Am J Clin Exp Immunol 2024;13(4):187-194
www.ajcei.us /ISSN:2164-7712/AJCEI0159093

Original Article

Characterization of tumor suppressors and oncogenes evaluated from TCGA cancers

Claire Shen[1,2*], Richard Geng[3*], Sissi Zhu[4*], Michael Huang[5], Jeffrey Liang[6], Binze Li[7], Yongsheng Bai[8,9]

[1]Johns Hopkins University, Baltimore, MD 21218, USA; [2]Jordan High School, Fulshear, TX 77441, USA; [3]Carmel High School, Carmel, IN 46032, USA; [4]Shady Side Academy, Pittsburgh, PA 15238, USA; [5]James E. Taylor High School, Katy, TX 77450, USA; [6]Daniel Hand High School, Madison, CT 06443, USA; [7]The University of California, Los Angeles, Los Angeles, CA 90095, USA; [8]Next-Gen Intelligent Science Training, Ann Arbor, MI 48105, USA; [9]Eastern Michigan University, Ypsilanti, MI 48197, USA. [*]Equal contributors.

Received July 9, 2024; Accepted August 14, 2024; Epub August 25, 2024; Published August 30, 2024

Abstract: Mutations in oncogenes and tumor suppressor genes can significantly impact cellular function during cancer development. A comprehensive analysis of their mutation patterns and significant gene ontology terms can provide insights into cancer emergence and suggest potential targets for drug development. This study analyzes twelve cancer subtypes by focusing on significant genetic and molecular factors. Two common genetic mutations associated with cancer are single nucleotide variants (SNVs) and copy number alterations (CNAs). Oncogenes, derived from mutated proto-oncogenes, disrupt normal cell functions and promote cancer, while tumor suppressor genes, often inactivated by mutations, regulate cell processes like proliferation and DNA damage response. This study analyzed datasets from The Cancer Genome Atlas (TCGA), which provides extensive genomic data across various cancers. In our analysis results, many genes with significant *p*-values based on Kaplan Meier gene expression data were identified in eight cancers (BRCA, BLCA, HNSC, KIRC, LUAD, KIRP, LUSC, STAD). Moreover, STAD is the only cancer for genes with both significant *p*-values and functional terms reported. Interestingly, we found that LIHC was the cancer reported with only one CNA mutated gene and its survival plot *p*-value being significant. Additionally, KICH has no reported significant genes at all. Our study proposed the relationship between tumor suppressor and oncogenes and shed light on cancer tumorigenesis due to genetic mutations.

Keywords: The Cancer Genome Atlas, computational biology, cancer, tumor suppressors, oncogenes

Introduction

Cancer is a disease where certain cells in the body multiply uncontrollably and can spread to other areas. In this study, we focus on the significant term analysis of twelve different cancer subtypes. Bladder urothelial carcinoma (BLCA) is one of the most common types of malignant tumors found in the urogenital system in adults. It predominantly originates in the urothelium, which is the epithelial tissue lining the inner surface of urinary organs [1]. Breast invasive carcinoma (BRCA) is the most common cancer diagnosed in women and it is the second most common cause of death from cancer among women in the world [2]. Cancers of the oral cavity and larynx such as head and neck squamous cell carcinoma (HNSC) are typically linked to tobacco use, excessive alcohol consumption, or both, while pharyngeal cancers are increasingly associated with human papillomavirus (HPV) infection [3]. Liver hepatocellular carcinoma (LIHC) ranks as the fifth most common cancer and is recognized as the second leading cause of cancer-related deaths. Despite advancements in screening and new discoveries, LIHC progresses quickly and has a high mortality rate. This is because patients with LIHC are often diagnosed at advanced stages due to the absence of specific symptoms [4].

Lung cancer is among the world's deadliest cancers. Lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), two prevalent subtypes, have significantly distinct biological characteristics. Despite this, they are frequently grouped together as non-small cell lung cancer (NSCLC) and often receive

https://doi.org/10.62347/XMZW6604

similar treatment [5]. Prostate adenocarcinoma (PRAD) is a common type of cancer in men, yet effective prognostic markers remain limited. Only a few indicators are used to predict the prognosis of PRAD patients, each with its own set of strengths and weaknesses [6]. Stomach adenocarcinoma (STAD) is a prevalent malignant tumor of the digestive tract. Identifying its survival predictors is essential for precision medicine but hasn't been thoroughly explored. The development of STAD is complex, influenced by multiple factors and stages, including genetic factors, H. pylori infection, smoking, and environmental factors [7]. Thyroid carcinoma (THCA) is the most prevalent malignant endocrine tumor, characterized by low mortality and generally favorable prognosis. Immune genes have garnered significant interest as molecular markers for THCA prognosis and potential targets for immunotherapy [8]. Renal cell carcinoma (RCC) is widespread globally and is the sixth most common cancer in the United States. The most prevalent RCC subtype is kidney renal clear cell carcinoma (KIRC). When one kidney is damaged, the other compensates, which often delays the detection of kidney function loss until later stages [9]. Kidney renal papillary cell carcinoma (KIRP) makes up 10%-15% of renal cell carcinoma (RCC) cases. Patients with KIRP typically have a poor prognosis, and there is a lack of effective prognostic markers for this cancer type [10]. Kidney chromophobe (KICH) is a rare subtype within renal cell carcinomas, a diverse group of cancers originating from the nephron [11].

An oncogene originates from a cellular gene (proto-oncogene) that becomes dysfunctional due to mutation, fusion with another gene, or overexpression. Oncogenes are understood to promote cancer by disrupting normal cell proliferation or by inhibiting the process of apoptosis. According to the cancer stem cell theory, cancers generally consist of a hierarchy of cells derived from a transformed tissue-specific stem cell [12]. These normal equivalents produce different cell types within a tissue, offering further insight into how oncogenes could contribute to the disorderly conduct of cancer cells [12]. The build-up of genetic alterations, such as the activation of proto-oncogenes and the deactivation of tumor-suppressor genes, propels the transformation of a normal cell into a cancerous one [13]. Tumor suppressor genes produce essential intracellular regulators, such as the retinoblastoma protein, controlling processes like cell proliferation, cell survival, and DNA damage response [14]. These genes are often mutated in various cancers. Research from numerous labs has demonstrated that while proto-oncogenes are activated through dominant gain-of-function mutations, tumor suppressor genes are typically inactivated by recessive loss-of-function mutations or epigenetic silencing [14]. By 1990, tumor suppressor genes were recognized as being as crucial to cancer development as oncogenes [14, 15].

In our study, we analyzed 12 cancer subtypes in The Cancer Genome Atlas (TCGA) to uncover key genetic factors and their regulatory mechanisms [16]. We used three main steps: First, we identified genes with significant single nucleotide variant (SNV) mutations or copy number alterations (CNA) mutations for each cancer subtype. Second, we determined the microRNAs (miRNAs) that target these mutated genes. Third, we classified these genes and miRNAs as oncogenes or tumor suppressors and visualized these classifications. Additionally, we identified significant Gene Ontology (GO) terms related to each cancer subtype to understand the broader biological implications. This integrated approach aimed to reveal insights into tumorigenesis and identify potential therapeutic targets.

Material and methods

Input gene selection from previous study

For every cancer subtype, we did a series of gene comparisons to find the genes with significant levels of mutations (genes with single nucleotide variants and copy number alterations). Supplementary Files 1-19 were retrieved from a published research study and used to generate gene lists selected from significant clusters with a q-value less than 0.1 [17]. The number of clusters and significant genes within each cancer subtype can be found in **Table 1**. The first gene comparison was between these gene lists and cBioPortal datasets of genes from each cancer subtype that contained SNVs and CNAs [18-20]. We then took these new common genes and compared them against genes present in cell lines. This was done by

Table 1. Cancer subtypes and their number of total clusters and significant genes

Cancer Subtype	Cluster Count	Significant Gene Count
BRCA	33	4238
BLCA	28	199
HNSC	96	556
KIRC	51	2551
LUAD	21	5682
KIRP	62	667
LUSC	39	62
THCA	57	311
STAD	39	2570
KICH	64	28
PRAD	52	504
LIHC	114	56

utilizing the online bioinformatics tool TumorComparer [21]. We compared these genes against TumorComparer genes found in cell lines with a CNA rank greater than 0.5.

Tumor suppressor and oncogene annotation

Furthermore, we generated and downloaded survival plots using the online bioinformatics tool UALCAN for genes with single nucleotide variants (SNVs) and copy number alterations (CNAs) [22, 23]. This was also done for microRNAs. Patients that demonstrated higher survival at low or medium expression of the gene are oncogenes and patients with higher survival at high expression of the gene are tumor suppressor genes. MiRNAs were determined to be oncogenic or tumor suppressors in the same way. The process by which we annotated the survival plots is demonstrated in **Figure 1**.

Functional annotation for targeted pairs

For every cancer type, we performed g:Profiler analysis (**Figure 2**) to find the present significant gene ontology (GO) terms [24]. This was done by inputting the gene lists created by our team from the expression file analysis using XENA into the online tool g:Profiler. The final gene lists were then processed to generate lists of miRNAs that target these mutated genes using our in-house developed pipeline. The miRNA and gene lists generated were then used for survival plot analysis.

Data collection from UCSC Xena

We then analyzed whether mutations or gene expression data provided more significant *p*-values in the context of our study. This involved comparing the impact of genetic mutations and gene expression on cancers in our study. We gathered mutation and gene expression data using UCSC Xena, a data visualization and analysis platform developed by the University of California, Santa Cruz [25]. We also downloaded Kaplan Meier Survival plots from Xena and annotated them (**Figure 3**). Since Xena allows researchers to explore large-scale genomic and clinical datasets from the Cancer Genome Atlas (TCGA), UCSC Xena was the best platform for our data collection.

P-value analysis using XenaAuto-Suv

To generate and analyze the *p*-values from our collected data. We developed XenaAuto-Suv, which analyzes gene expression and mutation survival data downloaded from UCSC Xena. Developed in 2024 in the R coding language, XenaAuto-Suv uses one main algorithm that analyzes and generates files with significant and non-significant genes determined by a specified *p*-value cutoff. In our study, we chose a *p*-value cutoff of .05 (<.05) to determine the significance of the genes. This threshold is widely accepted as it balances Type 1 and Type 2 errors. The R packages survival and survminer are utilized by the initial script in the XenaAuto-Suv pipeline [26, 27]. The survival package provides the fundamentals for survival analysis in R. Functions in this package include creating survival objects, estimating survival curves, and generating *p*-values associated with the expression data. Furthermore, the survminer package complements the survival package by providing tools to visualize the survival curves.

XenaAuto-Suv uses these packages and the Surv() function to build a standard survival object for each gene. Then the survfit() function is called to produce the Kaplan Meier estimates of the probability of the survival times along with the corresponding *p*-value for each gene. The second algorithm that XenaAuto-Suv uses generates a list of the genes that appear the most frequently across the cancers analyzed. The software is available at https://github.com/richito-g/XenaAnalysis/tree/main?tab=readme-ov-file.

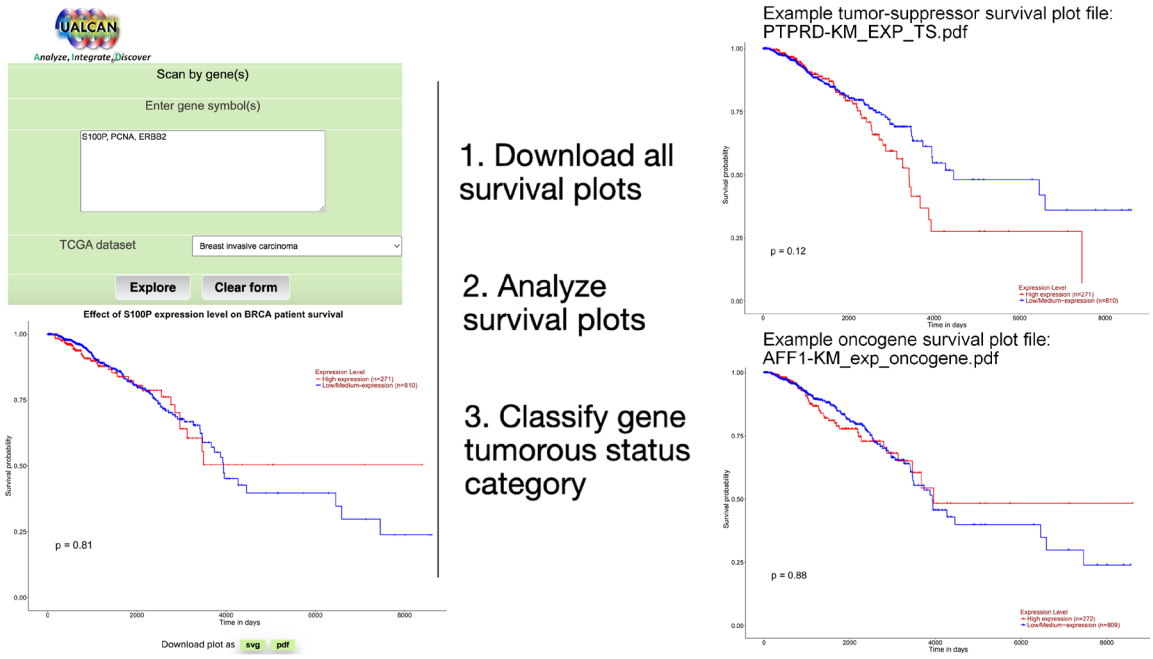

Figure 1. The workflow of identifying TS and oncogenes for 12 cancer types.

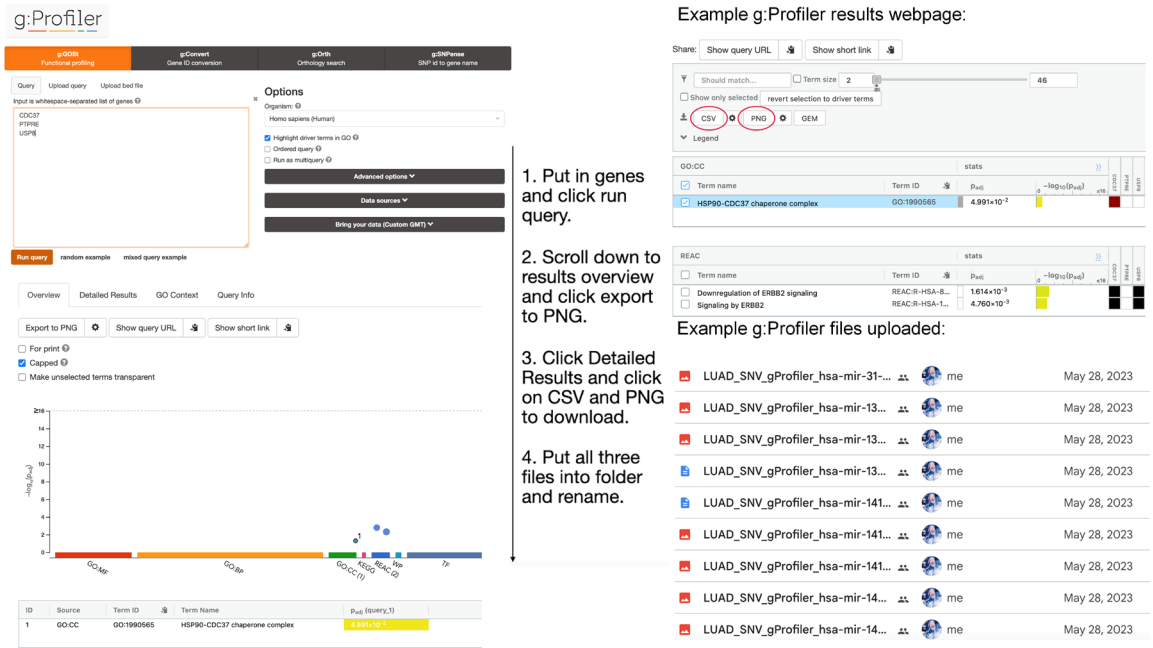

Figure 2. The workflow of downloading g:Profiler results for 12 cancer types.

Results

Significant gene information from clusters

Of the 12 cancer subtypes we studied, LIHC has the greatest number of detected clusters, and the cancer subtype with the greatest number of significant genes is LUAD. Although LIHC has the most clusters reported, it has the second-least number of significant genes. LIHC has 56 significant genes and KICH has the least, with 28 significant genes (**Table 1**).

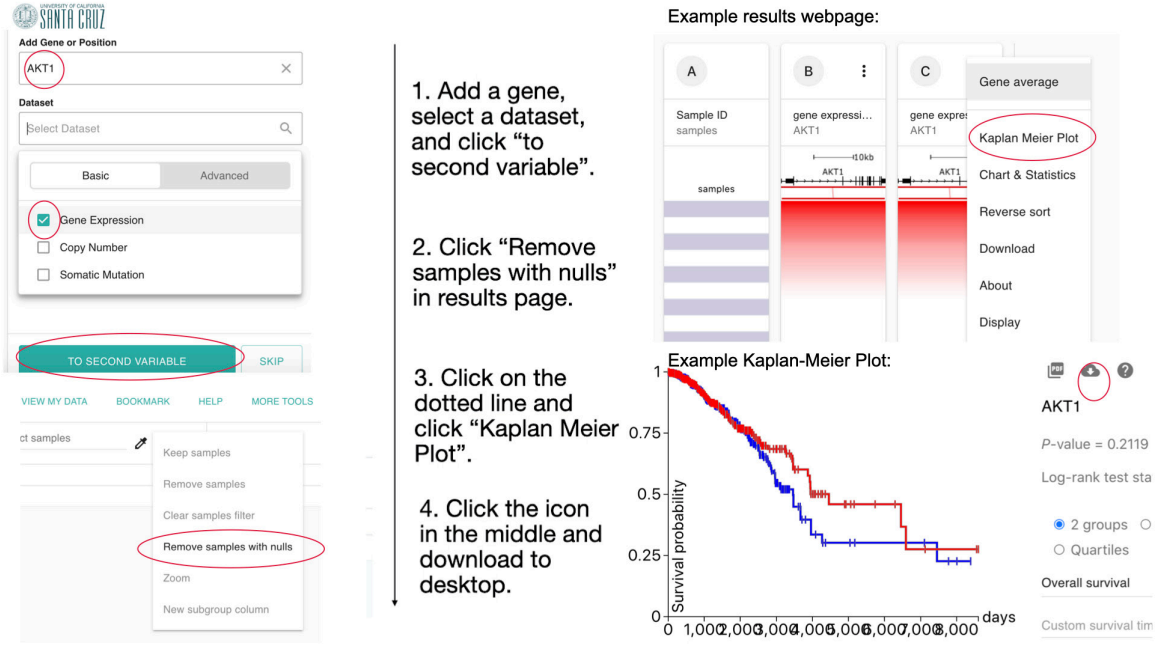

Figure 3. The workflow of downloading Xena Kaplan Meier plots.

Table 2. Survival plots statistics for studied genes

Cancer Subtype	SNV	CNA
BRCA	**932 (240)**	**251 (66)**
BLCA	**82 (28)**	5 (2)
HNSC	**61 (6)**	5 (0)
KIRC	**334 (126)**	**39 (15)**
LUAD	65 (7)	**394 (77)**
KIRP	**310 (80)**	13 (1)
LUSC	**27 (5)**	0 (0)
THCA	13 (0)	1 (0)
STAD	**1290 (222)**	**125 (22)**
KICH	4 (0)	1 (0)
PRAD	275 (8)	28 (1)
LIHC	40 (3)	1 (1)
Total	3433 (725)	863 (185)

*The number of significant genes is placed in parentheses. The bolded values have significant terms.

Statistics for studied genes

Table 2 reports that out of the 12 cancer subtypes, only BRCA, KIRC, and STAD have significant terms reported for both SNV and CNA categories. Moreover, the table shows that the cancer with the largest proportion of significant genes with SNVs is KIRC and the largest proportion of significant genes with CNAs is LIHC (Supplementary Figure 1). Since LIHC has only 1 significant CNA gene, it is worth noting that the next cancer subtype with the largest proportion of significant CNA genes is BLCA.

We noticed that BRCA has a significant term relating to TP53 regulation of gene transcription. In fact, a previous study has shown that somatic TP53 abnormalities are more frequently observed in breast cancers associated with BRCA1 or BRCA2 germ-line mutations compared to sporadic breast cancers [28].

Significant term analysis

After performing functional annotation using the online bioinformatics tool g:Profiler, we organized our significant term results into four different miRNA-gene targeting categories: oncogenic miRNAs that target oncogenes, oncogenic miRNAs that target tumor suppressor genes, tumor suppressor miRNAs that target tumor suppressor genes, and tumor suppressor miRNAs that target oncogenes (Supplementary Tables 1, 2). In these results, we then bolded the significant genes that were reported for each cancer subtype by our pipeline XenaAuto-Suv (Supplementary Table 3) that appeared to contain the significant terms.

XenaAuto-Suv software

After placing the downloaded data from UCSC Xena into their own designated folder. We used R-Studio to run XenaAuto-Suv which sorted the genes into files depending on the gene's significance. In these files, the genes' *p*-values are also displayed along with their significance (Supplementary Table 3). Next, to determine which genes appeared the most we used the pipeline's second script to remove the unnecessary information from the generated files. After isolating the gene names, we ran the third script in the XenaAuto-Suv pipeline to create a frequency table of all of the genes (Supplementary Table 4). This allowed us to identify which genes appeared the most across all of the cancers.

Discussion

Neurons have a direct influence on the behavior of normal and malignant cells by secreting neurotransmitters, neuropeptides, and protein-signaling ligands. Our research on head and neck squamous cell carcinoma showed 6 genes as significant: *DDHD2, FXR1, MTBP, RBS6KA5, SEMA3E, ZFAT* [29]. A study done in 2018 on the effects of the loss of *DDHD2* showed results that the loss of *DDHD2* promotes apoptosis of motor neurons, which suggests the significance of *DDHD2* [30]. Another study on aging shows evidence on how *MTBP* has been found over-expressed in many human malignancies and linked to poor patient outcomes [31].

The study of Glutamine metabolism genes prognostic signature for stomach adenocarcinoma (STAD) uses software R and Perl in order to find the relationship between STAD and GlnMgs [32]. To compare mRNA data with human survival data, the GlnMgs were sorted into different groups and a heatmap of GlnMgs was constructed and examined. The results suggested a possible correlation of STAD mutations with important gene dysregulation [32]. This process is similar to our study, as we obtained results using a similar process, by using STAD samples from TCGA datasets and using R software to run analysis.

We found that within KIRC, the PI3K-AKT-mTOR - vitamin D3 signaling pathway appeared as a significant term. This also aligns with previous studies that have shown that the phosphoinositide 3-kinase (PI3K)/protein kinase B (AKT)/mammalian target of rapamycin (mTOR) signaling pathway governs cell growth, differentiation, migration, survival, angiogenesis, and metabolism [33]. It is activated by growth factors, hormones, cytokines, and various extracellular signals [33]. Dysregulation of the PI3K/AKT/mTOR pathway is commonly observed in human cancers, including renal cell carcinoma (RCC), and is linked to aggressive tumor development and poor survival rates which are characteristic of KIRC [33].

Additionally, we found the significant term *EGFR* tyrosine kinase inhibitor resistance related to LUAD. The epidermal growth factor receptor (*EGFR*) and its three related proteins in the *ERBB* family are receptor tyrosine kinases crucial for normal physiological functions and cancer development [34]. When *EGFR* binds to its ligands, it undergoes dynamic conformational changes in both its extracellular and intracellular domains, leading to the transphosphorylation of tyrosine residues in the C-terminal regulatory domain [34]. These phosphorylated tyrosine residues act as docking sites for downstream molecules, promoting evasion of apoptosis, proliferation, invasion, and metastasis, which are all critical for the cancer phenotype [34]. In 2022, it was found that a mutation in the tyrosine kinase domain of the *EGFR* gene was discovered in a subset of lung cancers, including LUAD [34, 35]. Lung cancers harboring an *EGFR* mutation show high sensitivity to *EGFR* tyrosine kinase inhibitors, such as gefitinib and erlotinib [34, 35].

We chose this process to gain a comprehensive understanding of the genetic and molecular mechanisms driving different cancer subtypes. By identifying genes with significant SNV and CNA mutations, we aimed to pinpoint critical genetic alterations that contribute to cancer development. Understanding which miRNAs target these mutated genes allowed us to explore the regulatory networks involved in tumorigenesis. Classifying the identified genes and miRNAs as oncogenes or tumor suppressors was crucial for determining their roles in cancer progression. By distinguishing between these roles, we could better understand the molecular dynamics of each cancer subtype. Lastly, identifying significant Gene Ontology

(GO) terms further expanded our understanding by linking genetic alterations to specific biological processes, cellular components, and molecular functions. This broader context is essential for uncovering the pathways and cellular mechanisms affected by these mutations.

Acknowledgements

The results shown here are in whole or part based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

Disclosure of conflict of interest

None.

Address correspondence to: Claire Shen, Johns Hopkins University, Baltimore, MD 21218, USA. E-mail: claire.shen22@gmail.com; Yongsheng Bai, Next-Gen Intelligent Science Training, Ann Arbor, MI 48105, USA. E-mail: bioinformaticsresearchtomorrow@gmail.com

References

[1] Gilyazova I, Enikeeva K, Rafikova G, Kagirova E, Sharifyanova Y, Asadullina D and Pavlov V. Epigenetic and immunological features of bladder cancer. Int J Mol Sci 2023; 24: 9854.

[2] Menon G, Alkabban FM and Ferguson T. Breast cancer. Updated 2024 Feb 25. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024. Available from: https://www.ncbi.nlm.nih.gov/books/NBK482286/.

[3] Johnson DE, Burtness B, Leemans CR, Lui VWY, Bauman JE and Grandis JR. Head and neck squamous cell carcinoma. Nat Rev Dis Primers 2020; 6: 93.

[4] Kaur H, Bhalla S and Raghava GPS. Classification of early and late stage liver hepatocellular carcinoma patients from their genomics and epigenomics profiles. PLoS One 2019; 14: e0221476.

[5] Chen JW and Dhahbi J. Lung adenocarcinoma and lung squamous cell carcinoma cancer classification, biomarker identification, and gene expression analysis using overlapping feature selection methods. Sci Rep 2021; 11: 13323.

[6] Huang ZG, He RQ and Mo ZN. Prognostic value and potential function of splicing events in prostate adenocarcinoma. Int J Oncol 2018; 53: 2473-2487.

[7] Tong X, Yang X, Tong X, Zhai D and Liu Y. Complement system-related genes in stomach adenocarcinoma: prognostic signature, immune landscape, and drug resistance. Front Genet 2022; 13: 903421.

[8] Li J, Li Z and Zhao P. Diagnosis and prognosis of thyroid cancer by immune-related genes. Am J Clin Oncol 2024; 47: 1-10.

[9] Sun J, Chen F and Wu G. Role of NF-κB pathway in kidney renal clear cell carcinoma and its potential therapeutic implications. Aging (Albany NY) 2023; 15: 11313-11330.

[10] Lan H, Zeng J, Chen G and Huang H. Survival prediction of kidney renal papillary cell carcinoma by comprehensive LncRNA characterization. Oncotarget 2017; 8: 110811-110829.

[11] Rathmell KW, Chen F and Creighton CJ. Genomics of chromophobe renal cell carcinoma: implications from a rare tumor for pan-cancer studies. Oncoscience 2015; 2: 81-90.

[12] Brown G. Oncogenes, proto-oncogenes, and lineage restriction of cancer stem cells. Int J Mol Sci 2021; 22: 9667.

[13] Anderson MW, Reynolds SH, You M and Maronpot RM. Role of proto-oncogene activation in carcinogenesis. Environ Health Perspect 1992; 98: 13-24.

[14] Lipsick J. A history of cancer research: tumor suppressor genes. Cold Spring Harb Perspect Biol 2020; 12: a035907.

[15] Chen L, Liu S and Tao Y. Regulating tumor suppressor genes: post-translational modifications. Signal Transduct Target Ther 2020; 5: 90.

[16] Tomczak K, Czerwińska P and Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015; 19: A68-A77.

[17] Dai X, Ding L, Liu H, Xu Z, Jiang H, Handelman SK and Bai Y. Identifying interaction clusters for MiRNA and MRNA pairs in TCGA network. Genes (Basel) 2019; 10: 702.

[18] Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C and Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012; 2: 401-404.

[19] Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larsson E, Cerami E, Sander C and Schultz N. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013; 6: pl1.

[20] de Bruijn I, Kundra R, Mastrogiacomo B, Tran TN, Sikina L, Mazor T, Li X, Ochoa A, Zhao G, Lai B, Abeshouse A, Baiceanu D, Ciftci E, Dogrusoz U, Dufilie A, Erkoc Z, Garcia Lara E, Fu Z, Gross B, Haynes C, Heath A, Higgins D, Jagannathan P, Kalletla K, Kumari P, Lindsay J, Lisman A, Leenknegt B, Lukasse P, Madela D,

Madupuri R, van Nierop P, Plantalech O, Quach J, Resnick AC, Rodenburg SYA, Satravada BA, Schaeffer F, Sheridan R, Singh J, Sirohi R, Sumer SO, van Hagen S, Wang A, Wilson M, Zhang H, Zhu K, Rusk N, Brown S, Lavery JA, Panageas KS, Rudolph JE, LeNoue-Newton ML, Warner JL, Guo X, Hunter-Zinck H, Yu TV, Pilai S, Nichols C, Gardos SM and Philip J; AACR Project GENIE BPC Core Team, AACR Project GENIE Consortium; Kehl KL, Riely GJ, Schrag D, Lee J, Fiandalo MV, Sweeney SM, Pugh TJ, Sander C, Cerami E, Gao J and Schultz N. Analysis and visualization of longitudinal genomic and clinical data from the AACR Project GENIE biopharma collaborative in cBioPortal. Cancer Res 2023; 83: 3861-3867.

[21] Sinha R, Luna A, Schultz N and Sander C. A pan-cancer survey of cell line tumor similarity by feature-weighted molecular profiles. Cell Rep Methods 2021; 1: 100039.

[22] Chandrashekar DS, Karthikeyan SK, Korla PK, Patel H, Shovon AR, Athar M, Netto GJ, Qin ZS, Kumar S, Manne U, Creighton CJ and Varambally S. UALCAN: an update to the integrated cancer data analysis platform. Neoplasia 2022; 25: 18-27.

[23] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia 2017; 19: 649-658.

[24] Kolberg L, Raudvere U, Kuzmin I, Adler P, Vilo J and Peterson H. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res 2023; 51: W207-W212.

[25] Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, Zhu J and Haussler D. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol 2020; 38: 675-678.

[26] Therneau TM, Lumley T (original S->R port and R maintainer until 2009), Elizabeth A and Cynthia C. Survival: Survival Analysis. R-Packages. Published August 9, 2022. https://cran.r-project.org/web/packages/survival/index.html.

[27] Kassambara A, Kosinski M, Biecek P and Fabian S. Survminer: Drawing Survival Curves using “ggplot2.” R-Packages. Published March 9, 2021. https://cran.r-project.org/web/packages/survminer/index.html.

[28] Greenblatt MS, Chappuis PO, Bond JP, Hamel N and Foulkes WD. TP53 mutations in breast cancer associated with BRCA1 or BRCA2 germline mutations: distinctive spectrum and structural distribution. Cancer Res 2001; 61: 4092-4097.

[29] Venkatesh H and Monje M. Neuronal activity in ontogeny and oncology. Trends Cancer 2017; 3: 89-112.

[30] Maruyama T, Baba T, Maemoto Y, Hara-Miyauchi C, Hasegawa-Ogawa M, Okano HJ, Enda Y, Matsumoto K, Arimitsu N, Nakao K, Hamamoto H, Sekimizu K, Ohto-Nakanishi T, Nakanishi H, Tokuyama T, Yanagi S, Tagaya M and Tani K. Loss of DDHD2, whose mutation causes spastic paraplegia, promotes reactive oxygen species generation and apoptosis. Cell Death Dis 2018; 9: 797.

[31] Grieb BC and Eischen CM. MTBP and MYC: a dynamic duo in proliferation, cancer, and aging. Biology (Basel) 2022; 11: 881.

[32] Li H, Wu Z, Zhang Y, Lu X and Miao L. Glutamine metabolism genes prognostic signature for stomach adenocarcinoma and immune infiltration: potential biomarkers for predicting overall survival. Front Oncol 2023; 13: 1201297.

[33] Miricescu D, Balan DG, Tulin A, Stiru O, Vacaroiu IA, Mihai DA, Popa CC, Papacocea RI, Enyedi M, Sorin NA, Vatachki G, Georgescu DE, Nica AE and Stefani C. PI3K/AKT/mTOR signalling pathway involvement in renal cell carcinoma pathogenesis (Review). Exp Ther Med 2021; 21: 540.

[34] Mitsudomi T and Yatabe Y. Epidermal growth factor receptor in relation to tumor development: EGFR gene and cancer. FEBS J 2010; 277: 301-308.

[35] Tanaka K, Shimizu K, Kakegawa S, Ohtaki Y, Nagashima T, Kaira K, Horiguchi J, Oyama T and Takeyoshi I. Prognostic significance of aromatase and estrogen receptor beta expression in EGFR wild-type lung adenocarcinoma. Am J Transl Res 2016; 8: 81-97.

Candidate biomarkers tumorigenesis significance analysis

Supplementary Figure 1. Graphs showing gene significance statistics for different targeting categories in each cancer.

Supplementary Table 1. Significant gene ontology terms reported in each cancer

Supplementary Table 2. Gene significance statistics for Onco and TS categories

Supplementary Table 3. Significant genes reported in Onco and TS categories

Supplementary Table 4. MicroRNA-gene targeting pair summary table