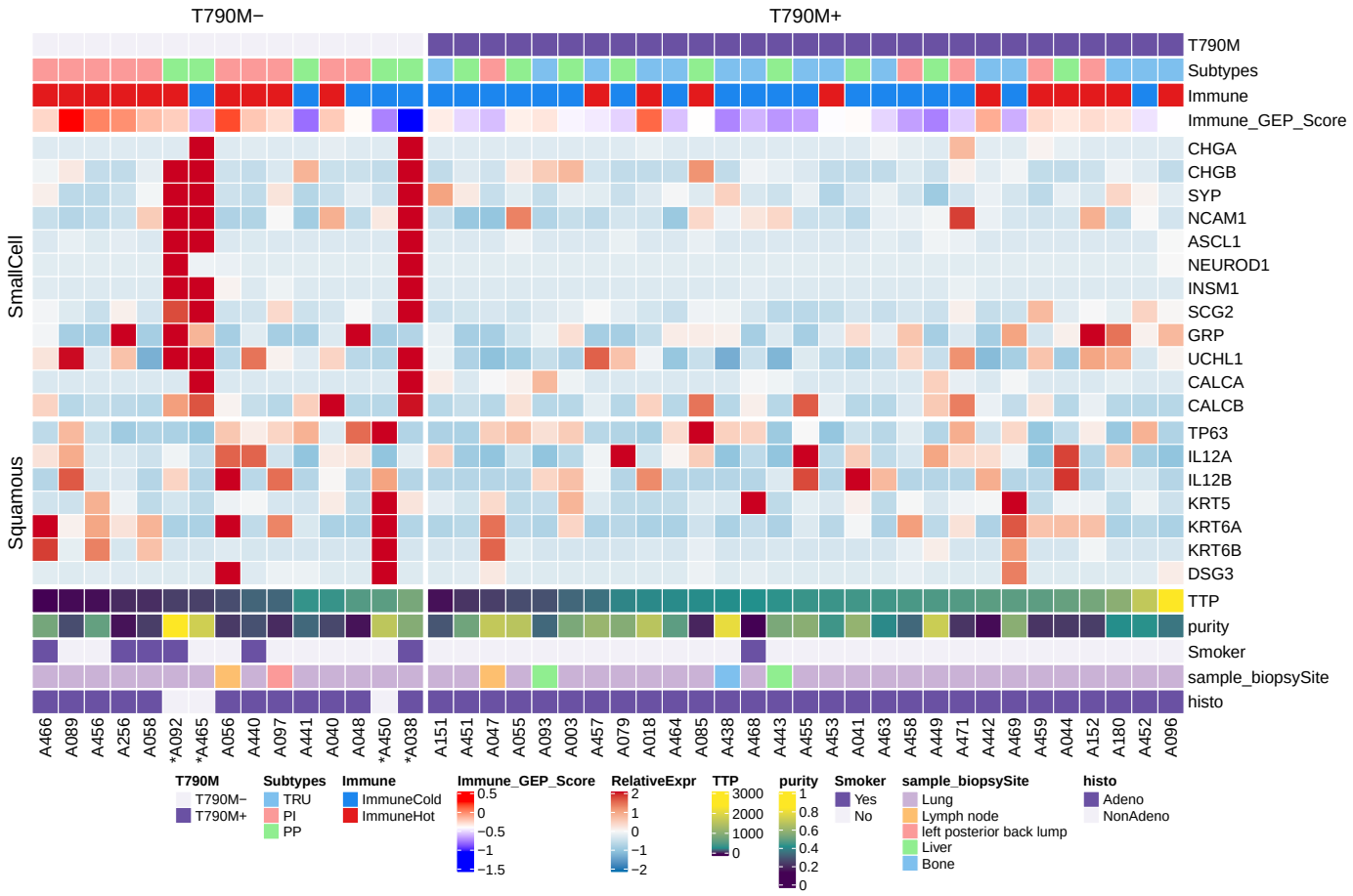
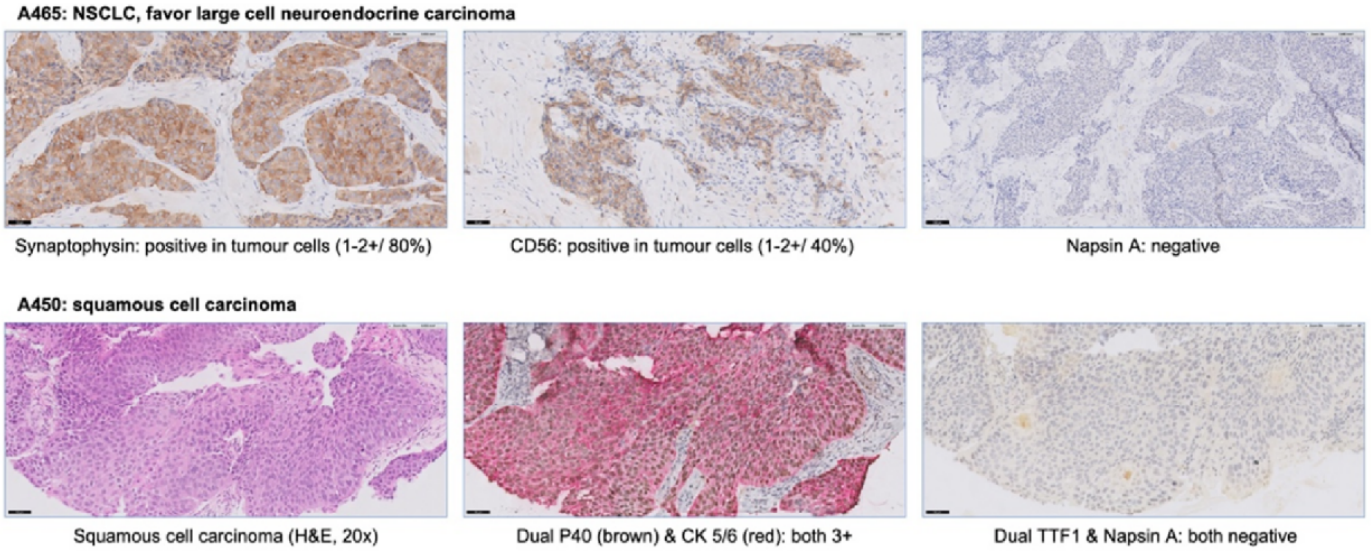
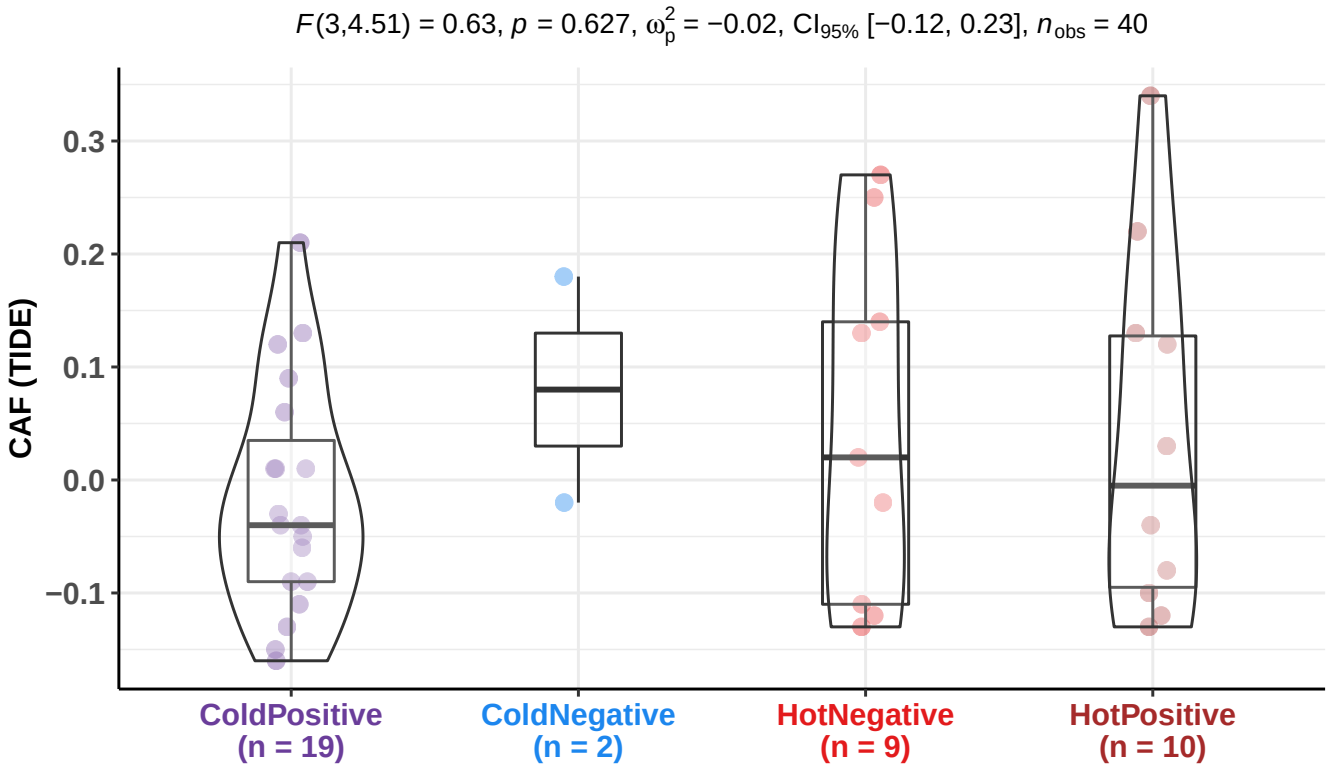


(a)



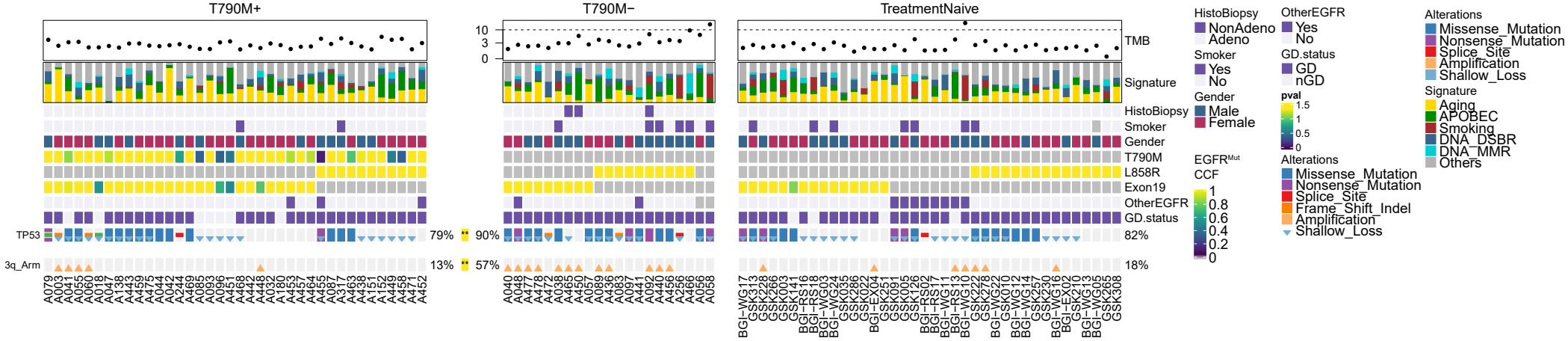
(b)

Figure S1: (a) Neuroendocrine and squamous carcinoma related markers expression of the TKI-resistant patients. Patients with asterisk (*) shows abnormally high expression of these genes and are excluded from TUMERIC deconvolution. (b) IHC review of A450 and A465. After the review, A465 is determined as a large-cell neuroendocrine tumor, showing the utility of RNA-seq in aiding tumor histology classification.



(a)

Figure S2: (a) Comparison of correlation to CAF (Derived with TIDE) for the different immune subtypes.

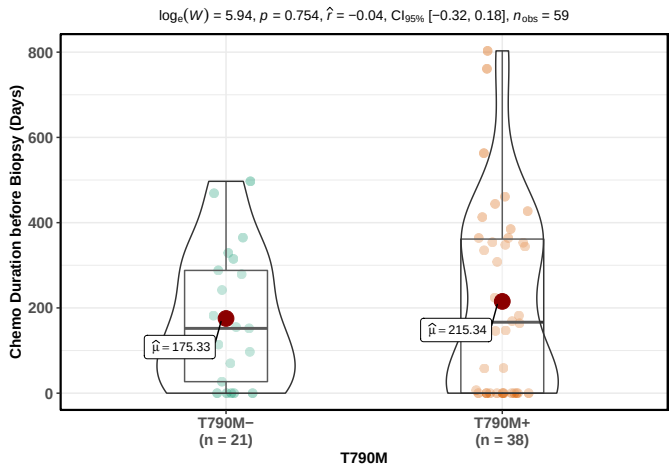


(a)

Figure S3: (a)TP53 mutation frequency including shallow copy loss. Shallow copy number loss is defined for samples with ”-1” in GISTIC’s analysis output file ”all_thresholded_by_gene.txt” file. As most of the treatment-naive samples do not have SNP-array data, we use whole-exome sequencing data for all samples here including the TKI-resistant patients to avoid sequencing bias in the analysis. WES data is segmentated using Sequenza and the segment files were used as input to GISTIC. TP53 mutations were subset to the deleterious mutation in an identical fashion the that of Figure 1 of the main manuscript.

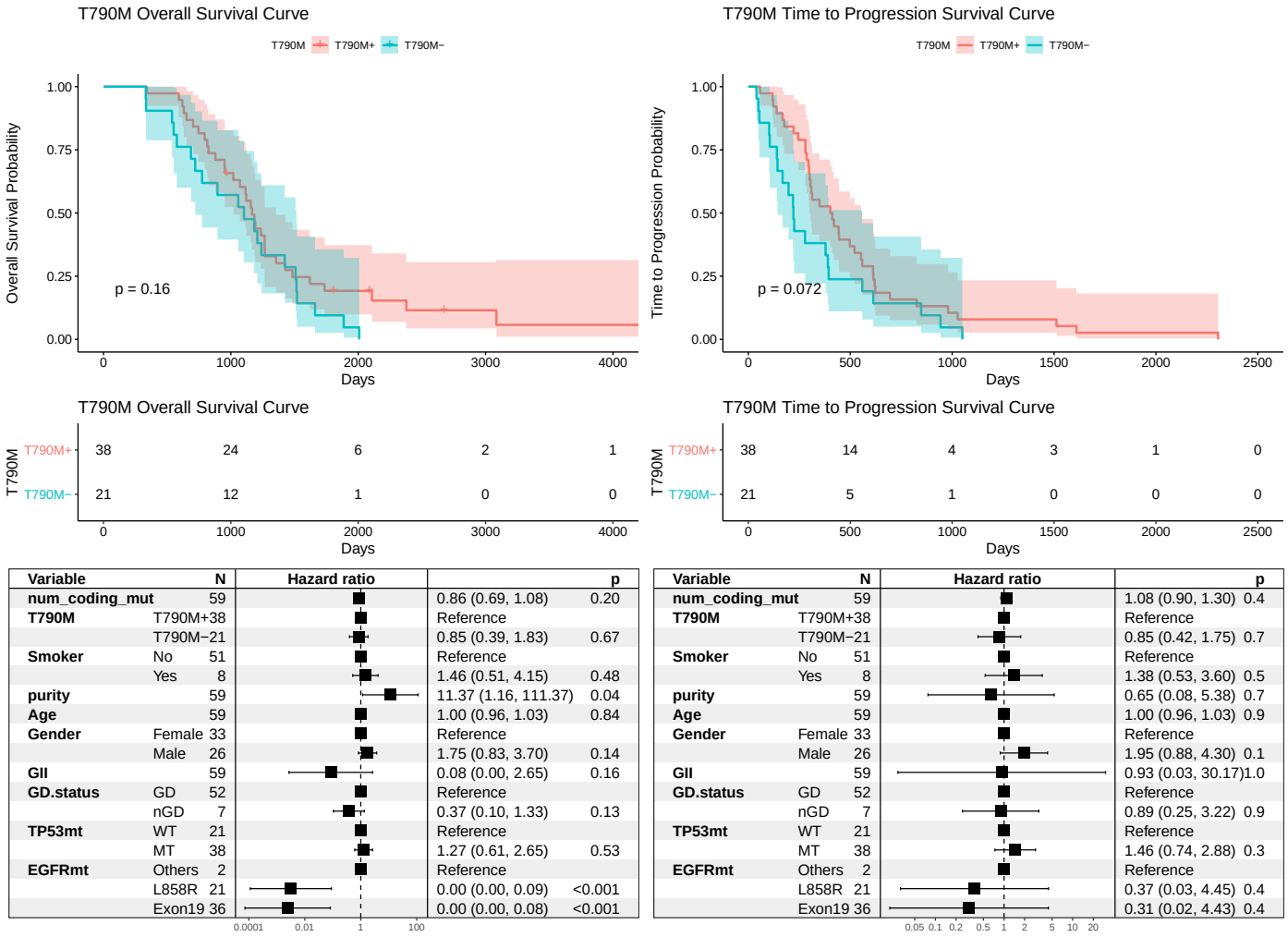
Characteristic	All Patients (n=59)	T790M Positive (n=38)	T790M Negative (n=21)
	N (%)	N (%)	N (%)
Age (years), median (range)	59 (41-79)	60 (41-79)	59 (45-70)
Gender			
Female	33 (56)	25 (66)	8 (38)
Male	26 (44)	13 (34)	13 (62)
Histology			
Adenocarcinoma	54 (92)	36 (95)	18 (86)
Adenosquamous	1 (2)	0 (0)	1 (5)
Carcinoma NOS	3 (5)	2 (5)	1 (5)
Mixed small cell/adenocarcinoma	1 (2)	0 (0)	1 (5)
Smoker			
Current smoker	6 (10)	1 (3)	5 (24)
Ex-smoker	2 (3)	1 (3)	1 (5)
Never smoker	51 (86)	36 (95)	15 (71)
Ethnicity			
Chinese	53 (90)	34 (89)	19 (90)
Malay	3 (5)	2 (5)	1 (5)
Other Asian	3 (5)	2 (5)	1 (5)
Baseline EGFR mutation			
Exon 19 deletion	35 (59)	27 (71)	8 (38)
Exon 19 insertion	1 (2)	0 (0)	1 (5)
L858R	22 (37)	11 (29)	11 (52)
L861Q	1 (2)	0 (0)	1 (5)
EGFR TKI received			
Gefitinib	39 (66)	27 (71)	12 (57)
Erlotinib	6 (10)	3 (8)	3 (14)
Afatinib	3 (5)	1 (3)	2 (10)
More than 1 type of EGFR TKI	11 (19)	7 (18)	4 (19)
Last line of therapy			
TKI alone	23 (39)	14 (37)	9 (43)
TKI + Combination	15 (25)	11 (29)	4 (19)
Chemo alone	21 (36)	13 (34)	8 (38)

(a)



(b)

Figure S4: (a) Clinical characteristics of the TKI-resistant cohort. (b) Comparison of chemotherapy duration prior to the biopsy taken for the study. Figure was generated using R package ggstatsplot using Wilcoxon non-parametric test.



Variable		N	Hazard ratio	p
num_coding_mut		59	0.86 (0.69, 1.08)	0.20
T790M	T790M+	38	Reference	
	T790M−	21	0.85 (0.39, 1.83)	0.67
Smoker	No	51	Reference	
	Yes	8	1.46 (0.51, 4.15)	0.48
purity		59	11.37 (1.16, 111.37)	0.04
Age		59	1.00 (0.96, 1.03)	0.84
Gender	Female	33	Reference	
	Male	26	1.75 (0.83, 3.70)	0.14
GII		59	0.08 (0.00, 2.65)	0.16
GD.status	GD	52	Reference	
	nGD	7	0.37 (0.10, 1.33)	0.13
TP53mt	WT	21	Reference	
	MT	38	1.27 (0.61, 2.65)	0.53
EGFRmt	Others	2	Reference	
	L858R	21	0.00 (0.00, 0.09)	<0.001
	Exon19	36	0.00 (0.00, 0.08)	<0.001

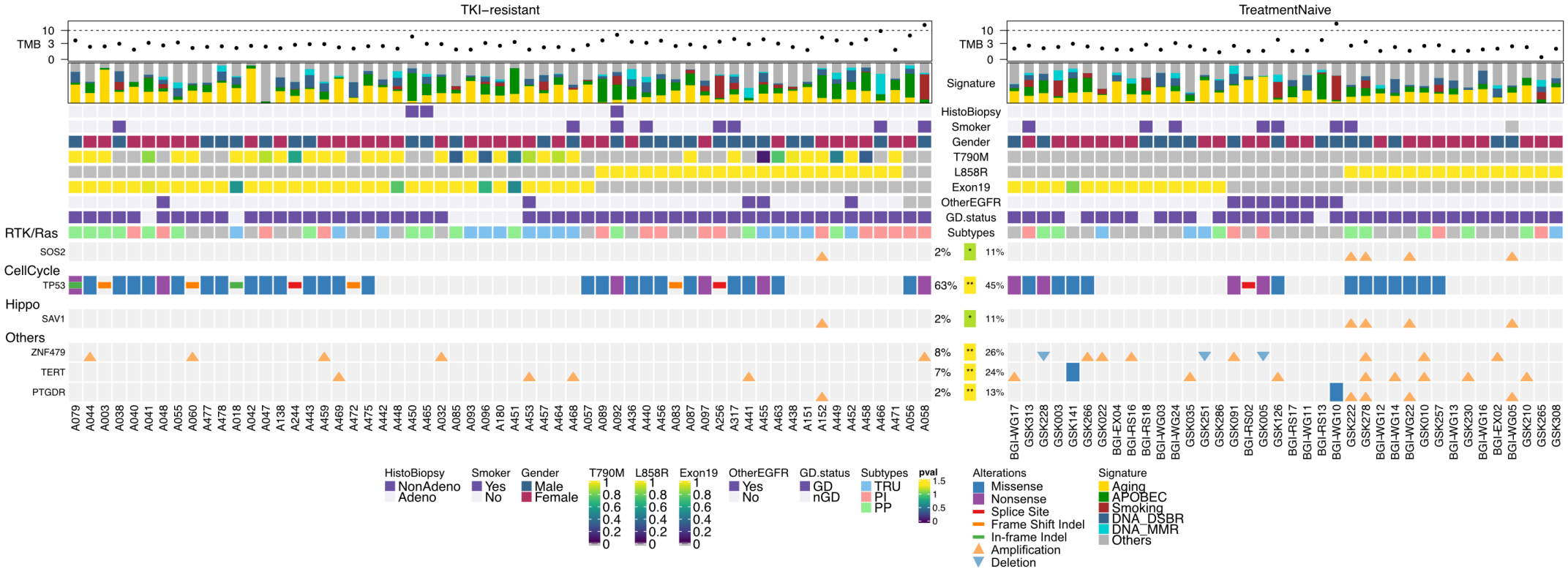
0.0001 0.01 1 100

Variable		N	Hazard ratio	p
num_coding_mut		59	1.08 (0.90, 1.30)	0.4
T790M	T790M+	38	Reference	
	T790M−	21	0.85 (0.42, 1.75)	0.7
Smoker	No	51	Reference	
	Yes	8	1.38 (0.53, 3.60)	0.5
purity		59	0.65 (0.08, 5.38)	0.7
Age		59	1.00 (0.96, 1.03)	0.9
Gender	Female	33	Reference	
	Male	26	1.95 (0.88, 4.30)	0.1
GII		59	0.93 (0.03, 30.17)	1.0
GD.status	GD	52	Reference	
	nGD	7	0.89 (0.25, 3.22)	0.9
TP53mt	WT	21	Reference	
	MT	38	1.46 (0.74, 2.88)	0.3
EGFRmt	Others	2	Reference	
	L858R	21	0.37 (0.03, 4.45)	0.4
	Exon19	36	0.31 (0.02, 4.43)	0.4

0.05 0.1 0.2 0.5 1 2 5 10 20

(a)

Figure S5: Survival analysis of T790M- compared to T790M+ patients. Left column represents overall survival analysis while right column represents time to progression (TTP) analysis. Figures on top show Kaplan-Meyer analysis of survival while tables on the bottom show multi-variate Cox-proportional hazard model including various co-variates as shown.



(a)

Figure S6: Comparing TKI-resistant patients to late-stage treatment-naive patients. Note that in order to avoid sequencing bias, we use whole-exome sequencing data to call all copy number alterations instead of using SNP-array for TKI-resistant patients. The genes that were found to be significantly different were then further interrogated to see if the differences remain if we use SNP-array to call CNA for TKI-resistant patients. The final set of genes that were found to be significantly different ($p <= 0.1$) in both comparisons mentioned were plotted in this figure. Copy number segmentation was done using Sequenza.

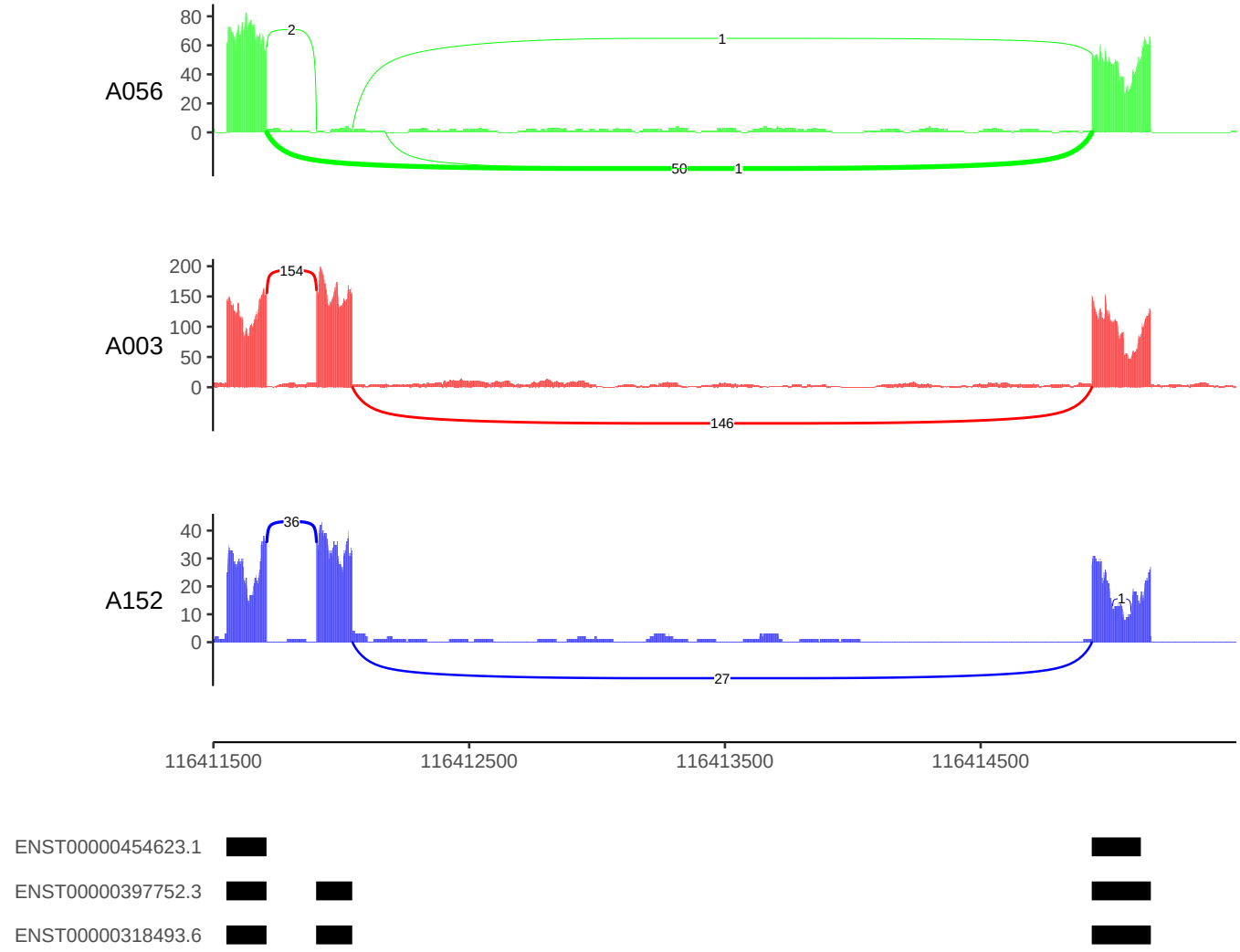
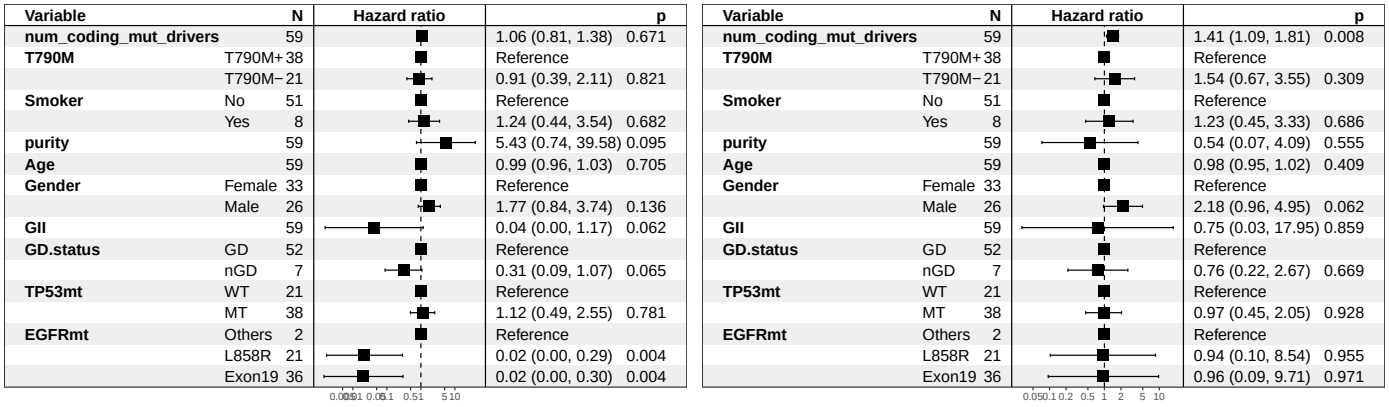
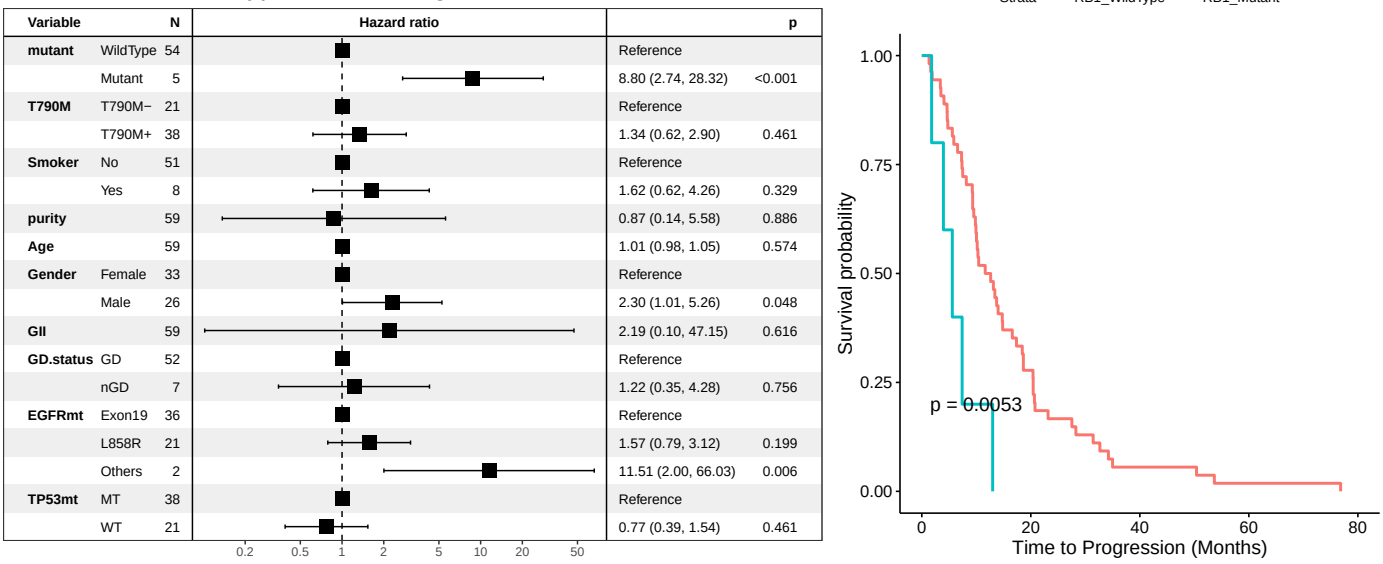


Figure S7: Sashimi plot with RNA transcripts for A056 on exon 13, 14 and 15 for MET gene. Splice site deletion was found for A056 in the exon 14 of MET gene and we confirm the effect of this deletion with transcriptomics data. A003 and A152 are used as the reference samples showing regular transcription of exon 14 (Middle exon).



(a)

RB1 Mutant vs WildType Time to Progression



(b)

Figure S8: (a) Multi-variate Cox-proportional hazard model survival analysis similar to Figure S5 but with number of coding mutations substituted with number of coding mutations in cancer driver genes. (b) Strong association of RB1 to time to progression.

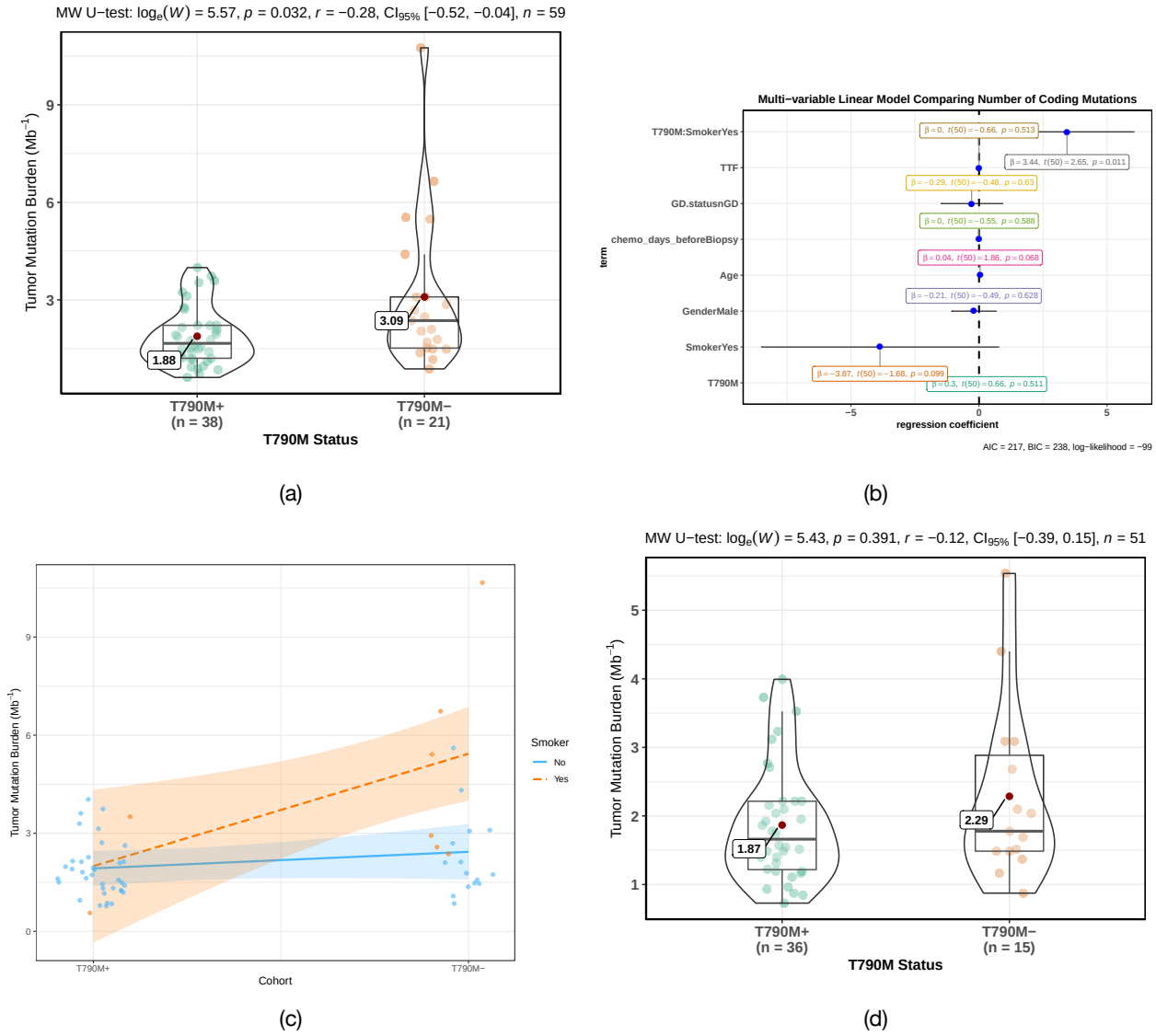


Figure S9: (a) Violin and box plot of tumor mutation burden (TMB, non-synonymous mutations) comparing T790M- to T790M+. Mann-Whitney U-test was carried out to compare the difference between groups. (b) Linear model estimate of TMB using multiple co-variates shown on the y-axis. We found strong interaction between T790M and smoking status. (c) The difference in TMB is mainly contributed by higher TMB in T790M- smoker. (d) After removing smokers, the difference in TMB becomes much less significant.

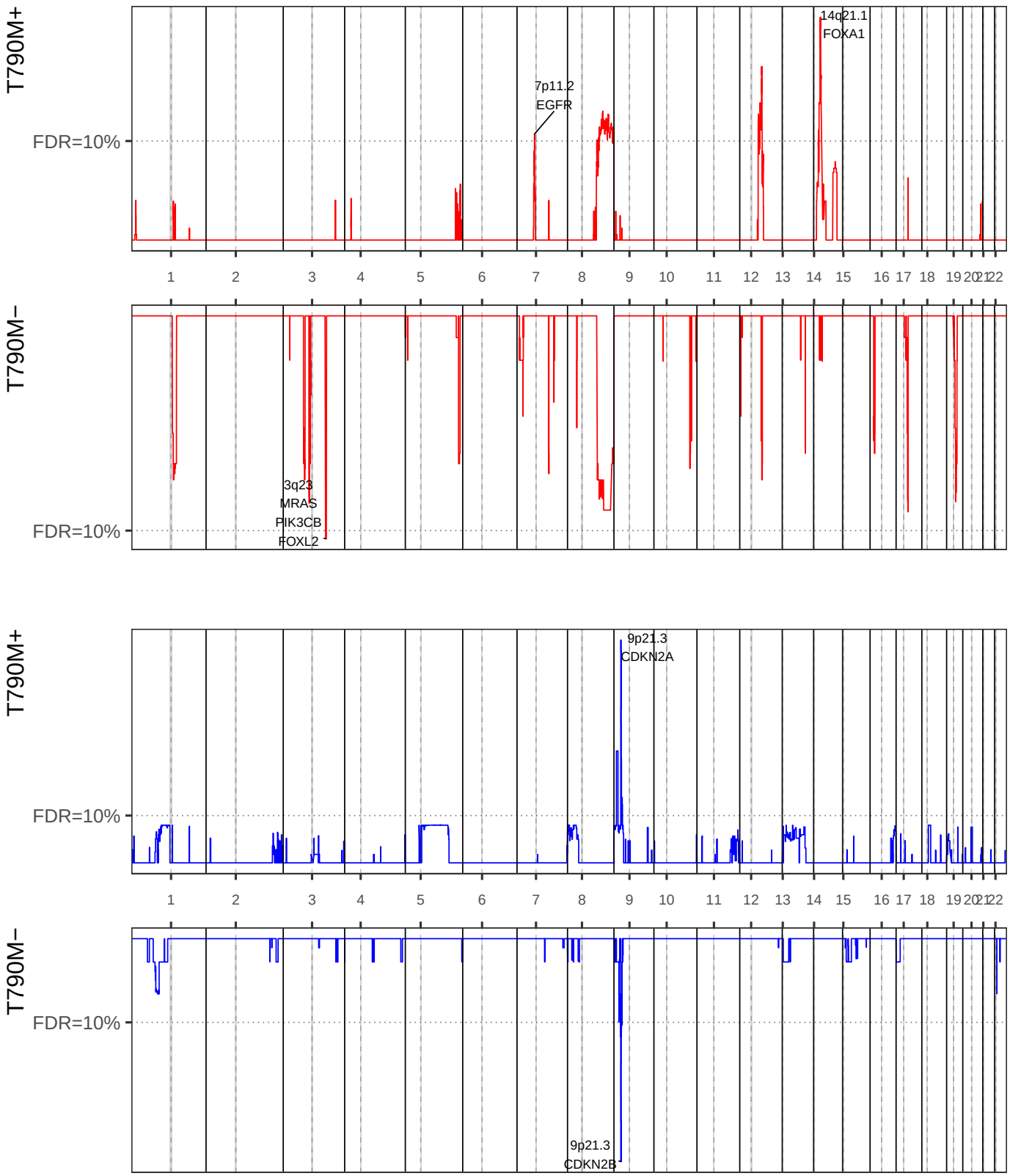
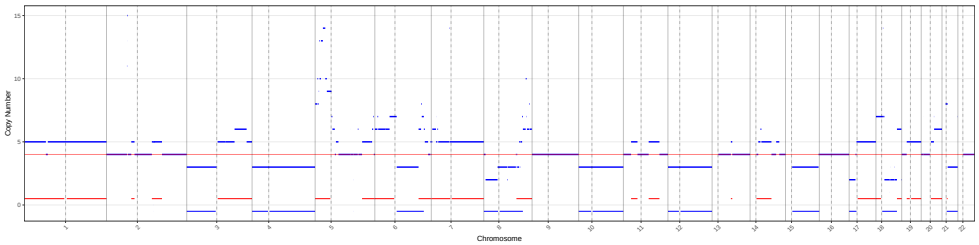
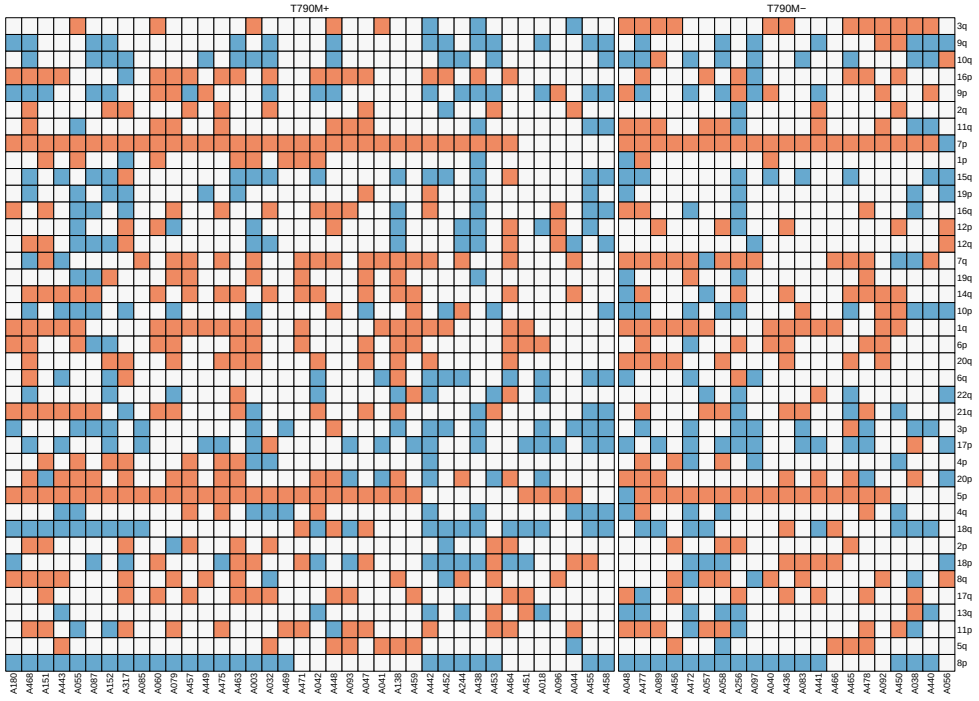


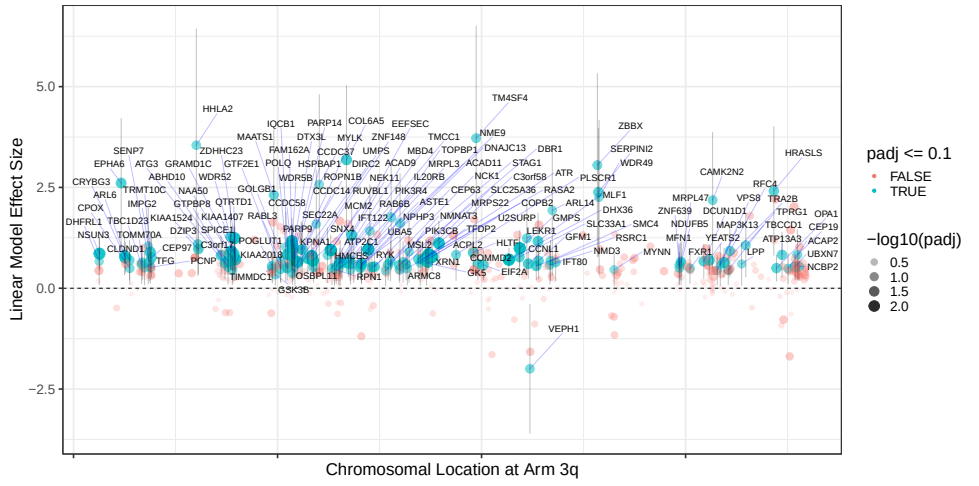
Figure S10: Recurrent focal copy number event in T790M+ and T790M- calculated by GISTIC. (a) Recurrent amplification. (b) Recurrent deletion.



(a)



(b)



(c)

Figure S11: (a) Example of segments joining. For each arm, the longest joined segments in either direction (Red for amplification, blue for deletion) are plotted below. (b) Arm event heatmap of all samples. Red and blue represents chromosomal arm amplification and deletion in more than 80% of the arm, respectively. (c) Expression level of genes on 3q comparing samples with amplifications of arm 3q to those without amplification. For each of the genes on 3q, we use linear model using the expression level as a response variable and the binary presence of 3q amplification on any individual sample as the predictive variable. The coefficient was used as the effect size. P-values were adjusted using the BH procedure to obtain the false discovery rate (FDR) and colored as blue if FDR is below 10%.

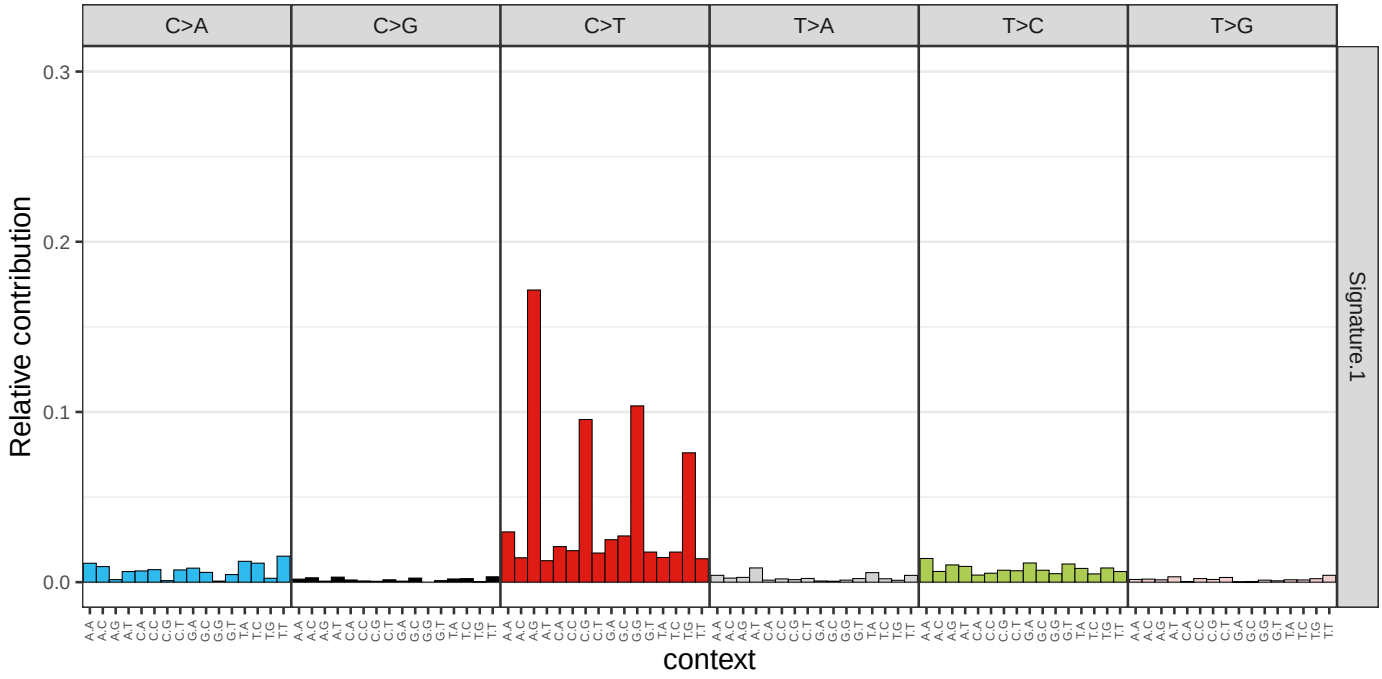
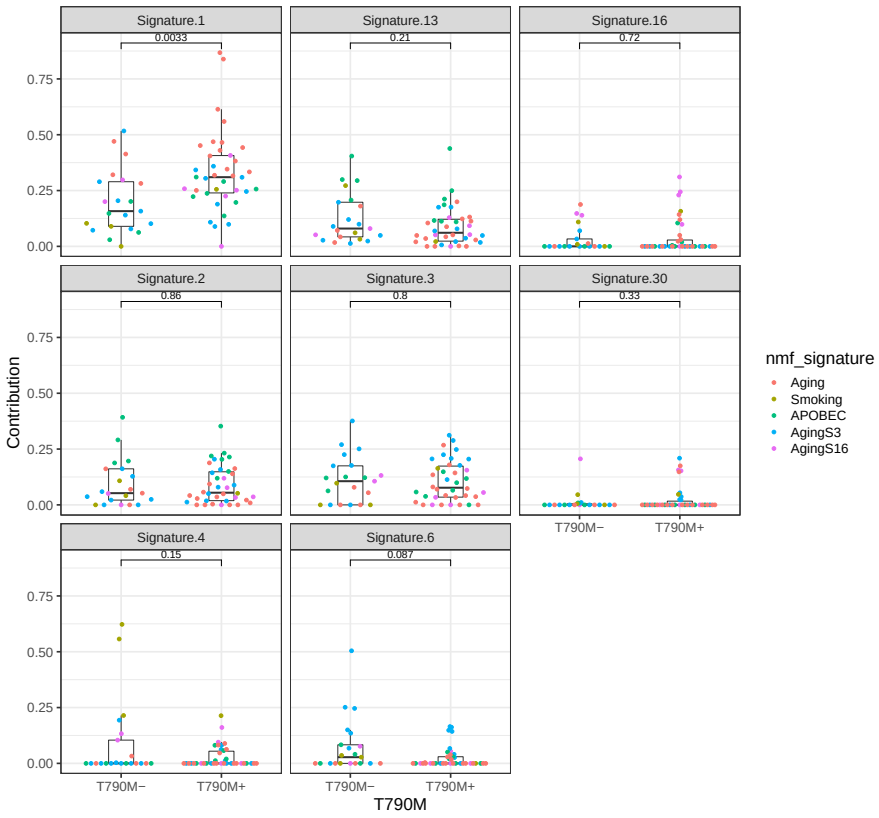
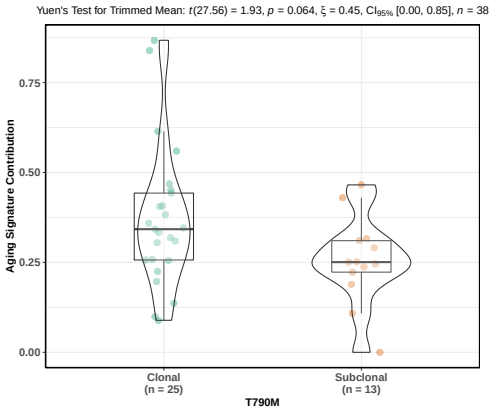


Figure S12: Trinucleotide mutational profile of Signature 1 from COSMIC database. Signature 1 has been associated as an aging signature. The tri-nucleotide context with the highest weight here is A[C>T]G, the same trinucleotide context of EGFR T790M mutation.



(a)



(b)

Linear Model Aging Absolute Count

Variable		N	Estimate	p
T790M	T790M+	38	Reference	
	T790M−	21	−9.07 (−23.87, 5.74)	0.2
Smoker	No	51	Reference	
	Yes	8	−0.83 (−22.97, 21.31)	0.9
Gender	Female	33	Reference	
	Male	26	−2.31 (−17.69, 13.06)	0.8
Age		59	0.22 (−0.58, 1.02)	0.6
purity		59	27.11 (−12.03, 66.26)	0.2
(Intercept)			21.78 (−30.69, 74.24)	0.4

−20 0 20 40 60

Linear Model Aging Proportion

Variable		N	Estimate	p
T790M	T790M+	38	Reference	
	T790M−	21	−0.12 (−0.22, −0.02)	0.02
Smoker	No	51	Reference	
	Yes	8	−0.08 (−0.22, 0.07)	0.30
Gender	Female	33	Reference	
	Male	26	−0.01 (−0.12, 0.09)	0.79
Age		59	−0.01 (−0.01, 0.00)	0.06
purity		59	−0.04 (−0.30, 0.21)	0.73
(Intercept)			0.66 (0.31, 1.01)	<0.001

−0.2 0 0.2 0.4 0.6 0.8 1

(c)

Figure S13: (a) Comparisons of the signatures with significant contributions. (b) Comparison of relative contribution of aging signatures in tumors with clonal T790M mutation (CCF $\geq$ 0.90) against those with subclonal T790M mutation. We used Yuen's test for trimmed mean to account for the two clonal T790M+ samples with much higher contribution of Aging signature. (c) Linear model accounting for other potential co-variates. T790M status remains a strong predictor for aging signature proportion but not aging mutation counts.

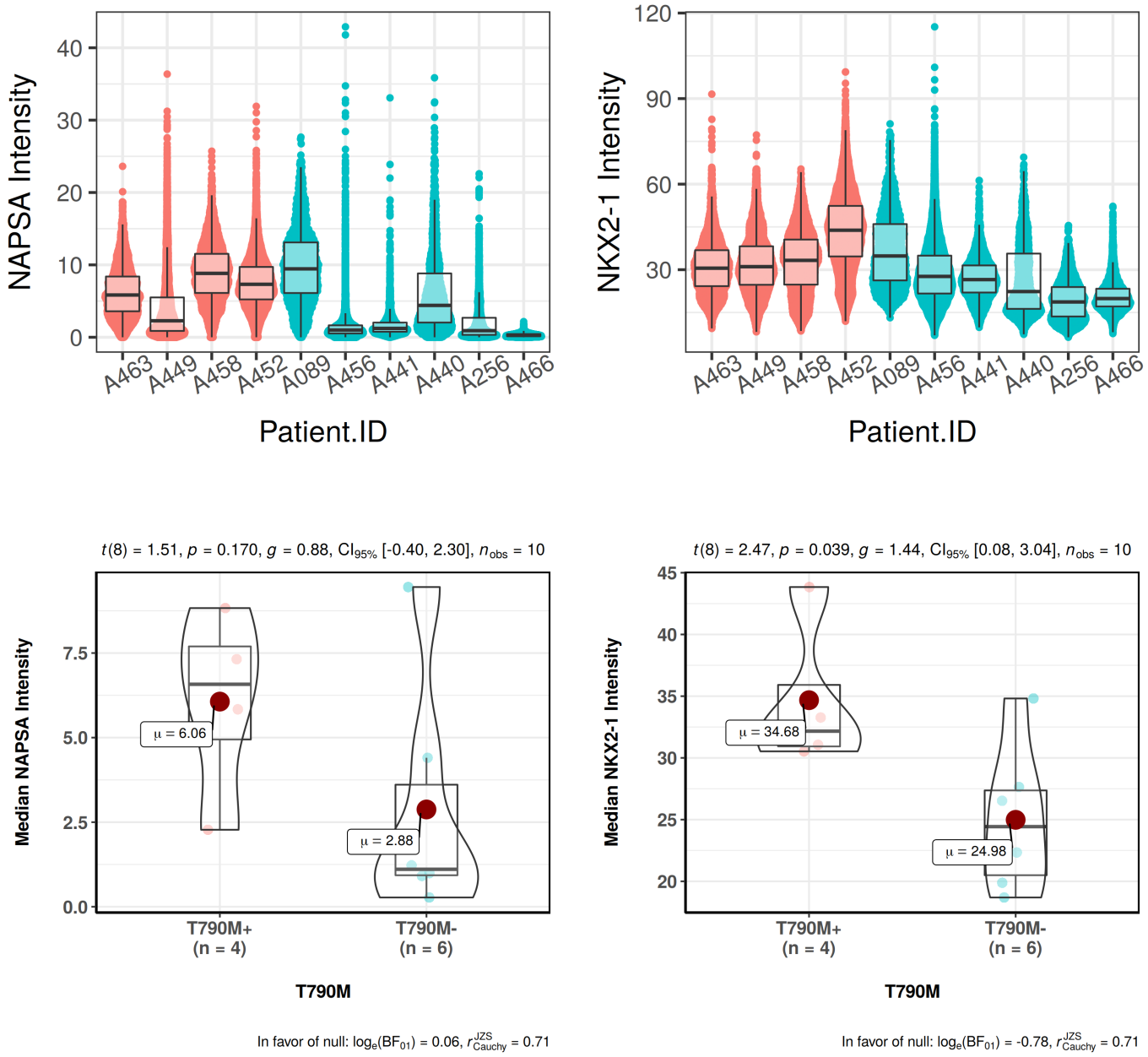
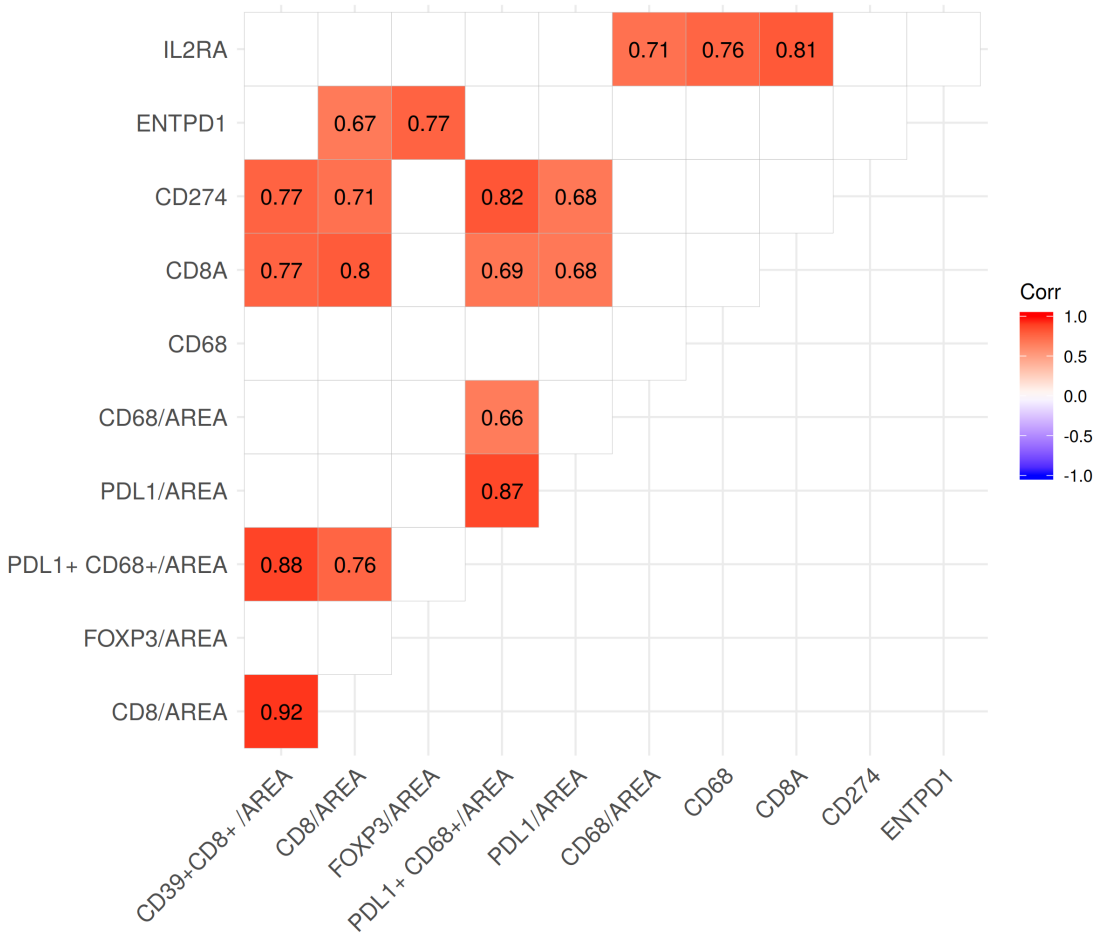
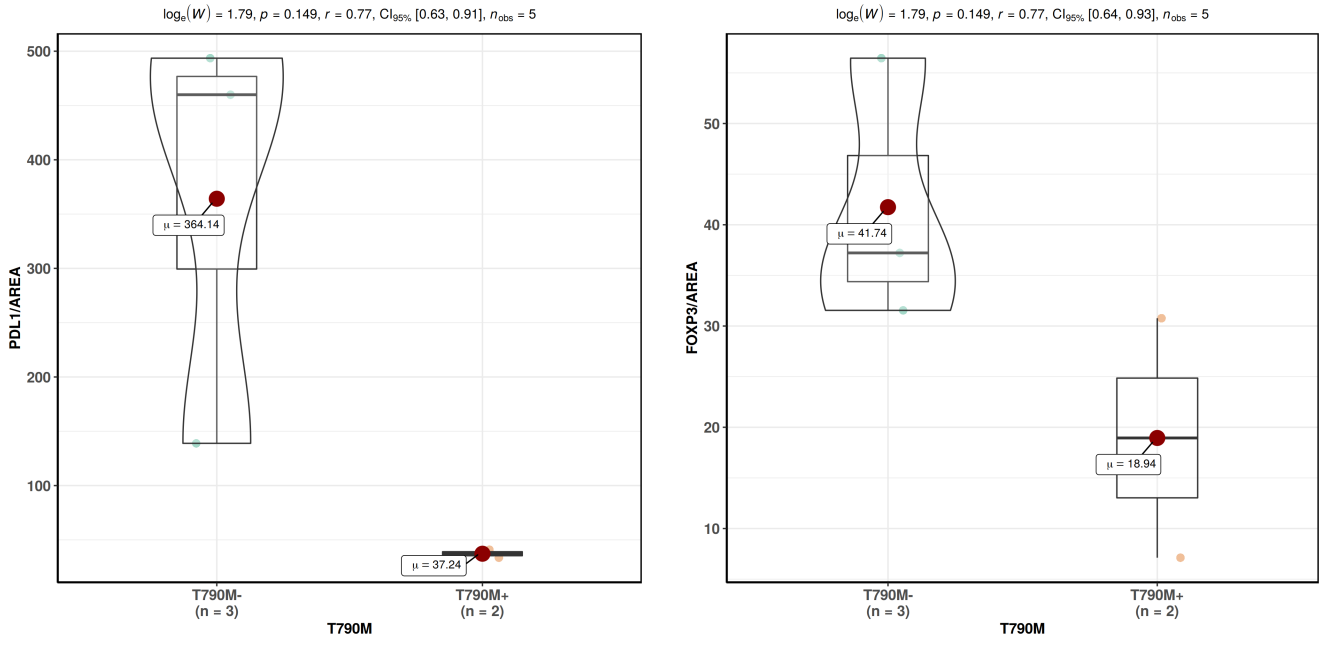


Figure S14: Comparison of multiplexed IHC staining intensity of NAPSA and NKX2-1 in cancer cells (stained using L858R antibody) between T790M-positive and T790M-negative patients. Top row shows the intensity level on an individual patient level. Bottom row compares the median intensity (for individual patient) between T790M-positive and T790M-negative. P-values were calculated using Student's T-test (Assuming equal variance).



(a)



(b) (c)

Figure S15: (a) Correlation of the average number of cells stained positive for different markers in the fields of view for each tissue to the RNA-seq expression of related genes. Note that CD274 is another name for PDL1. (b) and (c) Comparing CD274 (PD-L1) and FOXP3 between ImmuneHot T790M-positive to ImmuneHot T790M-negative. P-values were calculated using the non-parametric Wilcoxon test.

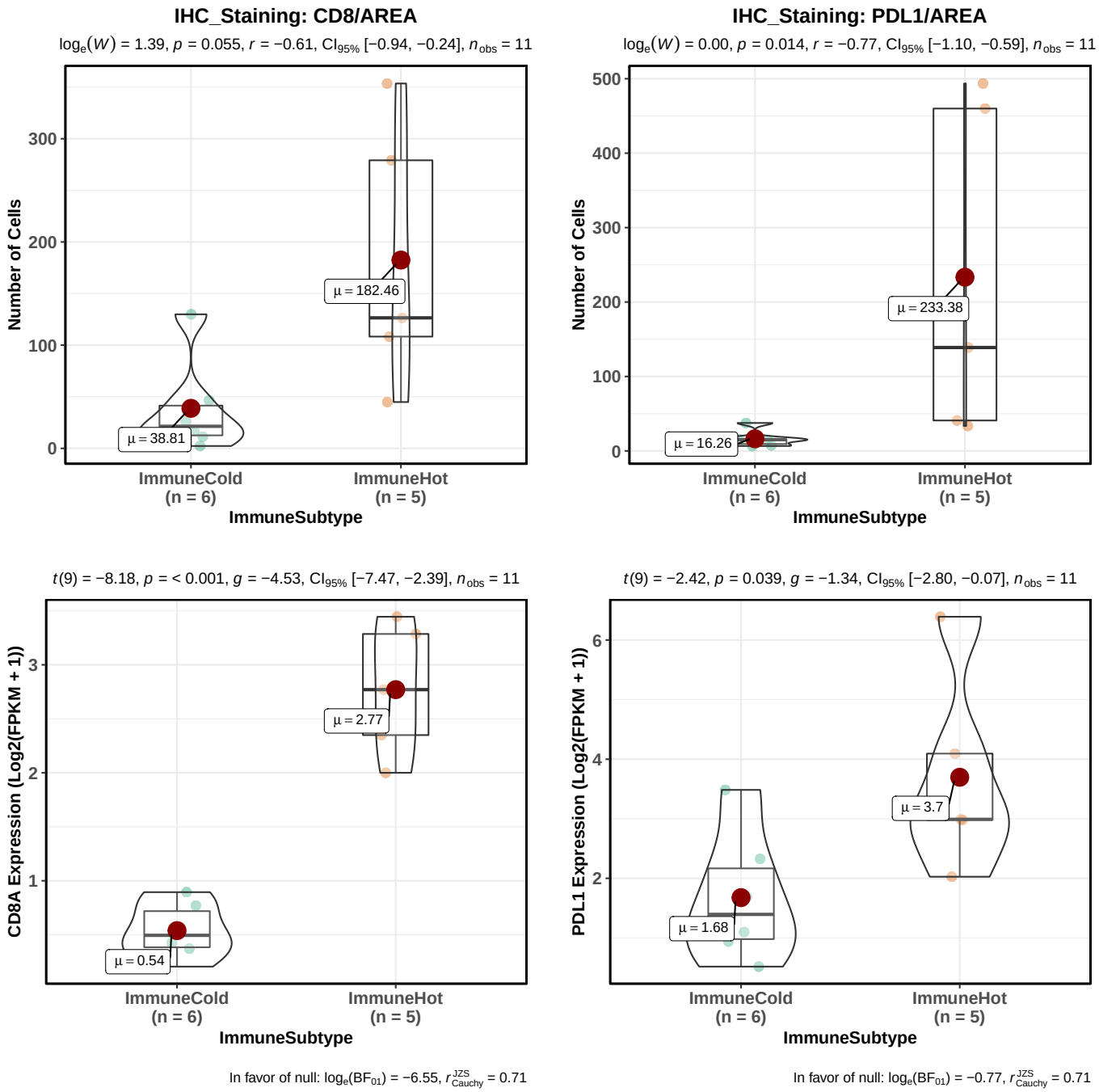
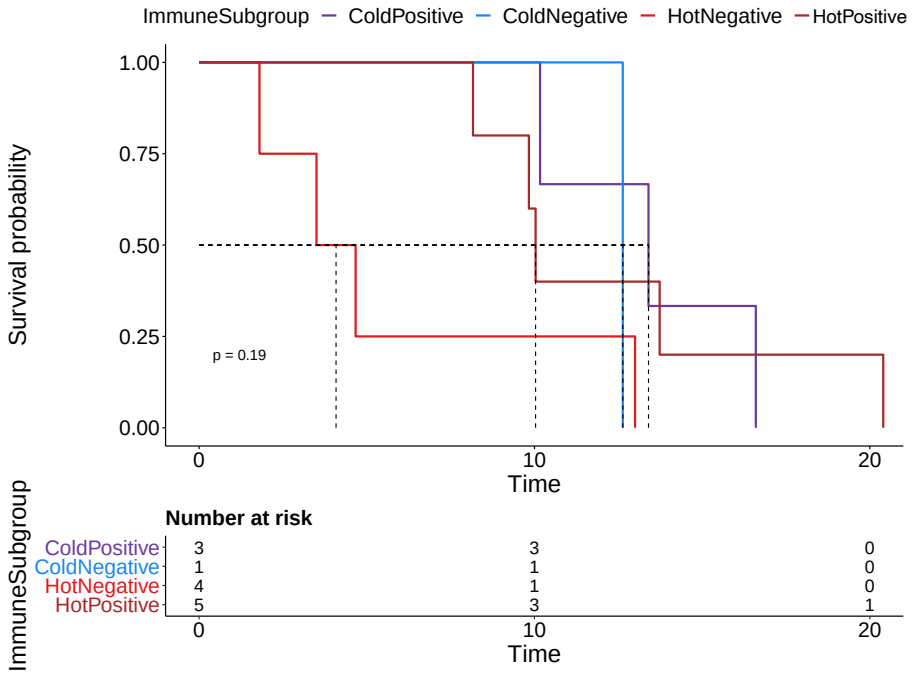
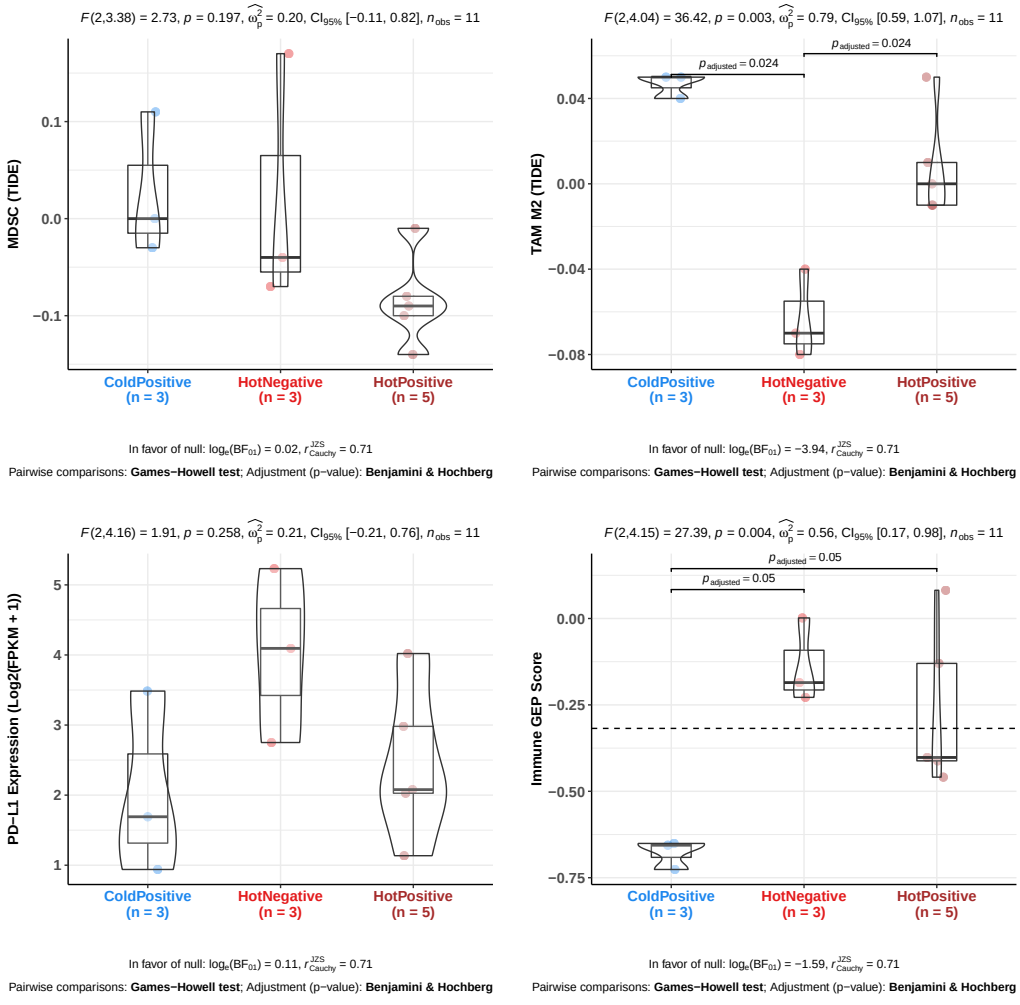


Figure S16: (Top) Comparison of number of cells (averaged over all fields of view) stained for CD8A and PDL1 using multiplexed IHC. P-values are calculated with Mann-Whitney U-test. (Bottom) Comparison of RNA-seq expression (Log_2 (FPKM + 1)) of PDL1 and CD8A. P-values were calculated using Student's T-test (Assuming equal variance).



(a)



(b)

Figure S17: Figure that aims to reproduce the trend in Main Figure 3 after filtering to only patients that were given TKI only in the three months prior to biopsy. Note that there's only 1 Cold T790M- tumor. (a) Kaplan Meier survival curve comparing different immune subtypes. (b) Comparing MDSC, TAM M2, PDL1 and Immune GEP score between the different immune subtypes. As there is only 1 Cold T790M- tumor, it was removed from the plot as statistic test cannot be carried out with a single datapoint.