Bioinformatics analysis of BUB1 expression and gene regulation network in lung adenocarcinoma

Luyao Wang1#, Xue Yang2#, Ning An3, Jia Liu1

1Department of Pharmacology, School of Pharmacy, Qingdao University, Qingdao, China; 2Department of Medical Oncology, The Affiliated Hospital of Qingdao University, Qingdao University, Qingdao, China; 3Department of Radiation Oncology, The Affiliated Hospital of Qingdao University, Qingdao University, Qingdao, China

Contributions: (I) Conception and design: L Wang, J Liu; (II) Administrative support: None; (III) Provision of study materials or patients: X Yang, N An; (IV) Collection and assembly of data: L Wang, J Liu; (V) Data analysis and interpretation: X Yang, N An; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

#These authors contributed equally to this work.

Correspondence to: Ning An, MD, PhD. Department of Radiation Oncology, The Affiliated Hospital of Qingdao University, Qingdao University, Qingdao 266000, China. Email: anning0921@outlook.com; Jia Liu, PhD. Associated Professor, Department of Pharmacology, School of Pharmacy, Qingdao University, Qingdao 266000, China. Email: dadaliujia@qdu.edu.cn.

Background: Lung adenocarcinoma is the most common type of lung cancer with high morbidity and mortality. Potential mechanisms and therapeutic targets of lung adenocarcinoma need further study. BUB1 (BUB1 mitotic checkpoint serine/threonine kinase) encodes a serine/threonine protein kinase which is critical in mitosis. It is associated with poor prognosis in multiple cancer types.

Methods: Oncomine database was used to determine the differential expression of BUB1 in normal and lung adenocarcinoma tissues, while UALCAN was used to perform analysis of the relative expression and survival of BUB1 between tumor and normal tissues in different tumor subgroups. We used the cBioPortal for Cancer Genomics to perform gene ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of the top 50 altered neighbor genes of BUB1. The LinkedOmics database was used to determine differential gene expression with BUB1 and to perform functional analysis. The kinase, miRNA and transcription factor target networks correlated with BUB1 were also analyzed by LinkedOmics database.

Results: The results revealed that BUB1 was highly expressed in lung adenocarcinoma patients. BUB1 involved multiple tumor-related pathways, such as cell cycle, oocyte meiosis and p53 signaling pathway. BUB1 is associated with tumor-associated kinases, microRNAs and transcription factors.

Conclusions: Our study analyzed BUB1 expression and potential gene regulation networks in lung adenocarcinoma based on bioinformatics analysis, guiding further study on the role and regulation of BUB1 in lung adenocarcinoma. BUB1 may hopefully become a novel marker and therapeutic target for lung adenocarcinoma.

Keywords: Bioinformatics analysis; BUB1; lung adenocarcinoma (LUAD)

doi: 10.21037/tcr-20-1045

View this article at: http://dx.doi.org/10.21037/tcr-20-1045

Introduction

Lung cancer is the most commonly diagnosed cancer and the most common mortality cause of cancer worldwide (1). Lung cancer patients are usually in advanced stage at the time of initial diagnosis and their 5-year net survival is low (10–20%) (2). Lung adenocarcinoma (LUAD), which is the most common form of lung cancer, is a multifactorial disease. Although significant progress has been made in the
treatment of LUAD in recent years (3,4), the survival rate is still very low. The pathogenesis of LUAD is complex and may involve the interaction of multiple genes in multiple processes. As data mining has provided a new method for finding tumor biomarkers and therapeutic targets, it may be also possible to identify novel therapeutic targets for LUAD by establishing gene networks associated with tumorigenesis and progression.

BUB1, whose official full name is BUB1 mitotic checkpoint serine/threonine kinase, encodes a serine/threonine protein kinase which is critical in the process of mitosis (5). The kinase BUB1 is essential in the spindle assembly checkpoint (SAC) and chromosome congression (6). BUB1 may also play a role in DNA damage response. Studies have shown that BUB1 is associated with the maintenance of cancer stem cells (7). Also, BUB1 may be a promising prognostic biomarker in cancer as high expression of BUB1 has been proved to be related to poor prognosis in multiple cancer types (8-10).

Our previous analysis of LUAD data in the Gene Expression Omnibus database revealed that BUB1 may be a key protein-coding gene in LUAD (11). Our studies in patients with LUAD have shown that BUB1 is overexpressed in tumor tissues than in adjacent normal tissues. Survival analysis has shown that overexpression of BUB1 was associated with poor prognosis of LUAD patients. Therefore, BUB1 may be a new proto-oncogene closely related to LUAD. Thus, we further studied the expression and mutations of BUB1 in LUAD patients in a variety of public databases. We also analyzed the genomic changes and functional networks associated with BUB1 in LUAD using multidimensional analysis methods. Our study may provide novel targets and novel approaches for diagnosis and treatment of LUAD.

We present the following article in accordance with the MDAR checklist (available at http://dx.doi.org/10.21037/tcr-20-1045).

Methods

Oncomine analysis

The expression of BUB1 between normal tissues and LUAD tissues was analyzed by Oncomine database (www.oncomine.org), which is the biggest oncogene chip database in the world (12). This analysis involved Garber Lung, Hou Lung, Landi Lung and Su Lung. In LUAD tissues, BUB1 expression was assessed in LUAD tissues relative to its expression in normal tissues. P<0.01 was considered significant. Oncomine database was also used to analyze survival of LUAD patients in various subgroups. P<0.01 was considered significant.

UALCAN analysis

UALCAN (http://ualcan.path.uab.edu), an open web-portal for tumor subgroup gene expression and survival analyzes, contains data from The Cancer Genome Atlas (TCGA) project (13). UALCAN was used to perform analysis of the relative expression of BUB1 between tumor tissues and normal tissues in different tumor subgroups based on gender, age, race, cancer stage and smoking habits. UALCAN was also used to analyze survival of LUAD patients in various subgroups.

c-BioPortal analysis

The cBioPortal for Cancer Genomics, which contains 225 cancer studies, provides resource for exploration and analysis of multidimensional cancer genomics data (14,15). The cBioPortal was used to analyze type and frequency of BUB1 alterations in LUAD. The OncoPrint shows an overview of genetic alterations each sample in BUB1. The cBioPortal is publicly available at http://cbioportal.org. STRING (http://string-db.org), an online tool for the retrieval of interacting genes and proteins, was used to establish PPI network of the top 50 altered neighbor genes of BUB1. We then performed gene ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of these 50 genes using database for annotation, visualization and integrated discovery (DAVID; http://david.ncifcrf.gov), which is a free and open-access website for researchers to analyze the biological significance behind genes (16). P<0.001 was considered significant.

LinkedOmics analysis

The LinkedOmics database is a multi-omics database that contains data of 32 cancer types and 11,158 patients from TCGA project (17). We used LinkCompare module of LinkedOmics to analyze genes differentially expressed in correlation with BUB1. Pearson’s correlation coefficient was used to perform statistical analysis. The tab LinkInterpreter of LinkedOmics performs enrichment analysis of differentially expressed genes using Gene Set Enrichment Analysis (GSEA).
GeneMANIA analysis

GeneMANIA (http://www.genemania.org) is an open-access web interface which contains hundreds of data sets from several databases (18). GeneMANIA was used to construct protein-protein interaction (PPI) network of genes that GSEA identified as being enriched in LUAD: ATR serine/threonine kinase (ATR) and miR-143.

Results

Expression of BUB1 in LUAD

The gene expression level of BUB1 was analyzed by Oncomine database. The results of Oncomine analysis revealed that the expression and DNA copy number variation (CNV) of BUB1 in LUAD tissues were significantly higher than that in normal tissues (P<0.01). Moreover, the fold differences were more than 2 and BUB1 ranked within the top 4% at the level of mRNA expression (Figure 1). The results of UALCAN analysis showed that BUB1 transcription level in LUAD patients was significantly higher than that in healthy people in subgroup analysis based on gender, age, race, cancer stage and smoking habits (P<0.001) (Figure 2). Therefore, the expression of BUB1 may be a potential diagnostic indicator of LUAD.
Survival analysis of LUAD patients based on various subgroups

Survival analysis based on UALCAN showed that the expression of BUB1 and patients’ race may have an impact on survival of LUAD patients (P<0.05) (Figure 3).

BUB1 alterations in LUAD

Then we used cBioPortal to determine the type and frequency of BUB1 alterations in LUAD according to the sequencing data of patients in TCGA database. BUB1 was altered in 59 (12%) of LUAD patients (Figure 4). The alteration frequency was shown in Figure 5. The most frequent alteration is mRNA upregulation.

Biological interaction network of BUB1 alterations in LUAD

The PPI network of the top 50 altered neighbor genes of BUB1 established by STRING online database was shown in Figure 6A. The GO analysis based on DAVID database showed that these 50 genes were enriched in cancer-associated biological pathways including protein binding, microtubule binding, ATP binding, microtubule motor activity, as well as protein serine/threonine kinase activity. For cellular component, these genes were enriched in condensed chromosome kinetochore, chromosome, centromeric region, cytosol, spindle and kinetochore (Figure 6B). GO enrichment analysis in the category biological process showed that these genes were enriched in cell division, mitotic nuclear division, sister chromatid
cohesion, chromosome segregation, as well as mitotic sister chromatid segregation (Figure 6B). The results of the KEGG analysis showed that these 50 genes were enriched in cell cycle, oocyte meiosis, progesterone-mediated oocyte maturation and p53 signaling pathway (Figure 6C). Above results indicated that biological interaction network of BUB1 alterations is mainly in chromosome and is mainly associated with cell cycle.

**Co-expression network of BUB1 in LUAD**

We used LinkedDomics Functional module to analyze the sequencing mRNA data of patients in the TCGA. Genes differentially expressed in correlation with BUB1 in LUAD was shown in Figure 7. This result suggests that BUB1 has a wide influence on the transcriptome. GO terms show that genes differentially expressed in correlation with BUB1 mainly concentrated in the condensed chromosome, chromosomal region and spindle (Figure 8A). These genes were enriched in the category biological process, including chromosome segregation, microtubule cytoskeleton organization involved in mitosis and DNA replication (Figure 8B). In addition, these genes were enriched in cancer-associated terms in the category molecular function.

---

**Figure 3** Survival analysis of BUB1 in subgroups of patients with lung adenocarcinoma (UALCAN). The BUB1 level and race have significant influence on the survival of lung adenocarcinoma patients (P<0.05), while gender and smoking habit not. (A) Effect of BUB1 expression level on lung adenocarcinoma patient survival; (B) effect of BUB1 expression level and race on lung adenocarcinoma patient survival; (C) effect of BUB1 expression level and gender on lung adenocarcinoma patient survival; (D) effect of BUB1 expression level and smoking habit on lung adenocarcinoma patient survival.
including histone deacetylase binding, catalytic activity acting on DNA and ATPase activity (Figure 8C). The results of the KEGG analysis showed that these genes were enriched in biological pathways including cell cycle, oocyte meiosis and DNA replication (Figures 8D,9).

Kinase, microRNA (miRNA) and transcription factor target networks correlated with BUB1 in LUAD

We then analyzed the kinase, miRNA and transcription factor target networks correlated with BUB1 based on GSEA. The miRNA, kinase and transcription factor-target networks of BUB1 in LUAD was shown in Figure 10. The top 5 miRNA-target network was associated with (TCATCTC) MIR-143, (GTAAGAT) MIR-200A, (GTACTGT) MIR-101, (ATGCAGT) MIR-217, (TATTATA) MIR-374. The top 5 most significant kinase-target networks were cyclin dependent kinase 2 (CDK2), cyclin dependent kinase 2 (CDK2), ATR, ATM serine/threonine kinase (ATM) and aurora kinase B (AURKB). The transcription factor-target network was significantly related to E2F family, such as VSE2F1_Q6, VSE2F_Q6, VSE2F_Q4, VSE2F4DP1_01 and VSE2F1DP1_01. PPI network of genes that GSEA identified as being enriched in LUAD (kinase ATR and miR-143) was shown in Figure 11.

Discussion

BUB1, which is related to kinase function and the spindle assembly checkpoint, has been reported to be highly expressed in the process of mitosis and cell proliferation (19). Our previous study has shown that BUB1 was over-expressed in LUAD tissue compared to normal tissue and high BUB1 expression was associated with poor prognosis of LUAD patients, indicating that BUB1 may be a key gene in LUADs. In order to understand the potential function of BUB1 in LUAD and its regulatory network in more detail, we used multiple databases to conduct bioinformatics analysis of public sequencing data to guide future study on LUAD.

LUAD is the most common form of lung cancer (20). Patients with lung cancer are usually in advanced stage at the time of first diagnosis and the survival rate is very low. New markers may be helpful in the diagnosis and treatment of LUAD. Analysis of sequencing data of LUAD from GEO database and TCGA database showed that BUB1 mRNA level and DNA CNV in LUAD were significantly higher than that in normal tissues. The fold differences were more than 2 and BUB1 ranked between top 1–4% at the level of mRNA expression. The over-expression of BUB1 in LUAD revealed that BUB1 may be a promising bio-marker of LUAD and it is worth further study. Our study also found that the most frequent change in BUB1 in LUAD was mRNA upregulation. Altered BUB1 expression may be critical for the occurrence and progress of LUAD since BUB1 plays an important role in DNA damage response and BUB1 is required for the maintain of cancer stem cells (7). Alteration of BUB1 may cause changes
in downstream signaling pathways. The top 50 altered neighbor genes of BUB1 were enriched in cell cycle, oocyte meiosis, progesterone-mediated oocyte maturation, as well as p53 signaling pathway. The biological interaction network of BUB1 alterations is mostly in the chromosome and is mainly associated with the cell cycle and p53 pathway. This result is consistent with the functions of BUB1 (21-23).

GSEA was used to analyze the co-expressed network of BUB1. The results showed that genes differentially expressed in correlation with BUB1 were enriched in cell cycle, oocyte meiosis and DNA replication. These genes were also enriched in tumor-associated terms in the category molecular function, biological process and cellular component, proving that BUB1 has a wide influence on the transcriptome. Genome instability and mutation are the basic characteristics of cancer cells (24,25). Kinases and their related signaling pathways contribute to the stabilization and repair of genomic DNA (26,27).

The results revealed that BUB1 in LUAD is associated with tumor-associated kinases, including CDK1, CDK2, ATR, ATM and AURKB. These kinases play a role in the cell cycle, mitosis and genomic stability (28-30).

The target network of transcription factors has significant relationship with E2F family, including V $E2F1 UQ6, V $E2F UQ6, V $E2F UQ4, V $E2F4DP1 U01 and V $E2F1DP1 U01. It has been reported that the E2F transcription factor is one of the key regulators of cell cycle progression and triggering apoptosis (31). Also, alterations of E2F may be associated with shorter survival of cancer patients (32). Moreover, the E2F1 member of the E2F family is associated with metastasis of cancer cells (33). Studies have shown that up-regulated transcription of E2F1 may promote the progression of non-small cell lung cancer (34). Thus, E2F transcription factor family may be a critical target of BUB1, and that BUB1 may regulate cell cycle through this transcription factor.

The significant miRNA-target networks of BUB1 were identified by GSEA. MiRNAs, endogenous RNAs...
Figure 7 Genes differentially expressed in correlation with BUB1 in lung adenocarcinoma (LinkedOmics). (A) Positively correlated genes; (B) negatively correlated genes; (C) positively correlated genes; (D) negatively correlated genes.
of about 22 nt, is an abundant class of gene regulators in multicellular organisms (35). miRNAs may affect the expression of many protein-coding genes. Studies have shown that the dysregulation of miRNAs may be the cause of many cancers, and thus miRNA-targeted therapeutics has attracted much attention (36). In our study, miR-143, miR-200A, miR-101, miR-217 and miR-374 were identified as miRNAs significantly associated with BUB1. MiR-143 has been considered to have potential as tumor marker for non-small cell lung cancer (37), while miR-200A has been
indicated to play a role in the progression of LUAD (38). MiR-101 may exert tumor inhibitory effect in non-small cell lung cancer (39). MiR-217 has been reported to be highly associated with survival time of LUAD patients (40). MiR-374 has been indicated to play a role in various cancers and it may be a promising prognostic biomarker and therapeutic target of cancer (41). The relationship between these miRNAs and BUB1 and their effects on LUAD need further experimental verification. They may provide novel tumor biomarkers and therapeutic targets for LUAD.

Figure 9 KEGG pathway (LinkedOmics; GSEA). (A) KEGG pathway annotations of the cell cycle pathway; (B) KEGG pathway annotations of the oocyte meiosis; (C) KEGG pathway annotations of the DNA replication. KEGG, Kyoto Encyclopedia of Genes and Genomes.
Our study used multiple databases which contain a large number of clinical data to study BUB1 expression and gene regulatory network in LUAD. Our study provides evidence for the importance of BUB1 in LUAD and its potential as a marker of LUAD. The results revealed that BUB1 was highly expressed in LUAD patients. The over-expression of BUB1 had a profound impact on genomic stability and involved multiple tumor-related pathways, such as cell cycle, oocyte meiosis and p53 signaling pathway. BUB1 is associated with several tumor-associated kinases (such as CDK1, CDK2, ATR, ATM and AURKB) miRNAs (such as miR-143, miR-200A, miR-101, miR-217 and miR-374) and...
transcription factors (such as E2F family).

In conclusion, BUB1 may hopefully become a new marker and therapeutic target for LUAD. In future study, we will further validate the regulatory network of BUB1 in LUAD.

Acknowledgments

Funding: This work was supported by National Natural Science Foundation of China (No. 81802271 and 81801734).

Footnote

Reporting Checklist: The authors have completed the MDAR checklist. Available at http://dx.doi.org/10.21037/tcr-20-1045

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr-20-1045). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References


Figure 11 Protein-protein interaction network of genes that GSEA identified as being enriched in lung adenocarcinoma (GeneMANIA). (A) PPI network of ATR kinase-target networks; (B) PPI network of miR-143 networks. GSEA, Gene Set Enrichment Analysis; PPI, protein-protein interaction.


