Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

KIF4A and STIL act as pan-cancer diagnostic and prognostic biomarkers by bioinformatics analysis and verifcation in osteosarcoma

Jiankang Pan
Second Xiangya Hospital
Yong Luo
Second Xiangya Hospital
Xiaohua Lei (leixiaohua@csu.edu.cn)
Second Xiangya Hospital
Xinzhan Mao
Second Xiangya Hospital

Research article

Keywords: bioinformatics analysis, chromosome instability biomarker, immunohistochemistry, KIF4A, osteosarcoma, STIL

Posted Date: February 5th, 2020

DOI: https://doi.org/10.21203/rs.2.22719/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: Cancer is a disease of abnormal cell proliferation that is caused by aberrantly expressed cancer-related genes. Transcriptome analysis to identify differential gene expression between tumors and paring normal tissue can help to reveal the key process of cell proliferation in cancer.

Methods: By screening RNA-seq data covering twelve types of cancer and their paring normal tissue from the cancer genome atlas (TCGA), we identified twenty-eight cell-cycle-related genes that were overexpressed in no less than ten types of cancers.

Results: Among them, kinesin family member 4A (KIF4A) and stromal tumor infiltrating lymphocyte (STIL) exhibited upregulated expression in eleven types of cancer, making them promising pan-cancer diagnostic markers. The high expression of either KIF4A or STIL indicates poor prognosis in four types of cancer. We further conducted the expression of KIF4A and STIL by immunohistochemistry（IHC） in osteosarcoma（OS） and paring normal tissue samples from a cohort of 57 osteosarcoma patients,which confirmed the results of bioinformatics analysis .

Conclusions: KIF4A and STIL can serve as pan-cancer diagnostic and prognostic markers. The fact that they both function in the mitotic chromosome segregation process further emphasizes the importance of mitosis regulation and chromosome stability in carcinogenesis and cancer development.

1. Introduction

Despite the rapid development of genomic medicine, cancer therapy is unsatisfactory, and cancer is still the main cause of death [1]. Essentially, cancer is a disease of abnormal cell proliferation caused by accumulated genomic mutations [2]. Previous studies in molecular medicine divided cancers into increasingly specific molecular subtypes according to their specific genomic status or transcriptome pattern, which greatly improved the prognosis of many cancer types [3, 4]. However, in many cases, it is still difficult to distinguish benign lesions from malignant ones, and effective therapies are still urgently needed. Thus, in accordance with the current trend of developing more precise medicines, pan-cancer analysis to identify similarities across cancer types is necessary. It is known that cancers share common biological behaviors like the epithelial–mesenchymal transition (EMT) [5–7]. Practically, understanding the shared characteristics of cancers is beneficial. For instance, it can help us to apply treatment known to be effective for one cancer to various other cancer types, as in the case of PD-L1 blockade therapy, which was originally applied for melanoma and was later found effective for multiple cancers [8].

In this paper, we focus on the shared genes or pathways that are differentially expressed in cancers and could serve as diagnostic and prognostic markers across cancer types. Transcriptome data from 615 pairs of tumors and paring normal tissue covering 12 types of cancers are used for differential expression gene (DEG) analysis. Twenty-eight DEGs are found to be overexpressed in no less than 10 types of cancers and are selected for further analysis. Through gene ontology (GO) molecular pathway analysis, all of them are found to be cell-cycle-related genes. Among them, KIAF4 and STIL, which both participate in chromosome segregation and are highly correlated with each other, are overexpressed in eleven types of cancers, and their elevated expression is positively correlated with poor survival in four cancer types. High expression of KIF4A and STIL could serve as a diagnostic and prognostic marker in various cancers.

2. Materials And Methods

2.1. The Cancer Genome Atlas (TCGA) Dataset

Data acquisition and analysis were conducted using R software (version 3.5.1 or above), unless otherwise mentioned. RNA-seq and clinical data were downloaded using the TCGAbioloinks R/Bioconductor package (version 2.10.5) [9]. Generally, we used TCGAbiolinks to download all available samples with Illumina HiSeq RNASeqV2 data from 33 cancer types. To ensure the validity of results, only those cancer types with more than 30 pairs of normal-cancer tissue data were included in our study. Thus, 12 types of cancer were selected from a total of 33 cancer types.

2.2. Data Analysis

The fragments per Kilobase of transcript per million fragments mapped (FPKM) is the most commonly used normalization method for RNA transcript reads. Upper-quantile normalized FPKM (FPKM-UQ) uses upper-quantile gene counts rather than total gene counts for normalization and is believed to have better sensitivity in gene differential expression identification [10]. In this study, FPKM-UQ RNA-seq data were downloaded and prepared using the GDCquery, GDCdownload, and GDCprepare functions.

2.3. Gene Network Analysis

GeneMANIA [11] was used to search for interacting genes of KIF4A and STIL based on physical interactions, coexpression, colocalization, pathways, shared protein domains, and predicted interactions.

2.4. Pathway Enrichment Analysis

The ClusterProfiler [12] R-package was used to analyze the GO biological pathway enrichment of the hub genes. A pathway-enrichment plot was drawn with the dotplot function.

2.5. Patient characteristics

In order to confirm the results of bioinformatics analysis, we further conducted the expression of KIF4A and STIL by immunohistochemistry in osteosarcoma.A total of 57 patients who underwent surgical treatment at our hospital and were diagnosed with osteosarcoma by post-operative pathological examination from June 2012 to March 2014 were included in this study. A total of 57 specimens of paracancerous tissue were also collected. This study was approved by the Ethics Committee of our hospital. Written informed consent was obtained from all patients prior to enrollment in the present study. The following general data were collected from patients: name, age, tumor size, enneking staging,and survival time(Table 1). Patients who received radiotherapy or chemotherapy before surgery were excluded from this study.

2.6. Immunohistochemistry(IHC) and scoring

Four micrometers specimen sections were deparaffinized and rehydrated. After treatment with endogenous peroxidase blocking solution (PBS), they were treated with specific antibodies against KIF4A(ab122227,Abcam) and STIL (ab222838, Abcam) overnight at 4 °C. After they were washed with PBS, the samples were treated with horseradish peroxidase-conjugated anti-rabbit IgG and were then stained with diaminobenzidine. Expression levels were scored by multiplying the percentage of positive cells by the staining intensity. Two experienced (blinded) pathologists independently interpreted the results. The percent positivity was scored as 0 if < 5% (negative), 1 if 5–30% (sporadic), 2 if 30–70% (focal) and 3 if > 70% (diffuse) of the cells were stained; and staining intensity was scored as 0 for no staining, 1 for weak to moderate staining and 2 for strong staining. A score ≥ 2 was regarded as ‘high’, and a score 1 was regarded as ‘low’ in immunohistochemical staining.

2.7. Statistical Analysis

All bioinformatic analyses were conducted using R software. A paired t test was used to assess differences in mRNA expression between tumors and matched normal tissue. The SPSS22.0 software program was used for the statistical analysis. The association of clinical variables was analyzed by the Pearson chi-square or Spearman-rho test. Univariate and multivariate analyses were performed by using the Cox proportional hazard model.the Kaplane-Meier curve was used for the survival analysis. A value of $P < 0.05$ was considered statistically significan.

2.8. Code Availability

All analysis codes used are freely available at https://github.com/hutaobo/prognosis.

3. Results

3.1. Cell-Cycle-Related Genes Are Extensively Overexpressed in Various Types of Cancers

The expression of 57,035 genes in 12 types of cancer and their paring normal tissue from TCGA were analyzed. The 12 analyzed types of cancers were breast invasive carcinoma (BRCA) [13], kidney renal clear cell carcinoma (KIRC) [14], lung adenocarcinoma

(LUAD) [15], stomach adenocarcinoma (STAD) [16], colon adenocarcinoma (COAD) [17], kidney renal papillary cell carcinoma (KIRP) [18], lung squamous cell carcinoma (LUSC) [19], thyroid carcinoma (THCA) [20], head and neck squamous cell carcinoma (HNSC) [21], liver hepatocellular carcinoma (LIHC) [22], prostate adenocarcinoma (PRAD) [23], and uterine corpus endometrial carcinoma (UCEC) [24]. The sample number for each cancer type is listed (Table 2).

For statistical analysis, only those reaching genome-wide significance were included and defined as DEGs p value less than 5×10^{-8}) [25]. About three-quarters of the genes did not show elevated expression in any cancer type. For the remaining genes, about 60% of them (n = 8434) had elevated expression in only one cancer type. Thus, only 9% of all the analyzed genes (n = 5343) showed elevated expression in more than one type of cancer. Among them, 28 genes were found to be overexpressed in no less than 10 types of cancer (Table 3). It was noted that, in stomach adenocarcinoma and thyroid carcinoma, only six and thirteen of the 28 pan-cancer DEGs were overexpressed, respectively, while in ten other cancer types, at least 20 DEGs were overexpressed. This difference was not caused by sample size, as thyroid carcinoma had the second largest sample size. Therefore, this indicates the existence of an intrinsic difference in stomach adenocarcinoma and thyroid carcinoma compared with other cancer types. The GO molecular pathway analysis showed that the 28 pan-cancer DEGs were enriched in 116 biological processes, most of which were cell-cycle-related processes (Table S1). This is no surprise, since cancer is essentially a disease of uncontrolled cell proliferation. However, among the 1263 genes that participate in the cell cycle, only those 28 DEGs appear to be extensively altered in various cancer types, indicating that those genes may be key for cell-cycle regulation in cancer. The 28 selected DEGs also have strong protein–protein interactions, as plotted using STRING (Fig. 1).

3.2. KIF4A and STIL Overexpression Can Serve as Diagnostic and Prognostic Markers in Various Cancers

Among all pan-cancer DEGs, two genes, KIF4A and STIL, showed elevated expression in eleven types of cancer (Fig. 2A,Fig. 2B). In a previous pan-cancer transcriptome study, KIF4A was also found to be upregulated in multiple cancer types, while STIL was not identified, probably due to the different p-values and fold-change thresholds applied in screening [26].

KIF4A encodes the chromokinesin protein KIF4A, an ATP-dependent molecular motor that promotes mitotic chromosome condensation and segregation [27]. KIF4A can also directly bind to chromatin and participate in DNA damage by associating with BRCA2 [28, 29]. Its roles in cancer are complicated. On the one hand, KIF4 knockout in mouse embryonic stem cells leads to the formation of tumors in nude mice [30]. On the other hand, many other studies have proven the upregulated expression of KIF4A in various cancer tissues and its positive correlation with poor prognosis [31–34]. Its overexpression can promote lung cancer resistance to cisplatin, while for breast cancer cells, its overexpression promotes cell apoptosis during treatment with doxorubicin [35, 36]. Our results strongly support the oncogenic role of KIF4A in various cancers. STIL encodes for SCL-interrupting locus protein (STIL), a centrosome protein located in the pericentriolar material. STIL is necessary for centriole duplication, as its absence could lead to centriole loss [37]. Meanwhile, its overexpression could generate extra copies of centrioles and promote chromosome missegregation [38, 39]. The deletion of most STIL exons and the formation of a TIA-1 and STIL recombinant is thought to be a cause of T-cell acute lymphoblastic leukemia; thus, the STIL gene was identified as an oncogene in NCG 6.0 [40, 41]. For solid tumors, upregulated expression of STIL has been found in lung cancers when compared with paring normal tissue, and its high expression was positively related to the mitotic index and metastatic activity [42]. Its mRNA expression amount was also positively correlated with the histological grade in ovarian cancers [43]. Our results prove the overexpression of STIL in multiple solid tumors and indicate that it can serve as a pan-cancer diagnostic marker.

The elevated expression of the KIF4A protein has also been verified in five types of cancers using immunohistochemistry data from the Human Protein Atlas (HPA; Fig. 2C). For all of the five cancer types, KIF4A showed moderate protein expression in normal tissue, while in some tumor samples, its expression was strong.

We next studied the roles of KIF4A and STIL expression in cancer prognosis. High expression of either gene was significantly correlated with poor prognosis in four kinds of cancer, KIRC, KIRP, LUAD, and LIHC, as shown by Kaplan–Meier survival analysis (Fig. 3). In other cancer types, the correlation was insignificant (p-value > 0.05). It is noteworthy that the cancer types in which

KIF4A and STIL showed prognostic values are perfectly overlaid with each other, implying a shared mechanism for cancer progress among these cancers.

3.3. KIF4A and STIL Coexpression and Interaction in Various Cancers

The similar characteristics between KIF4A and STIL encouraged us to explore the relations between them. In all 12 types of cancers, the mRNA expression of KIF4A and STIL was highly correlated, with a Pearson's correlation coefficient (r) of 0.61–0.88 (Fig. 4A). The GeneMANIA interaction network between KIF4A and STIL was plotted (Fig. 4B). Although no physical interaction between KIF4A and STIL was found, they are connected by eight other genes (KIF2C, KIF11, KIF14, KIF23, MELK, AURKB, CCNA2, and PRC1) and form a strong interaction hub. In the hub composed of ten genes, interactions exist among all pairs of genes. Moreover, all of them showed increased expression in multiple types of cancer tissue (Figure S1, Table 4). Five of them, KIF4A, KIF2C, KIF11, KIF14, and KIF23, belong to the kinesin family. KIF4A, KIF23, KIF2C, MELK, and CCNA2 have previously been identified as hub genes that have high prognostic value in kidney renal cell cancer [44]. The protein regulator of cytokinesis 1 (PRC1) was reported to directly interact with KIF4 and form a PRC1–KIF4 complex that tags the plus end of microtubule in a microtubule-length-dependent way. This tagging process is critical for cells to differentiate among microtubules of different lengths and is important for subsequent spindle formation [45, 46]. Aurora kinase B (AURKB) directly binds and activates KIF4A to promote binding between KIF4A and PRC1, subsequently promoting spindle organization [47]. Maternal embryonic leucine zipper kinase (MELK) has previously been identified, together with KIF4A, to be associated with poor prognosis in obese luminal A breast cancer patients [48]. Moreover, it was identified in another study as the top gene associated with elevated pan-cancer expression compared with normal tissue [26]. Pathway-enrichment analysis using GO was conducted on all ten hub genes based on GeneMANIA network analysis (Fig. 4C). As shown, all genes are strongly enriched in mitotic chromosome-segregation-related pathways, including mitotic nuclear division, mitotic spindle organization, and mitotic spindle elongation. The regulation of these processes is crucial for cell fate, as increased centrosome amplification could directly induce tumorigenesis [49].

4.1.The expression and clinical significance of KIF4A an STIL in osteosarcoma tissue and normal paracancerous tissue.

The expression of KIF4A and STIL were detected osteosarcoma tissue and normal paracancerous tissue tissue by immunohistochemistry. We found that expression rate of KIF4A and STIL in osteosarcoma tissues were 96.4% (55/57) and 96.4% (55/57) (Fig. 5A, Fig. 5B, Fig. 5D, Fig. 5E). Compared with 1.75% (1/57) and 3.50%(2/57)in normal paracancerous tissues(Fig. 5C, Fig. 5F), KIF4A and STIL were significantly over-expressed in osteosarcoma tissues (P□0.001). A significant correlation was indicated between KIF4A and STIL expression in osteosarcoma tissues(Table 5).

4.2. KIF4A and STIL were correlated with poor overall survival of osteosarcoma patients

The results of univariate Cox hazard analysis of overall survival of OS patients showed that Enneking stage and expression of KIF4A and STIL were significantly correlated with survival rate(Table 6, Fig. 6A, Fig. 6B). The results of multivariate Cox hazard analysis showed that KIF4A and STIL identified as independent prognostic factors of OS patients(Table 7).

4. Discussion

The pan-cancer expression upregulation of KIF4A and STIL is probably an indicator of the increased mitotic rate in cancer cells. It is also possible that their overexpression activates cell proliferation and cancer progress; however, this requires experimental validation. The strong prognostic power of the elevated expression of KIF4A and STIL in the four types of cancer indicates the important role of mitosis in the progress of these cancers. Dysregulation of mitosis could lead to aneuploidy and chromosome instability, which are both important characteristics of cancers. An increased mitotic rate and abnormal mitotic figures have currently been used for the diagnosis and prognosis of various cancers. It is a particularly essential marker for the grading of neuroendocrine tumors, as neuroendocrine carcinoma (NEC) is differentiated from benign neuroendocrine tumors (NETs) by the mitotic index [50]. In the practice of clinical pathology, the mitotic status of tumor tissue is assessed by Ki-67 immunostaining

and mitotic-figure counting under a high-powered field, which requires significant time and effort. Moreover, variation commonly exists in result interpretation among pathologists. Thus, it is important to identify new biomarkers for mitotic-status indication. Our results suggest that the overexpression of two cell-cycle genes, KIF4A and STIL, which function in the mitosis process, can be a potential novel diagnostic and prognostic marker for various cancers. We were the first to testify KIF4A and STIL by immunohistochemistry ,they were high expressed in osteosarcoma than that of normal paracancerous tissues and correlated with each other.High KIF4A and STIL expression was significantly correlated to poor overall survival and were identified as independent prognostic factors of OS patients.

Supplementary Materials

Figure S1: expression of eight other hub genes in all twelve types of cancers and paring normal tissue; Table S1: function enrichment of the 28 pan-cancer DEGs by STRING.

Declarations

Author Contributions: conceptualization, visualization, project administration and supervision, Zhanxin Mao; methodology, formal analysis, data collection,IHC experiment and writing original draft, Jiankang Pan; software, writing review and editing and validation, Xiaohua Lei; investigation, data curation,Yong Luo.

Funding: This research was funded by Hunan Provincial Health and Family Planning Commission Research Project, Grant No. 20180343. Health Commission of Hunan Province Reasearch Project Grant No.20200284.

Acknowledgments: The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov) is a cancer genomics collaboration that contributed to the collecting and sequencing of normal-tumor sample pairs.

Conflicts of Interest: The authors declare no conflict of interest.

Ethical Approval and Consent to participate: approved by the Ethics Committee of The Second Xiangya Hospital.

Consent for publication: All authors read and approved the final version of this manuscript.

Availability of supporting data: The datasets supporting the conclusions of this article are included within the article and its additional files.

References

1. Siegel RL, Miller KD, Jemal A: **Cancer statistics, 2019**. *Ca-Cancer J Clin* 2019, **69**(1):7-34.
2. Hu T, Kumar Y, Shazia I, Duan SJ, Li Y, Chen L, Chen JF, Yin R, Kwong A, Leung GK *et al*: **Forward and reverse mutations in stages of cancer development**. *Hum Genomics* 2018, **12**(1):40.
3. Yang C, Wang YT, Zheng CH: **A Random Walk Based Cluster Ensemble Approach for Data Integration and Cancer Subtyping**. *Genes (Basel)* 2019, **10**(1).
4. Hu T, Chen S, Ullah A, Xue H: **AluScanCNV2: An R package for copy number variation calling and cancer risk prediction with next-generation sequencing data**. *Genes Dis* 2019, **6**(1):43-46.
5. Thompson JA, Christensen BC, Marsit CJ: **Pan-Cancer Analysis Reveals Differential Susceptibility of Bidirectional Gene Promoters to DNA Methylation, Somatic Mutations, and Copy Number Alterations**. *Int J Mol Sci* 2018, **19**(8).
6. Crona J, Backman S, Welin S, Taieb D, Hellman P, Stalberg P, Skogseid B, Pacak K: **RNA-Sequencing Analysis of Adrenocortical Carcinoma, Pheochromocytoma and Paraganglioma from a Pan-Cancer Perspective**. *Cancers (Basel)* 2018, **10**(12).
7. Kahles A, Lehmann KV, Toussaint NC, Huser M, Stark SG, Sachsenberg T, Stegle O, Kohlbacher O, Sander C, Cancer Genome Atlas Research N *et al*: **Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients**. *Cancer Cell* 2018, **34**(2):211-224 e216.

8. Forde PM, Chaft JE, Smith KN, Anagnostou V, Cottrell TR, Hellmann MD, Zahurak M, Yang SC, Jones DR, Broderick S *et al*: **Neoadjuvant PD-1 Blockade in Resectable Lung Cancer**. *N Engl J Med* 2018, **378**(21):1976-1986.
9. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I *et al*: **TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data**. *Nucleic acids research* 2016, **44**(8):e71.
10. Bullard JH, Purdom E, Hansen KD, Dudoit S: **Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments**. *BMC Bioinformatics* 2010, **11**:94.
11. Zuberi K, Franz M, Rodriguez H, Montojo J, Lopes CT, Bader GD, Morris Q: **GeneMANIA prediction server 2013 update**. *Nucleic Acids Res* 2013, **41**(Web Server issue):W115-122.
12. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters**. *OMICS* 2012, **16**(5):284-287.
13. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, Fulton LL, Dooling DJ, Ding L, Mardis ER *et al*: **Comprehensive molecular portraits of human breast tumours**. *Nature* 2012, **490**(7418):61-70.
14. Cancer Genome Atlas Research N: **Comprehensive molecular characterization of clear cell renal cell carcinoma**. *Nature* 2013, **499**(7456):43-49.
15. Cancer Genome Atlas Research N: **Comprehensive molecular profiling of lung adenocarcinoma**. *Nature* 2014, **511**(7511):543-550.
16. Cancer Genome Atlas Research N: **Comprehensive molecular characterization of gastric adenocarcinoma**. *Nature* 2014, **513**(7517):202-209.
17. Muzny DM, Bainbridge MN, Chang K, Dinh HH, Drummond JA, Fowler G, Kovar CL, Lewis LR, Morgan MB, Newsham IF *et al*: **Comprehensive molecular characterization of human colon and rectal cancer**. *Nature* 2012, **487**(7407):330-337.
18. Linehan WM, Spellman PT, Ricketts CJ, Creighton CJ, Fei SS, Davis C, Wheeler DA, Murray BA, Schmidt L, Vocke CD *et al*: **Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma**. *New Engl J Med* 2016, **374**(2):135-145.
19. Hammerman PS, Lawrence MS, Voet D, Jing R, Cibulskis K, Sivachenko A, Stojanov P, McKenna A, Lander ES, Gabriel S *et al*: **Comprehensive genomic characterization of squamous cell lung cancers**. *Nature* 2012, **489**(7417):519-525.
20. Agrawal N, Akbani R, Aksoy BA, Ally A, Arachchi H, Asa SL, Auman JT, Balasundaram M, Balu S, Baylin SB *et al*: **Integrated Genomic Characterization of Papillary Thyroid Carcinoma**. *Cell* 2014, **159**(3):676-690.
21. Lawrence MS, Sougnez C, Lichtenstein L, Cibulskisl K, Lander E, Gabriel SB, Getz G, Ally A, Balasundaram M, Birol I *et al*: **Comprehensive genomic characterization of head and neck squamous cell carcinomas**. *Nature* 2015, **517**(7536):576-582.
22. Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, Holt RA, Jones SJM, Lee D, Ma Y *et al*: **Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma**. *Cell* 2017, **169**(7):1327-+.
23. Abeshouse A, Ahn J, Akbani R, Ally A, Amin S, Andry CD, Annala M, Aprikian A, Armenia J, Arora A *et al*: **The Molecular Taxonomy of Primary Prostate Cancer**. *Cell* 2015, **163**(4):1011-1025.
24. Cherniack AD, Shen H, Walter V, Stewart C, Murray BA, Bowlby R, Hu X, Ling SY, Soslow RA, Broaddus RR *et al*: **Integrated Molecular Characterization of Uterine Carcinosarcoma**. *Cancer Cell* 2017, **31**(3):411-423.
25. Panagiotou OA, Ioannidis JP, Genome-Wide Significance P: **What should the genome-wide significance threshold be? Empirical replication of borderline genetic associations**. *Int J Epidemiol* 2012, **41**(1):273-286.
26. Hu S, Yuan H, Li Z, Zhang J, Wu J, Chen Y, Shi Q, Ren W, Shao N, Ying X: **Transcriptional response profiles of paired tumor-normal samples offer novel perspectives in pan-cancer analysis**. *Oncotarget* 2017, **8**(25):41334-41347.
27. Mazumdar M, Sundareshan S, Misteli T: **Human chromokinesin KIF4A functions in chromosome condensation and segregation**. *The Journal of cell biology* 2004, **166**(5):613–620.
28. Wu G, Chen PL: **Structural requirements of chromokinesin Kif4A for its proper function in mitosis**. *Biochem Biophys Res Commun* 2008, **372**(3):454-458.
29. Wu G, Zhou L, Khidr L, Guo XE, Kim W, Lee YM, Krasieva T, Chen PL: **A novel role of the chromokinesin Kif4A in DNA damage response**. *Cell Cycle* 2008, **7**(13):2013-2020.

30. Mazumdar M, Lee JH, Sengupta K, Ried T, Rane S, Misteli T: **Tumor formation via loss of a molecular motor protein**. *Current biology : CB* 2006, **16**(15):1559-1564.
31. Hou G, Dong C, Dong Z, Liu G, Xu H, Chen L, Liu L, Wang H, Zhou W: **Upregulate KIF4A Enhances Proliferation, Invasion of Hepatocellular Carcinoma and Indicates poor prognosis Across Human Cancer Types**. *Sci Rep* 2017, **7**(1):4148.
32. Narayan G, Bourdon V, Chaganti S, Arias-Pulido H, Nandula SV, Rao PH, Gissmann L, Dürst M, Schneider A, Pothuri B *et al*: **Gene dosage alterations revealed by cDNA microarray analysis in cervical cancer: identification of candidate amplified and overexpressed genes**. *Genes, chromosomes & cancer* 2007, **46**(4):373–384.
33. Taniwaki M, Takano A, Ishikawa N, Yasui W, Inai K, Nishimura H, Tsuchiya E, Kohno N, Nakamura Y, Daigo Y: **Activation of KIF4A as a prognostic biomarker and therapeutic target for lung cancer**. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2007, **13**(22 Pt 1):6624–6631.
34. Minakawa Y, Kasamatsu A, Koike H, Higo M, Nakashima D, Kouzu Y, Sakamoto Y, Ogawara K, Shiiba M, Tanzawa H *et al*: **Kinesin family member 4A: a potential predictor for progression of human oral cancer**. *PloS one* 2013, **8**(12):e85951.
35. Xiao N, Fu CH, Lin YN, Dong ZX, Pan LN, Chen WW, Zhao J: **Overexpression of chromosome kinesin protein KIF4A enhance cisplatin resistance in A549/DDP cells**. *Int J Clin Exp Patho* 2016, **9**(3):3331-3339.
36. Wang H, Lu C, Li Q, Xie J, Chen T, Tan Y, Wu C, Jiang J: **The role of Kif4A in doxorubicin-induced apoptosis in breast cancer cells**. *Mol Cells* 2014, **37**(11):812-818.
37. Arquint C, Sonnen KF, Stierhof Y-D, Nigg EA: **Cell-cycle-regulated expression of STIL controls centriole number in human cells**. *Journal of cell science* 2012, **125**(Pt 5):1342–1352.
38. Patwardhan D, Mani S, Passemard S, Gressens P, El Ghouzzi V: **STIL balancing primary microcephaly and cancer**. *Cell Death Dis* 2018, **9**(2):65.
39. Arquint C, Sonnen KF, Stierhof YD, Nigg EA: **Cell-cycle-regulated expression of STIL controls centriole number in human cells**. *J Cell Sci* 2012, **125**(Pt 5):1342-1352.
40. D'Angio M, Valsecchi MG, Testi AM, Conter V, Nunes V, Parasole R, Colombini A, Santoro N, Varotto S, Caniglia M *et al*: **Clinical features and outcome of SIL/TAL1-positive T-cell acute lymphoblastic leukemia in children and adolescents: a 10-year experience of the AIEOP group**. *Haematologica* 2015, **100**(1):e10-13.
41. An O, Dall'Olio GM, Mourikis TP, Ciccarelli FD: **NCG 5.0: updates of a manually curated repository of cancer genes and associated properties from cancer mutational screenings**. *Nucleic acids research* 2016, **44**(D1):D992-999.
42. Erez A, Perelman M, Hewitt SM, Cojacaru G, Goldberg I, Shahar I, Yaron P, Muler I, Campaner S, Amariglio N *et al*: **Sil overexpression in lung cancer characterizes tumors with increased mitotic activity**. *Oncogene* 2004, **23**(31):5371–5377.
43. Rabinowicz N, Mangala LS, Brown KR, Checa-Rodriguez C, Castiel A, Moskovich O, Zarfati G, Trakhtenbrot L, Levy-Barda A, Jiang D *et al*: **Targeting the centriolar replication factor STIL synergizes with DNA damaging agents for treatment of ovarian cancer**. *Oncotarget* 2017, **8**(16):27380-27392.
44. Gu Y, Lu L, Wu L, Chen H, Zhu W, He Y: **Identification of prognostic genes in kidney renal clear cell carcinoma by RNAseq data analysis**. *Mol Med Rep* 2017, **15**(4):1661-1667.
45. Subramanian R, Ti S-C, Tan L, Darst SA, Kapoor TM: **Marking and measuring single microtubules by PRC1 and kinesin-4**. *Cell* 2013, **154**(2):377–390.
46. Hu CK, Coughlin M, Field CM, Mitchison TJ: **KIF4 regulates midzone length during cytokinesis**. *Curr Biol* 2011, **21**(10):815-824.
47. Nunes Bastos R, Gandhi SR, Baron RD, Gruneberg U, Nigg EA, Barr FA: **Aurora B suppresses microtubule dynamics and limits central spindle size by locally activating KIF4A**. *J Cell Biol* 2013, **202**(4):605-621.
48. Nuncia-Cantarero M, Martinez-Canales S, Andres-Pretel F, Santpere G, Ocana A, Galan-Moya EM: **Functional transcriptomic annotation and protein-protein interaction network analysis identify NEK2, BIRC5, and TOP2A as potential targets in obese patients with luminal A breast cancer**. *Breast Cancer Res Treat* 2018, **168**(3):613-623.
49. Levine MS, Bakker B, Boeckx B, Moyett J, Lu J, Vitre B, Spierings DC, Lansdorp PM, Cleveland DW, Lambrechts D *et al*: **Centrosome Amplification Is Sufficient to Promote Spontaneous Tumorigenesis in Mammals**. *Dev Cell* 2017, **40**(3):313-322 e315.

50. Klimstra DS, Modlin IR, Coppola D, Lloyd RV, Suster S: **The pathologic classification of neuroendocrine tumors: a review of nomenclature, grading, and staging systems**. *Pancreas* 2010, **39**(6):707-712.

Tables

Group		KIF4A		P	STIL		P
	55	Low	High		Low	High	
Gender							
Male	29	14	15	0.147	9	20	0.123
Femal	26	8	18		13	13	
Age							
20	40	14	26	0.117	17	23	0.382
20	15	8	7		5	10	
Tumor							
Site							
Femur	29	12	17	0.974	12	17	0.708
Tibal	18	7	11		6	12	
Others	8	3	5		4	4	
Enneking staging							
I							
II	7	6	1	0.016	7	0	0.001
III	41	15	26		15	26	
	7	1	6		0	7	
Distant metastasis							
YES	7	1	6	0.141	0	7	0.034
NO	48	21	27		22	26	

Table 1. Clinicopathological variables and the expression of KIF4A and STIL.

Abbr.	Cancer Type	N of Samples
TCGA-BRCA	Breast Invasive Carcinoma	112
TCGA-COAD	Colon Adenocarcinoma	41
TCGA-HNSC	Head and Neck Squamous Cell Carcinoma	43
TCGA-KIRC	Kidney Renal Clear Cell Carcinoma	72
TCGA-KIRP	Kidney Renal Papillary Cell Carcinoma	31
TCGA-LIHC	Liver Hepatocellular Carcinoma	50
TCGA-LUAD	Lung Adenocarcinoma	57
TCGA-LUSC	Lung Squamous Cell Carcinoma	49
TCGA-PRAD	Prostate Adenocarcinoma	52
TCGA-STAD	Stomach Adenocarcinoma	27
TCGA-THCA	Thyroid Carcinoma	58
TCGA-UCEC	Uterine Corpus Endometrial Carcinoma	23
		Total: 615

Table 2. Expression conditions of the top 28 DEGs in the 12 types of cancers investigated.

	BRCA	COAD	HNSC	KIRC	KIRP	LIHC	LUAD	LUSC	PRAD	STAD	THCA	UCEC	N
KIF4A	+	+	+	+	+	+	+	+	+	N.S.	+	+	11
STIL	+	+	+	+	+	+	+	+	+	+	N.S.	+	11
TMEM132A	+	+	+	+	N.S.	N.S.	+	+	+	+	+	+	10
TRIP13	+	+	+	+	+	+	+	+	+	+	N.S.	N.S.	10
GTSE1	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
UBE2T	+	+	+	+	+	+	+	+	N.S.	N.S.	+	+	10
AURKA	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
TPX2	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
BIRC5	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
ORC6	+	+	+	+	+	+	+	+	N.S.	N.S.	+	+	10
CLSPN	+	N.S.	+	+	+	+	+	+	N.S.	+	+	+	10
CDC45	+	+	+	+	+	+	+	+	N.S.	N.S.	+	+	10
CDC6	+	+	+	+	+	+	+	+	N.S.	N.S.	+	+	10
MMP11	+	+	+	+	+	+	+	+	N.S.	+	+	N.S.	10
CDKN3	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
MYBL2	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
E2F1	+	+	+	+	+	+	N.S.	+	N.S.	+	+	+	10
RNASEH2A	+	+	+	+	+	+	+	+	N.S.	N.S.	+	+	10
ASF1B	+	+	N.S.	+	+	+	+	+	+	N.S.	+	+	10
EZH2	+	+	N.S.	+	+	+	+	+	+	N.S.	+	+	10
FOXM1	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
CDCA3	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
KIF20A	+	+	N.S.	+	+	+	+	+	+	N.S.	+	+	10
CENPA	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
KIF14	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
NCAPH	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
HJURP	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
PKMYT1	+	+	+	+	+	+	+	+	N.S.	N.S.	+	+	10

“+”: specific gene is overexpressed in that cancer type compared to normal tissue. “N.S.”: not significant.

Table 3. Expression conditions of the 10 hub genes in all 12 types of cancer.

	BRCA	COAD	HNSC	KIRC	KIRP	LIHC	LUAD	LUSC	PRAD	STAD	THCA	UCEC	N
KIF4A	+	+	+	+	+	+	+	+	+	N.S.	+	+	11
STIL	+	+	+	+	+	+	+	+	+	+	N.S.	+	11
KIF14	+	+	+	+	+	+	+	+	+	N.S.	N.S.	+	10
KIF2C	+	+	+	+	N.S.	+	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	5
KIF23	+	+	+	N.S.	N.S.	+	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	4
KIF11	+	N.S.	N.S.	+	+	+	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	4
CCNA2	+	+	N.S.	N.S.	N.S.	+	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	3
MELK	+	N.S.	N.S.	+	N.S.	N.S.	N.S.	+	N.S.	N.S.	N.S.	N.S.	3
PRC1	+	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	1
AURKB	+	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	1

Table 4. List of full names and sample numbers for each type of cancer.

			STIL		P=0.017
			L	H	
KIF4A	L	22	13	9	
	H	33	9	24	
		55	22	33	

Table 5. The correlation of KIF4A and STIL expression (Spearman-rho test).

Characteristics		N	Overall survival		P
			HR	95% CI	
Sex	Female	26	1		0.342
	Male	29	1.307	0.753-2.270	
Age	≥ 20 years	15	1		0.497
	< 20 years	40	0.813	0.447-1.478	
tumor site	Femoral	29	1		0.498
	Tibia	18	1.412	0.777-2.566	
	Others	8	1.008	0.458-2.219	
Stage	I	7	1		<0.001
	II	41	3.726	1.512-9.181	
	III	7	83.066	18.639-370.178	
STIL	Low	22	1		<0.001
	High	33	5.733	2.897-11.348	
KIF4A	Low	22	1		0.001
	High	33	2.685	1.516-4.754	

Table 6. Clinicopathological factors associated with overall survival based on univariate Cox regression analysis.

		Overall survival		P
		HR	95% CI	
Stage	I	1		
	II	2.399	0.861-6.685	0.094
	III	53.608	9.901-290.265	<0.001
KIF4A	Low	1		
	High	1.899	1.004-3.590	0.049
STIL	Low	1		
	High	2.837	1.277-6.303	0.011

Table 7. Clinicopathological factors associated with overall survival based on multivariate Cox proportional regression analysis.

Figures

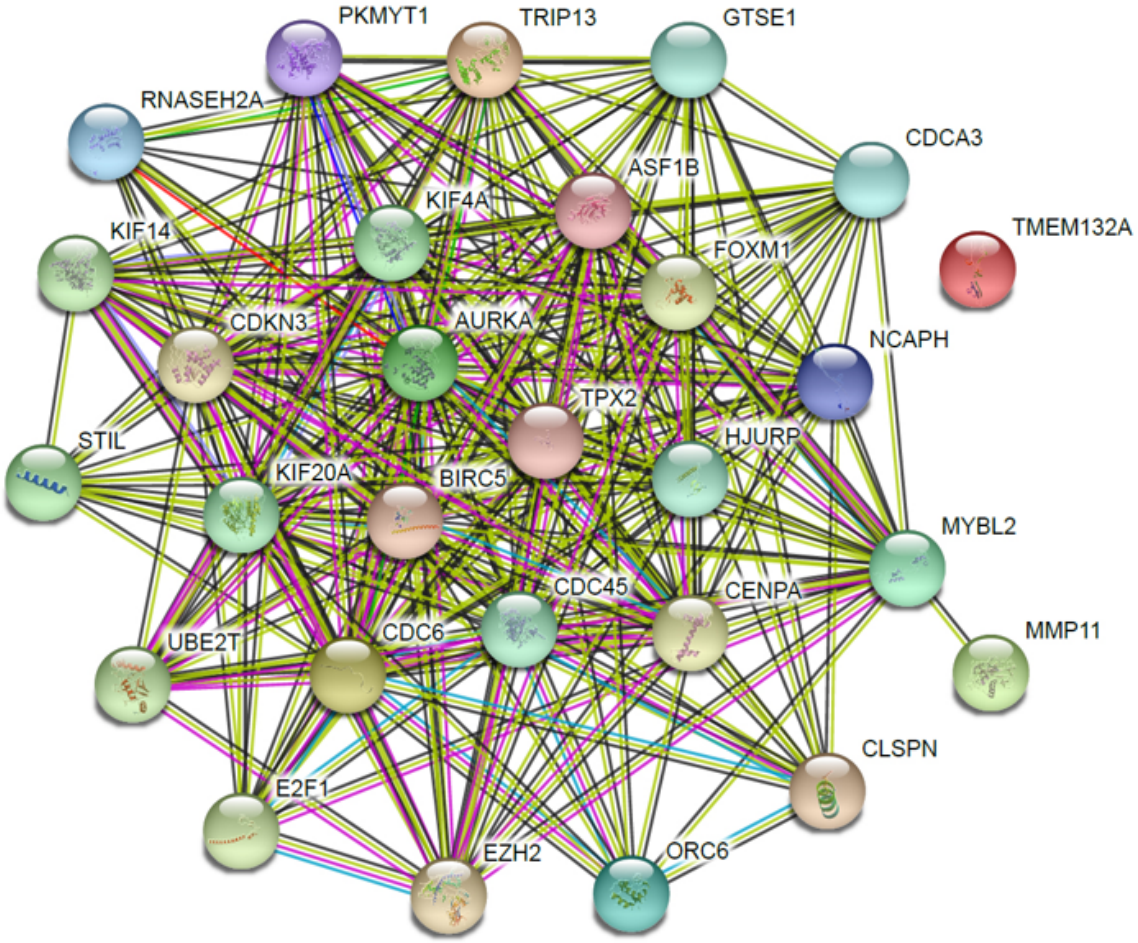

Figure 1

Figure 1

Protein–protein interactions (PPI) involving the 28 differentially expressed genes (DEGs) were identified using the STRING database. The “experiment”, “database”, and “coexpression” evidence channels were chosen for network construction. Clustering was performed using the MCL algorithm with inflation parameter 10. Different colors indicate different clusters and the line thickness indicates the strength of evidence.

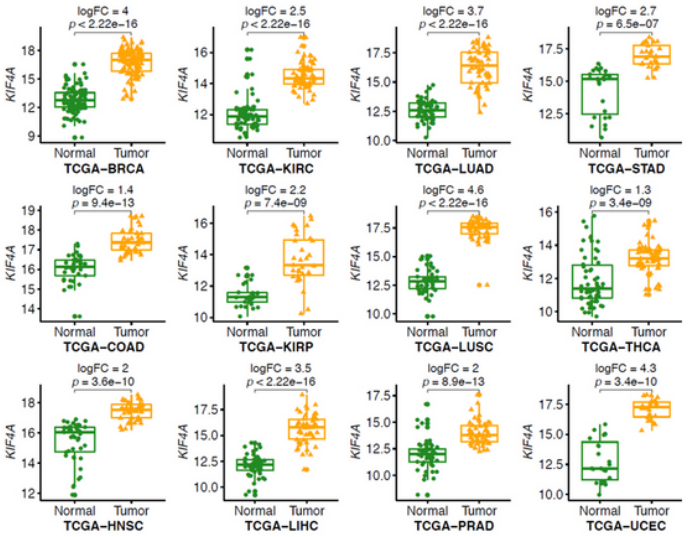

Figure 2A

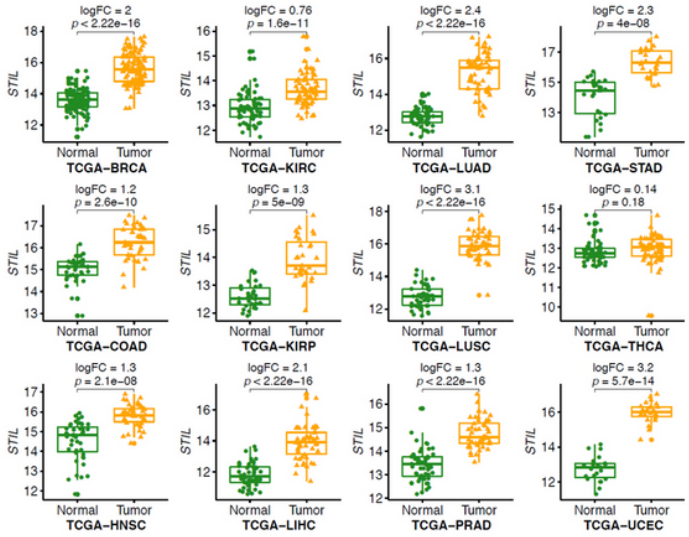

Figure 2B

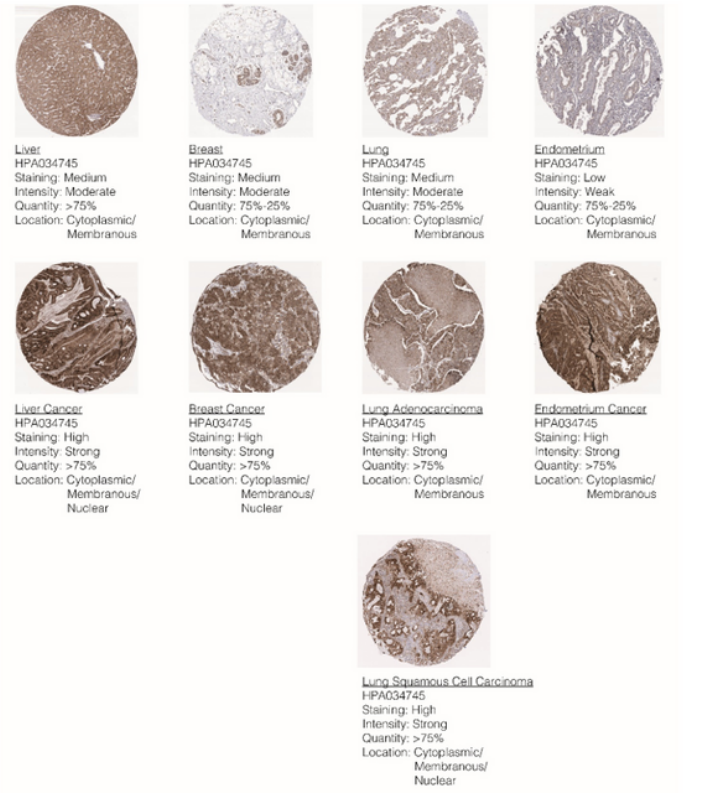

Figure 2C

Figure 2

A. Expression profile of KIF4A in cancer tissue and paring normal tissue. B. Expression profile of STIL in cancer tissue and paring normal tissue. C. Validation of KIF4A protein expression in various cancers and normal tissue using the Human Protein Atlas (HPA) database.

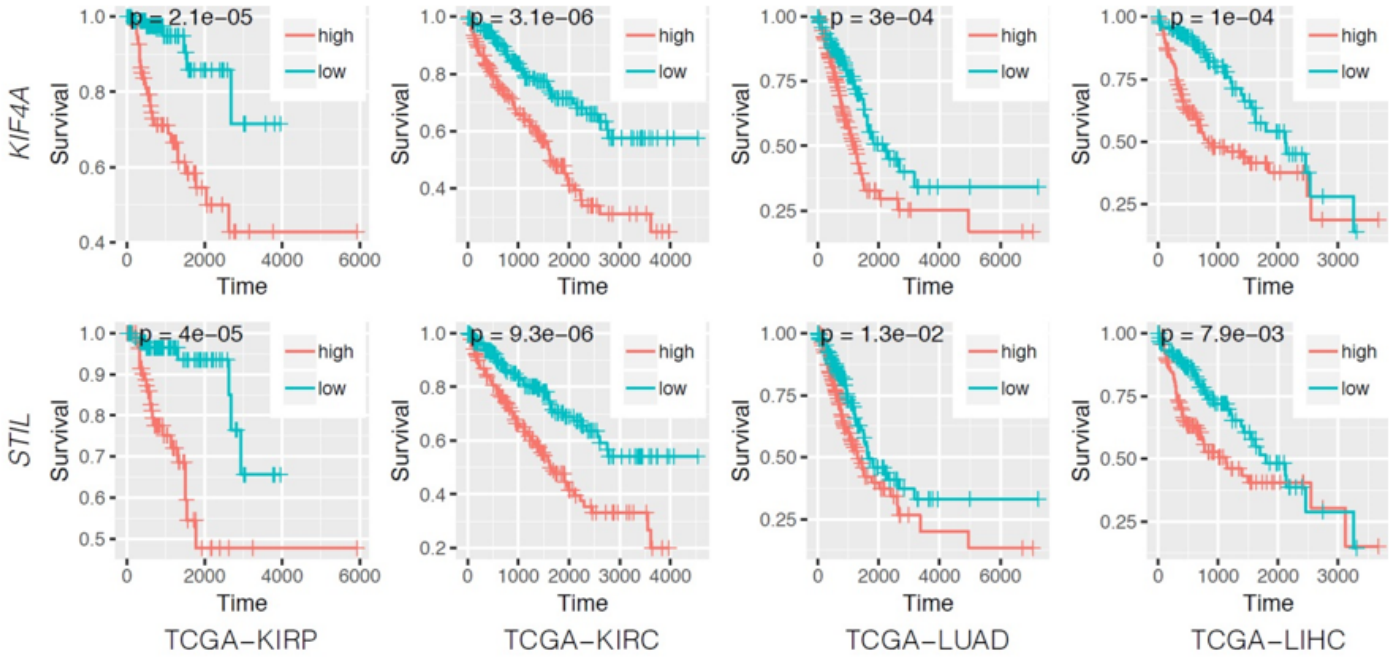

Figure 3

Figure 3

Kaplan–Meier survival curves for four cancer types in regard to KIF4A or STIL expression. Expression values of KIF4A or STIL were classified into high- or low-expression groups according to mean expression level; each group is comprised of the highest and lowest thirds of patients, respectively.

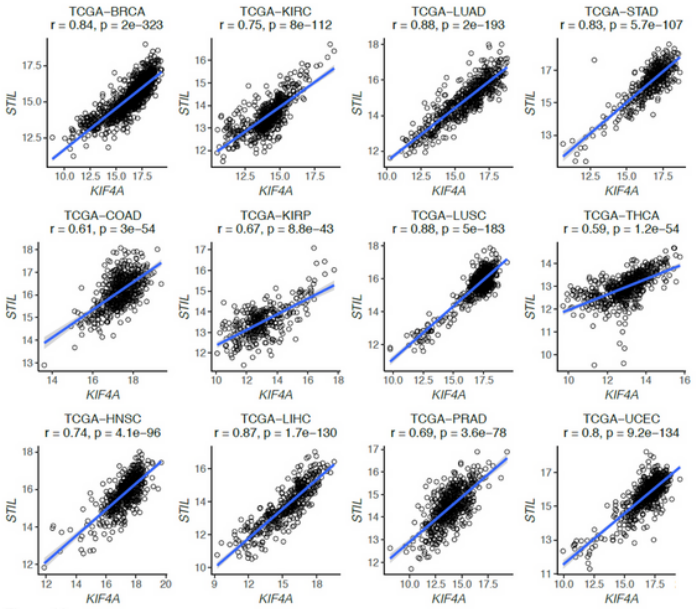

Figure 4A

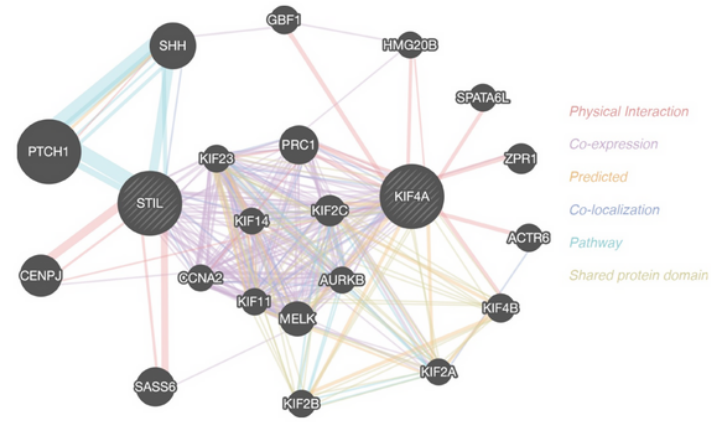

Figure 4B

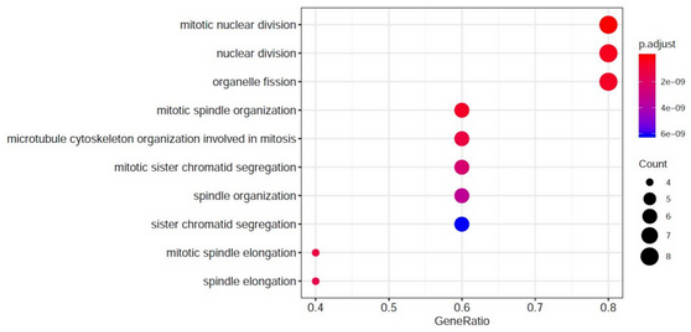

Figure 4C

Figure 4

A. KIF4A and STIL co-expression in 12 types of cancer. The Pearson correlation method was used. B. KIF4A and STIL gene-interaction network. GeneMANIA was used to search for interacting genes of KIF4A and STIL based on physical interactions, coexpression, colocalization, pathways, shared protein domains, and predicted interactions. C. Pathway-enrichment analysis of hub genes using GO. Dotspot was performed by clusterProfiler.

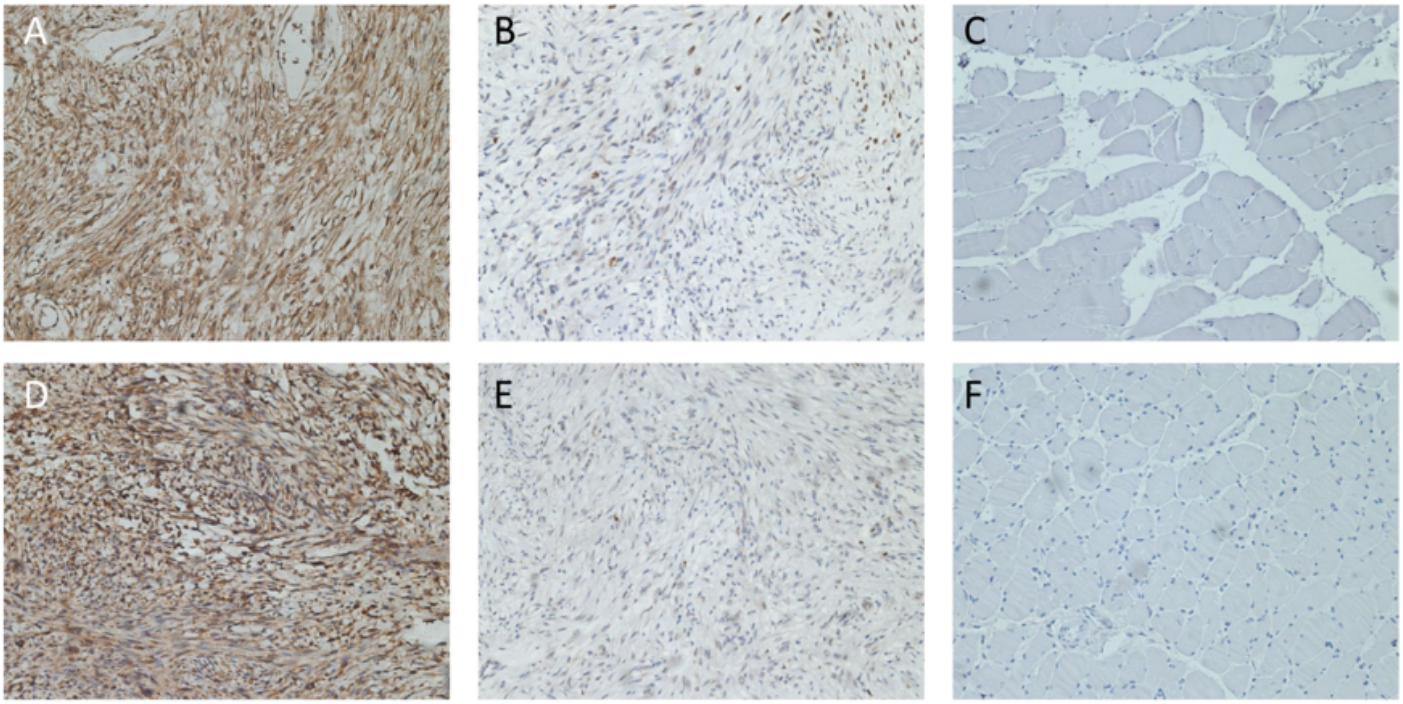

Figure 5

Representative immunohistochemical staining of KIF4A and STIL.A High expression of KIF4A in OS (x200).B Low expression of KIF4A in OS (x200).C Negative expression of KIF4A in normal paracancerous tissues(x200).D High expression of STIL in OS(x200).E Low expression of STIL in OS(x200).F Negative expression of STIL in normal paracancerous tissues(x200).

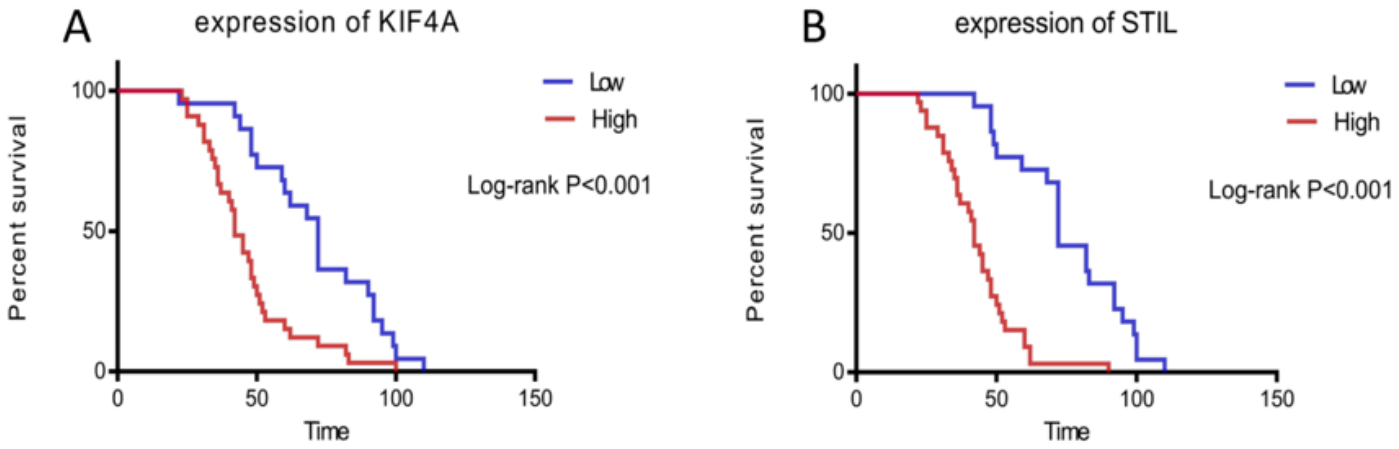

Figure 6

Kaplan–Meier survival curves of patients with OS.A The relationship between KIF4A expression and the five-year survival rate.B The relationship between STIL expression and the five-year survival rate.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx

- FigS1.pdf