Rare-variant collapsing analyses identified risk genes for neonatal acute respiratory distress syndrome

Huiyao Chen [a,1], Xiang Chen [b,1], Liyuan Hu [b], Chang Ye [b], Jiantao Zhang [b], Guoqiang Cheng [b], Lin Yang [a,d], Yulan Lu [c], Xinran Dong [c,*], Wenhao Zhou [a,b,c,*]

[a] Center for Molecular Medicine of Children's Hospital of Fudan University, Institutes of Biomedical Sciences, Fudan University, Shanghai, China
[b] Departments of Neonatology, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai 201102, China
[c] Center for Molecular Medicine, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai 201102, China
[d] Department of Pediatric Endocrinology and Inherited Metabolic Diseases, Children's Hospital of Fudan University, National Children's Medical Center, Shanghai 201102, China

ARTICLE INFO

ABSTRACT

Background: Acute respiratory distress syndrome (ARDS) could account for a considerable proportion of neonatal death, while the genetic etiology and pathophysiology of neonatal ARDS remain elusive. In this case-control study, 515 neonates were enrolled in the China Neonatal Genomes Project (CNGP, NCT03931707) from August 2016 to June 2021, including 196 ARDS and 319 non-ARDS matched by sex, gestational age, birth weight, perinatal asphyxia, pneumonia, sepsis, and necrotizing enterocolitis. Clinical exome sequencing was used to detect genetic variants. Collapsing analyses together with permutation tests were used to identify ARDS risk genes enriched for rare variants in ARDS samples. In silico functional interaction analysis and expression pattern studies at different stages of lung development were used to investigate the biological functions of the risk genes.
Results: Collapsing analyses identified that rare variants were significantly abundant in the genes associated with the precursor of the lamellar body and there were eight predicted risk genes with strong confidence ($P < 0.01$). Among them, the expression of *EDNRB* increased significantly in lung development and was up-regulated in ARDS ($P < 0.05$). In addition, 151 predicted transcriptional target proteins of *EDNRB* were highly enriched in the lamellar body responsible for pulmonary surfactant storage and secretion.
Conclusions: In our study, the genes associated with pulmonary surfactant storage and release were highly enriched with rare variants. A novel neonatal ARDS risk gene *EDNRB* may be a key gene for neonatal lung development and pulmonary surfactant homeostasis. Additional validation in independent patient populations and further exploration of underlying molecular mechanisms are warranted.

1. Introduction

Neonatal acute respiratory distress syndrome (ARDS), which has a prevalence of 1.5 % and overall mortality of 17 %–24 % in the neonatal intensive care unit [1], is characterized by extensive pulmonary inflammation and surfactant catabolism leading to pulmonary dysfunction [2]. It is considered to be triggered directly or indirectly by a range of etiologies [3], such as meconium aspiration syndrome, perinatal asphyxia, necrotizing enterocolitis, sepsis, and infectious pneumonia [4,5]. Surfactant is liquid-protein complexes that line the alveolar air–liquid interface, contributing to alveolar patency and effective gas exchange [6]. However, the lungs and immune systems of neonates are still rapidly developing, and the pathophysiology of respiratory insufficiency and the role of surfactant homeostasis remain unclear.

In addition to clinical risk factors, genetic predisposition has recently been identified as a potential etiology of respiratory distress in neonates [7,8]. In the Online Mendelian Inheritance in Man database, several genes have been reported to be responsible for surfactant catabolism and possibly cause clinical features of respiratory distress syndrome in term neonates, including *SFTPB*, *SFTPC*, and *ABCA3*. Moreover, transcription factor *NKX2-1* was also responsible for the onset of respiratory distress as it participates in the expression, catabolism [9], and function [10] of surfactant

* Corresponding authors at: Center for Molecular Medicine of Children's Hospital of Fudan University, Institutes of Biomedical Sciences, Fudan University, Shanghai, China (W. Zhou).
E-mail addresses: xrdong@fudan.edu.cn (X. Dong), zhouwenhao@fudan.edu.cn (W. Zhou).
[1] These authors contributed equally to the manuscript.

by regulating the expression of the above genes [11]. Pathogenic variants of *NKX2-1* can lead to choreoathetosis, hypothyroidism, and neonatal respiratory diseases [MIM:610978]. Researchers have found that single nucleotide polymorphisms having a higher allele frequency (AF) in the neonates with respiratory distress than in the control groups, such as *SFTPD* rs1923537 [12], *LPCAT1* rs9728 [13], and *NOS3* Glu298Asp [14], may also be another essential type of genetic risk factor for respiratory distress in neonates. However, only a few neonatal ARDS cases could be attributed to the genetic cause mentioned above and no studies have investigated their genetic risk of susceptibility.

To investigate potential genetic risk factors in extreme respiratory phenotypes, genome-wide association studies have been widely used for the common variants with a minor allele frequency (MAF) > 5 % [15–17]. Recently, empirical evidence suggested that rare variants (MAF < 1 %) may be a novel potential genetic risk factor associated with complex diseases [18]. Genes related to the onset of complex disease or trait variation could be identified by applying rare-variant collapsing analyses [19].

The variants of risk genes can be valuable markers for mapping different pathologies, allowing neonates to benefit from early diagnosis. We hypothesized that a high burden of rare variants may lead to a greater susceptibility of neonates to ARDS. Therefore, we conducted a case-control study to investigate the potential genetic etiologies in neonatal ARDS using rare-variant collapsing analyses and attempted to validate their biological plausibility through expression studies at different stages of lung development as well as in silico functional interaction analysis.

2. Methods

2.1. Study design and participants

This case-control study is based on the retrospective analysis of single-center data from the China Neonatal Genomes Project (CNGP, NCT03931707). In this project, clinical exome sequencing was performed on patients who have the suspicion of genetic diseases and underwent consultation with geneticists. We included 1,019 Han Chinese neonates admitted within 24h of birth at the Children's Hospital of Fudan University from August 2016 to June 2021. The diagnosis criteria of neonatal ARDS refer to the definition of Montreux [2] and the following points need to be satisfied: 1) acute onset (ie, within one week) from a known or suspected clinical insult but respiratory distress syndrome, transient tachypnoea of the neonate, or congenital anomalies is not a primary current acute respiratory condition; 2) diffuse, bilateral, and irregular opacities or infiltrates, or complete opacification on chest radiographs, which are not fully explained by local effusions, atelectasis, respiratory distress syndrome, transient tachypnoea of the neonate, or congenital anomalies; 3) absence of congenital heart disease explaining the pulmonary edema (this includes ductus arteriosus with pulmonary overflow if no acute pulmonary haemorrhage exists). Echocardiography is needed to verify the origin of edema; 4) oxygenation deficit expressed as oxygenation index (FiO_2*Paw*100/PaO_2) $\geq$ 4. Hospitalized neonates who did not diagnose with ARDS and without any type of respiratory support were classified as non-ARDS. Neonates with known genetic syndromes or chromosomopathies were excluded from the study. The neonatal ARDS diagnoses were evaluated by two neonatologists back-to-back.

Patient information was acquired from electronic medical records. The samples used in this study were collected with appropriate informed consent and the approval of the ethics committee of the Children's Hospital of Fudan University (2016–235).

2.2. Gene-based collapsing analyses

This study focused on predicted rare damaging variants (including potentially damaging variants) of 3,203 genes captured by the Agilent ClearSeq Inherited Disease Kit, which included 2,742 confirmed disease-causing genes [20]. The process of gene testing and damaging variant filtering are described in the eMethods in the Supplement. We performed the following steps to prepare the rare variant collapsing analyses. First, each variant was annotated to obtain its mutation type, which was further grouped into four types, including protein-truncating variants (PTVs), missense or non-synonymous variants (MISs), synonymous variants, and non-coding variants, according to the transfer table shown in eTable 1 in the Supplement. Next, the number of variants in each mutation type was calculated and summarized for each gene, and the genes with at least three PTV or MIS variants in the matched samples were included in the analysis. Fisher's exact test was used to test whether the number of variants from each mutation type was significantly higher in the ARDS group than in the non-ARDS group. An empirical (permutation-based) expected probability distribution was used for quantile–quantile plots (eMethods in the Supplement) to estimate risk genes with more robust statistics. We will consider the gene to be a risk gene if its observed *P* value is significantly smaller than the expected *P* value. In particular, tests of PTV and MIS variants were used to identify ARDS-related burden genes. Synonymous variants and non-coding variants were treated as the near-neutral background, and genes with significant differences found at the synonymous or non-coding level were filtered out.

2.3. Gene expression profiles and in silico functional interaction analyses

We downloaded the RNA transcriptome data of mouse and human samples from the GEO database. The mouse dataset GSE74243 contains lung tissue samples at 26 time points across all five canonical stages of lung development. The integrated human dataset (GSE14334 and GSE43767) included lung tissue samples from 6 to 24 postovulatory weeks (PWs), which is the initial branching period of airways. GSE14334 contains 38 samples from 53 to 154 days (6–24 weeks) post-conception, representing 29 distinct time points. GSE43767 contains samples from 6 to 8 PW (n = 10) and 16 to 24 PW (n = 9). The pediatric dataset GSE152980 contains 48 ARDS and 25 non-ARDS nasal cytology brushings. Next, we normalized the gene expression profiles of mice and humans to visualize the changes across lung development and analyze the expression differences between ARDS and non-ARDS in pediatrics. Finally, we constructed a lung-specific functional interaction network in silico study using the gene expression profile of human lung tissue (eMethods in the Supplement).

2.4. Statistical analysis

The propensity score matching strategy was used to screen the original collected ARDS cases and the non-ARDS controls using the R package MatchIt. Continuous variables were evaluated by the Kruskal–Wallis H or T-tests. Categorical variables were tested by a Chi-Squared and Fisher Exact Test. Gene expression correlation was performed using the Pearson correlation coefficient. The threshold of significance was *P* < 0.05. All analyses were performed using R software (version 4.0.3, http://cran.r-project.org).

3. Results

3.1. Study population

Altogether, we recruited a total of 1,019 neonates (eFigure 1 in the Supplement), including 374 ARDS and 645 non-ARDS. Their demographic and clinical characteristics were described in eTable 2. To investigate the genetic risk factors for ARDS, we performed propensity score analysis to establish a matching cohort. The matching features included basic information (sex, gestational age, and birth weight) and ARDS-related clinical manifestations (perinatal asphyxia, pneumonia, sepsis, and necrotizing enterocolitis). Based on propensity score matching, we established a cohort of 319 non-ARDS and 196 ARDS neonates (Fig. 1). No significant difference in clinical features was detected between the two groups (Table 1). All of these neonates underwent clinical exome sequencing, and no diagnosable genetic findings were detected in either the non-ARDS or ARDS group after curation by two geneticists in a blinded fashion.

3.2. Detection of ARDS burden gene sets and genes

In ARDS cases versus non-ARDS controls, we found that both PTV and MIS variants were significantly enriched in the set of genes associated with the late endosome (OR_{PTV} = 1.85, P = 0.03; OR_{MIS} = 1.22, P = 0.01; Fig. 2a), which is the precursor of the lamellar body [21]. In addition, we identified 55 candidate burden genes with more PTV or MIS variants in the ARDS group compared with the non-ARDS group (eTable 3 in the Supplement) from 171 and 2,050 genes that were included in the PTV and MIS variant collapsing analyses. Five of the genes (*MUTYH*, *ATXN3*, *HTT*, *TBP*, and *DCHS2*) had a higher PTV allele frequency, and 50 genes had a higher MIS allele frequency (P < 0.05) in the ARDS group. There was no overlap between these two types of candidate burden genes.

The permutation-based expected probability distribution helped us identify the significant reference thresholds (P values) of the statistics. For 55 candidate burden genes, P < 0.01 indicates a strong level of confidence, and 0.01 < P < 0.05 indicates a moderate level of confidence (eFigure 2a-b in the Supplement). The PTV burden genes all exhibited moderate confidence levels. Eight of the MIS burden genes (*LDLR*, *FANCA*, *NHS*, *SHROOM4*, *EDNRB*, *COL9A1*, and *CNGB3*, and *ATP6V0A1*) exhibited strong confidence levels (Table 2, eFigure 2b in the Supplement), in which the four variants (*LDLR* rs13306511, *FANCA* rs56369086, *SHROOM4* rs189694750, and *EDNRB* rs2070591) with high allele frequencies (AFs > 5/515) in CNGP did not show genetic heterogeneity (P > 0.05, eFigure 3a-d). Finally, we considered the eight MIS burden genes with strong confidence as strong confidence risk genes (SCGs).

3.3. Expression signature of strong confidence risk genes

To detect the potential biological function of these SCGs, we detected their expression level in human and mouse lung tissue. Human lung development at the early embryonic stage can be divided into two periods: the pseudoglandular stage (branching morphogenesis, 6–16 weeks) and the canalicular stage (formation of epithelial sacs and appearance of capillaries, 16–24 weeks) [22]. We found that the expression levels of three SCGs (*EDNRB*, *SHROOM4*, and *CNGB3*) significantly increased from the pseudoglandular stage to the canalicular stage (P < 0.05, Fig. 2b), and the expression levels of the other five genes (*LDLR*, *FANCA*, *NHS*, *COL9A1*, and *ATP6V0A2*) progressively decreased (P < 0.01, Fig. 2b). Among them, *EDNRB* had the highest relative expression compared with other SCGs (Fig. 2b). Similar results were shown in the transcriptional profiles of mice at two identical periods of lung development and *Ednrb* gene expression increased rapidly from embryonic to saccular stages in the prenatal mice and peaked at birth (Fig. 2c). In addition, *EDNRB* expression was significantly upregulated in patients with pediatric ARDS compared to the non-ARDS group (P = 0.035, Fig. 2d). This finding demonstrates that *EDNRB* has a more important role in the whole process of lung development and in the disease onset.
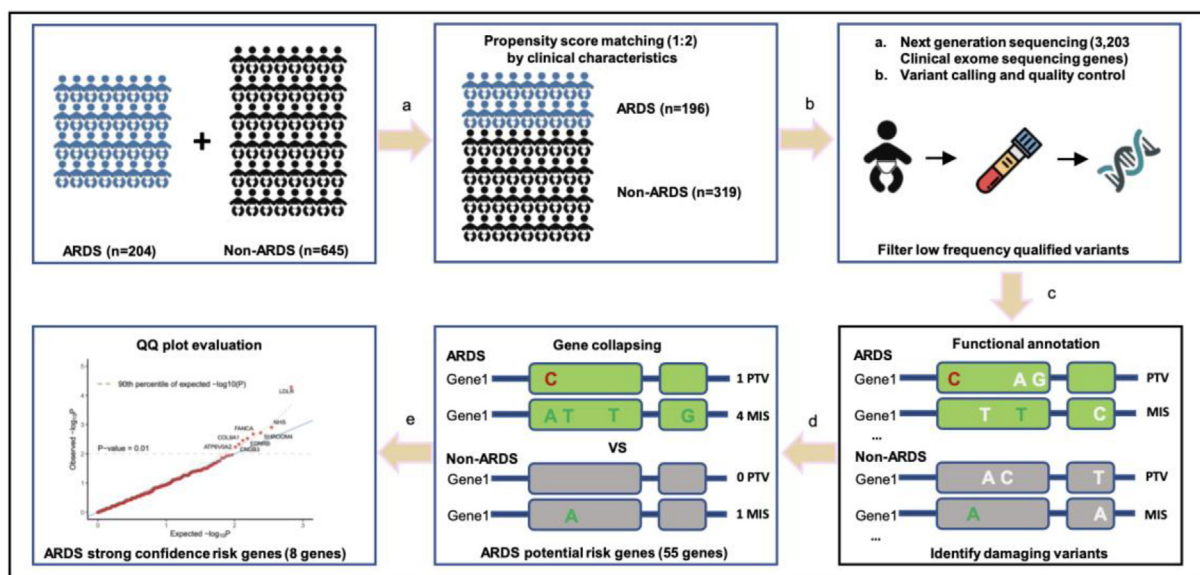

Fig. 1. Study design. Based on a case-control study with matched clinical information, we searched for strong confidence risk genes of neonatal ARDS through clinical exome sequencing and rare variant collapsing analyses. The specific process is divided into five steps. First, 515 samples were sequenced by clinical exome sequencing after propensity score matching by clinical characteristics (step a). Variant calling and quality control were performed by GATK and an in-home pipeline (step b). Ensembl Variant Effect Predictor (VEP) and ANNOVAR were used to annotate all variants, including protein-truncating variants (PTVs) and missense or non-synonymous variants (MISs) (step c). Rare-variant collapsing analyses were applied, and 55 ARDS potential risk genes were identified (P value < 0.05) (step d). Next, the results were evaluated using a quantile–quantile (QQ) plot (step e). Finally, we focused on eight strong confidence risk genes.

Table 1
Clinical information of both ARDS and non-ARDS groups.

| Clinical characteristics	All samples			Matched samples		
	ARDS	Non-ARDS	*P* value	ARDS	Non-ARDS	*P* value
No. (%) with data	374	645		196	319	
Sex, Male	126 (33.7)	381 (59.1)	<0.001	107 (54.6)	186 (58.3)	0.462
Gestational age, mean (SD), weeks	34.10 (2.73)	36.53 (1.94)	<0.001	35.92 (1.94)	36.03 (1.88)	0.501
Birth weight, mean (SD), kg	2.40 (0.66)	2.81 (0.67)	<0.001	2.71 (0.63)	2.63 (0.74)	0.220
Perinatal asphyxia	90 (24.1)	68 (10.5)	<0.001	33 (16.8)	42 (13.2)	0.309
Pneumonia	76 (20.3)	29 (4.5)	<0.001	24 (12.2)	28 (8.8)	0.264
Sepsis	135 (36.1)	136 (21.1)	<0.001	69 (35.2)	91 (28.5)	0.136
Necrotizing enterocolitis	33 (8.8)	24 (3.7)	0.001	5 (2.6)	8 (2.8)	1.000
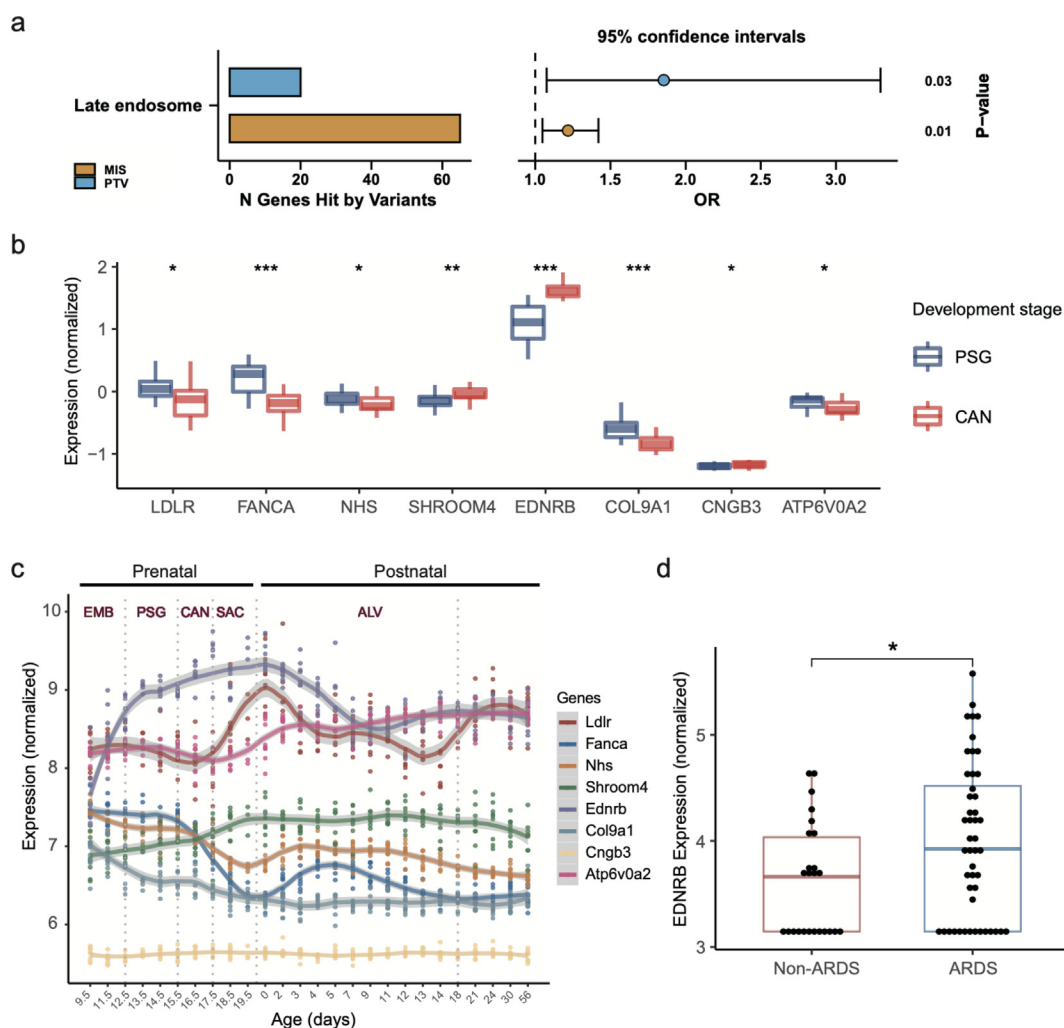
Fig. 2. Evaluation of the gene sets and strong confidence risk genes (SCGs) for ARDS. (a) The significantly enriched gene sets in rare variants. "N Genes Hit by Variants" refer to the number of genes in the set and the number of genes that had at least one rare variant hit, respectively. CI, confidence interval; OR, odds ratio. (b) Differential expression of eight SCGs in the pseudoglandular and canalicular stages of human lung development. (c) Expression patterns of eight SCGs in the five canonical stages of mouse lung development (EMB, embryonic; PSG, pseudoglandular; CAN, canalicular; SAC, saccular; ALV, alveolar). (d) Differential expression of *EDNRB* in the nasal cytology brushings of pediatric ARDS and non-ARDS. Expression levels were normalized for processing. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

3.4. Functional effects of variation on the in silico functional interaction network

In the in silico functional interaction network, *EDNRB* was estimated to positively regulate 101 target genes (Fig. 3a) and negatively regulate 50 target genes (Fig. 3b) at the transcriptional level. In addition, we found that *EDNRB* may positively regulate *SFTPB* and *SFTPC* expression (Fig. 3a), known as surfactant proteins. Moreover, *EDNRB* was highly correlated with the expression of *SFTPB* and *SFTPC* (R = 0.9, *P* < 2.2e-16, Fig. 3c). The enrichment analyses of the 151 target genes found that the lamellar body, the surfactant storage organelle in type II epithelial cells, was the top item

Table 2

The number of variants in eight SCGs and the results of rare-variant collapsing analyses.

Gene	ARDS PTV	ARDS MIS	Non-ARDS PTV	Non-ARDS MIS	PTV		MIS	
					P	**adjP**	**P**	**adjP**
LDLR	0	14	0	2	–	–	5.21E-05***	0.107
NHS	0	13	0	4	–	–	1.25E-03**	1
SHROOM4	0	20	0	11	–	–	1.91E-03**	1
FANCA	2	29	2	21	4.91E-01	1	2.11E-03**	1
EDNRB	0	13	2	5	–	–	2.97E-03**	1
COL9A1	1	16	1	8	–	–	3.49E-03**	1
CNGB3	0	10	0	3	–	–	4.66E-03**	1
ATP6V0A2	0	7	1	1	–	–	5.83E-03**	1

SCGs, strong confidence risk genes; *P* value is the significance of the Fisher's exact test (one-sided) and adj*P* is the adjusted p value with Bonferroni correction; * *P* < 0.05; ** *P* < 0.01; *** *P* < 0.001.
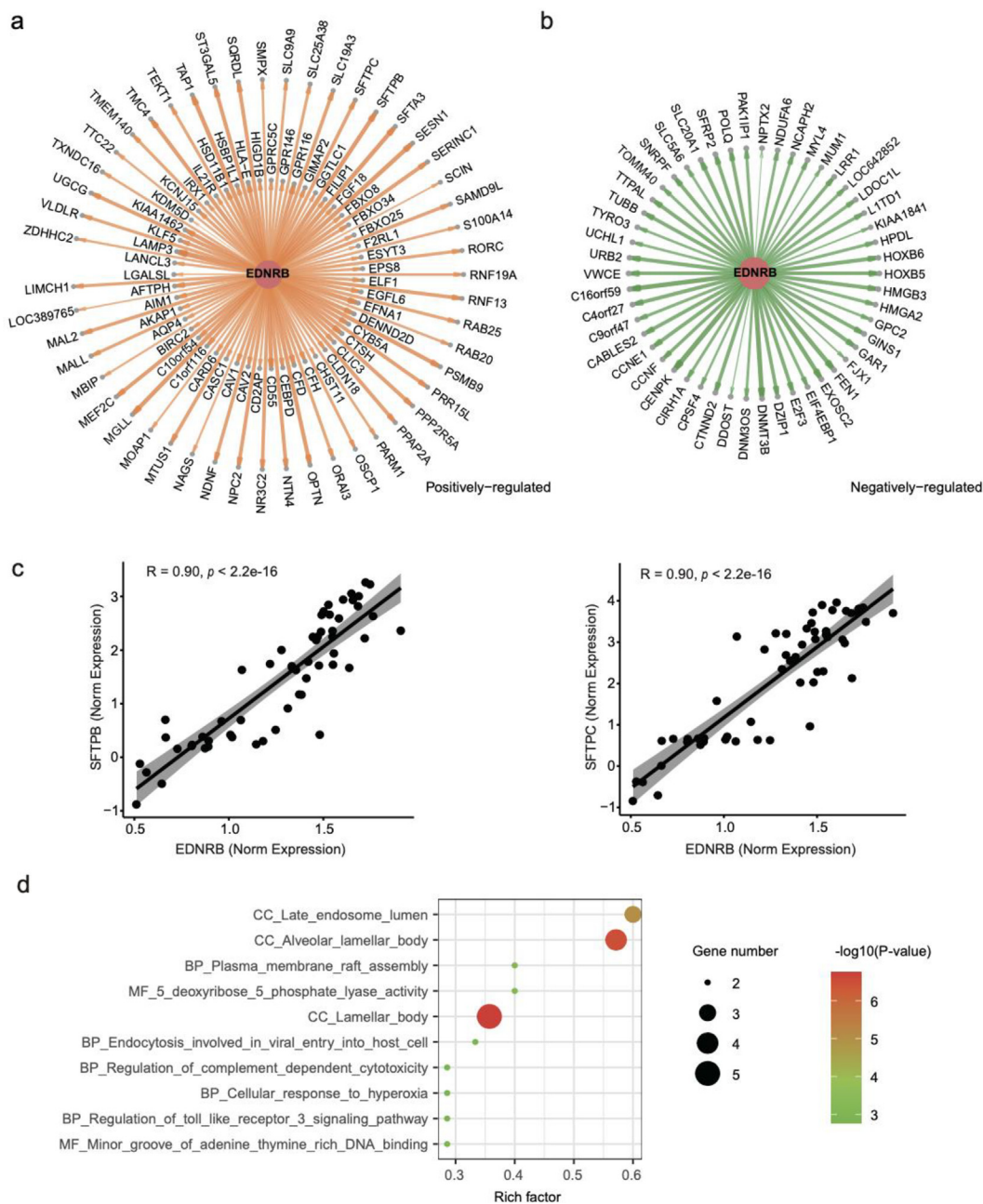

Fig. 3. EDNRB targets genes and their functions. (a-b) The 101 positively regulated genes (a) and 50 negatively regulated genes (b) of *EDNRB* were predicted by in silico functional interaction analysis. The thickness of the line represents the Spearman correlation. (c) Correlation of *EDNRB* with *SFTPB* and *SFTPC* expression in two stages of human lung development. (d) The top ten items of Gene Ontology enrichment analyses of the 151 *EDNRB* target genes were sorted by *P* value. The rich factor is the number of genes belonging to this term out of 151 *NDNRB* target genes divided by the number of genes belonging to this term out of 16,648 genes included in the human lung expression profile data.

in Gene Ontology analyses (P = 1.74e-7, Fig. 3d). These results may indicate the potential important function of *EDNRB* in the onset of ARDS.

4. Discussion

Neonatal ARDS is a secondary deficiency of pulmonary surfactant induced by multiple inflammatory causes both inside and outside the lung, in which alveolar epithelial cells and vascular endothelial cells are disrupted by pro-inflammatory factors, resulting in a dysfunctional pulmonary capillary barrier and increased permeability [23]. The risk genes neonates inherit may put them at an increased risk for ARDS and play a role in how the lungs respond to damage. The study of ARDS risk genes helped us to understand the variability in susceptibility to the development of lung lesions. With the widespread application of next-generation sequencing, there are>40 candidate genes associated with the development or outcome of ARDS have been identified, including *ACE*, *VEGF*, and *SFTPB* [24]. However, we found only a marginal significant burden of rare missense variants in *SFTPB* among the candidate genes (P_{MIS} = 0.056, eTable 4 in the Supplement), which may be due to the different pathogenesis of ARDS at different ages.

The essential attributes characterizing the biology of the lung were identified by analyzing the expression pattern at different stages of lung development [25]. Similarly, we highlighted the importance of *EDNRB* by outlining the expression patterns of the eight SCGs in the different periods of lung development in humans and mice. *EDNRB* is a non-selective type endothelin receptor of endothelin-1. Endothelin-1 signaling was found to mediate dynamic changes in fetal pulmonary blood flow [26]. The 151 predicted target genes of *EDNRB* were mainly enriched in the lamellar body-related gene set, and its precursors were also significantly enriched by rare variants. This organelle is found in type II alveolar cells in the lungs and contains densely packed concentric membrane layers for storing and secreting surfactants [27]. Dysfunction of the lamellar body can lead to neonatal respiratory distress syndrome and interstitial lung disease in children, such as a rare pathogenic variant of *ABCA3* [28]. Recent studies have found that lung inflammatory cell infiltration can also cause abnormal lamellar bodies [29]. Therefore, we hypothesized that *EDNRB* may participate in the surfactant secretion process by regulating the function of the lamellar body and late endosome. Regarding the *EDNRB* variants identified in this study, we found a missense variant annotated as a disease-causing mutation by the Human Gene Mutation Database and MutationTaster, namely, rs2070591, that was highly enriched in the ARDS neonates (P < 0.05, eTable 5 and eTable 6 in the Supplement). As a risk variant for Hirschsprung's disease [30], it may also be the main cause of *EDNRB* dysfunction in ARDS. Based on these results, *EDNRB* may be a potential ARDS risk gene and is involved in surfactant homeostasis.

Because neonatal ARDS is a definition that has been around for a short time, there are few studies have been conducted to identify candidate risk genes. However, several other studies have been published on the investigation of risk genes affecting other similar extreme respiratory phenotypes, such as neonatal respiratory distress syndrome and bronchopulmonary dysplasia. Several investigators have done statistical association analyses of individual patient characteristics with non-common variants, albeit in small cohorts (<50 patients) [31,32]. Another small cohort study using whole genome sequencing to identify rare variants in the genome identified 258 genes with rare nonsynonymous mutations in a cohort of 50 twin pairs [33]. Several other studies have begun to use whole-exome data from large cohorts of several hundred patients to identify novel genes and associated pathways that play an important role in the development of neonatal lung disease [34,35]. The relative strength of our study is that we used a large patient cohort and a systematic combination of multi-omics approaches to identify a more accurate and informative set of risk genes. However, there were few overlaps between our result and candidate risk genes from previous studies (overlap = 1.3 %, 7/55, eFigure 2c in the Supplement). This finding may be related to the differences of samples in the regional distribution and phenotypes, including gestational ages, disease types, and races of the study populations.

A major strength of this study is that it was the largest cohort of ARDS neonates to date to conduct a risk genetic study. The approach of using only genetic variation information - does not provide an integrated view. In our study, network-based in silico functional interaction analysis can elucidate the genes and pathways involved, as well as biomarkers and potential drug targets [36]. In silico functional interactions and gene-set enrichment analysis have been used to integrate the vast amount of complex data obtained from genomics and transcriptomics studies [37].

4.1. Limitations

Some limitations need to be improved in the present study. Firstly, our results require additional validation in an independent patient population and further exploration of the underlying molecular mechanisms. Secondly, given the limitations of clinical exome sequencing, we acknowledged that the systematic analyses of whole-genome sequencing in patients from different regions or different races would provide more comprehensive results.

5. Conclusions

In this study, we applied clinical exome sequencing in a cohort of neonatal ARDS and used multiple analytical approaches to identify *EDNRB* as a possible ARDS risk gene. Although *EDNRB* was not the most statistically significant gene identified in the initial analysis, we clarified several biologically plausible proofs by gene expression at different periods of lung development and by in silico functional interaction analysis.

Ethics Statement

The study was conducted ethically following the World Medical Association Declaration of Helsinki and approved by the Research Ethics Committee of the Children's Hospital of Fudan University (CHFudanU_NNICU11). Written informed consent was required from the legal guardian for genetic analysis.

Author Contributions

WZ and LY conceived and supervised the project. HC, XC, and XD designed the study and participated in data analysis and manuscript drafts. LH, XD, and GC participated in the design of the study and critical revision of the manuscript. CY, JZ, and YL contributed to data evaluation and data interpretation. All authors provided comments on the draft manuscript and approved the final manuscript.

Role of the Funder/Sponsor

Funders provide financial support for this study and do not participate in study design, data collection, data analyses, interpretation, or writing of the report.

Data Sharing Statement

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2022.08.055.

References

[1] De Luca D, Tingay DG, van Kaam AH, Courtney SE, Kneyber MCJ, Tissieres P, et al. Epidemiology of Neonatal Acute Respiratory Distress Syndrome: Prospective, Multicenter, International Cohort Study. Pediatr Crit Care Med 2022;23(7):524–34.

[2] De Luca D, van Kaam AH, Tingay DG, Courtney SE, Danhaive O, Carnielli VP, et al. The Montreux definition of neonatal ARDS: biological and clinical background behind the description of a new entity. Lancet Respir Med 2017;5 (8):657–66.

[3] De Luca D, Cogo P, Kneyber MC, Biban P, Semple MG, Perez-Gil J, et al. Surfactant therapies for pediatric and neonatal ARDS: ESPNIC expert consensus opinion for future research steps. Crit Care 2021;25(1):75.

[4] Manti S, Esper F, Alejandro-Rodriguez M, Leonardi S, Betta P, Cuppari C, et al. Respiratory syncytial virus seropositivity at birth is associated with adverse neonatal respiratory outcomes. Pediatr Pulmonol 2020;55(11):3074–9.

[5] You T, Zhou YR, Liu XC, Li LQ. Risk Factors and Clinical Characteristics of Neonatal Acute Respiratory Distress Syndrome Caused by Early Onset Sepsis. Front Pediatr 2022;10:847827.

[6] Hallman M. The surfactant system protects both fetus and newborn. Neonatology 2013;103(4):320–6.

[7] Somaschini M, Presi S, Ferrari M, Vergani B, Carrera P. Surfactant proteins gene variants in premature newborn infants with severe respiratory distress syndrome. J Perinatol 2018;38(4):337–44.

[8] Wang J, Fan J, Zhang Y, Huang L, Shi Y. ABCA3 gene mutations shape the clinical profiles of severe unexplained respiratory distress syndrome in late preterm and term infants. Transl Pediatr 2021;10(2):350–8.

[9] Barilli A, Rotoli BM, Visigalli R, Bussolati O, Gazzola GC, Kadija Z, et al. In Lysinuric Protein Intolerance system y+L activity is defective in monocytes and in GM-CSF-differentiated macrophages. Orphanet J Rare Dis 2010;26(5):32.

[10] Kalinichenko VV, Lim L, Stolz DB, Shin B, Rausa FM, Clark J, et al. Defects in pulmonary vasculature and perinatal lung hemorrhage in mice heterozygous null for the Forkhead Box f1 transcription factor. Dev Biol 2001;235 (2):489–506.

[11] Boggaram V. Thyroid transcription factor-1 (TTF-1/Nkx2.1/TITF1) gene regulation in the lung. Clin Sci (Lond) 2009;116(1):27–35.

[12] Hilgendorff A, Heidinger K, Bohnert A, Kleinsteiber A, Konig IR, Ziegler A, et al. Association of polymorphisms in the human surfactant protein-D (SFTPD) gene and postnatal pulmonary adaptation in the preterm infant. Acta Paediatr 2009;98(1):112–7.

[13] Shen W, Kuang P, Wang B, Zeng Q, Chen C, Lin X. Genetic Polymorphisms of LPCAT1, CHPT1 and PCYT1B and Risk of Neonatal Respiratory Distress Syndrome among a Chinese Han Population. Pediatr Neonatol 2020;61 (3):318–24.

[14] Demircubuk AG, Coskun MY, Demiryurek S, Dokuyucu R, Oztuzcu S, Taviloglu ZS, et al. Endothelial NOS gene Glu298Asp polymorphism in preterm neonates with respiratory distress syndrome. Pediatr Pulmonol 2013;48(10):976–80.

[15] Ambalavanan N, Cotten CM, Page GP, Carlo WA, Murray JC, Bhattacharya S, et al. Integrated genomic analyses in bronchopulmonary dysplasia. J Pediatr 2015;166(3):531–7 e13.

[16] Li W, Long C, Renjun L, Zhangxue H, Yin H, Wanwei L, et al. Association of SCNN1A Single Nucleotide Polymorphisms with neonatal respiratory distress syndrome. Sci Rep 2015;27(5):17317.

[17] Chang HY, Li F, Li FS, Zheng CZ, Lei YZ, Wang J. Genetic Polymorphisms of SP-A, SP-B, and SP-D and Risk of Respiratory Distress Syndrome in Preterm Neonates. Med Sci Monit 2016;24(22):5091–100.

[18] Gudmundsson J, Sulem P, Gudbjartsson DF, Masson G, Agnarsson BA, Benediktsdottir KR, et al. A study based on whole-genome sequencing yields a rare variant at 8q24 associated with prostate cancer. Nat Genet 2012;44 (12):1326–9.

[19] Gibson G. Rare and common variants: twenty arguments. Nat Rev Genet 2012;13(2):135–45.

[20] Hu X, Li N, Xu Y, Li G, Yu T, Yao RE, et al. Proband-only medical exome sequencing as a cost-effective first-tier genetic diagnostic test for patients without prior molecular tests and clinical diagnosis in a developing country: the China experience. Genet Med 2018;20(9):1045–53.

[21] Whitsett JA, Wert SE, Weaver TE. Diseases of pulmonary surfactant homeostasis. Annu Rev Pathol 2015;10:371–93.

[22] Bartman CM, Matveyenko A, Prakash YS. It's about time: clocks in the developing lung. J Clin Invest 2020;130(1):39–50.

[23] Wong JJ, Quek BH, Lee JH. Establishing the entity of neonatal acute respiratory distress syndrome. J Thorac Dis 2017;9(11):4244–7.

[24] Thompson BT, Chambers RC, Liu KD. Acute Respiratory Distress Syndrome. N Engl J Med 2017;377(6):562–72.

[25] Kho AT, Bhattacharya S, Tantisira KG, Carey VJ, Gaedigk R, Leeder JS, et al. Transcriptomic analysis of human lung development. Am J Respir Crit Care Med 2010;181(1):54–63.

[26] Lee JW, Gonzalez RF, Chapin CJ, Busch J, Fineman JR, Gutierrez JA. Nitric oxide decreases surfactant protein gene expression in primary cultures of type II pneumocytes. Am J Physiol Lung Cell Mol Physiol 2005;288(5):L950–7.

[27] Sever N, Milicic G, Bodnar NO, Wu X, Rapoport TA. Mechanism of Lamellar Body Formation by Lung Surfactant Protein B. Mol Cell. 2021 Jan 7;81(1):49-66 e8.

[28] Hu JY, Yang P, Wegner DJ, Heins HB, Luke CJ, Li F, et al. Functional characterization of four ATP-binding cassette transporter A3 gene (ABCA3) variants. Hum Mutat 2020;41(7):1298–307.

[29] Yu J, Tuo F, Luo Y, Yang Y, Xu J. Toxic effects of perinatal maternal exposure to nonylphenol on lung inflammation in male offspring rats. Sci Total Environ 2020;1(737):139238.

[30] Amiel J, Attie T, Jan D, Pelet A, Edery P, Bidaud C, et al. Heterozygous endothelin receptor B (EDNRB) mutations in isolated Hirschsprung disease. Hum Mol Genet 1996;5(3):355–7.

[31] Shulenin S, Nogee LM, Annilo T, Wert SE, Whitsett JA, Dean M. ABCA3 gene mutations in newborns with fatal surfactant deficiency. N Engl J Med 2004;350(13):1296–303.

[32] Carrera P, Di Resta C, Volonteri C, Castiglioni E, Bonfiglio S, Lazarevic D, et al. Exome sequencing and pathway analysis for identification of genetic variability relevant for bronchopulmonary dysplasia (BPD) in preterm newborns: A pilot study. Clin Chim Acta 2015;451(Pt A):39–45.

[33] Li J, Yu KH, Oehlert J, Jeliffe-Pawlowski LL, Gould JB, Stevenson DK, et al. Exome Sequencing of Neonatal Blood Spots and the Identification of Genes Implicated in Bronchopulmonary Dysplasia. Am J Respir Crit Care Med 2015;192 (5):589–96.

[34] Hamvas A, Feng R, Bi Y, Wang F, Bhattacharya S, Mereness J, et al. Exome sequencing identifies gene variants and networks associated with extreme respiratory outcomes following preterm birth. BMC Genet 2018;19(1):94.

[35] Bozkurt HB, Sahin Y. Whole exome sequencing identifies a novel variant in ABCA3 in an individual with fatal congenital surfactant protein deficiency. Turk J Pediatr 2021;63(4):703–7.

[36] Barabasi AL, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet 2011;12(1):56–68.

[37] Lachmann A, Ma'ayan A. Lists2Networks: integrated analysis of gene/protein lists. BMC Bioinf 2010;12(11):87.