Supplemental Material

Alternative splicing changes as drivers of cancer

Héctor Climente-González[1], Eduard Porta-Pardo[2], Adam Godzik[2], Eduardo Eyras[1,3]

[1]Universitat Pompeu Fabra (UPF), Barcelona, Spain

[2]Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, 92037, USA

[3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Supplemental Data

Related to Figure 1
Figure S1
Table S1

Related to Figure 2
Figure S2
Table S2
Table S3

Related to Figure 3
Figure S3
Figure S4
Figure S5
Table S4
Table S5
Table S6

Related to Figure 4
Figure S6

Supplemental Data

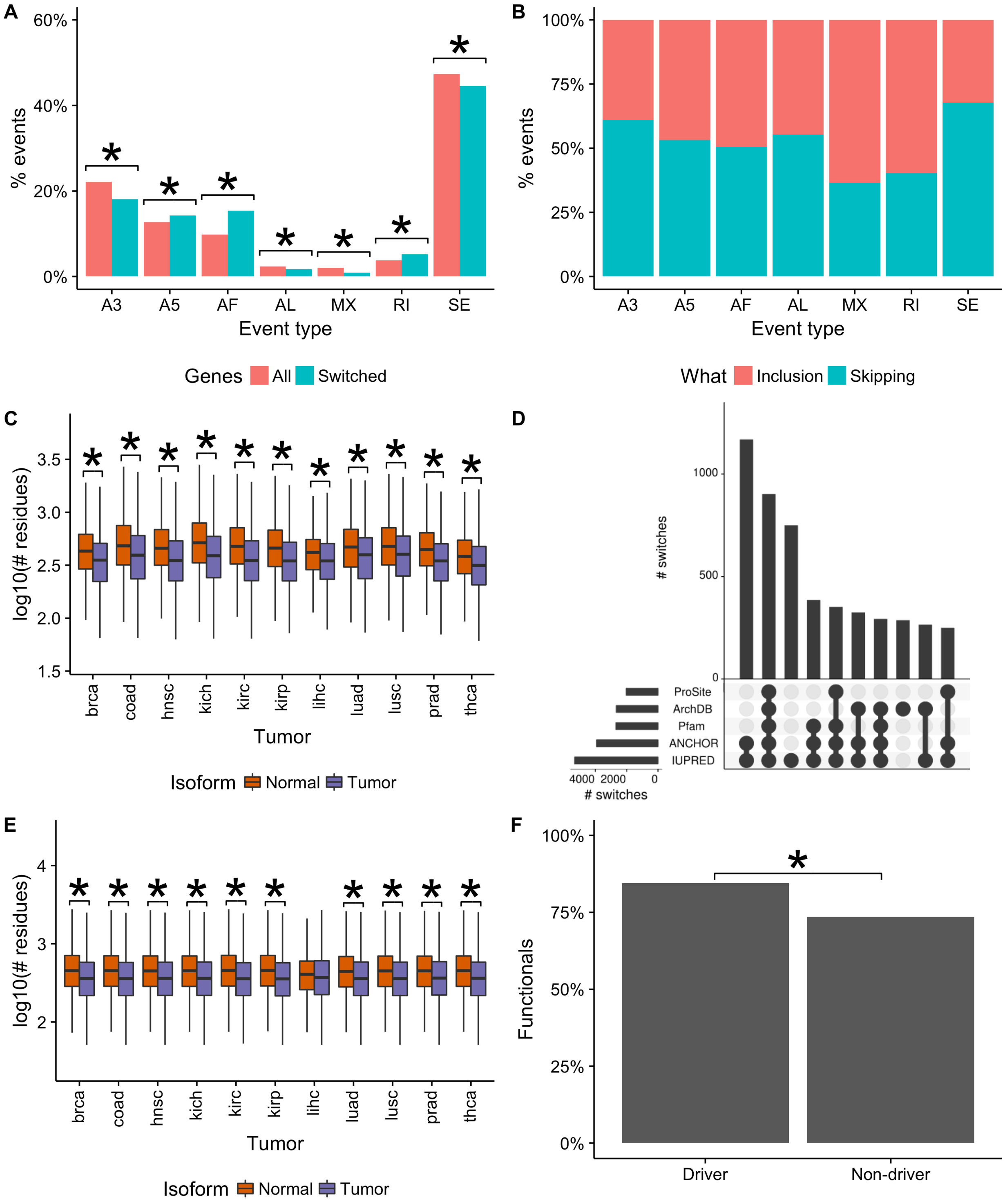

Figure S1. Properties of isoform switches Related to Figure 1. **(A)** Proportion of local alternative splicing event types (y-axis) described by the switches (blue) and by all genes in the annotation (red). These proportions are shown for events of type alternative 3’ (A3) and 5’ (A5) splice-site, alternative first (AF) and last (AL) exons, mutually exclusive exons (MX), intron

retention events (RI) and exon cassette events (SE). Significance of the difference was determined with a Fisher's exact test for each event type using a contingency table with the counts of each event type and the rest of events in the two sets: switches and annotation **(B)** For each set of local alternative splicing events from the same type mapped to isoform switches, we indicate the proportion of cases that correspond to either inclusion (red) or exclusion (blue). For instance, inclusion for the A3 and A5 events correspond to the longer form, for AF events to the most upstream exon, to the most downstream exon for AL events, to the inclusion of the exon with the lowest coordinates for MX events, to the retention of the intron for RI events, and to the inclusion of the cassette exon for SE events. Blue corresponds to the opposite configuration.Further details of the description of the events can be found in https://github.com/comprna/SUPPA (Alamancos et al., 2015). **(C)** Distributions of the lengths of the tumor (purple) and normal (red) protein isoforms in the calculated isoform switches. The *y*-axis indicates the number of residues in log10 scale. **(D)** Overlap graph (Conway et al., 2017) of protein features affected in functional switches: Prosite patterns (Prosite), protein loops (ArchDB), Pfam domains (Pfam), disordered regions with potential to mediate protein–protein interactions (ANCHOR), and general disordered regions (IUPRED). The horizontal bars indicate the number of switches affecting each feature. The vertical bars indicate the number of switches in each intersection indicated by connected bullet points. **(E)** Distributions of the lengths of the tumor (purple) and normal (red) protein isoforms in the simulated transcript isoform switches. **(F)** Enrichment of functional switches in cancer drivers. We separated all switches (from Table S1) according to whether they are cancer drivers or non-drivers (in any tumor type), and whether they have functional switches or not. From the 6004 functional switches, ~4% are drivers, whereas from the 2118 non-functional switches, ~2% are drivers. Similarly, from all considered 278 drivers, ~84% are functional, whereas ~73% of the 7844 non-driver switches are functional. A Fisher's exact test produced a p-value = 2.034e-05 and odds-ratio = 1.965563 for the enrichment of functional switches in drivers (95 percent confidence interval: 1.409, 2.799).

Table S1. Isoform switches. Related to Figure 1. Provided as a text file with tab-separated values (.tsv). This table contains the list of identified isoform switches used for this analysis, including functional and nonfunctional ones, and AS-drivers. The table provides the following information:

Column number	Column label	Description
1	GeneId	Entrez gene id
2	Symbol	HGNC gene symbol
3	Normal_transcript	UCSC transcript id
4	Tumor_transcript	UCSC transcript id
5	Normal_protein	Uniprot_ID (None if not known)
6	Tumor_protein	Uniprot_ID (None if not known)
7	DriverAnnotation	"Driver" if it's a driver, "d1" if it's an interactor of a driver, and "Nothing" otherwise
8	IsFunctional	1 if it is functional as defined in the article, 0 otherwise
9	Driver	1 if it is a driver, 0 otherwise
10	Druggable	1 if it is a target of a known drug according to DGIdb (http://dgidb.genome.wustl.edu/)
11	CDS_Normal	1 if the normal transcript has an annotated CDS, 0 otherwise
12	CDS_Tumor	1 if the tumor transcript has an annotated CDS, 0 otherwise
13	CDS_change	1 if the CDS changes between the tumor and normal transcripts
14	UTR_change	1 if the 5'3 or 3' UTRs change between the tumor and normal transcripts
15	Tumors	Tumor types in which the switch appears (brca, coad, etc…)
16	Number_samples	Number of samples in which the switch appears
17	Percentage_samples	Percentage of samples from the total studied across all tumor types in which the switch appears
18	Samples	IDs of samples in which the switch appears
19	Recurrence	1 if it is recurrent, 0 otherwise
20	PPI	1 if the switch affects a PPI in every tumor type where it appears; 0 otherwise. All PPIs affected by switches per tumor type are in Supp. File 3.
21	Affects_mutated_feature	1 if the switch leads to a gain or loss of a domain that is enriched in mutations in tumors, 0 otherwise
22	Pannegative	Number of cancer drivers from the same pathway with which the switch shows mutual exclusion
23	AS_driver	1 if 19,20,21 or 22 is equal to 1, 0 otherwise
24	MS.pam	Samples with co-occurrence of switch and PAM in the same gene
25	M.pam	Samples with PAMs only
26	S.pam	Samples with Switches
27	N.pam	Rest of samples
28	p.pam.me	p-value of the mutual exclusion test
29	MS.mut	Samples with co-occurrence of switch and WGS mutations
30	M.mut	Samples with WGS mutations only
31	S.mut	Samples with Switches
32	N.mut	Rest of samples
33	p.mut.o	p-value of the co-occurrence of mutations and switches

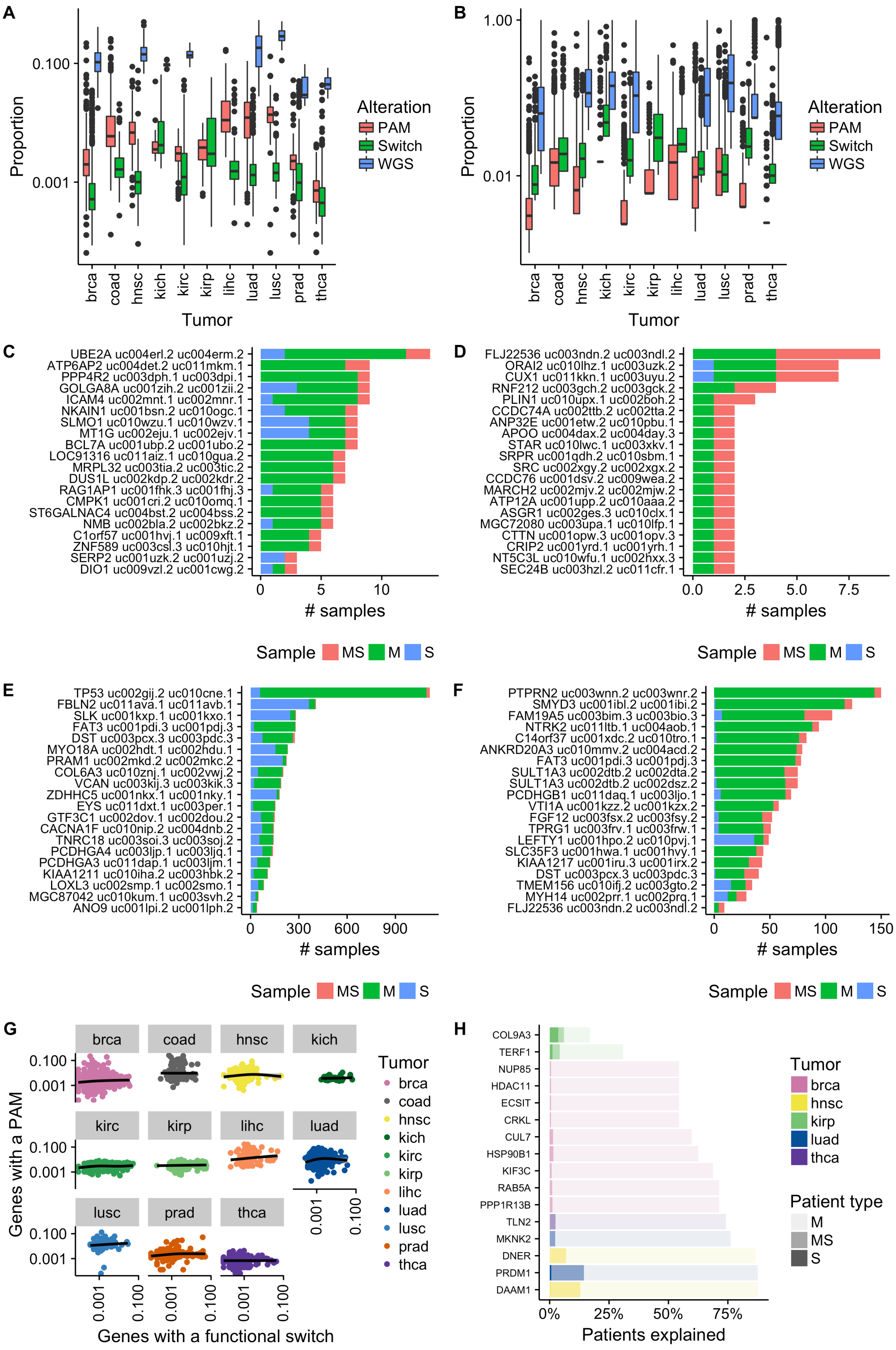
A
0.100
0.001
Proportion
brca
coad
hnsc
kich
kirc
kirp
lihc
luad
lusc
prad
thca
Tumor
Alteration
PAM
Switch
WGS
B
1.00
0.01
Proportion
brca
coad
hnsc
kich
kirc
kirp
lihc
luad
lusc
prad
thca
Tumor
Alteration
PAM
Switch
WGS
C
UBE2A uc004erl.2 uc004erm.2
ATP6AP2 uc004det.2 uc011mkm.1
PPP4R2 uc003dph.1 uc003dpi.1
GOLGA8A uc001zih.2 uc001zii.2
ICAM4 uc002mnt.1 uc002mnr.1
NKAIN1 uc001bsn.2 uc010ogc.1
SLMO1 uc010wzu.1 uc010wzv.1
MT1G uc002eju.1 uc002ejv.1
BCL7A uc001ubp.2 uc001ubo.2
LOC91316 uc011aiz.1 uc010gua.2
MRPL32 uc003tia.2 uc003tic.2
DUS1L uc002kdp.2 uc002kdr.2
RAG1AP1 uc001fhk.3 uc001fhj.3
CMPK1 uc001cri.2 uc010omq.1
ST6GALNAC4 uc004bst.2 uc004bss.2
NMB uc002bla.2 uc002bkz.2
C1orf57 uc001hvj.1 uc009xft.1
ZNF589 uc003csl.3 uc010hjt.1
SERP2 uc001uzk.2 uc001uzj.2
DIO1 uc009vzl.2 uc001cwg.2
0
5
10
# samples
Sample MS M S
D
FLJ22536 uc003ndn.2 uc003ndl.2
ORAI2 uc010lhz.1 uc003uzk.2
CUX1 uc011kkn.1 uc003uyu.2
RNF212 uc003gch.2 uc003gck.2
PLIN1 uc010upx.1 uc002boh.2
CCDC74A uc002ttb.2 uc002tta.2
ANP32E uc001etw.2 uc010pbu.1
APOO uc004dax.2 uc004day.3
STAR uc010lwc.1 uc003xkv.1
SRPR uc001qdh.2 uc010sbm.1
SRC uc002xgy.2 uc002xgx.2
CCDC76 uc001dsv.2 uc009wea.2
MARCH2 uc002mjv.2 uc002mjw.2
ATP12A uc001upp.2 uc010aaa.2
ASGR1 uc002ges.3 uc010clx.1
MGC72080 uc003upa.1 uc010lfp.1
CTTN uc001opw.3 uc001opv.3
CRIP2 uc001yrd.1 uc001yrh.1
NT5C3L uc010wfu.1 uc002hxx.3
SEC24B uc003hzl.2 uc011cfr.1
0.0
2.5
5.0
7.5
# samples
Sample MS M S
E
TP53 uc002gij.2 uc010cne.1
FBLN2 uc011ava.1 uc011avb.1
SLK uc001kxp.1 uc001kxo.1
FAT3 uc001pdi.3 uc001pdj.3
DST uc003pcx.3 uc003pdc.3
MYO18A uc002hdt.1 uc002hdu.1
PRAM1 uc002mkd.2 uc002mkc.2
COL6A3 uc010znj.1 uc002vwj.2
VCAN uc003kjj.3 uc003kik.3
ZDHHC5 uc001nkx.1 uc001nky.1
EYS uc011dxt.1 uc003per.1
GTF3C1 uc002dov.1 uc002dou.2
CACNA1F uc010nip.2 uc004dnb.2
TNRC18 uc003soi.3 uc003soj.2
PCDHGA4 uc003ljp.1 uc003ljq.1
PCDHGA3 uc011dap.1 uc003ljm.1
KIAA1211 uc010iha.2 uc003hbk.2
LOXL3 uc002smp.1 uc002smo.1
MGC87042 uc010kum.1 uc003svh.2
ANO9 uc001lpi.2 uc001lph.2
0
300
600
900
# samples
Sample MS M S
F
PTPRN2 uc003wnn.2 uc003wnr.2
SMYD3 uc001ibl.2 uc001ibi.2
FAM19A5 uc003bim.3 uc003bio.3
NTRK2 uc011ltb.1 uc004aob.1
C14orf37 uc001xdc.2 uc010tro.1
ANKRD20A3 uc010mmv.2 uc004acd.2
FAT3 uc001pdi.3 uc001pdj.3
SULT1A3 uc002dtb.2 uc002dta.2
SULT1A3 uc002dtb.2 uc002dsz.2
PCDHGB1 uc011daq.1 uc003ljo.1
VTI1A uc001kzz.2 uc001kzx.2
FGF12 uc003fsx.2 uc003fsy.2
TPRG1 uc003frv.1 uc003frw.1
LEFTY1 uc001hpo.2 uc010pvj.1
SLC35F3 uc001hwa.1 uc001hvy.1
KIAA1217 uc001iru.3 uc001irx.2
DST uc003pcx.3 uc003pdc.3
TMEM156 uc010ifj.2 uc003gto.2
MYH14 uc002prr.1 uc002prq.1
FLJ22536 uc003ndn.2 uc003ndl.2
0
50
100
150
# samples
Sample MS M S
G
brca
coad
hnsc
kich
kirc
kirp
lihc
luad
lusc
prad
thca
0.100
0.001
0.100
0.001
0.100
0.001
0.001
0.100
0.001
0.100
0.001
0.100
0.001
0.100
Genes with a PAM
Genes with a functional switch
Tumor
brca
coad
hnsc
kich
kirc
kirp
lihc
luad
lusc
prad
thca
H
COL9A3
TERF1
NUP85
HDAC11
ECSIT
CRKL
CUL7
HSP90B1
KIF3C
RAB5A
PPP1R13B
TLN2
MKNK2
DNER
PRDM1
DAAM1
0%
25%
50%
75%
Patients explained
Tumor
brca
hnsc
kirp
luad
thca
Patient type
M
MS
S

Figure S2. Properties of functional isoform switches in tumors. Related to Figure 2. **(A)** Proportion of genes in $\log_{10}$ scale (*y*-axis) with either of these three alterations: isoform switches (red), protein-affecting mutations (PAMs) from whole Exome sequencing (WES) data (green), and any mutation type from whole genome sequencing (WGS) data (blue). **(B)** Proportion of samples (*y*-axis) with either of these three alterations: isoform switches (red), PAMs from WES data (green), and any mutation type from WGS data (blue). **(C-F)** Potential associations between mutations and switches. We show the top 20 cases according to the Jaccard score for the association of mutations (M) and switches (S) using WES (C) and WGS (D) data. We also show the top 20 cases according to the number of MS samples for WES (E) and WGS (F) data. For each gene and isoform (y axis), we show the number of patients for which we observed a mutation only (M), a switch only (S), or the co-occurrence of both (MS). **(G)** Lack of correlation between mutations and switches. For each tumor type, each dot represents a sample according to the number of genes with a functional switch (*x*-axis) and the number of genes with protein-affecting mutations (PAMs) (*y*-axis). **(H)** Functional switches that potentially characterize pan-negative tumor samples. For each switch along the *y*-axis, we represent the proportion of patients from a given tumor type (*x*-axis) that harbor mutations in a tumor-specific mutational driver (M), have the switch (S), or have both (MS). The switches are ranked from the bottom of the *y*-axis according to the total number of patients explained. Only the top 30 cases are shown. Each case is color-coded according to tumor type.

Table S2. Mutation and domain gain/loss enrichments in protein domain families. Related to Figure 2. Provided as a text file with tab-separated values (.tsv). This table contains the information about the Protein domain families that are significantly enriched in mutations as well as gains or losses in isoform switches. The information provided for each domain family is the following:

Column number	Column label	Description
1	Pfam_id	PFAM ID for the domain family
2	Name	Name of the domain family
3	p_switch_gain	P-value for the gain-test
4	adjp_switch_gain	Adjusted P-value for the gain-test
5	p_switch_loss	P-value for the loss-test
6	adjp_switch_loss	Adjusted P-value for the loss-test
7	p_mutation	P-value for the mutation-test
8	adjp_mutation	Adjusted P-value for the mutation-test
9	Switches_where_gained	Number of switches where domain family is gained
10	Switches_where_lost	Number of switches where domain family is lost

Table S3. Mutual exclusion analysis between switches and cancer drivers. Related to Figure 2. Provided as a text file with tab-separated values (.tsv). This table contains the analysis of mutual exclusion between functional switches, *global mutual exclusion*, and mutational drivers in the same pathway, *local mutual exclusion*. Switches present global mutual exclusion if they exhibit an extreme mutually exclusive pattern (p_mut_ex < 0.05) with at least 3 of the most frequent tumor drivers for a certain cancer type (Number_ME_drivers >= 3). Switches present local mutual exclusion if they exhibit an extreme mutually exclusive pattern (p_me_pathway_driver < 0.05) with a driver from the same pathway (indicated in Same_pathway_driver). Switches that display both local and global mutual exclusion are considered Pan-negative AS-drivers.

Column number	Column label	Description
1	GeneId	Entrez gene ID
2	Symbol	HGNC gene symbol
3	Normal_transcript	UCSC transcript id
4	Tumor_transcript	UCSC transcript id
5	Tumor	Tumor type (brca, coad, etc…)
6	p_mut_ex	P-value for the test for mutual exclusion (ME) with mutational drivers
7	Number_ME_drivers	Number of drivers with mutual exclusion (ME)
8	MS_mut_ex	Number of samples with mutation (M) and switch (S)
9	M_mut_ex	Number of samples with only M
10	S_mut_ex	Number of samples with only S
11	N_mut_ex	Number of samples without M or S
12	ME_drivers	HGNC gene symbols for the ME drivers
13	Same_pathway_driver	Pathways shared with ME drivers
14	p_me_pathway_driver	P-value for the test for mutual exclusion (ME) with drivers in the same pathway
15	MS_me_pathway_driver	Number of samples with mutation (M) and switch (S)
16	M_me_pathway_driver	Number of samples with only M
17	S_me_pathway_driver	Number of samples with only S
18	N_me_pathway_driver	Number of samples without M or S

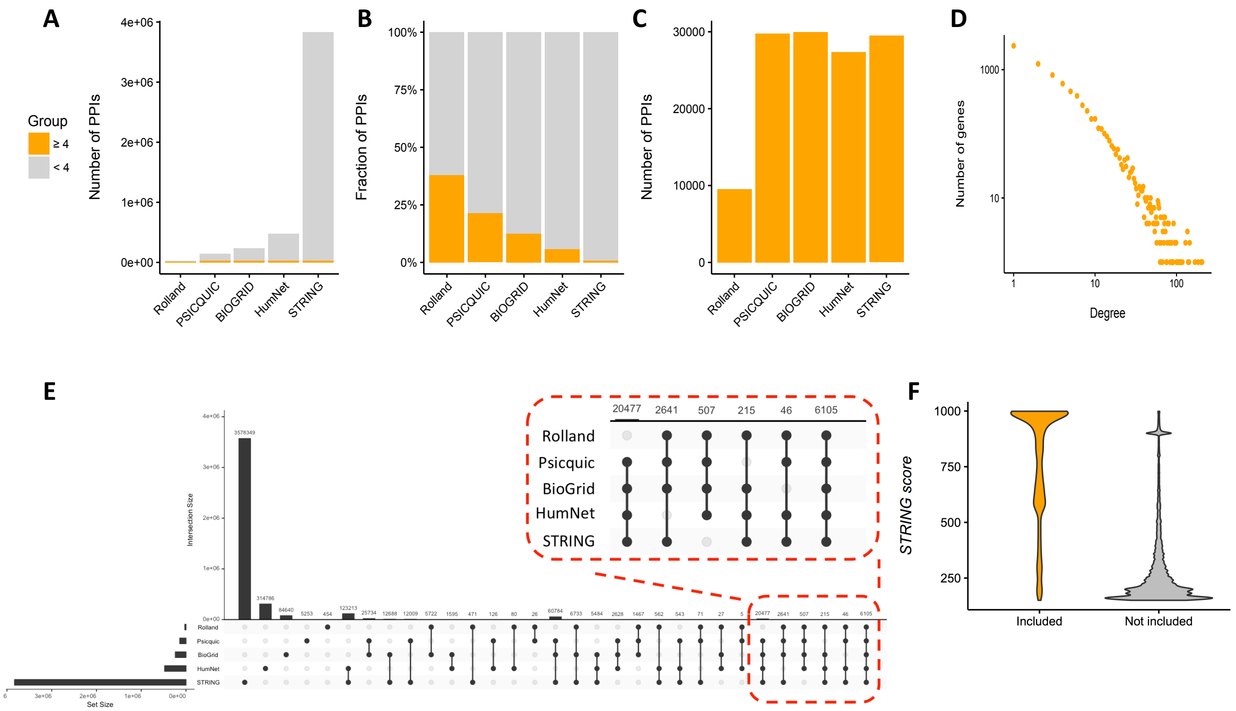

Figure S3. Protein-protein interaction network. Related to Figure 3**. (A)** Consensus protein–protein interaction (PPI) network. We used data from five different sources: PSICQUIC, BIOGRID, HumNet, STRING, and (Rolland et al., 2014). These networks vary in their size, connectivity, and origin, with PSICQUIC, BIOGRID, and Rolland being experimental networks and HumNet and STRING being functional networks. To build our consensus network, we used only those interactions that were defined in at least four different networks (shown in orange). **(B)** Fraction of each network included in the consensus network, with the data from (Rolland et al., 2014) having over 30% of its interactions and STRING less than 5%. **(C)** Number of interactions from each network included in the consensus network. **(D)** Degree distribution of the consensus network. For each number of PPI connections (*x*-axis), we give the number of genes with this degree (*y*-axis). **(E)** Highlighted in red are the PPIs considered for our analysis. Despite the fact that the dataset published in Rolland et al. was obtained through a search for new protein-protein interactions, many interactions in Rolland et al. are also present in the other PPI databases, with only 454 unique to Rolland et al. The plot also shows that even though many interactions are only present in STRING, most of them are not taken into account in our analysis. Plot performed with UpSetR (Conway et al., 2017). The horizontal bars indicate the number of switches for each property. The vertical bars indicate the number of switches in each of the intersections indicated by connected bullet points**. (F)** STRING PPIs included in our analysis (present in at least three other databases) are enriched for high-scoring interactions.

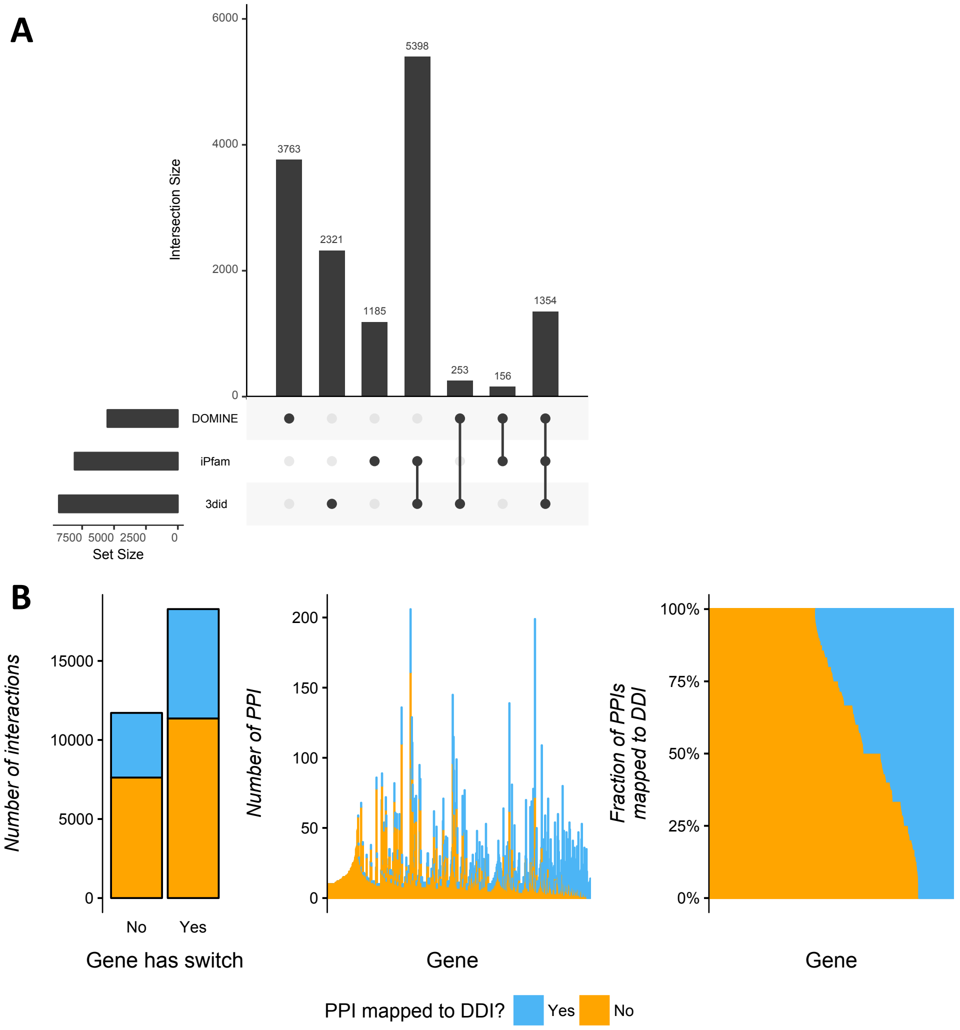

Figure S4. Protein-protein interactions assigned to functional isoform switches. Related to Figure 3**. (A)** Number of domain–domain interactions (DDIs) analyzed, separated by source: 3did, iPfam, DOMINE. The plot shows the number of cases in each source (horizontal bars) and the intersections between the sources (vertical bars), which are indicated by connected bullet points **(B)** Mapping of switches to protein-protein interactions (PPIs). Left panel: From a total of 29991 PPIs, 11008 of them were mapped to DDIs, 6917 of them in genes with switches whereas 4091 are in genes without switches. The rest of the 18983 PPIs did not map to DDIs: 11361 corresponded to genes with switches, and 7622 to genes without switches. Middle panel: Absolute number of PPI interactions mapped (blue) or not mapped (orange) to a DDI in each gene (only genes with at least 10 PPIs are depicted). Genes are sorted according to the fraction of interactions that could be mapped to DDIs. The picture shows no correlation between the degree of a gene and the fraction of interactions mapped. Right panel: Fraction of PPIs mapped to DDIs per gene. Genes are sorted according to the fraction of PPIs successfully mapped to DDIs.

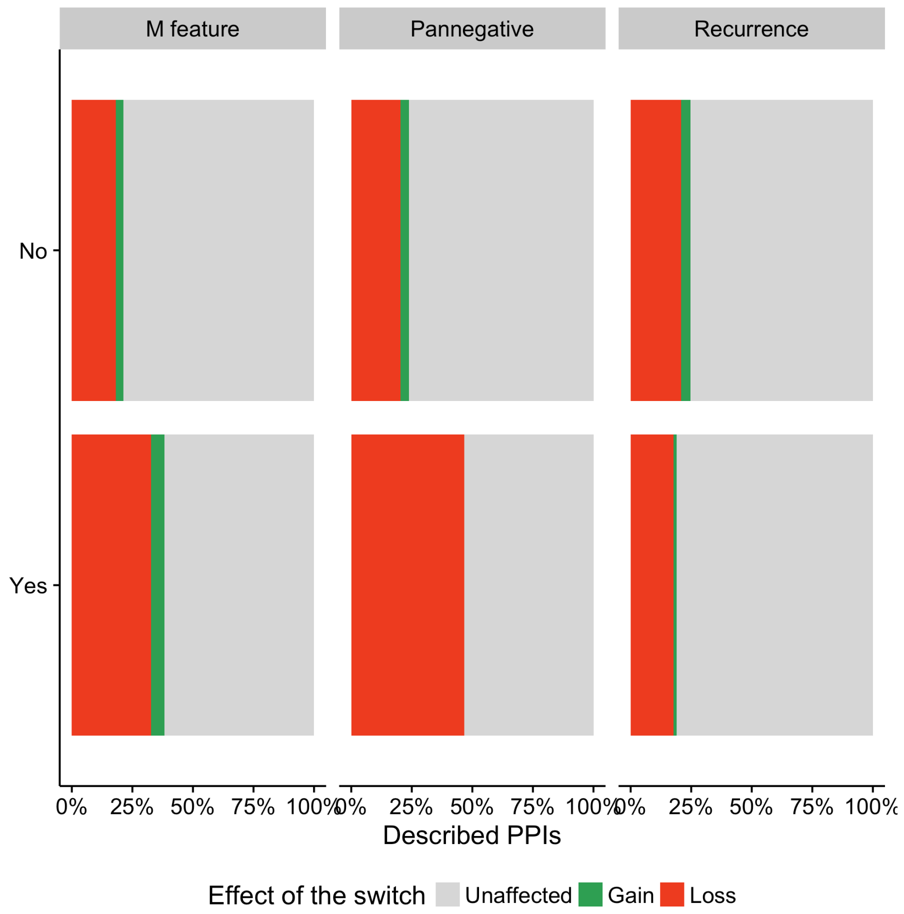

Figure S5. Properties of switches that affect protein-protein interactions. Related to Figure 3. Comparison of proportions of functional switches that affect protein-protein interactions (PPIs). In the left panel, functional switches are divided according to whether they affect domains frequently mutated in cancer (M feature) (Yes) or not (No). In the middle panel, functional switches are divided according to whether the switch has significant mutual exclusion with tumor-specific drivers (Pannegative). In the right panel, functional switches are divided according to whether they are recurrent (Yes) or not (No). In each subset we plot the proportion of PPIs that are kept unaffected (gray), lost (red), or gained (green). Using these three categories and the two values for each feature, M feature and Pannegative associate frequently with PPI-affecting switches (Chi-square test p-value < 2.2e-16 and p-value = 6.8e-08, respectively).

Table S4. Protein features and protein-protein interactions affected by isoform switches. Related to Figure 3. Provided as a text file with tab-separated values (.tsv). This table contains the proteins features and protein–protein interactions affected in each functional switch. The column descriptions are:

Column number	Column label	Description
1	Tumor	Tumor type (brca, coad, etc…)
2	GeneId	Entrez gene ID
3	Symbol	HGNC gene symbol
4	Normal_transcript	UCSC transcript id
5	Tumor_transcript	UCSC transcript id
6	Feature_type	Pfam, Prosite, IUPRED, ANCHOR
7	Feature_id	ID for the protein feature if available
8	Feature_name	Name of Feature if available, positions in protein for IUPRED and ANCHOR
9	Observation	Gained_in_tumor/Lost_in_tumor/No_change
10	Normal_isoform_order	Domain copy this corresponds to / total copies in normal isoform
11	Tumor_isoform_order	Domain copy this corresponds to / total copies in tumor isoform
12	GeneId_partner	Entrez ID of the protein-protein interaction partner
13	Symbol_partner	HGNC symbol of the protein-protein interaction partner
14	Transcript_partner	Transcripts identified as coding the interaction partner
15	Pfam_id_partner	PFAM ID for the domain mediating the interaction
16	Effect_on_interaction	Unaffected/Gain/Loss/NA(no interaction data)

Table S5. Pathways enriched in PPI-affecting switches. Related to Figure 3. Provided as a text file with tab-separated values (.tsv). This table contains the gene sets that are enriched in isoform switches that are predicted to affect protein-protein interactions. The enrichment tests is a Fisher's exact test based on the separations of switches being in the pathway or not, and affecting PPIs or not. We have tested Pathways, Complexes and gene sets-related to mRNA-metabolism. Only Pathways showed enrichment after multiple-test correction. The column descriptions are:

Column number	Column label	Description
1	Geneset_type	Pathway/Complex/mRNA_regulation
2	Geneset	Name of the gene set
3	Number_drivers	Number of drivers in the gene set.
4	p	Fisher's exact test p-value
5	adjp	p-value corrected for multiple testing
6	OR	Odds-ratio
7	eOR	Estimated odds-ration using with pseudocounts
8	Switched_genes	Genes in the gene set that have a PPI-affecting switch

Table S6. Gene modules with protein-protein interactions affected by isoform switches. Related to Figure 3. Provided as a text file with tab-separated values (.tsv). This table contains modules with high density of affected interactions: sets of genes that are connected in the network of protein-protein interactions and many of their interactions are affected by the isoform switches and separately from other genes in the PPI network. We provide a test for assigning a complex or pathway based on the intersection of the complex/pathway to the module (see Experimental Procedures for details). The column descriptions are:

Column number	Column label	Description
1	Module	Module number
2	Module_components	Genes in the module (calculated from the network of protein-protein interactions affected by isoform switches)
3	Geneset	Name of complex/pathway compared to the module (NA if none was assigned)
4	Geneset_size	Number of genes in the complex/pathway (NA if none was assigned)
5	p	p-value from binomial test for the intersection of the gene set (Complex/Pathway) to the module
6	Intersection	Number of genes from the gene set that are in the module
7	Number_drivers	Number of cancer drivers in the module
8	padj	p-value corrected for multiple testing

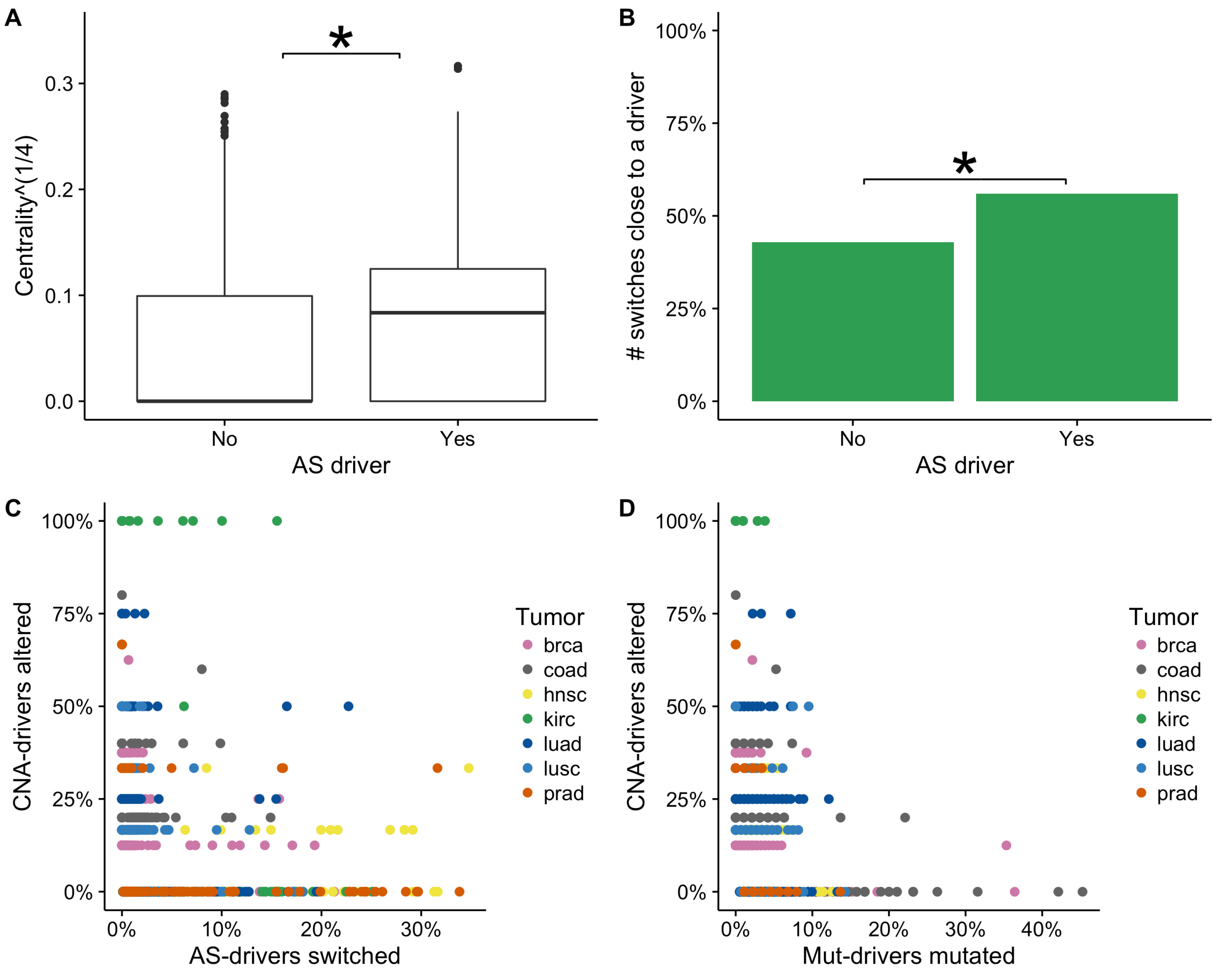

Figure S6. AS-drivers. Related to Figure 4. **(A)** We show the distribution of centrality values for switches predicted as AS-drivers (Yes) or not (No) (Mann-Whitney test p-value < 2.2e-16, W = 90999000). The y-axis shows the values of the 4th root of centrality (centrality)^(¼). **(B)** We show the proportion of AS-drivers and switches non-drivers that are separated according to the closest driver distance (CDD), calculated as the distance to the closest tumor-specific cancer gene driver in the consensus PPI network. Every switch with CDD<=3 was labelled as "Close to a driver". Otherwise, it was labelled "Far from a driver" otherwise. A Fisher's exact test on the proportion of switches AS-drivers or non AS-drivers that are close or far from a driver gives an enrichment of AS-drivers close to drivers (p-value < 2.2e-16, odds-ratio = 1.55). **(C)** Each patient is colored by tumor type and represented according to the percentage of tumor-specific copy number alteration (CNA) driver genes amplified in that sample (*y* axis) and the percentage of AS-drivers occurring in the same sample (*x* axis). **(D)** Each patient is colored by tumor type and represented according to the percentage of tumor-specific CNA driver genes amplified in that sample (*y* axis) and the percentage of mutational drivers mutated in the same sample (*x* axis)

References

Alamancos, G.P., Pagés, A., Trincado, J.L., Eyras, E., Pages, A., Trincado, J.L., Bellora, N., Eyras, E., 2015. Leveraging transcript quantification for fast computation of alternative splicing profiles. RNA 21, 1521–1531. doi:10.1261/rna.051557.115

Conway, J.R., Lex, A., Gehlenborg, N., 2017. UpSetR: An R Package for the Visualization of Intersecting Sets and their Properties 2–5. doi:10.1101/120600

Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S.J., Zhong, Q., Sahni, N., Yi, S., Lemmens, I., Fontanillo, C., Mosca, R., Kamburov, A., Ghiassian, S.D., Yang, X., Ghamsari, L., Balcha, D., Begg, B.E., Braun, P., Brehme, M., Broly, M.P., Carvunis, A.R., Convery-Zupan, D., Corominas, R., Coulombe-Huntington, J., Dann, E., Dreze, M., Dricot, A., Fan, C., Franzosa, E., Gebreab, F., Gutierrez, B.J., Hardy, M.F., Jin, M., Kang, S., Kiros, R., Lin, G.N., Luck, K., Macwilliams, A., Menche, J., Murray, R.R., Palagi, A., Poulin, M.M., Rambout, X., Rasla, J., Reichert, P., Romero, V., Ruyssinck, E., Sahalie, J.M., Scholz, A., Shah, A.A., Sharma, A., Shen, Y., Spirohn, K., Tam, S., Tejeda, A.O., Trigg, S.A., Twizere, J.C., Vega, K., Walsh, J., Cusick, M.E., Xia, Y., Barabási, A.L., Iakoucheva, L.M., Aloy, P., De Las Rivas, J., Tavernier, J., Calderwood, M.A., Hill, D.E., Hao, T., Roth, F.P., Vidal, M., 2014. A proteome-scale map of the human interactome network. Cell 159, 1212–1226. doi:10.1016/j.cell.2014.10.050