iScience, Volume *23*

Supplemental Information

Systematic Analysis of the Aberrances and Functional Implications of Ferroptosis in Cancer

Zekun Liu, Qi Zhao, Zhi-Xiang Zuo, Shu-Qiang Yuan, Kai Yu, Qingfeng Zhang, Xiaolong Zhang, Hui Sheng, Huai-Qiang Ju, Han Cheng, Feng Wang, Rui-Hua Xu, and Ze-Xian Liu

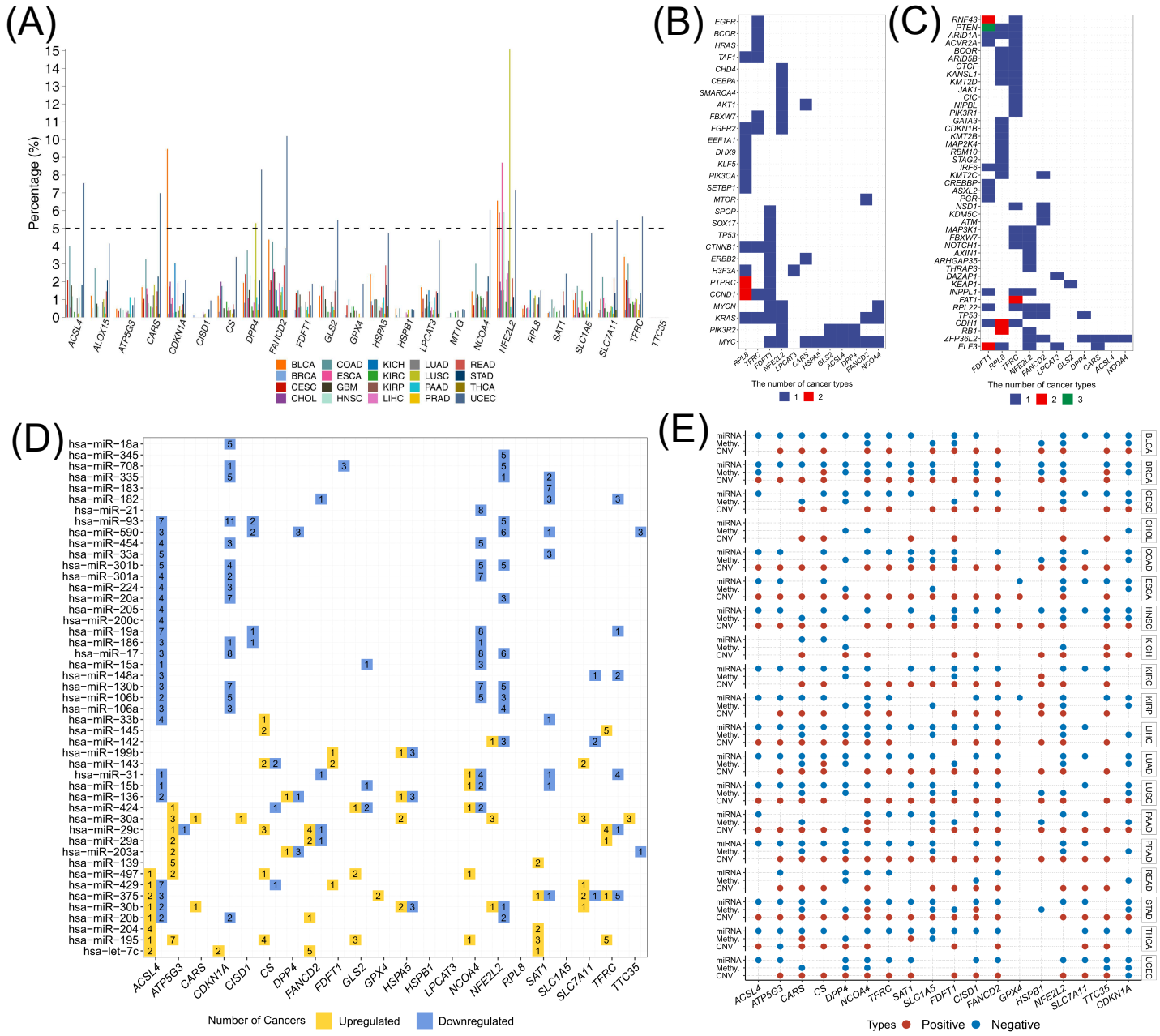

Figure S1. Molecular alterations of FRGs. Related to Figure 1.

The mutation frequencies (A), mutual exclusivity of FRGs and oncogenes (B) /tumor suppressor genes (C) among cancers. (D) The differential expressed FRGs-related miRNAs. The correlation between FRG expression and somatic copy number alternation, DNA methylation and miRNA expressions.

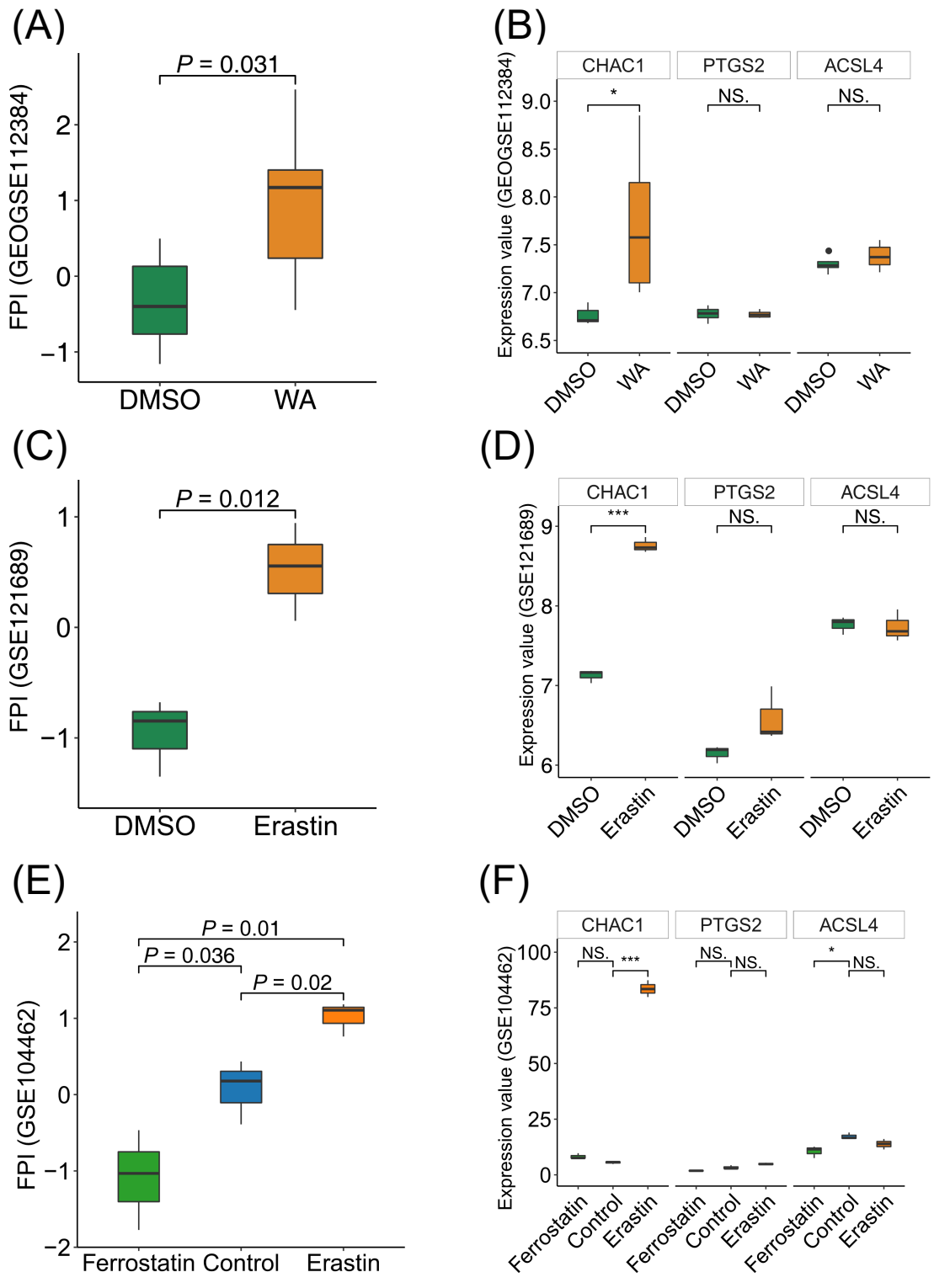

Figure S2. Established the FPI in cell lines and examined the performance of the FPI. Related to Figure 2.

The boxes showed the FPI between withaferin A (WA), erastin or ferrostatin and the control (A, D, G). The boxes showed the comparison of three genes between stimulators and control in neuroblastoma cells (B), clear cell carcinoma cells (E) and liver cancer cells (H). The correlation between the expression of *CHAC1* and the FPI in neuroblastoma cells (C), clear cell carcinoma cells (F) and liver cancer cells (I). The boxes in A-F mean the median values ± 1 quartile, their whiskers extending from the median to the smallest or biggest value which is 1.5 × IQR from the boundary of boxes. NS. indicates not significant (*p* value > 0.05), * *p* value < 0.05; ** *p* value < 0.01; *** *p* value < 0.001.

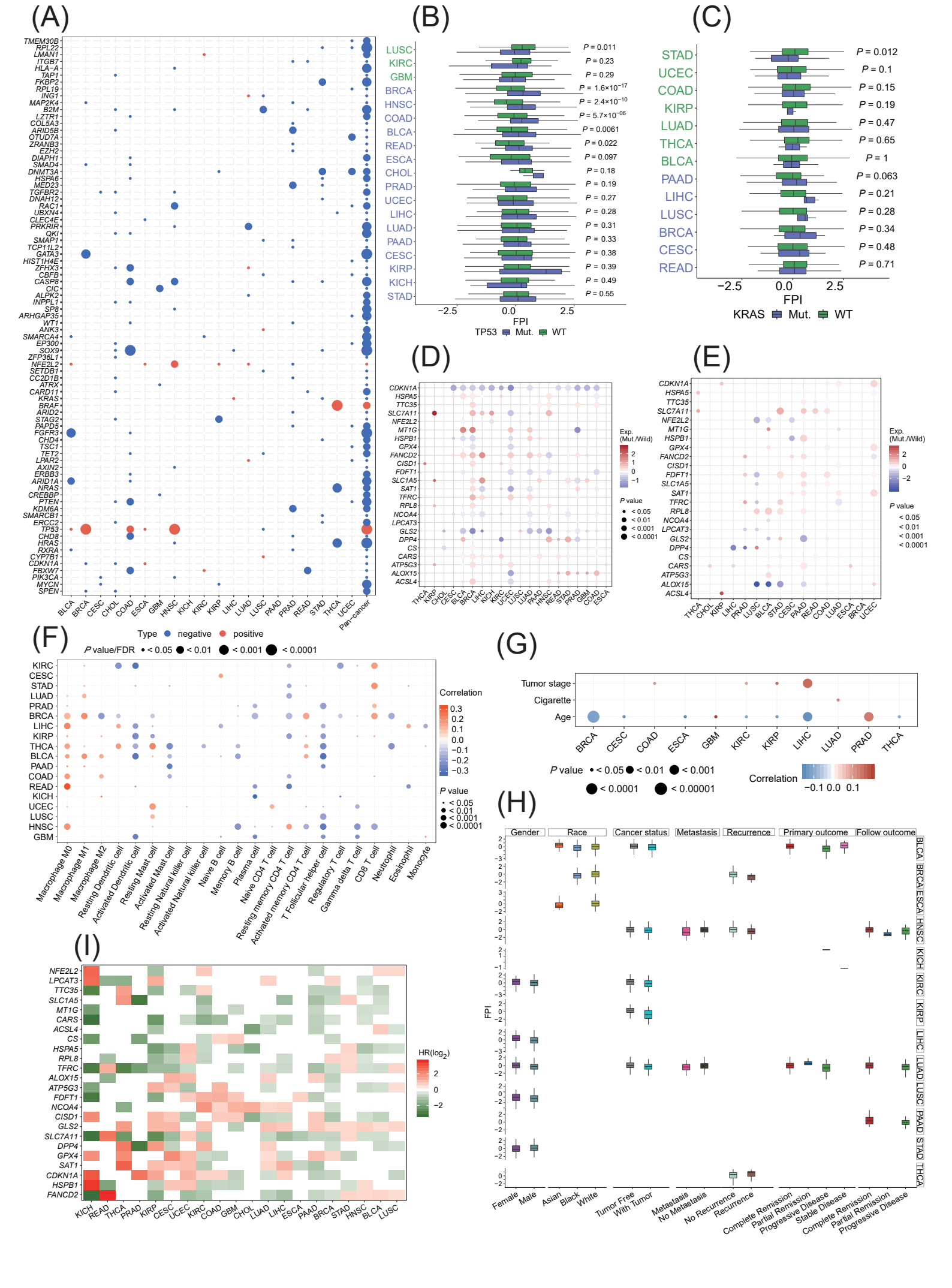
(A)
(B)
(C)
(D)
(E)
(F)
(G)
(H)
(I)
TP53 Mut. WT
KRAS Mut. WT
FPI
Exp. (Mut./Wild)
P value
Type negative positive
P value/FDR
Correlation
Tumor stage
Cigarette
Age
Gender
Race
Cancer status
Metastasis
Recurrence
Primary outcome
Follow outcome
HR(log2)
Pan-cancer

Figure S3. The relationship between other factors and the FPI or expression of FRGs. Related to Figure 2, Figure 3, and Figure 4.

The implication of the mutation status of driver genes on the FPI (A). The FPI between mutant and wild type tumors for *TP53* (B) and *KRAS* (C). The differential expression of FRGs between mutant and wild type tumors for *TP53* (D) and *KRAS* (E). The correlation between ferroptosis and immune cells (F) and clinical characteristics (G-H). The overall prognostic abilities of FRGs (I). The boxes in B, C, and H mean the median values $\pm$ 1 quartile, their whiskers extending from the median to the smallest or biggest value which is 1.5 $\times$ IQR from the boundary of boxes.

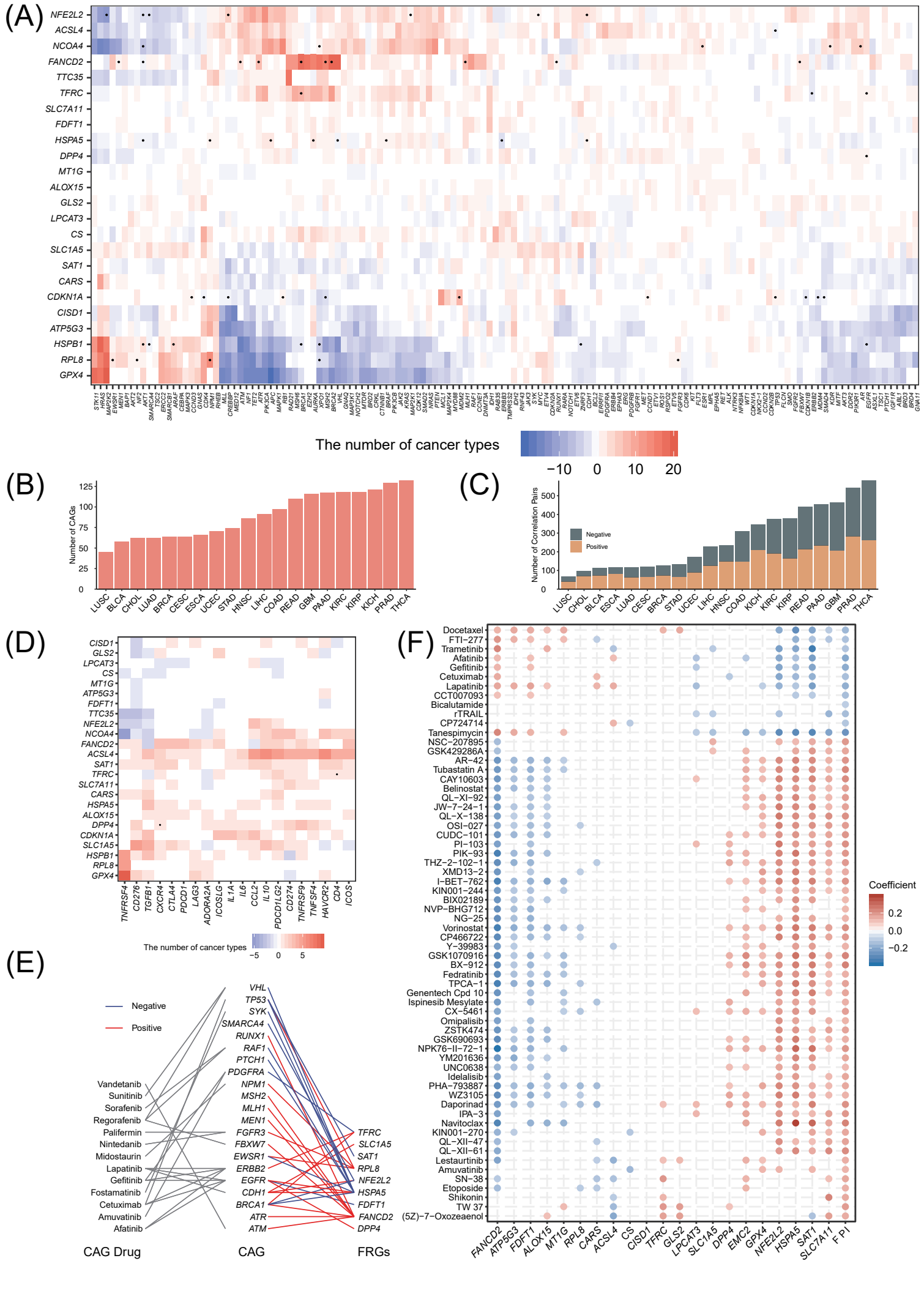
(A)
NFE2L2
ACSL4
NCOA4
FANCD2
TTC35
TFRC
SLC7A11
FDFT1
HSPA5
DPP4
MT1G
ALOX15
GLS2
LPCAT3
CS
SLC1A5
SAT1
CARS
CDKN1A
CISD1
ATP5G3
HSPB1
RPL8
GPX4
The number of cancer types
−10
0
10
20
(B)
Number of CAGs
125
100
75
50
25
0
LUSC
BLCA
CHOL
LUAD
BRCA
CESC
ESCA
UCEC
STAD
HNSC
LIHC
COAD
READ
GBM
PAAD
KIRC
KIRP
KICH
PRAD
THCA
(C)
Number of Correlation Pairs
500
400
300
200
100
0
Negative
Positive
LUSC
CHOL
BLCA
ESCA
LUAD
CESC
BRCA
STAD
UCEC
LIHC
HNSC
COAD
KICH
KIRC
KIRP
READ
PAAD
GBM
PRAD
THCA
(D)
CISD1
GLS2
LPCAT3
CS
MT1G
ATP5G3
FDFT1
TTC35
NFE2L2
NCOA4
FANCD2
ACSL4
SAT1
TFRC
SLC7A11
CARS
HSPA5
ALOX15
DPP4
CDKN1A
SLC1A5
HSPB1
RPL8
GPX4
TNFRSF4
CD276
TGFB1
CXCR4
CTLA4
PDCD1
LAG3
ADORA2A
ICOSLG
IL1A
IL6
CCL2
IL10
PDCD1LG2
CD274
TNFRSF9
TNFSF4
HAVCR2
CD4
ICOS
The number of cancer types
−5
0
5
(E)
Negative
Positive
Vandetanib
Sunitinib
Sorafenib
Regorafenib
Palifermin
Nintedanib
Midostaurin
Lapatinib
Gefitinib
Fostamatinib
Cetuximab
Amuvatinib
Afatinib
VHL
TP53
SYK
SMARCA4
RUNX1
RAF1
PTCH1
PDGFRA
NPM1
MSH2
MLH1
MEN1
FGFR3
FBXW7
EWSR1
ERBB2
EGFR
CDH1
BRCA1
ATR
ATM
TFRC
SLC1A5
SAT1
RPL8
NFE2L2
HSPA5
FDFT1
FANCD2
DPP4
CAG Drug
CAG
FRGs
(F)
Docetaxel
FTI−277
Trametinib
Afatinib
Gefitinib
Cetuximab
Lapatinib
CCT007093
Bicalutamide
rTRAIL
CP724714
Tanespimycin
NSC−207895
GSK429286A
AR−42
Tubastatin A
CAY10603
Belinostat
QL−XI−92
JW−7−24−1
QL−X−138
OSI−027
CUDC−101
PI−103
PIK−93
THZ−2−102−1
XMD13−2
I−BET−762
KIN001−244
BIX02189
NVP−BHG712
NG−25
Vorinostat
CP466722
Y−39983
GSK1070916
BX−912
Fedratinib
TPCA−1
Genentech Cpd 10
Ispinesib Mesylate
CX−5461
Omipalisib
ZSTK474
GSK690693
NPK76−II−72−1
YM201636
UNC0638
Idelalisib
PHA−793887
WZ3105
Daporinad
IPA−3
Navitoclax
KIN001−270
QL−XII−47
QL−XII−61
Lestaurtinib
Amuvatinib
SN−38
Etoposide
Shikonin
TW 37
(5Z)−7−Oxozeaenol
FANCD2
ATP5G3
FDFT1
ALOX15
MT1G
RPL8
CARS
ACSL4
CS
CISD1
TFRC
GLS2
LPCAT3
SLC1A5
DPP4
EMC2
GPX4
NFE2L2
HSPA5
SAT1
SLC7A11
F PI
Coefficient
0.2
0.0
−0.2

Figure S4. The pharmacogenomic interaction of the FPI and FRGs. Related to Figure 4.

(A) The co-expression of FRGs and 136 drug-targeted genes. (B) The number of CAGs correlated with FRGs. (C) The number of FRG-CAG correlation pairs. (D) The co-expression of FRGs and 20 genes for immunotherapy. (E) The interactions between drugs, drug targets and FRGs. (F) The correlation between the expression of FRGs and the area under the dose-response curve (AUCs) for drugs.

Transparent Methods

Datasets and Source

The mRNA expression data, copy number alteration thresholded data, masked copy number segmentation data, and DNA methylation 450K data of twenty cancers, including bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), cervical and endocervical cancers (CESC), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), kidney chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), pancreatic adenocarcinoma (PAAD), prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), uterine corpus endometrial carcinoma (UCEC), which had both tumors and normal samples were downloaded from Firehose (http://gdac.broadinstitute.org). Mutation data, miRNA-seq data, and clinical data were downloaded from the Xena Browser (https://xenabrowser.net/datapages/). Additional gene-centric RMA-normalized gene expression profiles and drug response data of over 1000 cancer cell lines were accessed from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/downloads) (Yang et al., 2013). Immune associated data, including immune cell type fractions and immunophenoscore were obtained from TCIA (https://tcia.at/home) (Charoentong et al., 2017), dysfunction and exclusion scores of tumor samples were obtained from TIDE (http://tide.dfci.harvard.edu/) (Fu et al., 2020; Jiang et al., 2018). Integrated protein-protein interaction data was obtained from the Human Protein Reference Database (http://www.hprd.org/) and BioGRID (https://thebiogrid.org) (Keshava Prasad et al., 2009; Peri et al., 2003). We obtained clinically actionable genes (CAGs) from a previous study (Van Allen et al., 2014) (https://software.broadinstitute.org/cancer/cga/target).

Differential expression analysis of mRNA

To test genes differentially expressed between tumor and normal tissue, gene expression data for 20502 genes across 20 cancer types were downloaded from TCGA at FireBrowse (http://gdac.broadinstitute.org, 2016 January). Then, the fold change and adjusted *P*-value were calculated by the edgeR package (Robinson et al., 2010). We defined genes with an adjusted *P*-value

less than 0.05 as the differential expression genes (DEGs).

Establishing the Ferroptosis Potential Index Model

The index to represent the ferroptosis level was establish based on the expression data for genes of ferroptosis core machine including positive components of *LPCAT3*, *ACSL4*, *NCOA4*, *ALOX15*, *GPX4*, *SLC3A2*, *SLC7A11*, *NFE2L2*, *NOX1*, *NOX3*, *NOX4*, *NOX5* and negative components of *FDFT1*, *HMGCR*, *COQ10A*, *COQ10B*. The enrichment score (ES) of gene set that positively or negatively regulated ferroptosis was calculated using single sample gene set enrichment analysis (ssGSEA) in the R package 'GSVA' (Hanzelmann et al., 2013), and the normalized differences between the ES of the positive components minus negative components was defined as the ferroptosis potential index (FPI) to computationally dissect the ferroptosis levels/trends of the tissue samples.

Somatic Copy-number Alteration (SCNA) and Mutation Analysis

The heterozygosity and homozygosity of amplification and deletion were included to elevate the copy-number alteration of each gene, in which over five percent was regarded as high-frequency SCNA. Pearson's correlation between expression values and copy number segment values of each gene was calculated to evaluate the association between SCNA and expression. The R package "DISCOVER" was employed to evaluate mutual exclusivity between FRGs and tumor suppressor genes or oncogenes across tumor samples in each cancer type (Canisius et al., 2016). The mutation and CNA events were integrated, while only homozygous amplification and deletion were included, and only protein-coding mutations were retained. For each cancer type, the genes were considered to be mutually exclusive if they had a q value of 0.05.

DNA methylation analysis

The R package "IlluminaHumanMethylation-450kanno.ilmn12.hg19" from Bioconductor was imported to annotate the methylation probe for the promoter of each gene. Differential methylation of each gene in tumor and normal samples was tested by the Wilcoxon signed rank test, and genes that were significantly hypomethylated or hypermethylated were identified using a *P*-value cutoff of 0.05. Pearson's correlation between the transcriptional expression of FRGs and the Beta value of the promoter DNA methylation were calculated and considered significant if the *P*-value < 0.05.

miRNA expression analysis

To investigate the mechanisms of dysregulation for ferroptosis regulator genes in cancer, we searched potential miRNAs which might regulate the FRGs based on miRNA-target intersections in starBase. The Spearman correlation between the expression of miRNA and FRGs was statistically evaluated (adjusted *P*-value < 0.1, rho < -0.1). Cytoscape software was used to visualize the high-frequency interaction networks among FRGs and miRNA (Shannon et al., 2003).

Multivariate Regression Analysis of Gene Expression

To assess which factors had significant effects on FRG expression, the expression of each FRG was modeled by linear regression as a function of the median miRNA expression, the median Beta value of promoter methylation, and the copy number of the genes.

Clinical Features Analysis

The R package “survival” was used to assess the prognosis potential of the FRGs and ferroptosis potential index among cancers. For survival analysis, the expression threshold was exhaustively tested and the one with most significant *P*-value was considered the best cut-off. To test the association between ferroptosis level and clinical features, the Pearson correlation was calculated between FPI and tumor stages, age, body mass index (BMI) and cigarette exposure per day, which were converted to numeric variables (“stage1” = 1, “stage2” = 2, etc.). Wilcoxon rank sum tests and Tukey’s tests were used to determine the impact on the ferroptosis potential index for other clinical characteristics including race, remission status, and alcohol. The influence of different MSI statuses(Liu et al., 2018), histologic types, and molecular subtypes on FPI were also considered(Ciriello et al., 2015).

Immune Features Analysis

To study the relationship of ferroptosis and immune microenvironments, we computed the Pearson correlation between FPI and immune parameters including immune cell fractions and 5 types of immunophenoscores.

Identifying the FPI associated significant driver gene mutation

A total of 375 driver genes identified in previous pan-cancer research were included for analysis(Lawrence et al., 2013). To test whether a driver gene's mutational status was significantly associated with ferroptosis among cancers, the rank-transformed FPI was modeled by linear regression as a function of the driver gene's mutational status, ignoring the synonymous variant. To diminish the confounding effects, the rank-transformed count of the total nonsynonymous mutations and the tumor type, which were encoded as virtual variables, were included. To further characterize each cancer, the Benjamini-Hochberg method was used to correct the *P* values across 375 genes. Genes with an adjusted *P* value less than 0.05 for mutation status variable were significantly associated with FPI.

Gene Set Enrichment Analysis

To identify the pathways associated with ferroptosis, the samples of each tumor type were divided into two groups according to the FPI, consisting of the top 30% and bottom 30%. Then, the gene set enrichment analysis (GSEA) was performed (Subramanian et al., 2005).

Correlation between Drug Sensitivity and FPI/FRG Expression

To test the correlation between small molecular drugs and FPI and FRGs, the Pearson correlation coefficients for FPI, the expression value of FRGs and the area under the dose-response curve (AUCs) values were calculated, the results with $|R| > 0.1$ and *P*-value < 0.05 were considered as significantly correlated. To further investigate the influence of drugs on ferroptosis, the significant association between the expression of clinically actionable genes and FRGs were finished across cancer cell lines, and the associations were filtered with PPIs. Then the drugs which target CAGs were selected according to DrugBank (Wishart et al., 2018) (https://www.drugbank.ca).

Supplemental References

Canisius, S., Martens, J.W., and Wessels, L.F. (2016). A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol *17*, 261.

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., Hackl, H., and Trajanoski, Z. (2017). Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep *18*, 248-262.

Ciriello, G., Gatza, M.L., Beck, A.H., Wilkerson, M.D., Rhie, S.K., Pastore, A., Zhang, H., McLellan, M., Yau, C., Kandoth, C., *et al.* (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. Cell *163*, 506-519.

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., and Liu, X.S. (2020). Large-scale public data reuse to model immunotherapy response and resistance. Genome Med *12*, 21.

Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 7.

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., Li, Z., Traugh, N., Bu, X., Li, B., *et al.* (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. Nat Med *24*, 1550-1558.

Keshava Prasad, T.S., Goel, R., Kandasamy, K., Keerthikumar, S., Kumar, S., Mathivanan, S., Telikicherla, D., Raju, R., Shafreen, B., Venugopal, A., *et al.* (2009). Human Protein Reference Database--2009 update. Nucleic Acids Res *37*, D767-772.

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.

Liu, Y., Sethi, N.S., Hinoue, T., Schneider, B.G., Cherniack, A.D., Sanchez-Vega, F., Seoane, J.A., Farshidfar, F., Bowlby, R., Islam, M., *et al.* (2018). Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. Cancer Cell *33*, 721-735 e728.

Peri, S., Navarro, J.D., Amanchy, R., Kristiansen, T.Z., Jonnalagadda, C.K., Surendranath, V., Niranjan, V., Muthusamy, B., Gandhi, T.K., Gronborg, M., *et al.* (2003). Development of human protein reference database as an initial platform for approaching systems biology in humans. Genome Res *13*, 2363-2371.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Van Allen, E.M., Wagle, N., Stojanov, P., Perrin, D.L., Cibulskis, K., Marlow, S., Jane-Valbuena, J., Friedrich, D.C., Kryukov, G., Carter, S.L., *et al.* (2014). Whole-exome sequencing and clinical interpretation of formalin-fixed, paraffin-embedded tumor samples to guide precision cancer medicine. Nat Med *20*, 682-688.

Wishart, D.S., Feunang, Y.D., Guo, A.C., Lo, E.J., Marcu, A., Grant, J.R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., *et al.* (2018). DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res *46*, D1074-D1082.

Yang, W., Soares, J., Greninger, P., Edelman, E.J., Lightfoot, H., Forbes, S., Bindal, N., Beare, D., Smith, J.A., Thompson, I.R., *et al.* (2013). Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res *41*, D955-961.