# Identification of four hub genes associated with adrenocortical carcinoma progression by WGCNA

**Wang-Xiao Xia** [1, 2, 3, 4], **Qin Yu** [1, 2, 3, 4], **Gong-Hua Li** [1, 2, 3], **Yao-Wen Liu** [1, 2, 3], **Fu-Hui Xiao** [1, 2, 3], **Li-Qin Yang** [1, 2, 3], **Zia Ur Rahman** [1, 2, 3, 4], **Hao-Tian Wang** [1, 2, 3, 4], **Qing-Peng Kong** [Corresp. 1, 2, 3]

[1] State Key Laboratory of Genetic Resources and Evolution/Key Laboratory of Healthy Aging Research of Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[2] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

[3] Kunming Key Laboratory of Healthy Aging Study, Kunming, China

[4] Kunming College of Life Science, University of Chinese Academy of Sciences, Beijing, China

Corresponding Author: Qing-Peng Kong
Email address: kongqp@mail.kiz.ac.cn

**Background.** Adrenocortical carcinoma (ACC) is a rare and aggressive malignant cancer in the adrenal cortex with poor prognosis. Though previous research has attempted to elucidate the progression of ACC, its molecular mechanism remains poorly understood.

**Methods.** Gene TPM (transcripts per million) data were downloaded from the UCSC Xena database, which included ACC (The Cancer Genome Atlas (TCGA), n = 77) and normal samples (Genotype Tissue Expression (GTEx), n = 128). We used weighted gene co-expression network analysis (WGCNA) to identify gene connections. Overall survival (OS) was determined using the univariate Cox model. A protein-protein interaction (PPI) network was constructed by the Search Tool for the Retrieval of Interacting Genes (STRING).

**Results.** To determine the critical genes involved in ACC progression, we obtained 2,953 significantly differentially expressed genes (DEGs) and nine modules. Among them, the blue module demonstrated significant correlation with the “Stage” of ACC. Enrichment analysis revealed that genes in the blue module were mainly enriched in cell division, cell cycle, and DNA replication. Combined with the PPI and co-expression networks, we identified four hub genes (i.e., *TOP2A*, *TTK*, *CHEK1*, and *CENPA*) that were highly expressed in ACC and negatively correlated with OS. Thus, these identified genes may play important roles in the progression of ACC and serve as potential biomarkers for future diagnosis.

# Identification of four hub genes associated with adrenocortical carcinoma progression by WGCNA

Wang-Xiao Xia[1,2,3,4], Qin Yu[1,2,3,4], Gong-Hua Li[1,2,3], Yao-Wen Liu[1,2,3], Fu-Hui Xiao[1,2,3], Li-Qin Yang[1,2,3], Zia Ur Rahman[1,2,3,4], Hao-Tian Wang[1,2,3,4], Qing-Peng Kong[1,2,3]

[1] State Key Laboratory of Genetic Resources and Evolution/Key Laboratory of Healthy Aging Research of Yunnan Province, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[2] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

[3] Kunming Key Laboratory of Healthy Aging Study, Kunming, China

[4] Kunming College of Life Science, University of Chinese Academy of Sciences, Beijing, China

Corresponding author:

Qing-Peng Kong, Telephone: +86-871-68125403; Fax: +86-871-68125403

Email address: kongqp@mail.kiz.ac.cn

## ABSTRACT

**Background.** Adrenocortical carcinoma (ACC) is a rare and aggressive malignant cancer in the adrenal cortex with poor prognosis. Though previous research has attempted to elucidate the progression of ACC, its molecular mechanism remains poorly understood.

**Methods.** Gene TPM (transcripts per million) data were downloaded from the UCSC Xena database, which included ACC (The Cancer Genome Atlas (TCGA), n = 77) and normal samples (Genotype Tissue Expression (GTEx), n = 128). We used weighted gene co-expression network analysis (WGCNA) to identify gene connections. Overall survival (OS) was determined using the univariate Cox model. A protein-protein interaction (PPI) network was constructed by the Search Tool for the Retrieval of Interacting Genes (STRING).

**Results.** To determine the critical genes involved in ACC progression, we obtained 2,953 significantly differentially expressed genes (DEGs) and nine modules. Among them, the blue module demonstrated significant correlation with the "Stage" of ACC. Enrichment analysis revealed that genes in the blue module were mainly enriched in cell division, cell cycle, and DNA replication. Combined with the PPI and co-expression networks, we identified four hub genes (i.e., *TOP2A*, *TTK*, *CHEK1*, and *CENPA*) that were highly expressed in ACC and negatively correlated with OS. Thus, these identified genes may play important roles in the progression of ACC and serve as potential biomarkers for future diagnosis.


## INTRODUCTION

Adrenocortical carcinoma (ACC) is a rare and aggressive malignant cancer found in the adrenal cortex (Fay et al., 2014). While this disease can occur at any age, it tends to show a bi-modal distribution with an initial peak in childhood (1–6 years old) and a second peak in middle-age (40–50 years old) (Kiseljak-Vassiliades et al., 2018). As ACC has no obvious phenotypic traits at

the early stage, almost 70% of patients are at stage III or IV when diagnosed (Bharwani et al., 2011; Fay et al., 2014). At these stages, ACC is invasive and metastatic, with patients at stage IV only having a five-year survival of 6%–13% (Else et al., 2014; Fassnacht et al., 2009; Fassnacht et al., 2013). Unfortunately, current ACC therapies, such as surgery, chemotherapy, and radiotherapy, exhibit poor performance and outcomes (Allolio et al., 2006). While next generation sequencing technology recently identified several genetic molecules associated with ACC (Soon et al., 2008; Assié et al., 2014; Greenhill et al., 2016; Zheng et al., 2016; Chortis et al., 2018), our understanding of ACC progression at each stage remains incomplete and treatment options are limited (Hoang et al., 2002; Cherradi, 2014). Thus, integrated analysis is required to further understand the molecular characterization of ACC gene expression, which may indicate stage and identify additional biomarkers for further research and clinical therapies.

Traditional methods of identifying the functional genes of ACC have focused on screening differentially expressed genes (DEGs) (Giordano et al., 2003; Slater et al., 2006; Lombardi et al., 2006), with limited attention paid to gene connections. Weighted gene co-expression network analysis (WGCNA) is a popular method in systems biology that can construct gene networks and detect gene modules (Clarke et al., 2013; Yang et al., 2014; Lee et al., 2015; Goldman et al., 2017; Sun et al., 2017). By analyzing the connectivity between modules and clinical traits, we can determine which modules are associated with which traits. Those genes found in the center of a regulation network usually exhibit more important functions. Thus, the degree of gene connectivity in one module can also be analyzed by the gene-gene interaction/regulation network, from which critical hub genes can be identified.

In this study, we identified genes involved in ACC progression via comprehensive transcriptome-wide analysis of ACC gene expression patterns. We systematically analyzed clusters of genes with similar expression patterns using WGCNA and found the MEblue module to be highly

related to clinical stage. Further analysis identified four hub genes (i.e., *TOP2A*, *TTK*, *CHEK1*, and *CENPA*) from the module that were associated with ACC progression and prognosis. Thus, these hub genes may serve as candidate biomarkers of ACC in clinical treatment and contribute to a greater understanding of ACC progression.

## MATERIALS & METHODS

### Data collection

We obtained gene expression TPM values (Table S1) from the UCSC Xena (https://xena.ucsc.edu/public-hubs/) database, which included 77 ACC samples from TCGA (https://cancergenome.nih.gov/) and 128 normal samples from GTEx (https://www.gtexportal.org/home/). The two databases raw sequencing reads were recalculated with a unifying pipeline. Clinical data were downloaded from TCGA using the 'cgdsr' package in R (v3.1.3) (Null et al., 2009; Jacobsen, 2015).

### DEG screening

Of the 60,498 genes in each sample, we removed genes with a mean TPM ≤ 2.5 (>1 is a common cutoff for determining if an isoform is expressed or not (Liu et al., 2016)) in the cancer and normal samples and thus retained 13,987 genes. For those genes in the samples that showed significant changes, we used analysis of variance (ANOVA) in R (v3.0.2) to determine the variance in genes between the two groups. ANOVA is a collection of statistical models useful for DEG analysis (Alabi et al., 2018; Simona et al., 2015). We obtained 2,953 significant DEGs (Table S2) in ACC with a $p < 0.001$ and |log2 (fold-change)| > 1 cutoff.

### Co-expression network construction by WGCNA

WGCNA (v1.49) can be applied to identify global gene expression profiles as well as co-expressed genes. Therefore, we installed WGCNA package for co-expression analysis using

Bioconductor (http://bioconductor.org/biocLite.R). We used the soft threshold method for Pearson correlation analysis of the expression profiles to determine the connection strengths between two transcripts to construct a weighted network. Average linkage hierarchical clustering was carried out to group transcripts based on topological overlap dissimilarity in network connection strengths. To obtain the correct module number and clarify gene interaction, we set the restricted minimum gene number to 30 for each module and used a threshold of 0.25 to merge the similar modules (see the detailed R script in Supplemental file S1).

**Identification of clinically significant modules**

We used two methods to identify modules related to clinical progression traits. Module eigengenes (MEs) are the major component for principal component analysis of genes in a module with the same expression profile. Thus, we analyzed the relationship between MEs and clinical traits and identified the relevant modules. We used log10 to transform the *p*-value from the linear regression between gene expression and clinical stage, which was defined as gene significance. Average gene significance in a module was defined as module significance.

**Functional and pathway enrichment analysis**

The Database for Annotation Visualization and Integrated Discovery (DAVID) (v6.8) (http://david.abcc.ncifcrf.gov/) was used for functional annotation of genes to better understand their biological functions. All genes in the blue module were uploaded for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, with cutoffs of $p < 0.01$ and $p < 0.05$ established for significant biological processes and pathways, respectively.

**PPI and co-expression analysis**

Genes were uploaded to the STRING (v10.5) (https://string-db.org/) database. Confidence was set to more than 0.4 and other parameters were set to default. We visualized the gene co-

expression network with Cytoscape (v2.7.0) (Shannon et al., 2003).

**Gene expression correlation with stage and survival analysis**

The correlation between gene expression and stage was determined using GEPIA (http://gepia.cancer-pku.cn/index.html) (Tang et al., 2017). The correlation between gene expression and overall survival (OS) was established using the Cox model. A hazard ratio p-value of <0.01 was considered significant. Each gene with higher expression in the ACC samples had corresponding lower survival expectation. The "limma" (Ritchie et al., 2015) R package was used to test significantly expressed gene in GSE10927.

## RESULTS

**Construction and analysis of gene co-expression network with DEGs in ACC**

Genes with a mean TPM ≤ 2.5 were removed from the two groups and the remaining 13,987 genes were used for differential expression analysis with ANOVA. In total, 2,953 significant DEGs were identified with a cutoff of $p < 0.001$ and |log2(fold-change)| > 1 (Fig. 1A), which included 1,181 up-regulated and 1,772 down-regulated genes (Fig. 1B). The 2,953 gene expression levels in ACC and normal samples are shown in the heat map in Fig. 1C (Table S2). Genes with similar expression patterns may participate in similar biological processes or networks (Mao et al., 2009). To better understand the gene expression network during ACC development, the co-expression network of the 2,953 DEGs was analyzed by WGCNA. First, to determine whether all 77 ACC samples were suitable for network analysis, the sample dendrogram and corresponding clinical traits were analyzed. We found that all samples were included in the clusters and passed the cutoff thresholds (Fig. 1D). The power value is a critical parameter that can affect the independence and average connectivity degree of the co-expression modules. Thus, network topology using different soft thresholding powers was screened, with β =

6 (scale free $R^2 = 0.928$) selected for later analysis (Figs. 1E, F). We then constructed the gene co-expression network using WGCNA based on the hierarchical clustering of the calculated dissimilarities, and nine modules were obtained (Fig. 1G, Table S3). We used eigengenes as representative profiles and quantified module similarity by eigengene correlation (Fig. 1H).

**Correlation of blue module with clinical stage and progression**

We investigated whether any module was correlated with clinical stage and tested the relevance between each module and ACC clinical traits. We found that module significance of the blue module was higher than that of any other, implying it had greater correlation with ACC stage (Fig. 2A). The blue module also displayed a positive correlation with ACC clinical stage (r = 0.5, *p* = 6e-06) and negative correlation with OS (r = -0.56, *p* = 3e-07) (Fig. 2B). The eigengene dendrogram and heat map indicated that the MEblue and MEyellow modules were highly correlated with clinical stage (Fig. 2C). Finally, gene significance and module membership were plotted for the blue module (Fig. 2D), which indicated that this module was significantly related to clinical stage.

To determine the function of the 650 genes in the blue module, GO and KEGG function and pathway enrichment analyses were performed by DAVID functional annotation (Huang et al., 2009). For GO biological processes, genes in the module were significantly enriched in cell division (*p* = 1.05e-26) (Fig. 2E, Table S4), whereas for KEGG pathway analysis, the genes were mainly enriched in cell cycle (*p* = 2.7e-19) and DNA replication (*p* = 8.27e-8) (Fig. 2F, Table S5) pathways. These processes and pathways all play critical roles in cancer progression (Tachibana et al., 2005), implying that genes in this module may be involved in ACC progression.

**PPI and co-expression networks to identify hub genes in ACC progression**

To clarify high confidence hub genes, we entered the blue module genes into the STRING

(Szklarczyk et al., 2015) database. The genes were ranked by the PPI nodes and the top 5% of genes (16 genes) were chosen as candidate hub genes (Fig. 3A, Table S6). As highly connected hub genes in a module play important roles in biological processes (Liu et al., 2016), genes in the blue module were ranked by their degree of gene co-expression connectivity (Table S7). To identify genes that may play notable roles in ACC progression, the top 5% of genes (31 genes) (Fig. 3B) in the blue module with the highest connectivity were classified as candidate hub genes for further analysis. Finally, four common genes (i.e., *TOP2A*, *TTK*, *CHEK1*, and *CENPA*) in the two analysis were identified as hub genes in ACC (Fig. 3C). These four genes were highly expressed in ACC samples compared with normal samples (Figs. 3D-3G), indicating that they likely act as oncogenes in ACC. Further analysis of the GSE10927 dataset, which included microarray data of 10 normal samples and 33 ACC samples (Human Genome U133A 2.0 Plus; Affymetrix, Santa Clara, CA, USA) (Giordano et al., 2009), demonstrated that the four genes showed significant high expression in ACC (Figs. S1A-S1D). Furthermore, based on immunoreactivity experiments, *TOP2A* is reported to be highly expressed in ACC (Giordano et al., 2003).

**Significant associations of hub genes with ACC stage and survival**

We investigated the four hub genes to better understand their functions. We found that *TOP2A*, *TTK*, *CHEK1*, and *CENPA* play critical roles in biological processes that are highly correlated with cancer (Dominguez-Brauer et al., 2015), such as DNA topological structure, cell cycle progression, and mitosis (Hoffmann et al., 2011; Liu et al., 2000; De et al., 2013; Thu et al., 2018), thereby suggesting their possible role in cancer development. Further exploration of their expression patterns during ACC clinical progression showed that the levels of these genes were significantly altered with clinical stage and markedly increased at stage III and IV (Figs. 4A-4D).

This correlation between the expression levels of the four genes and ACC progression may be useful in ACC diagnosis.

Tumor prognosis is an important feature in cancer and has attracted considerable attention. To assess the utility of WGCNA at identifying hub genes indicative of ACC, we conducted survival analysis (Figs. 4E-4H). We separated the samples into two groups according to median gene expression levels and performed survival analysis using the Cox model. Survival analysis showed that the expression of all four genes was significantly correlated with OS (Figs. 4E-4H), with higher expression associated with lower patient survival time. The correlation between the hub genes and ACC prognosis suggests that these four genes likely contribute to the progression of ACC.

## DISCUSSION

As ACC exhibits no obvious phenotypic traits during its early stages, diagnosis is often delayed in many patients (Bharwani et al., 2011; Fay et al., 2014). We systematically analyzed gene expression and found potential biomarker genes for ACC diagnosis. To identify genes that may play central roles in ACC progression, gene co-expression network analysis was conducted using WGCNA, which can describe correlation patterns among genes at the RNA level. Based on WGCNA, we obtained nine modules, with each module containing an average of 217 genes. Only 205 genes were unclassified in any module (in grey), accounting for 10.50% of DEGs. In comparison, previous studies have reported an average gene number in each module of 216 to 336 and percentage of genes not found in any module of 5.67%–33.61% of DEGs (Liu X et al., 2017; Liu Z et al., 2018; Yang Q et al., 2018; Zuo Z et al., 2018). In conclusion, our WGCNA results were comparable. We identified four hub genes (i.e., *TOP2A*, *TTK*, *CHEK1*, and *CENPA*) in the network center related to gene regulation and possible carcinogenesis.

Genes located in the central position of a gene-gene interaction network likely exhibit more important functions than other genes. Further investigation found that these four hub genes contribute to several tumor types indeed. For instance, *TOP2A* (topoisomerase II alpha), a specific marker of cell proliferation, is the primary molecular target of anthracyclines used for treating breast cancer (Villman et al., 2006; Wang et al., 2012). *TTK*, also known as monopolar spindle 1 (*MPS1*), plays a key role in cancer cell growth and proliferation, with its inhibition able to decrease tumor aggressiveness (Al-Ejeh et al., 2014; Maire et al., 2015; Zhu et al., 2018). *CHEK1* (checkpoint kinase 1), a conserved serine/threonine kinase, plays a key role in tumor growth promotion (Zhang and Hunter 2013). Furthermore, inhibition of *CHEK1* expression by UCN-01, CEP-3891 (Zhu et al., 2018), AZD7762, or LY2606368 inhibitors (Manic et al., 2017) can prevent the proliferation of cancer cells (Bryant et al., 2014; Schuler et al., 2017). *CENPA* (centromere protein A), a histone H3 variant, is highly expressed in cancers, including breast, colorectal, liver, lung, ovarian, and osteosarcoma (Athwal et al., 2015; Sun et al., 2016; Filipescu et al., 2017). In addition, inhibition of *CENPA* expression in cancer cells can reduce sphere forming ability, proliferation, and cell viability (Behnan et al., 2016). Here, our study revealed that the expression levels of all four hub genes were significantly correlated with ACC progression (Figs. 4A-4D) and OS (Figs. 4E-4H), suggesting their critical function in ACC. Our results indicated that these four genes may play key roles in ACC tumorigenesis. However, the specific functions of these genes that contribute to ACC cell proliferation, differentiation, and metastasis need further study.

## CONCLUSIONS

Based on gene co-expression network analysis, we identified four hub genes that likely contribute to the progression of ACC. The expressions of the four hub genes demonstrated significant

correlation with ACC clinical stage and prognosis (Figs. 4A-4H). Thus, these four genes may act as potential biomarkers in predicting clinical outcomes and diagnosis of ACC. Furthermore, inhibitors of *TOP2A*, *TTK*, and *CHEK1*, which are already used for treating certain cancers, could potentially be used in ACC treatment. Further experimental and clinical studies are required to extend these findings.

## ACKNOWLEDGMENTS

We thank Qiong-Hua Gao for suggestions in modifying the paper and Christine Watts for help in honing the manuscript.

## REFERENCES

Alabi N, Sheka D, Gupta M, Kannappan S. 2018. Identification of a Pathway- Based 5-Gene Expression Signature for Predicting Outcomes in Gastric Cancer. J Proteomics Bioinform. 11:161-168.

Allolio B, Fassnacht M. 2006. Adrenocortical Carcinoma: Clinical Update. The Journal of Clinical Endocrinology & Metabolism. **91**(6):2027-2037.

Al-Ejeh F, Simpson PT, Saunus JM, Klein K, Kalimutho M, Shi W, Miranda M, Kutasovic J, Raghavendra A, Madore J, Reid L, Krause L, Chenevix-Trench G, Lakhani SR, Khanna KK. 2014. Meta-analysis of the global gene expression profile of triple-negative breast cancer identifies genes for the prognostication and treatment of aggressive breast cancer. Oncogenesis. **3**(10):e124.

Assié G, Letouzé E, Fassnacht M, Jouinot A, Luscap W, Barreau O, Omeiri H, Rodriguez S, Perlemoine K, René-Corail F, Elarouci N, Sbiera S, Kroiss M, Allolio B, Waldmann J, Quinkler M, Mannelli M, Mantero F, Papathomas T, De Krijger R, Tabarin A, Kerlan V, Baudin E, Tissier F, Dousset B, Groussin L, Amar L, Clauser E, Bertagna X, Ragazzon B, Beuschlein F, Libé R, de

253 Reyniès A, Bertherat J. 2014. Integrated genomic characterization of adrenocortical carcinoma.
254 Nature Genetics. **46**(6):607-612.
255 Athwal RK, Walkiewicz MP, Baek S, Fu S, Bui M, Camps J, Ried T, Sung MH, Dalal Y. 2015.
256 CENP-A nucleosomes localize to transcription factor hotspots and subtelomeric sites in human
257 cancer cells. Epigenetics & Chromatin,8,1(2015-01-13) **8**(1):2.
258 Adrenocortical Carcinoma: A Footprint of a Rare Cancer. J Genomics. **5**(19):99-118.
259 Behnan J, Grieg Z, Joel M, Ramsness I, Stangeland B. 2016. Gene knockdown of CENPA
260 reduces sphere forming ability and stemness of glioblastoma initiating cells. Neuroepigenetics.
261 **7**(C):6-18.
262 Bharwani N, Rockall AG, Sahdev A, Gueorguiev M, Drake W, Grossman AB, Reznek RH. 2011.
263 Adrenocortical carcinoma: the range of appearances on CT and MRI. Ajr American Journal of
264 Roentgenology. **196**(6):706-714.
265 Bryant C, Rawlinson R and Massey AJ. 2014. Chk1 inhibition as a novel therapeutic strategy for
266 treating triple-negative breast and ovarian cancers. BMC Cancer,14,1(2014-08-07) **14**(1):570.
267 Cherradi N. 2014. microRNAs as Potential Biomarkers in Adrenocortical Cancer: Progress and
268 Challenges. Front Endocrinol. **6**:195.
269 Chortis V, Taylor AE, Doig CL, Walsh MD, Meimaridou E, Jenkinson C, Rodriguez-Blanco G,
270 Ronchi CL, Jafri A, Metherell LA, Hebenstreit D, Dunn WB, Arlt W, Foster PA. 2018.
271 Nicotinamide Nucleotide Transhydrogenase as a Novel Treatment Target in Adrenocortical
272 Carcinoma. Endocrinology. **159**(8):2836.
273 Clarke C, Madden SF, Doolan P, Aherne ST, Joyce H, O'Driscoll L, Gallagher WM, Hennessy
274 BT, Moriarty M, Crown J, Kennedy S, Clynes M. 2013. Correlating transcriptional networks to
275 breast cancer survival: a large-scale coexpression analysis. Carcinogenesis. **34**(10):2300-2308.

de Resende MF, Vieira S, Chinen LT, Chiappelli F, da Fonseca FP, Guimarães GC, Soares FA, Neves I, Pagotty S, Pellionisz PA, Barkhordarian A, Brant X, Rocha RM. 2013. Prognostication of prostate cancer based on TOP2A protein and gene assessment: TOP2A in prostate cancer. Journal of Translational Medicine. **11**(1):1-9.

Dominguez-Brauer C, Thu KL, Mason JM, Blaser H, Bray MR, Mak TW. 2015. Targeting Mitosis in Cancer: Emerging Strategies. Molecular Cell. **60**(4):524-536.

Else T, Williams AR, Sabolch A, Jolly S, Miller BS, Hammer GD. 2014. Adjuvant Therapies and Patient and Tumor Characteristics Associated With Survival of Adult Patients With Adrenocortical Carcinoma. The Journal of Clinical Endocrinology & Metabolism. **99**(2):455-461.

Fassnacht M, Johanssen S, Quinkler M, Bucsky P, Willenberg HS, Beuschlein F, Terzolo M, Mueller HH, Hahner S, Allolio B. 2009. Limited prognostic value of the 2004 International Union Against Cancer staging classification for adrenocortical carcinoma: proposal for a Revised TNM Classification. Cancer. **115**:243–250.

Fassnacht M, Kroiss M, Allolio B. 2013. Update in adrenocortical carcinoma. J Clin Endocrinol Metab. **98**:4551–4564.

Fay AP, Elfiky A, Teló GH, McKay RR, Kaymakcalan M, Nguyen PL, Vaidya A, Ruan DT, Bellmunt J, Choueiri TK. 2014. Adrenocortical carcinoma: the management of metastatic disease. Critical Reviews in Oncology/hematology. **92**(2):123-132.

Filipescu D, Naughtin M, Podsypanina K, Lejour V, Wilson L, Gurard-Levin ZA, Orsi GA, Simeonova I, Toufektchan E, Attardi LD, Toledo F, Almouzni G. 2017. Essential role for centromeric factors following p53 loss and oncogenic transformation. Genes & Development. **31**(5):463.

298 Giordano TJ, Thomas DG, Kuick R, Lizyness M, Misek DE, Smith AL, Sanders D, Aljundi RT,
299 Gauger PG, Thompson NW, Taylor JM, Hanash SM. 2003 Distinct transcriptional profiles of
300 adrenocortical tumors uncovered by DNA microarray analysis. Am. J. Pathol. 162(2):521-531.
301 Giordano TJ, Kuick R, Else T, Gauger PG, Vinco M, Bauersfeld J, Sanders D, Thomas DG,
302 Doherty G, Hammer G. 2009. Molecular classification and prognostication of adrenocortical
303 tumors by transcriptome profiling. Clinical Cancer Research. **15**(2):668-676.
304 Goldman M, Craft B, Brooks AN, Zhu J, Haussler D. 2017. Abstract 2584: The UCSC Xena
305 system for cancer genomics data visualization and interpretation. Cancer Research. **77**(13
306 Supplement):2584-2584.
307 Greenhill C. 2016. Adrenal gland: The genetics of adrenocortical carcinoma revealed. Nature
308 Reviews Endocrinology. **12**(8).
309 Hoang MP, Ayala AG and Albores-Saavedra J. 2002. Oncocytic adrenocortical carcinoma: a
310 morphologic, immunohistochemical and ultrastructural study of four cases. Mod Pathol.
311 **15**(9):973-978.
312 Hoffmann S, Dumont M, Barra V, Ly P, Nechemia-Arbely Y, McMahon MA, Hervé S,
313 Cleveland DW, Fachinetti D. 2011. CENP-A is dispensable for mitotic centromere function after
314 initial centromere/kinetochore assembly. Cell Reports. 17(9):2394-2404.
315 Huang DW, Sherman BT and Lempicki RA 2009. Bioinformatics enrichment tools: paths toward
316 the comprehensive functional analysis of large gene lists. Nucleic Acids Research. **37**(1):1.
317 Jacobsen A. 2015. cgdsr: R-Based API for Accessing the MSKCC Cancer Genomics Data Server
318 (CGDS).
319 Kiseljak-Vassiliades K, Zhang Y, Bagby SM, Kar A, Pozdeyev N, Xu M, Gowan K, Sharma V,
320 Raeburn CD, Albuja-Cruz M, Jones KL, Fishbein L, Schweppe RE, Somerset H, Pitts TM, Leong

S, Wierman ME. 2018. Development of new preclinical models to advance adrenocortical carcinoma research. Endocrine-related cancer: ERC-17-0447.

Koduru SV, Leberfinger AN and Ravnic DJ. 2017. Small Non-coding RNA Abundance in

Lee YS, Hwang SG, Kim JK, Park TH, Kim YR, Myeong HS, Kwon K, Jang CS, Noh YH, Kim SY. 2015. Topological network analysis of differentially expressed genes in cancer cells with acquired gefitinib resistance. Cancer Genomics & Proteomics. **12**(3):153.

Liu J, Jing L, Tu X. 2016. Weighted gene co-expression network analysis identifies specific modules and hub genes related to coronary artery disease. Bmc Cardiovascular Disorders. **16**(1):54.

Liu Q, Guntuku S, Cui XS, Matsuoka S, Cortez D, Tamai K, Luo G, Carattini-Rivera S, DeMayo F, Bradley A, Donehower LA, Elledge SJ. 2000. Chk1 is an essential kinase that is regulated by Atr and required for the G(2)/M DNA damage checkpoint. Genes Dev. 14(12):1448-1459.

Lombardi CP, Raffaelli M, Pani G, Maffione A, Princi P, Traini E, Galeotti T, Rossi ED, Fadda G, Bellantone R. 2006. Gene expression profiling of adrenal cortical tumors by cDNA macroarray analysis. Results of a preliminary study.Biomed. Pharmacother. **60**(4):186–190.

Liu X, Hu AX , Zhao JL, Chen FL. 2017. Identification of key gene modules for in human osteosarcoma by co-expression analysis weighted gene coexpression network analysis (wgcna).Journal of Cellular Biochemistry. 118(11), 3953.

Liu Z, Li M, Fang X, Shen L, Yao W, Fang Z, Chen J, Feng X, Hu, Zeng Z, Lin C, Weng J, Lai Y, Yi G. 2018. Identification of surrogate prognostic biomarkers for allergic asthma in nasal epithelial brushing samples by wgcna. Journal of Cellular Biochemistry.

Maire V, Baldeyron C, Richardson M, Tesson B, Vincent-Salomon A, Gravier E, Marty-Prouvost B, De Koning L, Rigaill G, Dumont A, Gentien D, Barillot E, Roman-Roman S, Depil S,

Cruzalegui F, Pierré A, Tucker GC, Dubois T. 2015. TTK/hMPS1 is an attractive therapeutic target for triple-negative breast cancer. Plos One. **8**(5):e63712.

Manic G, Signore M, Sistigu A, Russo G, Corradi F, Siteni S, Musella M, Vitale S, De Angelis ML, Pallocca M, Amoreo CA, Sperati F, Di Franco S, Barresi S, Policicchio E, De Luca G, De Nicola F, Mottolese M, Zeuner A, Fanciulli M, Stassi G, Maugeri-Saccà M, Baiocchi M, Tartaglia M, Vitale I, De Maria R. 2017. CHK1-targeted therapy to deplete DNA replication-stressed, p53-deficient, hyperdiploid colorectal cancer stem cells. Gut. **67**(5).

Mao L, Van Hemert JL, Dash S, Dickerson JA. 2009. Arabidopsis gene co-expression network and its functional modules. BMC Bioinformatics. **10**:346.

Null RCTR, Team R, Null RCT. 2009. R-a language and environment for statistical computing. Computing. **1**:12-21.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research. 43(7):e47.

Tachibana KE, Gonzalez MA, Coleman N. 2005. Cell cycle dependent regulation of dna replication and its relevance to cancer pathology. Journal of Pathology. 205(2):123-129.

Thu KL, Silvester J, Elliott MJ, Ba-Alawi W, Duncan MH, Elia AC, Mer AS, Smirnov P, Safikhani Z, Haibe-Kains B, Mak TW, Cescon DW. 2018. Disruption of the anaphase-promoting complex confers resistance to TTK inhibitors in triple-negative breast cancer. Proceedings of the National Academy of Sciences. 115(7):201719577.

Schuler F, Weiss JG, Lindner SE, Lohmüller M, Herzog S, Spiegl SF, Menke P, Geley S, Labi V, Villunger A. 2017. Checkpoint kinase 1 is essential for normal B cell development and lymphomagenesis. Nature Communications. **8**(1):1697.

Monterisi S, D'Ario G, Dama E, Rotmensz N, Confalonieri S, Tordonato C, Troglio F, Bertalot G, Maisonneuve P, Viale G, Nicassio F, Vecchi M, Di Fiore PP, Bianchi F. 2015. Mining cancer gene expression databases for latent information on intronic microRNAs. Molecular Oncology. **9**(2):473-487

Slater EP, Diehl SM, Langer P, Samans B, Ramaswamy A, Zielke A, Bartsch DK. 2006. Analysis by cDNA microarrays of gene expression patterns of human adrenocortical tumors. Eur. J. Endocrinol. **154**(4):587-598.

Soon PS, McDonald KL, Robinson BG, Sidhu SB. 2008. Molecular markers and the pathogenesis of adrenocortical cancer. Oncologist. **13**(5):548.

Sun Q, Zhao H, Zhang C, Hu T, Wu J, Lin X, Luo D, Wang C, Meng L, Xi L, Li K, Hu J, Ma D, Zhu T. 2017. Gene co-expression network reveals shared modules predictive of stage and grade in serous ovarian cancers. Oncotarget. **8**(26):42983-42996.

Sun X, Clermont PL, Jiao W, Helgason CD, Gout PW, Wang Y, Qu S. 2016. Elevated expression of the centromere protein-A (CENP-A)-encoding gene as a prognostic and predictive biomarker in human cancers. International Journal of Cancer. **139**(4):899-907.

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C. 2015. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Research. **43**(Database issue):D447.

Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. 2017. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Research. **45**(Web Server issue).

Villman K, Sjöström J, Heikkilä R, Hultborn R, Malmström P, Bengtsson NO, Söderberg M, Saksela E, Blomqvist C. 2006. TOP2A and HER2 gene amplification as predictors of response to anthracycline treatment in breast cancer. Acta Oncologica. **45**(5):590-596.

Wang J, Xu B, Yuan P, Zhang P, Li Q, Ma F, Fan Y. 2012. TOP2A amplification in breast cancer is a predictive marker of anthracycline-based neoadjuvant chemotherapy efficacy. Breast Cancer Research & Treatment. **135**(2):531-537.

Yang Q, Wang R, Wei B, Peng CG, Wang L, Hu GZ, kong DL, Du C. 2018. Candidate biomarkers and molecular mechanism investigation for glioblastoma multiforme utilizing wgcna. BioMed Research International. 2018:1-10.

Yang Y, Han L, Yuan Y, Li J, Hei N, Liang H. 2014. Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nature Communications. **5**(1):3231.

Zhang Y and Hunter T. 2013. Roles of Chk1 in cell biology and cancer therapy. International Journal of Cancer. **134**(5):1013-1023.

Zheng S, Cherniack AD, Dewal N, Moffitt RA, Danilova L, Murray BA, Lerario AM, Else T, Knijnenburg TA, Ciriello G, Kim S, Assie G, Morozova O, Akbani R, Shih J, Hoadley KA, Choueiri TK, Waldmann J, Mete O, Robertson AG, Wu HT, Raphael BJ, Shao L, Meyerson M, Demeure MJ, Beuschlein F, Gill AJ, Sidhu SB, Almeida MQ, Fragoso MCBV, Cope LM, Kebebew E, Habra MA, Whitsett TG, Bussey KJ, Rainey WE, Asa SL, Bertherat J, Fassnacht M, Wheeler DA. 2016. Comprehensive Pan-Genomic Characterization of Adrenocortical Carcinoma. Cancer Cell. **29**(5):723-736.

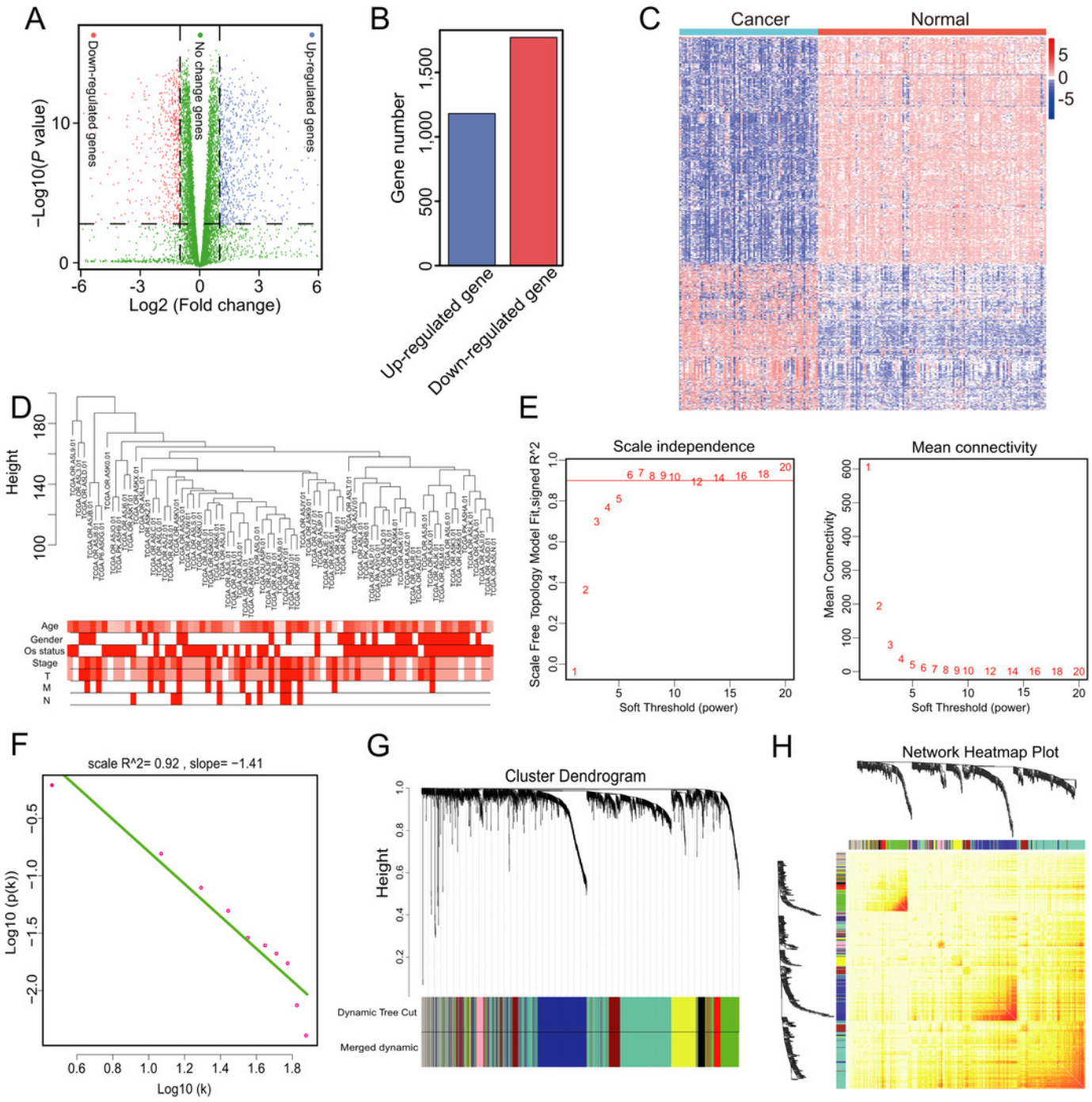
Zhu D, Xu S, Deyanat-Yazdi G, Peng SX, Barnes LA, Narla RK, Tran T, Mikolon D, Ning Y, Shi T, Jiang N, Raymon HK, Riggs JR, Boylan JF. 2018. Synthetic Lethal Strategy Identifies a

412 Potent and Selective TTK and CLK2 Inhibitor for Treatment of Triple-negative Breast Cancer
413 with a Compromised G1/S Checkpoint. Molecular Cancer Therapeutics.
414 Zuo Z, Shen JX, Pan Y, Pu J, Li YG, Shao XH, Wang WP. 2018. Weighted Gene Correlation
415 Network Analysis (WGCNA) Detected Loss of MAGI2 Promotes Chronic Kidney Disease (CKD)
416 by Podocyte Damage. Cell Physiol Biochem. 51(1):244-261.

# Figure 1

Nine modules obtained following WGCNA analysis of DEGs in ACC

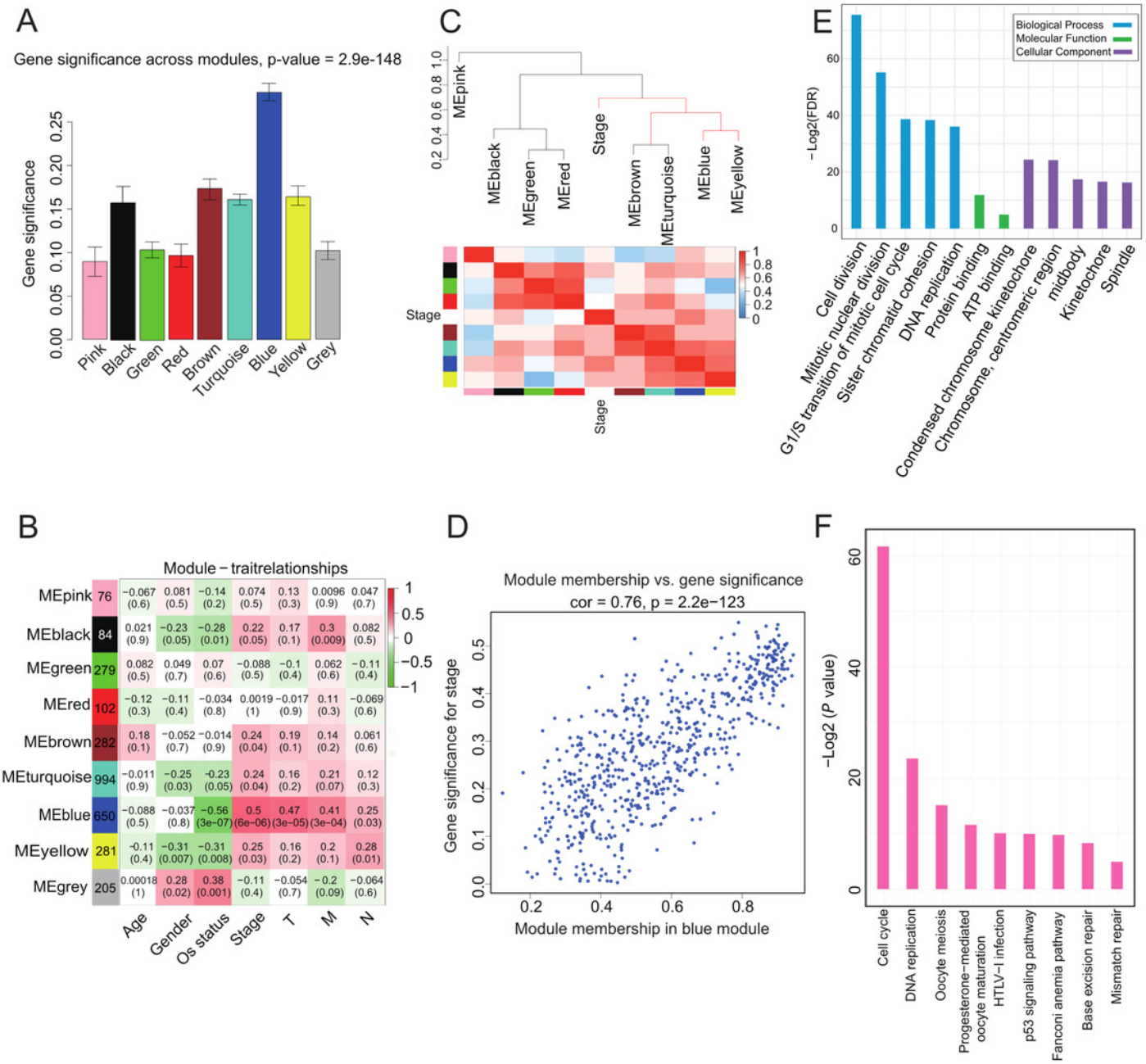
(A) X-axis represents log2 fold-changes and y-axis represents negative logarithm to the base 10 of the p-values. Black vertical and horizontal dashed lines reflect filtering criteria (FC = ±1 and *p*-value = 0.001). (B) Red and blue bars are number of significantly down-regulated (n = 1,772) or up-regulated genes (n = 1,181) in ACC compared with non-tumor samples. (C) Heat map shows all DEGs in ACC and GTEx. The Log2(TPM+0.001) expression level of each gene profile from each sample is represented by color. (D) Sample clustering was conducted to detect outliers. This analysis was based on the expression data of DEGs between tumor and non-tumor samples in ACC. All samples are located in the clusters and pass the cutoff thresholds. Color intensity is proportional to sample age, gender, status, and stage. (E) Soft-thresholding power analysis was used to obtain the scale-free fit index of network topology. (F) Scale free topology when β = 6. (G) Hierarchical cluster analysis was conducted to detect co-expression clusters with corresponding color assignments. Each color represents a module in the constructed gene co-expression network by WGCNA. (H) Heat map depicts the Topological Overlap Matrix (TOM) among 500 randomly selected genes from the DEG weighted co-expression network. Light color represents lower overlap and red represents higher overlap.


A
Down-regulated genes
No change genes
Up-regulated genes
−Log10(P value)
Log2 (Fold change)
B
Gene number
Up-regulated gene
Down-regulated gene
C
Cancer
Normal
D
Height
Age
Gender
Os status
Stage
T
M
N
E
Scale independence
Scale Free Topology Model Fit,signed R^2
Soft Threshold (power)
Mean connectivity
Mean Connectivity
Soft Threshold (power)
F
scale R^2= 0.92 , slope= −1.41
Log10 (p(k))
Log10 (k)
G
Cluster Dendrogram
Height
Dynamic Tree Cut
Merged dynamic
H
Network Heatmap Plot

# Figure 2

Correlation of Blue module with clinical stage

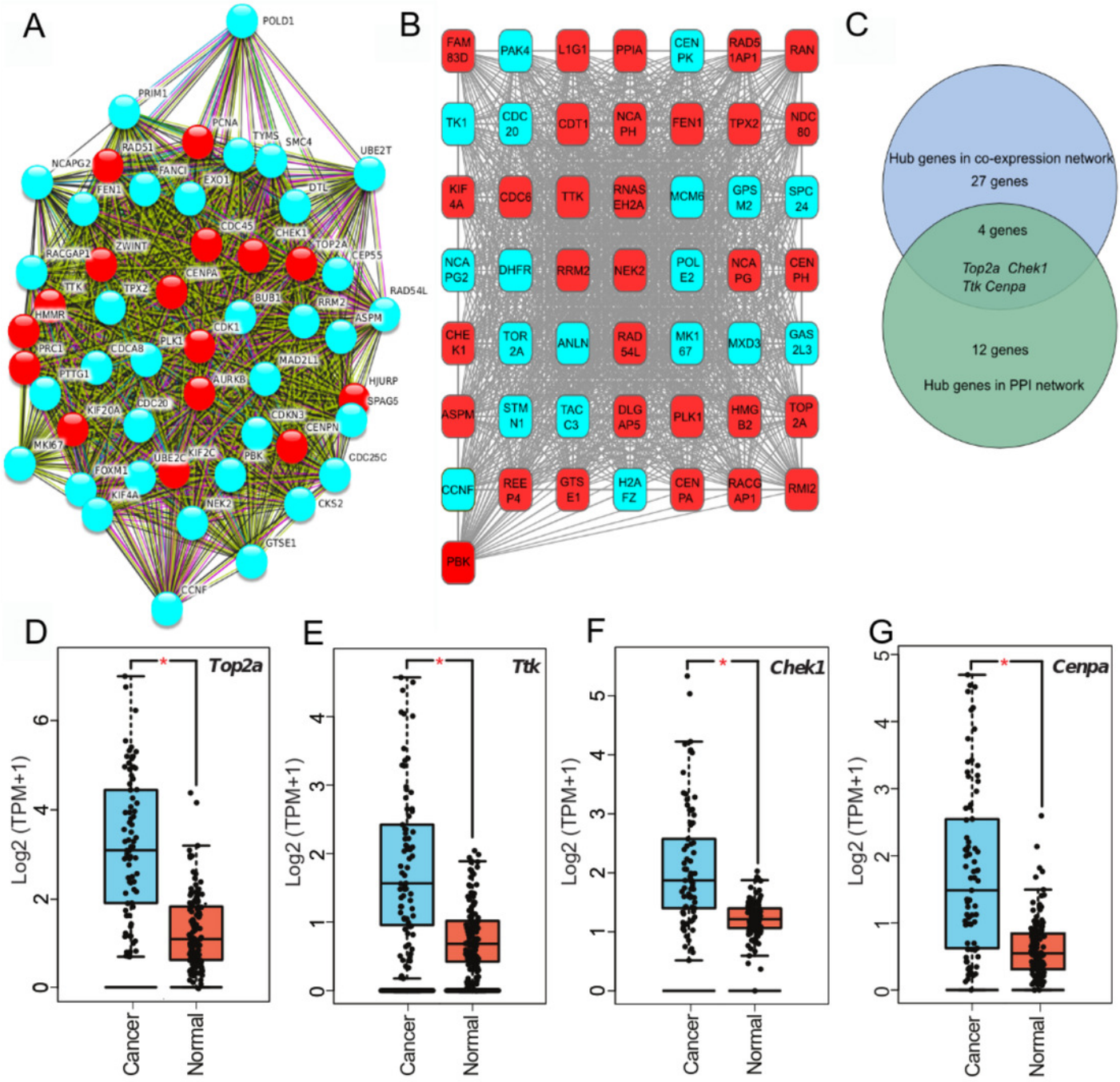
(A) Bar plot of mean gene significance across genes associated with ACC stage in the module. (B) Heat map with each cell containing the *p*-value correlation from the linear mixed-effects model. Row corresponds to module; column corresponds to ACC clinical traits. Results indicate that MEblue is highly related to patient stage. (C) The dendrogram shows the relation of modules with stage and the heatmap shows the eigengene adjacency. (D) Correlation between MEblue membership and gene significance. (E) GO enrichment analysis of 650 genes in MEblue identified biological processes related to cell proliferation. Y-axis represents significance of enrichment results transformed to '–log(P-value)'. (F) KEGG enrichment analysis of 650 genes in MEblue identified pathways related to cell cycle and DNA replication.


A
Gene significance across modules, p-value = 2.9e-148
Gene significance
0.00
0.05
0.10
0.15
0.20
0.25
Pink
Black
Green
Red
Brown
Turquoise
Blue
Yellow
Grey
C
MEpink
MEblack
MEgreen
MEred
Stage
MEbrown
MEturquoise
MEblue
MEyellow
Stage
Stage
E
Biological Process
Molecular Function
Cellular Component
−Log2(FDR)
Cell division
Mitotic nuclear division
G1/S transition of mitotic cell cycle
Sister chromatid cohesion
DNA replication
Protein binding
ATP binding
Condensed chromosome kinetochore
Chromosome, centromeric region
midbody
Kinetochore
Spindle
B
Module − traitrelationships
MEpink 76
MEblack 84
MEgreen 279
MERed 102
MEbrown 282
MEturquoise 994
MEblue 650
MEyellow 281
MEgrey 205
Age
Gender
Os status
Stage
T
M
N
D
Module membership vs. gene significance
cor = 0.76, p = 2.2e−123
Gene significance for stage
Module membership in blue module
F
−Log2 (P value)
Cell cycle
DNA replication
Oocyte meiosis
Progesterone−mediated oocyte maturation
HTLV−I infection
p53 signaling pathway
Fanconi anemia pathway
Base excision repair
Mismatch repair

# Figure 3

Four hub genes identified through PPI and gene-gene connection network

(A) PPI network of genes in MEblue. Intersection of top 50 genes in MEblue is shown, red nodes are hub genes of the network. (B) Co-expression network of top 50 genes in MEblue, red nodes are hub genes of the network. (C) Venn diagram shows common hub genes between co-expression and PPI network analyses. (D-G) Four hub genes significantly expressed in ACC samples compared with corresponding GTEx tissue samples.


A
POLD1
PRIM1
PCNA
TYMS
SMC4
UBE2T
RAD51
NCAPG2
FANCI
EXO1
FEN1
DTL
CDC45
CHEK1
TOP2A
CEP55
ZWINT
RACGAP1
CENPA
TTK
TPX2
BUB1
RRM2
RAD54L
ASPM
HMMR
CDK1
PLK1
PRC1
CDCA8
MAD2L1
PTTG1
AURKB
HJURP
SPAG5
CDC20
KIF20A
CDKN3
CENPN
MKI67
UBE2C
KIF2C
PBK
CDC25C
FOXM1
KIF4A
NEK2
CKS2
GTSE1
CCNF
B
FAM 83D
PAK4
L1G1
PPIA
CEN PK
RAD5 1AP1
RAN
TK1
CDC 20
CDT1
NCA PH
FEN1
TPX2
NDC 80
KIF 4A
CDC6
TTK
RNAS EH2A
MCM6
GPS M2
SPC 24
NCA PG2
DHFR
RRM2
NEK2
POL E2
NCA PG
CEN PH
CHE K1
TOR 2A
ANLN
RAD 54L
MK1 67
MXD3
GAS 2L3
ASPM
STM N1
TAC C3
DLG AP5
PLK1
HMG B2
TOP 2A
CCNF
REE P4
GTS E1
H2A FZ
CEN PA
RACG AP1
RMI2
PBK
C
Hub genes in co-expression network
27 genes
4 genes
Top2a Chek1
Ttk Cenpa
12 genes
Hub genes in PPI network
D
Top2a
E
Ttk
F
Chek1
G
Cenpa
Log2 (TPM+1)
Cancer
Normal

# Figure 4

Significant correlation between hub gene expression with pathological stage and survival

(A-D) Significant correlation between expression levels of *TOP2A*, *TTK*, *CHEK1*, and *CENPA* with ACC pathological stage. (E-H) Survival plot of OS in ACC. Higher expression (red line) of *TOP2A*, *TTK*, *CHEK1*, and *CENPA* indicates poorer prognosis. HR: hazard ratio.