SUPPLEMENTARY TABLE TITLES AND LEGENDS

Supplementary Table 1: Tumor Sample WES coverage

Exome Sequencing Coverage			
Patient ID	**Normal**	**Primary**	**Metastatic**
ST_001	148	244	251
ST_002	136	258	309
ST_004	79	388	225
ST_007	139	231	274
ST_009	97	325	267
ST_011	113	179	180
ST_013	131	365	234
ST_014	101	248	246
ST_015	259	309	559
ST_016	114	266	271
ST_017	107	309	272
ST_018	146	235	316
ST_019	157	279	295
ST_020	131	279	390
ST_021	145	231	306
ST_022	127	264	350
ST_023	142	384	309
ST_025	150	272	356
ST_026	112	250	266
ST_027	237	258	266
ST_028	117	279	339
ST_029	131	230	279
ST_030	65	331	139
LT_031	133	244	442
LT_033	226	300	316
LT_034	169	428	248
LT_035	142	278	355
LT_036	210	299	332
LT_037	213	263	336
LT_038	120	344	288
LT_039	175	389	218
LT_041	167	275	319
LT_042	186	338	181
LT_043	90	439	364
LT_044	241	331	278
LT_045	138	396	344
LT_046	187	283	331

Exome sequencing coverage for each patient's normal, primary, and metastatic tumor. Coverage is expressed as the number of times a base would be covered by sequencing reads.

Supplementary Table 2 BRCA1 Somatic Mutations

Patient ID	Clonal Variant	Variant Type	Nucleotide Position	Amino Acid Change	Primary VAF	Metastatic VAF	ClinVar Clinical Significance
ST_009	yes	missense	c.5344A>C	p.F1782V	72.48	46.46	not provided
ST_015	yes	frame-shift deletion	c.2419	p.A807fs	36.96	42.76	pathogenic
LT_039	no	missense	c.3419C>G	p.S1140T	56.67	NA	uncertain significance
LT_042	yes	frame-shift deletion	c.2284	p.R762fs	92.04	75.77	pathogenic

All BRCA1 somatic mutations identified in tumor cohort for WES with a VAF>30. Nucleotide position and amino acid changes are based on GRCh37.

Supplementary Table 3 BRCA1 Germline Mutations

Patient ID	Variant Type	Nucleotide Position	Amino Acid Change	VAF	ClinVar Clinical Significance
ST_021	missense	c.4798G>C	p.P1600A	53.77	conflicting interpretations of pathogenicity
ST_027	missense	c.1487C>T	p.R496H	43.96	benign
ST_027	missense	c.455A>G	p.L152P	59.51	conflicting interpretations of pathogenicity
ST_029	missense	c.4931G>C	p.A1644G	51.67	pathogenic/likely pathogenic
LT_031	nonsense	c.2035T>A	p.K679*	57.81	pathogenic
LT_031	missense	c.736A>C	p.L246V	55.29	benign
LT_037	missense	c.5396T>C	p.D1799G	42.86	conflicting interpretations of pathogenicity
LT_044	nonsense	c.2138G>C	p.S713*	37.01	pathogenic

All BRCA1 germline mutation identified in patient cohort from WES with a VAF>30. Nucleotide position and amino acid changes are based on GRCh37.

Supplementary Table 4 BRCA2 Germline Mutations

Patient ID	Variant Type	Nucleotide Position	Amino Acid Change	VAF	ClinVar Clinical Significance
LT_039	missense	c.3568C>T	p.R1190W	47.89	benign

All BRCA2 germline mutations identified in patient cohort from WES with a VAF>30. Nucleotide position and amino acid changes are based on GRCh37.

Supplementary Table 5 Somatic TP53 mutations

Patient ID	Clonal Variant	Variant Type	Nucleotide Position	Amino Acid Change	p53 Protein Domain	Primary VAF	Metastatic VAF
ST_001	yes	missense	c.536T>C	p.H179R	DNA-binding domain	55.70	23.85
ST_002	yes	missense	c.423G>C	p.C141W	DNA-binding domain	76.04	55.11
ST_007	yes	mssense	c.707T>C	p.Y236C	DNA-binding domain	49.59	36.27
ST_009	yes	missense	c.659T>C	p.Y220C	DNA-binding domain	72.30	50.00
ST_011	yes	missense	c.818C>T	p.R273H	DNA-binding domain	92.86	64.00
ST_013	yes	frame shift deletion	C.636A>-	p.R213fs	DNA-binding domain	92.54	78.82
ST_014	yes	missense	c.818C>T	p.R273H	DNA-binding domain	47.57	25.00
ST_015	yes	frame shift deletion	c.250del	p.A84fs	NA	38.18	29.93
ST_015	no	5' UTR	c.-7579816G>T	NULL	NA	6.59	NA
ST_016	yes	frame shift deletion	c.297_294GGAA>-	p.S99fs	DNA-binding domain	65.69	39.93
ST_017	yes	missense	c.487A>T	p.Y163N	DNA-binding domain	53.65	77.71
ST_018	yes	frame shift insertion	c.371_370->A	p.C124fs	DNA-binding domain	53.23	50.00
ST_019	yes	in-frame deletion	c.478_473TGGCGC >-	p.RAM158in_frame_delL	DNA-binding domain	NA	NA
ST_020	yes	missense	c.830C>A	p.C277F	DNA-binding domain	21.28	72.41
ST_021	yes	splice site	c.560-1C>T	e5-1	NA	30.63	25.90
ST_023	yes	nonsense	c.499G>A	p.Q167*	DNA-binding domain	18.60	50.00
ST_025	yes	missense	c.490T>C	p.K164E	DNA-binding domain	49.77	62.31
ST_026	yes	missense	c.580G>A	p.L194F	DNA-binding domain	53.37	55.75
ST_027	yes	frame shift deletion	c.1024G>-	p.R342fs	tetramer	51.16	54.20
ST_028	no	splice site	c.376-2T>C	e4-2	NA	95.05	NA
ST_029	yes	missense	c.764A>C	p.I255S	DNA-binding domain	76.10	89.09
ST_030	yes	missense	c.817G>A	p.R273C	DNA-binding domain	76.00	22.22
LT_031	yes	missense	c.722G>A	p.S241F	DNA-binding domain	63.24	46.01
LT_033	yes	missense	c.524C>T	p.R175H	DNA-binding domain	73.97	67.80
LT_033	yes	intronic	c.96+41_96+34 CTCCAGGT>-	e2+41	NA	NA	NA
LT_034	yes	missense	c.818C>T	p.R273H	DNA-binding domain	23.85	15.97
LT_035	yes	splice site	c.672+1C>1	e5+1	NA	57.79	45.45
LT_036	yes	missense	c.584A>G	p.I195T	DNA-binding domain	90.16	81.64
LT_037	yes	nonsense	c.493G>A	p.Q165*	DNA-binding domain	67.61	80.99

LT_038	yes	missense	c.743C>T	p.R248Q	DNA-binding domain	83.19	61.97
LT_039	yes	nonsense	c.592C>A	p.E198*	DNA-binding domain	70.54	41.90
LT_041	yes	missense	c.748G>A	p.P250S	DNA-binding domain	62.39	53.60
LT_041	yes	missense	c.747C>A	p.R249S	DNA-binding domain	62.39	53.28
LT_041	no	intronic	c.96+41_96+3CTCCAGGT>-	e2+41	NA	NA	NA
LT_042	yes	frame shift deletion	c.366_350 CACAGACTTGGCTGTCC>1	p.G117fs	DNA-binding domain	NA	NA
LT_043	yes	frame shift deletion	c.801C>-	p.N268fs	DNA-binding domain	82.26	83.78
LT_043	yes	intronic	c.96+41_96+34 CTCCAGGT>-	e2+41	NA	NA	NA
LT_044	yes	missense	c.488T>G	p.Y163S	DNA-binding domain	41.98	96.08
LT_045	yes	missense	c.638C>A	p.R213L	DNA-binding domain	81.82	77.00
LT_046	yes	frame shift insertion	c.853_852->T	p.E284fs	DNA-binding domain	66.67	17.59

Somatic TP53 mutations identified in tumor cohort from WES.

Supplementary Table 6 INTEGRATE Recurrent Gene Fusions

Recurrent gene fusions unique to ST survivors	Number of samples	Number of patients	Number of primary samples	Number of met samples
LINC01176>>ZNRF2	11	10	4	7
ZRANB1>>AL731577.2	10	9	4	6
CCDC32>>CBX3	9	9	6	3
AC060766.7>>SLFN12	9	9	5	4
AL035409.1>>ST6GALNAC5	9	8	5	4
TIAM2>>SCAF8	10	8	6	4
SLC35A1>>AL109930.1/AL354993.1	7	6	3	4
SLC35A1>>AL354993.1	6	5	2	4
AC025178.1>>GOLPH3	5	5	3	2
BMPR1B>>PDLIM5	4	4	3	1
AL049697.1/SLC35A1>>AL109930.1/AL354993.1	3	3	2	1
SLC35A1>>AL109930.1	4	3	1	3
BACH1-IT1>>BACH1	3	3	2	1
FP236383.1>>PDE3A	3	3	2	1
HYDIN>>AC239809.3	3	3	2	1
AL596220.1>>C1orf27	3	3	1	2
PPP1CB>>PLB1	3	2	2	1
COL14A1>>DEPTOR	3	2	1	2
CCDC6>>ANK3	2	2	1	1
SLC35A1>>AL354993.1/AL109930.1	2	2	1	1
TRIO>>ARHGEF3	3	2	2	1
FP236383.1>>COL6A1	2	2	0	2
RNU6-4P>>SNORA20/TCP1	2	2	1	1
ERBB4>>IKZF2	2	2	1	1
TPTE2>>MRPS31P2	2	2	2	0
CATSPERB>>TC2N	2	2	2	0
PSPC1>>ZMYM5	3	2	1	2
RNU6-2>>SNORD10	2	2	2	0
COL4A2>>COL4A1	2	2	1	1
NCALD>>GRHL2	2	2	2	0
LINC00032>>EQTN	2	2	1	1
MECOM>>PHC3	2	2	2	0
RAD9A>>KDM2A	2	2	2	0
SKIL>>PRKCI	2	2	1	1
CYP1B1-AS1>>RMDN2	2	2	2	0

Chart showing the recurrent gene fusions predicted by INTEGRATE for the ST survivor tumor samples. The chart displays a breakdown of number and types of samples each gene fusion was identified among the ST survivors.

Supplementary Table 7 Predicted Collagen Gene Fusions

Gene fusion (5'>>3')	Patient sample	Encompassing and spanning reads detected in sample	In-Frame prediction
COL1A1>>7SK/RN7SK	ST.015.M.FF	8	out-of-frame
COL1A1>>GSN	ST.001.M.FF	4	in-frame
BROX>>COL1A1	ST.015.M.FF	2	in-frame
MALAT1>>COL1A2	ST.018.M.FFPE	76	in-frame
COL1A2>>NEAT1	ST.029.M.FFPE	40	out-of-frame
C3>>COL1A2	ST.027.P.FF	20	in-frame
RPS11>>COL1A2	ST.018.P.FF	9	in-frame
COL1A2>>PABPC1	ST.001.M.FF	7	in-frame
COL1A2>>PCBP2/AC023509.1	ST.002.M.FF	4	in-frame
AC093752.1>>COL1A2	ST.001.M.FF	3	in-frame
COL1A2>>TBC1D4	ST.018.P.FF	3	out-of-frame
COL1A2>>PCBP2	ST.002.M.FF	2	in-frame
COL1A2>>TBCCD1	ST.001.M.FF	2	in-frame
COL1A2>>YWHAB	ST.029.P.FF	2	in-frame
HM13>>COL1A2	ST.021.P.FF	2	in-frame
MALAT1>>COL3A1	ST.028.M.FFPE	169	in-frame
FN1>>COL3A1	ST.018.P.FF	54	in-frame
COL3A1>>RN7SL2	ST.001.M.FF	52	out-of-frame
COL3A1>>RMRP	ST.021.P.FF	10	out-of-frame
CLU>>COL3A1	ST.029.P.FF	4	in-frame
STX7>>COL3A1	ST.029.P.FF	3	out-of-frame
HAX1>>COL3A1	ST.018.P.FF	3	in-frame
MAP2>>COL3A1	ST.018.P.FF	3	in-frame
COL4A2>>COL4A1	ST.001.M.FF ST.011.P.FF	5 2	in-frame in-frame
RPL15>>COL4A2	ST.030.M.FF	2	out-of-frame
FP236383.1>>COL6A1	ST.023.M.FFPE ST.028.M.FFPE	15 28	out-of-frame
COL6A3>>ARID1B	ST.021.P.FF	3	out-of-frame
COL6A2>>FP236383.3	LT.041.M.FFPE	7	out-of-frame
ORC4>>COL6A3	ST.011.M.FF	2	out-of-frame
COL6A6>>TTLL9	ST.027.P.FF	6	out-of-frame
RPS24>>COL12A1	ST.018.P.FF	4	out-of-frame

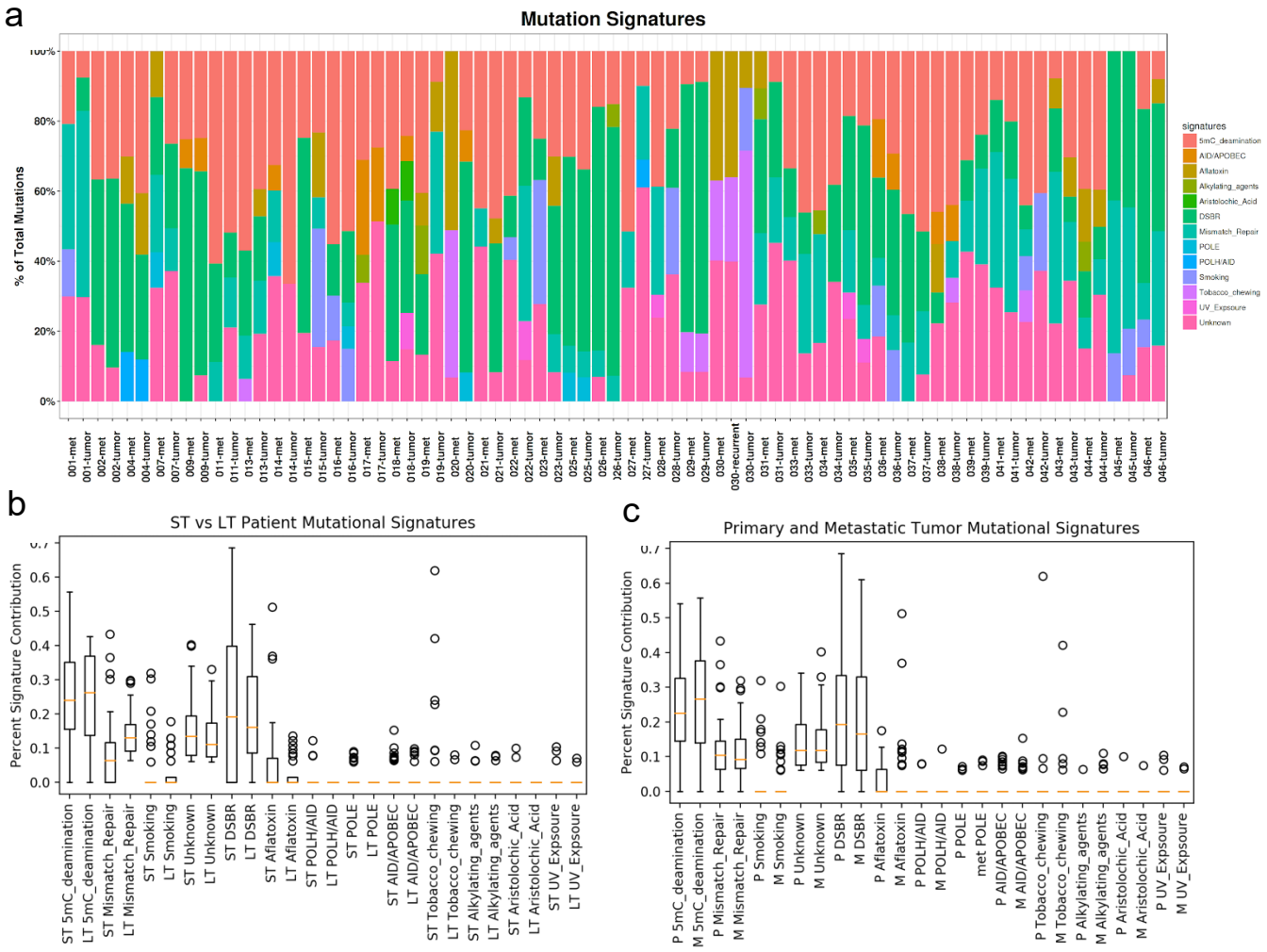
COL14A1>>DEPTOR	ST.007.P.FF ST.007.M.FF ST.030.M.FF	49 3 39	in-frame
TFDP2>>COL18A1	ST.013.M.FF	3	in-frame

Chart displaying the 33 gene fusions involving collagen genes, the sample the fusion was identified in, the number of encompassing and spanning reads that identified the fusion, and INTEGRATE's prediction if the reads were in- or out-of-frame.

Supplementary Table 8: CIBERSORT immune cell fraction statistics

	ST v LT survivors		Primary (P) v Metastatic (M)		ST survivors - PvM		LT survivors - PvM		Primary tumors - ST v LT		Metastatic tumors - ST v LT	
Immune cell type	Mann-Whitney Statistic	p-value	Wilcoxon Statistic	p-value	Wilcoxon Statistic	p-value	Wilcoxon Statistic	p-value	Mann-Whitney Statistic	p-value	Mann-Whitney Statistic	p-value
B cells naive	690	0.28222397	51	0.3793 74669	32	0.582 91955	3	0.465 20882	177	0.370 74399	167	0.5937 86144
B cells memory	647	0.722 85701	196	0.2035 40906	82	0.600 87101	24	0.132 95729	164	0.753 43757	158	0.9094 58667
Plasma cells	696.5	0.241 82612	52	0.4080 44206	32	0.582 91955	4	0.715 00065	189.5	0.118 64898	157	0.9229 74403
T cells CD8	733	0.176 80624	187	0.0360 35175	49	0.010 32448	45	0.972 12533	154.5	1	217.5	0.0407 89302
T cells CD4 naive	784	0.002 87582	23	0.6464 62203	23	0.646 4622	NA	NA	189	0.062 65679	203	0.0227 49719
T cells CD4 memory resting	690	0.395 8775	312	0.7414 58844	104	0.482 59115	24	0.078 49121	208	0.082 55079	133	0.5059 06885
T cells CD4 memory activated	658	0.379 63095	2	0.0250 61844	1	0.046 39946	0	0.179 71249	161	0.458 84883	169	0.4959 12321
T cells follicular helper	356	0.002 71976	259	0.3590 21078	100	0.590 0638	34	0.267 57813	84	0.024 11617	93	0.0495 61507
T cells regulatory (Tregs)	375	0.002 98062	144	0.2796 41534	43	0.861 30426	12	0.008 54492	53	0.000 4148	145	0.7706 23805
T cells gamma delta	658	0.166 19833	1	0.2850 49407	1	0.285 04941	NA	NA	161	0.458 84883	168	0.2699 00195
NK cells resting	581	0.658 98278	141	0.7970 68494	26	0.172 95492	4	0.009 92549	205.5	0.059 63849	80	0.0098 23086
NK cells activated	490.5	0.147 8352	259	0.9255 10561	75	0.420 90992	27	0.196 05096	105	0.114 21541	149	0.8837 53087
Monocytes	676	0.491 60656	301	0.6151 48209	119	0.823 72856	40	0.463 13477	168	0.660 61466	164	0.7578 58154
Macrophages M0	691	0.388 08951	202	0.0641 9212	67	0.054 24547	39	0.649 64403	181	0.389 21972	163.5	0.7690 31965
Macrophages M1	505	0.196 18414	223	0.3042 32069	68	0.277 24064	48	0.807 73926	109	0.133 85961	150.5	0.9222 738
Macrophages M2	499	0.178 38381	321	0.8504 63631	115	0.726 18389	41	0.501 58691	134	0.526 85427	115	0.2115 46541
Dendritic cells resting	664	0.413 30065	34	0.6948 86602	15	0.674 42407	5	1	162	0.707 75726	170	0.4666 76921
Dendritic cells activated	452	0.053 03165	227	0.9099 30768	72	0.556 57941	29	0.432 76758	134	0.511 25239	93	0.0464 126
Mast cells resting	606	0.901 55873	111	0.8757 10309	25	0.151 95619	6	0.092 89194	139.5	0.605 13707	166	0.6768 59087
Mast cells activated	625	0.916 40813	137	0.2115 55387	75	0.943 39206	6	0.028 41686	124.5	0.324 55326	183.5	0.3012 92617
Eosinophils	715	0.240 89382	180	0.8288 12374	85	0.982 62729	19	0.678 40238	174	0.516 89552	182	0.3542 47644
Neutrophils	587.5	0.71791202	78	0.0396 72486	21	0.047 98975	22	0.575 06213	155	0.984 13058	143	0.7187 26623

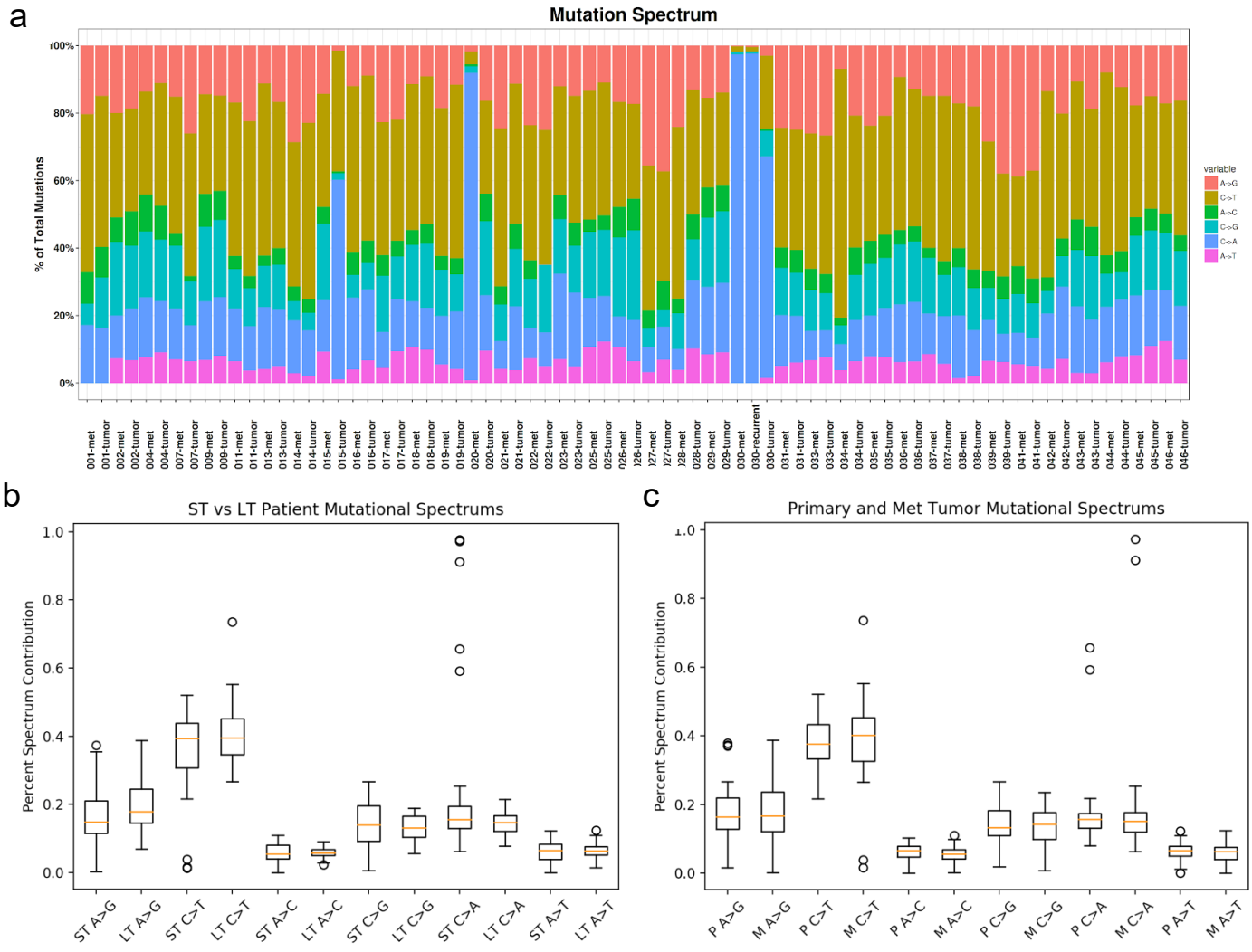
Table listing the Mann-Whitney or Wilcoxon Statistic and corresponding p-values for comparisons between ST and LT survivors and Primary and Metastatic tumor for each of the 22 immune cell types



Supplementary Fig 1 Mutational Signatures

a. Stacked bar plot displaying the percent of total mutations that fit into the mutational signatures shown on the right. Each patient is ordered by patient number and their metastatic and primary tumor with ST survivors on the left (001-met – 030 tumor) and LT survivors on the right (031-met – 046-tumor).

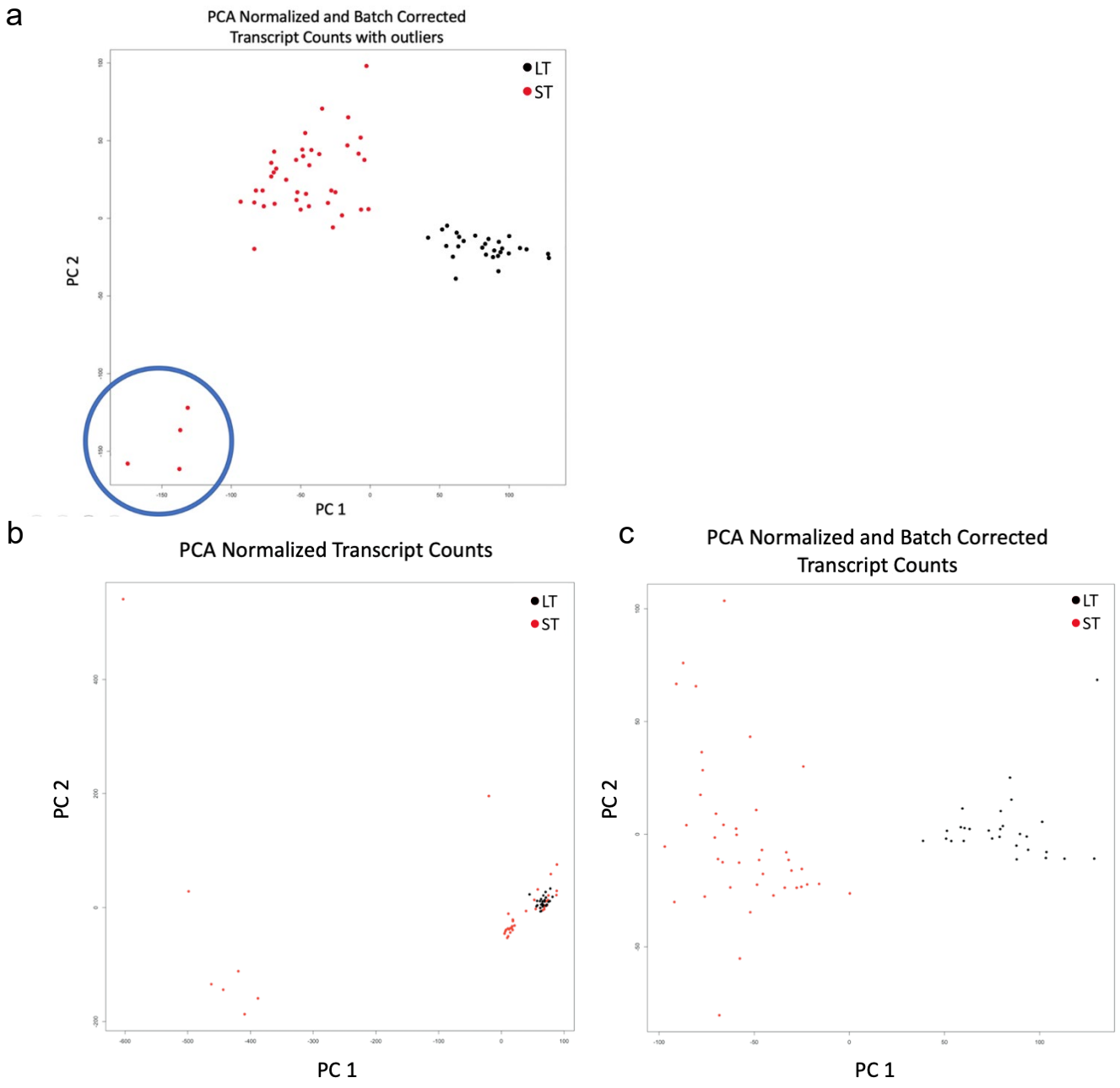
b-c. Box and whisker plots showing the percentage of total mutations for each signature, grouped by the patient's tumor as a ST or LT survivor (B) or as a primary (P) and metastatic tumor (M) (C). The boxplots define the range of the minimum to the maximum by lines, a box from the first quartile to the third quartile with the median as the center line. Circles outside the range of the data are outliers.



Supplementary Fig 2 Mutational Spectrum

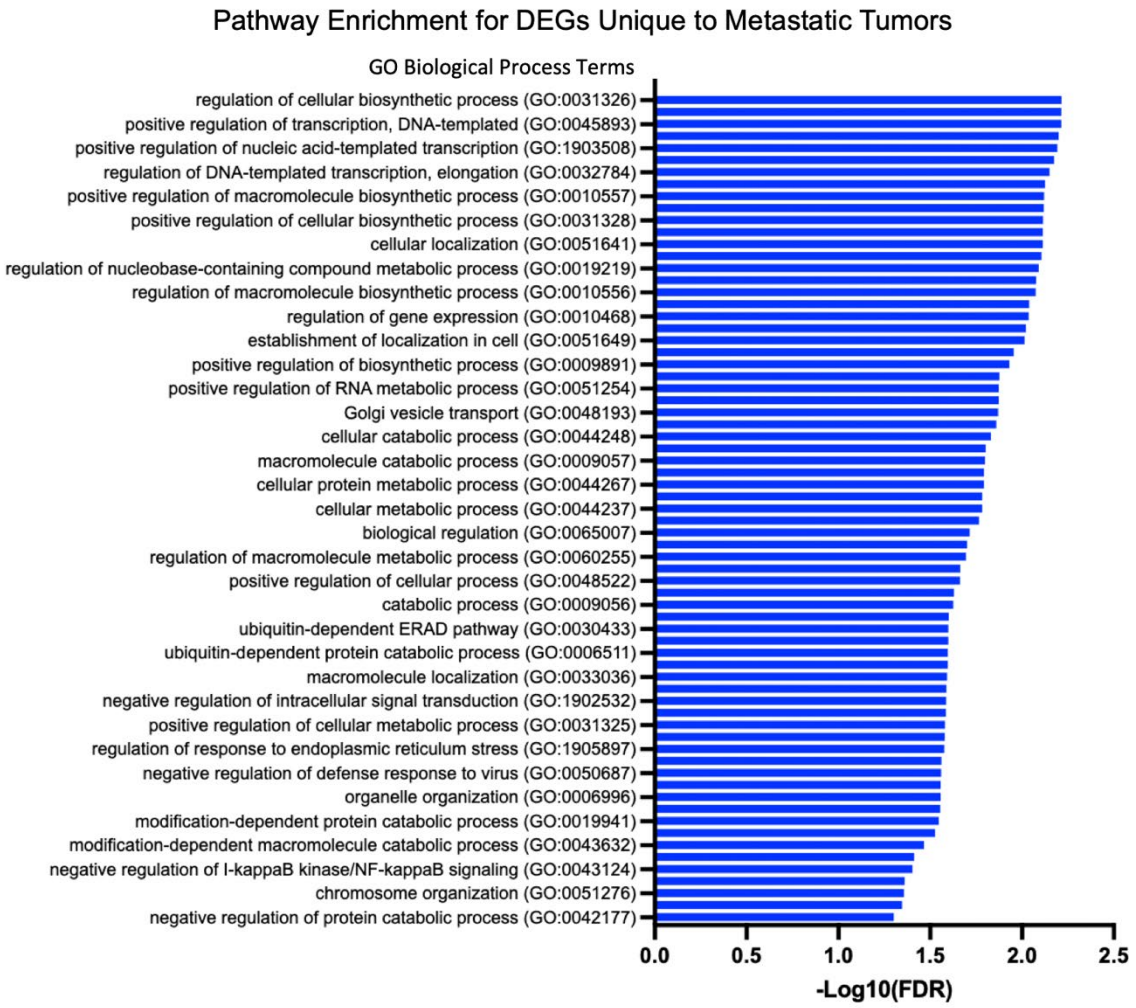
a. Stacked bar plot displaying the percent of total mutations that were nucleotide transitions or transversions, according to the key on the right. Each patient is ordered by patient number and their metastatic and primary tumor with ST survivors on the left (001-met – 030 tumor) and LT survivors on the right (031-met – 046-tumor).

b-c. Box and whisker plots showing the percentage of total mutations for each signature, grouped by the patient's tumor as a ST or LT survivor (B) or as a primary (P) and metastatic tumor (M) (C). The boxplots define the range of the minimum to the maximum by lines, a box from the first quartile to the third quartile with the median as the center line. Circles outside the range of the data are outliers.

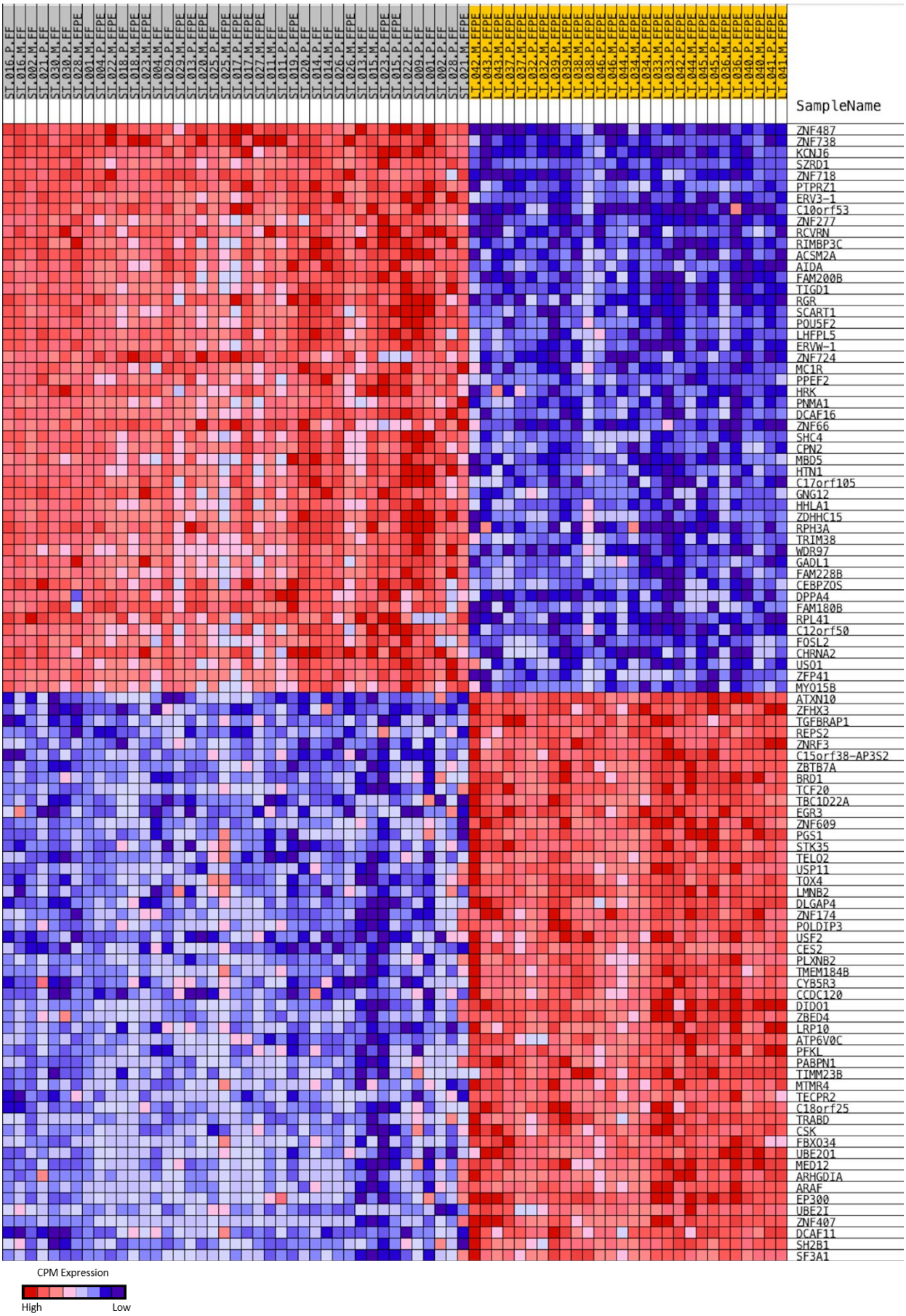


Supplementary Fig 3 PCA Plots of Log CPM counts in Differential Expression Analysis

a. PCA of cohort samples RNA-sequencing transcript counts with all tumor samples after normalization and batch effect correction with SVA. The 4 samples circled in the plot were removed as outliers.
b. PCA of cohort samples RNA-sequencing transcript counts after removal of the 4 outlier samples and normalization.
c. PCA of cohort samples RNA-sequencing transcript counts after removal of the 4 outlier samples, normalization, and batch effect correction with SVA.



Supplementary Fig 4 GO Biological Process Pathway Enrichment for Metastatic DEGs Barplot displaying the -Log10(FDR) values for each Biological Processes GO term that was statistically overrepresented among the DEGs unique to metastatic tumors

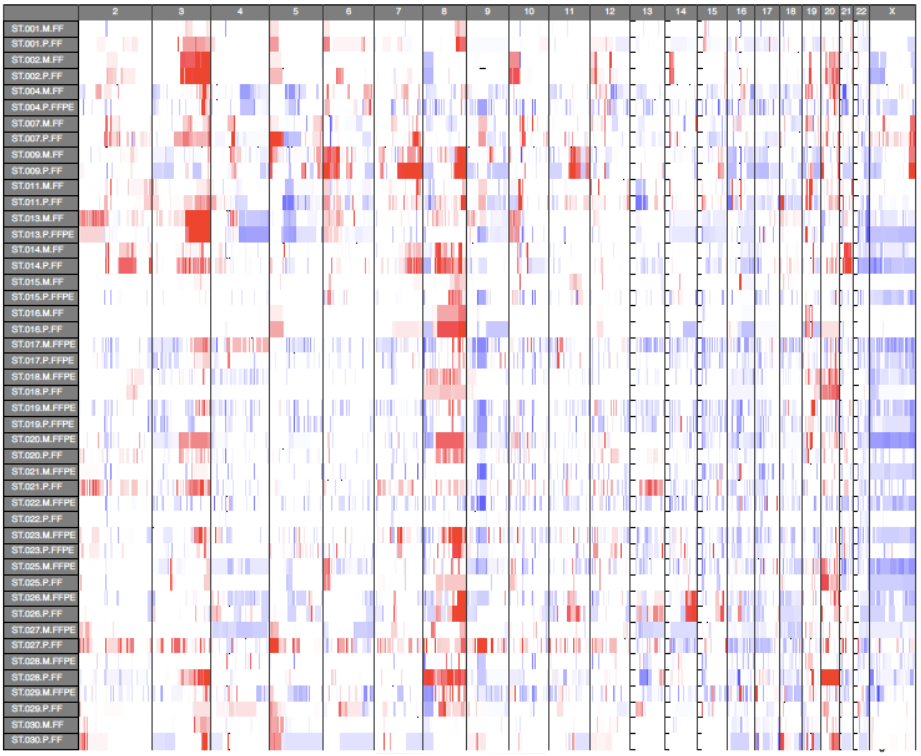

SampleName
ST.016.P.FF
ST.016.M.FF
ST.002.M.FF
ST.021.P.FF
ST.030.M.FF
ST.030.P.FF
ST.028.M.FFPE
ST.001.M.FF
ST.004.P.FFPE
ST.022.M.FFPE
ST.018.P.FF
ST.018.M.FFPE
ST.023.M.FFPE
ST.004.M.FF
ST.029.P.FF
ST.029.M.FFPE
ST.013.P.FFPE
ST.020.M.FFPE
ST.025.P.FF
ST.025.M.FFPE
ST.017.P.FFPE
ST.017.M.FFPE
ST.027.M.FFPE
ST.011.M.FF
ST.011.P.FF
ST.019.P.FFPE
ST.020.P.FF
ST.014.P.FF
ST.014.M.FF
ST.026.P.FF
ST.026.M.FFPE
ST.013.M.FF
ST.015.M.FF
ST.023.P.FFPE
ST.015.P.FFPE
ST.022.P.FF
ST.009.P.FF
ST.001.P.FF
ST.002.P.FF
ST.028.P.FF
ST.021.M.FFPE
LT.042.M.FFPE
LT.043.P.FFPE
LT.043.M.FFPE
LT.037.P.FFPE
LT.037.M.FFPE
LT.032.P.FFPE
LT.032.M.FFPE
LT.039.P.FFPE
LT.039.M.FFPE
LT.038.M.FFPE
LT.038.P.FFPE
LT.046.P.FFPE
LT.046.M.FFPE
LT.044.P.FFPE
LT.034.M.FFPE
LT.034.P.FFPE
LT.033.M.FFPE
LT.033.P.FFPE
LT.042.P.FFPE
LT.044.M.FFPE
LT.045.M.FFPE
LT.045.P.FFPE
LT.036.M.FFPE
LT.036.P.FFPE
LT.040.P.FFPE
LT.040.M.FFPE
LT.041.P.FFPE
LT.041.M.FFPE
ZNF487
ZNF738
KCNJ6
SZRD1
ZNF718
PTPRZ1
ERV3-1
C10orf53
ZNF277
RCVRN
RIMBP3C
ACSM2A
AIDA
FAM200B
TIGD1
RGR
SCART1
POU5F2
LHFPL5
ERVW-1
ZNF724
MC1R
PPEF2
HRK
PNMA1
DCAF16
ZNF66
SHC4
CPN2
MBD5
HTN1
C17orf105
GNG12
HHLA1
ZDHHC15
RPH3A
TRIM38
WDR97
GADL1
FAM228B
CEBPZOS
DPPA4
FAM180B
RPL41
C12orf50
FOSL2
CHRNA2
USO1
ZFP41
MYO15B
ATXN10
ZFHX3
TGFBRAP1
REPS2
ZNRF3
C15orf38-AP3S2
ZBTB7A
BRD1
TCF20
TBC1D22A
EGR3
ZNF609
PGS1
STK35
TELO2
USP11
TOX4
LMNB2
DLGAP4
ZNF174
POLDIP3
USF2
CES2
PLXNB2
TMEM184B
CYB5R3
CCDC120
DIDO1
ZBED4
LRP10
ATP6V0C
PFKL
PABPN1
TIMM23B
MTMR4
TECPR2
C18orf25
TRABD
CSK
FBXO34
UBE2O1
MED12
ARHGDIA
ARAF
EP300
UBE2I
ZNF407
DCAF11
SH2B1
SF3A1
CPM Expression
High
Low

Supplementary Fig 5 Top upregulated and downregulated DEGs between ST and LT survivors Heatmap displaying the top 50 upregulated and top 50 downregulated DEGs between the all ST and LT survivors generated by GSEA.

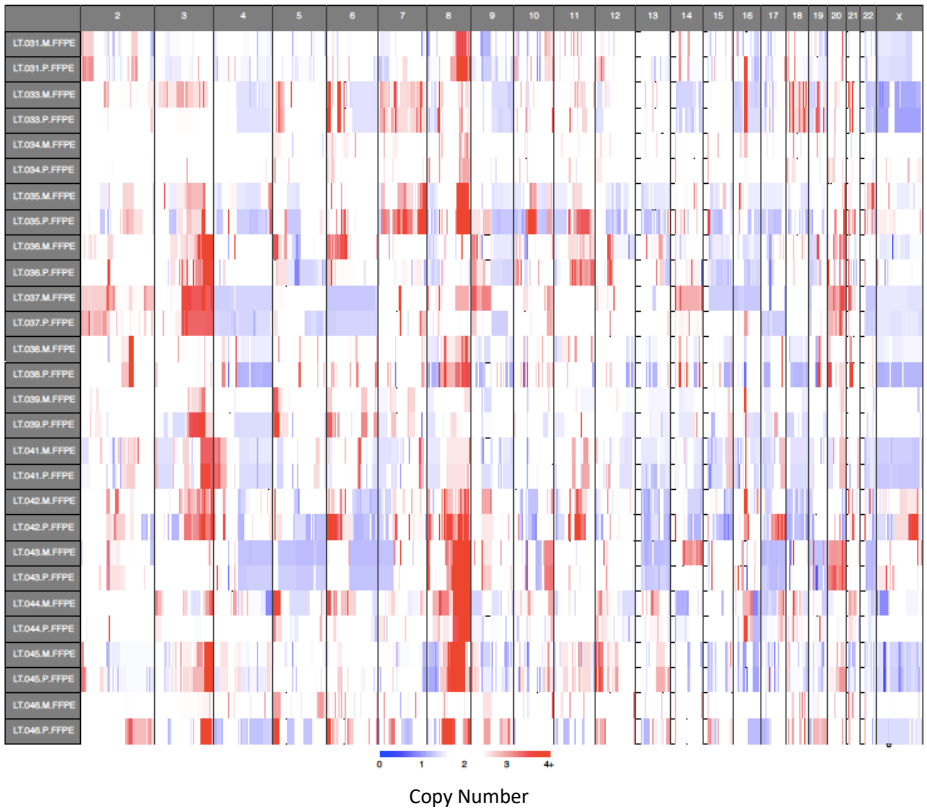
Category	Term	RT	Genes	Count	%	P-Value	Fold Enrichment	Benjamini	FDR	Fisher Exact
KEGG_PATHWAY	Adherens junction	RT		6	1.9	5.2E-3	5.3	6.2E-1	6.2E-1	8.9E-4
KEGG_PATHWAY	Protein processing in endoplasmic reticulum	RT		9	2.8	5.6E-3	3.3	6.2E-1	6.2E-1	1.6E-3
KEGG_PATHWAY	SNARE interactions in vesicular transport	RT		4	1.3	1.5E-2	7.6	1.0E0	1.0E0	1.8E-3
KEGG_PATHWAY	Alcoholic liver disease	RT		7	2.2	2.5E-2	3.1	1.0E0	1.0E0	7.4E-3
KEGG_PATHWAY	mTOR signaling pathway	RT		7	2.2	3.7E-2	2.8	1.0E0	1.0E0	1.2E-2
KEGG_PATHWAY	mRNA surveillance pathway	RT		5	1.6	6.7E-2	3.2	1.0E0	1.0E0	1.9E-2
KEGG_PATHWAY	Phospholipase D signaling pathway	RT		6	1.9	8.5E-2	2.5	1.0E0	1.0E0	3.1E-2
KEGG_PATHWAY	Insulin resistance	RT		5	1.6	9.1E-2	2.9	1.0E0	1.0E0	2.9E-2
KEGG_PATHWAY	MAPK signaling pathway	RT		9	2.8	9.4E-2	1.9	1.0E0	1.0E0	4.5E-2

Supplementary Fig 6 Kegg Pathway Enrichment
KEGG Pathways enriched in DEGs as identified by DAVID

a



b



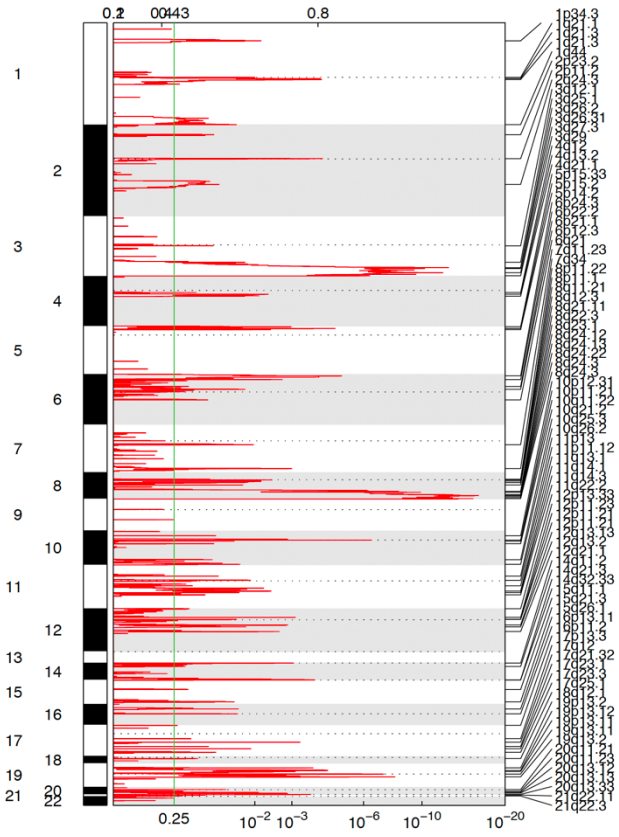
Copy Number

Supplementary Fig 7 Copy number alterations in tumors of ST and LT survivors

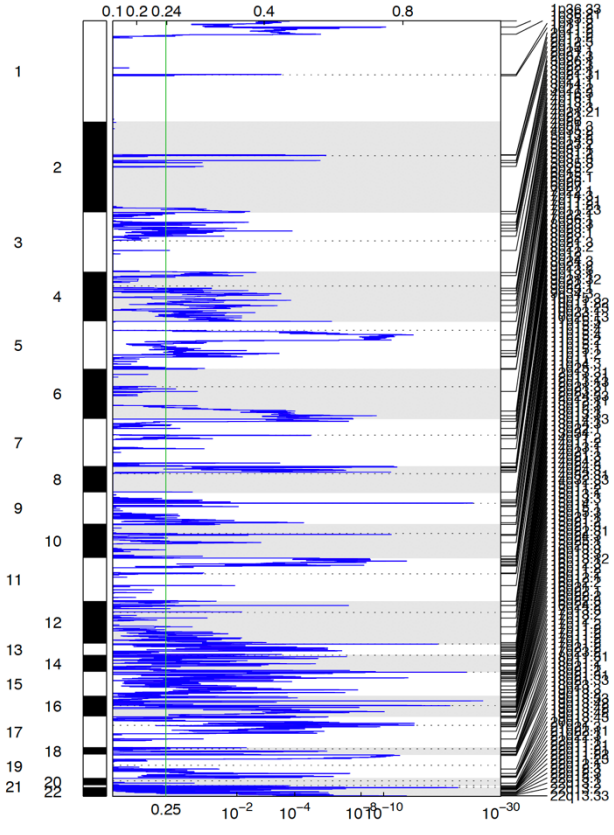
a. Copy-altered segments in primary and metastatic tumors of ST survivors across the genome. Amplifications are depicted in red while deletions are seen in blue.

b. Copy-altered segments in primary and metastatic tumors of LT survivors across the genome. Amplifications are depicted in red while deletions are seen in blue.

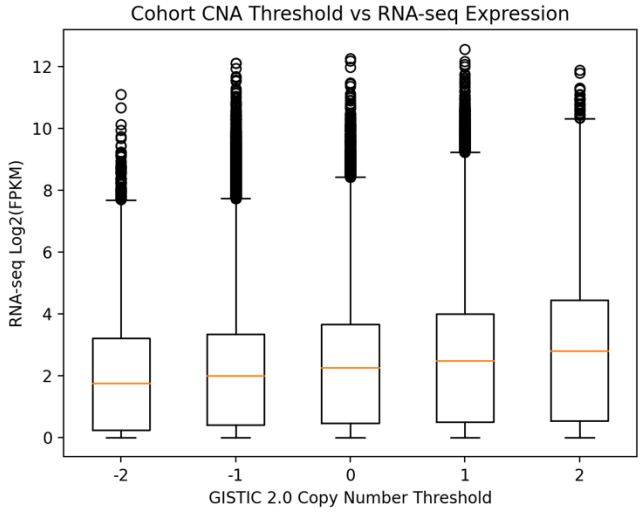
a



b



c

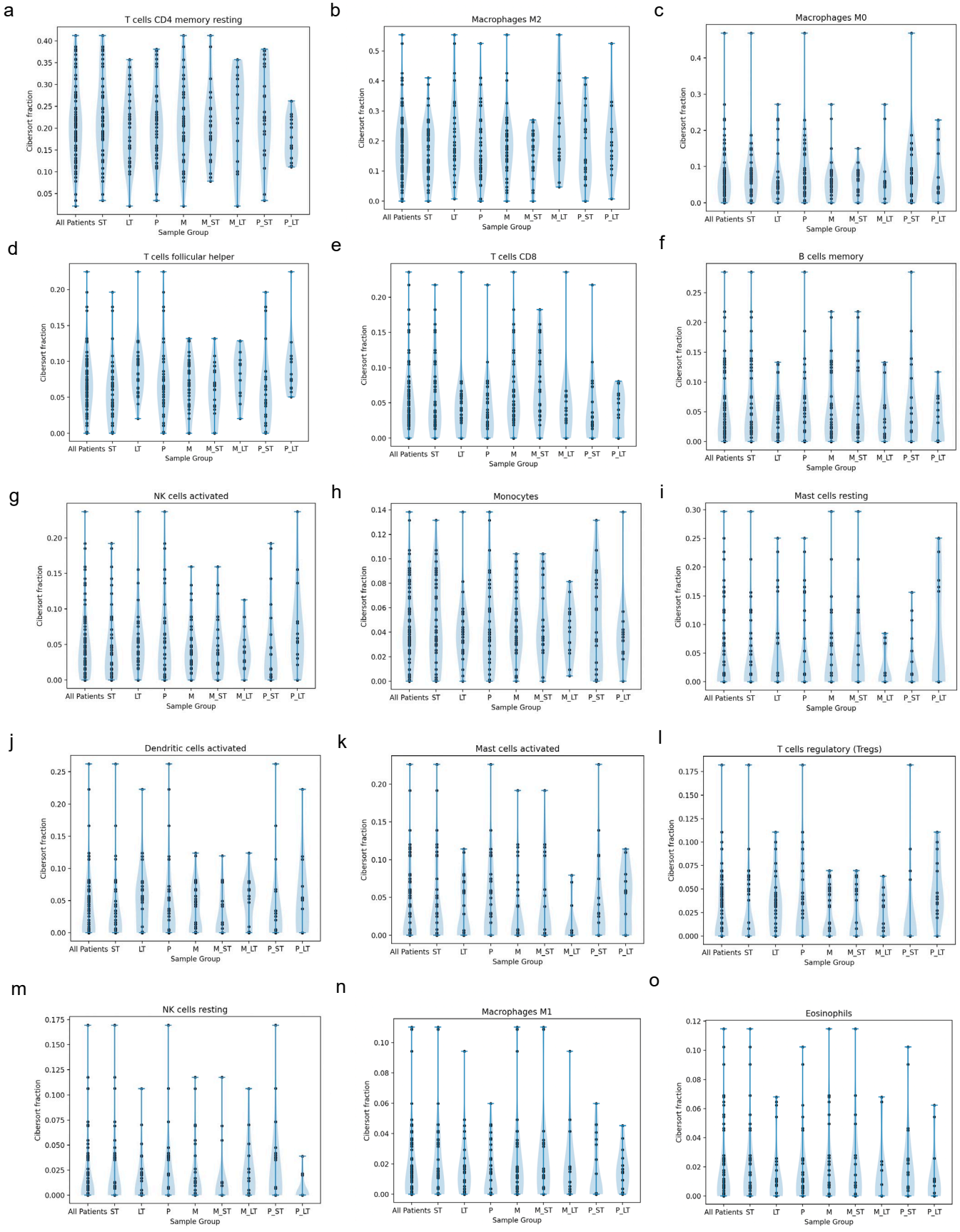

Cohort CNA Threshold vs RNA-seq Expression
RNA-seq Log2(FPKM)
GISTIC 2.0 Copy Number Threshold

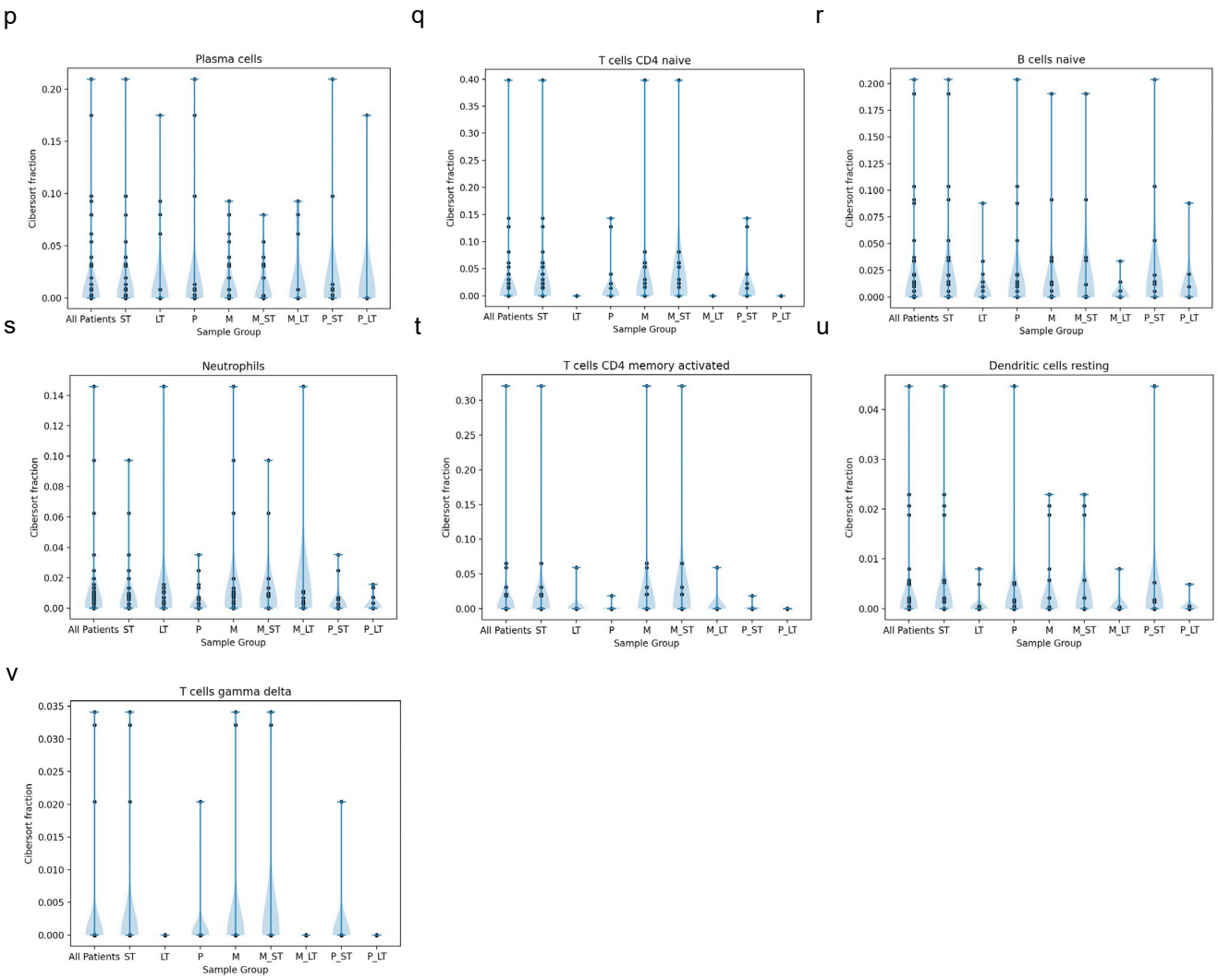
Supplementary Fig 8 GISTIC analysis for copy number alterations in tumors of ST and LT survivors

a. GISTIC Qplot of significantly amplified lesions across genome of all patient tumors (n=74) in cohort with 90% Confidence.

b. GISTIC Qplot of significantly deleted lesions across genome of all patient tumors (n=74) in cohort with 90% Confidence.

c. Boxplot of the GISTIC2.0 CNA value for every gene in every patient primary and metastatic sample plotted against their corresponding Log2(FPKM) value from RNA-seq. The medians of the FPKM values and the CNA threshold values has a spearmanr correlation value of 0.999 and p-value of 1.4x10^-24). (Number of genes in Copy Number Threshold groups from left to right: -2 n= 24316, -1 n= 416616, 0 n=280688, 1 n=346144 , 2 n=51856) The boxplot defines the range of the minimum to the maximum by lines, a box from the first quartile to the third quartile with the median as the center line. Circles outside the range of the data are outliers.


a
T cells CD4 memory resting
b
Macrophages M2
c
Macrophages M0
d
T cells follicular helper
e
T cells CD8
f
B cells memory
g
NK cells activated
h
Monocytes
i
Mast cells resting
j
Dendritic cells activated
k
Mast cells activated
l
T cells regulatory (Tregs)
m
NK cells resting
n
Macrophages M1
o
Eosinophils
Cibersort fraction
All Patients ST LT P M M_ST M_LT P_ST P_LT
Sample Group



Supplementary Fig 9 CIBERSORT immune cell fraction violin plots

a-v. Breakdown of the immune cell fractions calculated by CIBERSORT for each of the 22 immune cell signatures broken down by tumor type subgroup (all Patient tumors n= 72, ST survivors n= 44, LT survivors n= 28, Primary tumors n = 36, Metastatic tumors n= 36, ST survivor metastatic tumors n= 22, LT survivor metastatic tumors n=14, ST survivor primary tumors n= 22, and LT survivor metastatic tumors n= 14). The bars of the violin plot represent the range of the data (min and max). The shaded region on either side of the center line display the distribution of the data as a kernel density estimation. In other words, it shows the probability that a data point falls within that value range along the y-axis. Each individual data point is shown by a black dot.