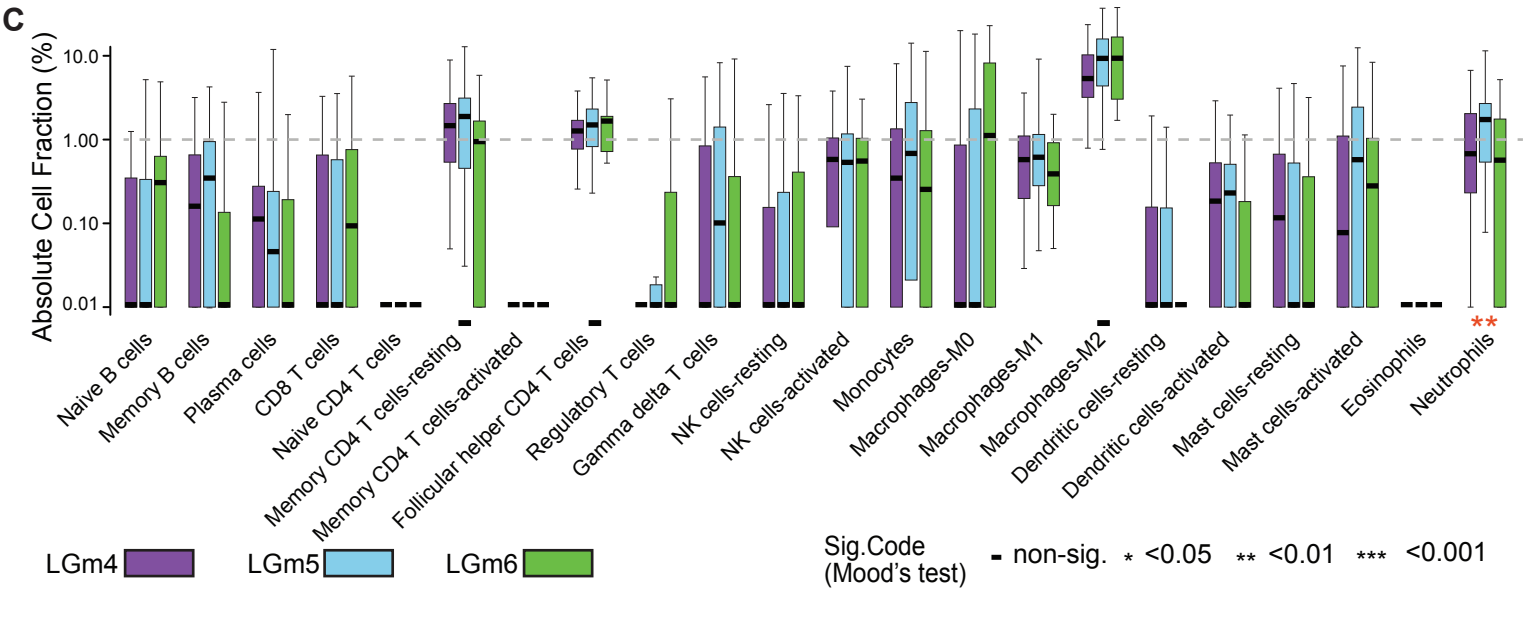

C
Absolute Cell Fraction (%)
10.0
1.00
0.10
0.01
Naive B cells
Memory B cells
Plasma cells
CD8 T cells
Naive CD4 T cells
Memory CD4 T cells-resting
Memory CD4 T cells-activated
Follicular helper CD4 T cells
Regulatory T cells
Gamma delta T cells
NK cells-resting
NK cells-activated
Monocytes
Macrophages-M0
Macrophages-M1
Macrophages-M2
Dendritic cells-resting
Dendritic cells-activated
Mast cells-resting
Mast cells-activated
Eosinophils
Neutrophils
**
LGm4
LGm5
LGm6
Sig.Code (Mood's test)
- non-sig. * <0.05 ** <0.01 *** <0.001

Figure S3. Related to Figure 3 and Figure 5.
(**A**) Representative images with immunohistochemical staining of the AIF1 and score map obtained by InForm image analysis tools in four matched pairs of primary and recurrent GBM. Thirty scan fields were unbiased selected for each tumor by Calipar Vectra pathology imaging system automatically. (**B**) IHC staining of the AIF1 in three additional matched pairs of primary and recurrent GBM. (**C**) Comparison of immune cell fractions among subtypes. Immune cell fractions were estimated using CIBERSORT and corrected using ABSOLUTE purity scores per sample. The distribution of immune cell fractions of 87 LGm4, 157 LGm5 and 33 LGm6 IDH-WT GBMs were shown by purple, skyblue and green box plots, respectively. Median value difference of cell fraction among subtypes was evaluated using Mood's test. (**A, C**) Boxplots represent 25^{th} and 75^{th} percentiles, with midlines indicating the median values and points within the boxes indicating the mean values. Whiskers extend to the lowest/highest values of the data sample exculding outliers.

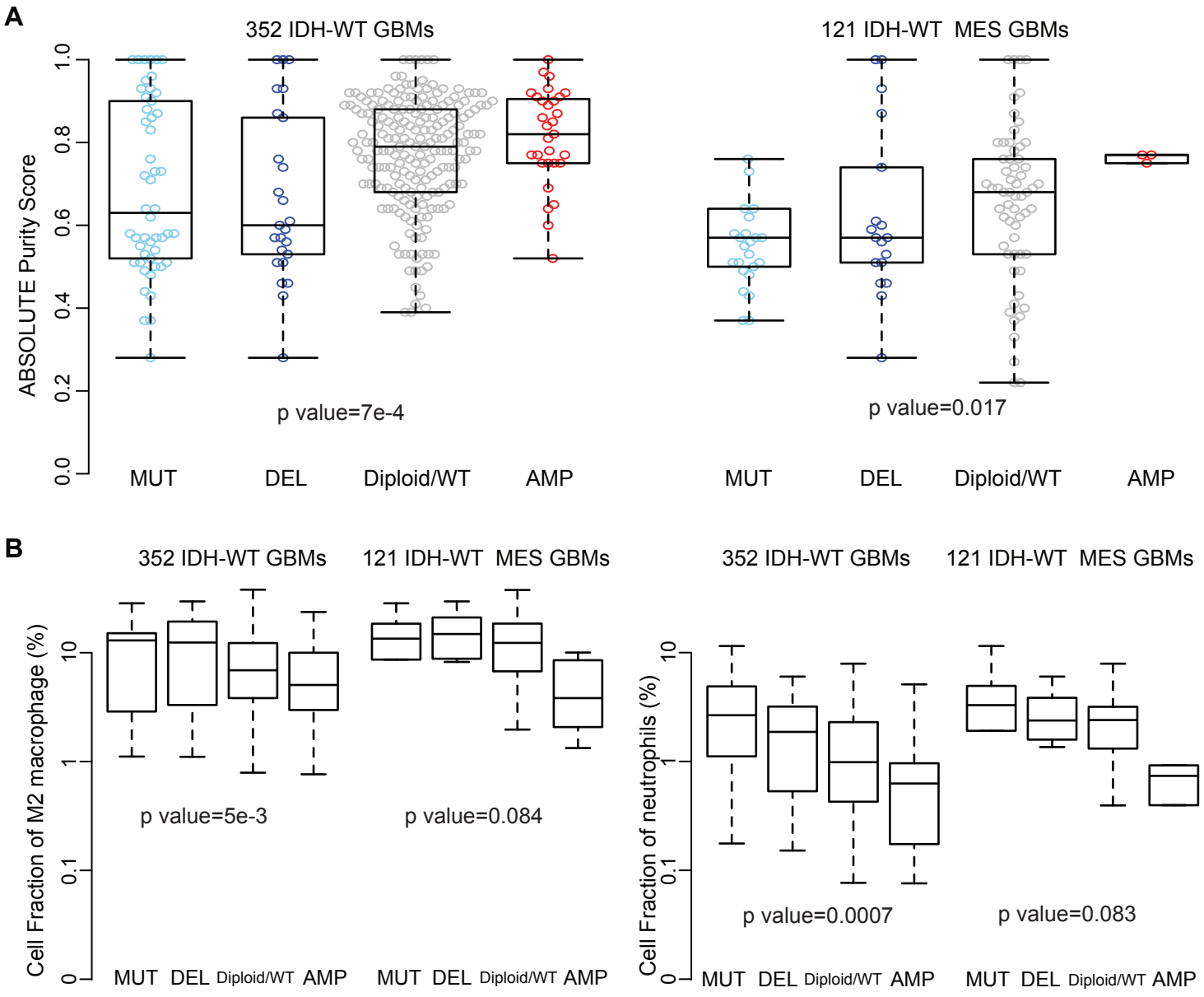


Figure S4. Related to Figure 4.
Comparison of tumor purity (**A**) and immune cell fraction (**B**) between GBMs with different *NF1* genomic status. (**A, B**) Wilcoxon rank test was used to test the difference between *NF1* deficient and the other cases, and p value was shown under the boxplots of each panel. Boxplots represent 25th and 75th percentiles, with midlines indicating the median values and points within the boxes indicating the mean values. Whiskers extend to the lowest/highest values of the data sample exculding outliers (**A, B**).

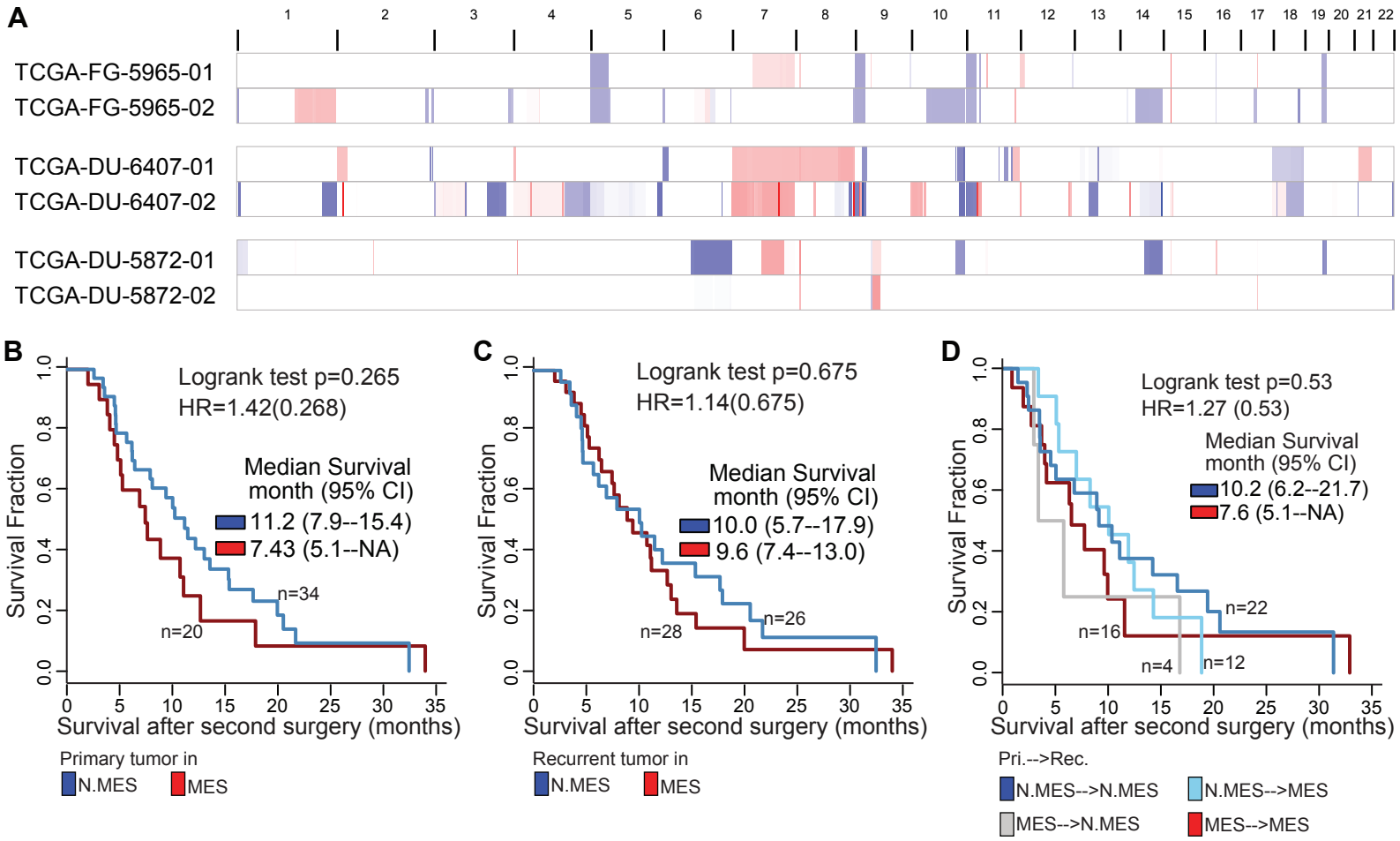


Figure S5. Related to Figure 5 and Figure 6.
(**A**) Copy number variation pattern comparison between primary and recurrent gliomas which switched methylation subtype. (**B-D**) Survival after secondary surgery comparison between primary (**B**) and recurrent (**C**) tumor subytpes, and different transition types (**D**). N.MES indicates non-mesenchymal case.

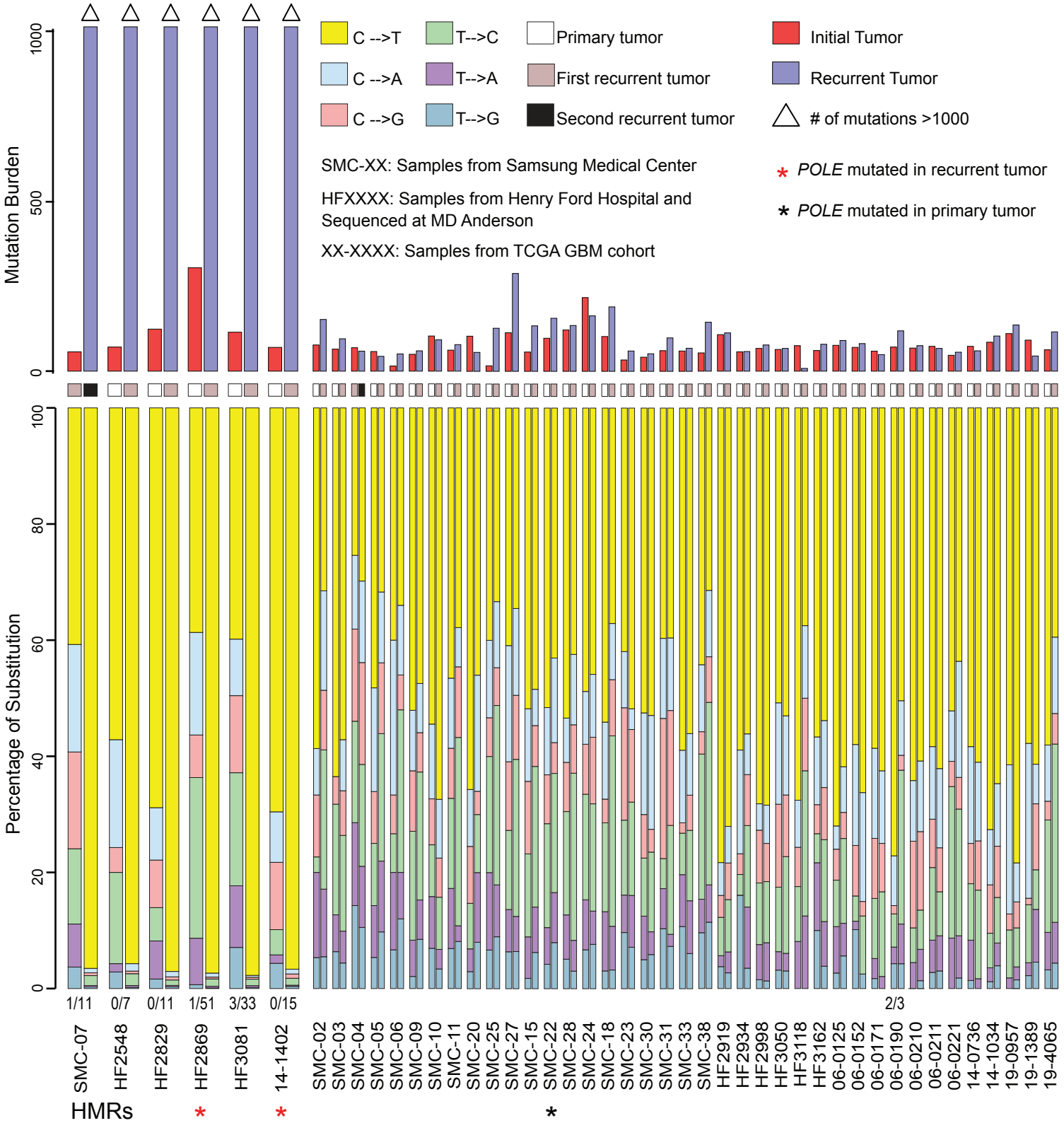


Figure S6. Related to Figure 7.
Mutation Spectrum of Hypermutated GBMs. Numbers below mutation spectrum indicated the mutation burden on the 150 DNA repairing genes which were compiled from MSigDB (Version 5.1) in the hallmark category. Only show number of mutation burden in patients more than 3 (2%) mutated DNA repairing genes. Numbers before and after the slash indicated mutation burden in primary and recurrent tumors, respectively.