

# Bioinformatic analysis of the molecular mechanism underlying bronchial pulmonary dysplasia using a text mining approach

Weitao Zhou, MD<sup>a</sup>, Fei Shao, MD<sup>b</sup>, Jing Li, MD<sup>c,d,\*</sup>

## Abstract

Bronchopulmonary dysplasia (BPD) is a common disease of premature infants with very low birth weight. The mechanism is inconclusive. The aim of this study is to systematically explore BPD-related genes and characterize their functions.

Natural language processing analysis was used to identify BPD-related genes. Gene data were extracted from PubMed database. Gene ontology, pathway, and network analysis were carried out, and the result was integrated with corresponding database.

In this study, 216 genes were identified as BPD-related genes with  $P < .05$ , and 30 pathways were identified as significant. A network of BPD-related genes was also constructed with 17 hub genes identified. In particular, phosphatidylinositol-3-enzyme-serine/threonine kinase signaling pathway involved the largest number of genes. Insulin was found to be a promising candidate gene related with BPD, suggesting that it may serve as an effective therapeutic target.

Our data may help to better understand the molecular mechanisms underlying BPD. However, the mechanisms of BPD are elusive, and further studies are needed.

**Abbreviations:** BP = biological process, BPD = bronchopulmonary dysplasia; bronchial pulmonary dysplasia, CC = cellular component, GO = gene ontology, IL = interleukin, INS = insulin, KEGG = Kyoto Encyclopedia of Genes and Genomes, MF = molecular function, NLP = natural language processing, TNF = tumor necrosis factor, TP53 = tumor protein P53, VEGF = vascular endothelial growth factor, VEGFA = vascular endothelial growth factor A.

**Keywords:** bronchial pulmonary dysplasia, PI3K-AKT signaling pathway, systematic analysis, text mining

## 1. Introduction

Bronchial pulmonary dysplasia (bronchopulmonary dysplasia [BPD]) is a common disease of premature infants with low birth

Editor: Jing Liu.

WZ and FS contributed equally to this work.

This study did not involve any human or animal experiment. Ethical approval was not required for secondary analysis of anonymous data in this study.

This study was supported in part by the National Natural Science Foundation of China

(No. 81101442).

The authors have no conflicts of interest to disclose.

<sup>a</sup> Department of Pediatrics, The First Affiliated Hospital of the University of Science and Technology of China, <sup>b</sup> Department of Oncology, Second Affiliated Hospital of Anhui Medical University, Hefei, <sup>c</sup> Department of Pediatric Intensive Care Unit, Children's Hospital of Chongqing Medical University; Ministry of Education Key Laboratory of Child Development and Disorders; National Clinical Research Center for Child Health and Disorders; China International Science and Technology Cooperation base of Child Development and Critical Disorders; Children's Hospital of Chongqing Medical University, <sup>d</sup> Chongqing Key Laboratory of Pediatrics, Chongqing, China.

\* Correspondence: Jing Li, Department of Pediatric Intensive Care Unit, Children's Hospital of Chongqing Medical University, Chongqing 400014, China (e-mail: lijingwangyi@126.com).

Copyright © 2019 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Zhou W, Shao F, Li J. Bioinformatic analysis of the molecular mechanism underlying bronchial pulmonary dysplasia using a text mining approach. *Medicine* 2019;98:52(e18493).

Received: 5 April 2019 / Received in final form: 13 November 2019 / Accepted: 22 November 2019

<http://dx.doi.org/10.1097/MD.00000000000018493>

weight, characterized by continuing oxygen therapy at least 28 days, accompanied by pulmonary imaging changes.<sup>[1]</sup> Pulmonary surfactant and small tidal volume aeration can alleviate lung injury and cut down premature infant mortality, but not decrease the incident rate of BPD. Recent studies show that 1 in every 10,000 children in the United States is diagnosed with BPD.<sup>[2]</sup> The etiology of BPD is mainly due to the genetic susceptibility, oxygen toxicity, air pressure injury, capacity injury, infection, immunity and so on, resulting in oxygen toxicity and pulmonary fibrosis. However, the pathogenic factor is not clear.<sup>[3]</sup> With the development of technique in molecular biology and genetic engineering, studies on molecular mechanism underlying the pathogenesis of BPD have been increasing in recent decades. The pathogenesis of BPD has been shown to be resulted from genetic susceptibility, infection, immature lung development, high concentration of oxygen injury, nutritional deficiency, and other factors.<sup>[4]</sup>

A large number of literatures show that the pathogenesis of BPD is attributed to some important protein and signaling pathways, such as angiotensin-converting enzyme, mannose-binding lectin, nuclear factor kappa B, phosphatidylinositol-3-enzyme-serine/threonine kinase (PI3K-AKT) signaling pathway, interleukin (IL), vascular endothelial growth factor (VEGF).<sup>[5-8]</sup> Although a large number of transcription factors or signaling proteins have been discovered, studies on these transcription factors tend to be more conservative. Furthermore, defining the most meaningful regulatory factors from the information network is an important project.<sup>[9]</sup>

Due to the development of high-throughput proteomic and transcriptomic approach, it is feasible to study tens of thousands of genes and proteins at the same time nowadays. However, the results of high-throughput data are often lack of

consistency due to the different choices of data platform and statistical criteria.<sup>[10]</sup> Moreover, most valuable information remains hidden within literatures through conventional gene-by-gene way. In recent years, text mining has been used as an effective method to explore underlying mechanisms for many diseases.<sup>[11,12]</sup> It provides a way to retrieve data from published researches automatically.<sup>[13]</sup> Text mining, as an effective method for the study of molecular mechanism, has been utilized in many diseases,<sup>[14]</sup> such as: glioblastoma,<sup>[15]</sup> endometrial cancer,<sup>[16]</sup> breast cancer,<sup>[17,18]</sup> ectopic pregnancy,<sup>[19]</sup> decidualization.<sup>[20]</sup>

In an attempt to provide a better understanding of BPD and pave a foundation for the development of novel therapeutic interventions in BPD, we performed a text mining analysis to identify genes related to BPD. The filtered gene set was subsequently carried out to systematically analyze their functions, and define signaling pathways and network involved.

**2. Methods**

**2.1. Natural Language processing (NLP) analysis of BPD**

PubMed database was used as the source of literature for text mining. The search was carefully conducted with the query term “bronchopulmonary dysplasia.” All relevant articles identified till March 2018 were retrieved and converted into XML format. All the genes and proteins associated with the query term were dug out and added to the list, followed by Gene mention tagging using a biomedical named entity recognizer software.<sup>[21]</sup> Conjunction resolution was also carried out to identify individual descriptions on the extracted genes. In this study, gene symbol in entrez gene database of the National Center for Biotechnology Information was commonly used.<sup>[22,23]</sup> The processing flow chart of the NLP analysis of bronchopulmonary-dysplasia was shown in Figure 1.

For each gene, the frequency of its occurrence was denoted. The higher frequency a gene has, the higher correlation between the certain gene and BPD. The total number of publications retrieved from the PubMed database was recorded as *N*. The frequency of the certain gene and BPD were denoted by *m* and *n*, respectively.

*K* was used to represent the occurrence of both the gene and BPD in actual situations. Then, we calculated the probability of the frequency greater than *k* simultaneously cited under the

completely random conditions by using hypergeometric distribution:

$$p = 1 - \sum_{i=0}^{k-1} p(i|n, m, N)$$

$$p(i|n, m, N) = \frac{n!(N - n)!m!(N - m)!}{(n - i)!i!(n - m)!(N - n - m + i)!N!}$$

The BPD-gene relations with *P*-value < .05 were retrieved for further use.

**2.2. Gene ontology (GO) analysis**

GO analysis was performed by using the GSEABase package from R (<http://www.r-project.org/>) statistical platform. Word cloud was generated by using the word cloud package from R.<sup>[24]</sup> The biological process (BP), cellular component (CC), and molecular function (MF) were characterized and evaluated in this study.

**2.3. Pathway analysis**

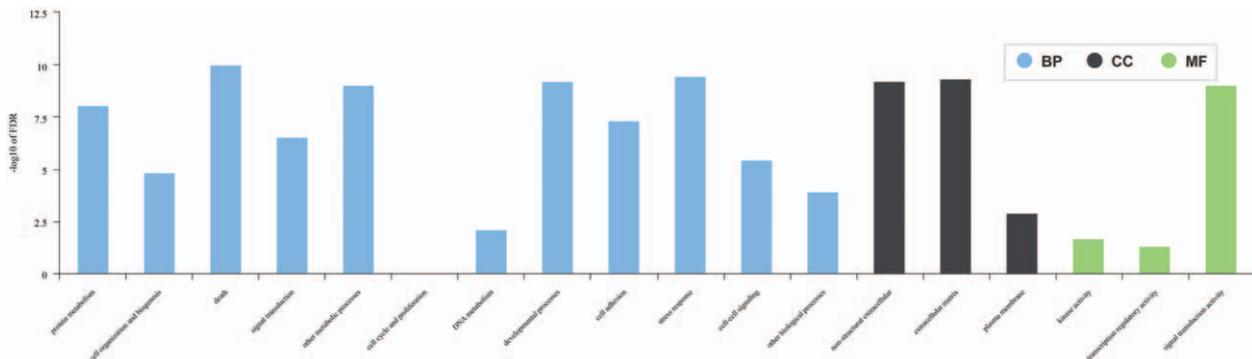
The Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used to analyze the interactions between genes. Genes were mapped to the database by using GenMAPP v2.1 (<http://www.genmapp.org/>). Then statistical tests were performed to identify enriched pathways. *P* < .05 was set as threshold.<sup>[25]</sup>

**2.4. Network analysis of BPD genes**

We integrated the relationships of involved genes into 3 different interaction categories:

- (1) protein interaction, gene regulation, protein modification listed in the KEGG database;
- (2) interaction data from existing high-throughput protein interaction experiments, such as protein-protein interactions (PPrel) confirmed by yeast 2-hybrid;
- (3) gene interaction demonstrated in previous publications.

Briefly, the pathway data were downloaded from the KEGG database, and then used to analyze the interaction among genes. We also downloaded KEGGSOAP package from R statistical



**Table 1****The top 20 frequently cited genes related to bronchopulmonary dysplasia.**

Official gene symbol	PubMed Count	P-value	Gene description
VEGFA	114	0	Vascular endothelial growth factor A
IL6	94	0	Interleukin 6 (interferon, beta 2)
TNF	61	0	Tumor necrosis factor (TNF superfamily, member 2)
IL1B	56	0	Interleukin 1, beta
ELN	55	0	Elastin
SFTPB	54	0	Surfactant protein B
IL10	37	0	Interleukin 10
SFTPC	37	0	Surfactant protein C
EPO	32	0	Erythropoietin
KDR	31	0	Kinase insert domain receptor (a type III receptor tyrosine kinase)
SFTPA1	29	0	Surfactant protein A1
SOD1	29	0	Superoxide dismutase 1, soluble
GRP	28	0	Gastrin-releasing peptide
CCL2	27	0	Chemokine (C-C motif) ligand 2
CTGF	21	0	Connective tissue growth factor
NOS3	20	4.77E-15	Nitric oxide synthase 3 (endothelial cell)
MMP9	18	6.76E-14	Matrix metalloproteinase 9 (gelatinase B, 92 kDa Gelatinase, 92 kDa type IV collagenase)
MPO	18	0	Myeloperoxidase
SFTPD	18	0	Surfactant protein D
CYP1A1	16	9.15E-14	Cytochrome P450, family 1, subfamily A, polypeptide 1

analysis platform (<http://www.bioconductor.org/packages/2.4/bioc/html/KEGGSOAP.html>). Three different types of relationships were analyzed: enzyme-enzyme relation (ECrel, indicating 2 enzymes catalyzing successive reaction steps), PPrel (such as binding and modification), gene expression interaction (GElrel, indicating relation of transcription factors and target gene products).<sup>[26]</sup> The PPrel data was obtained by using the mammalian protein-protein interaction database (<http://mips.helmholtz-muenchen.de/proj/ppi>).<sup>[27]</sup> For interactions that had already been published, we used co-citation matrices in PubMed. Using this algorithm, we identified the certain gene terms and co-occurred term variants within the sentences of an abstract. The frequency of the co-cited genes was analyzed as well. Finally, statistical analysis was performed using the hypergeometric distribution as described above. The resulting network of relationship was built and displayed by Medusa software.<sup>[28]</sup> Then, the genes with a large number of connections that play important roles in the network stability were identified.

### 3. Results

#### 3.1. NLP analysis of BPD

After the retrieval of documents from Pubmed, a total number of 5812 primary studies were identified, and 397 were obtained. Eventually, 216 genes were identified as BPD-related genes with  $P < .05$ . Among these genes, vascular endothelial growth factor A (VEGFA), IL-6, tumor necrosis factor (TNF), and interleukin 1, beta were mentioned most frequently. The top 20 frequently cited genes were listed in Table 1

#### 3.2. GO analysis

These 216 genes mentioned above were categorized in GO analysis. Enriched GO terms are classified according to BP, CC, and MF. In the BP category, 12 GO terms, namely protein metabolism, cell organization and biogenesis, death, signal transduction, other metabolic processes, cell cycle and prolifera-

tion, DNA metabolism, developmental processes, cell adhesion, stress response, cell-cell signaling, and other BP were found to be significantly enriched. GO terms related to nonstructural extracellular, extracellular matrix, and plasma membrane region were significantly enriched under the CC category. Enriched GO terms in the MF category included kinase activity, transcription regulatory activity, signal transduction activity (Fig. 1).

#### 3.3. Pathway analysis and gene network

To better understand the gene function related to BPD, we also performed pathway analysis by using DAVID online tools. All BPD-related genes were mapped to the pathways found in KEGG. As shown in Figure 2, 30 signaling pathways were identified as significant. Pathway assignment for the top 10 genes was recommended by the TM-rank algorithm and generated by Cytoscape software (Fig. 3).

We constructed a network of BPD-related genes that consists of 365 nodes connected via 6403 edges (Fig. 4A). Topological analysis proved that the network followed a power-law distribution (Fig. 4B). In this network, we identified 17 hub genes (Fig. 4C): tumor protein (TP53), v-akt murine thymoma viral oncogene homolog 1, jun oncogene, IL6, insulin (INS), B-cell CLL/lymphoma 2, VEGFA, epidermal growth factor receptor, TNF, transforming growth factor, beta 1, matrix metalloproteinase 9, matrix metalloproteinase 2, fibroblast growth factor 2, intercellular adhesion molecule 1, nitric oxide synthase 3, mitogen-activated protein kinase 8, and cyclin D1. Among these genes, TP53 was the BPD-related gene that exhibited the greatest number of interactions.

### 4. Discussion

A total number of 216 BPD-related genes were identified from the literature search of 5812 publications. Among the BPD-related genes we identified in the search, VEGFA, IL6 were the most frequently mentioned. Studies have already shown that there is a relevant association between BPD development and VEGFA

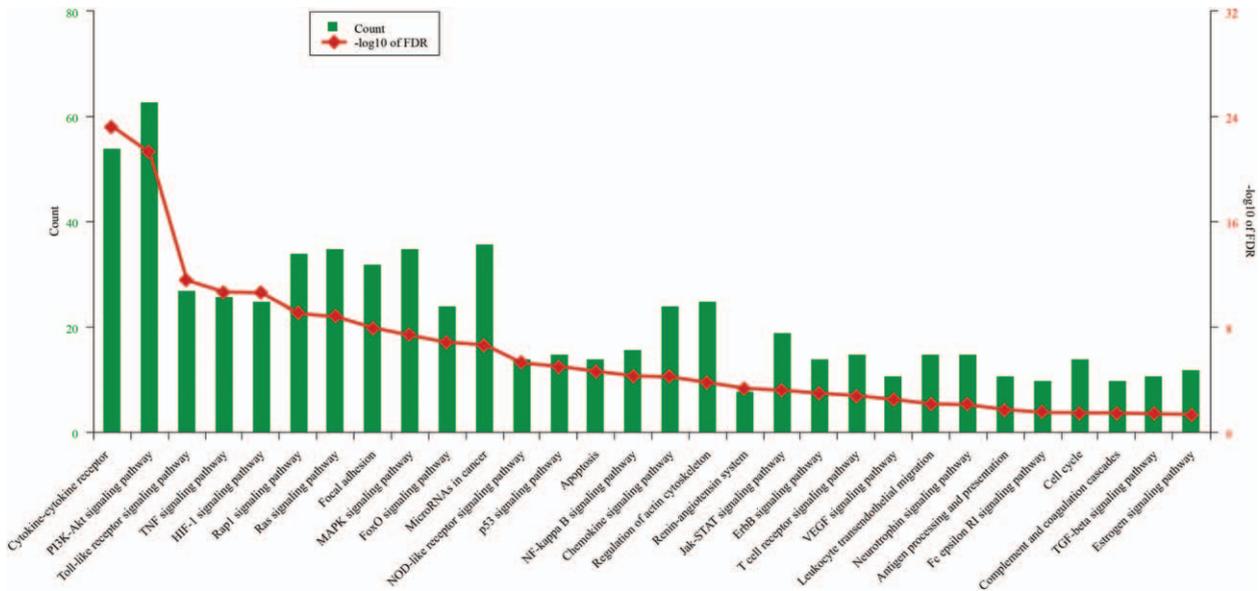


Figure 2. The significantly enriched pathways identified by using DAVID online tools.

activity in preterm infants.<sup>[29,30]</sup> VEGFA was shown to play vital roles in the repair of vascular lung, and the lack of VEGFA activity could lead to an impairing lung microvascular development in fetus.<sup>[31]</sup> IL6 expression was found to be increased with the histologic severity of chorioamnionitis as well.<sup>[32]</sup>

The role of BPD-related genes could be validated through wet experiments theoretically; however, due to the large number of genes involved, experimental work is not a feasible for us to get a comprehensive understanding of the gene set. Thus, we

performed pathway enrichment analysis using the complete list of 216 candidate genes by using the DAVID tools. A total of 30 enriched pathways were identified. Of particular interest was the PI3K-Akt signaling pathway. Notably, 67 BPD-related genes were enriched in this pathway. PI3K becomes activated by INS and is responsible for most of the metabolic actions mediated by INS. PI3K plays essential roles in many metabolic process such as cell growth, differentiation, survival, and protein synthesis.<sup>[33]</sup> Inhibition of the PI3K-Akt pathway disrupts normal lung

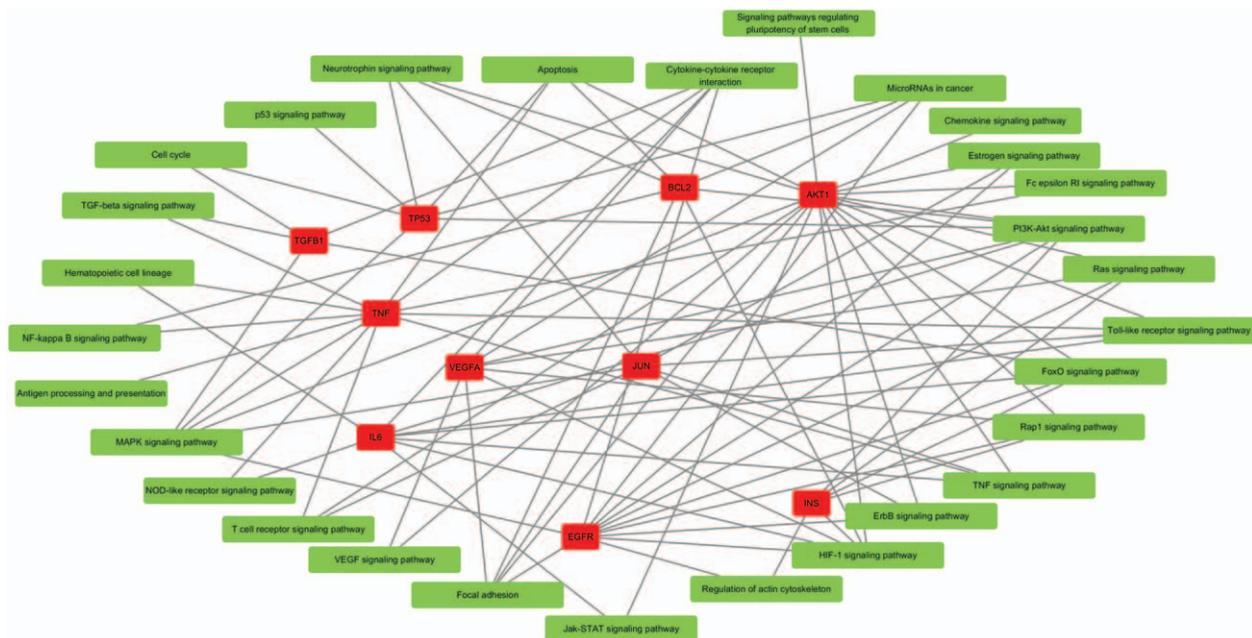
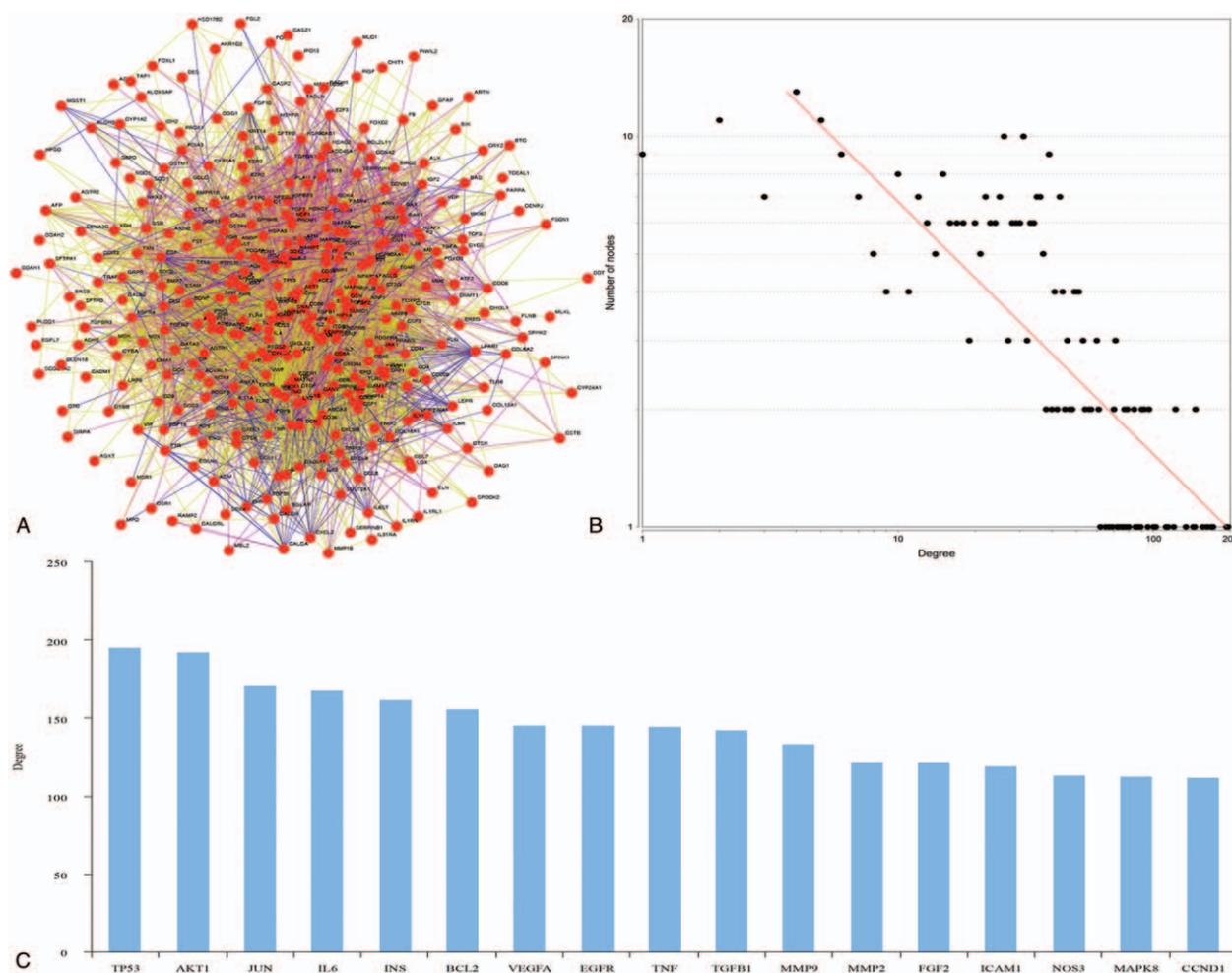


Figure 3. Pathway assignment for the top 10 genes recommended by the TM-rank algorithm. This graph was made by Cytoscape software. Ellipse nodes in red represent genes and rectangle nodes in blue represent pathways.



**Figure 4.** The gene network underlying all bronchopulmonary dysplasia candidate genes. (A) The structure of the gene network generated by using the MIPS database. (B) Degree distribution of the gene network. (C) The seventeen hub genes identified in the network. MIPS = mammalian protein-protein interaction.

development, whereas the activation of PI3K-Akt pathway preserves alveolar development.<sup>[34]</sup> PI3K-Akt signaling pathway could exert a cytoprotective role in cell survival during hyperoxia.<sup>[35]</sup>

We interpret that text mining result is usually based on the frequency calculation for each gene in publications. However, the most popular genes are not always the most important ones. It is known that, rather than working alone, gene products such as proteins usually form complexes to exert their functions.<sup>[36]</sup> Hence, the functional importance of a gene depends on its interaction with other partners. The hub genes are likely more important due to their key positions in the ECrel, PPrel and GERel network. In this study, we constructed a large network of BPD-related genes that consists of 365 nodes connected via 6403 edges. Seventeen hub genes were identified in this network. Even the noisiness and incompleteness of interaction data may cause the inaccuracy of our results, the network could still provide us a comprehensive and reasonable way to understand the gene set.

Text Mining is characterized as the way toward separating meaningful information from numerous unstructured text utilizing computational methods. However, these articles of high-throughput experiments generate a large amount of gene

information sometimes cannot be fully recognized and refined. For these articles, we downloaded full texts (as well as supplementary files if needed) and extracted gene mentions by hands. To reduce false-positive results, full texts of extracted papers contained hub-gene were downloaded, followed by manual confirmation of information. For most of the hub-genes, their relationships with BPD were easily understood and well-studied. However, when it comes to INS, most extracted papers are actually about insulin-like growth factor-I instead of INS. Only 2 true positive studies were found. One study mentioned that the serum INS is significantly increased in BPD infants after the administration of systemic corticosteroid treatment.<sup>[37]</sup> Postnatal application of glucocorticoids can prevent BPD in preterm infants. However, the treatment also has adverse effects such as hyperglycemia, hypertension, and intestinal perforation. Serum INS level is elevated as a result of hyperglycemia.<sup>[38]</sup> However, the role of INS in BPD is not clear. The other found that INS could increase cell function and NO production in normal fetal pulmonary artery endothelial cells (PAECs). PAECs from intrauterine growth restriction fetuses were less sensitive to INS which caused significantly decreased cell motility, growth, tube formation, and NO production. Impaired PAEC function may

contribute to the increased rate of BPD.<sup>[39]</sup> Furthermore, a related paper on Metformin reversing established lung fibrosis in a bleomycin model was published in 2008; it implicated INS in alleviating the later pathological change of BPD.<sup>[40]</sup> Usually, the limitation for text mining-based strategies is that there is no chance to discover new genes.<sup>[19]</sup> Interestingly, in this study we found INS as a novel gene related to BPD by accident. This could pave a foundation for further investigation about the molecular mechanisms between INS and BPD. As a matter of fact, our purpose of this work is to provide a foundation for identifying new molecules that are worthy perusing in the future, rather than summarizing previous works. Genes not in the central position of the network or do not have direct interactions with principal molecules could still be land for novel findings.

In conclusion, we systematically analyzed BPD-related genes using a text mining approach. These genes were further characterized by GO, pathway and network analysis. Our research provides a basis for a better understanding of the molecular mechanisms underlying BPD and paves a foundation for further studies as well.

### Acknowledgment

These authors wish to express their gratitude to the Shanghai Boyun Biotechnology Co. Ltd, for bioinformatics analysis.

### Author contributions

**Conceptualization:** Fei Shao, Jing Li.

**Data curation:** Weitao Zhou, Fei Shao.

**Formal analysis:** Weitao Zhou, Fei Shao.

**Funding acquisition:** Jing Li.

**Methodology:** Weitao Zhou, Fei Shao.

**Supervision:** Jing Li.

**Writing – original draft:** Weitao Zhou, Fei Shao.

**Writing – review and editing:** Weitao Zhou, Fei Shao.

Weitao Zhou orcid: 0000-0003-3618-5193.

### References

- [1] Poindexter BB, Feng R, Schmidt B, et al. Comparisons and limitations of current definitions of bronchopulmonary dysplasia for the prematurity and respiratory outcomes program. *Ann Am Thorac Soc* 2015;12:1822–30.
- [2] McEvoy CT, Jain L, Schmidt B, et al. Bronchopulmonary dysplasia: NHLBI workshop on the primary prevention of chronic lung diseases. *Ann Am Thorac Soc* 2014;11(Suppl 3):S146–53.
- [3] Nagiub M, Kanaan U, Simon D, et al. Risk factors for development of pulmonary hypertension in infants with bronchopulmonary dysplasia: systematic review and meta-analysis. *Paediatr Respir Rev* 2017;23:27–32.
- [4] Altit G, Dancea A, Renaud C, et al. Pathophysiology, screening and diagnosis of pulmonary hypertension in infants with bronchopulmonary dysplasia – a review of the literature. *Paediatr Respir Rev* 2017;23:16–26.
- [5] Reicherzer T, Haffner S, Shahzad T, et al. Activation of the NFkB pathway alters the phenotype of MSCs in the tracheal aspirates of preterm infants with severe BPD. *Am J Physiol Lung Cell Mol Physiol* 2018.
- [6] Wu D, Liang M, Dang H, et al. Hydrogen protects against hyperoxia-induced apoptosis in type II alveolar epithelial cells via activation of PI3K/Akt/Foxo3a signaling pathway. *Biochem Biophys Res Commun* 2018;495:1620–7.
- [7] McKenna S, Butler B, Jatana L, et al. Inhibition of IkappaBbeta/NFkappaB signaling prevents LPS-induced IL1beta expression without increasing apoptosis in the developing mouse lung. *Pediatr Res* 2017;82:1064–72.
- [8] Stouch AN, McCoy AM, Greer RM, et al. IL-1beta and inflammasome activity link inflammation to abnormal fetal airway development. *J Immunol* 2016;196:3411–20.
- [9] Spasic I, Livsey J, Keane JA, et al. Text mining of cancer-related information: review of current status and future directions. *Int J Med Inform* 2014;83:605–23.
- [10] Horcajadas JA, Pellicer A, Simon C. Wide genomic analysis of human endometrial receptivity: new times, new opportunities. *Hum Reprod Update* 2007;13:77–86.
- [11] Chun HW, Tsuruoka Y, Kim JD, et al. Automatic recognition of topic-classified relations between prostate cancer and genes using MEDLINE abstracts. *BMC Bioinformatics* 2006;7(Suppl 3):S4.
- [12] Deng X, Geng H, Bastola DR, et al. Link test—a statistical method for finding prostate cancer biomarkers. *Comput Biol Chem* 2006;30:425–33.
- [13] Rebholz-Schuhmann D, Oellrich A, Hoehndorf R. Text-mining solutions for biomedical research: enabling integrative biology. *Nat Rev Genet* 2012;13:829–39.
- [14] Natarajan J, Berrari D, Dubitzky W, et al. Text mining of full-text journal articles combined with gene expression analysis reveals a relationship between sphingosine-1-phosphate and invasiveness of a glioblastoma cell line. *BMC Bioinformatics* 2006;7:373.
- [15] Abuhusain HJ, Matin A, Qiao Q, et al. A metabolic shift favoring sphingosine 1-phosphate at the expense of ceramide controls glioblastoma angiogenesis. *J Biol Chem* 2013;288:37355–64.
- [16] Gao H, Zhang Z. Systematic analysis of endometrial cancer-associated hub proteins based on text mining. *BioMed Res Int* 2015;2015:615825.
- [17] Singhal A, Simmons M, Lu Z. Text mining genotype-phenotype relationships from biomedical literature for database curation and precision medicine. *PLoS Comput Biol* 2016;12:e1005017.
- [18] Kocbek S, Cavedon L, Martinez D, et al. Text mining electronic hospital records to automatically classify admissions against disease: measuring the impact of linking data sources. *J Biomed Inform* 2016;64:158–67.
- [19] Liu JL, Zhao M. Prioritization of susceptibility genes for ectopic pregnancy by gene network analysis. *Int J Mol Sci* 2016;17:E191.
- [20] Liu JL, Wang TS. Systematic analysis of the molecular mechanism underlying decidualization using a text mining approach. *PLoS One* 2015;10:e0134585.
- [21] Settles B. ABNER: an open source tool for automatically tagging genes, proteins and other entity names in text. *Bioinformatics* 2005;21:3191–2.
- [22] Smith L, Tanabe LK, Ando RJ, et al. Overview of BioCreative II gene mention recognition. *Genome Biol* 2008;9(Suppl 2):S2.
- [23] Morgan AA, Lu Z, Wang X, et al. Overview of BioCreative II gene normalization. *Genome Biol* 2008;9(Suppl 2):S3.
- [24] Bauer S, Robinson PN, Gagneur J. Model-based gene set analysis for bioconductor. *Bioinformatics* 2011;27:1882–3.
- [25] Dahlquist KD, Salomonis N, Vranizan K, et al. GenMAPP, a new tool for viewing and analyzing microarray data on biological pathways. *Nat Genet* 2002;31:19–20.
- [26] Ogata H, Goto S, Fujibuchi W, et al. Computation with the KEGG pathway database. *Biosystems* 1998;47:119–28.
- [27] Mewes HW, Albermann K, Heumann K, et al. MIPS: a database for protein sequences, homology data and yeast genome information. *Nucleic Acids Res* 1997;25:28–30.
- [28] Hooper SD, Bork P. Medusa: a simple tool for interaction graph analysis. *Bioinformatics* 2005;21:4432–3.
- [29] Bhatt AJ, Pryhuber GS, Huyck H, et al. Disrupted pulmonary vasculature and decreased vascular endothelial growth factor, Flt-1, and TIE-2 in human infants dying with bronchopulmonary dysplasia. *Am J Respir Crit Care Med* 2001;164:1971–80.
- [30] Lassus P, Turanlahti M, Heikkila P, et al. Pulmonary vascular endothelial growth factor and Flt-1 in fetuses, in acute and chronic lung disease, and in persistent pulmonary hypertension of the newborn. *Am J Respir Crit Care Med* 2001;164:1981–7.
- [31] Tang JR, Seedorf G, Balasubramaniam V, et al. Early inhaled nitric oxide treatment decreases apoptosis of endothelial cells in neonatal rat lungs after vascular endothelial growth factor inhibition. *Am J Physiol Lung Cell Mol Physiol* 2007;293:L1271–80.
- [32] Matsumura H, Ichiba H, Ohnishi S, et al. Histologic chorioamnionitis, amniotic fluid interleukin 6, krebs von den lungen 6, and transforming growth factor beta 1 for the development of neonatal bronchopulmonary dysplasia. *Jpn Clin Med* 2017;8:1179066017696076.
- [33] Taniguchi CM, Emanuelli B, Kahn CR. Critical nodes in signalling pathways: insights into insulin action. *Nat Rev Mol Cell Biol* 2006;7:85–96.
- [34] Alphonse RS, Vadivel A, Coltan L, et al. Activation of Akt protects alveoli from neonatal oxygen-induced lung injury. *Am J Respir Cell Mol Biol* 2011;44:146–54.

- [35] Xu D, Guthrie JR, Mabry S, et al. Mitochondrial aldehyde dehydrogenase attenuates hyperoxia-induced cell death through activation of ERK/MAPK and PI3K-Akt pathways in lung epithelial cells. *Am J Physiol Lung Cell Mol Physiol* 2006;291:L966–75.
- [36] Berggard T, Linse S, James P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* 2007;7:2833–42.
- [37] Ng PC, Lam CW, Lee CH, et al. Changes in serum leptin concentration after corticosteroid treatment in preterm infants. *Acta Paediatr* 2002;91:684–90.
- [38] Doyle LW, Cheong JL, Ehrenkranz RA, et al. Late (> 7 days) systemic postnatal corticosteroids for prevention of bronchopulmonary dysplasia in preterm infants. *Cochrane Database Syst Rev* 2017;10:CD001145.
- [39] Rozance PJ, Seedorf GJ, Brown A, et al. Intrauterine growth restriction decreases pulmonary alveolar and vessel growth and causes pulmonary artery endothelial cell dysfunction in vitro in fetal sheep. *Am J Physiol Lung Cell Mol Physiol* 2011;301:L860–71.
- [40] Rangarajan S, Bone NB, Zmijewska AA, et al. Metformin reverses established lung fibrosis in a bleomycin model. *Nat Med* 2018;24:1121–7.