Multi-gradient Permutation Survival Analysis Identifies Mitosis and Immune Signatures Steadily Associated with Cancer Patient Prognosis

Xinlei Cai[1], Yi Ye[2], Xiaoping Liu[1], Zhaoyuan Fang[4,5], Luonan Chen[1,2,6], Fei Li[3*], Hongbin Ji[1,2,6*]

[1]Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, 310024, China.

[2]State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China.

[3]Department of Pathology and Frontier Innovation Center, School of Basic Medical Sciences, Fudan University, Shanghai, China.

[4]Zhejiang University-University of Edinburgh Institute, Zhejiang University School of Medicine, Haining，314400, China

[5]The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, 310000, China

[6]School of Life Science and Technology, Shanghai Tech University, Shanghai 200120, China.

Running title: Mitosis and immune signatures in predicting cancer patient prognosis

***Corresponding author:**

Hongbin Ji, State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai, 200031, China, Phone: 86-21-54921108, Fax: 86-21-54921101, E-mail address: hbji@sibcb.ac.cn;

Fei Li, Department of Pathology and Frontier Innovation Center, School of Basic Medical

Sciences, Fudan University, 131 Dong An Road, Shanghai, 200032, China, E-mail address: li_fei@fudan.edu.cn.

Abstract

The inconsistency of the association between genes and cancer prognosis is often attributed to many variables that contribute to patient survival. Whether there exist the Genes Steadily Associated with Prognosis (GEARs) and what their functions are remain largely elusive. We have developed a novel method called "Multi-gradient Permutation Survival Analysis " (MEMORY) to screen the GEARs using RNA-seq data from the TCGA database. Then we employed a network construction approach to identify hub genes from GEARs, and utilized them for cancer classification. In the case of LUAD, the GEARs were found to be related to mitosis. Our analysis suggested that LUAD cell lines carrying PIK3CA mutations exhibit increased drug resistance. For BRCA, the GEARs were related to immunity. The analysis revealed that CDH1 mutation might influence immune infiltration through the EMT process in BRCA. We further explored the prognostic relevance of mitosis and immunity through their respective scores. This study offers significant biological insights into GEARs and highlights their potential as robust prognostic indicators across diverse cancer types.

Keywords: Multi-gradient Permutation Survival Analysis (MEMORY), Gene Steadily Associated with Prognosis (GEAR), Hub Genes, Mitosis, Immune

Introduction

Cancer is one of the leading causes of death worldwide, with over 200 types identified. Despite of huge advancements in the field, accurately predicting cancer patient prognosis still remains a significant challenge[1,2]. Previous studies have highlighted the complexity of this task, which is influenced by various factors including cancer type, clinical stage, therapeutic interventions, nursing care, unexpected comorbidities, and other non-cancer related illnesses that may interplay[3-6]. Furthermore, the gene expression plays a crucial role in predicting cancer patient prognosis, such as *HER2*, *VEGF*, *Ki67*, etc[7-12]. Nevertheless, there is still a need for further improvement in prognostic accuracy to better inform treatment decisions and patient outcomes.

Survival analysis is commonly used to assess the correlation between genes and cancer patient prognosis[13]. However, inconsistent findings are frequently observed, even within the same type of cancer. For instance, studies exploring the correlation between *CCND1* and prognosis in NSCLC reported contrasting results, including positive correlation, inverse correlation, or negligible influence[14-16]. These discrepancies can be attributed to various influential factors, such as differences in sample size, cohort characteristics (including cancer subtypes and tumor staging) and variations in therapy approaches[17]. Such observations raise an important question: whether there exist Genes Steadily Associated with Prognosis (GEARs) that consistently correlate with patient outcomes across different conditions, particularly considering the varying sample sizes used in different studies? Affirmative answers to this question would have significant implication not only for the development of more accurate prognostic models but also for enhancing our understanding of cancer biology.

The Cancer Genome Atlas (TCGA) is a comprehensive cancer genomics project initiated in the United States. It consists of transcriptome data, genomic data, and clinical information pertaining to 33 different cancer types, making it the largest cancer clinical sample database currently available. In this study, we developed a novel method called "Multi-gradient Permutation Survival Analysis" (MEMORY) with the utilization of TCGA RNA-seq data, and accessed the potential existence of GEARs across 15 cancer types, each comprising a cohort

106 of over 200 patients. Furthermore, we also evaluated the prognostic predictive power of these
107 GEARs and explored their potential biological functions in driving cancer progression.

108

Results

Multi-gradient permutation survival analysis identifies GEARs associated with mitosis and immune across multiple cancer types

GEARs are a group of genes consistently and significantly correlate with patient survival, independent of the sample size. To identify these GEARs, we developed a novel method called "Multi-gradient Permutation Survival Analysis " (MEMORY). This method allows us to assess the correlation between a specific gene and cancer patient prognosis using available transcriptomic dataset (**Figure 1A, Table 1**). Initially, we started with a sampling number at 10% of the cohort size initially, and gradually increased the sampling number with each 10% increment, which was analyzed with 1000 permutations. By calculating the statistical probability of each gene's association with patient survival, we can identify a group of GEARs for further analyses.

In this study, we utilized the TCGA datasets for several reasons, including the diversity of cancer types, the availability of gene expression profiles, and patient prognosis information. We set a minimum cohort size of 200 and included 15 eligible cancer types for analysis. These cancer types include bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), cervical squamous cell carcinoma and endocervical Adenocarcinoma (CESC), colon Adenocarcinoma (COAD), kidney renal papillary cell carcinoma (KIRC), brain lower grade glioma (LGG), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell sarcinoma (LUSC), ovarian serous cystadenocarcinoma (OV), pancreatic adenocarcinoma (PAAD), prostate adenocarcinoma (PRAD), stomach adenocarcinoma (STAD), thyroid carcinoma (THCA), and uterine corpus endometrial carcinoma (UCEC). As the number of samples increases, the survival probability of certain genes gradually approaches 1. Once this score reaches 0.8 and remained consistent with further sample gradient increase, the gene is considered a GEAR (**Figure 1B**). Remarkably, we successfully identified a set of GEARs across all 15 cancer types (**Supplementary Figure 1-2 , Table 2)**. The GEAR counts in most cancer types ranged from 100 to1000, with exceptions of CESC, KIRC, LGG, and PAAD with over 1000 GEARs, and THCA with only 22 GEARS (**Table 2)**.

In LUAD, the top 10 GEARs with the highest significance probabilities were *TLE1*, *GNG7*, *ERO1A*, *ANLN*, *DKK1*, *TMEM125*, *S100A16*, *KNL1*, *STEAP1*, and *BEX4*. Most of these genes are known to promote LUAD malignant progression except for S100A16[18-26]. In other cancer types, *BEX4* was identified as a common GEAR in KIRC, LGG, PAAD, and STAD (Table 2). *BEX4* is reported as a proto-oncogene promoting cancer onset and malignant progression in multiple cancers including LUAD, glioblastoma multiforme and oral squamous cell carcinoma [25,27,28]. We also identified the most significant GEARs in individual cancer types, such as *TLL1* (BLCA), *PGK1* (BRCA), *RFXANK* (CESC), *DPP7* (COAD), *VWA8* (KIRC), *SCMH1* (LGG), *HILPDA* (LIHC), *TLE1* (LUAD), *CD151* (LUSC), *ANKRD13A* (OV), *SOCS2* (PAAD), *DRG2* (PRAD), *ADAMTS6* (STAD), *PSMB8* (THCA), *ASS1* (UCEC) (**Table 3).** Many of these genes were previously reported to associate with tumorigenesis[18,29-40]. For example, TLE1 is known as a transcriptional repressor that promotes cell proliferation, migration, and inhibits apoptosis in LUAD [41,42]. Additionally, *PGK1*, a key enzyme in the glycolytic process, has been shown to promote cell proliferation, migration, and invasion in multiple cancers[43,44]. These findings demonstrate the intricate link between the functionality of GEARs and the initiation and progression of cancer.

To gain deep insights into the biological functions of these identified GEARs, we conducted Gene Ontology Enrichment Analysis (GOEA) (**Figure 1C**). Interestingly, we found the mitosis-related pathways were enriched in LICH, LUAD, LGG, and PAAD, and the immune-related pathways were enriched in BRCA and UCEC[45]. Additionally, other cancer types exhibited enrichment in various pathways, such as organic acid metabolism pathway in KIRC, oxidative phosphorylation pathway in CESC, organ development-related pathways in LUSC, neurogenesis-related pathways in BLCA, and sustenance metabolism pathways in THCA. Given the crucial role of mitosis in cancer progression and the significance of the immune system in cancer-host interactions, we specifically focused on the mitosis and immune-related GERAs, which accounted for approximately 40% of all the analyzed cancers.

Identification of hub genes in mitosis and immune-related cancers

To further identify crucial genes within the GEARs, we conducted a significant survival network (SSN) and extracted the higher-ranked edges to construct the core survival network

(CSN) (**Figure 2A, Table 4**). We created 6 CSNs with varying numbers of edges (250, 500, 1000, 2000, 4000, 8000). The top 10 GEARs with highest degree in these networks were defined as hub genes [46]. We then classified the samples using the hub genes derived from these networks and evaluated their predictive power for patient survival. We determined that the optimal number of edges for the network ranged from 1000 to 2000, as it demonstrated the best clustering effect of hub genes (**Figure 2B, Supplementary Figure 3A-O, Supplementary Figure 4A-O**). Furthermore, we conducted GOEA on the hub genes selected from the network with 1000 edges and found that the results were consistent with the GEAR analysis (**Table 5**). Specifically, the hub genes in LIHC, LGG, and LUAD were enriched in mitosis-related pathways, whereas the hub genes in BRCA and UCEC were enriched in immune-related pathways (**Figure 2C**). For instance, the 9 hub genes in LUAD were associated with functions related to mitosis, whereas the 8 hub genes of BRCA were associated with immune-related functions (**Figure 2D-E**).

Furthermore, we conducted survival-dependent analyses on the hub genes of LGG, LIHC, and LUAD. These analyses revealed that mitosis-related hub genes are closely associated with cancer cell viability, especially those hub genes that are correlated with multiple cancer types such as CDC20, TOP2A, BIRC5 and TPX2[47-50] (**Supplementary Figure 5A-C**). The importance of these genes is well established. For instance, TPX2, a hub gene in all three cancers, is known to play a crucial role in normal spindle assembly during mitosis and is essential for cell proliferation[51]. The significance of these hub genes for the survival of cancer cells suggests that the expression levels of these hub genes can be used to screen for inhibitors of tumor cell growth. By integrating the Genomics of Drug Sensitivity in Cancer (GDSC) and Connectivity Map (cMAP) databases, we identified 76 individual compounds that are able to effectively suppress the expression of the 10 hub genes in LUAD cell lines (**Figure 2F, Supplementary Figure 5D-E**). These compounds also significantly inhibit LUAD cell growth and may serve as potential therapeutic agents for the treatment of LUAD.

We then analyzed the LUAD dataset, which had a larger sample size compared to LGG and LIHC. Based on the expression of the hub genes, we classified the LUAD samples into three

subgroups: mitosis low (ML), mitosis medium (MM), and mitosis high (MH) (**Figure 2G, Table 6**). Previous study has reported three classic subgroups of LUAD known as the terminal respiratory unit (TRU), the proximal-proliferative (PP), and the proximal-inflammatory (PI) subgroups[52]. Our analysis revealed that the ML subgroup was primarily enriched with TRU subgroup and a small number of samples from the PP subgroup, but not the PI subgroup. The MH subgroup showed high expression of mitosis-related genes and predominantly encompassed PI and PP subgroups, whereas the MM group exhibited intermediate levels of mitosis-related gene expression (**Figure 2G-H**). This suggests that the new categorization method can help identify new factors that influence patient prognosis (**Supplementary Figure 5F-G**).

Distinct genetic mutation landscapes characterize different clusters of LUAD

The expression of hub genes provided crucial indicators for distinguishing various subgroups of LUAD. However, the underlying mechanisms driving these variations in expression patterns remain unknown. In our subsequent research, we aimed to explore the genomic differences, particularly in terms of gene mutations, among different LUAD subgroups. Initially, we analyzed the variations of genetic mutation landscape of different LUAD subgroups by assessing the tumor mutation burden (TMB). Interestingly, we found significant differences in TMB among the various subgroups of LUAD (**Figure 3A**). Analysis of common driver genes further revealed distinct proportions of tumors with ALK and ROS1 fusions in the favorable prognosis ML and MM subgroups, compared to tumors with mutations in KRAS, EGFR, BRAF, and ERBB2 (**Figure 3B**). Furthermore, we observed unique mutation characteristics in each of these 3 subgroups (**Figure 3C-E**). For example, TP53 mutation, a prevalent tumor suppressor gene mutation, was frequently observed in the MH subgroup with a mutation rate of 73%, compared to mutation rates of 14% in the ML subgroup and 39% in the MM subgroup. Additionally, genes such as *CSMD3*, *RP1L1*, *ZFHX4* exhibited similar trend in mutation frequency across the three subgroups. These findings indicate substantial genomic differences among these three LUAD subgroups based on the hub genes.

We identified gene mutations that showed significant changes through gene dependence analysis (**Figure 4A**). To further the functional implications of these mutations, we enriched them using a pathway system called Nested Systems in Tumors (NeST)[53]. The results revealed notable differences in multiple functional pathways, including androgen receptor, cell cycle, *TP53*, *EGFR*, IL-6, and *PIK3CA*-related pathways (**Figure 4B**). To gain further insights into the functional implications of these different mutations, we assessed the gene dependency of cells with these mutations using the DepMap database. Importantly, we observed that the pathway differences between the MM and MH clusters were primarily associated with *PIK3CA*-related pathways. Previous studies have reported that *PIK3CA* mutation confers resistance in colorectal cancer, lung cancer, and breast cancer[54-56]. Therefore, we aimed to validate the role of PIK3CA mutations in drug resistance using public databases. We further analyzed A549 cells (a lung adenocarcinoma cell line with KRAS mutation) and SW1573 cells (a lung adenocarcinoma cell line with both KRAS and PIK3CA mutations) using their respective GDSC data and identified five potentially effective inhibitors (BMS-345541, Dactinomycin, Epirubicin, Irinotecan, and Topotecan) from the previously screened set of 76 LUAD cell growth inhibitors. Strikingly, SW1573 cells exhibited increased resistance to all these 5 inhibitors when compared to A549 cells (**Figure 4C**). In line with this, clinical data also supports the notion that concurrent *PIK3CA* mutation is a poor prognostic factor for LUAD patients[57]. These findings suggest that *PIK3CA* mutation may contribute to drug resistance in LUAD.

Distinct clusters of BRCA exhibit different immune infiltration landscapes

We further focused on the immune-related cancer types, with BRCA standing out as a representative immune-related cancer. This prompted us to conduct a more comprehensive analysis (**Figure 1C**). We hierarchically grouped the BRCA subgroups into three distinct clusters based on the expression of hub genes: Immune low (IL), Immune medium (IM), and Immune high (IH) (**Figure 5A, Table 7**). Interestingly, all the classifications based on the widely used PAM50 classification contained samples from the IL, IM, and IH subgroups, although there were minor differences in proportions[58] (**Figure 5B, Supplementary Figure 6A-B**). In basal-type breast cancer, which generally has a poor prognosis, a set of samples

251 exhibited high immune infiltration and consequently better outcomes. Conversely, within the
252 HER2 subgroups, known for a better prognosis, there were samples with poor outcomes.
253 Integrating our method with traditional classification may enable a more detailed stratification
254 of breast cancer samples. Mutation analysis revealed distinct patterns of gene mutations
255 among the three subgroups, with significant differences in mutation frequencies of genes such
256 as *TP53*, *CDH1*, and *LRP2* (**Figure 5C-E**). We conducted a comparison of the proportions of
257 immune cells across these three subgroups and observed a strong association between the
258 overall immune cell proportion and the clustering results based on hub genes (**Figure 5F,**
259 **Supplementary Figure 6C-L**). Specifically, the IH subgroup exhibited significantly higher
260 proportions of $CD8^+$ T cells and Treg cells, with mean percentages at 5.0% and 6.0%,
261 respectively, compared to 1.4% and 3.2% in the IM subgroup, and 0.3% and 1.6% in the IL
262 subgroup. These findings suggest the presence of distinct immune responses in these three
263 subgroups, indicating a potential association between genetic mutation pattern and the
264 immune microenvironment landscape.

265 Next, we explored the specific genomic factors influencing immune infiltration in BRCA.
266 Mutation analysis revealed that *CDH1* gene mutation ranked as the second most prevalent
267 genetic alteration in BRCA samples, following TP53 mutation (**Figure 5C-E**). Previous study
268 has indicated that *CDH1* is involved in mechanisms regulating cell-cell adhesions, mobility,
269 and proliferation of epithelial cells[59]. We found that *CDH1*-mutant samples exhibited
270 significantly lower *CDH1* expression compared to *CDH1*-wildtype samples, and *CDH1*
271 expression showed a close correlation with immune cell infiltration (**Figure 5G,**
272 **Supplementary Figure 6M**). These results were consistent with clinical observations of
273 *CDH1* mutation and high immune infiltration in invasive lobular carcinoma of the breast [60].

274 We further investigated how *CDH1* mutation influenced immune infiltration. In BRCA, we
275 observed a significant association between the expression of *CDH1* and the expression of
276 EMT marker genes *VIM* and *TWIST2* (**Supplementary Figure 6N-O**)[61,62]. This suggests that
277 the regulation of *CDH1* is intricately linked with EMT process. Through calculating the EMT
278 score, we observed a positive correlation between the EMT score and the proportion of
279 immune infiltration (**Figure 5H**). Consequently, *CDH1* might influence immune infiltration

through the EMT process.

Mitotic and immune signatures predict patient prognosis at pan-cancer level

We further sought to explore the prognostic relevance of mitosis and immune by calculating their scores and assessing their correlation with patient outcomes at the pan-cancer level. The scores for these biological processes were computed using RNA-seq data from the TCGA database, and the median score was used as a cut-off to categorize patients into different groups. Out of the 33 cancer types analyzed, the prognosis of 17 cancer types showed a significant correlation with at least one of these two pathways (**Figure 6**). Specifically, 10 cancer types were exclusively associated with the mitosis score, 2 cancer types were exclusively associated with the immune score, and 5 cancer types exhibited a correlation with both mitosis and immune scores simultaneously. Overall, the identification of mitosis and immune-related biological processes as significant prognostic factors at the pan-cancer level suggests their potential utility as valuable biomarkers for predicting patient prognosis.

Discussion

Due to the multifaceted nature of variables affecting cancer patient prognosis, it remains uncertain whether there exist a set of genes steadily associated with cancer prognosis , regardless of sample size and other factors. Here, we utilized the MEMORY method to address this question and discovered that all the cancer types have GEARs. We observed significant variation in the number of GEARs among different cancer types, indicative of cancer type-specific patterns. The substantial heterogeneity in driver gene and mortality rates among various cancer types could potentially explain this phenomenon. For example, THCA, known for its low malignancy, displays the fewest genetic expression alterations among all the studied cancer types. This observation could be correlated with the relatively high five-year survival rate in THCA, which exceeds 50% even in advanced stage[63]. In contrast, PRAD, despite of a favorable prognosis similar to THCA, exhibits a significantly higher number of genetic expression alterations. We hypothesize that the positive prognosis in PRAD cases might be largely attributed to early diagnosis, which enables timely treatment for the majority of PRAD patients [64]. This discrepancy between the number of genetic alterations and

prognosis in THCA and PRAD highlights the intricate nature of cancer genetics and emphasizes the importance of personalized considerations in cancer treatment and prognosis evaluation. Furthermore, certain genes are commonly found in the GEARs of various cancer types, indicating their potential significance in cancer development. For instance, BEX4 is present in the GEARs of LUAD, KIRC, LGG, PAAD, and STAD, and known to play oncogenic role in inducing carcinogenic aneuploid transformation via modulating the acetylation of α-tubulin[65,66]. Aneuploidy is a hallmark characteristic of cancer and can lead to alterations in the dosage of oncogenes or tumor suppressor genes, thereby influencing tumor initiation and progression [67-70]. The regulation of aneuploid transformation by BEX4 may represent a common mechanism through which this gene impacts prognosis across different cancer types. Subsequently, we discovered that GEARs across different cancers displayed distinct functional characteristics. Notably, a recurrent theme was the prominence of mitosis-related and immune-related features. Specifically, the GEARs of LGG, LIHC, and LUAD were enriched in mitosis-related pathways, whereas BRCA and UCEC showed the enrichment of immune-related pathways. Mitosis and immune processes have been widely recognized as having a significant impact on patient prognosis[71-75]. However, current clinical guidelines do not yet recommend the use of transcriptomic data to assess scores related to mitotic and immune pathways for predicting patient outcomes, despite the well-established association between these pathways and prognosis in cancer[76-81]. Cancers enriched with mitosis pathways often exhibit heterogeneous tumor growth kinetics across individuals, with tumor size being one of the crucial factors influencing patient survival [82-84]. For instance, in the case of LGG, the growth rate of tumors directly impacts the patient's prognosis due to the primary location in the brain. Consequently, alterations in the expression levels of genes related to mitosis are often indicative of the prognosis of LGG [85]. Cancer types that are enriched in immune-related pathways, such as BRCA and UCEC, are closely associated with deficiencies in DNA mismatch repair (MMR) [86,87]. Cancers enriched in immune-related pathways often exhibit heterogeneous tumor growth kinetics across individuals, where tumor size is closely correlated with patient survival.

To explore the underlying differences between samples associated with distinct mitotic and

337 immune-related pathways, we employed GEAR analysis to identify hub genes for further
338 molecular classification and mutation analysis. Specifically, we focused on LUAD and
339 BRCA, two representative cancer types exhibiting enriched mitosis and immune signatures in
340 GEARs. By classifying the hub gene, we divided LUAD into three subgroups: ML, MM and
341 MH, which displayed significant difference in survival outcomes. Subsequently, we utilized
342 NeST to analyze the mutations within these subgroups. Notably, there were significant
343 differences observed in cell cycle and signal transduction-related pathways among ML, MM
344 and MH subgroups. Of particular importance, the *PIK3CA*-related pathway emerged as a key
345 differentiating pathway between the MH and MM subgroups. Our analyses suggest that
346 LUAD cell line carrying *PIK3CA* mutation may exhibit increased drug resistance. This aligns
347 with previous studies demonstrating that targeting the PI3K pathway can overcome drug
348 resistance[88,89]. Of course, future research is warranted to elucidate the exact functional role of
349 *PIK3CA* mutations in LUAD.

350 To investigate the factors influencing immune infiltration in BRCA, we classified the samples
351 based on hub genes and analyzed the differentially mutated genes. Interestingly, we found that
352 *CDH1* mutation occurred at a high frequency in the IH subgroups. This led us to speculate
353 that *CDH1* plays a crucial role in the regulation of immune infiltration in BRCA.
354 Furthermore, previous studies have reported that *CDH1* inactivation promotes immune
355 infiltration in breast cancer[60]. *CDH1* is a vital gene associated with EMT, and various studies
356 have demonstrated the close relationship between EMT and the tumor immune
357 microenvironment in different cancers[90]. Consistently, our results also supported the
358 correlation between EMT score and immune infiltration in BRCA. These findings suggest
359 that the mutations in *PIK3CA* and *CDH1*, identified through GEAR analysis, have significant
360 impacts on cancer development and hold potential value in improving clinical therapies. This
361 further emphasizes the importance of GEARs in understanding cancer biology and guiding
362 treatment strategies.

363 Lastly, we investigated the prognostic predictive capabilities of the mitosis score and immune
364 score at the pan-cancer level. Surprisingly, we found that approximately half of the cancer
365 types exhibited significant correlations between these two scores and patient prognosis.

Interestingly, even for cancers originating from the same primary location, their correlations with these scores could differ, indicating the potential diversity of mechanisms underlying cancer-related mortality. For instance, both LGG and GBM are brain tumors, but the primary risk factor for patient prognosis in LGG is closely linked to tumor diameter, whereas the main risk factor for GBM patients lies in its high invasiveness and challenge of surgical resection[91,92].

Despite of the strong association between GEAR and patient prognosis, the edges of the CSN constructed based on GEARs was undirected. As a result, the hub genes identified from GEAR in the CSN may primarily serve as stable and effective biological markers. Additionally, through multi-omics analysis, we obtained some functional mutations, but the therapeutic significance of these mutations remains to be elucidated. For example, further study is needed to understand the exact role of PIK3CA mutations in promoting tumor cell proliferation and drug resistance. Similarly, the association of *CDH1* mutations with the infiltration of multiple immune cell types also warrants additional experimental investigation. In our future studies, we plan to utilize protein-protein interaction networks and pathway databases to construct a novel network based on the CSN. This approach will allow us to directly screen genes from GEARs that could potentially serve as therapeutic targets. Undoubtedly, future efforts are still required to utilize protein-protein interaction networks and pathway databases to construct a new network based on CSN.

In conclusion, our study utilized the MEMORY algorithm to identify GEARs in 15 cancer types, highlighting the significance of mitosis and immunity in cancer prognosis. Our findings demonstrate that GEARs possess substantial biological significance beyond their role as prognostic biomarkers. This study provides valuable guidance for establishing standards for survival analysis evaluation and holds potential for the development of novel therapeutic strategies.

Methods

Datasets

The gene expression profiles and clinical information of 33 cancers were obtained from the TCGA database and downloaded by the GDC data website (https: //portal.gdc.cancer.gov/). All gene expression data and survival data were integrated and normalized.

Multi-gradient permutation survival analysis

Gene expression data of cancer patients were obtained from the RNA-seq expression matrix of TCGA after TPM normalization. We randomly sampled the gene expression data according to the gradient. The sampling strategy was as follows: Ten gradient increases in sample size were pre-set, ranging from approximately 10% to about 100%, with intervals of 10%. Random samples were taken from total samples of each cancer 1000 times, based on each pre-set sample size. Multiple sampling matrix was obtained after the sampling strategy was performed in each gradient. The survival analysis was performed for every sampling matrix by “survival” and “survminer”. “survival” and “survminer” are R packages for survival analysis. The survival analysis method was as follows according to the median expression value of every gene, every sampling matrix was divided into a high and low expression group and performed survival analysis by “survival” and “survminer”. We used 1 for significant survival analysis results ($P < 0.05$) and 0 for non-significant results ($P > 0.05$). The survival analysis matrices were obtained after these processes.

Construction of core survival network

The survival analysis matrices were integrated to a significant-probability matrix by a formula:

$$A_{ij} = \frac{\sum_{n=1}^{1000} k_{j\times i\times n}}{1000} \qquad (1)$$

where the A_{ij} is the value from row i and column j in the significant-probability matrix. We defined the sampling size k_j reached saturation when the max value of column j was equal to 1 in a significant-probability matrix. The least value of k_j was selected, and the genes with their corresponding A_{ij} greater than 0.8 were extracted as GEAR.

We also defined survival analysis similarity (*SAS*) as the similarity of the effect on patient prognosis in two genes. The *SAS* computational formula is as follows:

$$SAS = \frac{A \cap B}{A \cup B - 2(A \cap B) + 1} \qquad (2)$$

Where *A* and *B* are results of *1000* survival analyses at the first sampling gradient, which are sequences of *0* and *1* of length *1000*. $A \cap B$ means the number of events that are simultaneously 1 in both A and B. $A \cup B$ means the size of the union of *A* and *B*. The *SAS* of significant genes with total genes was calculated, and the significant survival network was constructed. Then some *SAS* values at the top of the rankings were extracted, and the *SAS* was visualized to a network by Cytoscape[93]. The network was named core survival network (CSN). The degree of each node was calculated, and the nodes with the top10 degrees were defined hub genes. The effect of molecular classification by hub genes is indicated that 1000 to 2000 was a range that the result of molecular classification was best. One thousand was chosen as the number of *SAS*. Therefore, one thousand was selected as the number of *SAS* values for constructed significant survival network.

Gene ontology enrichment analysis

Gene ontology has been used to classify genes based on functions. The gene functions were divided three types, includes molecular function (MF), biological process (BP), and cellular component (CC). ClusterProfiler is an R package for gene set enrichment analysis[94]. The gene ontology enrichment analysis (GOEA) was processed in significant gene sets and hub genes by ClusterProfiler.

Hub gene classification

Tumor samples were genotyped. First, the expression matrix of hub genes corresponding to cancer types was extracted from RNA-seq data. The matrix M is log-transformed after adding a pseudo count of 1: $M' = log_2(M + 1)$. Finally, M' was grouped by hierarchical clustering.

Calculation of tumor mutation burden (TMB) and differential mutation

The TMB and differential mutation gene analysis were carried out to explore the differences between different genotyped samples at the genomic level. The data are from the gene

mutation data of TCGA tumor samples, and the grouping information is from the hub gene hierarchical clustering results. Maftools is an R package for the analysis of somatic variant data, which can export results in form of charts and graph[95]. The calculation of TMB and difference mutation analysis were used by the maftools.

Quantifying the effect of gene mutations for tumor cell viability

Gene dependence refers to the extent to which genes are essential for cell proliferation and survival. The Cancer Dependency Map (Depmap, https://depmap.org/portal/) database provides genome-wide gene dependence data for large number of tumor cell lines[96]. Gene mutation data of cell lines were obtained from the CCLE database (https://sites.broadinstitute.org/ccle). The genes appearing in the gene mutation data of cell lines were extracted and sorted into mutation list. Each gene in the list was then analyzed for survival-dependent differences. For each gene in the mutation list, we compared the gene dependence score (S) between cell lines with and without the gene mutation. Specifically, cell lines from the CCLE mutation dataset harboring a mutation in a given gene were classified into one group, while cell lines without the mutation constituted the control group. S for the mutation (S_m) and wild-type (S_{wt}) cell lines were then extracted from the DepMap database.

A two-sample t-test was used to assess the statistical significance of differences between S_m and S_{wt}. The test yielded a P-value, with an alpha level set at 0.05, to determine the significance of the difference in means. D_m was defined as gene mutation dependence, the computational formula was as follows:

$$D_m = |\bar{s}_m - \bar{s}_{wt}| \qquad (3)$$

Subsequently, we conducted a standardization assessment of D_m, employing the following formula:

$$S_{dm} = \frac{2(\bar{s}_m - \bar{s}_{wt})}{\bar{s}_m + \bar{s}_{wt}} \qquad (4)$$

In this study, we categorized mutations identified in cell lines into three groups based on their impact on gene dependency. Functional mutations refer to the mutations that exhibit

significant changes in gene dependency scores, with a P-value < 0.05 and S_{dm} > 0.1. Non-functional mutations were those without significant changes in gene dependency scores compared to wildtype. Subtype-associated functional mutations were a subset of functional mutations that show statistically significant differences in occurrence frequency across different LUAD subtypes.

Identification of drugs downregulating hub gene expression

The IC50 data of 76 compounds of LUAD cell lines were obtained from Genomics of Drug Sensitivity in Cancer (GDSC) database[97]. The hub gene expression data of A549 cell lines after treatment with 76 compounds was obtained from Connectivity MAP (cMAP, https://clue.io/) database, and the data were grouped by hierarchical clustering[98]. The drugs which were presented only in the hub gene expression suppression group were considered to effective against the LUAD cell.

Immune infiltration analysis

Immune infiltration data was calculated by quantiseq method[99]. Immunedeconv was an R package for unified access to computational methods for estimating immune Cell fractions from bulk RNA-seq data[100]. The data were from the gene expression data of TCGA tumor samples after TPM was standardized. Then the immune cell fractions of tumor samples were calculated by quantiseq method invoking immunedeconv.

Biological function score calculation

GSVA is an R package for calculate biological function score based on a single sample[101]. The data was from TCGA in the calculation process, and the software package was GSVA. The gene set of the immune score was from other study (**Table 8**) [99]. The gene sets that calculated the mitosis score were obtained from gene ontology (GO:0140014).

Acknowledgment

This work was supported by the National Key Research and Development Program of China (grants 2022YFA1103900 to H.J., F.L.; 2020YFA0803300 to H.J.); the National Natural Science Foundation of China (grants 82341002 to H.J., 32293192 to H.J., 82030083 to H.J.,

82173340 to L.H., 82273400 to Y.J., 32100593 to X.T., 82203306 to W.H., 82372763 to X.W., 82303916 to C.G., 82303039 to Z.Q., 82141101 to F.L., 2022hwyq16 to F.L., 82372794 to F.L.); the Basic Frontier Scientific Research Program of Chinese Academy of Science (ZDBS-LY-SM006 to H.J.); the Innovative research team of high-level local universities in Shanghai (SSMU-ZLCX20180500 to H.J.); Science and Technology Commission of Shanghai Municipality (21ZR1470300 to L.H.).

Author Contributions

H.J. and F.L. conceived the idea and designed the analysis. X.C. performed all bioinformatics analyses. Y.Y., X.L. and L.C. provided technical assistance and helpful comments. H.J., F.L. and X.C. wrote the manuscript. All authors approved the final version.

Disclosure of conflicts of interest

The authors declare that there is no conflict of interests.

References

1. Tomczak, K., Czerwinska, P., and Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) *19*, A68-77. 10.5114/wo.2014.47136.
2. Cancer Genome Atlas Research, N., 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. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat Genet *45*, 1113-1120. 10.1038/ng.2764.
3. Liu, J., Lichtenberg, T., Hoadley, K.A., Poisson, L.M., Lazar, A.J., Cherniack, A.D., Kovatich, A.J., Benz, C.C., Levine, D.A., Lee, A.V., et al. (2018). An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. Cell *173*, 400-416 e411. 10.1016/j.cell.2018.02.052.
4. Smeltzer, M.P., Faris, N.R., Ray, M.A., and Osarogiagbon, R.U. (2018). Association of Pathologic Nodal Staging Quality With Survival Among Patients With Non-Small Cell Lung Cancer After Resection With Curative Intent. JAMA Oncol *4*, 80-87. 10.1001/jamaoncol.2017.2993.
5. Miyauchi, E., Morita, S., Nakamura, A., Hosomi, Y., Watanabe, K., Ikeda, S., Seike, M., Fujita, Y., Minato, K., Ko, R., et al. (2022). Updated Analysis of NEJ009: Gefitinib-Alone Versus Gefitinib Plus Chemotherapy for Non-Small-Cell Lung Cancer With Mutated EGFR. J Clin Oncol, JCO2102911. 10.1200/JCO.21.02911.

6. Teh, S.H., Uong, S., Lin, T.Y., Shiraga, S., Li, Y., Gong, I.Y., Herrinton, L.J., and Li, R.A. (2021). Clinical Outcomes Following Regionalization of Gastric Cancer Care in a US Integrated Health Care System. J Clin Oncol *39*, 3364-3376. 10.1200/JCO.21.00480.
7. Gimotty, P.A., Van Belle, P., Elder, D.E., Murry, T., Montone, K.T., Xu, X., Hotz, S., Raines, S., Ming, M.E., Wahl, P., and Guerry, D. (2005). Biologic and prognostic significance of dermal Ki67 expression, mitoses, and tumorigenicity in thin invasive cutaneous melanoma. J Clin Oncol *23*, 8048-8056. 10.1200/JCO.2005.02.0735.
8. Wu, Y., Lv, M., Qian, T., and Shen, Y. (2021). Correlation analysis of Ki67 and CK7 expression with clinical characteristics and prognosis of postoperative cervical adenocarcinoma patients. Ann Palliat Med *10*, 9544-9552. 10.21037/apm-21-1974.
9. Mouabbi, J.A., Singareeka Raghavendra, A., Bassett, R.L., Jr., Hassan, A., Tripathy, D., and Layman, R.M. (2023). Survival Outcomes in Patients With Hormone Receptor-Positive Metastatic Breast Cancer With Low or No ERBB2 Expression Treated With Targeted Therapies Plus Endocrine Therapy. JAMA Netw Open *6*, e2313017. 10.1001/jamanetworkopen.2023.13017.
10. Kang, Y.J., Oh, S.J., Bae, S.Y., Kim, E.K., Lee, Y.J., Park, E.H., Jeong, J., Park, H.K., Suh, Y.J., and Kim, Y.S. (2023). Predictive biological factors for late survival in patients with HER2-positive breast cancer. Sci Rep *13*, 11008. 10.1038/s41598-023-38200-y.
11. Cai, C., Wang, X., Fu, Q., and Chen, A. (2022). The VEGF expression associated with prognosis in patients with intrahepatic cholangiocarcinoma: a systematic review and meta-analysis. World J Surg Oncol *20*, 40. 10.1186/s12957-022-02511-7.
12. Zhang, C., Wang, L., Xiong, C., Zhao, R., Liang, H., and Luo, X. (2021). The role of vascular endothelial growth factor as a prognostic and clinicopathological marker in osteosarcoma: a systematic review and meta-analysis. J Orthop Surg Res *16*, 738. 10.1186/s13018-021-02888-3.
13. Lai, Y., OuYang, G., Sheng, L., Zhang, Y., Lai, B., and Zhou, M. (2021). Novel prognostic genes and subclasses of acute myeloid leukemia revealed by survival analysis of gene expression data. BMC Med Genomics *14*, 39. 10.1186/s12920-021-00888-0.
14. Anton, R.C., Coffey, D.M., Gondo, M.M., Stephenson, M.A., Brown, R.W., and Cagle, P.T. (2000). The expression of cyclins D1 and E in predicting short-term survival in squamous cell carcinoma of the lung. Mod Pathol *13*, 1167-1172. 10.1038/modpathol.3880215.
15. Esposito, V., Baldi, A., De Luca, A., Tonini, G., Vincenzi, B., Santini, D., Persichetti, P., Mancini, A., Citro, G., Baldi, F., et al. (2005). Cell cycle related proteins as prognostic parameters in radically resected non-small cell lung cancer. J Clin Pathol *58*, 734-739. 10.1136/jcp.2004.023531.
16. Dworakowska, D., Jassem, E., Jassem, J., Boltze, C., Wiedorn, K.H., Dworakowski, R., Skokowski, J., Jaskiewicz, K., and Czestochowska, E. (2005). Prognostic value of cyclin D1 overexpression in correlation with pRb and p53 status in non-small cell lung cancer (NSCLC). J Cancer Res Clin Oncol *131*, 479-485. 10.1007/s00432-004-0661-9.
17. Freedman, L.S. (1982). Tables of the number of patients required in clinical trials using the logrank test. Stat Med *1*, 121-129. 10.1002/sim.4780010204.
18. Yao, X., Ireland, S.K., Pham, T., Temple, B., Chen, R., Raj, M.H., and Biliran, H. (2014). TLE1 promotes EMT in A549 lung cancer cells through suppression of E-cadherin. Biochem Biophys Res Commun *455*, 277-284. 10.1016/j.bbrc.2014.11.007.

575 19. Zheng, H., Tian, H., Yu, X., Ren, P., and Yang, Q. (2021). G protein gamma 7 suppresses
576 progression of lung adenocarcinoma by inhibiting E2F transcription factor 1. Int J Biol
577 Macromol *182*, 858-865. 10.1016/j.ijbiomac.2021.04.082.
578 20. Chen, G., Wang, Q., and Wang, K. (2022). MicroRNA-218-5p affects lung adenocarcinoma
579 progression through targeting endoplasmic reticulum oxidoreductase 1 alpha. Bioengineered
580 *13*, 10061-10070. 10.1080/21655979.2022.2063537.
581 21. Xu, J., Zheng, H., Yuan, S., Zhou, B., Zhao, W., Pan, Y., and Qi, D. (2019). Overexpression of
582 ANLN in lung adenocarcinoma is associated with metastasis. Thorac Cancer *10*, 1702-1709.
583 10.1111/1759-7714.13135.
584 22. Zhang, P., Li, S., Lv, C., Si, J., Xiong, Y., Ding, L., Ma, Y., and Yang, Y. (2018). BPI-9016M, a
585 c-Met inhibitor, suppresses tumor cell growth, migration and invasion of lung adenocarcinoma
586 via miR203-DKK1. Theranostics *8*, 5890-5902. 10.7150/thno.27667.
587 23. Wang, T., Zhai, R., Lv, X., Wang, K., and Xu, J. (2020). LINC02418 promotes malignant
588 behaviors in lung adenocarcinoma cells by sponging miR-4677-3p to upregulate KNL1
589 expression. BMC Pulm Med *20*, 217. 10.1186/s12890-020-01229-0.
590 24. Huo, S.F., Shang, W.L., Yu, M., Ren, X.P., Wen, H.X., Chai, C.Y., Sun, L., Hui, K., Liu, L.H.,
591 Wei, S.H., et al. (2020). STEAP1 facilitates metastasis and epithelial-mesenchymal transition
592 of lung adenocarcinoma via the JAK2/STAT3 signaling pathway. Biosci Rep *40*.
593 10.1042/BSR20193169.
594 25. Zhao, Z., Li, J., Tan, F., Gao, S., and He, J. (2018). mTOR up-regulation of BEX4 promotes
595 lung adenocarcinoma cell proliferation by potentiating OCT4. Biochem Biophys Res
596 Commun *500*, 302-309. 10.1016/j.bbrc.2018.04.064.
597 26. Fan, T., Liu, Y., Liu, H., Wang, L., Tian, H., Zheng, Y., Zheng, B., Xue, L., Li, C., and He, J.
598 (2022). Transmembrane Protein-Based Risk Model and H3K4me3 Modification
599 Characteristics in Lung Adenocarcinoma. Frontiers in Oncology *12*.
600 10.3389/fonc.2022.828814.
601 27. Lee, S., Kang, H., Shin, E., Jeon, J., Youn, H., and Youn, B. (2021). BEX1 and BEX4 Induce
602 GBM Progression through Regulation of Actin Polymerization and Activation of YAP/TAZ
603 Signaling. Int J Mol Sci *22*. 10.3390/ijms22189845.
604 28. Gao, W., Li, J.Z., Chen, S.Q., Chu, C.Y., Chan, J.Y., and Wong, T.S. (2016). Decreased brain-
605 expressed X-linked 4 (BEX4) expression promotes growth of oral squamous cell carcinoma. J
606 Exp Clin Cancer Res *35*, 92. 10.1186/s13046-016-0355-6.
607 29. He, Y., Luo, Y., Zhang, D., Wang, X., Zhang, P., Li, H., Ejaz, S., and Liang, S. (2019). PGK1-
608 mediated cancer progression and drug resistance. Am J Cancer Res *9*, 2280-2302.
609 30. Matsuura, K., Sawai, H., Ikeo, K., Ogawa, S., Iio, E., Isogawa, M., Shimada, N., Komori, A.,
610 Toyoda, H., Kumada, T., et al. (2017). Genome-Wide Association Study Identifies TLL1
611 Variant Associated With Development of Hepatocellular Carcinoma After Eradication of
612 Hepatitis C Virus Infection. Gastroenterology *152*, 1383-1394. 10.1053/j.gastro.2017.01.041.
613 31. Poplawski, P., Alseekh, S., Jankowska, U., Skupien-Rabian, B., Iwanicka-Nowicka, R.,
614 Kossowska, H., Fogtman, A., Rybicka, B., Boguslawska, J., Adamiok-Ostrowska, A., et al.
615 (2023). Coordinated reprogramming of renal cancer transcriptome, metabolome and
616 secretome associates with immune tumor infiltration. Cancer Cell Int *23*, 2. 10.1186/s12935-
617 022-02845-y.
618 32. Guo, M., Li, F., Zhao, L., Fang, Z., Yu, H., Songyang, Z., and Xiong, Y. (2023). Pan-cancer

619 investigation of C-to-U editing reveals its important role in cancer development and new
620 targets for cancer treatment. Front Oncol *13*, 1097667. 10.3389/fonc.2023.1097667.
621 33. Povero, D., Johnson, S.M., and Liu, J. (2020). Hypoxia, hypoxia-inducible gene 2
622 (HIG2)/HILPDA, and intracellular lipolysis in cancer. Cancer Lett *493*, 71-79.
623 10.1016/j.canlet.2020.06.013.
624 34. Erfani, S., Hua, H., Pan, Y., Zhou, B.P., and Yang, X.H. (2021). The Context-Dependent
625 Impact of Integrin-Associated CD151 and Other Tetraspanins on Cancer Development and
626 Progression: A Class of Versatile Mediators of Cellular Function and Signaling, Tumorigenesis
627 and Metastasis. Cancers (Basel) *13*. 10.3390/cancers13092005.
628 35. Won, M., Park, K.A., Kim, S., Ju, E., Ko, Y., Yoo, H., Ro, H., Lee, J., Oh, J., Lee, E.G., et al.
629 (2022). ANKRD13a controls early cell-death checkpoint by interacting with RIP1 independent
630 of NF-kappaB. Cell Death Differ *29*, 1152-1163. 10.1038/s41418-021-00906-9.
631 36. Xu, J., Chen, Q., Tian, K., Liang, R., Chen, T., Gong, A., Mathy, N.W., Yu, T., and Chen, X.
632 (2020). m6A methyltransferase METTL3 maintains colon cancer tumorigenicity by
633 suppressing SOCS2 to promote cell proliferation. Oncol Rep *44*, 973-986.
634 10.3892/or.2020.7665.
635 37. Xu, C., Li, H., Zhang, L., Jia, T., Duan, L., and Lu, C. (2016). MicroRNA-1915-3p prevents
636 the apoptosis of lung cancer cells by downregulating DRG2 and PBX2. Mol Med Rep *13*,
637 505-512. 10.3892/mmr.2015.4565.
638 38. Mead, T.J. (2022). ADAMTS6: Emerging roles in cardiovascular, musculoskeletal and cancer
639 biology. Front Mol Biosci *9*, 1023511. 10.3389/fmolb.2022.1023511.
640 39. Yang, B.Y., Song, J.W., Sun, H.Z., Xing, J.C., Yang, Z.H., Wei, C.Y., Xu, T.Y., Yu, Z.N.,
641 Zhang, Y.N., Wang, Y.F., et al. (2018). PSMB8 regulates glioma cell migration, proliferation,
642 and apoptosis through modulating ERK1/2 and PI3K/AKT signaling pathways. Biomed
643 Pharmacother *100*, 205-212. 10.1016/j.biopha.2018.01.170.
644 40. Keshet, R., Lee, J.S., Adler, L., Iraqi, M., Ariav, Y., Lim, L.Q.J., Lerner, S., Rabinovich, S.,
645 Oren, R., Katzir, R., et al. (2020). Targeting purine synthesis in ASS1-expressing tumors
646 enhances the response to immune checkpoint inhibitors. Nat Cancer *1*, 894-908.
647 10.1038/s43018-020-0106-7.
648 41. Palaparti, A., Baratz, A., and Stifani, S. (1997). The Groucho/Transducin-like Enhancer of
649 split Transcriptional Repressors Interact with the Genetically Defined Amino-terminal
650 Silencing Domain of Histone H3. Journal of Biological Chemistry *272*, 26604-26610.
651 10.1074/jbc.272.42.26604.
652 42. Lin, W., Wang, J., Ge, J., Zhou, R., Hu, Y., Xiao, L., Peng, Q., and Zheng, Z. (2023). The
653 activity of cuproptosis pathway calculated by AUCell algorithm was employed to construct
654 cuproptosis landscape in lung adenocarcinoma. Discover Oncology *14*. 10.1007/s12672-023-
655 00755-7.
656 43. Li, X., Jiang, Y., Meisenhelder, J., Yang, W., Hawke, D.H., Zheng, Y., Xia, Y., Aldape, K., He,
657 J., Hunter, T., et al. (2016). Mitochondria-Translocated PGK1 Functions as a Protein Kinase to
658 Coordinate Glycolysis and the TCA Cycle in Tumorigenesis. Mol Cell *61*, 705-719.
659 10.1016/j.molcel.2016.02.009.
660 44. Fu, Q., and Yu, Z. (2020). Phosphoglycerate kinase 1 (PGK1) in cancer: A promising target for
661 diagnosis and therapy. Life Sci *256*, 117863. 10.1016/j.lfs.2020.117863.
662 45. Waldman, A.D., Fritz, J.M., and Lenardo, M.J. (2020). A guide to cancer immunotherapy:

663 from T cell basic science to clinical practice. Nat Rev Immunol *20*, 651-668. 10.1038/s41577-
664 020-0306-5.
665 46. Zhou, Z., Li, Y., Hao, H., Wang, Y., Zhou, Z., Wang, Z., and Chu, X. (2019). Screening Hub
666 Genes as Prognostic Biomarkers of Hepatocellular Carcinoma by Bioinformatics Analysis.
667 Cell Transplant *28*, 76S-86S. 10.1177/0963689719893950.
668 47. Hwang, L.H., Lau, L.F., Smith, D.L., Mistrot, C.A., Hardwick, K.G., Hwang, E.S., Amon, A.,
669 and Murray, A.W. (1998). Budding yeast Cdc20: a target of the spindle checkpoint. Science
670 *279*, 1041-1044. 10.1126/science.279.5353.1041.
671 48. Uuskula-Reimand, L., and Wilson, M.D. (2022). Untangling the roles of TOP2A and TOP2B
672 in transcription and cancer. Sci Adv *8*, eadd4920. 10.1126/sciadv.add4920.
673 49. Lamers, F., van der Ploeg, I., Schild, L., Ebus, M.E., Koster, J., Hansen, B.R., Koch, T.,
674 Versteeg, R., Caron, H.N., and Molenaar, J.J. (2011). Knockdown of survivin (BIRC5) causes
675 apoptosis in neuroblastoma via mitotic catastrophe. Endocr Relat Cancer *18*, 657-668.
676 10.1530/ERC-11-0207.
677 50. Wittmann, T., Wilm, M., Karsenti, E., and Vernos, I. (2000). TPX2, A novel xenopus MAP
678 involved in spindle pole organization. J Cell Biol *149*, 1405-1418. 10.1083/jcb.149.7.1405.
679 51. Kufer, T.A., Sillje, H.H., Korner, R., Gruss, O.J., Meraldi, P., and Nigg, E.A. (2002). Human
680 TPX2 is required for targeting Aurora-A kinase to the spindle. J Cell Biol *158*, 617-623.
681 10.1083/jcb.200204155.
682 52. Cancer Genome Atlas Research, N. (2014). Comprehensive molecular profiling of lung
683 adenocarcinoma. Nature *511*, 543-550. 10.1038/nature13385.
684 53. Zheng, F., Kelly, M.R., Ramms, D.J., Heintschel, M.L., Tao, K., Tutuncuoglu, B., Lee, J.J.,
685 Ono, K., Foussard, H., Chen, M., et al. (2021). Interpretation of cancer mutations using a
686 multiscale map of protein systems. Science *374*, eabf3067. 10.1126/science.abf3067.
687 54. Wang, Q., Shi, Y.L., Zhou, K., Wang, L.L., Yan, Z.X., Liu, Y.L., Xu, L.L., Zhao, S.W., Chu,
688 H.L., Shi, T.T., et al. (2018). PIK3CA mutations confer resistance to first-line chemotherapy in
689 colorectal cancer. Cell Death Dis *9*, 739. 10.1038/s41419-018-0776-6.
690 55. Shibata, T., Kokubu, A., Tsuta, K., and Hirohashi, S. (2009). Oncogenic mutation of PIK3CA
691 in small cell lung carcinoma: a potential therapeutic target pathway for chemotherapy-resistant
692 lung cancer. Cancer Lett *283*, 203-211. 10.1016/j.canlet.2009.03.038.
693 56. Hanker, A.B., Pfefferle, A.D., Balko, J.M., Kuba, M.G., Young, C.D., Sanchez, V., Sutton,
694 C.R., Cheng, H., Perou, C.M., Zhao, J.J., et al. (2013). Mutant PIK3CA accelerates HER2-
695 driven transgenic mammary tumors and induces resistance to combinations of anti-HER2
696 therapies. Proc Natl Acad Sci U S A *110*, 14372-14377. 10.1073/pnas.1303204110.
697 57. Eng, J., Woo, K.M., Sima, C.S., Plodkowski, A., Hellmann, M.D., Chaft, J.E., Kris, M.G.,
698 Arcila, M.E., Ladanyi, M., and Drilon, A. (2015). Impact of Concurrent PIK3CA Mutations on
699 Response to EGFR Tyrosine Kinase Inhibition in EGFR-Mutant Lung Cancers and on
700 Prognosis in Oncogene-Driven Lung Adenocarcinomas. J Thorac Oncol *10*, 1713-1719.
701 10.1097/JTO.0000000000000671.
702 58. Thorsson, V., Gibbs, D.L., Brown, S.D., Wolf, D., Bortone, D.S., Ou Yang, T.H., Porta-Pardo,
703 E., Gao, G.F., Plaisier, C.L., Eddy, J.A., et al. (2018). The Immune Landscape of Cancer.
704 Immunity *48*, 812-830 e814. 10.1016/j.immuni.2018.03.023.
705 59. Meigs, T.E., Fedor-Chaiken, M., Kaplan, D.D., Brackenbury, R., and Casey, P.J. (2002).
706 Galpha12 and Galpha13 negatively regulate the adhesive functions of cadherin. J Biol Chem

277, 24594-24600. 10.1074/jbc.M201984200.

60. An, Y., Adams, J.R., Hollern, D.P., Zhao, A., Chang, S.G., Gams, M.S., Chung, P.E.D., He, X., Jangra, R., Shah, J.S., et al. (2018). Cdh1 and Pik3ca Mutations Cooperate to Induce Immune-Related Invasive Lobular Carcinoma of the Breast. Cell Rep *25*, 702-714 e706. 10.1016/j.celrep.2018.09.056.

61. Zeisberg, M., and Neilson, E.G. (2009). Biomarkers for epithelial-mesenchymal transitions. J Clin Invest *119*, 1429-1437. 10.1172/JCI36183.

62. Yang, J., Mani, S.A., Donaher, J.L., Ramaswamy, S., Itzykson, R.A., Come, C., Savagner, P., Gitelman, I., Richardson, A., and Weinberg, R.A. (2004). Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell *117*, 927-939. 10.1016/j.cell.2004.06.006.

63. Egner, J.R. (2010). AJCC cancer staging manual. Jama *304*, 1726-1727.

64. Rebello, R.J., Oing, C., Knudsen, K.E., Loeb, S., Johnson, D.C., Reiter, R.E., Gillessen, S., Van der Kwast, T., and Bristow, R.G. (2021). Prostate cancer. Nat Rev Dis Primers *7*, 9. 10.1038/s41572-020-00243-0.

65. Lee, J.K., Lee, J., Go, H., Lee, C.G., Kim, S., Kim, H.S., Cho, H., Choi, K.S., Ha, G.H., and Lee, C.W. (2016). Oncogenic microtubule hyperacetylation through BEX4-mediated sirtuin 2 inhibition. Cell Death Dis *7*, e2336. 10.1038/cddis.2016.240.

66. Lee, J.K., Ha, G.H., Kim, H.S., and Lee, C.W. (2018). Oncogenic potential of BEX4 is conferred by Polo-like kinase 1-mediated phosphorylation. Exp Mol Med *50*, 1-12. 10.1038/s12276-018-0168-0.

67. Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M., et al. (2010). The landscape of somatic copy-number alteration across human cancers. Nature *463*, 899-905. 10.1038/nature08822.

68. Davoli, T., Xu, A.W., Mengwasser, K.E., Sack, L.M., Yoon, J.C., Park, P.J., and Elledge, S.J. (2013). Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. Cell *155*, 948-962. 10.1016/j.cell.2013.10.011.

69. Girish, V., Lakhani, A.A., Thompson, S.L., Scaduto, C.M., Brown, L.M., Hagenson, R.A., Sausville, E.L., Mendelson, B.E., Kandikuppa, P.K., Lukow, D.A., et al. (2023). Oncogene-like addiction to aneuploidy in human cancers. Science *381*, eadg4521. 10.1126/science.adg4521.

70. Bosco, N., Goldberg, A., Zhao, X., Mays, J.C., Cheng, P., Johnson, A.F., Bianchi, J.J., Toscani, C., Di Tommaso, E., Katsnelson, L., et al. (2023). KaryoCreate: A CRISPR-based technology to study chromosome-specific aneuploidy by targeting human centromeres. Cell *186*, 1985-2001 e1919. 10.1016/j.cell.2023.03.029.

71. Ha, S.Y., Choi, M., Lee, T., and Park, C.K. (2016). The Prognostic Role of Mitotic Index in Hepatocellular Carcinoma Patients after Curative Hepatectomy. Cancer Res Treat *48*, 180-189. 10.4143/crt.2014.321.

72. Zhou, C., Wang, Y., Liu, X., Liang, Y., Fan, Z., Jiang, T., Wang, Y., and Wang, L. (2018). Molecular profiles for insular low-grade gliomas with putamen involvement. J Neurooncol *138*, 659-666. 10.1007/s11060-018-2837-1.

73. Makinen, J.M., Laitakari, K., Johnson, S., Makitaro, R., Bloigu, R., Paakko, P., Lappi-Blanco, E., and Kaarteenaho, R. (2017). Histological features of malignancy correlate with growth patterns and patient outcome in lung adenocarcinoma. Histopathology *71*, 425-436.

751 10.1111/his.13236.
752 74. Byrne, A., Savas, P., Sant, S., Li, R., Virassamy, B., Luen, S.J., Beavis, P.A., Mackay, L.K.,
753 Neeson, P.J., and Loi, S. (2020). Tissue-resident memory T cells in breast cancer control and
754 immunotherapy responses. Nat Rev Clin Oncol *17*, 341-348. 10.1038/s41571-020-0333-y.
755 75. Liu, W., Sun, L., Zhang, J., Song, W., Li, M., and Wang, H. (2021). The landscape and
756 prognostic value of immune characteristics in uterine corpus endometrial cancer. Biosci Rep
757 *41*. 10.1042/BSR20202321.
758 76. Ducreux, M., Cuhna, A.S., Caramella, C., Hollebecque, A., Burtin, P., Goere, D., Seufferlein,
759 T., Haustermans, K., Van Laethem, J.L., Conroy, T., et al. (2015). Cancer of the pancreas:
760 ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol *26*
761 *Suppl 5*, v56-68. 10.1093/annonc/mdv295.
762 77. Cardoso, F., Kyriakides, S., Ohno, S., Penault-Llorca, F., Poortmans, P., Rubio, I.T.,
763 Zackrisson, S., Senkus, E., and clinicalguidelines@esmo.org, E.G.C.E.a. (2019). Early breast
764 cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-updagger. Ann
765 Oncol *30*, 1194-1220. 10.1093/annonc/mdz173.
766 78. Oaknin, A., Bosse, T.J., Creutzberg, C.L., Giornelli, G., Harter, P., Joly, F., Lorusso, D., Marth,
767 C., Makker, V., Mirza, M.R., et al. (2022). Endometrial cancer: ESMO Clinical Practice
768 Guideline for diagnosis, treatment and follow-up. Ann Oncol *33*, 860-877.
769 10.1016/j.annonc.2022.05.009.
770 79. Vogel, A., Cervantes, A., Chau, I., Daniele, B., Llovet, J.M., Meyer, T., Nault, J.C., Neumann,
771 U., Ricke, J., Sangro, B., et al. (2018). Hepatocellular carcinoma: ESMO Clinical Practice
772 Guidelines for diagnosis, treatment and follow-up. Ann Oncol *29*, iv238-iv255.
773 10.1093/annonc/mdy308.
774 80. Stupp, R., Brada, M., van den Bent, M.J., Tonn, J.C., Pentheroudakis, G., and Group, E.G.W.
775 (2014). High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and
776 follow-up. Ann Oncol *25 Suppl 3*, iii93-101. 10.1093/annonc/mdu050.
777 81. Postmus, P.E., Kerr, K.M., Oudkerk, M., Senan, S., Waller, D.A., Vansteenkiste, J., Escriu, C.,
778 Peters, S., and Committee, E.G. (2017). Early and locally advanced non-small-cell lung cancer
779 (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann
780 Oncol *28*, iv1-iv21. 10.1093/annonc/mdx222.
781 82. Gui, C., Kosteniuk, S.E., Lau, J.C., and Megyesi, J.F. (2018). Tumor growth dynamics in
782 serially-imaged low-grade glioma patients. J Neurooncol *139*, 167-175. 10.1007/s11060-018-
783 2857-x.
784 83. Alvarez, M., Benhammou, J.N., Darci-Maher, N., French, S.W., Han, S.B., Sinsheimer, J.S.,
785 Agopian, V.G., Pisegna, J.R., and Pajukanta, P. (2022). Human liver single nucleus and single
786 cell RNA sequencing identify a hepatocellular carcinoma-associated cell-type affecting
787 survival. Genome Med *14*, 50. 10.1186/s13073-022-01055-5.
788 84. Infante, M., Berghmans, T., Heuvelmans, M.A., Hillerdal, G., and Oudkerk, M. (2013). Slow-
789 growing lung cancer as an emerging entity: from screening to clinical management. Eur
790 Respir J *42*, 1706-1722. 10.1183/09031936.00186212.
791 85. Hoshino, T. (1984). A commentary on the biology and growth kinetics of low-grade and high-
792 grade gliomas. J Neurosurg *61*, 895-900. 10.3171/jns.1984.61.5.0895.
793 86. Sajjadi, E., Venetis, K., Piciotti, R., Invernizzi, M., Guerini-Rocco, E., Haricharan, S., and
794 Fusco, N. (2021). Mismatch repair-deficient hormone receptor-positive breast cancers:

795 Biology and pathological characterization. Cancer Cell Int *21*, 266. 10.1186/s12935-021-
796 01976-y.
797 87. Doghri, R., Houcine, Y., Boujelbene, N., Driss, M., Charfi, L., Abbes, I., Mrad, K., and
798 Sellami, R. (2019). Mismatch Repair Deficiency in Endometrial Cancer:
799 Immunohistochemistry Staining and Clinical Implications. Appl Immunohistochem Mol
800 Morphol *27*, 678-682. 10.1097/PAI.0000000000000641.
801 88. Zhang, Q., Sun, T.Y., Kang, P.M., Qian, K., Deng, B., Zhou, J.H., Wang, R.W., Jiang, B., Li,
802 K., Liu, F., et al. (2016). Combined analysis of rearrangement of ALK, ROS1, somatic
803 mutation of EGFR, KRAS, BRAF, PIK3CA, and mRNA expression of ERCC1, TYMS,
804 RRM1, TUBB3, EGFR in patients with non-small cell lung cancer and their clinical
805 significance. Cancer Chemoth Pharm *77*, 583-593. 10.1007/s00280-016-2969-y.
806 89. Donev, I.S., Wang, W., Yamada, T., Li, Q., Takeuchi, S., Matsumoto, K., Yamori, T., Nishioka,
807 Y., Sone, S., and Yano, S. (2011). Transient PI3K inhibition induces apoptosis and overcomes
808 HGF-mediated resistance to EGFR-TKIs in EGFR mutant lung cancer. Clin Cancer Res *17*,
809 2260-2269. 10.1158/1078-0432.CCR-10-1993.
810 90. Wang, G., Xu, D., Zhang, Z., Li, X., Shi, J., Sun, J., Liu, H.Z., Li, X., Zhou, M., and Zheng, T.
811 (2021). The pan-cancer landscape of crosstalk between epithelial-mesenchymal transition and
812 immune evasion relevant to prognosis and immunotherapy response. NPJ Precis Oncol *5*, 56.
813 10.1038/s41698-021-00200-4.
814 91. Brown, T.J., Bota, D.A., van Den Bent, M.J., Brown, P.D., Maher, E., Aregawi, D., Liau,
815 L.M., Buckner, J.C., Weller, M., Berger, M.S., and Glantz, M. (2019). Management of low-
816 grade glioma: a systematic review and meta-analysis. Neurooncol Pract *6*, 249-258.
817 10.1093/nop/npy034.
818 92. Chen, L., Ma, J., Zou, Z., Liu, H., Liu, C., Gong, S., Gao, X., and Liang, G. (2022). Clinical
819 characteristics and prognosis of patients with glioblastoma: A review of survival analysis of
820 1674 patients based on SEER database. Medicine (Baltimore) *101*, e32042.
821 10.1097/MD.0000000000032042.
822 93. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,
823 Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated
824 models of biomolecular interaction networks. Genome Res *13*, 2498-2504.
825 10.1101/gr.1239303.
826 94. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing
827 biological themes among gene clusters. OMICS *16*, 284-287. 10.1089/omi.2011.0118.
828 95. Mayakonda, A., Lin, D.C., Assenov, Y., Plass, C., and Koeffler, H.P. (2018). Maftools:
829 efficient and comprehensive analysis of somatic variants in cancer. Genome Res *28*, 1747-
830 1756. 10.1101/gr.239244.118.
831 96. Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S.,
832 Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a Cancer Dependency
833 Map. Cell *170*, 564-576 e516. 10.1016/j.cell.2017.06.010.
834 97. Yang, W., Soares, J., Greninger, P., Edelman, E.J., Lightfoot, H., Forbes, S., Bindal, N., Beare,
835 D., Smith, J.A., Thompson, I.R., et al. (2013). Genomics of Drug Sensitivity in Cancer
836 (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res *41*,
837 D955-961. 10.1093/nar/gks1111.
838 98. Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J.,

Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell *171*, 1437-1452 e1417. 10.1016/j.cell.2017.10.049.

99. Finotello, F., Mayer, C., Plattner, C., Laschober, G., Rieder, D., Hackl, H., Krogsdam, A., Loncova, Z., Posch, W., Wilflingseder, D., et al. (2019). Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med *11*, 34. 10.1186/s13073-019-0638-6.

100. Sturm, G., Finotello, F., and List, M. (2020). Immunedeconv: An R Package for Unified Access to Computational Methods for Estimating Immune Cell Fractions from Bulk RNA-Sequencing Data. Methods Mol Biol *2120*, 223-232. 10.1007/978-1-0716-0327-7_16.

101. Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics *14*, 7. 10.1186/1471-2105-14-7.

Figure Legends

Figure 1: MEMORY uncovers the enrichment of mitosis and immune signatures in multiple cancers. (A) Sample sizes, ranging from 10% to 100% with 10% intervals ("gradient"), were used for 1000 permutations of survival analyses of all 15 cancer types Each matrix was divided into high and low expression groups based on median gene expression values, and survival analyses were performed. Log-rank test significance results were coded as 1 for significant ($P < 0.05$) and 0 for non-significant ($P > 0.05$) outcomes, forming survival analysis matrices and the summarized significant probability matrix, which allowed the identification of GEARs. (B) The maximum of significant probability for each gradient sample size of every cancer type. Sample number ratio refers to the percentage of samples in each sampling gradient compared to the total number of samples. (C) The pathways were enriched by GOEA based on the GEARs of each cancer type. The displayed pathways represent the top 5 most significant pathways for each cancer type. Mitosis-related pathways were marked in red whereas immune-related pathways were marked in blue.

Figure 2: Identification of hub genes in mitosis or immune-related cancers. (A) The SAS of all GEARs was calculated, and CSN was constructed. (B) The P-value of survival analysis was determined based on hub genes grouped by edge numbers (250, 500, 1000, 2000, 4000 and 8000 edges) from six CSNs. (C) The left nodes represent five different cancers, while the right nodes detail the functional types of hub-gene pathways. The height of the edges in the middle represents the proportion of hub-gene pathways corresponding to a specified functional type. (D-E) The CSN of LUAD and BRCA. The enlarged section represents the hub gene. (F) LUAD clustering based on hub gene expression, compared with inhibitor-based classification[52]. (G) LUAD clustering based on hub gene expression in comparison with the classic classification.

Figure 3: Different LUAD subgroups were characterized by unique genetic mutation profiles. (A) TMB analysis in three groups of LUAD samples. ***$P < 0.0005$. (B) Proportion of different oncogenic drivers including KRAS, EGFR, BRAF, ERBB2 mutations, and ALK, ROS1 fusions in three LUAD subgroups. (C-E) Comparison of top gene mutations in three

LUAD clusters, including ML vs. MM (C), MM vs. MH (D), ML vs MH (E).

Figure 4: PIK3CA mutation associates with mitosis and drug resistance in LUAD. (A) Cancer-related gene mutations were annotated based on gene dependency scores data from Depmap database. Functional mutations were indicated in green; functional (subtype-associated) mutations were highlighted in red; non-functional mutations were indicated in gray. (B) The analysis of the NeST differential pathways across three groups including ML vs. MM, MM vs. MH, ML vs MH. (C) The heat map showed the hub gene expression of A549 after compounds treatment from cMAP database[98]. Blue annotations meant 76 compounds that can inhibit hub gene expression. (D) Comparison of the IC50 z-score of five tumor cell growth inhibitors for A549 (*KRAS* mutation) and SW1573 (*KRAS* and *PIK3CA* mutation).

Figure 5: The hub gene classification of BRCA revealed a significant association between EMT and immune infiltration. (A) The hub gene expression heatmap containing the result of hub gene classification and PAM50 classification of BRCA[58]. (B) Comparison of molecular classification based on hub genes with PAM50 classification. (C-E) Comparison of gene mutations in three groups of BRCA samples, including IL vs. IM (C), IM vs. IH (D), IL vs. IH (E). (F) Comparison of the total immune cell rate in three BRCA clusters. ***$P < 0.0005$. (G) Comparison of *CDH1* expression in BRCA samples with wildtype or mutant *CDH1*. ***$P < 0.0005$. (H) A schematic diagram illustrating the correlation between the EMT score and immune cell rate. The red line represents the fitting curve. The correlation analysis was performed using Pearson correlation coefficient.

Figure 6: Mitosis and immune signatures predict patient prognosis at the pan-cancer level. The mitosis and immune-related pathway scores of 33 cancer types were analyzed, and the median score was utilized as a threshold to categorize patients for survival analysis. Among the 33 cancer types examined, 10 cancer types were exclusively associated with the mitosis score, 2 cancer types were exclusively associated with the immune score, and 5 cancer types showed a correlation with both mitosis and immunity scores concurrently.

Figure 1

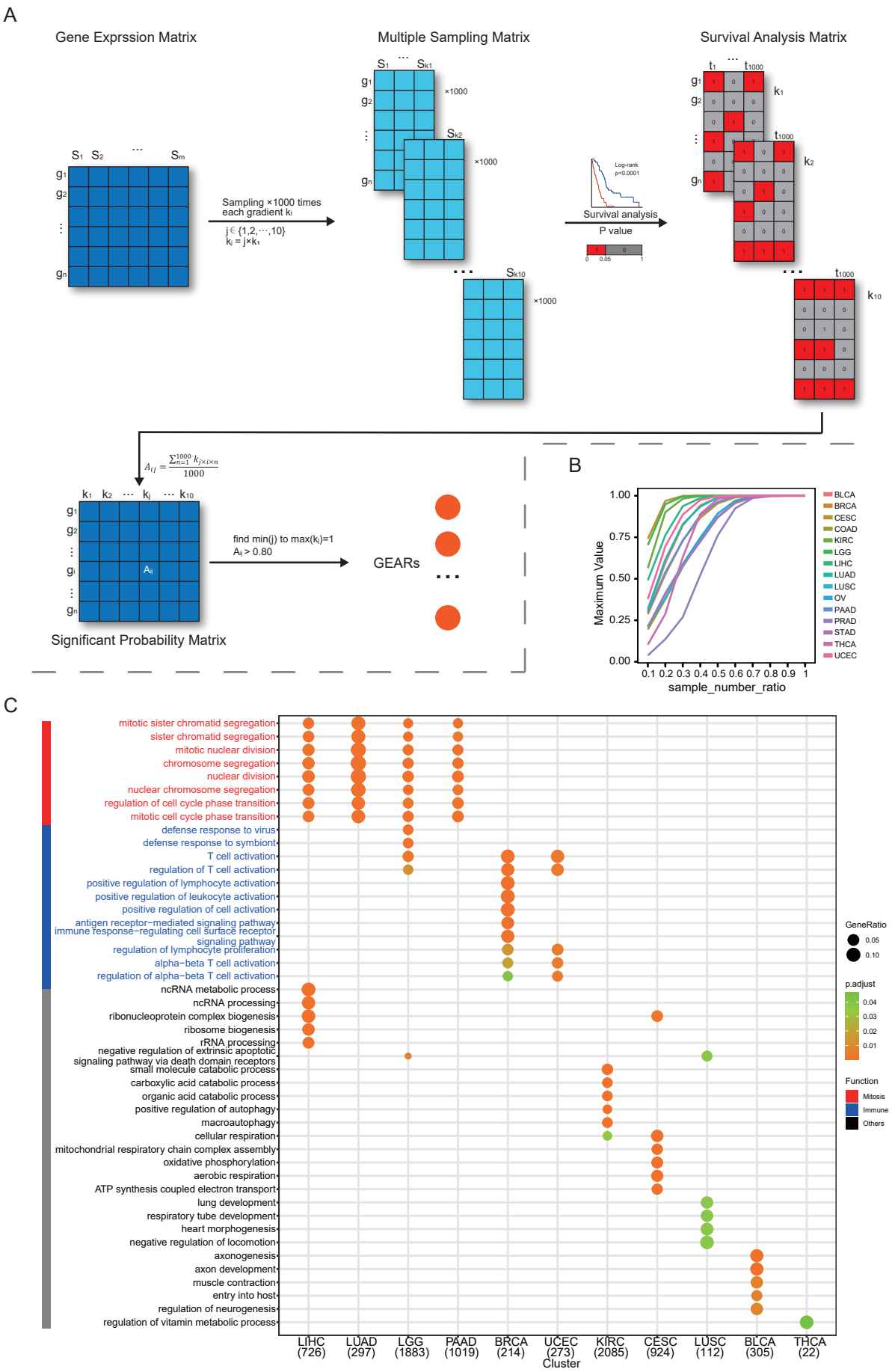
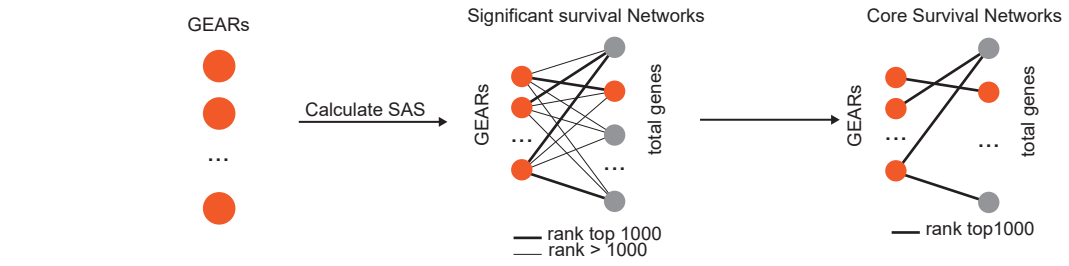
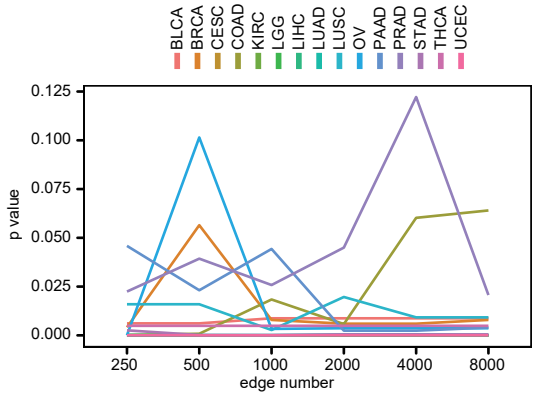

A
Gene Exprssion Matrix
Multiple Sampling Matrix
Survival Analysis Matrix
Sampling ×1000 times each gradient kᵢ
j ∈ {1,2,⋯,10}
kⱼ = j×k₁
×1000
Log-rank p<0.0001
Survival analysis
P value
Significant Probability Matrix
find min(j) to max(kⱼ)=1
Aᵢⱼ > 0.80
GEARs
B
Maximum Value
sample_number_ratio
BLCA
BRCA
CESC
COAD
KIRC
LGG
LIHC
LUAD
LUSC
OV
PAAD
PRAD
STAD
THCA
UCEC
C
mitotic sister chromatid segregation
sister chromatid segregation
mitotic nuclear division
chromosome segregation
nuclear division
nuclear chromosome segregation
regulation of cell cycle phase transition
mitotic cell cycle phase transition
defense response to virus
defense response to symbiont
T cell activation
regulation of T cell activation
positive regulation of lymphocyte activation
positive regulation of leukocyte activation
positive regulation of cell activation
antigen receptor−mediated signaling pathway
immune response−regulating cell surface receptor signaling pathway
regulation of lymphocyte proliferation
alpha−beta T cell activation
regulation of alpha−beta T cell activation
ncRNA metabolic process
ncRNA processing
ribonucleoprotein complex biogenesis
ribosome biogenesis
rRNA processing
negative regulation of extrinsic apoptotic signaling pathway via death domain receptors
small molecule catabolic process
carboxylic acid catabolic process
organic acid catabolic process
positive regulation of autophagy
macroautophagy
cellular respiration
mitochondrial respiratory chain complex assembly
oxidative phosphorylation
aerobic respiration
ATP synthesis coupled electron transport
lung development
respiratory tube development
heart morphogenesis
negative regulation of locomotion
axonogenesis
axon development
muscle contraction
entry into host
regulation of neurogenesis
regulation of vitamin metabolic process
LIHC (726)
LUAD (297)
LGG (1883)
PAAD (1019)
BRCA (214)
UCEC (273)
KIRC (2085)
CESC (924)
LUSC (112)
BLCA (305)
THCA (22)
Cluster
GeneRatio
p.adjust
Function
Mitosis
Immune
Others

Figure 2

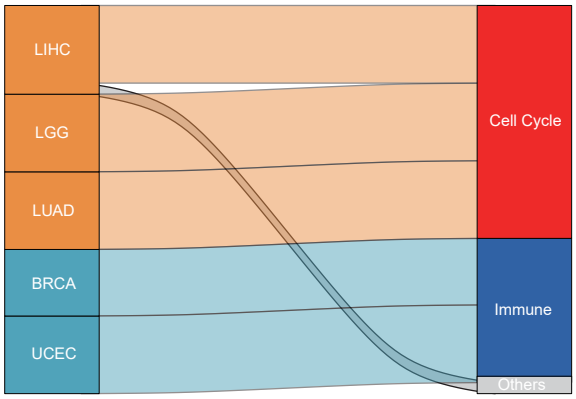
A



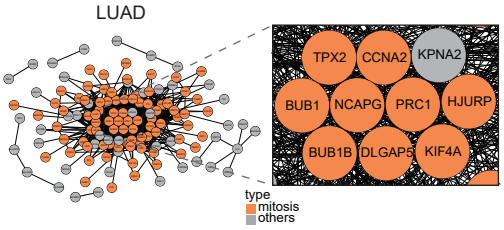
B



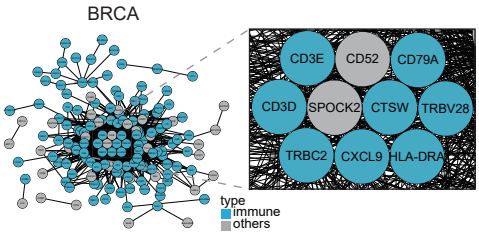
C



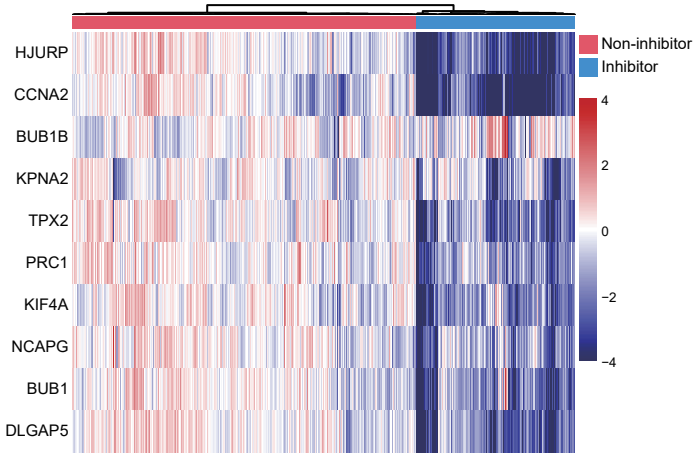
D



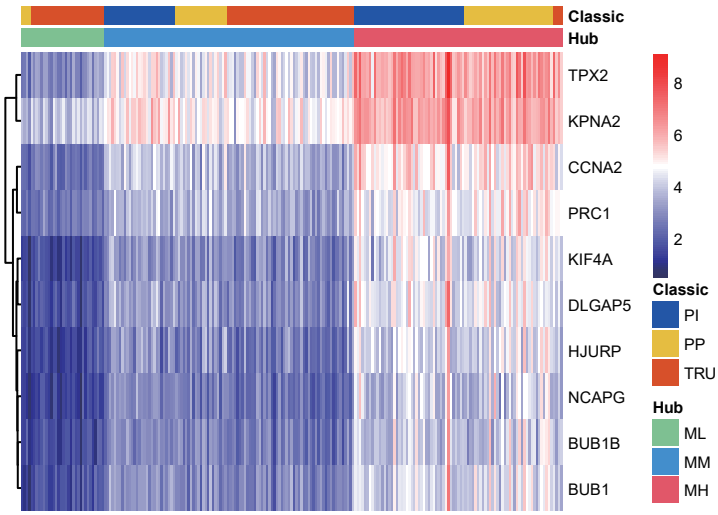
E



F



G



H

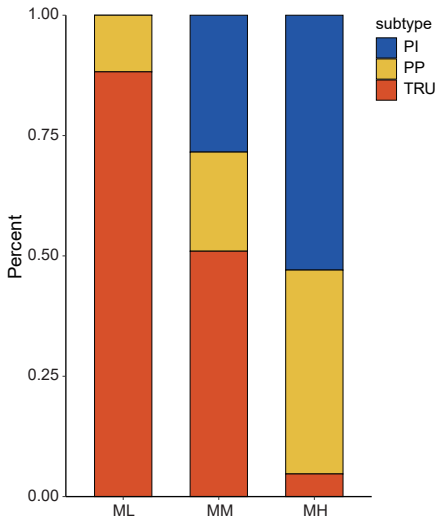


Figure 3

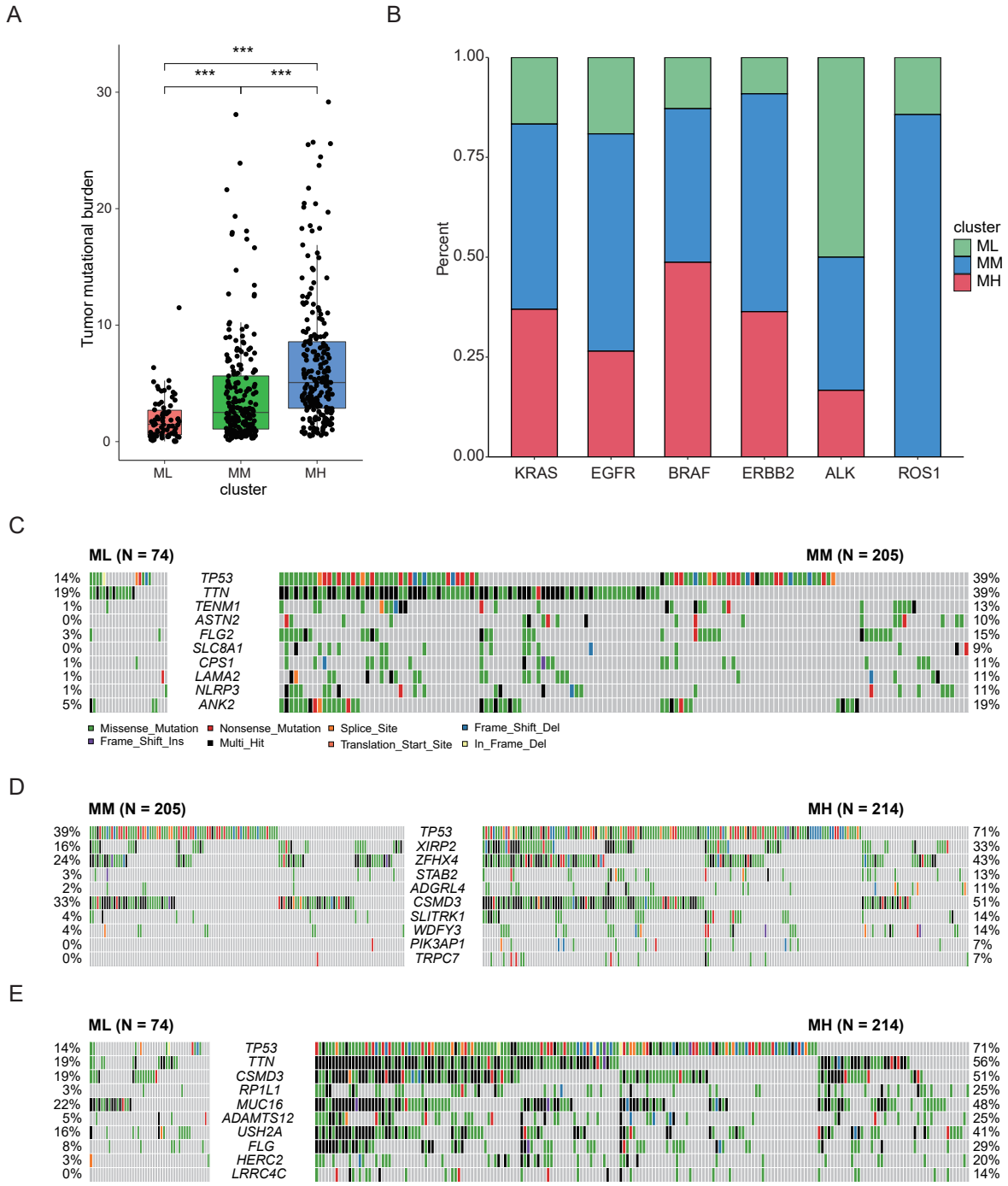
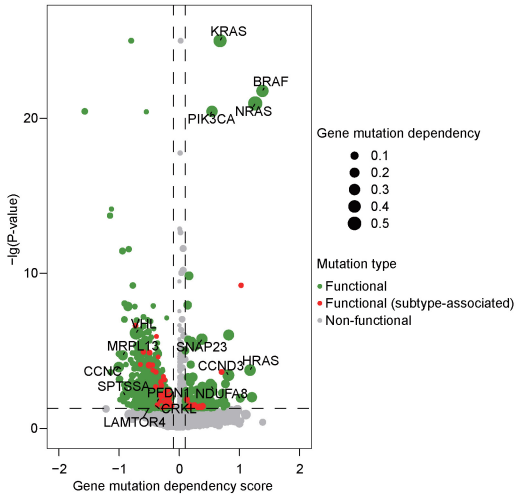
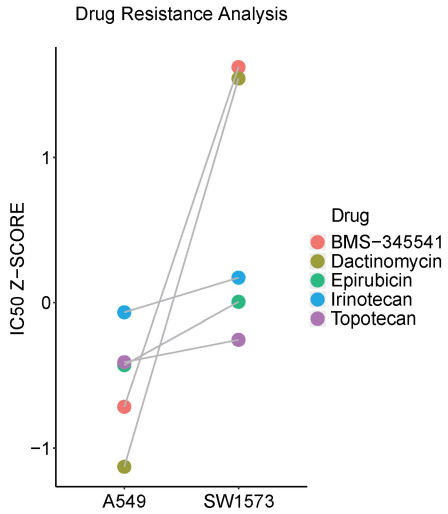

A
***
***
***
Tumor mutational burden
30
20
10
0
ML
MM
MH
cluster
B
Percent
1.00
0.75
0.50
0.25
0.00
KRAS
EGFR
BRAF
ERBB2
ALK
ROS1
cluster
ML
MM
MH
C
ML (N = 74)
MM (N = 205)
14% TP53 39%
19% TTN 39%
1% TENM1 13%
0% ASTN2 10%
3% FLG2 15%
0% SLC8A1 9%
1% CPS1 11%
1% LAMA2 11%
1% NLRP3 11%
5% ANK2 19%
Missense_Mutation
Nonsense_Mutation
Splice_Site
Frame_Shift_Del
Frame_Shift_Ins
Multi_Hit
Translation_Start_Site
In_Frame_Del
D
MM (N = 205)
MH (N = 214)
39% TP53 71%
16% XIRP2 33%
24% ZFHX4 43%
3% STAB2 13%
2% ADGRL4 11%
33% CSMD3 51%
4% SLITRK1 14%
4% WDFY3 14%
0% PIK3AP1 7%
0% TRPC7 7%
E
ML (N = 74)
MH (N = 214)
14% TP53 71%
19% TTN 56%
19% CSMD3 51%
3% RP1L1 25%
22% MUC16 48%
5% ADAMTS12 25%
16% USH2A 41%
8% FLG 29%
3% HERC2 20%
0% LRRC4C 14%

Figure 4

A


KRAS
BRAF
NRAS
PIK3CA
VHL
MRPL13
SNAP23
CCND3
HRAS
CCNC
SPTSSA
PFDN1
NDUFA8
GRIK1
LAMTOR4
−lg(P-value)
Gene mutation dependency score
Gene mutation dependency
0.1
0.2
0.3
0.4
0.5
Mutation type
Functional
Functional (subtype-associated)
Non-functional

C


Drug Resistance Analysis
IC50 Z−SCORE
A549
SW1573
Drug
BMS−345541
Dactinomycin
Epirubicin
Irinotecan
Topotecan

B

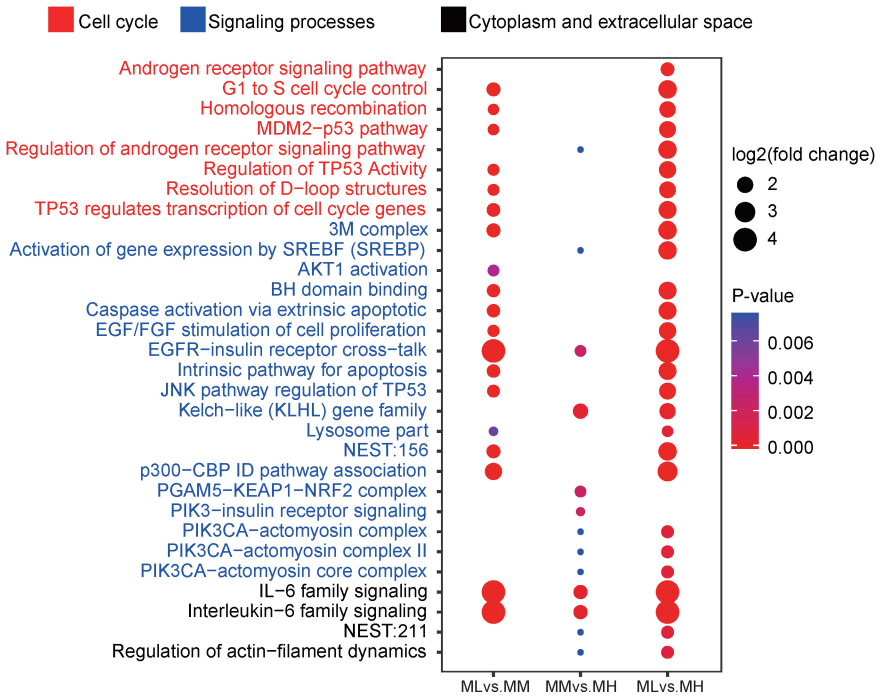

Cell cycle
Signaling processes
Cytoplasm and extracellular space
Androgen receptor signaling pathway
G1 to S cell cycle control
Homologous recombination
MDM2−p53 pathway
Regulation of androgen receptor signaling pathway
Regulation of TP53 Activity
Resolution of D−loop structures
TP53 regulates transcription of cell cycle genes
3M complex
Activation of gene expression by SREBF (SREBP)
AKT1 activation
BH domain binding
Caspase activation via extrinsic apoptotic
EGF/FGF stimulation of cell proliferation
EGFR−insulin receptor cross−talk
Intrinsic pathway for apoptosis
JNK pathway regulation of TP53
Kelch−like (KLHL) gene family
Lysosome part
NEST:156
p300−CBP ID pathway association
PGAM5−KEAP1−NRF2 complex
PIK3−insulin receptor signaling
PIK3CA−actomyosin complex
PIK3CA−actomyosin complex II
PIK3CA−actomyosin core complex
IL−6 family signaling
Interleukin−6 family signaling
NEST:211
Regulation of actin−filament dynamics
MLvs.MM
MMvs.MH
MLvs.MH
log2(fold change)
2
3
4
P-value
0.006
0.004
0.002
0.000

Figure 5

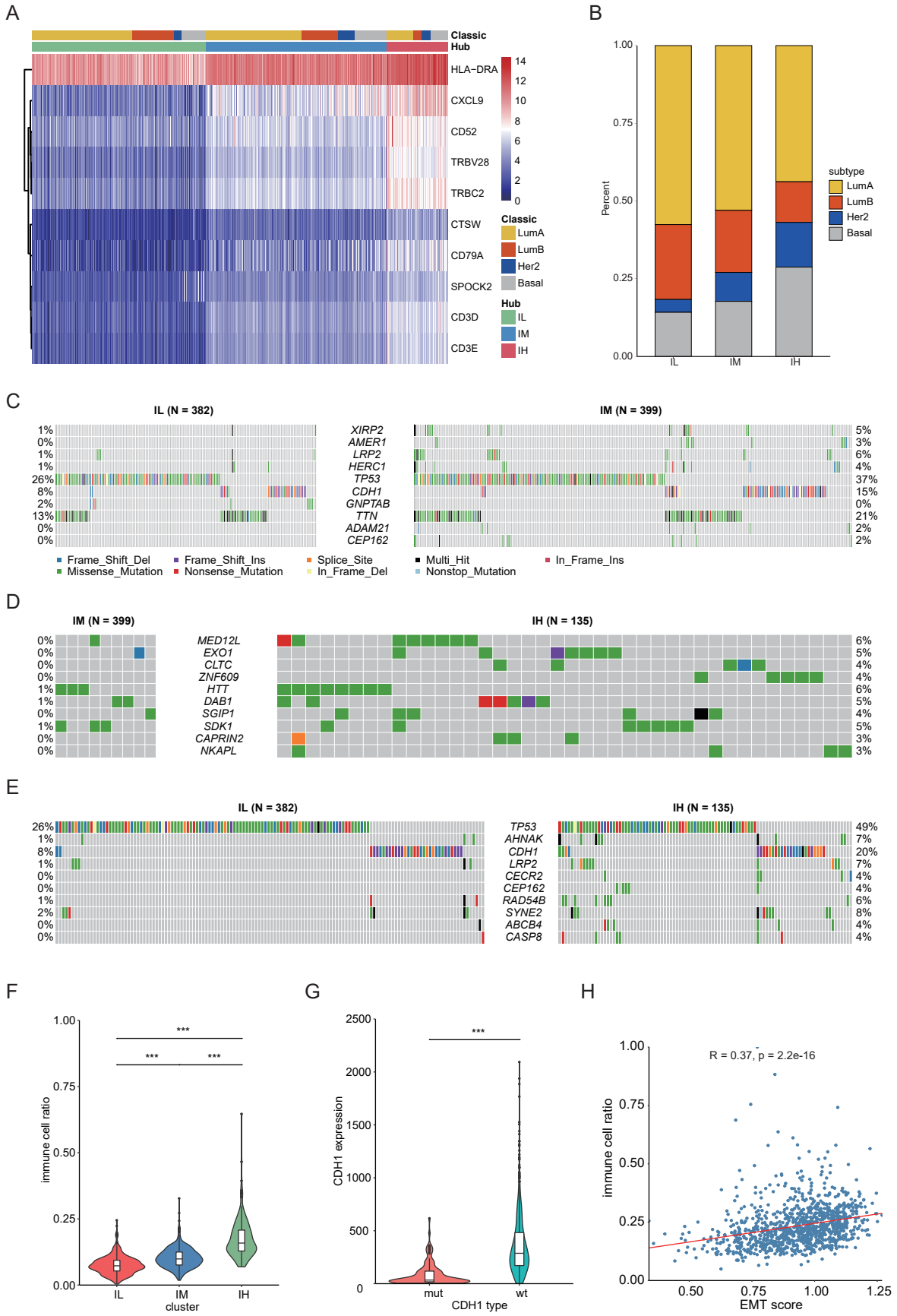
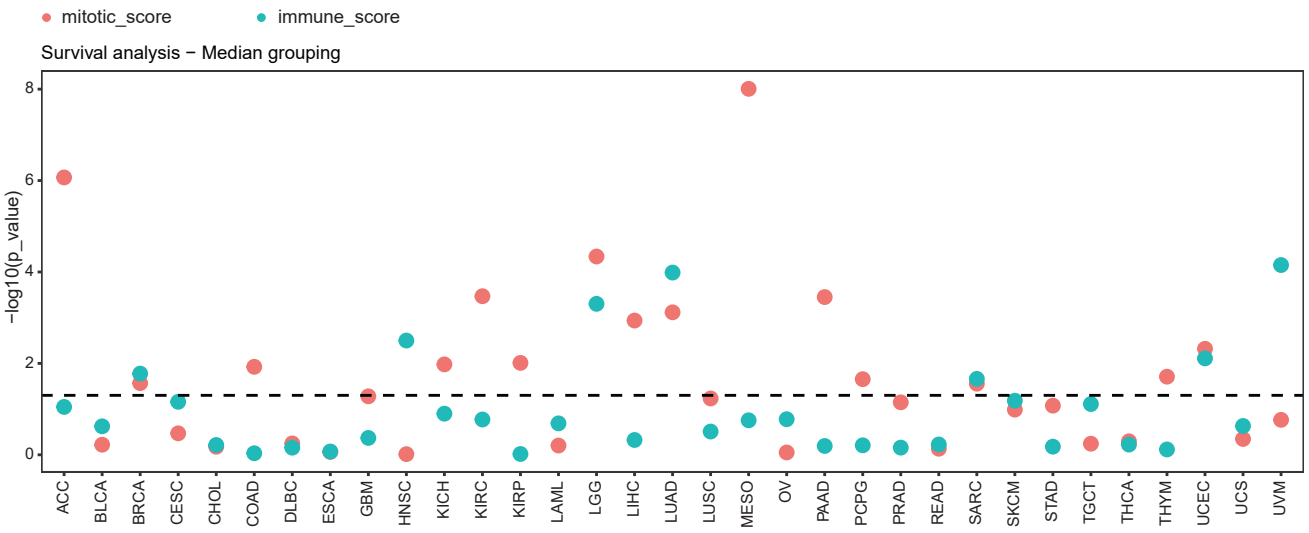

A
Classic
Hub
HLA−DRA
CXCL9
CD52
TRBV28
TRBC2
CTSW
CD79A
SPOCK2
CD3D
CD3E
14
12
10
8
6
4
2
0
Classic
LumA
LumB
Her2
Basal
Hub
IL
IM
IH
B
1.00
0.75
0.50
0.25
0.00
Percent
IL
IM
IH
subtype
LumA
LumB
Her2
Basal
C
IL (N = 382)
IM (N = 399)
1%
0%
1%
1%
26%
8%
2%
13%
0%
0%
XIRP2
AMER1
LRP2
HERC1
TP53
CDH1
GNPTAB
TTN
ADAM21
CEP162
5%
3%
6%
4%
37%
15%
0%
21%
2%
2%
Frame_Shift_Del
Frame_Shift_Ins
Splice_Site
Multi_Hit
In_Frame_Ins
Missense_Mutation
Nonsense_Mutation
In_Frame_Del
Nonstop_Mutation
D
IM (N = 399)
IH (N = 135)
0%
0%
0%
0%
1%
1%
0%
1%
0%
0%
MED12L
EXO1
CLTC
ZNF609
HTT
DAB1
SGIP1
SDK1
CAPRIN2
NKAPL
6%
5%
4%
4%
6%
5%
4%
5%
3%
3%
E
IL (N = 382)
IH (N = 135)
26%
1%
8%
1%
0%
0%
1%
2%
0%
0%
TP53
AHNAK
CDH1
LRP2
CECR2
CEP162
RAD54B
SYNE2
ABCB4
CASP8
49%
7%
20%
7%
4%
4%
6%
8%
4%
4%
F
1.00
0.75
0.50
0.25
0.00
immune cell ratio
***
***
***
IL
IM
IH
cluster
G
2500
2000
1500
1000
500
0
CDH1 expression
***
mut
wt
CDH1 type
H
1.00
0.75
0.50
0.25
0.00
immune cell ratio
R = 0.37, p = 2.2e-16
0.50
0.75
1.00
1.25
EMT score

Figure 6


mitotic_score
immune_score
Survival analysis − Median grouping
−log10(p_value)
0
2
4
6
8
ACC
BLCA
BRCA
CESC
CHOL
COAD
DLBC
ESCA
GBM
HNSC
KICH
KIRC
KIRP
LAML
LGG
LIHC
LUAD
LUSC
MESO
OV
PAAD
PCPG
PRAD
READ
SARC
SKCM
STAD
TGCT
THCA
THYM
UCEC
UCS
UVM

Multi-gradient Permutation Survival Analysis Identifies Mitosis and Immune Signatures Steadily Associated with Cancer Patient Prognosis

Xinlei Cai[1], Yi Ye[2], Xiaoping Liu[1], Zhaoyuan Fang[4,5], Luonan Chen[1,2,6], Fei Li[3*], Hongbin Ji[1,2,6*]

[1]Key Laboratory of Systems Health Science of Zhejiang Province, School of Life Science, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou, 310024, China.

[2]State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, Shanghai 200031, China.

[3]Department of Pathology and Frontier Innovation Center, School of Basic Medical Sciences, Fudan University, Shanghai, China.

[4]Zhejiang University-University of Edinburgh Institute, Zhejiang University School of Medicine, Haining，314400, China

[5]The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, 310000, China

[6]School of Life Science and Technology, Shanghai Tech University, Shanghai 200120, China.

Running title: Mitosis and immune signatures in predicting cancer patient prognosis

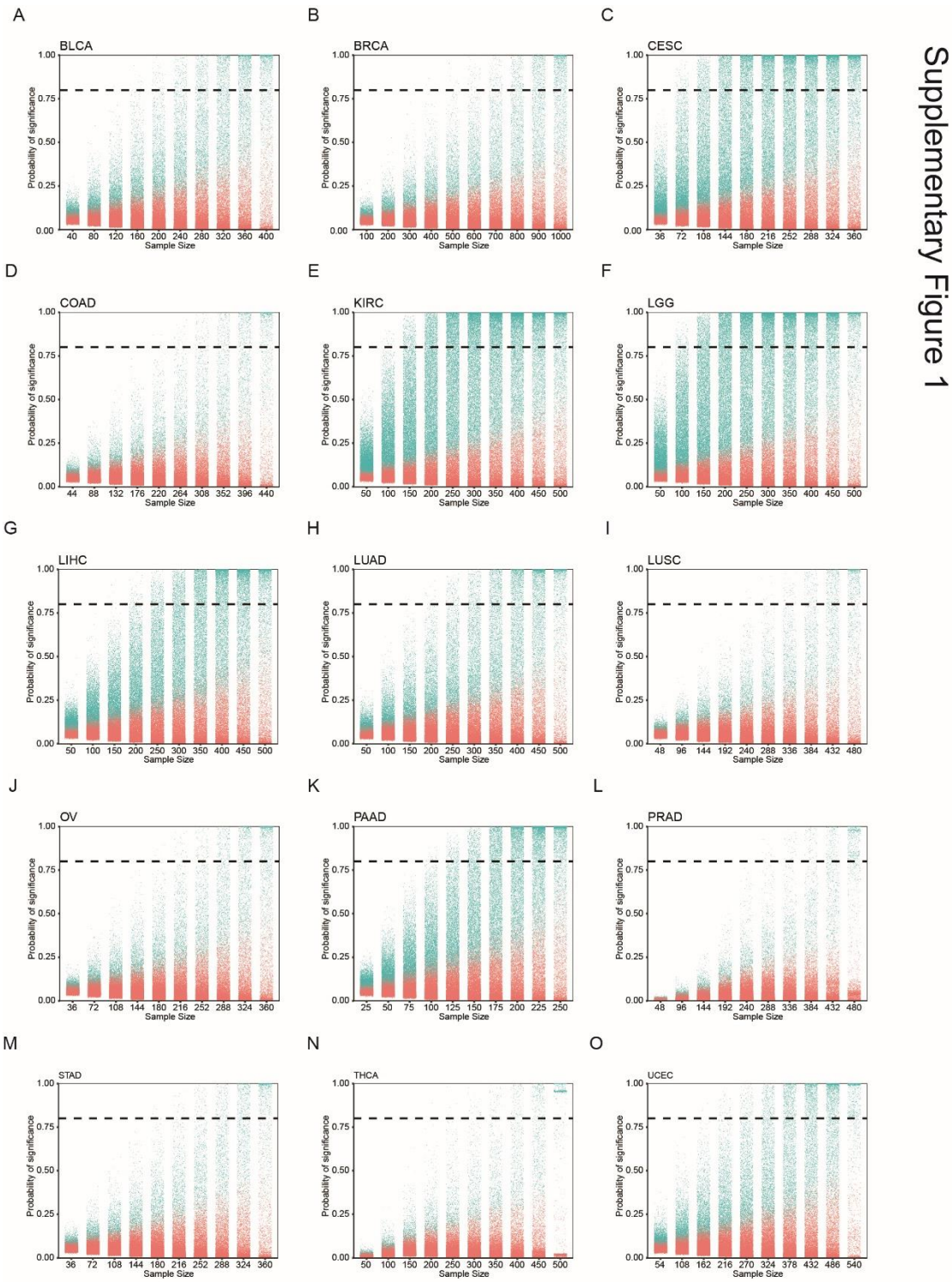
***Corresponding author:**

Hongbin Ji, State Key Laboratory of Cell Biology, Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai, 200031, China, Phone: 86-21-54921108, Fax: 86-21-54921101, E-mail address: hbji@sibcb.ac.cn;

Fei Li, Department of Pathology and Frontier Innovation Center, School of Basic Medical

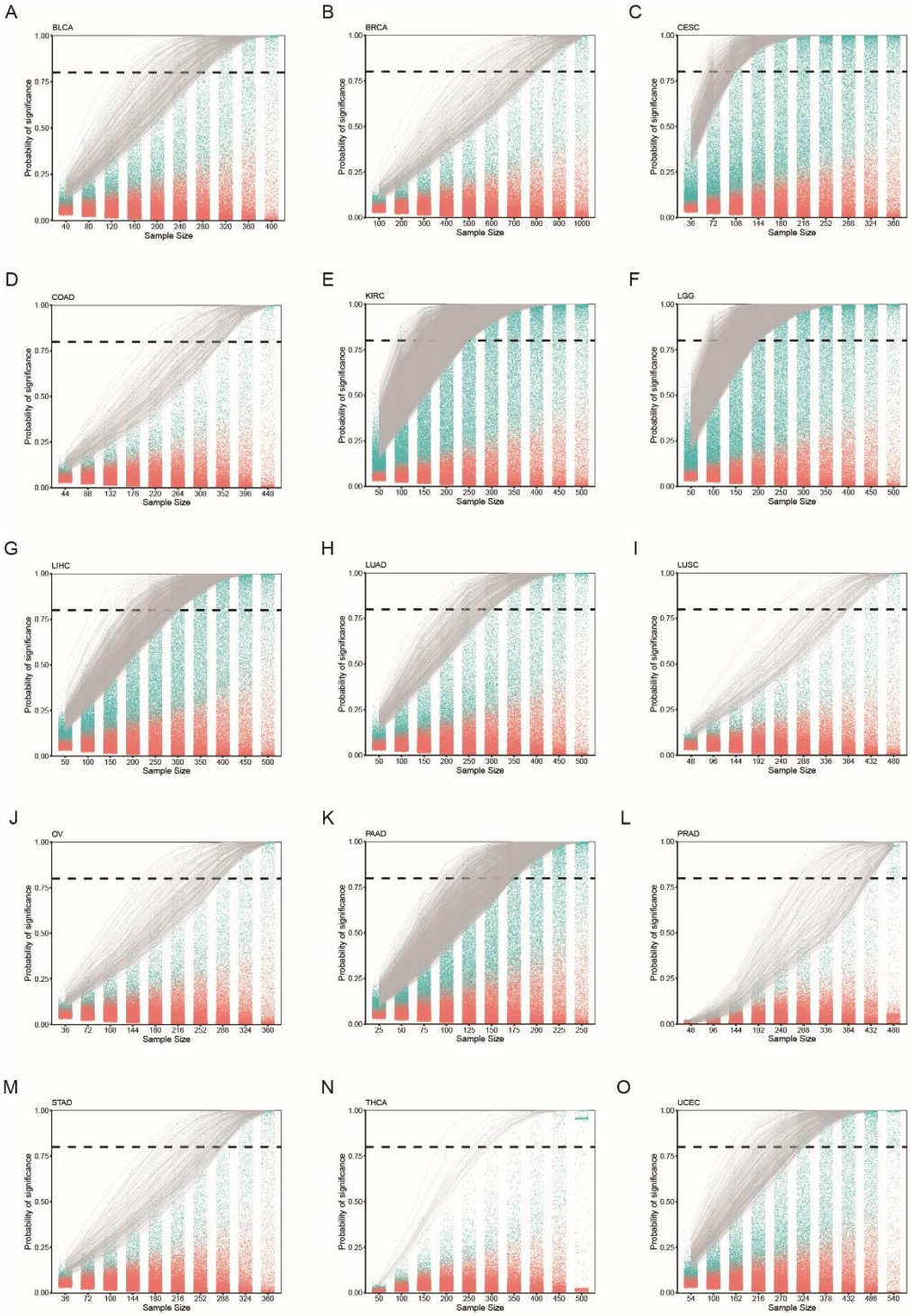
Sciences, Fudan University, 131 Dong An Road, Shanghai, 200032, China, E-mail address: li_fei@fudan.edu.cn.

Supplementary Figure 1

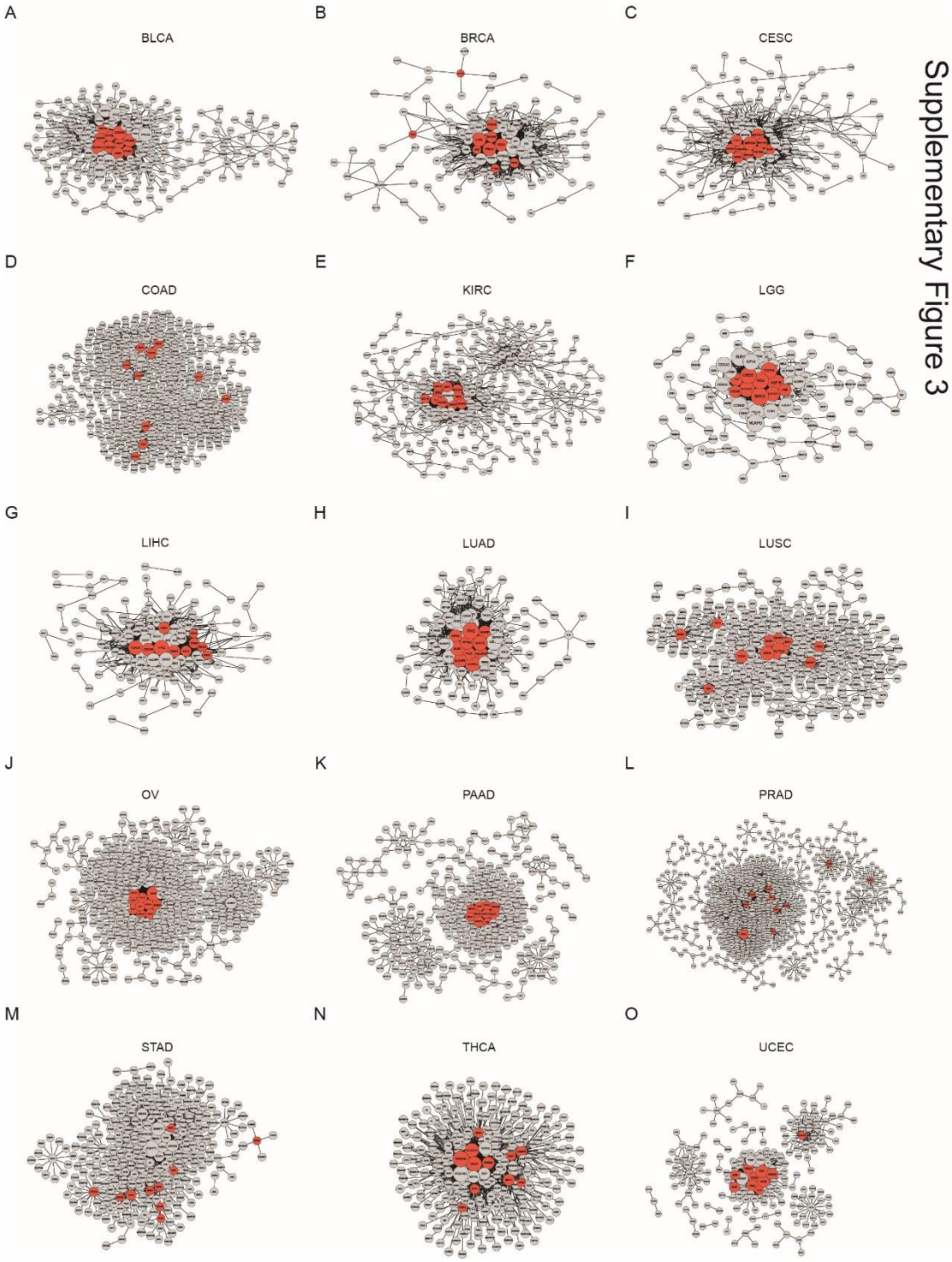


Supplementary Figure 1: (A) to (O) The number of genes significantly related to prognosis which were obtained through MEMORY was positively correlated with sample size. The vertical value of the dotted line is 0.8.

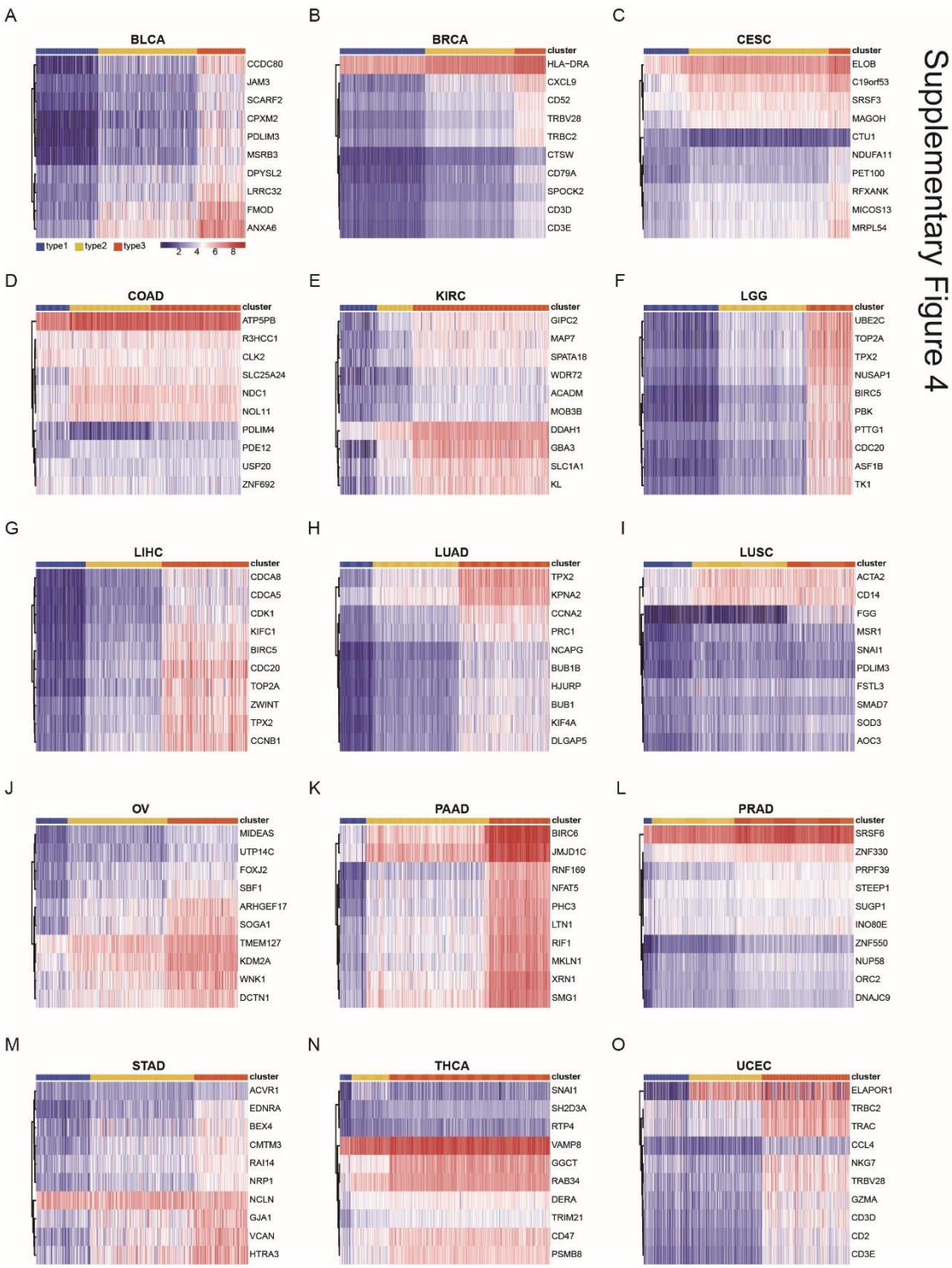
Supplementary Figure 2



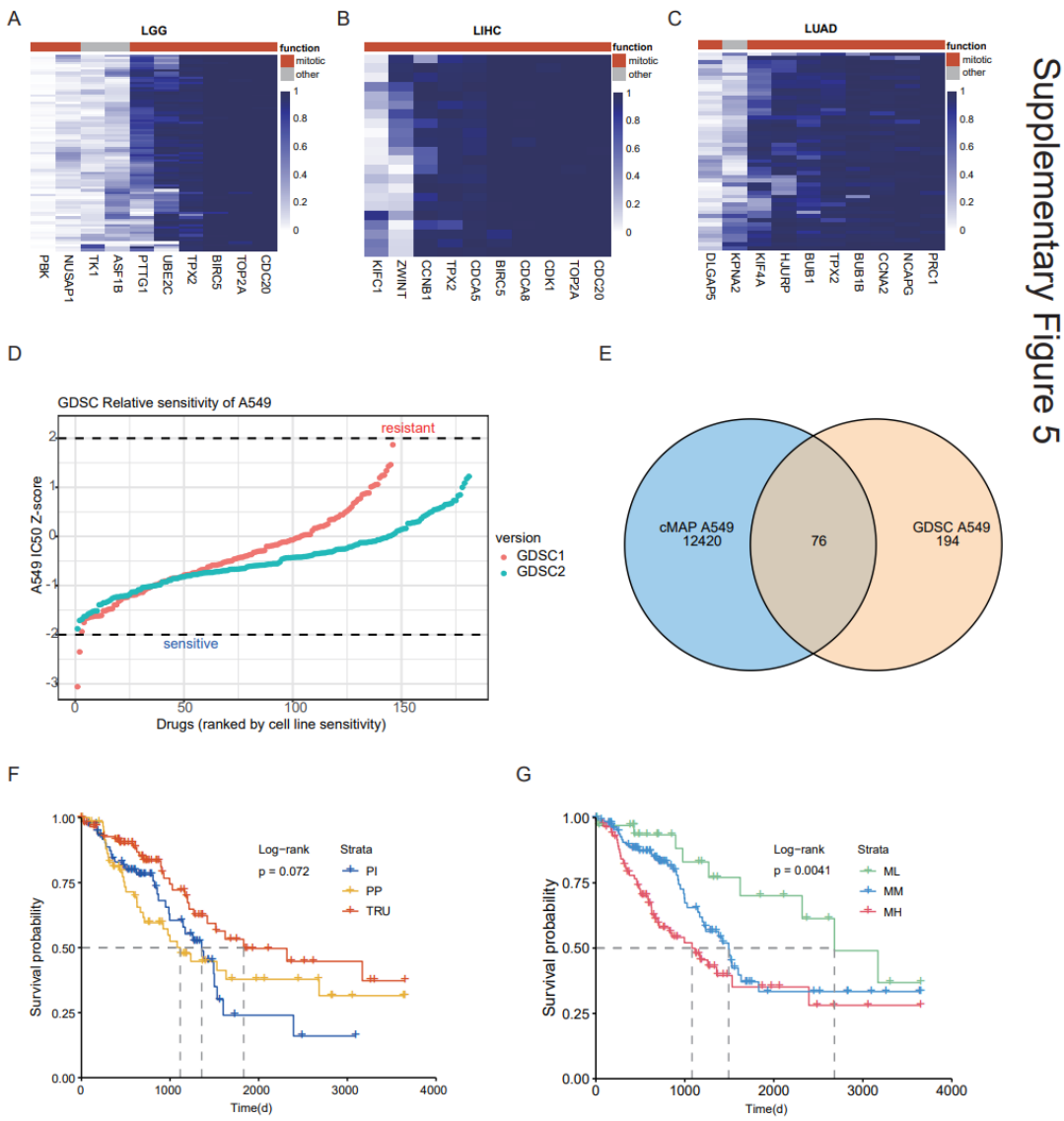
Supplementary Figure 2: (A) to (O) The significant probability of GEAR at each gradient. The vertical value of the dotted line is 0.8.



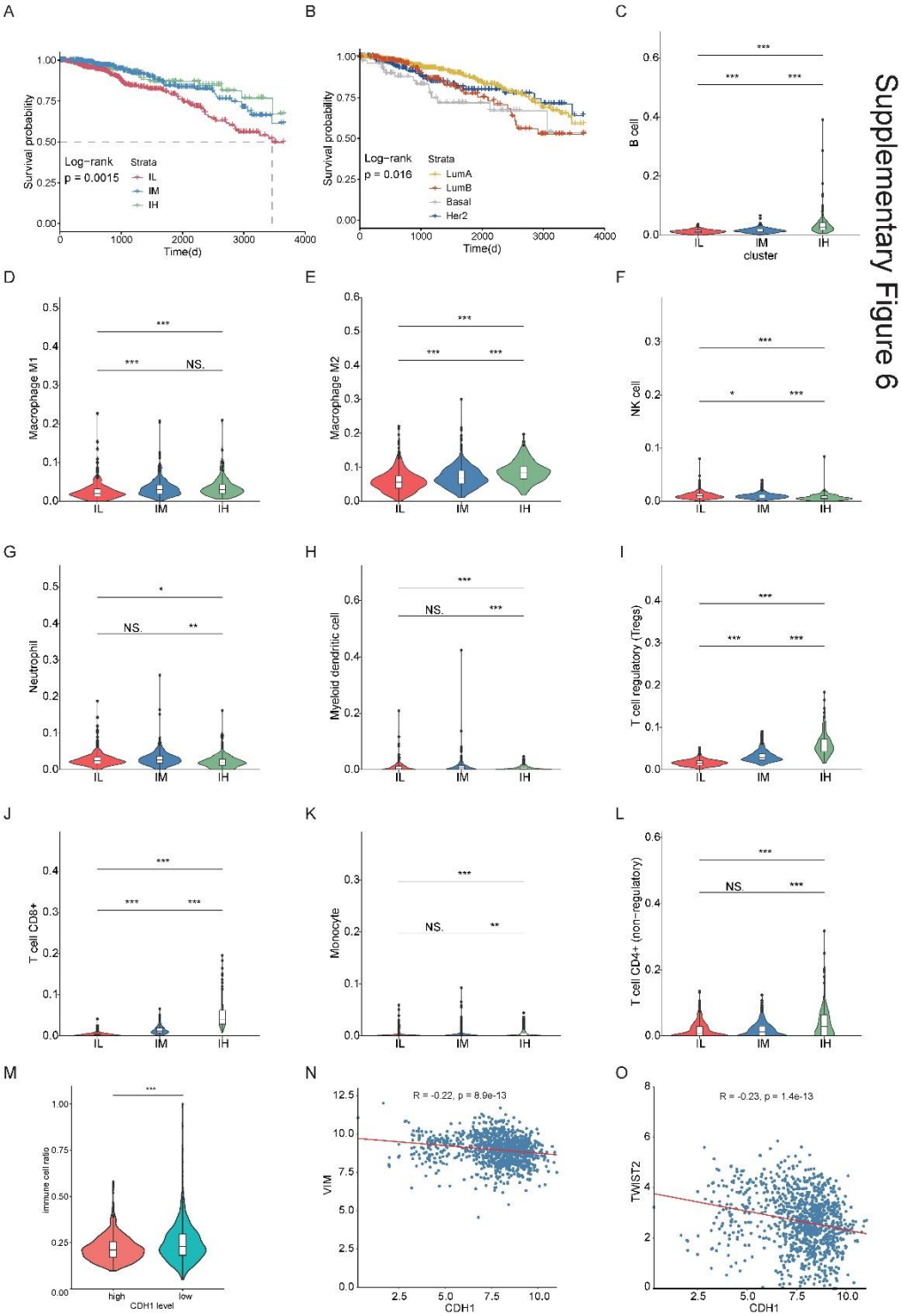
Supplementary Figure 3: (A) to (O) The core survival networks of 15 cancer types. Red nodes are hub genes.



Supplementary Figure 4: (A) to (O) 15 cancer types were genotyped through hierarchical clustering based on the expression of hub genes, and the samples of each cancer were divided into three clusters



Supplementary Figure 5: (A-C) The gene dependency analysis of hub genes for LGG (A), LIHC (B), LUAD (C). (D) The number of compounds associated with A549 in the cMAP and GDSC databases. (E) Relative sensitivity of A549 cell line to all compounds from the GDSC database, calculated based on IC50 values. (F) Kaplan-Meyer overall survival curves for LUAD classic subgroups. (G) Kaplan-Meyer overall survival curves for LUAD hub gene subgroups.



Supplementary Figure 6: (A) Kaplan-Meyer overall survival curves for BRCA classic subgroups. (B) Kaplan-Meyer overall survival curves for BRCA PAM50 subgroups. (C-L) Comparison of infiltration of different types of immune cells in three group BRCA samples. ***P < 0.0005, **P < 0.005 and *P < 0.05. (M) Comparison of the proportion of immune

cells in the total cells in BRCA samples with high *CDH1* expression and low *CDH1* expression. ***P < 0.0005. (N) Schematic diagram of the correlation between *CDH1* expression and *VIM* expression. The red line was a fitting curve. Pearson correlation coefficient was used for correlation analysis. (O) Schematic diagram of the correlation between *CDH1* expression and *TWIST2* expression. The red line was a fitting curve. Pearson correlation coefficient was used for correlation analysis.