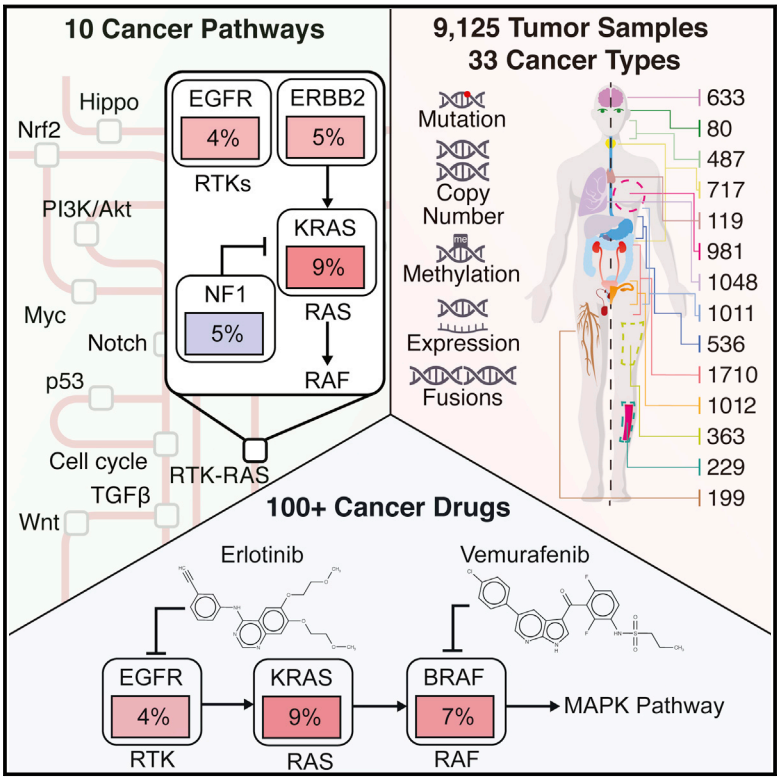
Article

Cell

Oncogenic Signaling Pathways in The Cancer Genome Atlas

Graphical Abstract



Highlights

- Alteration map of 10 signaling pathways across 9,125 samples from 33 cancer types
- Reusable, curated pathway templates that include a catalogue of driver genes
- 57% of tumors have at least one potentially actionable alteration in these pathways
- Co-occurrence of actionable alterations suggests combination therapy opportunities

Authors

Francisco Sanchez-Vega, Marco Mina, Joshua Armenia, ..., Giovanni Ciriello, Chris Sander, Nikolaus Schultz

Correspondence

Giovanni.Ciriello@unil.ch (G.C.), sander.research@gmail.com (C.S.), schultz@cbio.mskcc.org (N.S.)

In Brief

An integrated analysis of genetic alterations in 10 signaling pathways in >9,000 tumors profiled by TCGA highlights significant representation of individual and co-occurring actionable alterations in these pathways, suggesting opportunities for targeted and combination therapies.

Sanchez-Vega et al., 2018, Cell *173*, 321–337
April 5, 2018 © 2018 Elsevier Inc.
https://doi.org/10.1016/j.cell.2018.03.035

CellPress

Cell

Oncogenic Signaling Pathways in The Cancer Genome Atlas

Francisco Sanchez-Vega,[1,2,30] Marco Mina,[4,30] Joshua Armenia,[1,2,30] Walid K. Chatila,[1] Augustin Luna,[5] Konnor C. La,[1] Sofia Dimitriadoy,[6] David L. Liu,[7] Havish S. Kantheti,[8] Sadegh Saghafinia,[4] Debyani Chakravarty,[1] Foysal Daian,[1] Qingsong Gao,[9] Matthew H. Bailey,[9] Wen-Wei Liang,[9] Steven M. Foltz,[9] Ilya Shmulevich,[10] Li Ding,[9,11] Zachary Heins,[1] Angelica Ochoa,[1] Benjamin Gross,[1] Jianjiong Gao,[1] Hongxin Zhang,[1] Ritika Kundra,[1] Cyriac Kandoth,[1] Istemi Bahceci,[12] Leonard Dervishi,[12] Ugur Dogrusoz,[12] Wanding Zhou,[13] Hui Shen,[13] Peter W. Laird,[13] Gregory P. Way,[24] Casey S. Greene,[24] Han Liang,[25] Yonghong Xiao,[26] Chen Wang,[27] Antonio Iavarone,[28] Alice H. Berger,[14] Trever G. Bivona,[15] Alexander J. Lazar,[16] Gary D. Hammer,[17] Thomas Giordano,[18] Lawrence N. Kwong,[19]

(Author list continued on next page)

[1]Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[2]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[3]Departments of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[4]Department of Computational Biology, University of Lausanne (UNIL), 1011 Lausanne, Vaud, Switzerland and Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland
[5]cBio Center, Dana-Farber Cancer Institute, Boston, MA; Department of Cell Biology, Harvard Medical School, Boston, MA
[6]Princeton University, Princeton, NJ, USA
[7]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA; Cancer Program, Broad Institute of MIT and Harvard, Cambridge, MA, US
[8]University of Texas at Dallas, Richardson, TX 75080, USA
[9]Department of Medicine and McDonnell Genome Institute, Washington University in St. Louis, St. Louis, Missouri, 63110, USA
[10]Institute for Systems Biology, Seattle, WA, USA
[11]Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO 63110, USA
[12]Computer Engineering Department, Bilkent University, Ankara 06800, Turkey
[13]Van Andel Research Institute, 333 Bostwick Ave NE, Grand Rapids Michigan, 49503, USA
[14]Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
[15]UCSF Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, 1450 3rd Street, San Francisco, California 94143, USA
[16]Departments of Pathology, Genomic Medicine & Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd-Unit 85, Houston, Texas 77030, USA

(Affiliations continued on next page)

SUMMARY

Genetic alterations in signaling pathways that control cell-cycle progression, apoptosis, and cell growth are common hallmarks of cancer, but the extent, mechanisms, and co-occurrence of alterations in these pathways differ between individual tumors and tumor types. Using mutations, copy-number changes, mRNA expression, gene fusions and DNA methylation in 9,125 tumors profiled by The Cancer Genome Atlas (TCGA), we analyzed the mechanisms and patterns of somatic alterations in ten canonical pathways: cell cycle, Hippo, Myc, Notch, Nrf2, PI-3-Kinase/Akt, RTK-RAS, TGFβ signaling, p53 and β-catenin/Wnt. We charted the detailed landscape of pathway alterations in 33 cancer types, stratified into 64 subtypes, and identified patterns of co-occurrence and mutual exclusivity. Eighty-nine percent of tumors had at least one driver alteration in these pathways, and 57% percent of tumors had at least one alteration potentially targetable by currently available drugs. Thirty percent of tumors had multiple targetable alterations, indicating opportunities for combination therapy.

INTRODUCTION

Cancer is a disease in which cells have acquired the ability to divide and grow uncontrollably (Hanahan and Weinberg, 2000, Hanahan and Weinberg, 2011), usually through genetic alterations in specific genes. Advances in DNA sequencing over the past decade have made it possible to systematically study these genetic changes, and we now have a better understanding of the commonly involved processes and signaling pathways (Garraway and Lander, 2013; Vogelstein et al., 2013). As more genetic alterations become targetable by specific drugs, DNA sequencing is becoming part of routine clinical care (Hartmaier et al., 2017; Schram et al., 2017; Sholl et al., 2016; Zehir et al., 2017). However, there is considerable variation in the genes and pathways altered across different tumor types and individual tumor samples, and a complete understanding of the genes and

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Grant McArthur,[20,29] Chenfei Huang,[21] Aaron D. Tward,[22] Mitchell J. Frederick,[21] Frank McCormick,[23] Matthew Meyerson,[7] and The Cancer Genome Atlas Research Network, Eliezer M. Van Allen,[7,31] Andrew D. Cherniack,[7,31] Giovanni Ciriello,[4,31,*] Chris Sander,[5,31,*] Nikolaus Schultz[1,2,3,31,32,*]
[17]Department of Internal Medicine, Division of Metabolism, Endocrinology and Diabetes, Endocrine Oncology Program, University of Michigan, Ann Arbor, Michigan, MI 48105, USA
[18]Department of Pathology, University of Michigan Medical School, Ann Arbor, MI; Department of Internal Medicine, Division of Metabolism, Endocrinology & Diabetes, University of Michigan Medical School, Ann Arbor, MI; Comprehensive Cancer Center, Michigan Medicine, Ann Arbor, MI, USA
[19]Department of Translational Molecular Pathology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[20]Peter MacCallum Cancer Centre, Melbourne, VIC, Australia
[21]Dept. of Otolaryngology, Baylor College of Medicine, USA
[22]University of California, San Francisco Department of Otolaryngology-Head and Neck Surgery. 2233 Post Street, San Francisco, CA, 94143, USA
[23]UCSF Helen Diller Family Comprehensive Cancer Center, University of California San Francisco, 1450 3rd Street, San Francisco, CA 94143, USA
[24]Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania, Philadelphia, PA, 19104, USA
[25]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX, 77030, USA
[26]TESARO Inc., Waltham, MA, 02451, USA
[27]Department of Health Sciences Research and Department of Obstetrics and Gynecology, Mayo Clinic College of Medicine, 200 First Street SW, Rochester, MN, 55905, USA
[28]Institute for Cancer Genetics, Department of Neurology and Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY, 10032, USA
[29]University of Melbourne, Melbourne, VIC, Australia
[30]These authors contributed equally
[31]Co-senior author
[32]Lead contact
*Correspondence: Giovanni.Ciriello@unil.ch (G.C.), sander.research@gmail.com (C.S.), schultz@cbio.mskcc.org (N.S.)
https://doi.org/10.1016/j.cell.2018.03.035

pathways altered in all cancer types is essential to identify potential therapeutic options and vulnerabilities.

Several important signaling pathways have been identified as frequently genetically altered in cancer, including the RTK/RAS/MAP-Kinase (hereafter also called RTK-RAS for brevity) pathway, PI3K/Akt signaling, and others (Vogelstein and Kinzler, 2004). Members of these pathways and their interactions have been captured in a number of pathway databases, such as Pathway Commons (Cerami et al., 2011), which aggregates a number of databases, including REACTOME (Joshi-Tope et al., 2005) and KEGG (Kanehisa and Goto, 2000). Genes in key pathways are not altered at equal frequencies, with certain genes recurrently altered and well-known in cancer, while others are only rarely or never altered.

The detection of recurrence of rare alterations often requires large numbers of samples (Lawrence et al., 2014). This is confounded by the challenge to distinguish between functionally relevant (or "driver" alterations) and non-oncogenic "passenger" events (Gao et al., 2014), especially in tumor types with a high background mutation burden (Alexandrov et al., 2013; Lawrence et al., 2013). In these cases, many mutations, even when they occur in cancer genes, may have no functional effect. This topic is further addressed in Bailey et al. (Bailey et al., 2018).

Previous studies by The Cancer Genome Atlas (TCGA) have incrementally mapped out the alteration landscape in signaling pathways. Certain pathways, such as RTK-RAS signaling or the cell-cycle pathway, are altered at high frequencies across many different tumor types, whereas other pathways are altered in more specific subsets of malignancies (e.g., alterations in the oxidative stress response pathway are strongly associated with squamous histologies). With >10,000 samples characterized by TCGA, there is an opportunity to systematically characterize and define the alterations within well-known cancer pathways across all tumor types and map out commonalities and differences across pathways. The existence of shared genomic features across histologies has been highlighted by several studies (Ciriello et al., 2013; Hoadley et al., 2014, 2018), but these studies traditionally used a gene-centric, as opposed to pathway-centric, approach. Identifying relationships of inter- and intra-pathway recurrence, co-occurrence or mutual exclusivity across different types of cancers can help elucidate functionally relevant mechanisms of oncogenic pathway alterations that might inform treatment options.

Here, we worked within the framework of the TCGA PanCancer Atlas initiative (Cancer Genome Atlas Research Network et al., 2013c) to build a uniformly processed dataset and a unified data analysis pipeline aimed at exploring similarities and differences in canonical cancer pathway alterations across 33 cancer types. The focus of this effort is on mitogenic signaling pathways with vidence for functional alterations; other oncogenic processes, such as alterations in DNA repair (Knijnenburg et al., 2018), the spliceosome (Seiler et al., 2018), ubiquitination (Ge et al., 2018), or metabolic pathways (Peng et al., 2018), as well as the effects of splicing mutations (Jayasinghe et al., 2018), are covered by other efforts of the TCGA PanCancer Atlas project.

RESULTS

Dataset

We evaluated all samples in the TCGA PanCancer Atlas collection for which the following data types were available: somatic mutations (whole-exome sequencing), gene expression levels

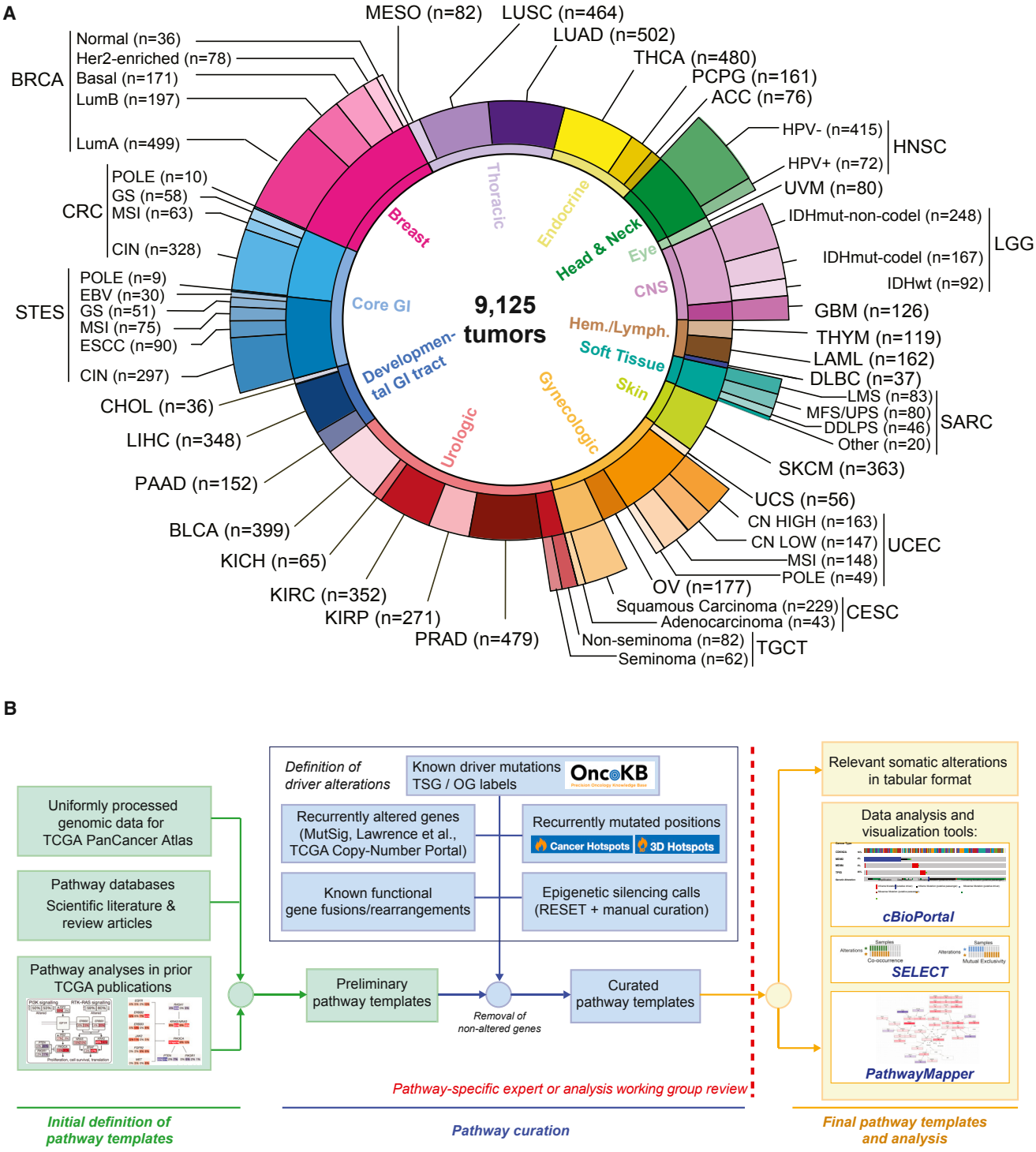


Figure 1. TCGA PanCancer Atlas Pathways Dataset and Workflow

(A) Distribution of cancer types in the cohort, including molecular subtypes analyzed.

(B) Workflow for pathway curation and analysis. Genes were curated from previous TCGA efforts and the scientific literature. Only genes with evidence for statistically recurrent or known driver alterations in the uniformly processed TCGA PanCancer Atlas dataset were included in the curated pathway templates. TCGA disease codes and abbreviations: AML, acute myeloid leukemia; ACC, adrenocortical carcinoma; BRCA, breast cancer; CESC, cervical cancer; KICH, chromophobe renal cell carcinoma; KIRC: clear cell kidney carcinoma; CRC, colorectal adenocarcinoma; SKCM, cutaneous melanoma; DLBC, diffuse large B cell lymphoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; LGG, lower grade glioma;

(legend continued on next page)

(RNA-Seq), DNA copy-number alterations (Affymetrix SNP6 arrays), and DNA methylation (Infinium arrays). This resulted in a final set of 9,125 samples from 33 different cancer types (Figure 1A, Table S1). In order to account for molecular or histological subtypes, these cancer types were further stratified into a total of 64 genomically distinct tumor subtypes, as previously defined by the individual TCGA analysis working groups (Figure 1A, Table S1). All genomic data and clinical attributes per sample can be visualized through the cBioPortal for Cancer Genomics at http://www.cbioportal.org/ (Cerami et al., 2012).

Definition of Pathways and Alterations

We evaluated 10 canonical signaling pathways with frequent genetic alterations, starting with key cancer genes explored in these pathways in previous TCGA publications, and focused on pathway members likely to be cancer drivers (functional contributors) or therapeutic targets. The pathways analyzed are: (1) cell cycle, (2) Hippo signaling, (3) Myc signaling, (4) Notch signaling, (5) oxidative stress response/Nrf2, (6) PI-3-Kinase signaling, (7) receptor-tyrosine kinase (RTK)/RAS/MAP-Kinase signaling, (8) TGFβ signaling, (9) p53 and (10) β-catenin/Wnt signaling (Figures 2 and S1, Table S2). Alterations in DNA repair pathways, epigenetic modifiers, splicing, and other cellular processes frequently altered in cancer were not included, as these primarily provide a background of genomic instability, rather than specifically proliferative potential.

We began by compiling and reviewing the full set of cancer-type specific pathway diagrams from the compendium of TCGA manuscripts published between 2008 and 2017 (Brennan et al., 2013; Cancer Genome Atlas Network, 2012a, 2012b, 2015a, 2015b, Cancer Genome Atlas Research Network, 2008, 2011, 2013a, 2013b, 2014a, 2014b, 2014c, 2014d, 2017a, 2017b; Davis et al., 2014), each of which included the pathway genes found to be genetically altered in the individual tumor types. These pathway diagrams are publicly available as predefined network templates within the www.PathwayMapper.org visualization tool (Bahceci et al., 2017). By taking the union of pathway members across multiple TCGA studies, we produced a consolidated list of candidate member genes for each of the ten pathways. These were then further curated based on updated literature (including but not limited to the references in Table S2), public pathway databases, and expert opinion (Figure 1B).

The selected genes in the ten pathways were then assessed for recurrent alterations within and across different tumor types as follows (Figure 1B): Alterations of pathway members were classified as activating events (usually specific recurrent missense mutations, i.e., *hotspot* mutations, amplifications, or fusions involving oncogenes) or inactivating events (truncating mutations, specific recurrent missense or inframe mutations, deletions, as well as fusions and promoter hypermethylation of tumor suppressor genes). Individual alterations were also scrutinized for two features: statistical recurrence across sets of tumor samples and presumed functional impact. We first assessed statistical recurrence using MutSigCV (Lawrence et al., 2014) for mutations and GISTIC 2.0 (Mermel et al., 2011) for copy-number alterations. In order to identify likely functional variants, we then used recurrence across tumor samples at the residue level (linear and 3D mutational hotspots; Chang et al., 2016, 2018; Gao et al., 2017; see STAR Methods) and prior knowledge about specific variants via the OncoKB knowledge base, which contains information about the oncogenic effects and treatment implications of variants in > 400 cancer genes (Chakravarty et al., 2017a). Epigenetic silencing through promoter DNA hypermethylation of tumor suppressor genes was evaluated using the RESET algorithm (see STAR Methods). Gene fusions and structural rearrangements were called from RNA-Seq data using a combination of the STAR-Fusion, EricScript and BreakFast algorithms (Gao et al., 2018, see STAR Methods), and likely passenger events were filtered out based on OncoKB annotation. Through this process, genes without evidence for recurrent or previously known oncogenic alterations were removed from the preliminary pathway templates. The resulting curated pathway templates and the identified genetic alterations were vetted for functional importance by individual pathway experts or the corresponding TCGA PanCancer Atlas pathway-specific analysis working groups, when applicable (Figure 1B). The pathway member genes and the genetic alterations considered as oncogenic are listed in Table S3, and binary genomic alteration matrices are provided as Table S4 (see STAR Methods).

The resulting comprehensive dataset of different types of alterations across many tumor types form the basis of all subsequent analyses regarding pathways, patterns of co-occurrence and mutual exclusivity, as well as potential therapeutic implications. The simplified pathway diagrams in Figure 2 show the most frequently altered genes in the ten pathways, including alteration frequencies as well as the types of oncogenic alterations identified in each of the genes.

Pathway Alteration Frequencies per Tumor Type

For each tumor type and subtype, we computed the fraction of samples with at least one alteration in each of the 10 signaling pathways (Figure 3). A tumor sample was considered as altered in a given pathway if one or more genes in the pathway contained a recurrent or known driver alteration (as described above). Despite the fact that non-recurrent and not previously known alterations were filtered out as likely passenger events and were not included in the alteration frequencies, the microsatellite instability (MSI) and polymerase ε (POLE) mutant subtypes of

LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OV, ovarian serous cystadenocarcinoma; KIRP, papillary kidney carcinoma; THCA, papillary thyroid carcinoma; STAD, stomach adenocarcinoma; PRAD, prostate adenocarcinoma; BLCA, urothelial bladder cancer; UCS, uterine carcinosarcoma; UCEC, uterine corpus endometrial carcinoma; ESCA, esophageal cancer; PCPG, pheochromocytoma & paraganglioma; PAAD, pancreatic ductal adenocarcinoma; MESO, mesothelioma; UVM, uveal melanoma; SARC, sarcoma; CHOL, cholangiocarcinoma; TGCT, testicular germ cell cancer; THYM, thymoma; STES, stomach and esophageal cancer; EBV, Epstein-Barr virus; HPV, human papillomavirus; DDLPS, dedifferentiated liposarcoma; LMS, leiomyosarcoma; MFS/UPS, myxofibrosarcoma/undifferentiated pleomorphic sarcoma; ESCC, esophageal squamous cell carcinoma; GS, genomically stable; CIN, chromosomal instability; MSI, microsatellite instability.

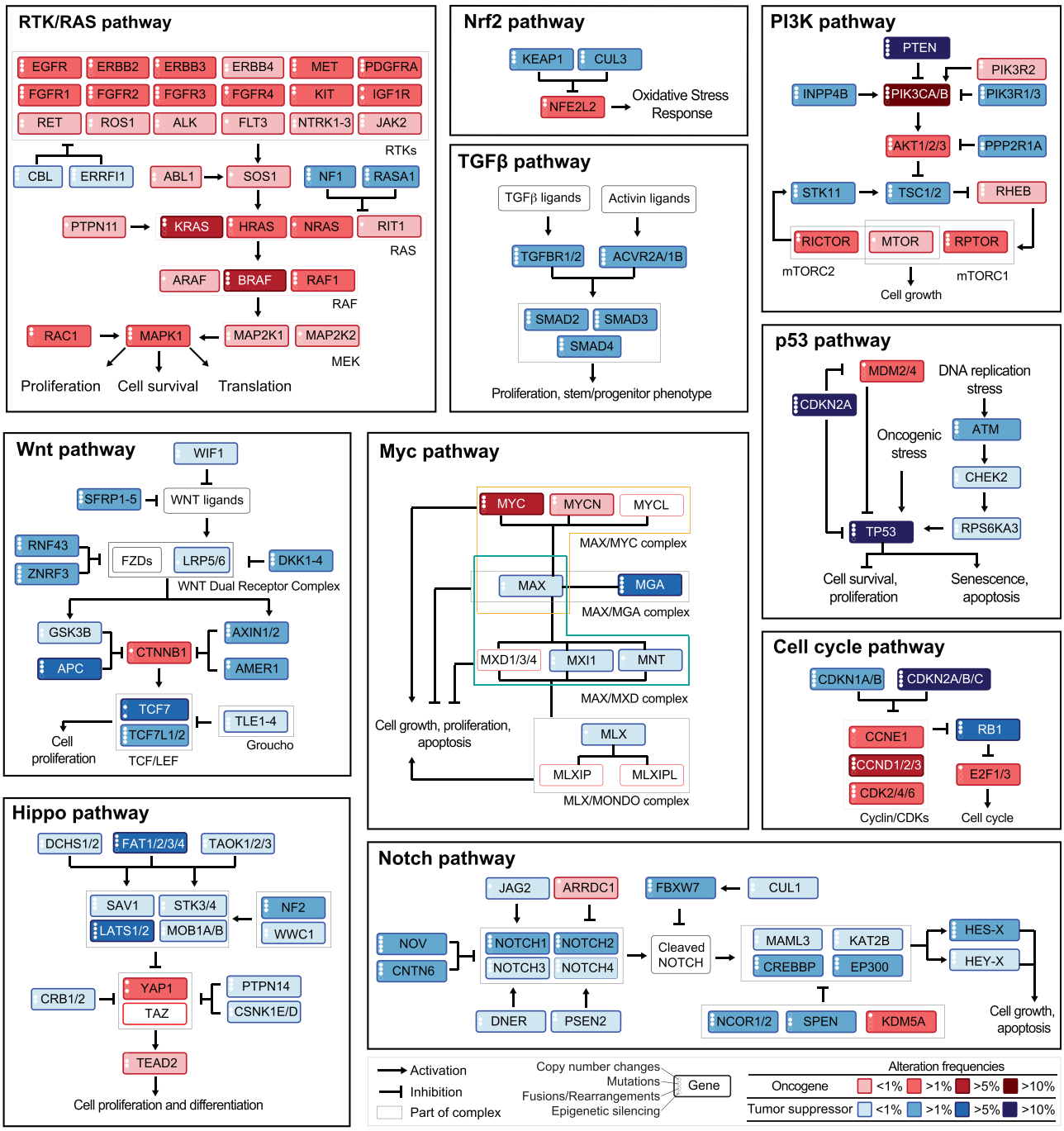


Figure 2. Curated Pathways

Pathway members and interactions in the ten selected pathways. Genes are altered at different frequencies (color intensity indicates the average frequency of alteration within the entire dataset) by oncogenic activations (red) and tumor suppressor inactivations (blue). The types of somatic alteration considered for each gene (copy-number alterations, mutations, fusions or epigenetic silencing) are specified using a set of four vertical dots on the left of each gene symbol. An expanded version including cross-pathways interactions is provided as Figure S1.

gastrointestinal and uterine tumors, which had the highest mutation burden, also had the highest overall frequencies of pathway alterations. This is possibly due to the frequent inactivating mutations introduced by the predominant mutation mechanisms in these tumor types (Boland and Goel, 2010; Rayner et al., 2016).

The RTK-RAS pathway was the signaling pathway with the highest median frequency of alterations (46% of samples)

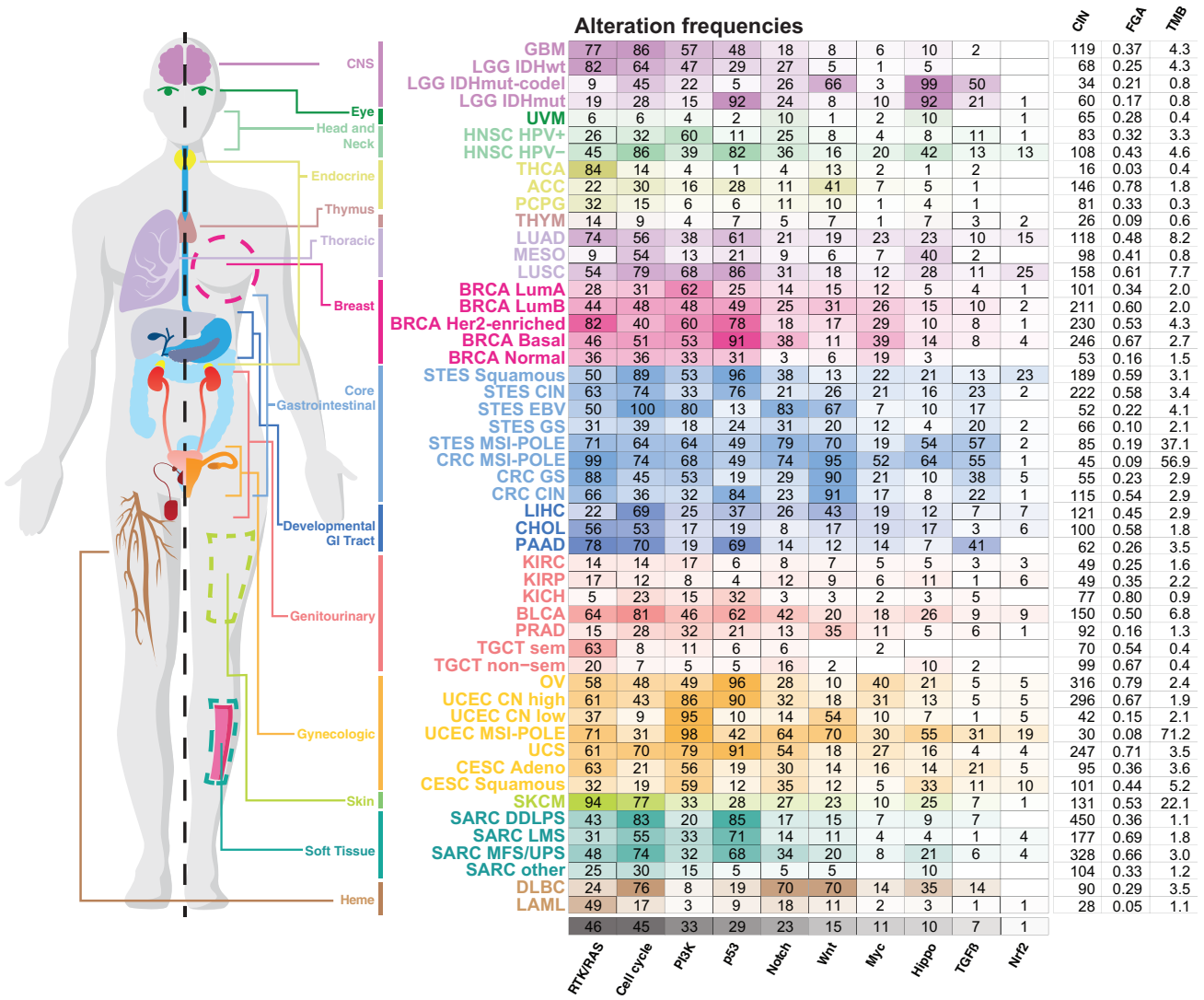
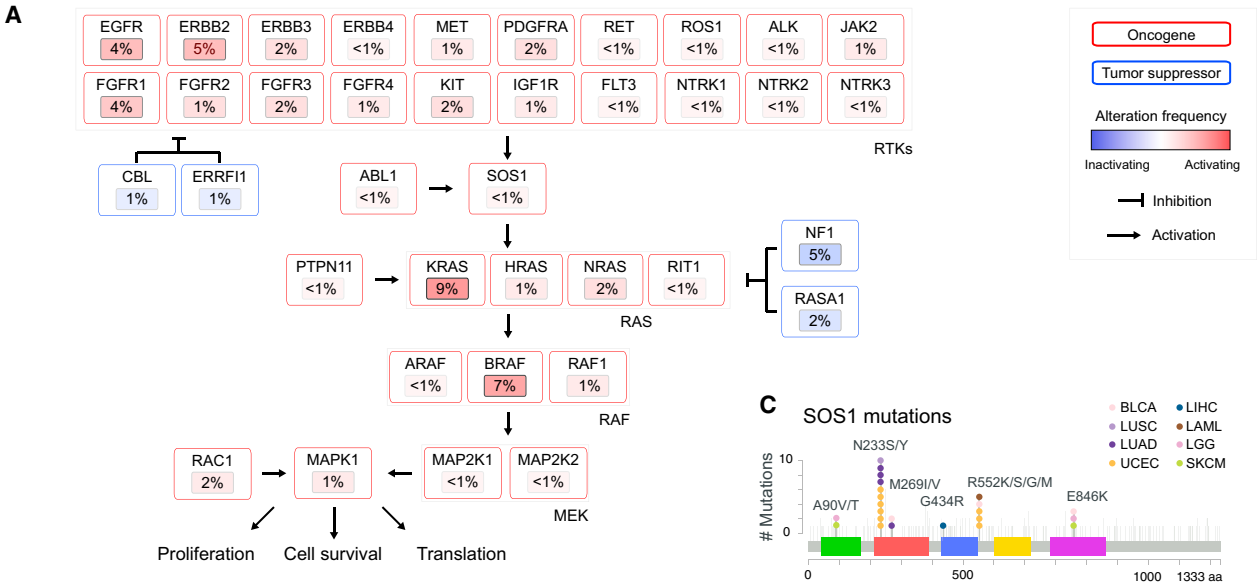


Figure 3. Pathway Alteration Frequencies

Fraction of altered samples per pathway and tumor subtype. Pathways are ordered by decreasing median frequency of alterations. Increasing color intensities reflect higher percentages. Average mutation count, as well as number of unbalanced segments and fraction genome altered (two measures of the degree of copy-number alterations) per cancer subtype are also provided. The MSI and POLE subtypes were grouped in this figure in colorectal, stomach and endometrial cancers.

across all cancer types. The tumor subtypes with the highest fraction of alterations in this pathway were (in descending order): melanoma (SKCM, 94% altered), the genomically-stable subtype of colorectal cancer (CRC GS, 88%), Her2-enriched breast cancer (BRCA Her2-enriched, 82%), pancreatic cancer (PAAD, 78%), IDH1-wild-type glioma (LGG IDHwt, 82%), lung adenocarcinoma (LUAD, 74%), and thyroid carcinoma (THCA, 84%). Some tumor types, such as lung squamous cell carcinoma (LUSC), EBV-positive esophagogastric cancer (STES EBV), and non-hypermutated uterine cancer (UCEC CN high and CN low), had high rates of alterations in the PI3K pathway, altered in 68%, 80%, 86%, and 95% of samples, respectively. While cell-cycle alterations were common in many tumor types, the pathway was only rarely altered in uveal melanoma (UVM), thymoma (THYM), testicular cancer (TGCT), and acute myeloid leukemia (AML). Alterations in the Wnt pathway were the most variable across cancer types. Colorectal cancer had near universal activation of this pathway, while others, such as renal cell carcinomas and breast cancer, had very low frequencies of alteration in genes in this pathway. The oxidative stress response/Nrf2 pathway had the lowest overall frequency of alteration (1% of samples), and it was altered most frequently in lung squamous (25% altered) and esophagogastric squamous cell carcinoma (STES ESCC, 23%).

Particularly interesting gene alterations across tumor types were observed in the RTK-RAS pathway. *KRAS* was the most frequently altered gene (9% across all samples), followed by *BRAF* (7%) and *EGFR* (4%) (Figure 4A). *KRAS* alterations were



(legend on next page)

most common in pancreatic carcinoma (PAAD, 72%), genomically stable colorectal cancer (69%), and lung adenocarcinoma (33%) (Figure 4B). *BRAF* alterations (predominantly known hotspot mutations) were found in melanoma and thyroid carcinoma, altered in 51% and 62% of samples, respectively. *EGFR* alterations were predominantly found in glioblastoma (GBM, 50%), low grade glioma IDHwt (52%), HPV-negative head and neck cancer (HNSC HPV-, 13%), lung adenocarcinoma (13%), and esophagogastric squamous carcinoma (14%), while *ERBB2* alterations were found most commonly in breast cancer and chromosomally unstable esophagogastric carcinoma (STES CIN 26% altered), as well as cervical carcinoma (CESC 23% altered). While most of the alterations described here were previously reported as functional contributors, we identified relatively rare potentially oncogenic alterations in *SOS1* (<1%). *SOS1* encodes a guanine-nucleotide-exchange factor (GEF) involved in the activation of Ras proteins. Specific germline mutations in this gene are involved in Noonan syndrome (Lepri et al., 2011), and recurrent somatic mutations in *SOS1* were recently identified in otherwise RAS-pathway driver-negative lung adenocarcinoma samples (Campbell et al., 2016). We identified recurrent (hotspot) mutations (A90V/T, N233Y/S) and other known activating mutations (M269I/V, G434R, R552S/K/G/M, E846K) in *SOS1* in a total of 1% of lung adenocarcinoma samples, 1% of uterine carcinomas, independent of subtype, and at lower frequencies in several other cancer types (Figure 4C). This finding suggests that rarely altered novel cancer genes can be identified as more tumor samples are profiled. A more detailed analysis of RAS pathway alterations is published separately, including a description of downstream transcriptional changes due to malfunctioning Ras signaling and results suggesting that multiple hits in the Ras pathway are capable of increasing overall Ras activity in *RAS* wild-type tumors (Way et al., 2018).

The alteration frequencies of the most commonly altered genes in the other nine pathways are in Figure 5 (full heatmaps providing frequencies of alteration for every gene in each pathway are in Figures S2 and S3). In some pathways, the alterations are distributed over many genes (e.g., cell cycle, PI3K), while in others the alterations mainly affect only a few genes (Wnt, Myc, Nrf2). Several pathways are featured in more detail as separate publications, including: (1) PI3K pathway (Zhang et al., 2017): aberrations in the PI3K pathway were found predominantly as activating events in *PIK3CA* (less commonly in *PIK3CB*) and inactivating events in *PTEN* or *PIK3R1* with *PIK3CA* and *PTEN* alterations being most commonly found in head and neck cancer, breast cancer, gastrointestinal and gynecological tumors; (2) TGFβ pathway (personal communication): The TGFβ pathway had the highest alteration rate in pancreatic and gastrointestinal cancers, while renal and brain cancers, among others, had almost no alterations in this pathway; and (3) Myc pathway (Schaub et al., 2018): Myc pathway alterations were most common in tumor types with amplification of chromosome 8, which contains *MYC*, such as breast cancer, ovarian cancer (OV), and others.

Mutual Exclusivity and Co-occurrence among Pathway Alterations

Individual tumors typically have multiple functional alterations affecting more than a single pathway. Some pathways may be the target of more than one alteration per patient or distinct pathways may be co-altered in one tumor. Patterns of mutual exclusivity between alterations across large patient cohorts have been associated either with functional redundancy, indicating that once one occurred and is selected the second will not provide a further selective advantage, or with synthetic lethality, indicating that cells cannot survive with both alterations (Etemadmoghadam et al., 2013; Mina et al., 2017). Patterns of co-occurrence of alterations in many tumor samples, on the other hand, indicate functional synergies and, importantly, may reflect resistance to therapy targeting one of the alterations (Nissan et al., 2014)

To explore significantly co-occurring and mutually exclusive alterations by pathway or by gene, we used the SELECT method (Mina et al., 2017). Among the 410 alterations characterized, we identified 156 pairs of mutually exclusive alterations and 117 pairs of co-occurring alterations (Table S5).

Upon mapping these significant pairs to the affected pathways, we found numerous mutually exclusive pairs within the p53, cell-cycle, RAS, and PI3K pathways, suggesting that one alteration is sufficient to functionally alter each of these pathways or that more than one might be disadvantageous (Figure 6A). On the other hand, the Hippo, RTK, and, to a lesser extent, Wnt pathways often had multiple alterations per tumor sample, suggesting co-occurring events that mediate synergistic activation of each pathway (Figure 6A). The SELECT method also identified several significant interdependencies between distinct pathways (Figure S4). For example, activation of RTKs was significantly mutually exclusive with alterations that promote either RAS or PI3K signaling, consistent with RTKs being able to activate either pathway without the need for additional alterations. Notable exceptions were significant co-occurrent alterations in

Figure 4. RTK-RAS Pathway Alterations

(A) Altered genes and their functional relationships in the RTK-RAS pathway. Shades of red indicate frequencies of activating events (known or likely activating mutations or fusions, amplifications) and shades of blue indicate frequencies of inactivating events (known or likely inactivating mutations or fusions, homozygous losses).

(B) Detailed heatmap of alteration frequencies in members of the RTK-RAS pathway. Only known or likely oncogenic alterations in each gene are considered, as described in STAR Methods. The individual gene alteration frequencies may add up to more than the total for each tumor type, as some tumor samples may have multiple alterations. Color side bars show the fraction of samples affected by each type of somatic alteration (or a combination of them) for each pathway gene. Top color bars show the proportion of different types of alterations for each cancer subtype.

(C) Recurrent or known functional mutations in *SOS1*. Recurrent or known mutations are color-coded by tumor type, all other mutations observed in the gene are considered variants of unknown significance (gray). Three singleton mutations involved in a 3D hotspot are not shown for space reasons: D89A in a UCS sample, A93D in CRC, and S92P in UCEC.

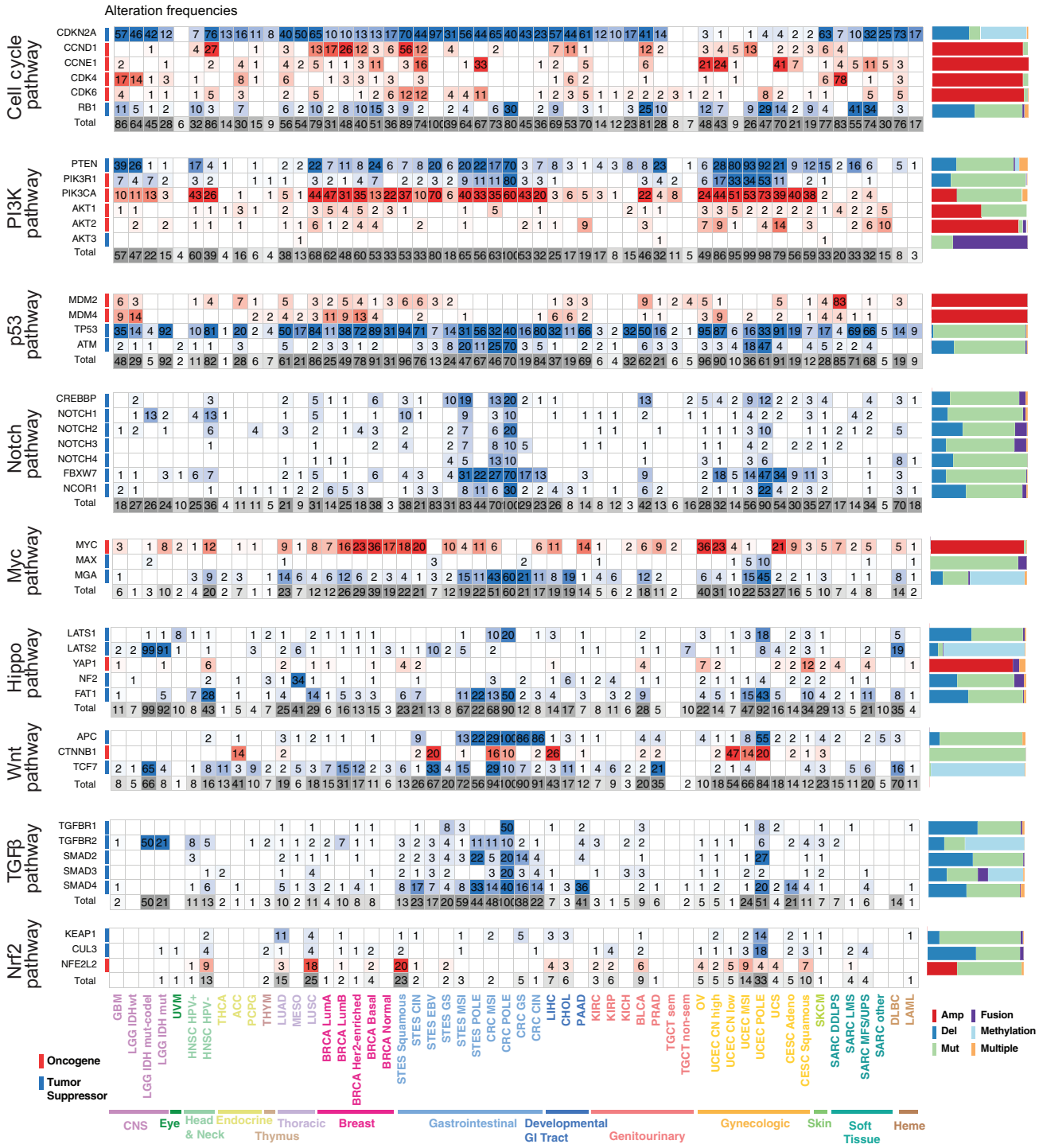
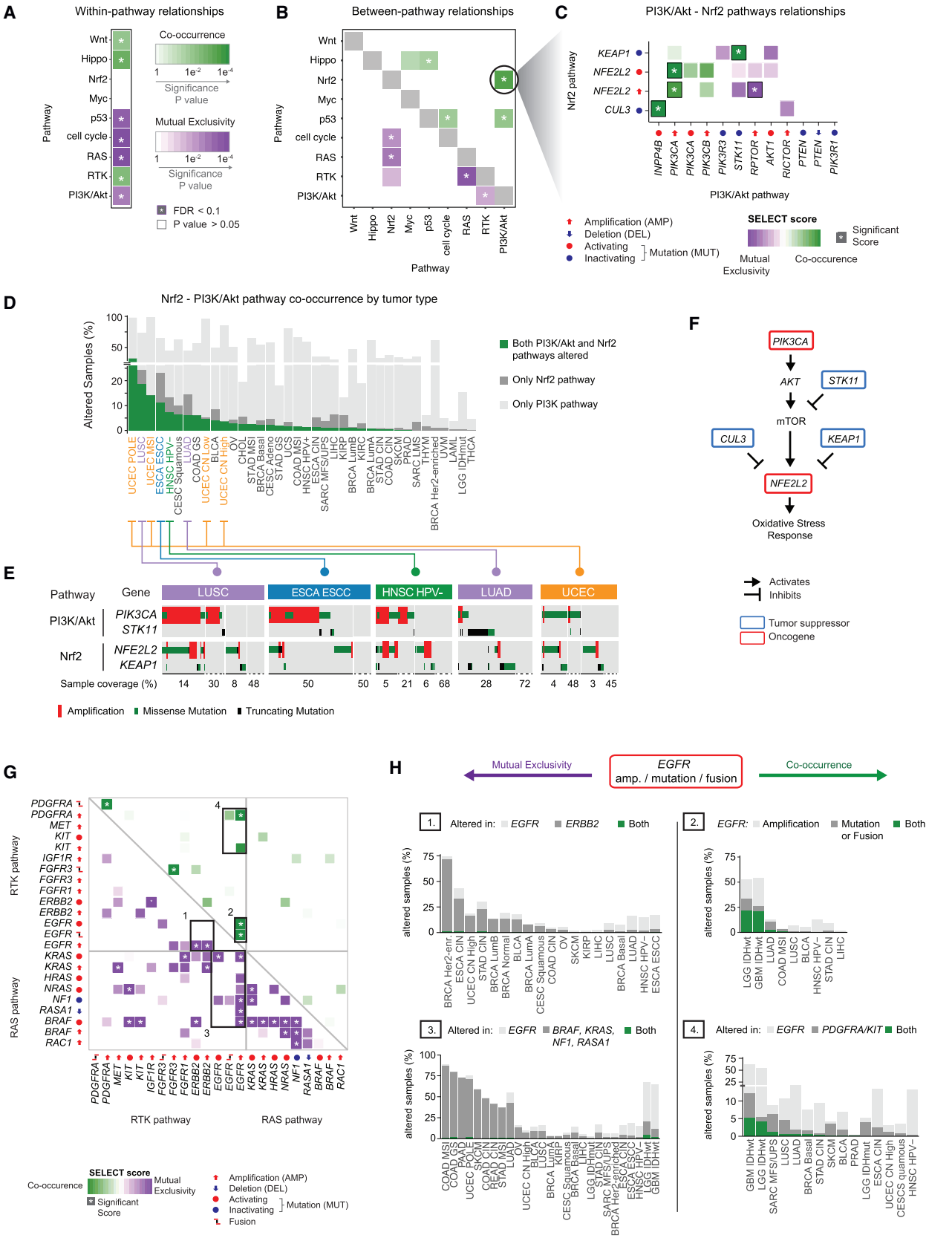


Figure 5. The Most Commonly Altered Genes in Nine Signaling Pathways

Oncogenic alteration frequencies per gene and tumor subtype for the most frequently altered genes in each pathway (for RTK-RAS see Figure 4). Red: activating events; blue: inactivating events; frequency of occurrence scale with color intensity. Last row for each pathway: overall alteration frequency of that pathway per tumor type. The individual gene alteration frequencies may add up to more than the total for each tumor type, as some tumor samples have multiple alterations. Color side bars show the fraction of samples affected by each type of somatic alteration (or a combination of them) for each pathway gene. Comprehensive heatmaps with alterations for every gene in each pathway are in Figure S2 and Figure S3.



(legend on next page)

the FGF receptors *FGFR2* and *FGFR3* and alterations in the PI3K pathway (Figure S4A). The p53 and cell-cycle pathways were frequently co-altered. Indeed, numerous alterations affecting Rb-mediated cell-cycle control were found co-occurring with *TP53* mutations. These included amplification of *CCNE1*, mutation of *CDKN2A*, *RB1* loss, and amplification of *CDK6* and *E2F3* (Figure S4B). Interestingly, *TP53* mutations were found mutually exclusive with *CDKN2A* deletion, consistent with the latter invariably affecting both p16, regulating the cell cycle, and ARF, promoting p53-dependent apoptosis. Similarly, *MDM2* amplification was significantly mutually exclusive with *RB1* and *CDKN2A* loss. However, *MDM2* is proximal to *CDK4* in the genome and the two genes were almost always co-amplified. Overall, these results indicate that p53 signaling and cell-cycle control are frequently co-altered across multiple tumor types, either through two independent events (e.g., mutations of *TP53* and *RB1*), or through a single alteration that is able to affect both pathways (e.g., *CDKN2A* deletion).

The strongest co-occurrence among pathways was found between alterations of the PI3K and Nrf2 pathways. Here, gain of function mutations and amplifications of the *NFE2L2* gene (encoding for NRF2) significantly co-occurred with *PIK3CA* amplification and tended to co-occur with *PIK3CA* mutations and *PIK3CB* amplification (Figure 6C). Interestingly, even though *NFE2L2* amplification was largely mutually exclusive with loss of *STK11* (a.k.a. *LKB1*), the latter significantly co-occurred with loss of function mutations of *KEAP1*, a negative regulator of *NFE2L2*. Co-occurring Nrf2-PI3K pathway alterations were most frequent in lung tumors (both squamous cell and adenocarcinoma), esophageal carcinomas, head and neck squamous cell carcinoma and uterine carcinoma, independent of subtype (Figure 6D). In these tumor types, alterations in *NFE2L2* and *KEAP1* were recurrent and almost perfectly mutually exclusive, and they frequently co-occurred with *PIK3CA* activation or *STK11* loss (Figure 6E). PI3K pathway activation promotes NRF2 accumulation, which, in turn, mediates metabolic pathways required to sustain cell proliferation and protection from reactive oxidative species (Mitsuishi et al., 2012). NRF2 is however kept in check by inhibitory molecules such as KEAP1 and CUL3 (Figure 6F). The observed co-occurrence between alterations of the PI3K and Nrf2 pathways suggests that bypassing these inhibitory mechanisms (either by loss of *KEAP1* or *CUL3*, or by direct over-activation of *NFE2L2*) is synergistic with active PI3K signaling. Importantly, tumors with these events might depend on NRF2 activity to tolerate PI3K pathway over-activation.

The RTK-RAS pathway contained numerous and some of the most significantly mutually exclusively altered genes, as well as several gene pairs that were significantly concurrently altered (Figure 6G). In particular, alterations promoting EGFR activation (gain of function mutations, fusion, and amplification) were involved in the highest number of significant pairs. *EGFR* amplification was significantly mutually exclusive with activation of its paralog growth factor receptor Her2 (*ERBB2,* Figure 6H, Box 1) and with key drivers of the RAS pathway, including oncogenic mutations in *BRAF* and *KRAS* as well as loss of *NF1* and *RASA1* (Figure 6G). Since oncogenic *EGFR* can be synthetically lethal with mutated *KRAS* and can mediate resistance to BRAF inhibition in colon cancer and melanoma (Sun et al., 2014; Unni et al., 2015), these results suggest a similar antagonistic interaction with loss of *NF1* or *RASA1*. Overall, alterations of either of these genes were recurrent across multiple tumor types, although almost never in the same patient (Figure 6H, Box 2). On the other hand, in glioblastoma and IDH wild-type low grade glioma, *EGFR* amplifications were highly co-occurrent with either *EGFR* mutations or gene fusions (Figure 6H, Box 3) or with focal amplifications of chromosome 4q12, where both *KIT* and *PDGFRA* are located (Figure 6H, Box 4). It should be noted that the majority of *EGFR* and *PDGFRA* fusions were found coincident with amplifications in these genes, indicating that, potentially, in these cases the same structural variant was detected as both a copy number gain and a fusion (Alaei-Mahabadi et al., 2016). Interestingly, co-amplification of *EGFR* and *PDGFRA* has been proposed to be an early event in glioblastoma development, where the two receptors heterodimerize under EGF stimulation and respond to EGFR-inhibitors (Chakravarty et al., 2017b).

Overall, these results provide a map of the cross-talk between pathways and pathway components, reflecting functional interactions and dependencies that could be therapeutically exploited.

Therapeutic Actionability

DNA sequencing has been used routinely to inform the choice of targeted therapy in specific cancer types for several years, and some institutions now apply it more broadly to guide clinical trial enrollment for many additional cancer types. A relatively small

Figure 6. Co-occurrence and Mutual Exclusivity of Pathway Alterations

(A and B) Mutual exclusivity (purple) and co-occurrence (green) of gene alterations within pathways (A) and between pathways (B). Asterisks indicate significant relationships (Q value < 0.1).

(C) Co-occurrence and mutual exclusivity of individual gene alterations in the PI3K and Nrf2 pathways.

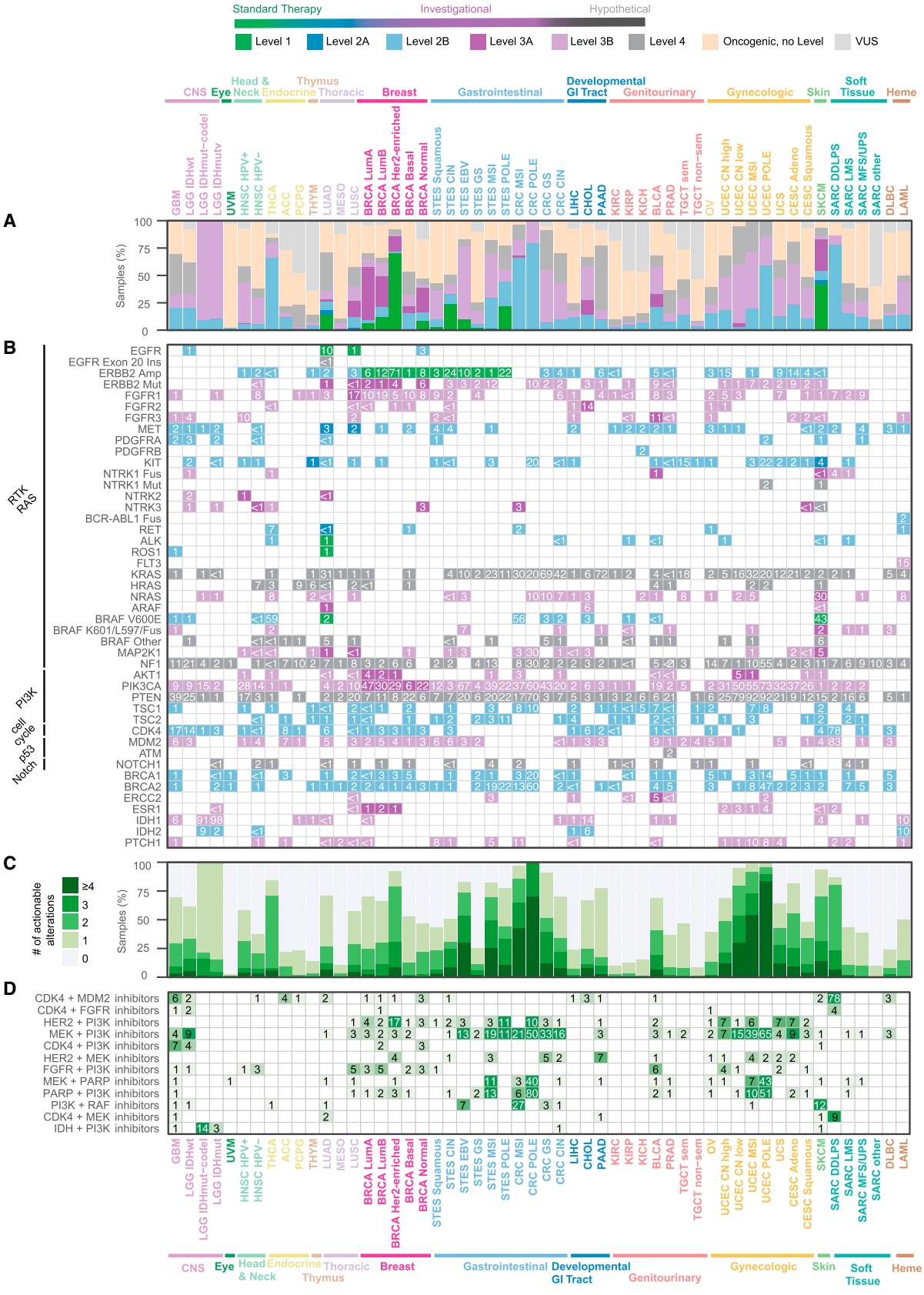
(D) Breakdown of the co-occurrence of gene alterations in the PI3K and Nrf2 pathways by tumor subtype. Green bars: percentage of samples with alterations in both PI3K and Nrf2 pathways. Pathways are sorted by decreasing percentage of samples with alterations in both pathways.

(E) Details of gene alterations in select genes (*PIK3CA*, *STK11*, *NFE2L2*, and *KEAP1*) within and between PI3K and Nrf2 pathways, with co-occurrence and mutual exclusivity between alterations. Samples are shown from left to right, and the number of samples in each group (bottom; note: the changing scale, indicated by solid and dashed lines).

(F) Pathway representation of the link between the PI3K and Nrf2 pathways.

(G) Dependencies between single alterations in the RTK and RAS/ERK pathways. Only the 22 alterations with at least one significant interaction (average sum correction, ASC score > 0.24; Mina et al., 2017) included.

(H) Breakdown of the interactions involving *EGFR* amplifications and mutations, corresponding to the bounding boxes in panel G. Left side: mutually exclusive interactions. Right side: co-occurring interactions.



(legend on next page)

number of alterations in a subset of tumor types are currently biomarkers for standard care targeted therapies, and a larger number are potential biomarkers for investigational therapies, some with promising clinical results. Using the OncoKB knowledge base of clinically actionable alterations (Chakravarty et al., 2017a), we systematically assessed all alterations in each sample of each cancer type, distinguishing between standard care actionability (Levels 1 or 2) and investigational therapies (Levels 3 and 4). Overall, 51% of tumors had at least one potentially actionable alteration in the ten signaling pathways, and 57% had at least one actionable alteration when including genes outside of these pathways, most notably BRCA1/2 and IDH1/2 (all numbers referenced below include these additional genes). Apart from the Her2-enriched breast cancer samples, most of which have a standard care targeted therapy, melanoma was the tumor type with the highest fraction of tumors with a Level 1 or 2A alteration (46%) (Figure 7A), mainly due to frequent *BRAF* mutations (Figure 7B), followed by esophagogastric cancers (*ERBB2* amplifications). Luminal A breast cancer was the tumor type with the highest frequency of biomarkers with promising investigational data (Level 3A), driven by the high prevalence of *PIK3CA*, *AKT1* and *ERBB2* mutations. Several tumor types had frequent mutations that are biomarkers for drug sensitivity in other cancer types (Level 3B), including endometrial cancer, where *PIK3CA* mutations are common. Uveal melanoma and testicular non-seminoma had the lowest percentage of potentially targetable samples (2.5% and 8.5%, respectively); thymoma, mesothelioma (MESO), and renal clear cell carcinoma (KIRC) also had low frequencies of potentially actionable alterations.

Thirty percent of tumor samples had two or more potentially targetable alterations (Figure 7C). Among these, the MSI-H and *POLE*-mutated tumor subtypes had the highest proportion of samples with multiple potentially actionable alterations (not considering the fact that patients with MSI-H tumors are now eligible for immunotherapy). Other tumor types with a high frequency of samples with multiple targetable alterations included non-hypermutated endometrial cancer (64%), colorectal cancer (37%), and breast cancer (28%).

Finally, we searched for candidate drug combinations that could prove effective across different tumor types based on the occurrence of actionable alterations. Hypermutant MSI and *POLE* subtypes had a high fraction of samples of actionable alterations corresponding to various drug combinations. In other tumor subtypes, a combination of CDK4 and MDM2 inhibitors was the most commonly indicated combination (1% total), in particular in dedifferentiated liposarcomas (SARC DDLPS), in which 78% of the cases had co-amplification of the two targets (Figure 7D). By a similar consideration linking actionable alterations of targets to their inhibitors, a combination of HER2 and PI3K inhibitors might be beneficial across multiple tumor types, in particular Her2-enriched breast cancer (17%), uterine carcinosarcoma (UCS, 7%), chromosomally unstable endometrioid carcinoma (UCEC CN high, 7%), and cervical adenocarcinoma (7%) (Figure 7D). Additional candidate combination therapies include PI3K and MEK inhibitors in EBV+ stomach tumors (10%), CDK4 and PI3K inhibitors in glioblastoma multiforme (7%), HER2 and MEK inhibitors in pancreatic cancer (7%), PI3K and RAF inhibitors in melanoma (SKCM, 12%), and IDH and PI3K inhibitors in IDH-mutant low grade glioma (14%) (Figure 7D). While there are many steps from the observation of combinations of genetic alterations to valid combination therapies, this survey indicates the wide landscape of potential tumor-type specific novel therapeutic combinations that can be explored in experimental and clinical contexts.

DISCUSSION

Signaling pathways are somatically altered in cancer at varying frequencies and in varying combinations across different organs and tissues, indicative of complex interplay and pathway crosstalk. Understanding the extent, detailed mechanisms, and co-occurrence of the oncogenic alterations in these pathways is critical for the development of new therapeutic approaches that can improve patient care.

Here we performed a comprehensive characterization of 10 selected signaling pathways across the 33 cancer types analyzed by TCGA. This report constitutes the first pan-cancer exploration that uses a uniformly processed dataset and a standardized set of pathway templates, curated through a combination of computational methods and expert review (Figures 1 and 2). The results highlight similarities and differences in frequencies of alteration of individual pathways in different cancer types and specific molecular subtypes (Figure 3). They also underscore the potential for discovering previously uncharacterized alterations in pathway genes that occur at low frequencies and might otherwise remain statistically unnoticeable (see *SOS1*, Figure 4). More generally, even though a small set of critical genes contains a very large fraction of alterations in these pathways (Figures 4 and 5), there is a complex interplay of co-occurring and mutually exclusive alterations within and across pathways (Figure 6). In spite of the accumulating wealth of biological knowledge and the accepted oncogenic relevance of these pathways, the number of currently approved biomarkers

Figure 7. Therapeutic Actionability and Drug Combinations

(A) Frequencies of clinical actionability by cancer subtype, broken down by level of evidence (Levels 1–4). Samples are classified by the alteration that carries the highest level of evidence. Tumor type–specific samples are analyzed by variants considered actionable, oncogenic but not actionable, or variants of unknown significance (VUS).

(B) Frequencies of actionable alterations per gene across cancer subtypes. For genes with different levels for different alterations, multiple rows are shown. Genes are grouped by pathway. Six additional genes not in the ten pathways (*BRCA1*, *BRCA2*, *ERCC2*, *IDH1*, *IDH2*, *ESR1*) are included and taken into account in the overall frequencies.

(C) Fraction of samples with a given number of actionable alterations per tumor type.

(D) Frequencies of possible drug combinations indicated by the co-alteration of actionable variants in each tumor type for the most frequent drug class combinations.

linked to standard of care therapies remains sparse (Figure 7), but additional drug targets in these pathways will hopefully emerge, and candidates for combination therapy will be explored.

This analysis of targetable alterations only included currently approved therapies or investigational therapies with reported promising results. These predominantly target the RTK-RAS, PI3K, cell-cycle, and p53 pathways. While some of these therapies are standard care, many are still investigational, and further testing is required to assess how effective different targeted therapies will be across tumor types and in tumors with different co-mutation spectra. Efforts are underway to develop therapies that target additional pathways, some of which are in clinical trials (Table S6) (Park and Guan, 2013), (Whitfield et al., 2017), (Whitfield et al., 2017), (Aster and Blacklow, 2012), (Takebe et al., 2014), (Buijs et al., 2012), (Sheen et al., 2013), (Pai et al., 2017). In the Wnt signaling pathway, for example, two approaches involve drugs targeting *PORCN*, which is involved in the processing of wingless proteins, and monoclonal antibodies directed at proteins in the Frizzled gene family. While the Nrf2 pathway does not have therapies directly targeting any of the pathway members included in this study, alterations in Nrf2 pathway members (*NFE2L2* and *KEAP1*) are used as part of the inclusion criteria in the Phase 2 trial of a TORC1/2 inhibitor. Clinical trials involving these pathways exemplify opportunities in precision medicine to associate additional functional alterations as part of inclusion criteria (Table S6). Not all apparently functional mutations, however, represent therapeutic targets, as illustrated, e.g., by the unusually large number of mutations in the MSI-H and POLE-mutated tumor subtypes, of which only a small fraction plausibly dominate oncogenesis. The observed co-occurrence patterns indicate a potential for combination therapies in some tumor types. The development of targeted combination therapies has been challenging for several reasons, including lack of safety data for combinations, the relatively slow pace of adoption of clinically approved multi-panel gene tests and of clinical trials testing combinations of multiple targeted therapies. However, there is a growing corpus of promising preclinical data indicating such combinations can be effective, such as the combination of MDM2 and CDK4 inhibitors (Laroche-Clary et al., 2017), and the combination of PI3K inhibitors and HER2 inhibitors in HER2-positive/*PIK3CA* mutant breast cancer patients, even when single gene-therapy approaches (e.g., PI3K monotherapy for *PIK3CA* mutant tumors) have thus far not had definitive clinical impact.

Although we cover a diverse range of oncogenic processes that spans most tissues and organ systems (Figures 1 and 3), we did not include some tumor types in the scope of this TCGA project, including most hematologic cancers. Furthermore, in spite of the relatively large set of samples, this effort is still underpowered to reliably discover tumor-type specific alterations that occur at very low frequencies; these will require further exploration using larger tumor-type specific sample sets.

The original aim and scope of TCGA was to genomically characterize primary, untreated tumors with a basic set of genetic alterations and transcript profiles. As the program is now completed, a future challenge is to expand these analyses to larger sample sets, additional data types, such as metabolite levels, a wider range of epigenetic states, post-translational modifications of proteins, and to investigate metastatic disease and genomic alterations that arise in post-treatment samples, as well as analyzing the role of a wider range of germline alterations and their interplay with somatic events. These new avenues of research will benefit from pathway-level analysis for which the templates and template curation pipelines presented here constitute a promising starting point. Similarly, as the catalog of clinically actionable alterations continues to grow, understanding intra- and inter-pathway dependencies, such as the ones considered here, will be crucial for the development of effective combination therapies that address or prevent resistance to initially successful single agent therapies.

The curated pathway templates and the uniformly processed dataset of alteration calls in 9,125 tumor samples are publicly available (Tables S3 and S4) and can be easily accessed through the PathwayMapper tool (http://pathwaymapper.org/), which allows alteration frequencies to be visually overlaid on the pathway templates; and, via the cBioPortal for Cancer Genomics (http://www.cbioportal.org/). This pathway landscape in The Cancer Genome Atlas is meant to provide a valuable resource for clinical oncologists, for cancer researchers and for a broad scientific community interested in cancer precision medicine.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Sample Selection and Exclusions
- METHOD DETAILS
  - Somatic mutation calling
  - Pathway Template Curation
  - Epigenetic silencing
  - Gene fusion detection and filtering
  - Generation of Genomic Alteration Matrices (GAMs)
  - Analysis of conditional selection between alterations
  - Pathway-level analysis of conditional selection
  - Curation of Clinical Trials
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and eight tables and can be found with this article online at https://doi.org/10.1016/j.cell.2018.03.035.

ACKNOWLEDGMENTS

This work was supported by NIH Grants U54 HG003273, U54 HG003067, U54 HG003079, U24 CA143799, U24 CA143835, U24 CA143840, U24 CA143843, U24 CA143845, U24 CA143848, U24 CA143858, U24 CA143866, U24 CA143867, U24 CA143882, U24 CA143883, U24 CA144025, P30 CA016672.

AUTHORS CONTRIBUTIONS

F.S.-V., M.M., J.A., W.K.C., A.L., E.M.V.A., A.D. C., G.C., C. S. and N. S. designed the study; F.S.-V., M.M., J.A., W.K.C., A.L., J.G., K.L., S.D., H.S.K., Z.H., A.O., B.G., J.G., H.Z., R.K., I.B., L.D., U.D., C.K., Q.G., M.H.B,W.-W.L., S.M.F., I.S., H.L., L.D., A.D.C., G.C., and N.S. collected and annotated the data; F.S.-V., M.M., J.A., W.K.C., A.L., D.C., W.Z., H.S., P.W.L., F.D., D.L.L., S.S., G.P.W., C.S.G., Y.X., C.W., A.I., A.H.B., T.G.B., A.J.L., G.D.H., T.G., L.N.K., G.M., C.H., A.D.T., M.J.F., F.M., M.M., E.V.A, A.D.C., G.C., C.S. and N.S. analyzed and interpreted the data. The Cancer Genome Atlas Research Network was responsible for the project administration. F.S.-V., M.M., J.A., W.K.C., A.L., E.M.V.A., A.D.C., G.C., C.S. and N.S. drafted the manuscript. All authors participated in editing or reviewing of the manuscript, and all authors approved the submitted manuscript.

DECLARATION OF INTERESTS

Michael Seiler, Peter G. Smith, Ping Zhu, Silvia Buonamici, and Lihua Yu are employees of H3 Biomedicine, Inc. Parts of this work are the subject of a patent application: WO2017040526 titled "Splice variants associated with neomorphic sf3b1 mutants." Shouyoung Peng, Anant A. Agrawal, James Palacino, and Teng Teng are employees of H3 Biomedicine, Inc. Andrew D. Cherniack, Ashton C. Berger, and Galen F. Gao receive research support from Bayer Pharmaceuticals. Gordon B. Mills serves on the External Scientific Review Board of Astrazeneca. Anil Sood is on the Scientific Advisory Board for Kiyatec and is a shareholder in BioPath. Jonathan S. Serody receives funding from Merck, Inc. Kyle R. Covington is an employee of Castle Biosciences, Inc. Preethi H. Gunaratne is founder, CSO, and shareholder of NextmiRNA Therapeutics. Christina Yau is a part-time employee/consultant at NantOmics. Franz X. Schaub is an employee and shareholder of SEngine Precision Medicine, Inc. Carla Grandori is an employee, founder, and shareholder of SEngine Precision Medicine, Inc. Robert N. Eisenman is a member of the Scientific Advisory Boards and shareholder of Shenogen Pharma and Kronos Bio. Daniel J. Weisenberger is a consultant for Zymo Research Corporation. Joshua M. Stuart is the founder of Five3 Genomics and shareholder of NantOmics. Marc T. Goodman receives research support from Merck, Inc. Andrew J. Gentles is a consultant for Cibermed. Charles M. Perou is an equity stock holder, consultant, and Board of Directors member of BioClassifier and GeneCentric Diagnostics and is also listed as an inventor on patent applications on the Breast PAM50 and Lung Cancer Subtyping assays. Matthew Meyerson receives research support from Bayer Pharmaceuticals; is an equity holder in, consultant for, and Scientific Advisory Board chair for OrigiMed; and is an inventor of a patent for EGFR mutation diagnosis in lung cancer, licensed to LabCorp. Eduard Porta-Pardo is an inventor of a patent for domainXplorer. Han Liang is a shareholder and scientific advisor of Precision Scientific and Eagle Nebula. Da Yang is an inventor on a pending patent application describing the use of antisense oligonucleotides against specific lncRNA sequence as diagnostic and therapeutic tools. Yonghong Xiao was an employee and shareholder of TESARO, Inc. Bin Feng is an employee and shareholder of TESARO, Inc. Carter Van Waes received research funding for the study of IAP inhibitor ASTX660 through a Cooperative Agreement between NIDCD, NIH, and Astex Pharmaceuticals. Raunaq Malhotra is an employee and shareholder of Seven Bridges, Inc. Peter W. Laird serves on the Scientific Advisory Board for AnchorDx. Joel Tepper is a consultant at EMD Serono. Kenneth Wang serves on the Advisory Board for Boston Scientific, Microtech, and Olympus. Andrea Califano is a founder, shareholder, and advisory board member of DarwinHealth, Inc. and a shareholder and advisory board member of Tempus, Inc. Toni K. Choueiri serves as needed on advisory boards for Bristol-Myers Squibb, Merck, and Roche. Lawrence Kwong receives research support from Array BioPharma. Sharon E. Plon is a member of the Scientific Advisory Board for Baylor Genetics Laboratory. Beth Y. Karlan serves on the Advisory Board of Invitae.

Received: November 17, 2017
Revised: February 28, 2018
Accepted: March 15, 2018
Published: April 5, 2018

REFERENCES

Alaei-Mahabadi, B., Bhadury, J., Karlsson, J.W., Nilsson, J.A., and Larsson, E. (2016). Global analysis of somatic structural genomic alterations and their impact on gene expression in diverse human cancers. Proc. Natl. Acad. Sci. USA *113*, 13768–13773.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., et al.; Australian Pancreatic Cancer Genome Initiative; ICGC Breast Cancer Consortium; ICGC MMML-Seq Consortium; ICGC PedBrain (2013). Signatures of mutational processes in human cancer. Nature *500*, 415–421.

Aster, J.C., and Blacklow, S.C. (2012). Targeting the Notch pathway: twists and turns on the road to rational therapeutics. J. Clin. Oncol. *30*, 2418–2420.

Bahceci, I., Dogrusoz, U., La, K.C., Babur, Ö., Gao, J., and Schultz, N. (2017). PathwayMapper: a collaborative visual web editor for cancer pathways and genomic data. Bioinformatics *33*, 2238–2240.

Bailey, M.H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Betrand, D., Weerasinghe, A., Colaprico, A., Kim, J., Wendl, M.C., Reardon, B., et al. (2018). Comprehensive Discovery and Characterization of Driver Genes and Mutations in Human Cancers. Cell *173*, this issue, 371–385.

Benelli, M., Pescucci, C., Marseglia, G., Severgnini, M., Torricelli, F., and Magi, A. (2012). Discovering chimeric transcripts in paired-end RNA-seq data by using EricScript. Bioinformatics *28*, 3232–3239.

Boland, C.R., and Goel, A. (2010). Microsatellite instability in colorectal cancer. Gastroenterology *138*, 2073–2087.e3.

Brennan, C.W., Verhaak, R.G.W., McKenna, A., Campos, B., Noushmehr, H., Salama, S.R., Zheng, S., Chakravarty, D., Sanborn, J.Z., Berman, S.H., et al.; TCGA Research Network (2013). The somatic genomic landscape of glioblastoma. Cell *155*, 462–477.

Buijs, J.T., Stayrook, K.R., and Guise, T.A. (2012). The role of TGF-β in bone metastasis: novel therapeutic perspectives. Bonekey Rep. *1*, 96.

Campbell, J.D., Alexandrov, A., Kim, J., Wala, J., Berger, A.H., Pedamallu, C.S., Shukla, S.A., Guo, G., Brooks, A.N., Murray, B.A., et al.; Cancer Genome Atlas Research Network (2016). Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat. Genet. *48*, 607–616.

Cancer Genome Atlas Network (2012a). Comprehensive molecular characterization of human colon and rectal cancer. Nature *487*, 330–337.

Cancer Genome Atlas Network (2012b). Comprehensive molecular portraits of human breast tumours. Nature *490*, 61–70.

Cancer Genome Atlas Network (2015a). Comprehensive genomic characterization of head and neck squamous cell carcinomas. Nature *517*, 576–582.

Cancer Genome Atlas Network (2015b). Genomic Classification of Cutaneous Melanoma. Cell *161*, 1681–1696.

Cancer Genome Atlas Research Network (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature *455*, 1061–1068.

Cancer Genome Atlas Research Network (2011). Integrated genomic analyses of ovarian carcinoma. Nature *474*, 609–615.

Cancer Genome Atlas Research Network (2012). Comprehensive genomic characterization of squamous cell lung cancers. Nature *489*, 519–525.

Cancer Genome Atlas Research Network (2013a). Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature *499*, 43–49.

Cancer Genome Atlas Research Network, Kandoth, C., Schultz, N., Cherniack, A.D., Akbani, R., Liu, Y., Shen, H., Robertson, A.G., Pashtan, I., Shen, R., Benz, C.C., et al. (2013b). Integrated genomic characterization of endometrial carcinoma. Nature *497*, 67–73.

Cancer Genome Atlas Research Network,, Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013c). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.

Cancer Genome Atlas Research Network (2014a). Comprehensive molecular characterization of urothelial bladder carcinoma. Nature *507*, 315–322.

Cancer Genome Atlas Research Network (2014b). Comprehensive molecular profiling of lung adenocarcinoma. Nature *511*, 543–550.

Cancer Genome Atlas Research Network (2014c). Comprehensive molecular characterization of gastric adenocarcinoma. Nature *513*, 202–209.

Cancer Genome Atlas Research Network (2014d). Integrated genomic characterization of papillary thyroid carcinoma. Cell *159*, 676–690.

Cancer Genome Atlas Research Network (2017a). Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma. Cell *169*, 1327–1341.e23.

Cancer Genome Atlas Research Network; Analysis Working Group: Asan University; BC Cancer Agency; Brigham and Women's Hospital; Broad Institute; Brown University; Case Western Reserve University; Dana-Farber Cancer Institute; Duke University; Greater Poland Cancer Centre; Harvard Medical School; et-al (2017b). Integrated genomic characterization of oesophageal carcinoma. Nature *541*, 169–175.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. *2*, 401–404.

Cerami, E.G., Gross, B.E., Demir, E., Rodchenkov, I., Babur, O., Anwar, N., Schultz, N., Bader, G.D., and Sander, C. (2011). Pathway Commons, a web resource for biological pathway data. Nucleic Acids Res. *39*, D685–D690.

Chakravarty, D., Gao, J., Phillips, S.M., Kundra, R., Zhang, H., Wang, J., Rudolph, J.E., Yaeger, R., Soumerai, T., Nissan, M.H., et al. (2017a). OncoKB: A Precision Oncology Knowledge Base. JCO Precis. Oncol. 2017.

Chakravarty, D., Pedraza, A.M., Cotari, J., Liu, A.H., Punko, D., Kokroo, A., Huse, J.T., Altan-Bonnet, G., and Brennan, C.W. (2017b). EGFR and PDGFRA co-expression and heterodimerization in glioblastoma tumor sphere lines. Sci. Rep. *7*, 9043.

Chang, M.T., Asthana, S., Gao, S.P., Lee, B.H., Chapman, J.S., Kandoth, C., Gao, J., Socci, N.D., Solit, D.B., Olshen, A.B., et al. (2016). Identifying recurrent mutations in cancer reveals widespread lineage diversity and mutational specificity. Nat. Biotechnol. *34*, 155–163.

Chang, M.T., Bhattarai, T.S., Schram, A.M., Bielski, C.M., Donoghue, M.T.A., Jonsson, P., Chakravarty, D., Phillips, S., Kandoth, C., Penson, A., et al. (2018). Accelerating Discovery of Functional Mutant Alleles in Cancer. Cancer Discov. *8*, 174–183.

Ciriello, G., Miller, M.L., Aksoy, B.A., Senbabaoglu, Y., Schultz, N., and Sander, C. (2013). Emerging landscape of oncogenic signatures across human cancers. Nat. Genet. *45*, 1127–1133.

Davis, C.F., Ricketts, C.J., Wang, M., Yang, L., Cherniack, A.D., Shen, H., Buhay, C., Kang, H., Kim, S.C., Fahey, C.C., et al.; The Cancer Genome Atlas Research Network (2014). The somatic genomic landscape of chromophobe renal cell carcinoma. Cancer Cell *26*, 319–330.

Ellrott, K., Bailey, M.H., Saksena, G., Covington, K.R., Kandoth, C., Stewart, C., Hess, J., Ma, S., McLellan, M., Sofia, H.J., et al. (2018). Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines. Cell Syst *6*. https://doi.org/10.1016/j.cels.2018.03.002.

Etemadmoghadam, D., Weir, B.A., Au-Yeung, G., Alsop, K., Mitchell, G., George, J., Davis, S., D'Andrea, A.D., Simpson, K., Hahn, W.C., and Bowtell, D.D.; Australian Ovarian Cancer Study Group (2013). Synthetic lethality between CCNE1 amplification and loss of BRCA1. Proc. Natl. Acad. Sci. USA *110*, 19489–19494.

FANTOM Consortium and the RIKEN PMI and CLST (DGT), Forrest, A.R., Kawaji, H., Rehli, M., Baillie, J.K., de Hoon, M.J., Haberle, V., Lassmann, T., Kulakovskiy, I.V., Lizio, M., Itoh, M., et al. (2014). A promoter-level mammalian expression atlas. Nature *507*, 462–470.

Gao, J., Ciriello, G., Sander, C., and Schultz, N. (2014). Collection, integration and analysis of cancer genomic profiles: from data to insight. Curr. Opin. Genet. Dev. *24*, 92–98.

Gao, J., Chang, M.T., Johnsen, H.C., Gao, S.P., Sylvester, B.E., Sumer, S.O., Zhang, H., Solit, D.B., Taylor, B.S., Schultz, N., and Sander, C. (2017). 3D clusters of somatic mutations in cancer reveal numerous rare mutations as functional targets. Genome Med. *9*, 4.

Gao, Q., Liang, W.-W., Foltz, S.M., Mutharasu, G., Jayasinghe, R.G., Cao, S., Liao, W.-W., Reynolds, S.M., Wyczalkowski, M.A., Yao, L., et al. (2018). Driver Fusions and Their Implications in the Development and Treatment of Human Cancers. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.03.050.

Garraway, L.A., and Lander, E.S. (2013). Lessons from the cancer genome. Cell *153*, 17–37.

Ge, Z., Leighton, J., Wang, Y., Peng, X., Chen, Z., Chen, H., Sun, Y., Yao, F., Li, J., Zhang, H., et al. (2018). Integrated Genomic Analysis of the Ubiquitin Pathway across Cancer Types. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.03.047.

Hanahan, D., and Weinberg, R.A. (2000). The hallmarks of cancer. Cell *100*, 57–70.

Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. Cell *144*, 646–674.

Hartmaier, R.J., Albacker, L.A., Chmielecki, J., Bailey, M., He, J., Goldberg, M.E., Ramkissoon, S., Suh, J., Elvin, J.A., Chiacchia, S., et al. (2017). High-Throughput Genomic Profiling of Adult Solid Tumors Reveals Novel Insights into Cancer Pathogenesis. Cancer Res. *77*, 2464–2475.

Hoadley, K.A., Yau, C., Wolf, D.M., Cherniack, A.D., Tamborero, D., Ng, S., Leiserson, M.D.M., Niu, B., McLellan, M.D., Uzunangelov, V., et al.; Cancer Genome Atlas Research Network (2014). Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. Cell *158*, 929–944.

Hoadley, K.A., Yau, C., Hinoue, T., Wolf, M.D., Lazar, J.A., Drill, E., Shen, R., Taylor, M.A., Cherniack, D.A., Akbani, R., et al. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell *173*. https://doi.org/10.1016/j.cell.2018.03.022.

Jayasinghe, R.G., Cao, S., Gao, Q., Wendl, M.C., Vo, N.S., Reynolds, S.M., Zhao, Y., Climente-González, H., Chai, S., Wang, F., et al. (2018). Systematic analysis of splice site-creating mutations in cancer. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.03.052.

Joshi-Tope, G., Gillespie, M., Vastrik, I., D'Eustachio, P., Schmidt, E., de Bono, B., Jassal, B., Gopinath, G.R., Wu, G.R., Matthews, L., et al. (2005). Reactome: a knowledgebase of biological pathways. Nucleic Acids Res. *33*, D428–D432.

Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. *28*, 27–30.

Korkut, A., Zaidi, S., Kanchi, R., Berger, A., Robertson, G., Kwong, L., Datto, M., Roszik, J., Ling, S., Ravikumar, V., et al. (2017). A PanCancer molecular analysis of disruptions in the TGF-β pathway. Cancer Cell.

Laroche-Clary, A., Chaire, V., Algeo, M.-P., Derieppe, M.-A., Loarer, F.L., and Italiano, A. (2017). Combined targeting of MDM2 and CDK4 is synergistic in dedifferentiated liposarcomas. J. Hematol. Oncol. *10*, 123.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature *499*, 214–218.

Lawrence, M.S., Stojanov, P., Mermel, C.H., Robinson, J.T., Garraway, L.A., Golub, T.R., Meyerson, M., Gabriel, S.B., Lander, E.S., and Getz, G. (2014). Discovery and saturation analysis of cancer genes across 21 tumour types. Nature *505*, 495–501.

Lepri, F., De Luca, A., Stella, L., Rossi, C., Baldassarre, G., Pantaleoni, F., Cordeddu, V., Williams, B.J., Dentici, M.L., Caputo, V., et al. (2011). SOS1 mutations in Noonan syndrome: molecular spectrum, structural insights on pathogenic effects, and genotype-phenotype correlations. Hum. Mutat. *32*, 760–772.

Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhim, R., and Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. *12*, R41.

Mina, M., Raynaud, F., Tavernari, D., Battistello, E., Sungalee, S., Saghafinia, S., Laessle, T., Sanchez-Vega, F., Schultz, N., Oricchio, E., and Ciriello, G. (2017). Conditional Selection of Genomic Alterations Dictates Cancer Evolution and Oncogenic Dependencies. Cancer Cell *32*, 155–168.e6.

Mitsuishi, Y., Taguchi, K., Kawatani, Y., Shibata, T., Nukiwa, T., Aburatani, H., Yamamoto, M., and Motohashi, H. (2012). Nrf2 redirects glucose and glutamine into anabolic pathways in metabolic reprogramming. Cancer Cell *22*, 66–79.

Nissan, M.H., Pratilas, C.A., Jones, A.M., Ramirez, R., Won, H., Liu, C., Tiwari, S., Kong, L., Hanrahan, A.J., Yao, Z., et al. (2014). Loss of NF1 in cutaneous melanoma is associated with RAS activation and MEK dependence. Cancer Res. *74*, 2340–2350.

Pai, S.G., Carneiro, B.A., Mota, J.M., Costa, R., Leite, C.A., Barroso-Sousa, R., Kaplan, J.B., Chae, Y.K., and Giles, F.J. (2017). Wnt/beta-catenin pathway: modulating anticancer immune response. J. Hematol. Oncol. *10*, 101.

Park, H.W., and Guan, K.-L. (2013). Regulation of the Hippo pathway and implications for anticancer drug development. Trends Pharmacol. Sci. *34*, 581–589.

Peng, X., Chen, Z., Farshidfar, F., Xu, X., Lorenzi, P., Wang, Y., Cheng, F., Tan, L., Mojumdar, K., Du, D., et al. (2018). Molecular Characterization and Clinical Relevance of Metabolic Expression Subtypes in Human Cancers. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.03.077.

Rayner, E., van Gool, I.C., Palles, C., Kearsey, S.E., Bosse, T., Tomlinson, I., and Church, D.N. (2016). A panoply of errors: polymerase proofreading domain mutations in cancer. Nat. Rev. Cancer *16*, 71–81.

Schaub, F.X., Dhankani, V., Berger, A.C., Trivedi, M., Richardson, A.B., Shaw, R., Zhao, W., Zhang, X., Ventura, A., Liu, Y., et al. (2018). Alterations of the MYC oncogene and its Proximal Network across The Cancer Genome Atlas. Cell Syst *6*. https://doi.org/10.1016/j.cels.2018.03.003.

Schram, A.M., Berger, M.F., and Hyman, D.M. (2017). Precision oncology: Charting a path forward to broader deployment of genomic profiling. PLoS Med. *14*, e1002242.

Seiler, M., Peng, S., Agrawal, A.A., Palacino, J., Teng, T., Zhu, P., Smith, P.G., Cancer Genome Atlas Research Network, The, Buonamici, S., Yu, L., et al. (2018). Survey of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.01.088.

Sheen, Y.Y., Kim, M.-J., Park, S.-A., Park, S.-Y., and Nam, J.-S. (2013). Targeting the Transforming Growth Factor-β Signaling in Cancer Therapy. Biomol. Ther. (Seoul) *21*, 323–331.

Sholl, L.M., Do, K., Shivdasani, P., Cerami, E., Dubuc, A.M., Kuo, F.C., Garcia, E.P., Jia, Y., Davineni, P., Abo, R.P., et al. (2016). Institutional implementation of clinical tumor profiling on an unselected cancer population. JCI Insight *1*, e87062.

Sun, C., Wang, L., Huang, S., Heynen, G.J.J.E., Prahallad, A., Robert, C., Haanen, J., Blank, C., Wesseling, J., Willems, S.M., et al. (2014). Reversible and adaptive resistance to BRAF(V600E) inhibition in melanoma. Nature *508*, 118–122.

Takebe, N., Nguyen, D., and Yang, S.X. (2014). Targeting notch signaling pathway in cancer: clinical development advances and challenges. Pharmacol. Ther. *141*, 140–149.

Knijnenburg, T., Wang, L., Zimmermann, M., Chambwe, N., Gao, G., Cherniack, A., Fan, H., Shen, H., Way, G., Greene, C., et al. (2018). Genomic and Molecular Landscape of DNA Damage Repair Deficiency Across The Cancer Genome Atlas. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.03.076.

Triche, T.J., Jr., Weisenberger, D.J., Van Den Berg, D., Laird, P.W., and Siegmund, K.D. (2013). Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. *41*, e90.

Unni, A.M., Lockwood, W.W., Zejnullahu, K., Lee-Lin, S.-Q., and Varmus, H. (2015). Evidence that synthetic lethality underlies the mutual exclusivity of oncogenic KRAS and EGFR mutations in lung adenocarcinoma. eLife *4*, e06907.

Vogelstein, B., and Kinzler, K.W. (2004). Cancer genes and the pathways they control. Nat. Med. *10*, 789–799.

Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr., and Kinzler, K.W. (2013). Cancer genome landscapes. Science *339*, 1546–1558.

Wang, Z., Yang, B., Zhang, M., Guo, W., Wu, Z., Wang, Y., Jia, L., Li, S., The Cancer Genome Atlas Research Network, Xie, W., and Yang, D. (2018). LncRNA epigenetic landscape analysis identifies EPIC1 as an oncogenic lncRNA that interacts with MYC and promotes cell cycle progression in cancer. Cancer Cell *33*. https://doi.org/10.1016/j.ccell.2018.03.006.

Way, G.P., Sanchez-Vega, F., La, K., Armenia, J., Chatila, W.K., Luna, A., Sander, C., Cherniack, A.D., Mina, M., Ciriello, G., et al. (2018). Pan-cancer Ras Pathway Activation in The Cancer Genome Atlas. Cell Rep *23*. https://doi.org/10.1016/j.celrep.2018.03.046.

Whitfield, J.R., Beaulieu, M.-E., and Soucek, L. (2017). Strategies to Inhibit Myc and Their Clinical Applicability. Front. Cell Dev. Biol. *5*, 10.

Zehir, A., Benayed, R., Shah, R.H., Syed, A., Middha, S., Kim, H.R., Srinivasan, P., Gao, J., Chakravarty, D., Devlin, S.M., et al. (2017). Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat. Med. *23*, 703–713.

Zhang, Y., Kwok-Shing Ng, P., Kucherlapati, M., Chen, F., Liu, Y., Tsang, Y.H., de Velasco, G., Jeong, K.J., Akbani, R., Hadjipanayis, A., et al. (2017). A Pan-Cancer Proteogenomic Atlas of PI3K/AKT/mTOR Pathway Alterations. Cancer Cell *31*, 820–832.e3.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited Data		
Raw and processed clinical, array and sequence data.	NCI Genomic Data Commons	https://portal.gdc.cancer.gov/ https://gdc.cancer.gov/about-data/publications/pancanatlas cBioPortal: http://www.cbioportal.org
Digital Pathology Images	Cancer Digital Slide Archive	http://cancer.digitalslidearchive.net/
Software and Algorithms		
PathwayMapper	(Bahceci et al., 2017)	http://www.pathwaymapper.org/
SELECT	(Mina et al., 2017)	http://ciriellolab.org/select/select.html
GISTIC 2.0	(Mermel et al., 2011)	http://archive.broadinstitute.org/cancer/cga/gistic
MutSigCV	(Lawrence et al., 2014)	http://software.broadinstitute.org/cancer/software/genepattern/modules/docs/MutSigCV
STAR-Fusion	Hass et al., bioRxiv https://doi.org/10.1101/120295	https://github.com/STAR-Fusion/STAR-Fusion/wiki
Breakfast	See link	https://github.com/annalam/breakfast
EricScript	(Benelli et al., 2012)	https://sites.google.com/site/bioericscript/
RESET	(Saghafinia, Mina et al., manuscript under review)	http://ciriellolab.org/
Other		
OncoKB	(Chakravarty et al., 2017a)	www.oncokb.org
CancerHotspots	(Chang et al., 2016)	www.cancerhotspots.org
3D Hotspots	(Gao et al., 2017)	www.3dhotspots.org
cBioPortal	Cerami et al., 2012	http://www.cbioportal.org/
TCGA Batch Effects	The University of Texas MD Anderson Cancer Center	http://bioinformatics.mdanderson.org/tcgambatch/
Pathway Commons	(Cerami et al., 2011)	http://www.pathwaycommons.org/

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Nikolaus Schultz (schultz@cbio.mskcc.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

TCGA Project Management has collected necessary human subjects documentation to ensure the project complies with 45-CFR-46 (the ''Common Rule''). The program has obtained documentation from every contributing clinical site to verify that IRB approval has been obtained to participate in TCGA. Such documented approval may include one or more of the following:

- An IRB-approved protocol with Informed Consent specific to TCGA or a substantially similar program. In the latter case, if the protocol was not TCGA-specific, the clinical site PI provided a further finding from the IRB that the already-approved protocol is sufficient to participate in TCGA.
- A TCGA-specific IRB waiver has been granted.
- A TCGA-specific letter that the IRB considers one of the exemptions in 45-CFR-46 applicable. The two most common exemptions cited were that the research falls under 46.102(f)(2) or 46.101(b)(4). Both exempt requirements for informed consent, because the received data and material do not contain directly identifiable private information.
- A TCGA-specific letter that the IRB does not consider the use of these data and materials to be human subjects research. This was most common for collections in which the donors were deceased.

Sample Selection and Exclusions

We started from the set of 11,276 patients that were included in the final whitelist for the TCGA PanCanAtlas project. We only used samples that had available data across these four genomic platforms: mutations, copy number, DNA methylation and mRNA expression. Our analyses excluded certain molecular platforms that have previously been used in TCGA, such as protein levels from reverse-phase protein arrays (RPPA), microRNA, and lncRNA, as their inclusion would have implied a sharp decrease in the total number of samples with data available across all platforms. Additionally, we excluded samples that had been flagged during pathology review by an expert committee or due to quality control (QC) issues identified by the individual tumor-type or PanCanAtlas analysis working groups. After these filters had been applied, a total of 9,125 patients were used. Samples consisted of primary solid tumors for a large majority of these patients (8602/9125, 94%), plus a small number of blood tumors corresponding to the AML subset (162/9125, 2%) and a small subset of metastatic samples from melanoma patients (361/9125, 4%).

METHOD DETAILS

Somatic mutation calling

We used version 2.8 of the mutation annotation format (MAF) file provided by the MC3 ("Multi-Center Mutation Calling in Multiple Cancers") group within the TCGA Network (Ellrott et al., 2018). The mutation data can be found here (https://gdc.cancer.gov/about-data/publications/mc3-2017). We augmented this file in collaboration with the MC3 group and included all validated mutation calls from the original AML publication. The final MAF that was used for our analyses, including OncoKB annotations, can be downloaded from the TCGA PanCancer Atlas publication page (https://gdc.cancer.gov/about-data/publications/pancanatlas).

Pathway Template Curation

We manually curated the gene annotation of the ten selected pathways using the following workflow.

Selection and classification of genes in pathways

Genes were assigned to pathways based on a combined revision of pathway analyses in previous TCGA marker papers published between 2008 and 2017, a review of the scientific literature (including but not limited to the references in Table S2) and expert curation. We applied two different kinds of expert curation. 1) several of the pathways, such as TGF-Beta, Myc and PI3K, had specific analysis working groups. These groups were led by experts in each pathway and published separate manuscripts (Ge et al., 2018; Korkut et al., 2017; Peng et al., 2018; Schaub et al., 2018; Wang et al., 2018; Way and Greene, 2017). 2) for some of the pathways, we consulted experts from outside of TCGA in order to validate or improve our curated pathway templates (e.g., Frank McCormick for RTK-RAS or Mitchell Frederick for Notch). After the lists of pathway members were finalized, each gene was annotated as Tumor Suppressor (TSG) or Oncogene (OG) using OncoKB and prior knowledge from the scientific literature. The final gene lists that were selected for each pathway are provided in Table S3.

Identification of mutational hotspots

The cancer hotspots algorithm that we used identifies recurrent alterations based on a cohort of 24,592 tumor samples (Chang et al., 2016, 2018). Identification of 3D hotspots was based on recurrence of mutations in the context of spatial neighborhoods in protein structures (Gao et al., 2017).

Annotation of functionally relevant mutations

We used information about oncogenic and clinically actionable mutations from the OncoKB database (Chakravarty et al., 2017a), which provides information on variants in more than 400 genes. For template curation, we used OncoKB to filter out putative passenger mutations and copy number changes, by discarding somatic alterations that were not labeled as *oncogenic*, *likely oncogenic* or *predicted oncogenic* in the database. For the analysis of therapeutic implications, we used annotations about different levels of clinical actionability as described in the text. These had originally been compiled and curated by OncoKB by combining a diverse set of sources, including FDA-, NCCN- and other guidelines, ClinicalTrials.gov and the scientific literature.

Annotation of functionally relevant CNVs

We applied a two step procedure to determine whether the annotated genes were functionally amplified or deleted in each specific sample. First, a collection of functional relevant amplifications and deletions was curated by integrating the GISTIC 2.0 analysis of the PancanAtlas dataset and the OncoKB database. GISTIC was run separately on each individual tumor type, and then globally on the entire PanCanAtlas dataset, yielding a list of recurrently amplified and deleted regions of interest (ROIs). Default parameters of GISTIC 2.0 were used, with the confidence level set to 0.95. For Genes within ROIs, copy number variants consistent with the role of the gene (amplification of OGs and deletions of TSGs) were retained. Gene-specific copy number variants that were labeled as oncogenic, likely oncogenic or predicted oncogenic in OncoKB were also retained, yielding a list of gene-level functional CNVs. As an additional validation step, we individually inspected each of the gene level calls to ensure that there was a good correlation between copy-number status and gene expression, and we excluded calls in genes for which this correlation was non-existent. Thresholded gene-level amplification/deletion values produced by GISTIC were used for pathway analysis, considering only amplifications (+2) and deep deletions (−2). In total, 7,532 gene amplifications and 5,602 deletions were selected, for a total of 13,134 occurrences.

Epigenetic silencing

Curated analysis of CDKN2A promoter hypermethylation

CDKN2A promoter methylation was assessed using Illumina Infinium HumanMethylation450 probe cg13601799 located within Exon 1a of *CDKN2A* ($p16^{INK4a}$). We described the selection of this probe for *CDKN2A* methylation calling in a prior report (Cancer Genome Atlas Research Network, 2012). We introduced a further refinement of DNA methylation calling to avoid artifactual hypermethylation calls due to deep deletion in a gene. In brief, we used Level 1 IDAT files to calculate out-of-band ('oob') probe intensities as a surrogate for background intensity, superior to internal negative controls (Triche et al., 2013). cg13601799 is a Type I probe with both methylated (M) and unmethylated (U) versions in the red color channel, and therefore dye bias is not a concern. We compared the foreground intensities (M and U) to the empirical distribution of the background intensities (as measured by the 'oob' probes). We first called a sample to be methylated when the methylated (M) signal was higher than the 95th percentile of the background ('oob') probes (FDR = 5%). As this locus is unmethylated in normal tissues, the U signals are generally higher than the M signal due to the presence of contaminating normal cell types. We required a Log2(foreground/background) log-ratio of 2 or greater for the U probe to ensure that the U signal was derived from tumor cells and not from contaminating normal cells in the case of a tumor with CDKN2A deletion. If Log2(foreground/background) was < 2 for U and < 0 for M for this probe, then we concluded that this locus was deleted in the tumor cells, and we then denoted these cases as containing "no signal" (Table S4). We identified 681 such samples, and we had GISTIC copy number change data for 627 out of these 681. Out of these 627, 471 were called to have high-level deletion for CDKN2A (−2 in GISTIC calls) and 120 had low level deletion for this gene (−1 in GISTIC calls), validating this approach.

Analysis of DNA hypermethylation at the promoters of other tumor suppressor genes

Epigenetic DNA hypermethylation events at promoters of tumor suppressor genes that are associated with decreased gene expression were systematically identified using the RESET bioinformatic tool (Saghafinia, Mina et al. manuscript in preparation). RESET extracts probes that (i) map to gene promoter regions, (ii) are significantly hypermethylated compared to normal tissue samples, and (iii) are associated with lower transcript levels of the corresponding gene. More specifically, only probes overlapping gene promoter regions extracted from the FANTOM5 cohort of robust promoters are considered (FANTOM Consortium and the RIKEN PMI and CLST (DGT) et al., 2014). The status of a probe (dichotomized in hypermethylated and not hypermethylated) is determined by comparing its beta value to the beta value distribution from adjacent normal tissue samples available in the TCGA sample collection. Finally, RESET determines whether a hypermethylation event is associated with mRNA downregulation by checking whether the mRNA expression of the associated gene is significantly decreased in hypermethylated tumors, compared to the not hypermethylated ones. To avoid biases due to intrinsic gene expression and methylation differences between tumor samples of different origins, we separately applied RESET within each tumor type. For tumor types without normal adjacent samples, the entire set of normal samples from the TCGA cohort was used to define the background beta value distribution.

In this study, we evaluated all tumor suppressors in ten pathway templates (Table S3). We considered as significant only silencing events with a false discovery rate FDR < 10% and a RESET score > 1. The results were further manually curated to exclude cases where the methylation event might be tissue-associated, leading to a list of 15 genes silenced by DNA methylation (Table S7) Consistently with the procedure used for copy number calls, all hypermethylation occurrences for these 15 genes in all tumor samples were retained, even if the silencing event was only significantly recurrent in a subset of tumor types. This pancan set of occurrences was further filtered to increase the likelihood of functional relevance: only the hypermethylation occurrences with a gene expression lower than the 25 percentile of the gene expression distribution from the unmethylated samples were retained as functional and considered in the downstream analyses. The sample-specific epigenetic silencing calls are provided as part of the genomic alteration matrix described below (Table S4).

Gene fusion detection and filtering

TCGA RNA-Seq data were downloaded from Cancer Genomics Hub (CGHub, http://cghub.ucsc.edu) and analyzed using Google cloud. For each sample, the fastq file was mapped to the human genome (build 38) followed by fusion calling using STAR-Fusion (parameters:–annotation –coding-effect), EricScript (default parameters) and BREAKFAST (two different minimum distance cutoffs were used: 5 kb and 100 kb). STAR-Fusion showed higher sensitivity in detecting the fusions reported in previous TCGA studies. Therefore, we focused on the STAR-Fusion output and integrated EricScript and BREAKFAST output in one of the following filtering steps: 1) an exclusion list of genes was curated, including uncharacterized genes, immunoglobin genes, mitochondrial genes, etc. Fusions involving these genes were filtered; 2) Fusions from the same gene or paralog genes (downloaded from https://github.com/STAR-Fusion/STAR-Fusion_benchmarking_data/tree/master/resources) were filtered; 3) Fusions reported in normal samples were filtered, including the ones from TCGA normal samples, GTEx tissues (reported in STAR-Fusion output), and non-cancer cell study (Babiceanu et al.); 4) For the fusions reported by only STAR-Fusion but not EricScript, a minimum value of FFPM (fusion fragments per million total reads) was required, as suggested by the author; For the fusions reported by both callers, no requirement. 5) Finally, fusions with exactly the same breakpoints in ≥ 10 samples across different cancer types were removed unless they were reported in previous TCGA studies (e.g., FGFR3-TACC3).

For our pathway analyses, we included only the fusions that (a) involved at least one gene labeled as TSG in one of our pathway templates, or (b) involved at least one gene labeled as OG in one of our pathway templates and such that the fusion is labeled as *oncogenic*, *likely oncogenic* or *predicted oncogenic* in OncoKB. We also included a small set of additional fusions (MAML3-UBTF, NOTCH2-SEC22B and PIK3CA-TBL1XR) based on recent evidence from the literature. Any fusion failing to satisfy at least

one of these requirements was excluded from subsequent pathway analyses (although some additional fusions that are clinically actionable based on OncoKB where included in Figure 7 for completeness). The final set of all fusion calls used in our manuscript is provided as Table S8.

Generation of Genomic Alteration Matrices (GAMs)

To integrate all the genomic data in a format readily usable in the downstream analyses, the complexity of mutation and CNV data was summarized into a binary Genomic Alteration Matrix (GAM) representing the occurrence of gene alterations across samples, provided as Table S4. This matrix includes the set of functionally relevant mutations and CNVs selected for each gene and summarized in the onco-query language column provided as part of each pathway template in Table S3. In the alteration level version of this matrix, copy number events and point mutation events affecting the same gene were kept distinct. We also included epigenetic silencing of CDKN2A based on DNA methylation analysis of the gene promoter and the epigenetic silencing of 15 additional genes uncovered by RESET. The resulting table has entries for 9,125 samples and 411 alterations, for a total of 33,324 occurrences. For completeness, in Table S4 we also provide a version of the GAM where alterations are aggregated at the gene level and a third version were alterations are aggregated at the pathway level for the ten pathways in our analysis.

Analysis of conditional selection between alterations

SELECT, a method that infers conditional selection dependencies between alterations from occurrence patterns (Mina et al., 2017), was run on the PancanPathway GAM. The default parameters of the R package implementation were used, with 5,000 random permutations. SELECT analysis was performed at alteration level, considering as separate features the point mutations, copy number changes, silencing and fusion events affecting the same gene. Alteration type, tumor type and tumor subtype were used as covariates in the analysis. Only alterations with more than 5 occurrences were considered (0.05% of the samples). In total, SELECT produced a list of 273 high-scoring motifs between 315 alterations.

Pathway-level analysis of conditional selection

The dependency motifs were summarized at pathway level by considering independently (i) the sum of motif scores between each pair of pathways, and (ii) the number of significant motifs. The significance of pathway-level interactions was empirically estimated by comparing the observed sum of motif scores and number of significant motifs to the null distribution obtained by randomly permuting the pathway annotation of the genes. The two metrics were first tested independently, and the two P values were then combined using Stouffer's method. Combined P values were then corrected with the Benjamini-Hochberg method. Corrected P values smaller than 0.25 were deemed to be significant.

Curation of Clinical Trials

The list of clinical trials for genes in pathways not represented in OncoKB was manually curated from ClinicalTrials.gov (http://clinicaltrials.gov). Clinical trials with drug compounds targeting pathway members or which described pathway members in their inclusion or exclusion criteria are reported. Focus was given to ongoing clinical trials. A description of data retrieved from particular clinical trials is in the README worksheet of Table S6. If available, PubChem Compound IDs (https://pubchem.ncbi.nlm.nih.gov/) are given for drug compounds.

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantitative and statistical methods are described above within the context of individual analyses in the Method Details section.

DATA AND SOFTWARE AVAILABILITY

The raw data, processed data and clinical data can be found at the legacy archive of the GDC (https://portal.gdc.cancer.gov/legacy-archive/search/f) and the TCGA PanCancer Atlas publication page (https://gdc.cancer.gov/about-data/publications/pancanatlas). The mutation data can be found here (https://gdc.cancer.gov/about-data/publications/mc3-2017).

Data can also be visualized and downloaded using a dedicated section of the cBioPortal for Cancer Genomics (http://www.cbioportal.org/).

ADDITIONAL RESOURCES

Pathway diagrams were curated using PathwayMapper (Bahceci et al., 2017), a tool that allows visualization and design of pathway diagrams stylized as in classical TCGA publications. This tool is publicly available online at www.pathwaymapper.org. Our curated templates provided in Table S3 are accessible as pre-defined pathway diagrams that have been incorporated to the PahtwayMapper interface. PathwayMapper also acts as an interactive resource that allows to easily overlay user-inputed alteration frequencies on top of these predefined diagrams.

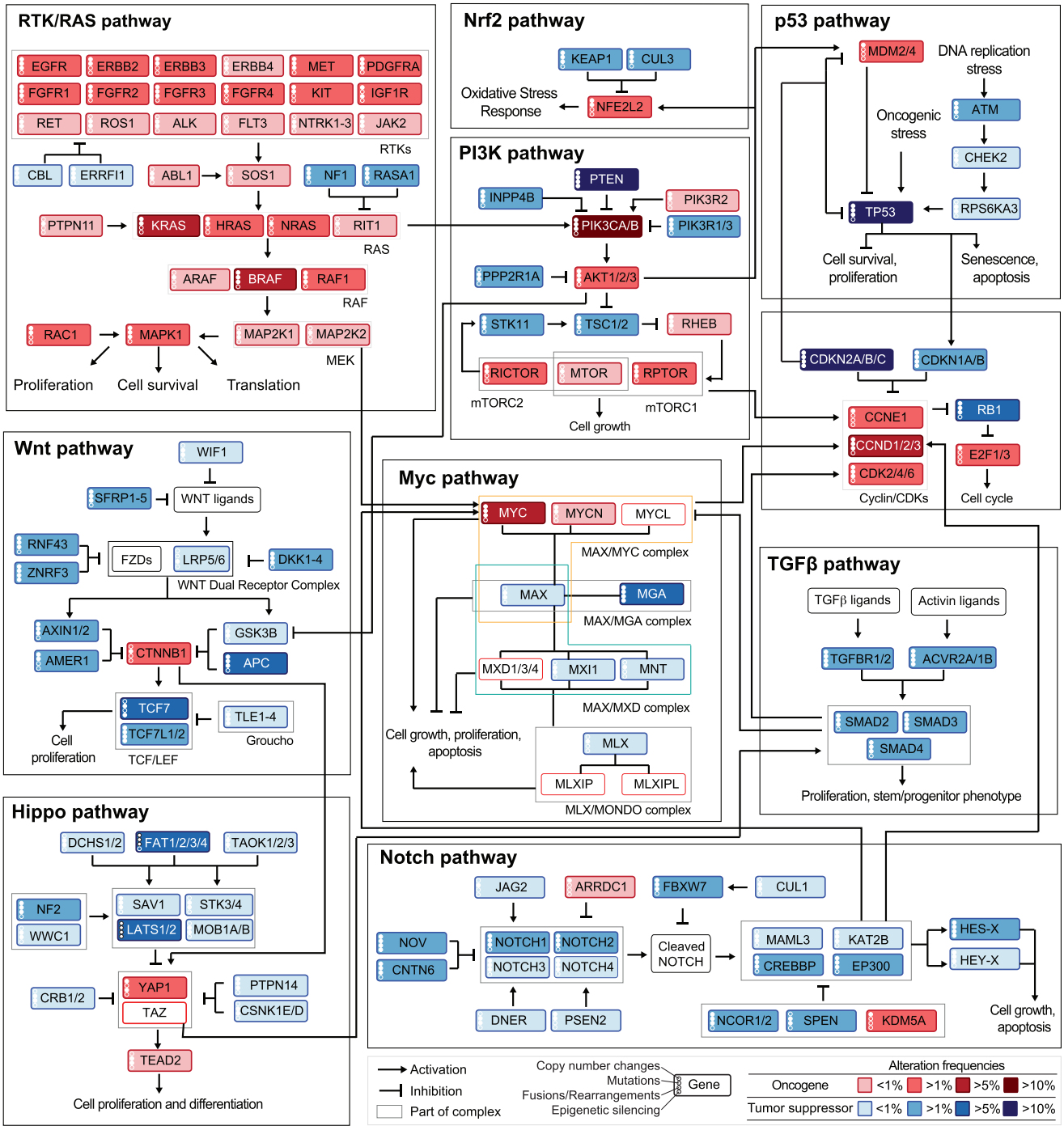
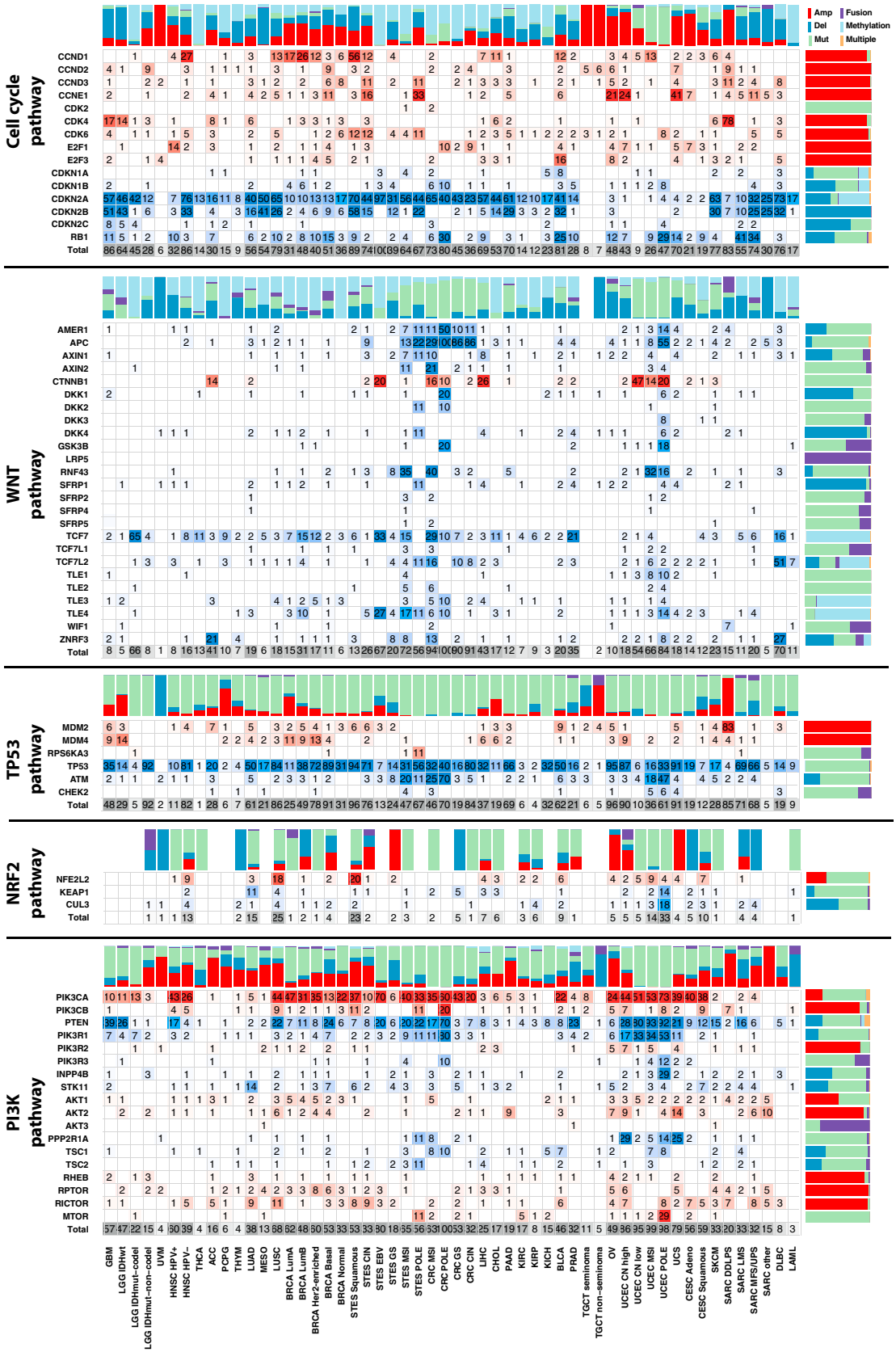


Figure S1. Curated Pathways Including Cross-Cross Pathway Interactions, Related to Figure 2

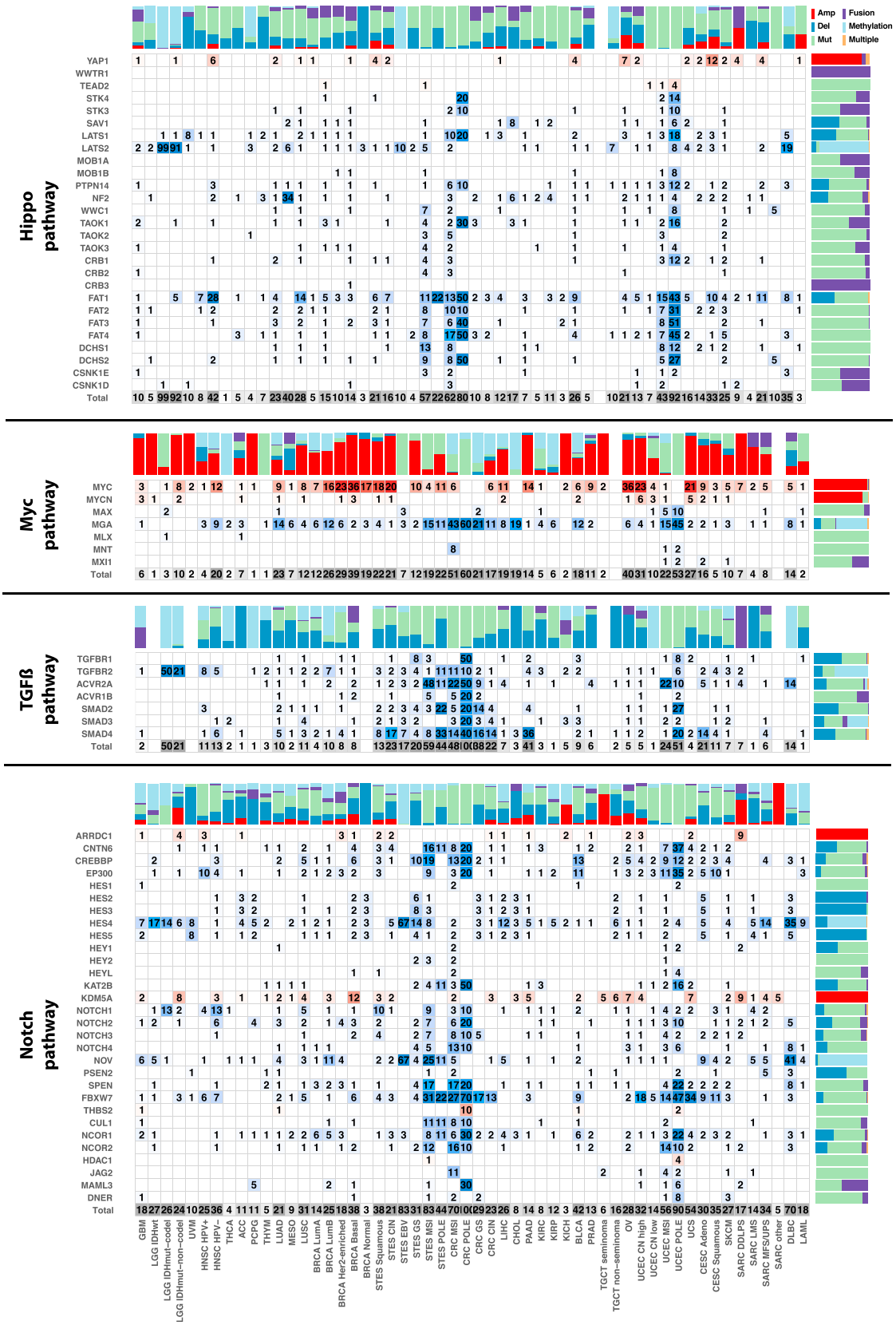
Same as Figure 2, but including cross-pathway interactions.



(legend on next page)

Figure S2. Cell-Cycle, Wnt, p53, Nrf2 and PI3K Pathway Alterations, Related to Figure 5

Detailed heatmap of alteration frequencies in members of the Cell-cycle, WNT, TP53 and NRF2 pathways. Shades of red indicate activating event (mutation, amplification, activating fusion) and shades of blue indicate inactivating event (mutation, homozygous loss, inactivating fusion, epigenetic silencing). Color side bars show the fraction of samples affected by each type of somatic alteration (or a combination of them) for each pathway gene. Top color bars show the proportion of different types of alterations for each cancer subtype.



(legend on next page)

Figure S3. Hippo, Myc, TGFβ and Notch Pathway Alterations, Related to Figure 5

Detailed heatmap of alteration frequencies in members of the Hippo, Myc, TGFβ and Notch pathways. Shades of red indicate activating event (mutation, amplification, activating fusion) and shades of blue indicate inactivating event (mutation, homozygous loss, inactivating fusion, epigenetic silencing). Color side bars show the fraction of samples affected by each type of somatic alteration (or a combination of them) for each pathway gene. Top color bars show the proportion of different types of alterations for each cancer subtype.

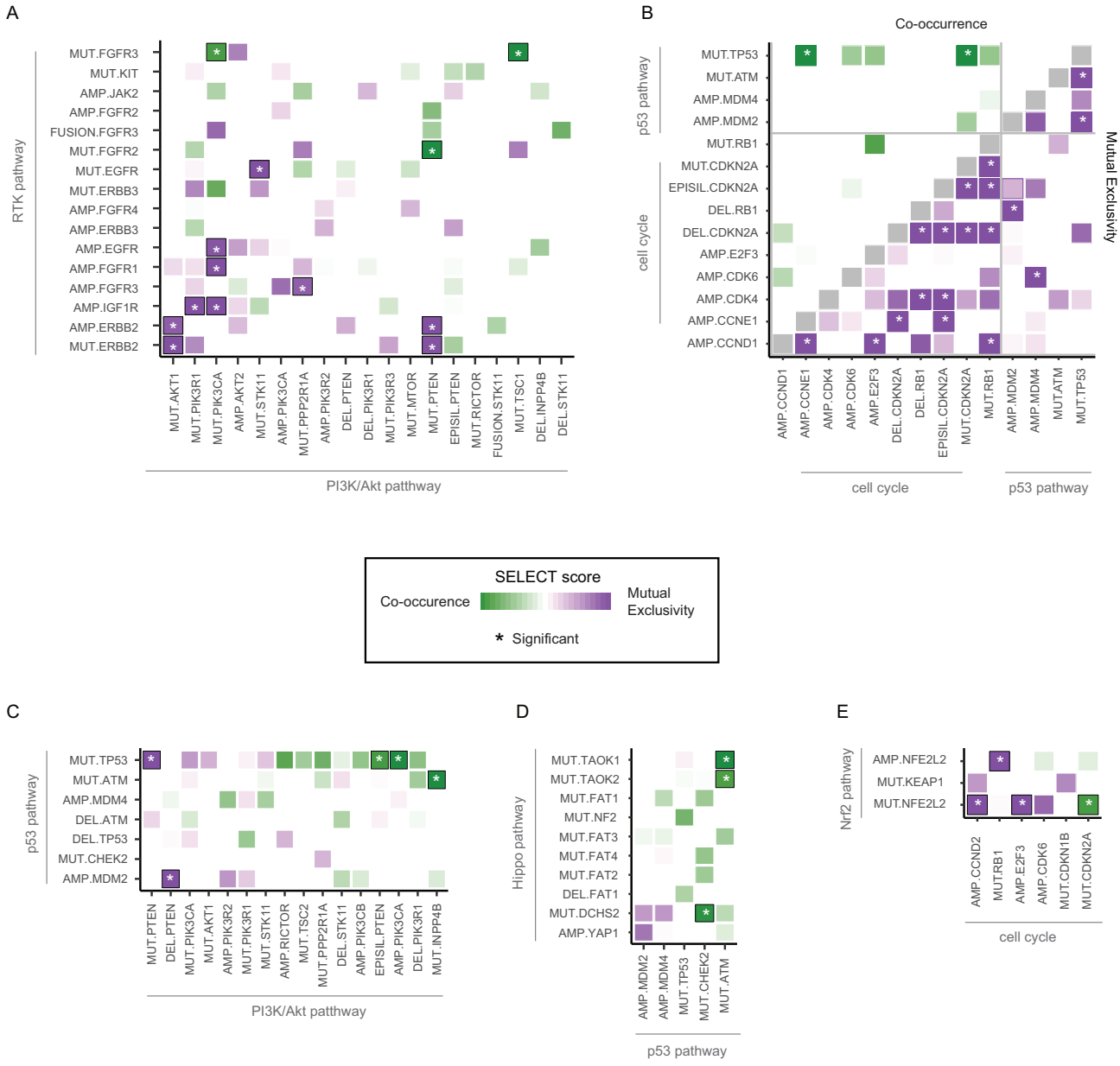


Figure S4. Additional Results for Conditional Selection, Related to Figure 6

(A–E) Mutual exclusivity (in purple) and Co-occurrence (in green) between alterations in (A) PI3K and RTK pathways, (B) Cell-cycle and p53 pathways, (C) p53 and PI3K pathways, (D) p53 and Hippo pathways, and (E) cell-cycle and Nrf2 pathways.