



Nucleotide polymorphism in ARDS outcome: a whole exome sequencing association study

Jing-Yuan Xu¹, Ai-Ran Liu¹, Zong-Sheng Wu¹, Jian-Feng Xie², Xiao-Xiao Qu³, Cai-Hua Li³, Shan-Shan Meng², Song-Qiao Liu², Cong-Shan Yang², Ling Liu², Ying-Zi Huang¹, Feng-Mei Guo¹, Yi Yang^{1#}, Hai-Bo Qiu^{1#}

¹Jiangsu Provincial Key Laboratory of Critical Care Medicine, Department of Critical Care Medicine, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China; ²Jiangsu Provincial Key Laboratory of Critical Care Medicine, Department of Critical Care Medicine, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China; ³Center for Genetics & Genomics Analysis, Genesky Biotechnologies Inc., Shanghai, China

Contributions: (I) Conception and design: JY Xu, SQ Liu, CS Yang, L Liu, YZ Huang, Y Yang, HB Qiu; (II) Administrative support: Y Yang, HB Qiu; (III) Provision of study materials or patients: AR Liu, ZS Wu, JF Xie, SS Meng; (IV) Collection and assembly of data: JY Xu, JF Xie; (V) Data analysis and interpretation: XX Qu, SS Meng, CH Li; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Prof. Yi Yang; Prof. Hai-Bo Qiu. Jiangsu Provincial Key Laboratory of Critical Care Medicine, Department of Critical Care Medicine, Zhongda Hospital, School of Medicine, Southeast University, Nanjing 210009, China. Email: yiyiyang2004@163.com; haiboq2000@163.com.

Background: Genetic locus were identified associated with acute respiratory distress syndrome (ARDS). Our goal was to explore the associations between genetic variants and ARDS outcome, as well as subphenotypes.

Methods: This was a single-center, prospective observational trial enrolling adult ARDS patients. After baseline data were collected, blood samples were drawn to perform whole exome sequencing, single nucleotide polymorphism (SNP)/insertion-deletion to explore the quantitative and functional associations between genetic variants and ICU outcome, clinical subphenotypes. Then the lung injury burden (LIB), which was defined as the ratio of nonsynonymous SNP number per megabase of DNA, was used to evaluate its value in predicting ARDS outcome.

Results: A total of 105 ARDS patients were enrolled in the study, including 70 survivors and 35 nonsurvivors. Based on the analysis of a total of 65,542 nonsynonymous SNP, LIB in survivors was significantly higher than nonsurvivors [1,892 (1,848–1,942)/MB versus 1,864 (1,829–1,910)/MB, $P=0.018$], while GO analysis showed that 60 functions were correlated with ARDS outcome, KEGG enrichment analysis showed that SNP/InDels were enriched in 13 pathways. Several new SNPs were found potentially associated with ARDS outcome. Analysis of LIB was used to determine its outcome predicting ability, the area under the ROC curve of which was only 0.6103, and increase to 0.712 when combined with APACHE II score.

Conclusions: Genetic variants are associated with ARDS outcome and subphenotypes; however, their prognostic value still need to be verified by larger trials.

Trial registration: Clinicaltrials.gov NCT02644798. Registered 20 April 2015.

Keywords: Acute respiratory distress syndrome prognosis (ARDS prognosis); single nucleotide polymorphism (SNP); whole-exome sequencing

Submitted Aug 06, 2020. Accepted for publication Feb 07, 2021.

doi: 10.21037/atm-20-5728

View this article at: <http://dx.doi.org/10.21037/atm-20-5728>

Introduction

Acute respiratory distress syndrome (ARDS) is characterized by increased pulmonary vascular permeability and reduced aerated lung tissue (1). With an extremely high hospital mortality among 35–46% (2), current therapeutic strategies to increase ARDS survival are still limited. Advances in etiology and pathology of ARDS are urging. Clinical factors and protocolized therapeutic strategy poorly explain ARDS outcome, the role of genetic locus in the pathogenesis of ARDS is increasingly recognized (3–6).

Numerous genetic variants were identified associated with ARDS outcome. Morrell (7) found that genetic variation in MAP3K1 associated with ventilator-free days in ARDS, while Wei (8) showed that the missense genetic variant in LRRRC16A/CARMIL1 improved survival by attenuating platelet count decline in ARDS patients. However, as a heterogeneous disease with multiple and interactive pathogenic processes, the effect of genetics contributing differently (7–11), meanwhile, the racial and ethnic differences in mortality also exist (12,13). More than five categories of genes were found to associate with ARDS outcome: genes influencing immune regulation, genes influencing endothelial barrier function, genes influencing respiratory epithelial function, genes influencing coagulation, genes influencing injury and oxidative stress and so forth. Then a few genetic risk factors have been discovered by large-scale genotyping approaches, from *in vivo* or *in vitro* models of lung injury, which highlight the importance of identifying genetic biomarkers of ARDS outcome to further improve stratification. The mutational landscape and variability at single nucleotide polymorphisms (SNP) with ARDS outcome in Chinese is unknown, not to mention their associations. By whole-exome sequencing association study, our goal was to explore the associations between genetic variants and ARDS outcome.

We present the following article in accordance with the MDAR reporting checklist (available at <http://dx.doi.org/10.21037/atm-20-5728>).

Methods

Setting

This was an investigator-initiated, single-center, prospective observational trial that was conducted in the intensive care unit of a tertiary care teaching hospital. The study protocol was approved by the Ethics Committee (Approval Number: 2015ZDSYLL014.0) of Zhongda Hospital, School of

Medicine, Southeast University, and written informed consent was obtained from each patient or their next of kin. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Patients

Adult ARDS (according to Berlin definition) patients were enrolled in the trial. The diagnostic criteria included: (I) within one week of a known clinical insult or new or worsening respiratory symptoms; (II) chest imaging showing that bilateral opacities-not fully explained by effusions, lobar/lung collapse, or nodules; (III) respiratory failure not fully explained by cardiac failure or fluid overload; and (IV) arterial partial pressure of oxygen/fraction of inspiration oxygen ($\text{PaO}_2/\text{FiO}_2$ ratio, P/F ratio) less than or equal to 300 mmHg.

Data collection

Baseline-recorded data included demographic characteristics, comorbidities, and the origin and etiology of ARDS were collected by trained investigators. Severity of illness was assessed with the Acute Physiology and Chronic Health Evaluation (APACHE) II score within 24 hours on enrollment. Sequential Organ Failure Assessment (SOFA) and Murray lung injury score within 24 hours on enrollment were also calculated.

Predisposing conditions of ARDS were collected, and subphenotypes of ARDS were determined. Severe ARDS group and non-severe group were divided according to the severity of lung injury (Berlin definition). Patients with risk factors of pneumonia (pulmonary sepsis), pulmonary contusion, inhalation and drowning were categorized as having pulmonary ARDS, whereas patients with risk factors of non-pulmonary sepsis or pancreatitis were categorized as having extrapulmonary ARDS. Patients with sepsis on enrollment after enrolled were recorded as ARDS with sepsis. Patients with shock on enrollment were recorded as ARDS with shock. Sepsis was defined by Sepsis 3.0.

Peripheral blood samples were drawn. Prognosis was recorded as the survivors and non-survivors in ICU.

Study methods

Whole-exome sequencing was performed by the sequencing platform Illumina, the data were compared with reference genome UCSC hg19. Firstly, genomic DNA was isolated

Table 1 Baseline patient characteristics and subphenotypes of ARDS

Characteristic	Patients (n=105)
Age, year	59 [46–73]
Sex, male/female	83/22
APACHE II score	23 [17–27]
SOFA score	9 [6–12]
Murray lung injury score	2.7 [2.1–3.3]
Severe ARDS, n (%)	52 (50.0)
Pulmonary ARDS, n (%)	91 (87.0)
ARDS combined with sepsis on enrollment, n (%)	89 (85.0)
ARDS combined with shock on enrollment, n (%)	66 (63.0)
Non-survivors in ICU, n (%)	35 (34.0)

Data presented as median [interquartile range]. ARDS, acute respiratory distress syndrome; APACHE II score, Acute Physiology and Chronic Health Evaluation II score; SOFA score, Sequential Organ Failure Assessment score.

from the peripheral blood samples taken from individuals by following the manufacturer's standard procedure using QIAamp DNA Blood kits (Qiagen, Hilden, Germany). Then exome sequence capture was performed on SureSelect Human All Exon V6 (Agilent). DNA library was subjected to 2×150 bp paired-end massively parallel sequencing using a HiSeq2000 Sequencing System (Illumina, San Diego, CA, USA). Before variant calling, sequence alignment files were generated to duplicate removal, local realignment around known Indels and base quality recalibration using the Genome Analysis Toolkit (GATK) (14). Variations that included single-nucleotide variants (SNVs) and insertions or deletions (Indels) were identified using both the VarScan 2.2.7 software package (<http://www.ncbi.nlm.nih.gov/pubmed/22300766>) (15) as well as the variant quality score recalibration (VQSR) protocol in GATK, and further filtered using a recommended threshold value (mapping quality >30, base quality >15, and read numbers >3). Then, SNP available at dbSNP130 (hg19) as well as those reported by the 1000 Genomes Project were filtered out from the output files using the ANNOVAR (<http://nar.oxfordjournals.org/content/38/16/e164>) (16).

After identifying a newly number of coding SNPs potentially associated with ARDS, SNP/InDel were tested by plink method to understand the difference between the

outcome and subphenotypes of ARDS. While detecting the number and function of nonsynonymous SNV, the lung injury burden (LIB) was calculated by the ratio of nonsynonymous SNP number per megabase (MB) of DNA. The area under the receiver operating characteristic curve (ROC) was used for evaluating the predictive values of LIB in predicting outcome and subphenotype of patients with ARDS.

Statistical analysis

Data were presented as number (%) for categorical variables, and median (interquartile range) for continuous variables. Fisher's exact test or χ^2 test were used for categorical variables, and Student's *t*-test or Mann-Whitney U test were used for continuous variables, as appropriate. The value of predictive ability was evaluated by the area under the curve (AUC) in the receiver operating characteristic (ROC) analysis. A P value <0.05 was considered statistically significant. Statistical analyses were carried out by the SPSS 16.0 software (IBM, Somers, NY, USA).

Results

There were 105 patients enrolled in the study, including 70 survivors and 35 non-survivors. The characteristics of outcome and subphenotype are presented in *Table 1*. The median age was 59 years old, while the median APACHE II score was 23, the median SOFA score and Murray lung injury score was 9 and 2.7, respectively. Among them, 91 patients were categorized as having pulmonary ARDS, 89 patients were diagnosed as sepsis and 66 patients as shock on enrollment.

SNP/InDel data by whole-exome sequencing

By whole-exome sequencing, the number of SNP/InDel were 471,131 (*Table 2*). Among them, 120,830 SNP/InDel were in exonic region. The number of nonsynonymous SNV were 65,542, with 436 of frameshift-insertion for InDel and 897 of frameshift-deletion for InDel. GO analysis showed that 52 functions were correlated with ARDS development ($P<0.01$), and KEGG enrichment analysis showed that these SNP/InDel were in 10 pathways, such as cGMP-PKG signaling pathway, Platelet activation ($P<0.05$).

SNP/InDel data between ARDS patients with different outcome

LIB was tested to determine the quantitative differences

between survivors and non-survivors. LIB of survivors was significantly higher than non-survivors [1,892 (1,848–1,942)/MB versus 1,864 (1,829–1,910)/MB, P=0.018]. GO and KEGG analysis were performed to determine the

functional difference of SNP with outcome. GO analysis showed that 60 functions were correlated with ARDS outcome (P<0.01) (Figure 1), and KEGG enrichment analysis showed these SNP/InDel were in 13 pathways (Table 3), such as ECM-receptor interaction pathway, Platelet activation pathway and cGMP-PKG signaling pathway (P<0.01).

Table 2 Functional type of SNP/InDel count by whole-exome sequencing

Functional type	SNP/InDel count
Nonsynonymous SNV	64,452
Synonymous SNV	49,590
Unknown	2,613
Frameshift insertion	436
Frameshift deletion	897
Nonframeshift insertion	792
Nonframeshift deletion	1,269
Stopgain	1,125
Stoploss	66

SNP, single nucleotide polymorphism; SNV, single-nucleotide variant.

Association of genetic polymorphisms with ARDS outcome

To identify the novel SNPs which associated with ARDS outcome, the genotype distribution in different gene were summarized in Table 4, conformed to Hardy-Weinberg equilibrium. Although no strong evidence of stratification has been reported, several SNPs which potentially associated with ARDS outcome were found (Figure 2).

The value of LIB in predicting ARDS outcome

To determine the ability of LIB to predict ARDS outcome, analysis was carried out on LIB, P/F ratio, APACHE II score, SOFA score and Murray lung injury score with the

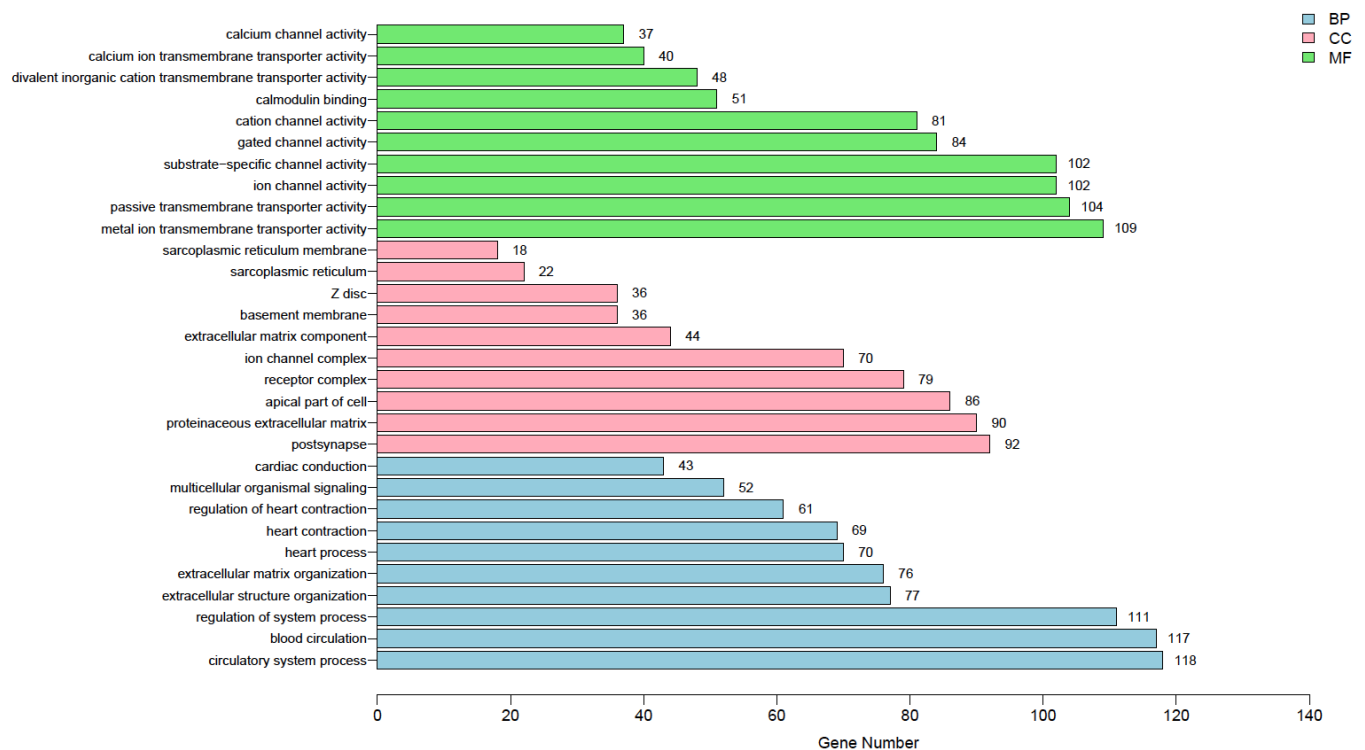


Figure 1 GO analysis showed that the top 30 of the 60 functions were correlated with ARDS outcome. ARDS, acute respiratory distress syndrome; BP, biological process; CC, cellular component; MF, molecular function.

Table 3 KEGG enrichment analysis for SNP/InDel data between different ARDS outcome

KEGG enrichment pathway	Number of genes	P value
ECM-receptor interaction	29	0.00144
Purine metabolism	48	0.00279
Protein digestion and absorption	29	0.00336
Platelet activation	36	0.00336
Calcium signaling pathway	48	0.00336
Thyroid hormone signaling pathway	34	0.00367
Insulin secretion	27	0.00386
Oxytocin signaling pathway	41	0.00495
Phospholipase D signaling pathway	39	0.00685
cGMP-PKG signaling pathway	42	0.00782
Glutamatergic synapse	32	0.00782
Adrenergic signaling in cardiomyocytes	38	0.00807
Long-term depression	20	0.00821

SNP, single nucleotide polymorphism; ARDS, acute respiratory distress syndrome.

area under the ROC curve of 0.6103 ($P=0.0807$), 0.568 ($P=0.3124$), 0.6763 ($P=0.0053$), 0.6204 ($P=0.1002$), 0.6614 ($P=0.0581$), respectively. The predicting value of LIB could increase to 0.712 ($P=0.001$) when combined with APACHE II score.

SNP/InDel data between ARDS patients with different subphenotypes

ARDS patients were divided into different subphenotypes (see online supplemental material in [Appendix 1](#)).

Severe ARDS group and non-severe group were divided according to the severity of lung injury. Compared with non-severe group, LIB was lower in severe ARDS group, with the ROC of predictive value of 0.727 ($P<0.0001$). GO analysis showed that 25 functions were correlated with ARDS severity ($P<0.01$), and KEGG enrichment analysis showed that these SNP/InDel were in 4 pathways, such as PI3K-Akt signaling pathway, ECM-receptor interaction ($P<0.05$).

ARDS patients were divided into pulmonary ARDS and extrapulmonary ARDS group. LIB was not significantly altered between the pulmonary and extrapulmonary ARDS. GO analysis showed that 19 functions were correlated

with pulmonary and extrapulmonary ARDS ($P<0.01$), and KEGG enrichment analysis showed that these SNP/InDel were in 8 pathways, such as ECM-receptor interaction ($P<0.05$).

ARDS patients were divided into ARDS combined with sepsis and ARDS without sepsis on enrollment. Compared with patients without sepsis, the LIB was lower in ARDS combined with sepsis, with the ROC of predictive value of 0.6803 ($P=0.0084$). GO analysis showed that 24 functions were correlated with ARDS combined with sepsis ($P<0.01$), and KEGG enrichment analysis showed that these SNP/InDel were in 3 pathways, such as ECM-receptor interaction, Focal adhesion ($P<0.05$).

ARDS patients were divided into ARDS combined with shock and ARDS without shock on enrollment. Compared with patients without shock, the LIB was lower in ARDS combined with shock, with the ROC of predictive value of 0.6915 ($P=0.0008$). GO analysis showed that 46 functions were correlated with ARDS combined with shock ($P<0.01$), and KEGG enrichment analysis showed that these SNP/InDel were in 10 pathways, such as cAMP signaling pathway, ECM-receptor interaction ($P<0.05$).

Discussion

In this single-center, prospective observational trial which enrolled adult ARDS patients, whole exome-sequencing was performed to understand the difference between ARDS outcome and subphenotypes. The highlight of the study is the integrated framework of genetic variability of ARDS displayed through ARDS survivors and non-survivors. As defined by LIB, the mutational landscape of ARDS showed the overall genetic variability between survivors and non-survivors, while the detailed specific genetic polymorphisms which have an influence on outcome which finally showed genetic factors play a role in ARDS outcome.

As the role of genetics in the pathogenesis of ARDS is increasingly recognized, numerous genes and genetic variants were identified to proclaim their association with ARDS outcome. However, most were single genetic polymorphisms, little studies focus on the whole mutational landscape and its influence on ARDS outcome. To build an integrated framework, we classified different categories of genes, and try to observe their association with the outcome of ARDS, which are genes influencing immune regulation, genes influencing endothelial barrier function, genes influencing respiratory epithelial function, genes influencing coagulation, genes influencing injury and oxidative stress

Table 4 Single nucleotide polymorphisms and their genotype distribution between ARDS outcome

SNP	Model	Genotype	Case	Control	OR (95% CI)	P value	Gene
rs3735041	Codominant	G/G	11	2	24.57 (4.45–135.7)	0.0002402	<i>AKR1B10</i>
rs3735041	Recessive	A/A-A/G	24	67	16.17 (3.293–79.4)	0.0006081	<i>AKR1B10</i>
rs10904402	Dominant	A/A	21	18	0.2144 (0.08731–0.5264)	0.0007792	<i>AKR1C3</i>
rs10904402	Additive	–	–	–	0.2955 (0.1446–0.6039)	0.0008289	<i>AKR1C3</i>
rs12239311	Additive	–	–	–	4.247 (2.021–8.923)	0.0001349	<i>CFAP57</i>
rs12239311	Dominant	C/C	10	46	5.935 (2.331–15.11)	0.0001882	<i>CFAP57</i>
rs11210812	Additive	–	–	–	4.098 (1.94–8.655)	0.0002179	<i>CFAP57</i>
rs11210805	Additive	–	–	–	3.876 (1.824–8.237)	0.0004277	<i>CFAP57</i>
rs513009	Additive	–	–	–	3.876 (1.824–8.237)	0.0004277	<i>CFAP57</i>
rs2453412	Additive	–	–	–	3.876 (1.824–8.237)	0.0004277	<i>CFAP57</i>
rs612626	Additive	–	–	–	3.876 (1.824–8.237)	0.0004277	<i>CFAP57</i>
rs616045	Additive	–	–	–	4.006 (1.844–8.703)	0.0004544	<i>CFAP57</i>
rs663824	Additive	–	–	–	3.839 (1.808–8.153)	0.0004634	<i>CFAP57</i>
rs598336	Additive	–	–	–	3.839 (1.808–8.153)	0.0004634	<i>CFAP57</i>
rs75528102	Additive	–	–	–	3.797 (1.763–8.179)	0.0006557	<i>CFAP57</i>
rs663336	Additive	–	–	–	3.797 (1.763–8.179)	0.0006557	<i>CFAP57</i>
rs499839	Codominant	A/C	25	23	5.999 (2.265–15.89)	0.0003115	<i>CLSTN2</i>
rs499839	Dominant	A/A	8	42	5.603 (2.144–14.65)	0.0004392	<i>CLSTN2</i>
rs533657	Additive	–	–	–	3.966 (1.772–8.877)	0.0008043	<i>CLSTN2</i>
rs495109	Codominant	T/A	22	20	4.58 (1.852–11.32)	0.0009866	<i>CLSTN2</i>
rs589819	Additive	–	–	–	3.797 (1.763–8.179)	0.0006557	<i>FAM183A</i>
rs626842	Additive	–	–	–	3.659 (1.735–7.717)	0.0006571	<i>FAM183A</i>
rs60192064	Additive	–	–	–	3.606 (1.686–7.712)	0.0009434	<i>FAM183A</i>
rs2453416	Additive	–	–	–	3.606 (1.686–7.712)	0.0009434	<i>FAM183A</i>
rs291083	Dominant	A/A	14	53	4.97 (2.047–12.06)	0.0003956	<i>FCMR</i>
rs167082	Dominant	C/C	14	53	4.97 (2.047–12.06)	0.0003956	<i>FCMR</i>
rs291083	Additive	–	–	–	4.099 (1.848–9.093)	0.0005207	<i>FCMR</i>
rs167082	Additive	–	–	–	4.099 (1.848–9.093)	0.0005207	<i>FCMR</i>
rs4645915	Additive	–	–	–	4.586 (1.944–10.82)	0.0005038	<i>KAT5</i>
rs4244812	Codominant	A/A	10	4	12.36 (2.809–54.38)	0.0008799	<i>KAT5</i>
rs4645915	Dominant	T/T	13	50	4.523 (1.851–11.05)	0.0009311	<i>KAT5</i>
rs2274344	Dominant	T/T	20	62	6.589 (2.35–18.47)	0.0003372	<i>MIPEP</i>
rs2274344	Additive	–	–	–	6.589 (2.35–18.47)	0.0003372	<i>MIPEP</i>
rs12866705	Dominant	T/T	21	62	5.802 (2.057–16.36)	0.0008895	<i>MIPEP</i>
rs12866705	Additive	–	–	–	5.802 (2.057–16.36)	0.0008895	<i>MIPEP</i>

Table 4 (continued)

Table 4 (continued)

SNP	Model	Genotype	Case	Control	OR (95% CI)	P value	Gene
rs17007214	Dominant	C/C	31	35	0.1298 (0.04112–0.4094)	0.0004959	MYF5
rs10862184	Dominant	T/T	31	35	0.1298 (0.04112–0.4094)	0.0004959	MYF5
rs17007214	Additive	–	–	–	0.1456 (0.04786–0.4428)	0.0006859	MYF5
rs10862184	Additive	–	–	–	0.1456 (0.04786–0.4428)	0.0006859	MYF5
rs2227294	Codominant	G/G	13	4	15.65 (3.737–65.54)	0.0001672	MYL3
rs2227294	Additive	–	–	–	3.548 (1.789–7.038)	0.0002891	MYL3
rs2227294	Recessive	T/T-T/G	22	65	10.24 (2.88–36.42)	0.0003253	MYL3
rs875956	Codominant	C/T	15	9	6.269 (2.259–17.39)	0.0004228	PIEZO2
rs875956	Additive	–	–	–	3.585 (1.749–7.347)	0.000487	PIEZO2
rs73943314	Dominant	T/T	16	55	4.699 (1.929–11.45)	0.0006602	PIEZO2
rs9954308	Dominant	T/T	16	55	4.699 (1.929–11.45)	0.0006602	PIEZO2
rs2277858	Dominant	T/T	15	53	4.561 (1.888–11.02)	0.0007457	PIEZO2
rs2291822	Dominant	A/A	19	14	0.2147 (0.0879–0.5243)	0.0007319	TLL1
rs2291822	Additive	–	–	–	0.3165 (0.1614–0.6208)	0.0008169	TLL1
rs1061495	Dominant	T/T	20	61	5.805 (2.124–15.86)	0.0006064	TNC
rs1061495	Dominant	T/T	20	61	5.805 (2.124–15.86)	0.0006064	TNC
rs2274750	Dominant	C/C	22	63	6.827 (2.23–20.9)	0.0007659	TNC
rs2274750	Additive	–	–	–	6.827 (2.23–20.9)	0.0007659	TNC
rs79003972	Dominant	C/C	22	62	6.712 (2.194–20.54)	0.000848	TNC
rs79003972	Additive	–	–	–	6.712 (2.194–20.54)	0.000848	TNC
rs2094794	Dominant	C/C	22	63	7.476 (2.274–24.57)	0.0009218	TNC
rs2094794	Additive	–	–	–	7.476 (2.274–24.57)	0.0009218	TNC
rs58968019	Additive	–	–	–	7.044 (2.464–20.14)	0.0002694	UBXN4
rs16831997	Dominant	C/C	20	61	5.937 (2.142–16.45)	0.0006149	UBXN4
rs16831997	Additive	–	–	–	5.937 (2.142–16.45)	0.0006149	UBXN4
rs4851890	Codominant	T/C	24	21	5.373 (2.048–14.1)	0.0006357	VWA3B
rs6967385	Dominant	T/T	21	19	0.2065 (0.08322–0.5126)	0.0006723	VWDE
rs6967385	Codominant	T/G	8	37	0.1735 (0.06266–0.4801)	0.0007455	VWDE

ARDS, acute respiratory distress syndrome.

and so forth. However, as multiple different pathogenic processes, all these genes could interrelate.

Tumor mutation burden (TMB) is a marker which calculated as the nonsynonymous mutation number of per MB of DNA in tumor tissue (17). High TMB often correlates with a higher probability of tumor neoantigens, which could be recognized by lymphocytes (18,19), so it is

hypothesized that the tumors with the highest TMB might be more likely to respond to immune checkpoint blockade therapy. Previous studies showed that patients with high TMB response better to immune checkpoint blockade therapy (20–22) and might have a better outcome (23). However, little data observed the mutation burden in ARDS, which might make a rough estimate on the whole

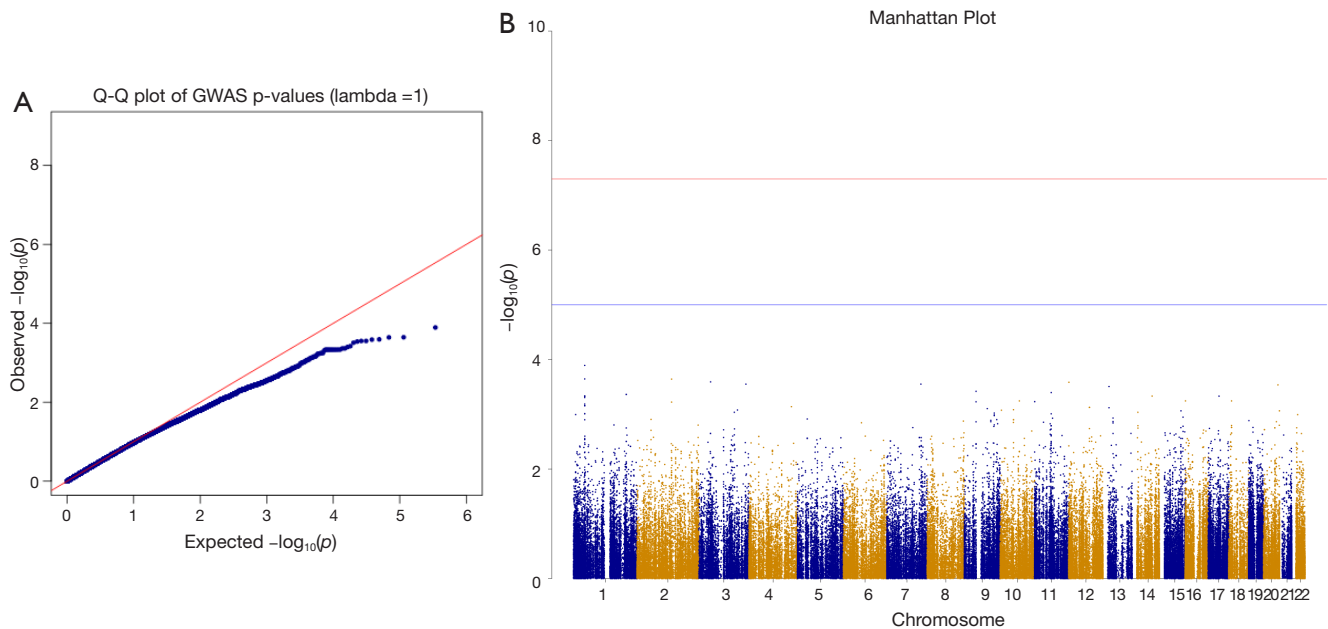


Figure 2 Genome-wide association study showed that several SNPs potentially associated with ARDS outcome. (A) The QQ-plots of the results of whole exome sequencing of ARDS obtained in the analyses. (B) Corresponding Manhattan plots for the same analysis on the left panel. SNP, single nucleotide polymorphism; ARDS, acute respiratory distress syndrome.

mutational landscape of ARDS. In this study, we found when combined with clinical characteristics, burden could predict prognosis of ARDS.

We acknowledge some limitations in our study. Firstly, there was no validation group to study the association of functional SNPs with ARDS outcome which found by whole exome sequencing. Secondly, functional studies are needed to evaluate the mechanisms that underlie the associations between all the genetic variants and ARDS outcome and the mediating pathway. Thirdly, there was no healthy control group, the data were compared with reference genome UCSC hg19. In addition, the findings in the study were mainly pertinent to patients in single center who developed ARDS, and should be validated before generalization in cohorts.

Conclusions

Genetic variants are associated with ARDS outcome; however, their prognostic value still need to be verified by larger trials.

Acknowledgments

Funding: This work is partially supported by grants from the

National Natural Science Foundations of China (81501705, 81571874, 81930058, 81671892, 81971888), grants from the Scientific Research Foundation of Graduate School of Southeast University (YBPY1604), grants from the Jiangsu Provincial Medical Youth Talent (QNRC2016808). This work is partially supported by grants from the Jiangsu Provincial Special Program of Medical Science (BL2013030, BE2018743, BE2019749), grants from Ministry of Science and Technology of China (2020YFC0843700), grants from Jiangsu Provincial Medical Priority Majors (ZDXKA2016025).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at <http://dx.doi.org/10.21037/atm-20-5728>

Data Sharing Statement: Available at <http://dx.doi.org/10.21037/atm-20-5728>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-5728>). 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. The study protocol was approved by the Ethics Committee (Approval Number: 2015ZDSYLL014.0) of Zhongda Hospital, School of Medicine, Southeast University, and written informed consent was obtained from each patient or their next of kin. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

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

1. Fan E, Brodie D, Slutsky AS. Acute respiratory distress syndrome: advances in diagnosis and treatment. *JAMA* 2018;319:698-710.
2. Bellani G, Laffey JG, Pham T, et al. Epidemiology, patterns of care, and mortality for patients with acute respiratory distress syndrome in intensive care units in 50 countries. *JAMA* 2016;315:788-800.
3. Reilly JP, Christie JD. Linking genetics to ARDS pathogenesis: the role of the platelet. *Chest* 2015;147:585-6.
4. Reilly JP, Christie JD, Meyer NJ. Fifty years of research in ARDS. Genomic contributions and opportunities. *Am J Respir Crit Care Med* 2017;196:1113-21.
5. Grigoryev DN, Cheranova DI, Chaudhary S, et al. Identification of new biomarkers for Acute Respiratory Distress Syndrome by expression-based genome-wide association study. *BMC Pulm Med* 2015;15:95.
6. Shortt K, Chaudhary S, Grigoryev D, et al. Identification of novel single nucleotide polymorphisms associated with acute respiratory distress syndrome by exome-seq. *PLoS One* 2014;9:e111953.
7. Morrell ED, O'Mahony DS, Glavan BJ, et al. Genetic variation in MAP3K1 associates with ventilator-free days in Acute Respiratory Distress Syndrome. *Am J Respir Cell Mol Biol* 2018;58:117-25.
8. Wei Y, Tejera P, Wang Z, et al. A Missense genetic variant in LRRC16A/CARMIL1 improves Acute Respiratory Distress Syndrome survival by attenuating platelet count decline. *Am J Respir Crit Care Med* 2017;195:1353-61.
9. Dötsch A, Eisele L, Rabeling M, et al. Hypoxia Inducible Factor-2 Alpha and Prolinhydroxylase 2 Polymorphisms in patients with Acute Respiratory Distress Syndrome (ARDS). *Int J Mol Sci* 2017;18:1266.
10. Sapru A, Liu KD, Wiemels J, et al. Association of common genetic variation in the protein C pathway genes with clinical outcomes in acute respiratory distress syndrome. *Crit Care* 2016;20:151.
11. Kangelaris KN, Sapru A, Calfee CS, et al. The association between a Darc gene polymorphism and clinical outcomes in African American patients with acute lung injury. *Chest* 2012;141:1160-9.
12. Moss M, Mannino DM. Race and gender differences in acute respiratory distress syndrome deaths in the United States: an analysis of multiple-cause mortality data (1979-1996). *Crit Care Med* 2002;30:1679-85.
13. Erickson SE, Shlipak MG, Martin GS, et al. Racial and ethnic disparities in mortality from acute lung injury. *Crit Care Med* 2009;37:1-6.
14. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297-303.
15. Koboldt D C, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 2012;22:568-76.
16. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010;38:e164.
17. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 2013;499:214-8.
18. Matsushita H, Vesely MD, Koboldt DC, et al. Cancer exome analysis reveals a T-cell-dependent mechanism of cancer immunoediting. *Nature* 2012;482:400-4.
19. Maleki Vareki S. High and low mutational burden tumors versus immunologically hot and cold tumors and response to immune checkpoint inhibitors. *J Immunother Cancer* 2018;6:157.
20. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J*

- Med 2017;377:2500-1.
21. Cristescu R, Mogg R, Ayers M, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018;362:eaar3593.
 22. Chan TA, Yarchoan M, Jaffee E, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol* 2019;30:44-56.
 23. Devarakonda S, Rotolo F, Tsao MS, et al. Tumor mutation burden as a biomarker in resected non-small-cell lung cancer. *J Clin Oncol* 2018;36:2995-3006.

Cite this article as: Xu JY, Liu AR, Wu ZS, Xie JF, Qu XX, Li CH, Meng SS, Liu SQ, Yang CS, Liu L, Huang YZ, Guo FM, Yang Y, Qiu HB. Nucleotide polymorphism in ARDS outcome: a whole exome sequencing association study. *Ann Transl Med* 2021;9(9):780. doi: 10.21037/atm-20-5728

ARDS patients were divided into different subphenotypes.

Severe ARDS group and non-severe group were divided according to the severity of lung injury.

Compared with non-severe group, LIB was lower in severe ARDS group [1,911.17 (1,870.37–1,944.03) vs. 1,855.11 (1,823.79–1,878.54), $P=0.000$], with the ROC of predictive value of 0.727 ($P<0.0001$). GO analysis showed that 25 functions were correlated with ARDS severity ($P<0.01$) (Figure S1), and KEGG enrichment analysis showed that these SNP/InDel were in 4 pathways, such as PI3K-Akt signaling pathway, ECM-receptor interaction ($P<0.05$) (Figure S2).

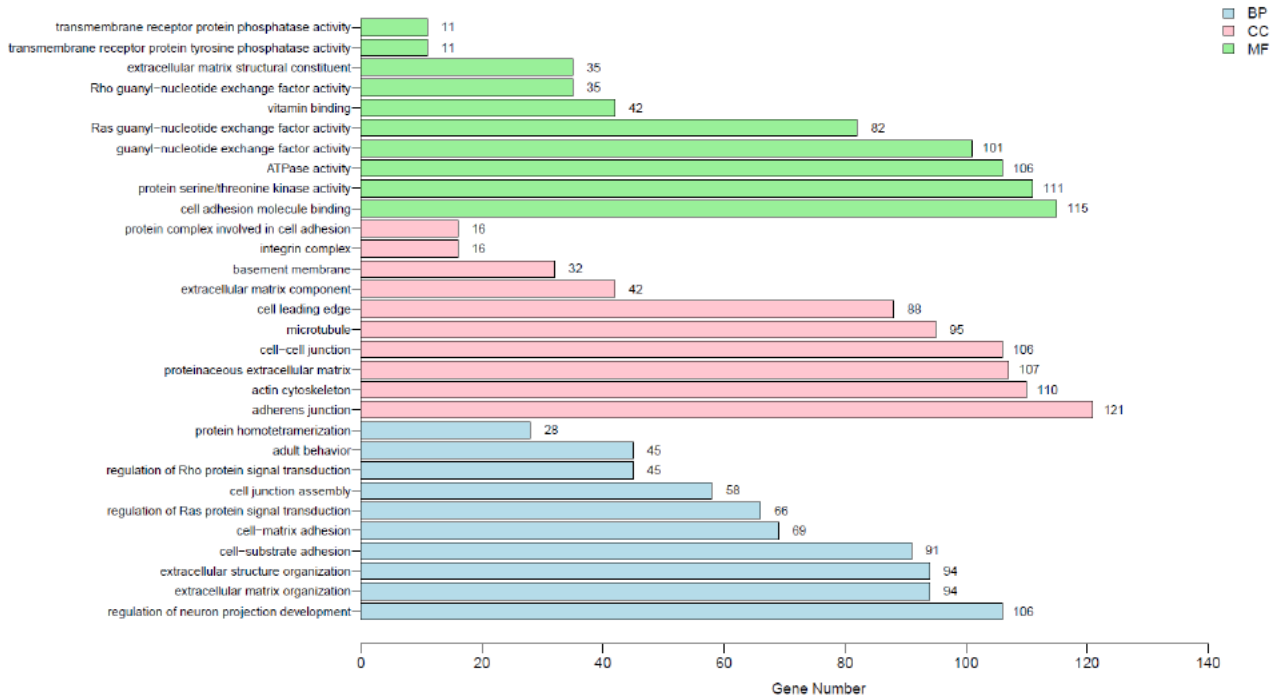


Figure S1 GO analysis showed that the Top 30 of the 60 functions were correlated with ARDS severity. BP, Biological process; CC, Cellular component; MF, Molecular function.

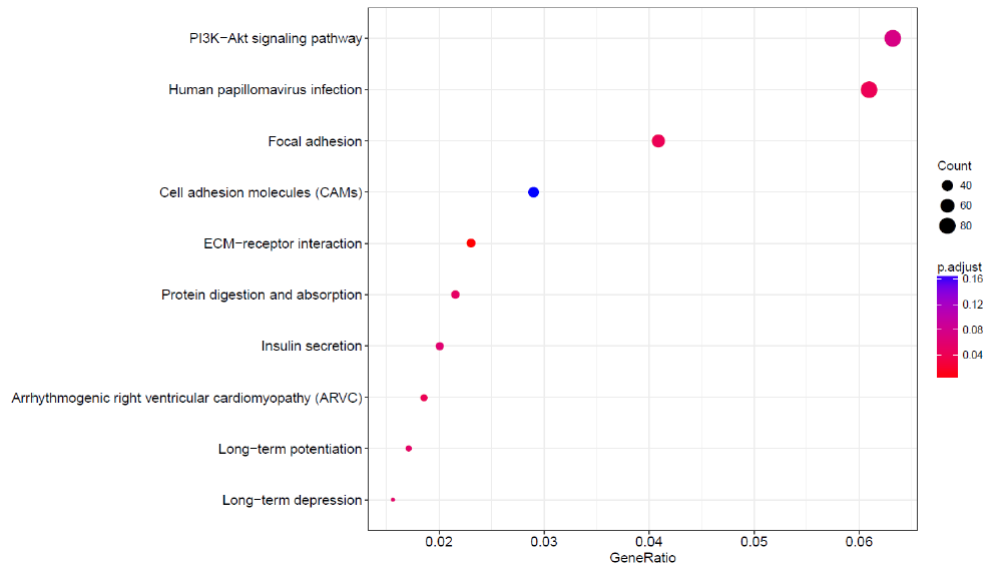


Figure S2 KEGG enrichment pathway analysis showed pathways correlated with ARDS severity. ARDS patients were divided into pulmonary ARDS and extrapulmonary ARDS group. LIB was not significantly altered between the pulmonary and extrapulmonary ARDS [1,877.49 (1,846.03–1,933.11) vs. 1,880.49 (1,835.22–1,919.92), $P=0.717$]. GO analysis showed that 19 functions were correlated with pulmonary and extrapulmonary ARDS ($P<0.01$) (Figure S3), and KEGG enrichment analysis showed that these SNP/InDel were in 8 pathways, such as ECM-receptor interaction ($P<0.05$) (Figure S4).

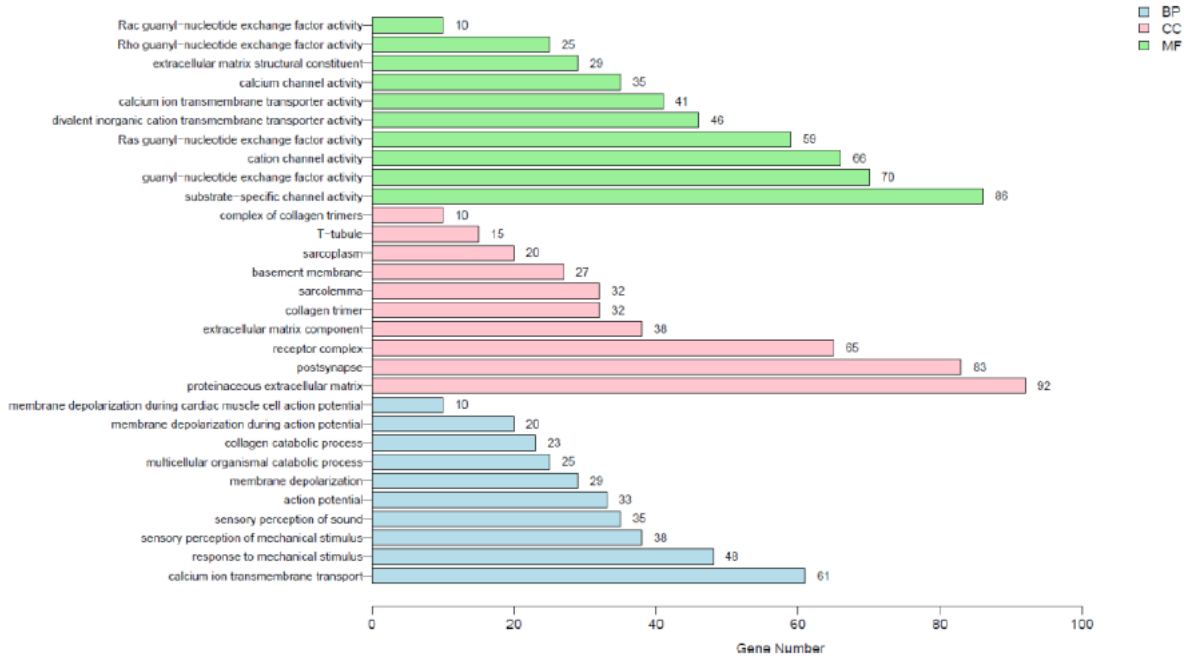


Figure S3 GO analysis showed that the Top 30 of the 60 functions were correlated with pulmonary and extrapulmonary ARDS. BP, Biological process; CC, Cellular component; MF, Molecular function.

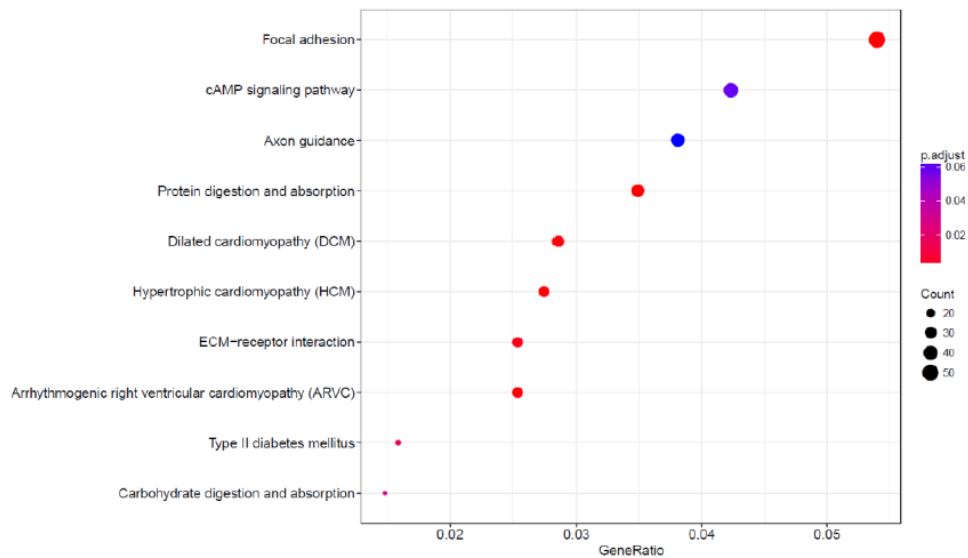


Figure S4 KEGG enrichment pathway analysis showed pathways correlated with pulmonary and extrapulmonary ARDS. ARDS patients were divided into ARDS combined with sepsis and ARDS without sepsis on enrollment. Compared with patients without sepsis, the LIB was lower in ARDS combined with sepsis [1911.82 (1880.22–1944.17) vs. 1870.13 (1833.88–1928.44), $P=0.044$], with the ROC of predictive value of 0.6803 ($P=0.0084$). GO analysis showed that 24 functions were correlated with ARDS combined with sepsis ($P<0.01$) (Figure S5), and KEGG enrichment analysis showed that these SNP/InDel were in 3 pathways, such as ECM-receptor interaction, Focal adhesion ($P<0.05$) (Figure S6).

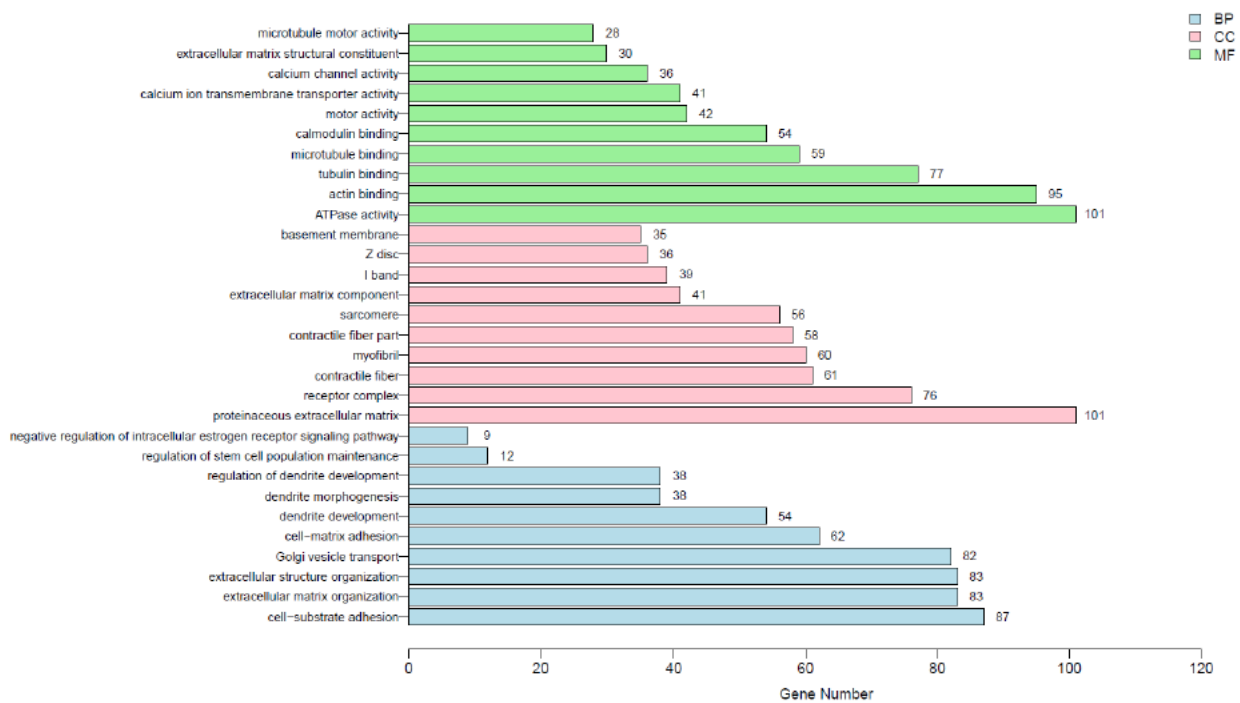


Figure S5 GO analysis showed that the Top 30 of the 60 functions were correlated with ARDS and sepsis. BP, Biological process; CC, Cellular component; MF, Molecular function.

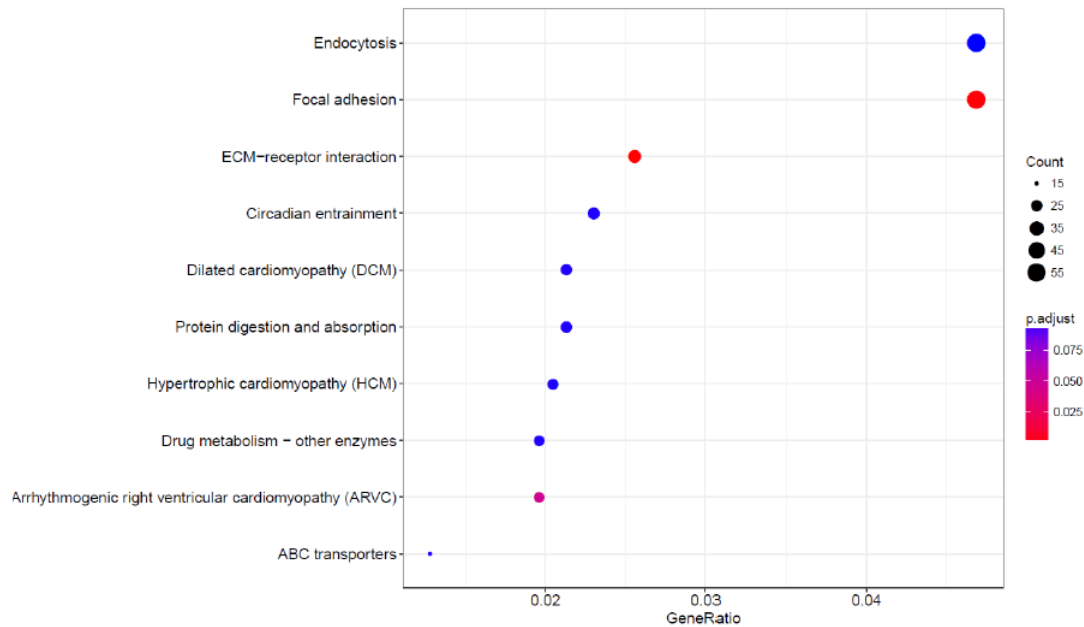


Figure S6 KEGG enrichment pathway analysis showed pathways correlated with ARDS and sepsis. ARDS patients were divided into ARDS combined with shock and ARDS without shock on enrollment. Compared with patients without shock, the LIB was lower in ARDS combined with shock [1,911.05 (1,874.18–1,940.98) vs. 1,858.70 (1,825.16–1,908.99), $P=0.004$], with the ROC of predictive value of 0.6915 ($P=0.0008$). GO analysis showed that 46 functions were correlated with ARDS combined with shock ($P<0.01$) (Figure S7), and KEGG enrichment analysis showed that these SNP/InDel were in 10 pathways, such as cAMP signaling pathway, ECM-receptor interaction ($P<0.05$) (Figure S8).

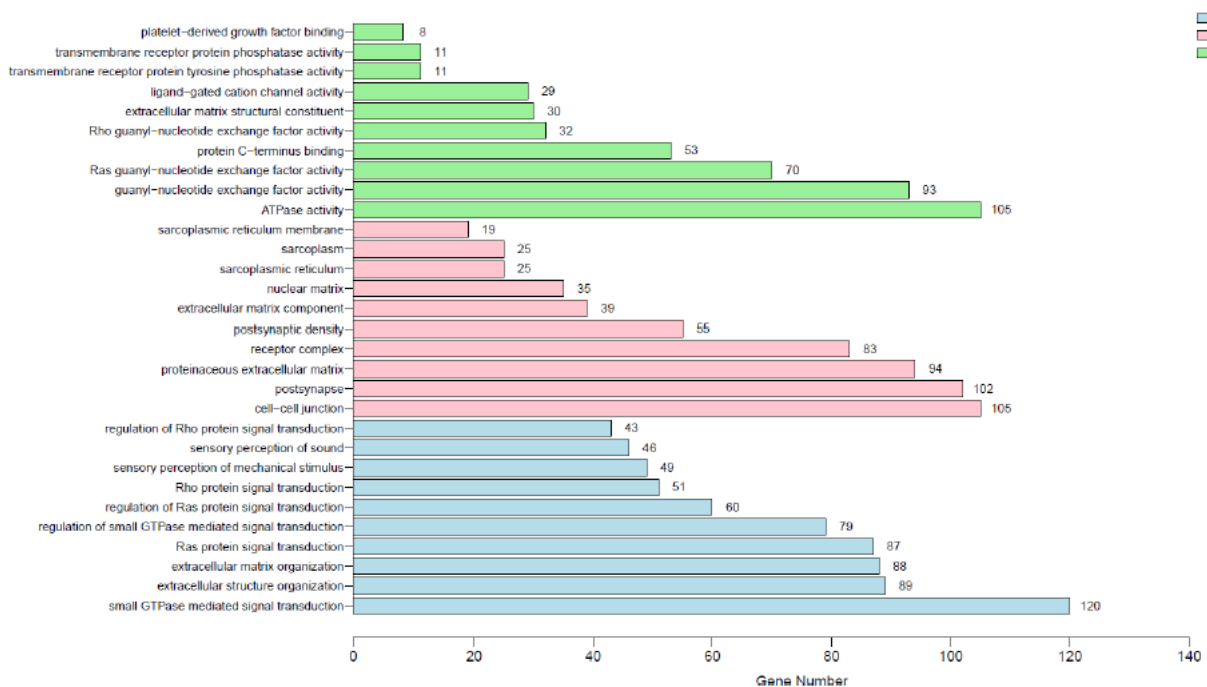


Figure S7 GO analysis showed that the Top 30 of the 60 functions were correlated with ARDS and shock. BP, Biological process; CC, Cellular component; MF, Molecular function.

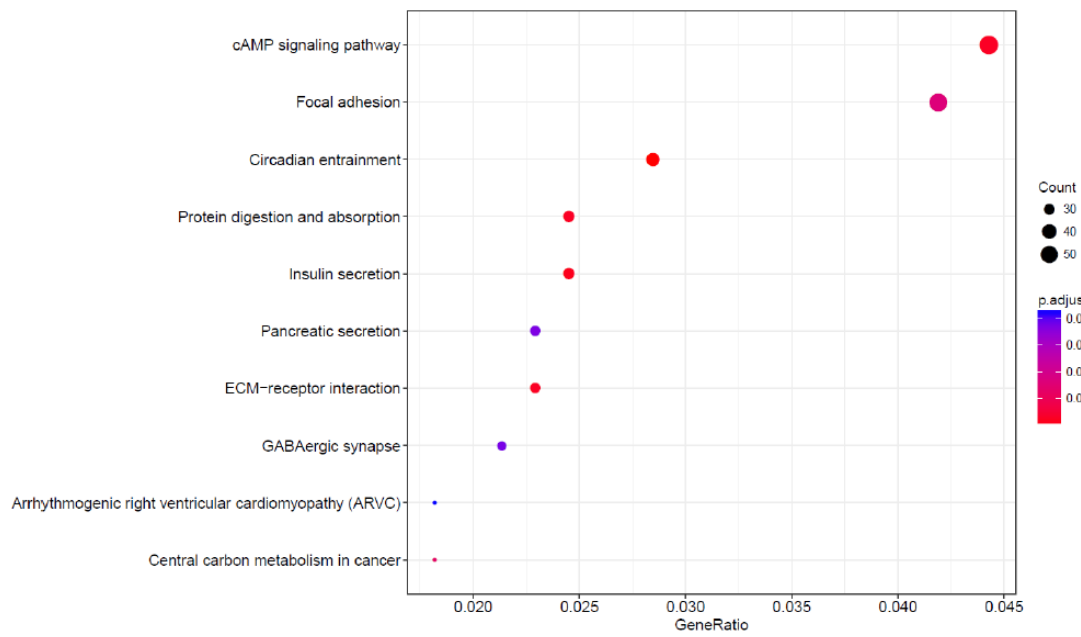


Figure S8 KEGG enrichment pathway analysis showed pathways correlated with ARDS and shock.