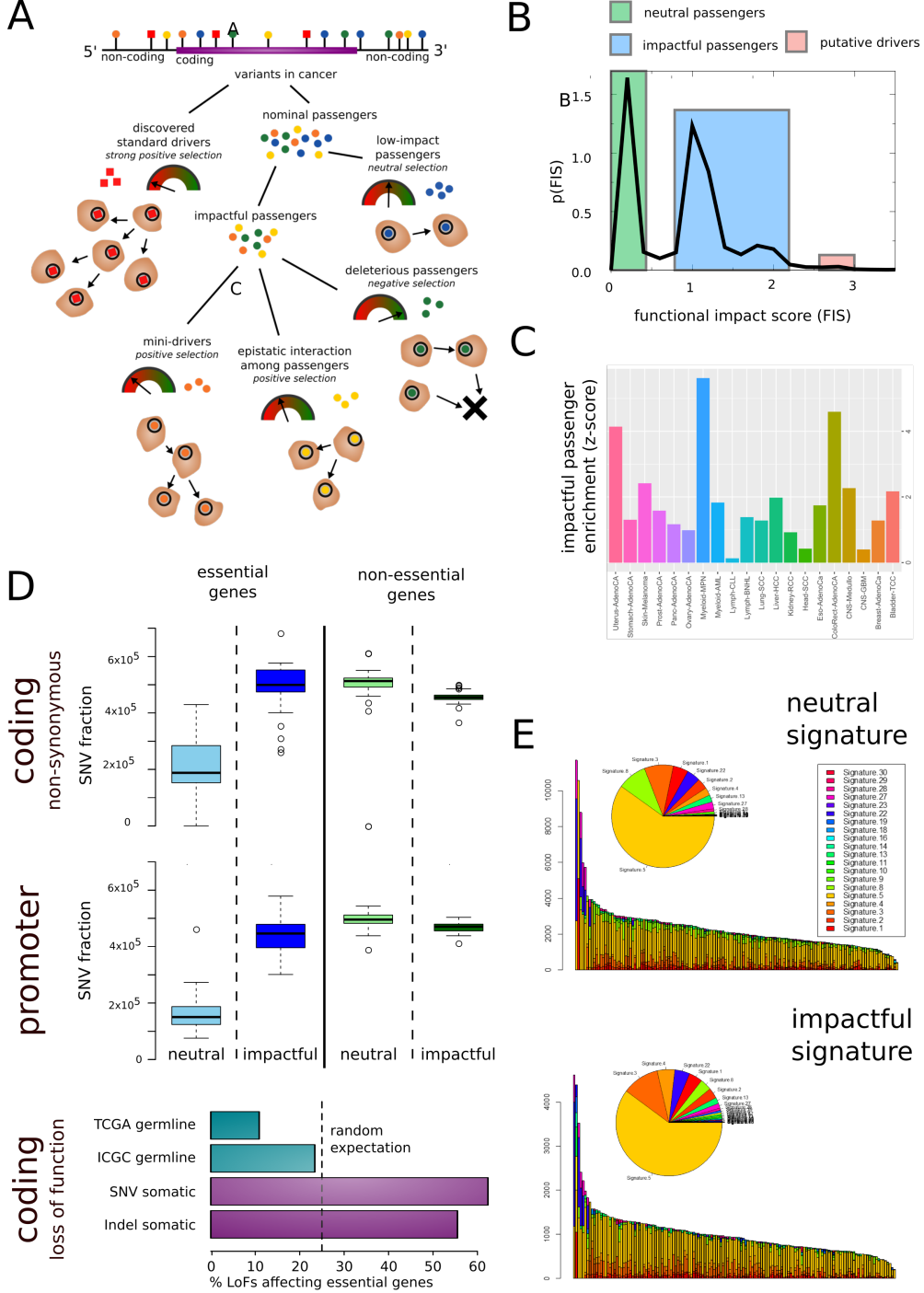
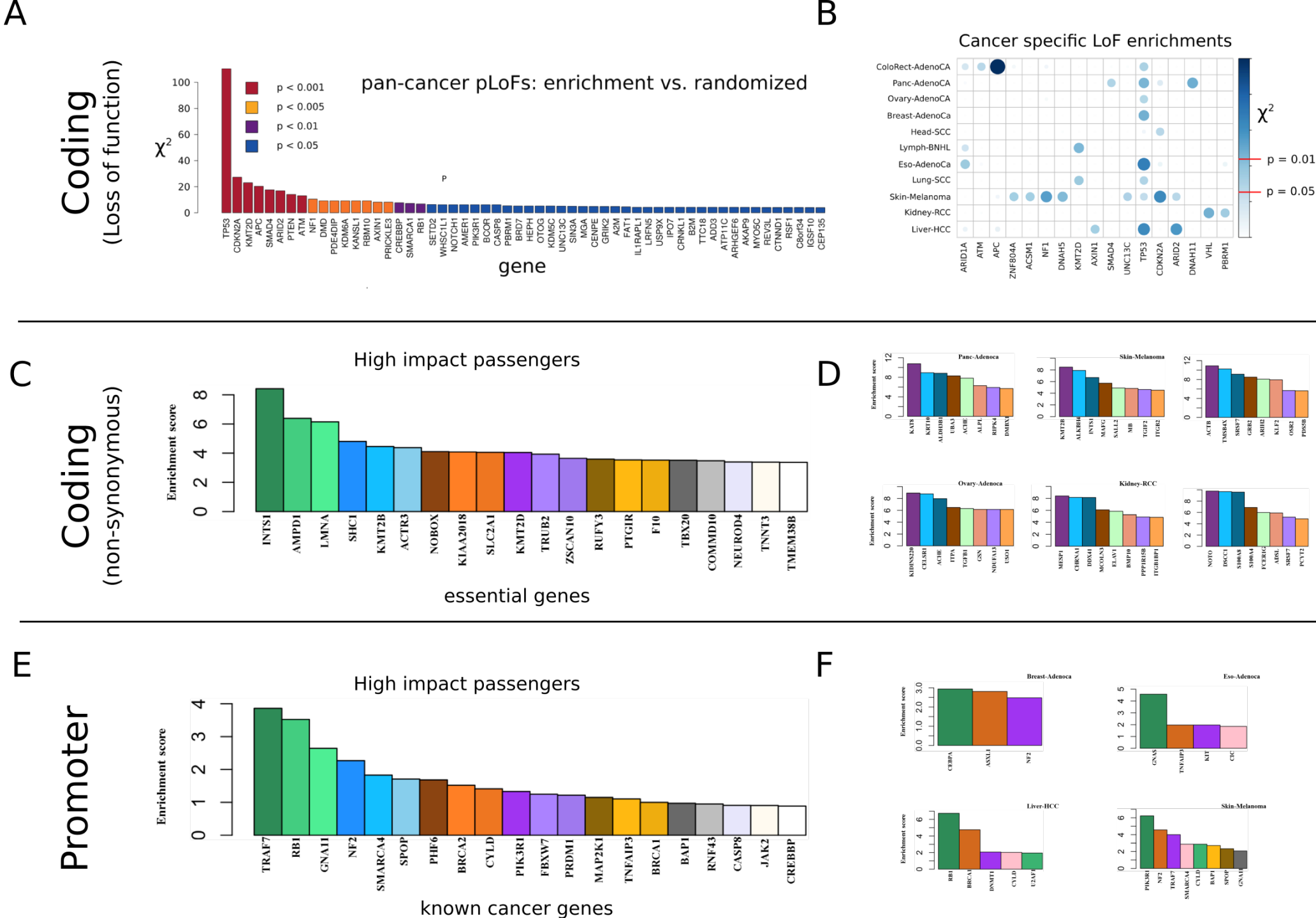
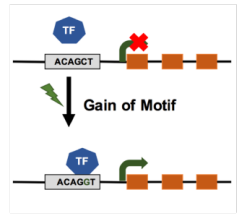
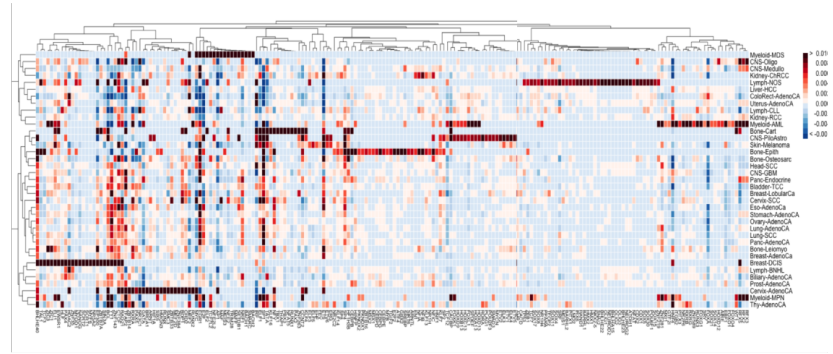

A
5'
non-coding
coding
non-coding
3'
variants in cancer
nominal passengers
discovered
standard drivers
strong positive selection
low-impact
passengers
neutral selection
impactful passengers
deleterious passengers
negative selection
C
mini-drivers
positive selection
epistatic interaction
among passengers
positive selection
B
neutral passengers
impactful passengers
putative drivers
p(FIS)
functional impact score (FIS)
C
impactful passenger
enrichment (z-score)
D
essential
genes
non-essential
genes
coding
non-synonymous
SNV fraction
promoter
SNV fraction
neutral
impactful
neutral
impactful
coding
loss of function
TCGA germline
ICGC germline
SNV somatic
Indel somatic
random
expectation
% LoFs affecting essential genes
E
neutral
signature
impactful
signature


A
Coding
(Loss of function)
pan-cancer pLoFs: enrichment vs. randomized
p < 0.001
p < 0.005
p < 0.01
p < 0.05
χ2
gene
B
Cancer specific LoF enrichments
ColoRect-AdenoCA
Panc-AdenoCA
Ovary-AdenoCA
Breast-AdenoCa
Head-SCC
Lymph-BNHL
Eso-AdenoCa
Lung-SCC
Skin-Melanoma
Kidney-RCC
Liver-HCC
p = 0.01
p = 0.05
C
Coding
(non-synonymous)
High impact passengers
Enrichment score
essential genes
D
Panc-Adenoca
Skin-Melanoma
Ovary-Adenoca
Kidney-RCC
E
Promoter
High impact passengers
Enrichment score
known cancer genes
F
Breast-Adenoca
Eso-Adenoca
Liver-HCC
Skin-Melanoma

A

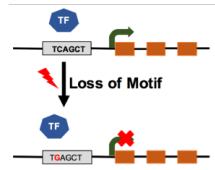
Motif Gain

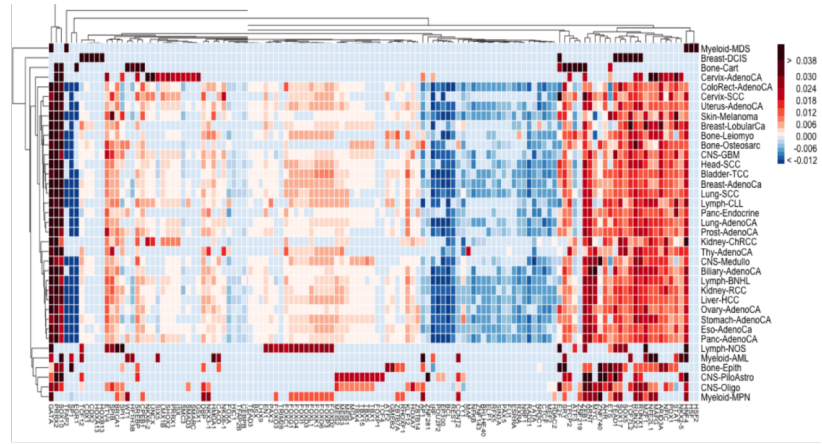




B

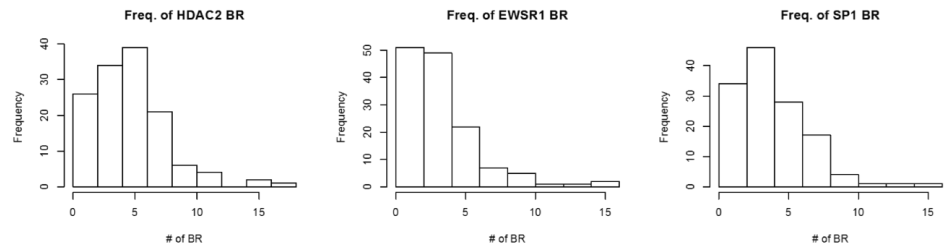
Motif Break

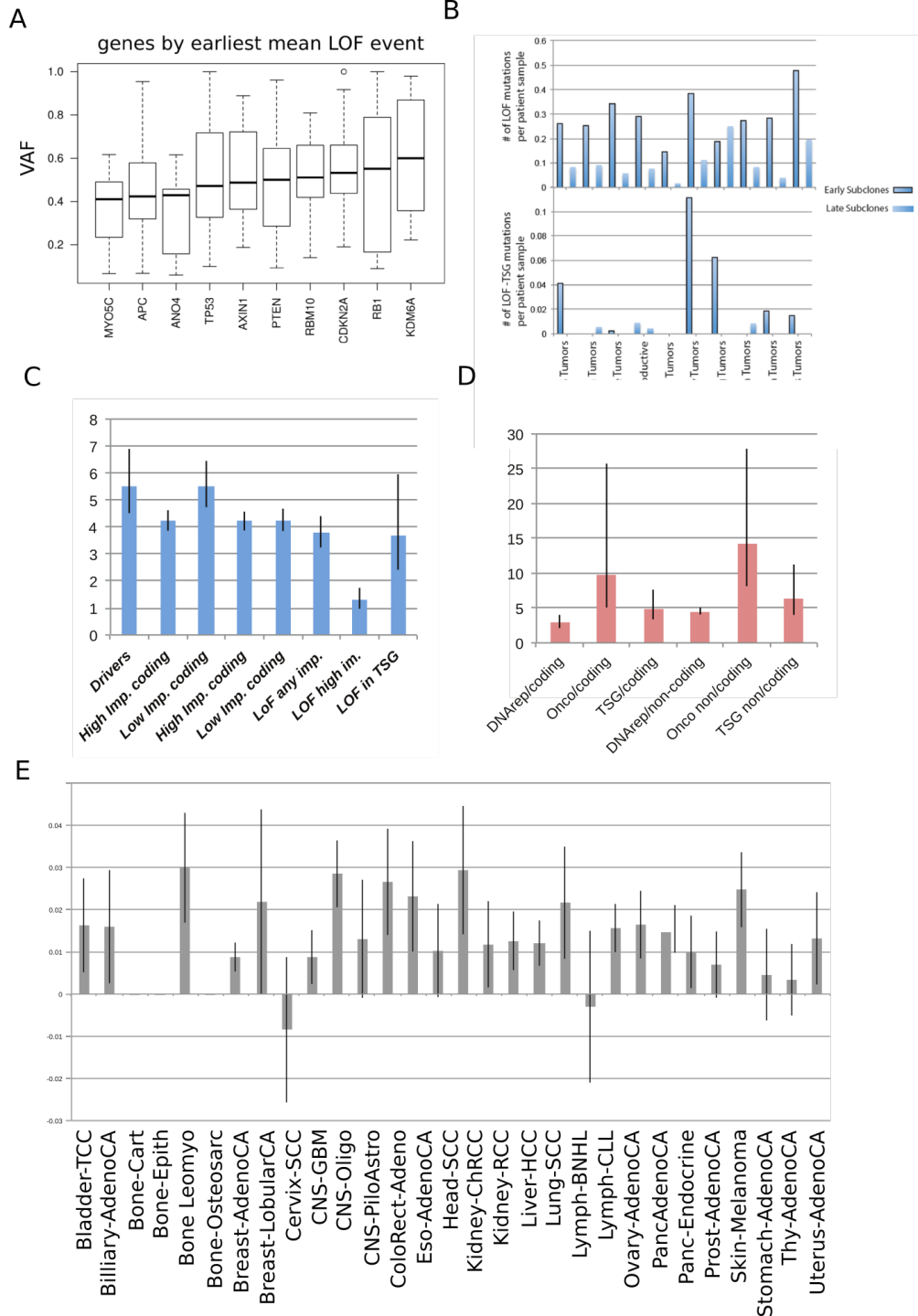




C

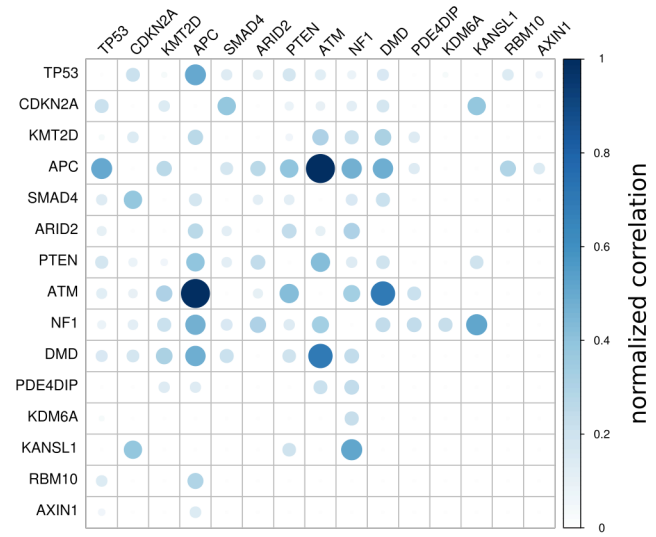
TF breaking signatures




A
genes by earliest mean LOF event
VAF
1.0
0.8
0.6
0.4
0.2
MYO5C
APC
ANO4
TP53
AXIN1
PTEN
RBM10
CDKN2A
RB1
KDM6A
B
# of LOF mutations per patient sample
# of LOF -TSG mutations per patient sample
Early Subclones
Late Subclones
C
Drivers
High imp. coding
Low imp. coding
High imp. coding
Low imp. coding
LoF any imp.
LOF high im.
LOF in TSG
D
DNArep/coding
Onco/coding
TSG/coding
DNArep/non-coding
Onco non/coding
TSG non/coding
E
Bladder-TCC
Billiary-AdenoCA
Bone-Cart
Bone-Epith
Bone Leomyo
Bone-Osteosarc
Breast-AdenoCA
Breast-LobularCA
Cervix-SCC
CNS-GBM
CNS-Oligo
CNS-PiloAstro
ColoRect-Adeno
Eso-AdenoCA
Head-SCC
Kidney-ChRCC
Kidney-RCC
Liver-HCC
Lung-SCC
Lymph-BNHL
Lymph-CLL
Ovary-AdenoCA
PancAdenoCA
Panc-Endocrine
Prost-AdenoCA
Skin-Melanoma
Stomach-AdenoCA
Thy-AdenoCA
Uterus-AdenoCA

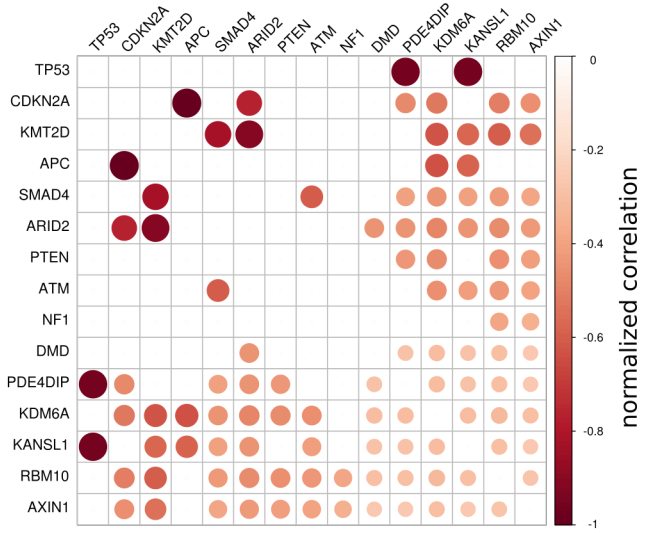
A

correlated LoFs



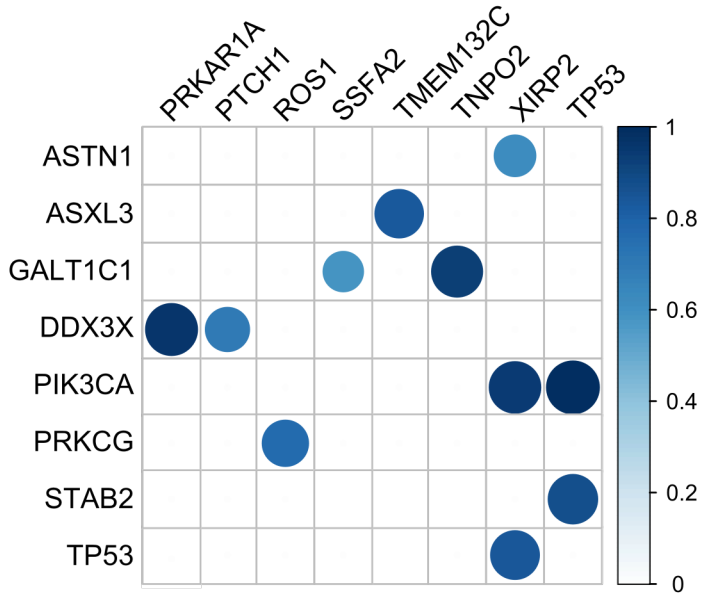
B

anticorrelated LoFs

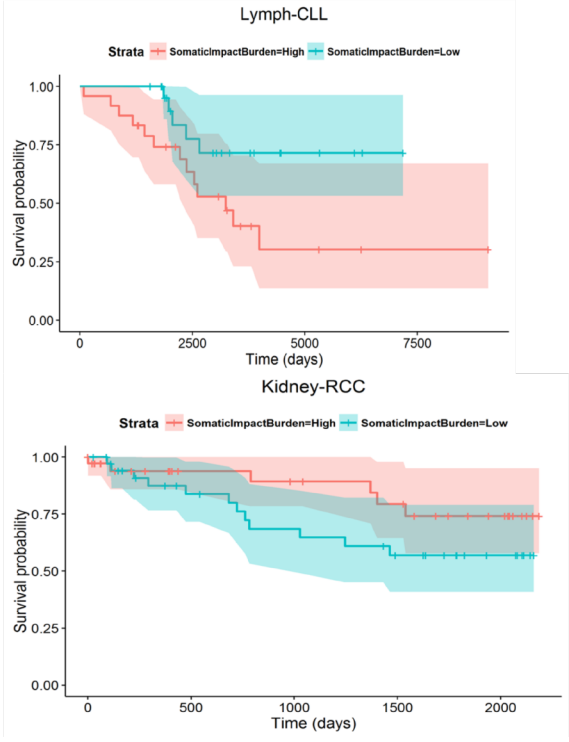


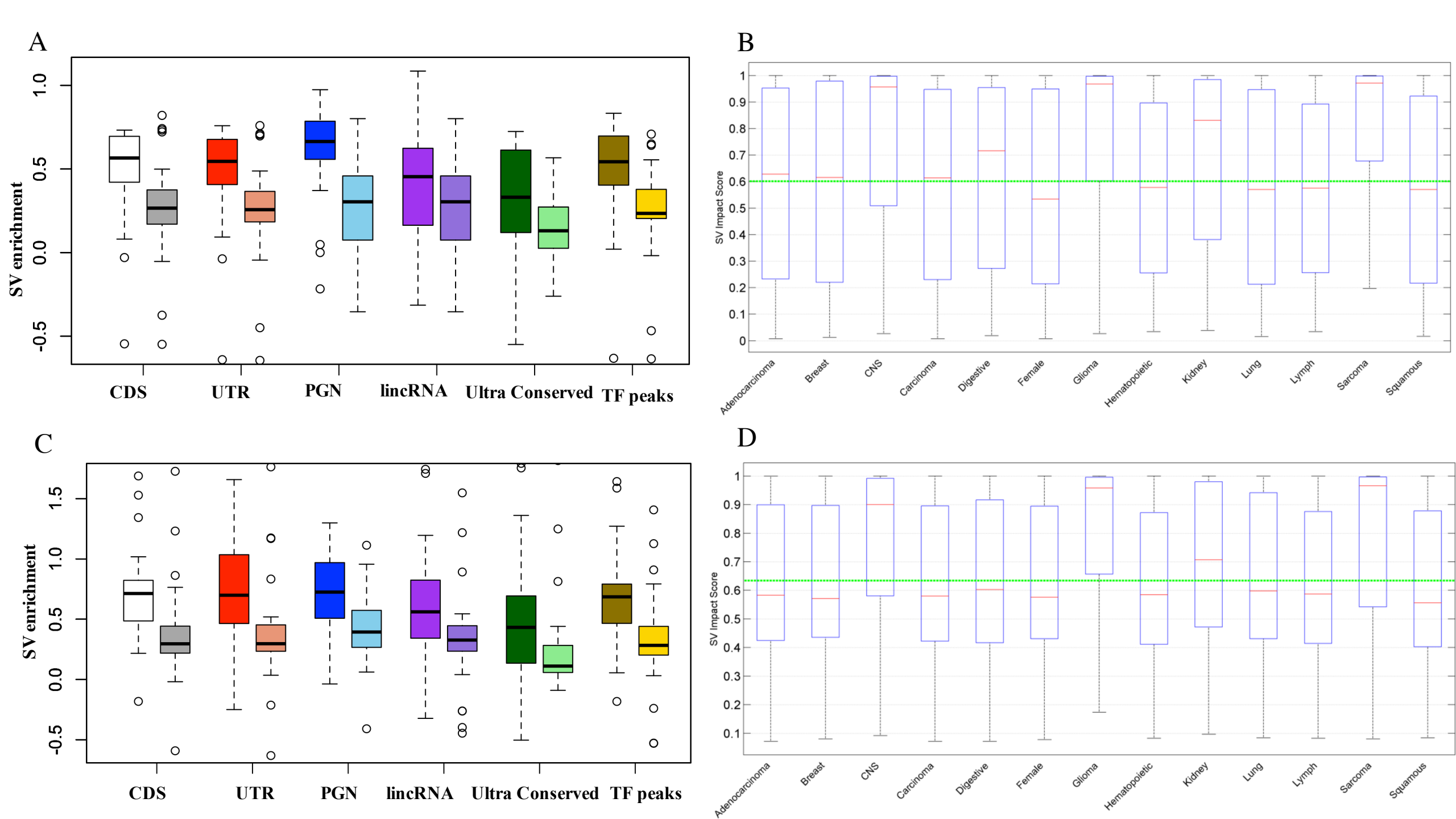
C

gene co-mutation



D




A
SV enrichment
1.0
0.5
0.0
-0.5
CDS
UTR
PGN
lincRNA
Ultra Conserved
TF peaks
B
SV Impact Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Adenocarcinoma
Breast
CNS
Carcinoma
Digestive
Female
Glioma
Hematopoietic
Kidney
Lung
Lymph
Sarcoma
Squamous
C
SV enrichment
1.5
1.0
0.5
0.0
-0.5
CDS
UTR
PGN
lincRNA
Ultra Conserved
TF peaks
D
SV Impact Score
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Adenocarcinoma
Breast
CNS
Carcinoma
Digestive
Female
Glioma
Hematopoietic
Kidney
Lung
Lymph
Sarcoma
Squamous

Relevant tables

Somatic Functional Burden & Survival [1]

Subtype	HR	CI95	p.value
Breast-AdenoCa	0.94	[0.66-1.34]	0.746
CNS-GBM	1.03	[0.75-1.4]	0.871
Eso-AdenoCa	1.01	[0.78-1.31]	0.939
Kidney-RCC	0.73	[0.54-0.98]	0.038
Liver-HCC	1.02	[0.83-1.26]	0.842
Lymph-CLL	1.48	[1.09-2.02]	0.012
Ovary-AdenoCA	1.04	[0.87-1.24]	0.698
Panc-AdenoCA	1	[0.88-1.15]	0.956
Skin-Melanoma	1.14	[0.87-1.5]	0.338

Table <Surv1> Primary findings: Relation of Somatic Impact Burden to survival.
The hazard ratio experienced for every quartile increase in Somatic Impact Burden (of observed minus randomized impact scores) over baseline, with associated 95% confidence intervals and p-values, including Low-Impact Load and patient age at diagnosis as covariates. Elevated Somatic Impact Burden is associated with shorter survival in Kidney-RCC and longer survival in Lymph-CLL.

Somatic Functional Burden & Survival [2]

Subtype	HR	CI95	p.value
Breast-AdenoCa	0.87	[0.57-1.33]	0.51897
CNS-GBM	0.64	[0.42-0.97]	0.03602
Eso-AdenoCa	0.91	[0.7-1.2]	0.50601
Kidney-RCC	1.06	[0.76-1.47]	0.74877
Liver-HCC	1.09	[0.87-1.35]	0.46297
Lymph-CLL	0.65	[0.47-0.88]	0.00567
Ovary-AdenoCA	0.73	[0.61-0.87]	0.00054
Panc-AdenoCA	0.96	[0.84-1.1]	0.58399
Skin-Melanoma	0.9	[0.69-1.16]	0.40197

Table <Surv2> Secondary findings: Relation of Low-Impact Load to survival.
The hazard ratio experienced for every quartile increase in Low-impact Load from baseline, with associated 95% confidence intervals and p-values including Somatic Impact Burden and patient age at diagnosis as covariates. Elevated load of low-impact SNVs is associated with longer survival in CNS-GBM, Lymph-CLL, and Ovary-AdenoCA.

SUBTYPE	P-value for Enrichment of High-Impact Variants among High-VAF Genes	P-value for Enrichment of High-Impact Variants among Low-VAF Genes
Pan-Cancer	0.165	0.01
Prost-AdenoCA	1.73E-08	0.023
Stomach-AdenoCA	0.016	7.60E-05
Ovary-AdenoCA	3.48E-04	0.001
CNS-Medullo	0.005	0.094
Skin-Melanoma	0.006	0.05
Kidney-RCC	0.145	0.014
Liver-HCC	0.022	0.016
Lymph-CLL	0.016	0.286
CNS-PiloAstro	0.02	0.985
Bone-Leiomyo	0.29	0.028
ColoRect-AdenoCA	0.462	0.032
Panc-AdenoCA	0.076	0.05
Eso-AdenoCa	0.052	0.707
Kidney-ChRCC	0.811	0.057
Thy-AdenoCA	0.096	0.732
Uterus-AdenoCA	0.125	0.353
Lymph-BNHL	0.146	0.128
Biliary-AdenoCA	0.982	0.136
Head-SCC	0.291	0.138
CNS-GBM	0.196	0.303
Bladder-TCC	0.21	0.286
Lung-AdenoCA	0.219	0.243
Breast-AdenoCa	0.273	0.278
Breast-LobularCa	0.307	0.851
Panc-Endocrine	0.482	0.321
Bone-Epith	0.383	0.459
Bone-Osteosarc	0.41	0.608
Cervix-SCC	0.515	0.75
CNS-Oligo	0.549	0.971
Lung-SCC	0.556	0.564
Myeloid-MPN	0.715	0.715
Myeloid-AML	0.768	0.733

Bullets

- Assessed whether genes with high VAF or low VAF are statistically enriched in high-impact variants compared with random, which would serve as evidence of latent drivers and deleterious passengers, respectively
- Pan-cancer analysis found statistical evidence of deleterious passengers but not latent drivers
- Subtype-specific analysis found some subtypes with evidence of deleterious passengers, some subtypes with evidence of latent drivers, and some subtypes with evidence for both kinds of mutation classes
- Intriguingly, Lymph-CLL was one of few subtypes with simultaneous evidence of latent drivers and lack of evidence of deleterious passengers, and Renal-RCC was one of few subtypes with the opposite pattern, which could explain the survival trends we observe in these cancer subtypes.

Germline functional burden predisposes patients to earlier age of cancer diagnosis

Subtype	Coef	P.value
Biliary-AdenoCA	-3.397	0.068
Breast-AdenoCa	-2.911	0.020
CNS-GBM	1.320	0.509
CNS-Medullo	-1.352	0.024
CNS-PiloAstro	-0.044	0.953
ColoRect-AdenoCA	-0.638	0.805
Eso-AdenoCa	1.366	0.112
Head-SCC	-3.899	0.125
Kidney-ChRCC	-1.692	0.354
Kidney-RCC	-0.058	0.947
Liver-HCC	-0.146	0.822
Lung-AdenoCA	1.892	0.263
Lung-SCC	0.009	0.994
Lymph-BNHL	0.458	0.807
Lymph-CLL	1.165	0.244
Ovary-AdenoCA	1.609	0.120
Panc-AdenoCA	-0.072	0.925
Panc-Endocrine	-4.260	0.002
Prost-AdenoCA	-0.089	0.877
Skin-Melanoma	-0.195	0.883
Stomach-AdenoCA	0.123	0.958
Thy-AdenoCA	-5.580	0.072
Uterus-AdenoCA	5.064	0.106

Linear regression predicting age in years of cancer diagnosis as a function of combined coding and noncoding germline burden. Patient race, and low-impact somatic load used as additional covariates.