

Identification of differentially expressed protein-coding genes in lung adenocarcinomas

LUYAO WANG^{1*}, SHICHENG LI^{2*}, YUANYONG WANG²,
ZHENXUE TANG¹, CHAOLONG LIU¹, WENJIE JIAO² and JIA LIU¹

¹Department of Pharmacology, School of Pharmacy, Qingdao University;

²Department of Thoracic Surgery, Affiliated Hospital of Qingdao University, Qingdao, Shandong 266000, P.R. China

Received August 31, 2018; Accepted November 1, 2019

DOI: 10.3892/etm.2019.8300

Abstract. Lung adenocarcinoma accounts for a high proportion of lung cancers. Though efforts have been made to develop new and effective treatments for this disease, the mortality rate remains high. Gene expression microarrays facilitate the study of lung cancer at the molecular level. The present study aimed to detect differentially expressed protein-coding genes to identify novel biomarkers and therapeutic targets for lung adenocarcinoma. Aberrations in gene expression in lung adenocarcinoma were determined by analysis of mRNA microarray datasets from the Gene Expression Omnibus database. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, protein-protein interaction (PPI) networks and statistical analysis were used to identify the biological functions of the differentially expressed genes (DEGs). The results of the bioinformatics analysis were subsequently validated using reverse transcription-quantitative PCR. A total of 303 DEGs were identified in lung adenocarcinomas, and they were enriched in a number of cancer-associated GO terms and KEGG pathways. DNA topoisomerase 2 α (TOP2A), cell division cycle protein homolog 20 (CDC20), mitotic checkpoint serine/threonine protein kinase BUB1 (BUB1) and mitotic spindle assembly checkpoint protein MAD2A (MAD2L1) exhibited the highest degree of interaction in the PPI network. Survival analysis performed using Kaplan-Meier curves

and Cox regression indicated that these four genes were all significantly associated with the survival of patients with lung adenocarcinomas. In conclusion, TOP2A, CDC20, BUB1 and MAD2L1 may be key protein-coding genes that may serve as biomarkers and therapeutic targets in lung adenocarcinomas.

Introduction

Lung cancer presents a major health threat to men and women worldwide (1). The mortality rate of lung cancer is the highest among all human malignancies. Patients generally present at an advanced stage at initial diagnosis and their estimated five-year survival rate is 10-20% (2,3). Lung adenocarcinoma, which is a typical form of non-small cell lung cancer (NSCLC), accounts for a high proportion of lung cancer cases, and the majority of primary pulmonary adenocarcinomas are highly heterogeneous (4). Although new management strategies for NSCLC, including chemotherapy and targeted therapy, have been developed (5), the prognosis of affected patients remains poor (6).

Gene expression microarrays facilitate the identification of differentially expressed genes (DEGs) between normal and tumor tissues. These genes may be promising biomarkers and/or therapeutic targets for the diagnosis and treatment of cancer (7). However, critical molecular targets for lung adenocarcinomas have yet to be identified.

With the aim of improving the diagnosis and treatment of lung adenocarcinomas, in the present study, data mining of the public Gene Expression Omnibus (GEO) database and bioinformatics analysis was performed to identify differentially expressed protein-coding genes between lung cancer and normal tissue samples. The DEGs were assessed using Gene Ontology (GO) functional enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis to identify their biological functions. A protein-protein interaction (PPI) network of the DEGs was also constructed to identify hub genes. The results indicated that DNA topoisomerase 2 α (TOP2A), cell division cycle protein 20 homolog (CDC20), mitotic checkpoint serine/threonine kinase BUB1 (BUB1) and mitotic spindle assembly checkpoint MAD2A (MAD2L1) may be potential biomarkers and therapeutic targets in lung adenocarcinomas. In addition, reverse transcription-quantitative PCR (RT-qPCR) analysis of

Correspondence to: Professor Wenjie Jiao, Department of Thoracic Surgery, Affiliated Hospital of Qingdao University, 16 Jiangsu Road, Qingdao, Shandong 266000, P.R. China
E-mail: xwkjiao@126.com

Professor Jia Liu, Department of Pharmacology, School of Pharmacy, Qingdao University, Qingdao, Shandong 266000, P.R. China
E-mail: dadaliujia@qdu.edu.cn

*Contributed equally

Key words: lung adenocarcinoma, bioinformatics analysis, biomarkers, DNA topoisomerase 2 α , cell division cycle protein homolog 20, mitotic checkpoint serine/threonine protein kinase BUB1, mitotic spindle assembly checkpoint protein MAD2L1

the key genes identified in 2 patient samples was used to verify the results of the bioinformatics analysis.

Materials and methods

Bioinformatics prediction. The microarray dataset GSE27262, including 25 normal tissue samples and 25 lung adenocarcinoma tissue samples, was obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>) (8-11). The GPL570 platform (Affymetrix; Thermo Fisher Scientific, Inc.) was used to analyze the microarray data. The platform was an Affymetrix Human Genome U133 Plus 2.0 Array. The annotation file was also acquired.

Statistical computing and graphical representation were performed using R software version 3.3.3 (The R Project for Statistical Computing; The R Foundation). The R linear models for microarray data (limma) package was used to accomplish all of the data processing. After background subtraction and normalization using robust multi-chip averaging, the samples from the GEO dataset were divided into two groups: A disease group (25 tumor tissues) and a control group (25 tumor-adjacent tissues). The limma algorithm was used for differential analysis in disease. The criteria to classify a gene as a DEG were a log fold change >2 and a significance of P<0.05 between tumor adjacent tissue samples and tumor samples.

Co-expression network. The PPI network was established using the Search Tool for the Retrieval of Interacting Genes/proteins (STRING; version 11.0; <https://string-db.org>) and Cytoscape software (version 3.6.1; www.cytoscape.org; The Cytoscape Consortium). Hub genes were identified from the PPIs between differentially expressed protein-coding genes and the edge lengths between nodes from these hub genes were subsequently obtained.

Functional enrichment and pathway analysis. DEG enriched GO terms and KEGG pathways were determined to identify their biological function based on the Database for Annotation, Visualization and Integrated Discovery (DAVID; version 6.8; <https://david.ncifcrf.gov>).

Survival analysis. Kaplan-Meier curve and Cox regression analyses were performed to identify the prognostic value of the key DEGs. The patients from the dataset were divided into pairs of subgroups, namely the ‘high expression group’ and the ‘low expression group’, prior to performing the survival analysis. The Oncomine database (version 4.5; <https://www.oncomine.org>) and Kaplan-Meier Plotter database (www.kmplot.com) were used to perform the survival analyses (12,13). Student's t-test was used to examine the statistical significance between 2 groups. P<0.05 was considered a statistically significant difference. The Oncomine and Human Protein Atlas (HPA) databases (version 19; <http://www.proteinatlas.org>) are free open-access databases (14), which were used to validate the expression of TOP2A and BUB1 at the protein level in lung adenocarcinoma tissues.

Verification of bioinformatics in patient samples. Paired lung cancer tissues and adjacent normal tissues were collected in

Table I. Primer sequences of β-actin, TOP2A, BUB1, CDC20 and MAD2L1 used for reverse transcription-quantitative PCR.

Gene	Primer sequence (5'→3')
β-actin	F: AAGAGAGGCATCCTGACCCCT R: TACATGGCTGGGGTGTGTGAA
TOP2A	F: ACCATTGCAGCCTGTAAATGA R: GGGCGGAGCAAAATATGTTCC
BUB1	F: AGCCCAGACAGTAACAGACTC R: GTTGGCAACCTTATGTGTTTCAC
CDC20	F: GCTTTGAACCTGAACGGTTTTG R: TCTGGCGCATTTTGTGGTTTT
MAD2L1	F: GGACTCACCTTGCTTGTAACCTAC R: GATCACTGAACGGATTTTCATCCT

BUB1, mitotic checkpoint serine/threonine protein kinase BUB1; CDC20, cell division cycle protein 20 homolog; F, forward; MAD2L1, mitotic spindle assembly checkpoint protein MAD2A; TOP2A, topoisomerase 2α; R, reverse.

January 2018 from two patients with lung cancer, 1 male (age, 64; BMI, 18.9; no smoking history) and 1 female (age, 74; BMI, 19.9; no smoking history), that were treated at the Affiliated Hospital of Qingdao University (Qingdao, China). The patients included in the study were lung adenocarcinoma patients who had received surgical treatment for lung adenocarcinoma. To confirm the results of the present bioinformatics analysis the expression levels of TOP2A, CDC20, BUB1 and MAD2L1 in tumor tissues of two lung adenocarcinoma patients were quantified using RT-qPCR. Specimens were frozen in liquid nitrogen immediately after resection and stored at -80°C. Total RNA was extracted using NcmZol reagent (NCM Biotech), according to the manufacturer's protocol. A FastQuant RT kit (with gDNase; Tiangen Biotech Co., Ltd.) was used to reverse transcribe total RNA to cDNA; the temperature of the RT process was 42°C. qPCR was performed in an ABI StepOnePlus Thermocycler (Applied Biosystems; Thermo Fisher Scientific, Inc.) using a SYBR Green PCR kit (Qiagen GmbH). The primer sequences are listed in Table I. The following thermocycling conditions were used for the qPCR: Initial denaturation temperature is 95°C for 15 min; 40 cycles of 95°C for 10 sec, 55°C for 30 sec and 72°C for 30 sec. All of the analyses were performed in triplicate and the relative expression levels were calculated based on the comparative 2^{-ΔΔCq} method (15).

Statistical analysis. Statistical analysis for comparison between the tumor group and the control group was performed using unpaired Student' t-test. The fold change >2 and P<0.05 were considered to indicate a statistically significant difference. Student's t-test was also used to examine the statistical significance of survival analysis and the identification of hub genes. P<0.05 was considered to indicate a statistically significant difference. The expression of TOP2A and BUB1 in different stages of lung cancer was analyzed using an F-test. F test was performed as part of an ANOVA. Student-Newman-Keuls was used as a post hoc test after ANOVA.

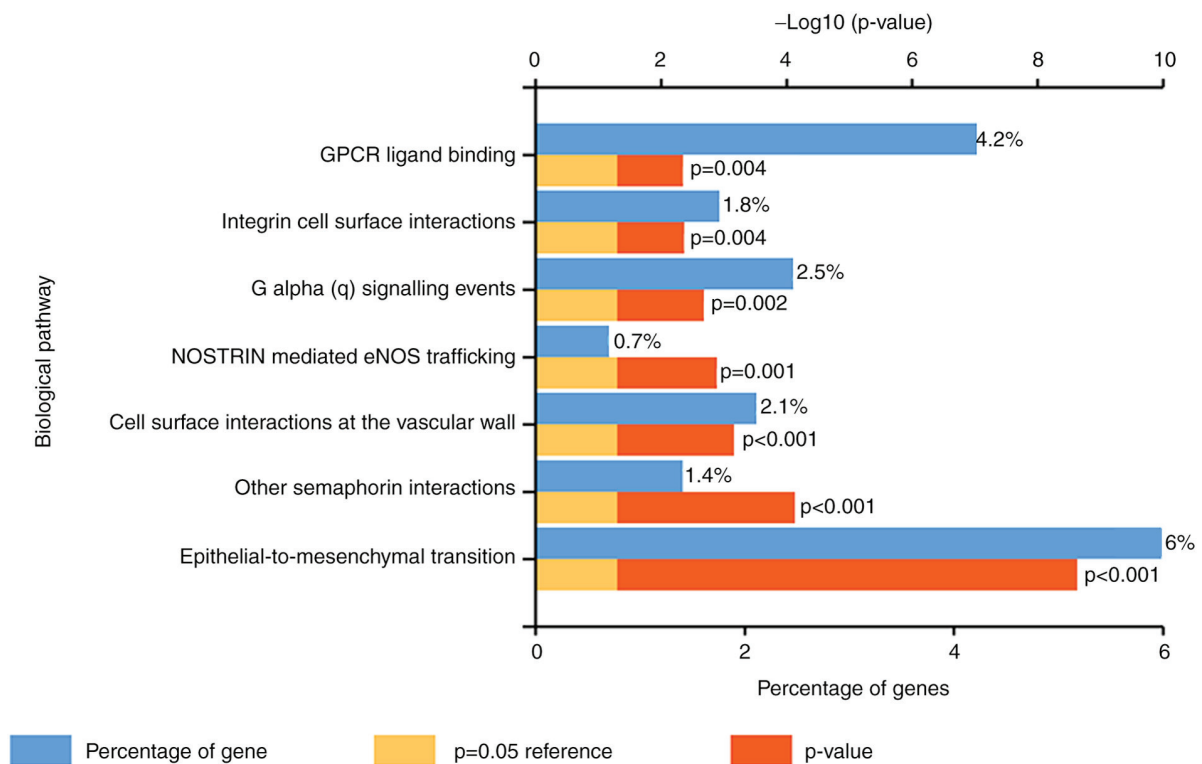


Figure 1. Top seven Kyoto Encyclopedia of Genes and Genomes pathways enriched by the differentially expressed protein-coding genes in lung adenocarcinoma. eNOS, nitric oxide synthase 3; NOSTRIN, nitric oxide synthase trafficking; GPCR, G protein-coupled receptor.

Results

Identification of DEGs and bioinformatics analysis. In the present database study, a total of 303 DEGs were identified in lung adenocarcinomas, including 247 downregulated (Table SI) and 56 upregulated DEGs (Table SII). The results of KEGG analysis indicated that these DEGs were enriched in numerous cancer-associated biological pathways, including epithelial-to-mesenchymal transition and G protein-coupled receptor (GPCR) ligand binding (Fig. 1). The GO analyses indicated that the DEGs were enriched in numerous cancer-associated terms in the category biological process, including signal transduction, cell communication and cell growth and/or maintenance (Fig. 2A). GO enrichment analysis in the category cellular component indicated that the DEGs are concentrated in the plasma membrane, and are integral to the plasma membrane and extracellular matrix (Fig. 2B). In addition, the DEGs were enriched in various cancer-associated GO terms in the category molecular function, including cell adhesion molecule activity and receptor activity (Fig. 2C).

TOP2A, CDC20, BUB1 and MAD2L1 are significant hub genes in lung adenocarcinoma. To identify the critical hub genes in lung adenocarcinoma, a PPI network was generated for all 303 DEGs (Fig. 3). Among the 303 DEGs, TOP2A, CDC20, BUB1 and MAD2L1 were identified to have the highest degree of interaction and were considered to be hub genes. In the present study, Kaplan-Meier curve and Cox regression analyses were performed to determine the prognostic value of the four hub genes identified from the PPI. The results indicated that TOP2A, CDC20, BUB1 and MAD2L1

were all significantly associated with the survival of patients with lung adenocarcinomas (hazard ratios: 1.55, 1.82, 1.83 and 2.23, respectively) and were also highly expressed in lung adenocarcinoma tissue (Fig. 4). Therefore, these four DEGs may be potential biomarkers to predict the survival of lung adenocarcinoma patients. In addition, TOP2A and BUB1 were identified to be highly expressed in lung adenocarcinoma tissues in patient samples from the Affiliated Hospital of Qingdao University (Fig. 5). The expression of TOP2A and BUB1 in different stages of lung cancer is presented in Fig. 6. The Oncomine and HPA databases were used to analyze the validation of the expression of TOP2A and BUB1 at the protein level in lung adenocarcinoma tissues. The analysis results are presented in Fig. 7.

Subsequent verification based on RT-qPCR. RT-qPCR was performed to experimentally confirm the expression of TOP2A, CDC20, BUB1 and MAD2L1 in tumor tissues of patients with lung adenocarcinoma, in order to verify the results of the present bioinformatic analysis. As presented in Fig. 8, the mRNA expression of TOP2A, BUB1, CDC20 and MAD2L1 in lung adenocarcinoma samples was significantly increased compared with that in non-tumor tissues (all, $P<0.05$).

Discussion

Despite advances in chemotherapy and targeted therapies for NSCLC, lung adenocarcinoma remains a threat to human life (16). The purpose of the present study was to identify reliable biomarkers in lung adenocarcinomas to improve the diagnosis and predicted survival of patients and to provide

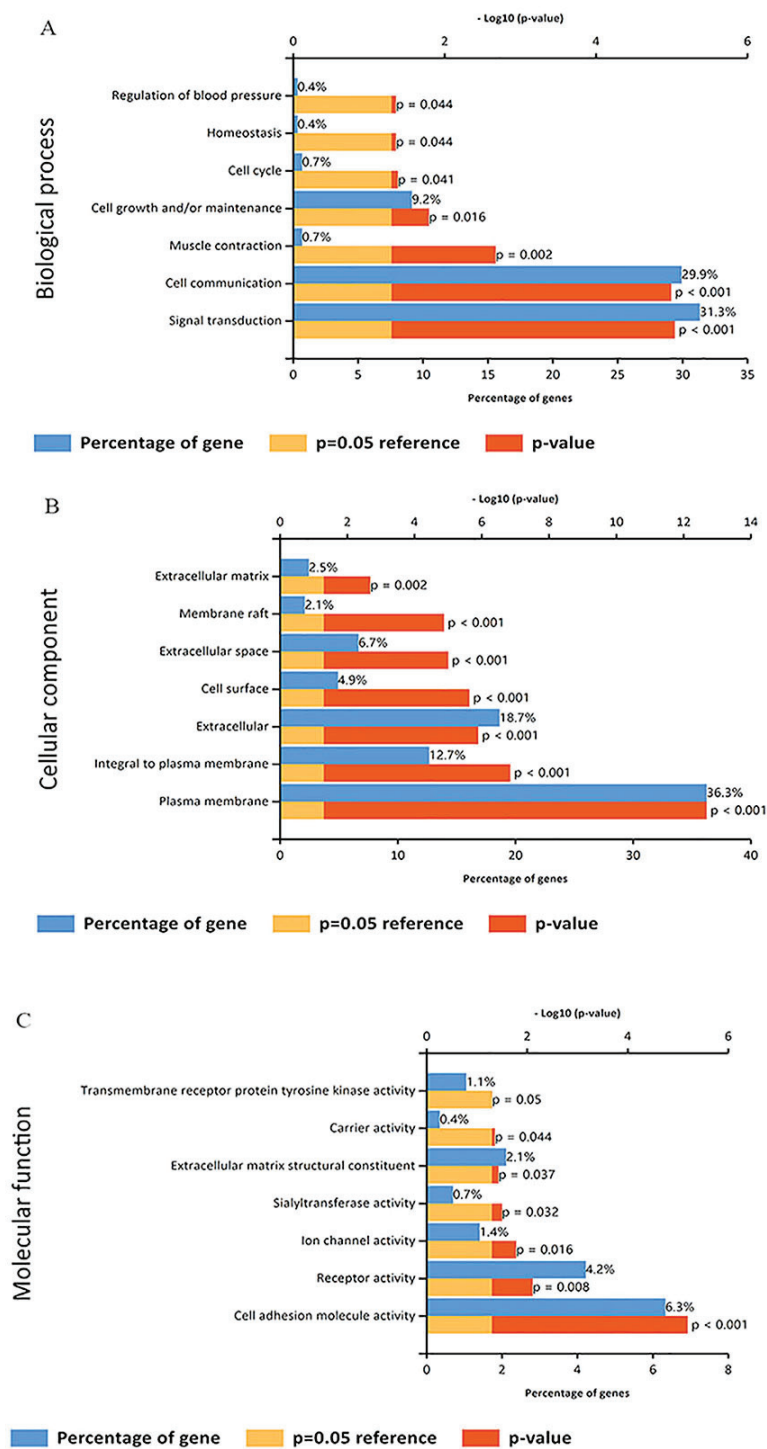
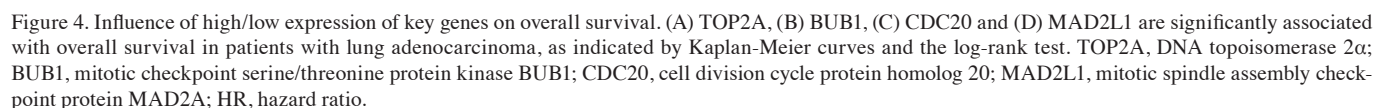
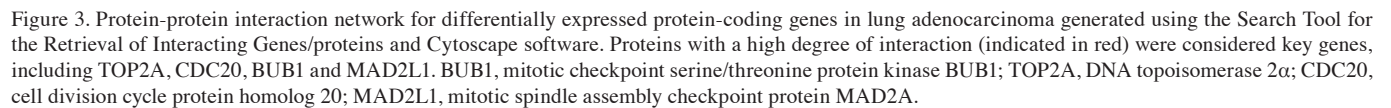


Figure 2. Gene Ontology terms enriched by the differentially expressed protein-coding genes in lung adenocarcinoma in the categories (A) biological process, (B) cellular component and (C) molecular function.

novel therapeutic targets. The gene expression profiles between lung cancer and normal tissues (25 pairs) from a GEO dataset were compared and 303 DEGs in lung NSCLC tissues were identified; however, due to the number of samples in the microarray dataset analyzed being limited to 25 pairs, future studies should use a larger number of samples to verify the findings of the present study. GO and KEGG pathway analyses may be used to identify the biological functions of DEGs (17).

The KEGG analysis performed in the present study indicated that the DEGs were enriched in a number of cancer-associated pathways, including epithelial-to-mesenchymal transition and GPCR ligand binding. Furthermore, the GO enrichment analysis in the category cellular component indicated that the DEGs were concentrated in or/and are integral to the plasma membrane or extracellular matrix. In the category biological process, the DEGs were enriched in GO terms including signal



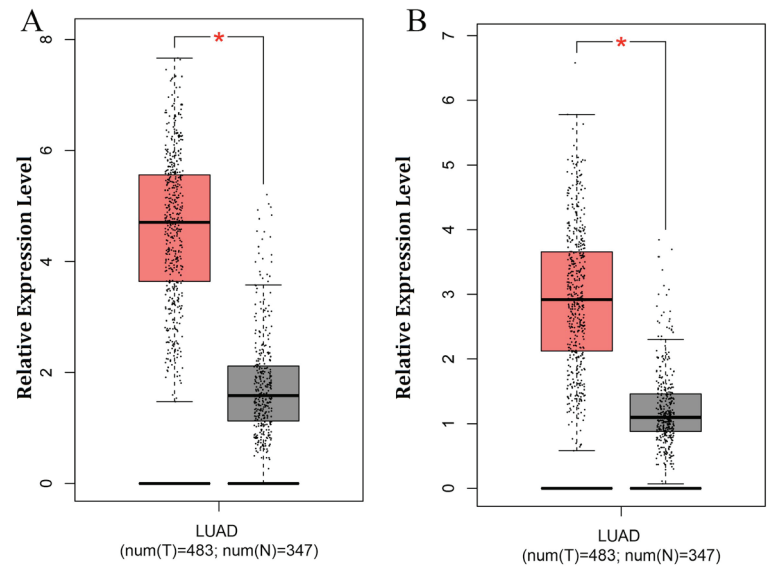


Figure 5. (A) DNA topoisomerase 2α and (B) mitotic checkpoint serine/threonine protein kinase BUB1 were identified to be overexpressed in LUAD tissues. Red corresponds to the tumor group and gray corresponds to the normal group. *P<0.05. LUAD, lung adenocarcinoma; T, tumor tissues; N, normal lung tissues.

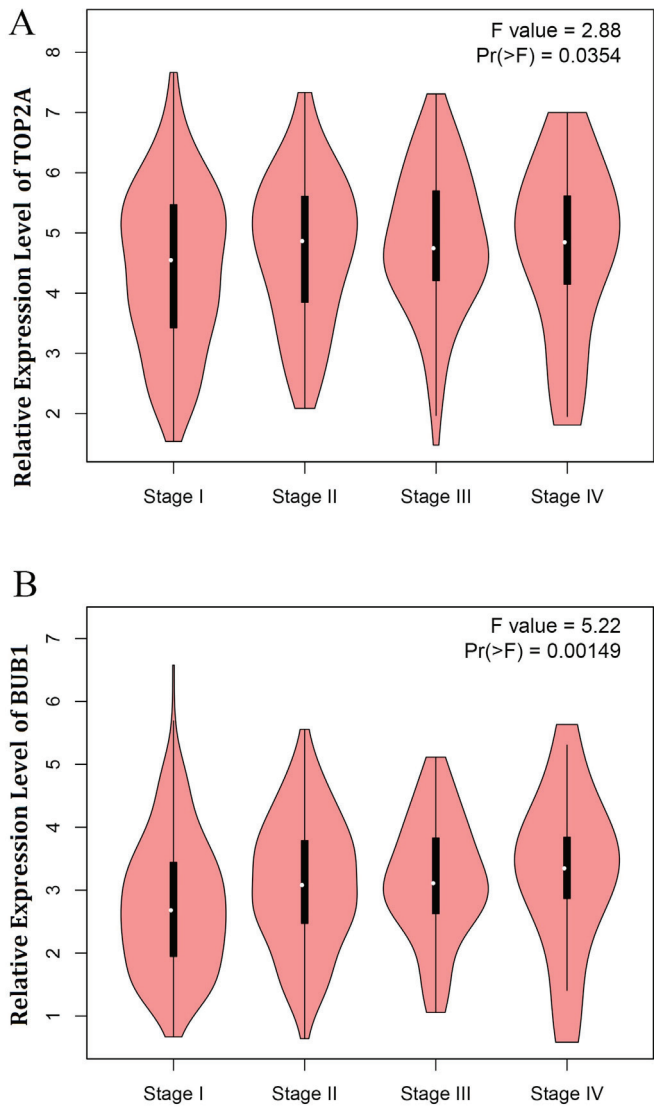


Figure 6. Expression of (A) TOP2A and (B) mitotic checkpoint serine/threonine protein kinase BUB1 in different stages of lung cancer. TOP2A, DNA topoisomerase 2α.

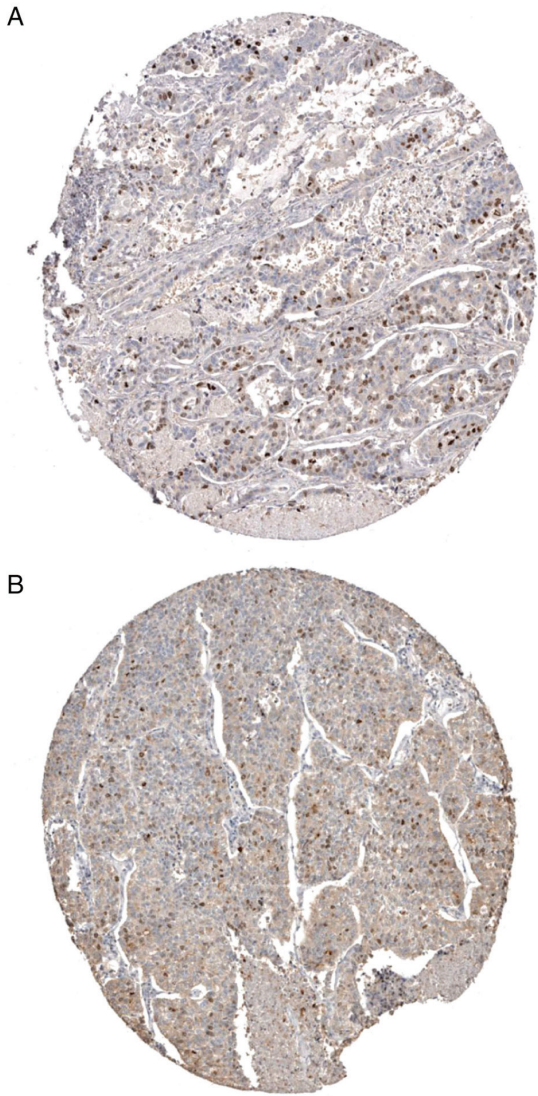


Figure 7. Validation of the protein expression levels of (A) DNA topoisomerase 2α and (B) mitotic checkpoint serine/threonine protein kinase BUB1 in lung adenocarcinoma tissues using the Oncomine and HPA databases.

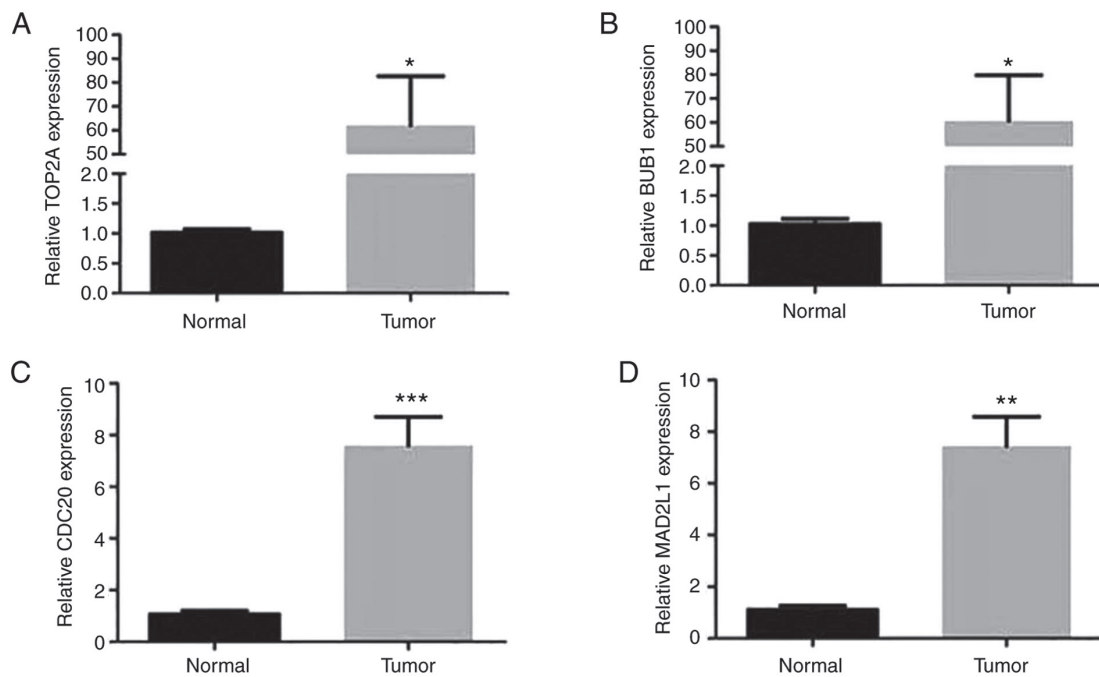


Figure 8. (A) TOP2A, (B) BUB1, (C) CDC20 and (D) MAD2L1 were confirmed to be significantly upregulated in lung cancer tissues compared to normal tissues based on reverse transcription-quantitative PCR analysis. * $P<0.05$, ** $P<0.01$, *** $P<0.001$ vs. normal group. TOP2A, DNA topoisomerase 2 α ; BUB1, mitotic checkpoint serine/threonine protein kinase BUB1; CDC20, cell division cycle protein homolog 20; MAD2L1, mitotic spindle assembly checkpoint protein MAD2A.

transduction, cell communication, and cell growth and/or maintenance, and in the category molecular function, they were enriched in terms including adhesion molecule activity and receptor activity. A PPI network was generated to determine the critical hub genes and provide valuable information for the analysis of cellular functions and biological processes. TOP2A, CDC20, BUB1 and MAD2L1 displayed a high degree of interaction in the PPI network and were considered as hub genes. Survival analysis also indicated that these four key genes were highly expressed in lung cancer and that their expression was significantly associated with overall survival. Information, including the TNM stage, performance status scores and the degree of tumor histological differentiation for the patients was not included in the dataset used to perform survival analysis, which is a limitation of the present study. Taken together, the results of the KEGG, GO and survival analyses demonstrated that TOP2A, CDC20, BUB1 and MAD2L1 are significantly associated with lung adenocarcinomas and may potentially serve as effective biomarkers and therapeutic targets in this malignancy.

TOP2A is a highly evolutionarily conserved enzyme with serves an important role in centromere biology (18). Previous studies have indicated that TOP2A promotes tumor development and progression and may be used as an early prognostic marker in numerous types of cancer, particularly in breast cancer and ovarian cancer (19,20). TOP2A protein has been considered as a marker that is associated with increased tumor grade, since it is highly expressed in proliferating cells (21). Topoisomerase-targeting anti-cancer drugs have been successfully used in the clinic, and their efficacy is based on topoisomerase poisoning (22). Recently, it has been reported that abnormal alterations and modifications of TOP2A may be critical to chromosome instability in some cancer types (23).

Another recent study indicated that miR-139 targets TOP2A and that defective miR-139 expression induces increased expression of TOP2A, and the increased TOP2A then promotes the transcription of EMT-associated genes through activating the β -catenin signaling pathway in cancer (24).

BUB1 is a serine/threonine kinase protein that is associated with kinase function and the spindle assembly checkpoint (25,26). During mitosis and cell proliferation, BUB1 is highly expressed. BUB1 is able to phosphorylate CDC20 in a number of different ways. In early mitosis, BUB1 accumulates at unattached kinetochores to promote the recruitment of core mitotic checkpoint components, including mitotic spindle assembly checkpoint MAD1, mitotic spindle assembly checkpoint MAD2A, mitotic checkpoint serine/threonine kinase BUB1 β , mitotic checkpoint protein BUB3 and centromere protein E, which leads to inhibition of CDC20 (27). It has been reported that the expression of BUB1 may be a promising prognostic biomarker in a number of different cancer types, including gastric adenocarcinoma, breast cancer and ovarian cancer (20,28,29).

CDC20 is encoded by the CDC20 gene and serves an essential role in the regulation of cell division (30). The most important function of CDC20 is thought to be the activation of a master cell cycle regulator called the anaphase promoting complex/cyclosome (APC/C) by APC/C E3 ubiquitin ligase (31,32). Depletion of CDC20 has been indicated to enhance the cytotoxicity of paclitaxel, indicating that CDC20 is a potential marker that is associated with cancer (33). CDC20 has been reported to be overexpressed in numerous human cancer cell lines and carcinoma tissue types (34,35). Previous studies have also indicated that high CDC20 expression is associated with poor prognosis in multiple types of cancer (36-38). Gene expression profiling studies have iden-

tified CDC20 as an important gene that is associated with numerous types of cancer (39,40). This evidence indicates that CDC20 may be utilized as an effective potential biomarker and therapeutic target in cancer.

MAD2L1 protein is a component of the mitotic spindle assembly checkpoint and inhibits the onset of the anaphase by sequestering CDC20 until all chromosomes are aligned at the metaphase plate (41). It has been reported that MAD2L1 is highly expressed in patients with lung cancer (42), and furthermore, MAD2L1 may have the potential to be used as a diagnostic marker and therapeutic target in breast cancer (41). In addition, high intratumoral MAD2L1 levels may lead to reduced recurrence-free survival rates of patients with cancer (41).

In the present study, a total of 56 upregulated genes and 247 downregulated genes among a total of 303 DEGs were identified using bioinformatics analysis. TOP2A, CDC20, BUB1 and MAD2L1 were indicated to be significantly associated with survival of patients with lung adenocarcinoma and may serve as potential prognostic biomarkers for these patients; however, whether they can accurately predict patient prognosis requires further research because there are many other factors that can influence survival rates of patients with cancer, including age and comorbidity (43). The results of the present study may provide valuable indications for further studies on lung cancers. The absence of cell line-based analysis to confirm the biological implication of altered expression of these genes is a limitation of the present study. The absence of western blot to measure the protein expressions of these genes and immunohistochemistry to observe the expressions of these genes is also a limitation. TOP2A and BUB1 were more significantly upregulated than CDC20 and MAD2L1 in lung adenocarcinoma, according to the results of RT-qPCR, so only the results of TOP2A and BUB1 were investigated in further experiments. The absence of further data on CDC20 and MAD2L1 is an additional limitation of the current study. The absence of data on the expression of these genes in normal tissues and the expression of phospho-CDC20 are also limitations of the present study. In the present study, the relationship between the four identified hub genes and lung adenocarcinoma was validated using RT-qPCR.

The present study may use the same bioinformatics analysis tools as other microarray papers studying lung adenocarcinoma, but the present study has focused on datasets that have not, to the best of our knowledge, been analyzed previously in this way and led to the identification of unique DEGs. Analysis of further datasets may lead us to different conclusions. As the EGFR-mutation in lung cancer may be a discrete entity, the relationship between these four genes and the EGFR mutation should be a focus of future study. In addition, the expression levels of a number of genes in lung adenocarcinoma and non-lung adenocarcinoma NSCLC are different; for example, the gene expression of thymidylate synthase and excision repair cross complementing 1 in lung adenocarcinoma patients are significantly higher compared with patients with non-adenocarcinoma NSCLC (44). The data for bioinformatics analysis in the present study is from patients with lung adenocarcinoma and only four DEGs in patients with lung adenocarcinoma were detected; however, there may be differences between lung adenocarcinoma and non-adenocarcinoma NSCLC regarding the importance of these four DEGs. Thus, the present study is limited because

it lacks a comparison of the expression of these four DEGs between lung adenocarcinoma and non-lung adenocarcinoma NSCLC. Because there are only two samples, the experimental results may be contingent, which is another limitation of the present study. Finally, the Oncomine and HPA database did not provide control samples of TOP2A and BUB1 expression levels, which is another limitation of the present study.

In conclusion, TOP2A, CDC20, BUB1 and MAD2L1 may be used as biomarkers and therapeutic targets in patients with lung adenocarcinoma.

Acknowledgements

Not applicable.

Funding

The present study was supported by The Zhejiang Provincial Natural Science Foundation of China (grant no. LQ16C090001).

Availability of data and materials

The dataset used and/or analyzed during the present study is available in the GEO database (accession no. GSE27262; <https://www.ncbi.nlm.nih.gov/gds>). All other datasets used/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

SL performed the data mining. LW performed RT-qPCR and contributed to writing the manuscript. YW preserved and extracted RNA and assisted with manuscript writing. ZT and CL analyzed data and assisted with RT-qPCR. JL designed the experiments and revised the manuscript. WJ recruited patients to the study and harvested tissue. All of the authors read and approved the final manuscript.

Ethics approval and consent to participate

The present study was approved by the Ethics Committee of the Affiliated Hospital of Qingdao University (Qingdao, China; no. 2018020). Each patient was informed of the use of their tissue specimens and provided a statement of written informed consent.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Centers for Disease Control and Prevention. Lung Cancer Statistics. cdc.gov/cancer/lung/statistics/index.htm. Accessed March 26, 2014.
- Jakobsen E, Rasmussen TR and Green A: Mortality and survival of lung cancer in Denmark: Results from the Danish lung cancer group 2000-2012. *Acta Oncol* 55 (Suppl 2): S2-S9, 2016.

3. Richards TB, Henley SJ, Puckett MC, Weir HK, Huang B, Tucker TC and Allemani C: Lung cancer survival in the United States by race and stage (2001-2009): Findings from the CONCORD-2 study. *Cancer* 123 (Suppl 24): S5079-S5099, 2017.
4. Travis WD: Pathology of lung cancer. *Clin Chest Med* 32: 669-692, 2011.
5. Reck M, Heigener DF, Mok T, Soria JC and Rabe KF: Management of non-small-cell lung cancer: Recent developments. *Lancet* 382: 709-719, 2013.
6. Zhou C, Wu YL, Chen G, Feng J, Liu XQ, Wang C, Zhang S, Wang J, Zhou S, Ren S, *et al*: Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): A multicentre, open-label, randomised, phase 3 study. *Lancet Oncol* 12: 735-742, 2011.
7. Liang Y, Ma YY, Li LL, Shen XY, Xin T, Zhao YW and Ma R: Effect of long non-coding RNA LINC01116 on biological behaviors of non-small cell lung cancer cells via the hippo signaling pathway. *J Cell Biochem* 119, 2018 DOI: 10.1002/jcb.26711.
8. Barrett T and Edgar R: Mining microarray data at NCBI's Gene Expression Omnibus (GEO)[®]. *Methods Mol Biol* 338: 175-190, 2006.
9. Edgar R, Domrachev M and Lash AE: Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30: 207-210, 2002.
10. Wei TY, Juan CC, Hsia JY, Su LJ, Lee YC, Chou HY, Chen JM, Wu YC, Chiu SC, Hsu CP, *et al*: Protein arginine methyltransferase 5 is a potential oncoprotein that upregulates G1 cyclins/cyclin-dependent kinases and the phosphoinositide 3-kinase/AKT signaling cascade. *Cancer Sci* 103: 1640-1650, 2012.
11. Wei TY, Hsia JY, Chiu SC, Su LJ, Juan CC, Lee YC, Chen JM, Chou HY, Huang JY, Huang HM and Yu CT: Methylosome protein 50 promotes androgen- and estrogen-independent tumorigenesis. *Cell Signal* 26: 2940-2950, 2014.
12. Beer DG, Kardias SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, *et al*: Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 8: 816-824, 2002.
13. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A and Chinnaiyan AM: ONCOMINE: A cancer microarray database and integrated data-mining platform. *Neoplasia* 6: 1-6, 2004.
14. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, *et al*: Proteomics. Tissue-based map of the human proteome. *Science* 347: 1260419, 2015.
15. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2⁻(Delta Delta C(T)) method. *Methods* 25: 402-408, 2001.
16. Lemjabbar-Alaoui H, Hassan OU, Yang YW and Buchanan P: Lung cancer: Biology and treatment options. *Biochim Biophys Acta* 1856: 189-210, 2015.
17. Xing Z, Chu C, Chen L and Kong X: The use of Gene Ontology terms and KEGG pathways for analysis and prediction of oncogenes. *Biochim Biophys Acta* 1860: 2725-2734, 2016.
18. Wu KZ, Wang GN, Fitzgerald J, Quachthithu H, Rainey MD, Cattaneo A, Bachi A and Santocanale C: DDK dependent regulation of TOP2A at centromeres revealed by a chemical genetics approach. *Nucleic Acids Res* 44: 8786-8798, 2016.
19. Klintman M, Buus R, Cheang MC, Sheri A, Smith IE and Dowsett M: Changes in expression of genes representing key biologic processes after neoadjuvant chemotherapy in breast cancer, and prognostic implications in residual disease. *Clin Cancer Res* 22: 2405-2416, 2016.
20. Ocaña A, Pérez-Peña J, Alcaraz-Sanabria A, Sánchez-Corrales V, Nieto-Jiménez C, Templeton AJ, Seruga B, Pandiella A and Amir E: In silico analyses identify gene-sets, associated with clinical outcome in ovarian cancer: Role of mitotic kinases. *Oncotarget* 7: 22865-22872, 2016.
21. Chen T, Sun Y, Ji P, Kopetz S and Zhang W: Topoisomerase IIα in chromosome instability and personalized cancer therapy. *Oncogene* 34: 4019-4031, 2015.
22. Delgado JL, Hsieh CM, Chan NL and Hiasa H: Topoisomerases as anticancer targets. *Biochem J* 475: 373-398, 2018.
23. An X, Xu F, Luo R, Zheng Q, Lu J, Yang Y, Qin T, Yuan Z, Shi Y, Jiang W and Wang S: The prognostic significance of topoisomerase II alpha protein in early stage luminal breast cancer. *BMC Cancer* 18: 331, 2018.
24. Pei YF, Yin XM and Liu XQ: TOP2A induces malignant character of pancreatic cancer through activating β-catenin signaling pathway. *Biochim Biophys Acta* 1864: 197-207, 2018.
25. Lara-Gonzalez P, Westhorpe FG and Taylor SS: The spindle assembly checkpoint. *Curr Biol* 22: R966-R980, 2012.
26. Raaijmakers JA, van Heesbeen RGHP, Blomen VA, Janssen LME, van Diemen F, Brummelkamp TR and Medema RH: BUB1 is essential for the viability of human cells in which the spindle assembly checkpoint is compromised. *Cell Rep* 22: 1424-1438, 2018.
27. Ricke RM, Jeganathan KB, Malureanu L, Harrison AM and van Deursen JM: Bub1 kinase activity drives error correction and mitotic checkpoint control but not tumor suppression. *J Cell Biol* 199: 931-949, 2012.
28. Mukherjee A, Joseph C, Craze M, Chrysanthou E and Ellis IO: The role of BUB and CDC proteins in low-grade breast cancers. *Lancet* 385 (Suppl 1): S72, 2015.
29. Stahl D, Braun M, Gentles AJ, Lingohr P, Walter A, Kristiansen G and Gütgemann I: Low BUB1 expression is an adverse prognostic marker in gastric adenocarcinoma. *Oncotarget* 8: 76329-76339, 2017.
30. Hartwell LH, Culotti J and Reid B: Genetic control of the cell-division cycle in yeast. I. Detection of mutants. *Proc Natl Acad Sci USA* 66: 352-359, 1970.
31. Quek LS, Grasset N, Jasmen JB, Robinson KS and Bellanger S: Dual role of the anaphase promoting Complex/Cyclosome in regulating Stemness and differentiation in human primary keratinocytes. *J Invest Dermatol* 138: 1851-1861, 2018.
32. Kapanidou M, Curtis NL and Bolanos-Garcia VM: Cdc20: At the Crossroads between chromosome segregation and mitotic exit. *Trends Biochem Sci* 42: 193-205, 2017.
33. Taniguchi K, Momiyama N, Ueda M, Matsuyama R, Mori R, Fujii Y, Ichikawa Y, Endo I, Togo S and Shimada H: Targeting of CDC20 via small interfering RNA causes enhancement of the cytotoxicity of chemoradiation. *Anticancer Res* 28: 1559-1563, 2008.
34. Jiang J, Jedinak A and Sliva D: Ganodermanontriol (GDNT) exerts its effect on growth and invasiveness of breast cancer cells through the down-regulation of CDC20 and uPA. *Biochem Biophys Res Commun* 415: 325-329, 2011.
35. Chang DZ, Ma Y, Ji B, Liu Y, Hwu P, Abbruzzese JL, Logsdon C and Wang H: Increased CDC20 expression is associated with pancreatic ductal adenocarcinoma differentiation and progression. *J Hematol Oncol* 5: 15, 2012.
36. Choi JW, Kim Y, Lee JH and Kim YS: High expression of spindle assembly checkpoint proteins CDC20 and MAD2 is associated with poor prognosis in urothelial bladder cancer. *Virchows Arch* 463: 681-687, 2013.
37. Kato T, Daigo Y, Aragaki M, Ishikawa K, Sato M and Kaji M: Overexpression of CDC20 predicts poor prognosis in primary non-small cell lung cancer patients. *J Surg Oncol* 106: 423-430, 2012.
38. Karra H, Repo H, Ahonen I, Löyttyneimi E, Pitkänen R, Lintunen M, Kuopio T, Söderström M and Kronqvist P: Cdc20 and securin overexpression predict short-term breast cancer survival. *Br J Cancer* 110: 2905-2913, 2014.
39. Han Y, Jin X, Zhou H and Liu B: Identification of key genes associated with bladder cancer using gene expression profiles. *Oncol Lett* 15: 297-303, 2018.
40. He X, Zhang C, Shi C and Lu Q: Meta-analysis of mRNA expression profiles to identify differentially expressed genes in lung adenocarcinoma tissue from smokers and non-smokers. *Oncol Rep* 39: 929-938, 2018.
41. Wang Z, Katsaros D, Shen Y, Fu Y, Canuto EM, Benedetto C, Lu L, Chu WM, Risch HA and Yu H: Biological and clinical significance of MAD2L1 and BUB1, genes frequently appearing in expression signatures for breast cancer prognosis. *PLoS One* 10: e0136246, 2015.
42. Kato T, Daigo Y, Aragaki M, Ishikawa K, Sato M, Kondo S and Kaji M: Overexpression of MAD2 predicts clinical outcome in primary lung cancer patients. *Lung Cancer* 74: 124-131, 2011.
43. Asmis TR, Ding K, Seymour L, Shepherd FA, Leighl NB, Winton TL, Whitehead M, Spaans JN, Graham BC and Goss GD: National Cancer Institute of Canada Clinical Trials Group: Age and comorbidity as independent prognostic factors in the treatment of non small-cell lung cancer: A review of National Cancer Institute of Canada Clinical Trials Group trials. *J Clin Oncol* 26: 54-59, 2008.
44. Zou ZQ, Du YY, Sui G and Xu SN: Expression of TS, RRM1, ERCC1, TUBB3 and STMN1 genes in tissues of non-small cell lung cancer and its significance in guiding postoperative adjuvant chemotherapy. *Asian Pac J Cancer Prev* 16: 3189-3194, 2015.