Revealing hub pathway cross-talk for premature newborns with bronchopulmonary dysplasia by the integration of pathway analysis and Monte Carlo Cross-Validation

CHENGBIN WANG*, BIN ZHU*, MING CHEN, GAOYAN CHEN, MUZHEN XU and RUI PAN

Department of Pediatrics, Xiangyang Central Hospital, Hubei University of Arts and Science, Xiangyang, Hubei 441021, P.R. China

Received July 6, 2018; Accepted January 22, 2019

DOI: 10.3892/etm.2019.7257

Abstract. The objective of this study was to reveal hub pathway cross-talk for premature newborns with bronchopulmonary dysplasia (BPD) based on the pathway enrichment analysis and Monte Carlo Cross-Validation (MCCV) method. The inference of key pathway cross-talk consisted of four parts: i) Identifying differentially expressed genes (DEGs); ii) detecting differentially expressed pathways (DEPs); iii) computing discriminating score (DS) for each pair of DEPs or cross-talk and investigating seed cross-talk through the random forest (RF) algorithm and iv) extracting hub cross-talk dependent on the MCCV method. The results showed that a total of 132 DEGs and 137 DEPs were obtained across BPD patients and normal controls. Using the DS and RF algorithm, 10 seed DEP cross-talk were detected. By conducting the MCCV on seed cross-talk, 3 hub cross-talk for BPD were uncovered: i) The pair of pathways role of interleukin-17F (IL-17F) in allergic inflammatory airway diseases and role of IL-17A in psoriasis; ii) the pair of pathways role of IL random forest 17A in psoriasis and IL-17A signaling in fibroblasts and ii) the pair of pathways IL-17A signaling in airway cells and role of hypercytokinemia/hyperchemokinemia in the pathogenesis of influenza. These 3 hub cross-talk among DEPs might give an insight to reveal the molecular mechanism underlying the premature newborns with BPD.

Introduction

Bronchopulmonary dysplasia (BPD), one of the most common complications of prematurity, is a lung injury in preterm

Correspondence to: Dr Rui Pan, Department of Pediatrics, Xiangyang Central Hospital, Hubei University of Arts and Science, 136 Jingzhou Street, Xiangyang, Hubei 441021, P.R. China E-mail: ruipanhb@163.com

*Contributed equally

Key words: bronchopulmonary dysplasia, pathway, crosstalk, Monte Carlo Cross-Validation

infants resulting from oxygen and mechanical ventilation (1,2). Although progress has been made in neonatal care, the incidence of BPD is still increasing, mainly because the survival rate of younger and more premature babies has increased (3). Hence, effective prevention and treatments for BPD are on urgent need, and their necessary precondition is a comprehensive understanding of molecular and pathological mechanisms underlying this disease.

Previous literature has demonstrated that prenatal factors have important roles in the progression of BPD, such as surfactant deficiency and maternal infection (4,5). Recently, genetic variants were discovered in preterm infants of BPD (6,7). For instance, mutations of the lipid transporter ATP-binding cassette subfamily A member 3 (*ABCA3*), which is involved in surfactant synthesis in alveolar type II cells, have been revealed in BPD premature newborns (7). Moreover, these studies have highlighted a potential role for molecular pathways involved in inflammation and lung development in affected infants. Cross-talk among gene pathways can be means of regulatory interaction among different pathways or can express the gene overlap among pathways (8,9).

In this study, we proposed to reveal hub pathway cross-talk for premature newborns with BPD based on the integration of pathway related analyses and Monte Carlo Cross-Validation (MCCV) algorithm. Ultimately, hub cross-talk was extracted from seed cross-talk dependent on the MCCV method. The study of cross-talk involved in BPD could help in defining the molecular mechanisms leading to the onset and progression of the pathology (10).

Materials and methods

Data collection. Gene expression data with accession number E-GEOD-8586 (11) was collected from the ArrayExpress database for premature newborns at risk of BPD related studies. In E-GEOD-8586, there were 54 newborns with a gestational age from 23 to 27 weeks. These 54 newborns included 20 infants who later developed BPD (BPD group) and 34 infants who did not develop BPD (normal controls). There were 23 females and 31 males. Subsequently, in order to control the quality of the data, background corrections and standard normalizations were conducted. Identification of differentially expressed genes (DEGs). For investigating the differentially expressed levels of genes across BPD samples and normal controls, the empirical bayes (EB) approach was implemented to identify DEGs. In particular, the EB approach is equivalent to shrinkage of the estimated sample variances towards a pooled estimate, resulting in far more stable inference when the number of arrays is small (12). During this process, P-value for each gene was adjusted by Benjamini-Hochberg (BH) test (13). Importantly, only genes that met the thresholds of false discovery rate (FDR) <0.01 and $llog_2FoldChangel >2$ simultaneously were considered as DEGs in BPD and normal control groups.

Determination of differentially expressed pathways (DEPs). In the present study, all human biological pathways were derived from the Ingenuity Pathways Analysis (IPA) tool. Subsequently, DEGs of BPD were mapped to these biological pathways with an attempt to make them more correlated with the disease, and we gained pathways that were responsible for coordinating DEG activities, named as background pathways. Then we evaluated the significance between DEGs and genes in background pathways using Fisher's exact test (14). P-values were adjusted by BH test (13). In addition, the cut-off for DEPs across BPD patients and normal controls was P<0.01.

DEP cross-talk investigation. Based on the DEPs, we further focused on the relationships among any two of them. If a correlation existed between two DEPs, the DEP pair would be defined as a cross-talk for DEPs (10). Hence, we selected the discriminating score (DS) to access the strength of cross-talk among DEPs in the present study (10). Supposing that x and y were two DEPs in cross-talk, the DS was computed as following formula:

$$DS(x,y) = \frac{|u_x - u_y|}{\sigma_x + \sigma_y}$$

Seed cross-talk extraction. With an attempt to evaluate the functional activities of DEP cross-talk more deeply, a random forest (RF) classification model (15) was utilized to access the classified performance of these pairs between BPD and normal control groups. The method was comprised of three sections, first, drawing Ntree bootstrap samples dependent on the DS for each DEP cross-talk (N_{tree} = 500); secondly, growing a regression tree for each of the bootstrap samples; and finally, aggregating the predictions of the N_{tree} trees. The top 10 DEP cross-talk in descending order of AUC were defined as seed cross-talk.

MCCV for seed cross-talks. As mentioned above, seed cross-talk was identified, and then the MCCV method was adopted to validate the biological functions of seed cross-talk in BPD samples (16). The 54 samples were randomly split into two sets, the calibration set (Sc), contained nc samples for fitting the models; and the validation set (Sv), included nv samples for validating the model.

$$MCCV_{n_{\nu}} = \frac{1}{Rn_{\nu}} \sum_{k=1}^{R} ||E||^{2}$$

Of which E represents the squared prediction error, R represents the procedure repeated times (R=50). Theoretically, the fewer samples used in model calibration, the more repeat times were needed. For each bootstrap, the DEGs, DEPs, cross-talk and their DS values were carried out. The repeat times of DEP cross-talk were statistically calculated. More repeat times might imply more significance of this cross-talk.

Results

DEGs. In the present study, there were 20,514 genes in the pretreated E-GEOD-8586. After performing EB algorithm, a total of 132 DEGs which satisfied the cut-off (P<0.01 and $llog_2FoldChangel > 2$) between BPD premature newborns and normal controls were obtained. All DEGs were ranked in descending order of their P-values (Table I). The smaller the P-value, the more significant the expression of the DEG was. The most significant five DEGs across BPD group and normal controls were *EMILIN1* (P=4.29Ex10⁻⁶), *MFAP51* (P=8.77Ex10⁻⁶), *SNCA1* (P=9.23Ex10⁻⁶), interleukin-62 (*IL-62*) (P=3.14Ex10⁻⁵) and *STRADB1* (P=7.56Ex10⁻⁵).

DEPs. As described above, the IPA tool with 589 biological pathways (covering 5,169 genes) was implemented to execute pathway enrichment analysis for 132 DEGs of BPD infants, and the Fisher's exact test was used to evaluate significance of the pathways. Consequently, a total of 137 DEPs with P<0.01 were identified between BPD samples and normal controls, of which role of IL-17A in psoriasis (P=5.96Ex10⁻¹³) was the most significant one. Furthermore, DEPs with larger number of DEGs might be more important than that of the smaller ones. The DEPs with DEG number ≥ 4 are listed in Table II. Among these DEPs, flucocorticoid receptor signaling (P=7.11Ex10⁻⁷) contains 6 DEGs. Interestingly, the number of DEGs enriched in 6 DEPs was 5, including granulocyte adhesion and diapedesis (P=1.55Ex10⁻⁶), role of IL-17A in psoriasis (P=5.96Ex10⁻¹³), glucocorticoid receptor signaling (P=5.32Ex10⁻⁶), role of IL-17A in arthritis (P=2.24E-10-8), IL-17A signaling in airway cells (P=4.54Ex10⁻⁸) and flucocorticoid receptor signaling (P=4.70Ex10x10⁻⁵).

Seed cross-talk. In this section, first of all, DS was assigned to any pair of DEPs or cross-talk by comparing their gene expression levels. In consequence, 609 cross-talk among 137 DEPs were identified, and the DS distribution is illustrated in Fig. 1. Two of the cross-talk had higher DS than the others: Role of IL-17A in psoriasis; IL-17A signaling in fibroblasts (DS=6.328) and LXR/RXR activation; role of IL-17A in psoriasis (DS=6.134). Subsequently, we applied the RF classification of cross-talk for each sample based on the DS, and chose the AUC to evaluate the classification performance of each cross-talk. The cross-talk was ranked according to AUC in a descending order and the top 10 cross-talk were denoted as seed cross-talks (Table III). The best cross-talk was the pair of pathways acute phase response signaling and gepatic fibrosis/hepatic stellate cell activation (AUC=0.915). Besides, the AUC for the following two cross-talks were ranged from 0.8 to 0.9: The pair of pathways role of IL-17F in allergic inflammatory airway diseases and role of IL-17A in

Rank	Genes	Rank	Genes	Rank	Genes	Rank	Genes
1	EMILIN1	34	LINC00702	67	TRIM58	100	CXCL3
2	MFAP51	35	SLC25A391	68	BPGM	101	MORC4
3	SNCA1	36	STRADB2	69	SLC19A2	102	<i>C3</i>
4	IL-62	37	XK1	70	ZBTB16	103	SERPINE11
5	STRADB1	38	EPB422	71	SNCA	104	NPTN-IT1
6	CXCL12	39	AHSP1	72	DMTN	105	PTGS2
7	BPGM1	40	KRT13	73	EPB42	106	CXCL6
8	CXCL32	41	S100A91	74	PTX31	107	SGK1
9	TNFAIP32	42	SOD22	75	IL-61	108	CXCL8
10	SOD21	43	BOK	76	HEMGN	109	S100A8
11	EPB421	44	GYS1	77	KRT1	110	TNFAIP3
12	PTX32	45	CCL22	78	BLVRB	111	LIF
13	CCL21	46	S100A81	79	STRADB	112	CXCL1
14	HBM1	47	KRT141	80	FAM46C	113	FOSB1
15	CXCL62	48	MFAP52	81	UCP2	114	CCL20
16	KRT14	49	ZBTB163	82	SERPINE12	115	IL-6
17	PLEKHO1	50	CCL202	83	XK	116	SOD2
18	UCP21	51	CXCL33	84	AHSP	117	TNFAIP6
19	KAT2B1	52	HBM4	85	GCH1	118	PTX3
20	POC1B	53	CXCL63	86	CHRDL2	119	S100A9
21	ADIPOR1	54	PTX33	87	LIF1	120	CCL2
22	ID1	55	CXCL82	88	CXCL31	121	S100A12
23	HBQ1	56	CXCL13	89	TNFAIP61	122	MFAP5
24	FAM46C1	57	SQRDL	90	CCL201	123	EGR1
25	DCAF12	58	MAP3K8	91	SLC25A39	124	FOS
26	BPGM2	59	DCK	92	HBM	125	FOSB
27	C15orf48	60	SUSD6	93	CSF3	126	SERPINE1
28	PLA2G2A1	61	KAT2B	94	TNFAIP31	127	FOSB2
29	SERPINB2	62	GYPA	95	CXCL11	128	EGR11
30	ZBTB162	63	SGK11	96	CXCL61	129	FOS1
31	HBM3	64	KRT5	97	CXCL81	130	ZBTB161
32	HEMGN1	65	YOD1	98	PLA2G2A	131	HBM2
33	SNCA2	66	PIP4K2A	99	SELE	132	RPS4Y1

Table I. Differentially expressed genes (DEGs).

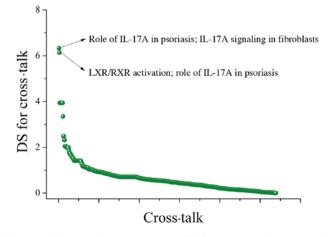


Figure 1. DS was assigned to any pair of DEPs or cross-talk by comparing their gene expression levels. In consequence, 609 cross-talk among 137 DEPs were identified, and the DS distribution are illustrated. DEP, differentially expressed pathway; DS, computing discriminating score.

psoriasis (AUC=0.857), and the pair of pathways LXR/RXR activation and hepatic fibrosis/hepatic stellate cell activation (AUC=0.812).

Hub cross-talk. Herein, we divided 20 BPD samples and 30 normal controls of total 54 samples in the gene expression data into two sets according to the ratio of 6:4. Among the 10 seed cross-talk, we uncovered that 3 cross-talk were repeated many times. The pair of pathways role of IL-17F in allergic inflammatory airway diseases and role of IL-17A in psoriasis was repeated 29 times. The pair of pathways role of IL-17A in psoriasis and IL-17A signaling in fibroblasts was repeated 28 times. The pair of pathways IL-17A signaling in airway cells and role of hypercytokinemia/hyperchemokinemia in the pathogenesis of influenza was repeated 28 times. Hence we defined these three cross-talk as hub cross-talk for BPD patients, which might play more crucial roles in the progression of BPD samples than the others.

DEPs	P-value	Number of genes in pathway	Number of DEGs
Glucocorticoid receptor signaling	7.11x10 ⁻⁷	255	6
Granulocyte adhesion and diapedesis	1.55x10 ⁻⁶	163	5
Role of IL-17A in psoriasis	5.96x10 ⁻¹³	13	5
Glucocorticoid receptor signaling	5.32x10 ⁻⁶	255	5
Role of IL-17A in arthritis	2.24x10 ⁻⁸	54	5
IL-17A signaling in airway cells	4.54x10 ⁻⁸	63	5
Glucocorticoid receptor signaling	4.70x10 ⁻⁵	255	5
Role of IL-17A in arthritis	6.74x10 ⁻⁷	54	4
IL-17A signaling in airway cells	1.18x10 ⁻⁶	63	4
Agranulocyte adhesion and diapedesis	5.89x10 ⁻⁵	173	4
Role of IL-17A in arthritis	2.72x10 ⁻⁷	54	4
IL-17A signaling in airway cells	4.78x10 ⁻⁷	63	4
Glucocorticoid receptor signaling	2.25x10 ⁻⁵	255	4
Atherosclerosis signaling	6.60x10 ⁻⁸	119	4
Role of IL-17A in psoriasis	2.63x10-9	13	4
Granulocyte adhesion and diapedesis	$1.07 \mathrm{x} 10^{-4}$	163	4
Agranulocyte adhesion and diapedesis	1.36x10 ⁻⁴	173	4

Table II. Differentially expressed pathways (DEPs) with DEG number ≥ 4 .

Table III. Seed cross-talk with AUC value for random forest (RF) classification.

ID	Cross-talk	AUC
1	Acute phase response signaling; hepatic fibrosis/hepatic stellate cell activation	0.915
2	Role of IL-17F in allergic inflammatory airway diseases; role of IL-17A in psoriasis	0.857
3	LXR/RXR activation; hepatic fibrosis/hepatic stellate cell activation	0.812
4	LXR/RXR activation; acute phase response signaling	0.799
5	Role of IL-17A in psoriasis; IL-17A signaling in fibroblasts	0.790
6	IL-17A signaling in airway cells; role of hypercytokinemia/hyperchemokinemia in the pathogenesis of influenza	0.786
7	IL-17 signaling; differential regulation of cytokine production in intestinal epithelial cells by IL-17A and IL-17F	0.786
8	Role of IL-17F in allergic inflammatory airway diseases; atherosclerosis signaling	0.782
9	LXR/RXR activation; glucocorticoid receptor signaling	0.782
10	IL-17A signaling in gastric cells; role of hypercytokinemia/hyperchemokinemia	
	in the pathogenesis of influenza	0.778

Discussion

In the present study, we investigated hub cross-talk of DEPs for BPD samples by integrating a series of bioinformatics analyses, identification of DEGs, extraction of DEPs, exploration of seed cross-talk and investigation of hub cross-talk. Although the scenario might be intuitive, the biological functions and activities of DEPs have not been well studies (17,18). In order to increase the confidence and feasibility of our study, we implemented the MCCV to screen the hub cross-talk (10,19).

Consequently, a total of 3 hub cross-talk for premature newborns with BPD were gained. IL-17F is a cytokine that shares sequence similarity with IL-17 and is expressed by activated

T cells (20). It inhibits angiogenesis of endothelial cells and induces endothelial cells to produce IL-2, TGFB1/TGFB, and signaling and function in inflammation (21). The involvement of IL-17A/IL-17F in inflammatory immune responses and host defense mechanisms, and the correlation between IL-17A/IL-17F with other cytokines has been reviewed previously (22,23). The review indicated that IL-17A and IL-17F possibly work together in human disease. Coincidently, in our results, the hub cross-talk (role of IL-17F in allergic inflammatory airway diseases and role of IL-17A in psoriasis) also revealed this phenomenon in BPD development.

IL-17A has been demonstrated to take vital part in allergic airway inflammation, and promote inflammation via inducing various proinflammatory cytokines and chemokines, recruiting neutrophils, increasing antibody production, and activating T cells (24). We found that IL-17A was correlated to all of the three hub cross-talk of BPD infants, including the pair of pathways role of IL-17F in allergic inflammatory airway diseases and role of IL-17A in psoriasis; the pair of pathways role of IL-17A in psoriasis and IL-17A signaling in fibroblasts; the pair of pathways IL-17A signaling in airway cells and role of hypercytokinemia/hyperchemokinemia in the pathogenesis of influenza. Importantly, IL-17A has been reported to exhibit a vital role in the development of pulmonary fibrosis, suggesting it is a potential treatment target for inflammation-induced fibrosis (25). Taken together, we inferred that the three hub cross-talk closely correlated to BPD.

In summary, 3 hub cross-talk among DEPs were identified in our study, which might give insight into revealing the molecular mechanism underlying the premature newborns with BPD.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets used in this study are available from the corresponding author on reasonable request.

Authors' contributions

CW, RP and BZ conceived the study. MC and GC acquired, and analyzed the data. CW wrote the manuscript. RP, GC and MX explained the results and revised the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- 1. Bancalari E and Claure N: Bronchopulmonary dysplasia: Definitions and epidemiology. In: Bronchopulmonary Dysplasia. Bhandari V (ed). Springer International Publishing, Switzerland,
- pp167-182, 2016. Smith VC, Zupancic JA, McCormick MC, Croen LA, Greene J, Escobar GJ and Richardson DK: Rehospitalization in the first year of life among infants with bronchopulmonary dysplasia. J Pediatr 144: 799-803, 2004.
- 3. Bhandari A and Panitch H: An update on the post-NICU discharge management of bronchopulmonary dysplasia. Semin Perinatol 42: 471-477, 2018.

- 4. Watterberg KL, Demers LM, Scott SM and Murphy S: Chorioamnionitis and early lung inflammation in infants in whom bronchopulmonary dysplasia develops. Pediatrics 97: 210-215, 1996.
- 5. Glaser K and Speer CP: Pre and postnatal inflammation in the pathogenesis of bronchopulmonary dysplasia. In: Bronchopulmonary Dysplasia. Bhandari V (ed). Springer International Publishing, Switzerland, pp55-77, 2016.
- 6. Gadhia MM, Cutter GR, Abman SH and Kinsella JP: Effects of early inhaled nitric oxide therapy and vitamin A supplementation on the risk for bronchopulmonary dysplasia in premature newborns with respiratory failure. J Pediatr 164: 744-748, 2014.
- 7. Lavoie PM: Genetics of bronchopulmonary dysplasia. In: Bronchopulmonary Dysplasia. Bhandari V (ed). Springer International Publishing, Switzerland, pp109-127, 2016.
- 8. Aksamitiene E, Kiyatkin AB and Kholodenko BN: Cross-talk between mitogenic Ras/MAPK and survival PI3K/Akt pathways: A fine balance. Biochem Soc Trans 40: 139-146, 2012.
- 9. Bernards R: A missing link in genotype-directed cancer therapy. Cell 151: 465-468, 2012.
- 10. Colaprico A, Cava C, Bertoli G, Bontempi G and Castiglioni I: Integrative analysis with Monte Carlo cross-validation reveals miRNAs regulating pathways cross-talk in aggressive breast cancer. Biomed Res Int 2015: 831314, 2015.
- 11. Cohen J, Van Marter LJ, Sun Y, Allred E, Leviton A and Kohane IS: Perturbation of gene expression of the chromatin remodeling pathway in premature newborns at risk for bronchopulmonary dysplasia. Genome Biol 8: R210, 2007.
- 12. Smyth GK: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3: Article3, 2004.
- 13. Bogdan M, Ghosh JK and Tokdar ST: A comparison of the Benjamini-Hochberg procedure with some Bayesian rules for multiple testing. In: Beyond Parametrics in Interdisciplinary Research: Festschrift in Honor of Professor Pranab K. Sen. Vol 1. Institute of Mathematical Statistics, Beachwood, OH, pp211-230, 2008
- 14. Routledge R: Fisher's exact test. In: Encyclopedia of Biostatistics. John Wiley Publishing, New York, NY, 2005.
- 15. Liaw A and Wiener M: Classification and regression by random-Forest. R News 2: 18-22, 2002.
- 16. Xu QS and Liang YZ: Monte Carlo cross validation. Chemom Intell Lab Syst 56: 1-11, 2001.
 17. Donato M, Xu Z, Tomoiaga A, Granneman JG, Mackenzie RG,
- Bao R, Than NG, Westfall PH, Romero R and Draghici S: Analysis and correction of cross-talk effects in pathway analysis. Genome Res 23: 1885-1893, 2013.
- 18. Yang H, Cheng C and Zhang W: Average rank-based score to measure deregulation of molecular pathway gene sets. PLoS One 6: e27579, 2011.
- 19. Xu QS, Liang YZ and Du YP: Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. J Chemometr 18: 112-120, 2004.
- 20. Besnard AG, Sabat R, Dumoutier L, Renauld JC, Willart M, Lambrecht B, Teixeira MM, Charron S, Fick L, Erard F, et al: Dual role of IL-22 in allergic airway inflammation and its cross-talk with IL-17A. Am J Respir Crit Care Med 183: 1153-1163, 2011. 21. Chang SH and Dong C: IL-17F: Regulation, signaling and
- function in inflammation. Cytokine 46: 7-11, 2009.
- 22. Jin W and Dong C: IL-17 cytokines in immunity and inflammation. Emerg Microbes Infect 2: e60, 2013.
- 23. Puel A, Döffinger R, Natividad A, Chrabieh M, Barcenas-Morales G, Picard C, Cobat A, Ouachée-Chardin M, Toulon A, Bustamante J, et al: Autoantibodies against IL-17A, IL-17F, and IL-22 in patients with chronic mucocutaneous candidiasis and autoimmune polyendocrine syndrome type I. J Exp Med 207: 291-297, 2010.
- 24. Iwakura Y, Nakae S, Saijo S and Ishigame H: The roles of IL-17A in inflammatory immune responses and host defense against pathogens. Immunol Rev 226: 57-79, 2008.
- 25. Wilson MS, Madala SK, Ramalingam TR, Gochuico BR, Rosas IO, Cheever AW and Wynn TA: Bleomycin and IL-1β-mediated pulmonary fibrosis is IL-17A dependent. J Exp Med 207: 535-552, 2010.



This work is licensed under a Creative Commons International (CC BY-NC-ND 4.0) License.