Supplementary Material

Distinct and replicable genetic risk factors for acute respiratory distress syndrome of pulmonary or extrapulmonary origin

Paula Tejera, Nuala Meyer, Feng Chen, Rui Feng, Yang Zhao, D. Shane O'Mahony, Lin Li, Chau-Chyun Sheu, Rihong Zhai, Zhaoxi Wang, Li Su, Ed Bajwa, Amy M. Ahasic, Peter Clardy, Michelle N. Gong, Angela J. Frank, Paul N. Lanken, B. Taylor Thompson, Jason D. Christie, Mark Wurfel, Grant O'Keefe, David C. Christiani.

PATIENTS AND METHODS

1. Study design and participants populations

We intended a multi-stage approach for our analyses(1, 2) using the most heterogeneous population (Boston/MGH Cohort) with mixed risk factors to detect potential SNPs associated with ARDS secondary to direct or indirect insults to the lung. Our results were validated in two separated replication populations with extrapulmonary (Population I: UW trauma cohort) or pulmonary injury (Population II, ARDSnet/MGH: pneumonia/pulmonary sepsis) as causes for ARDS/ALI. We used a third replication population (HUP trauma cohort) to confirm associations with trauma-associated ALI (supplementary Figure 1, study overview).

1.1 Discovery phase population

All subjects used in the initial phase of the study were recruited from adult intensive care units (ICUs) at the Massachusetts General Hospital (MGH, Boston, MA) from September 1999 to March 2009 and the Beth Israel Deaconess Medical Center (BIDMC, Boston, MA) from January 2007 to February 2009 as part of the Molecular Epidemiology of ARDS Study. Details of the study-design have been described previously.(3) Briefly, patients admitted to the ICUs with pre-depositing conditions for ARDS, including bacteremia, sepsis, pneumonia, trauma, aspiration, or multiple transfusions as defined previously(4) and without any of the exclusion criteria (age < 18, diffuse alveolar hemorrhage, chronic lung diseases other than chronic obstructive pulmonary disease or asthma, directive to withhold intubation, immunosupression not secondary to corticosteroid, and treatment with granulocyte colony-stimulating factor) were enrolled and followed daily for the development of ARDS, as defined by the American-European Consensus Committee (AECC) criteria for ARDS.(5) At-risk patients who did not meet the criteria for ARDS during the ICU hospitalization were classified as non-ARDS.

1.1.1 Data collection and pulmonary and extrapulmonary injury definitions

Patient demographic information and baseline clinical characteristics were recorded upon enrollment. Acute Physiology Age and Chronic Health Evaluation (APACHE) III scores were calculated based on the data within the first 24 hrs of ICU admission.

Patients with pneumonia, aspiration, localized pulmonary contusion, or sepsis and/or bacteremia from pulmonary sources as their risk factor for ARDS were categorized as having pulmonary injury. Patients with extrapulmonary injury were those with trauma (other than pulmonary contusion), multiple transfusion or sepsis and/or bacteremia originating from the abdomen or other extrapulmonary sources. Causes of ARDS were determined by the treating physicians upon ARDS diagnosis. The classification of patients into two categories of lung injury was retrospectively and independently made by two investigators according to causes of ARDS. Patients with both types of lung injury were excluded from the study. We further restricted analysis to Caucasians (> 90% of the study subjects). The flowchart of study design in the discovery population is illustrated in supplementary Figure 2.

1.2 Replication populations

Stage II of our study consisted of two independent replication populations. Population I consisted of patients admitted to the Harborview Medical Center (HMC, Seattle, Washington) ICU for 48 hours or longer following major trauma without isolated traumatic brain injury, burn injury, or a perceived low probability of survival due to the injury.(6) Cases and controls from this population were shared with the Trauma-associated ALI SNP Consortium (TASC), a multicenter effort to perform a genome-wide association study (GWAS) of trauma-associated ALI.(6, 7) The imbalance of cases to controls in this population (597 cases and 168 non-ARDS) resulted from the TASC design, which utilized at-risk controls only in the replication phase.(6)

Subjects were followed until hospital discharge or death, clinical information was abstracted from the electronic medical record, and ARDS/ALI determination was made according to AECC criteria.(5) Population II consisted of ARDS cases collected as part of the Fluid and Catheter Treatment Trial (FACTT), Albuterol for the treatment of ALI (ALTA) and EDEN-Omega trials conducted by the NHLBI ARDS Network (http://www.ardsnet.org/clinicians/studies). Patients were eligible for inclusion in our study if they have pneumonia/ sepsis from pulmonary sources as a risk factor for ARDS. Controls in this population were non-ARDS patients with pulmonary injury from the discovery set.

Stage III consisted of subjects admitted to the surgical ICU of Hospital of the University of Pennsylvania after a major trauma and with an injury severity score (ISS) $\geq$ 16, corresponding to severe trauma.(8) Isolated brain injury or pre-existing lung disease were major exclusion criteria. Details of this cohort have been published.(9, 10) To be classified as ALI case, subjects have to meet all AECC definition criteria within a 24-hour period while tracheally intubated and mechanically ventilated.(5)

At each site, the Institutional Review Board and/or Human Subjects Committee reviewed and approved the study. For the Stage I and Stage II (Population II) of the study signed informed-consent was obtained from all study participants or their appropriate surrogates. Stage II (Population I) and Stage III were granted waiver of informed consent in accordance with institutional and federal regulations.

2. ***Genotyping strategy and quality control***

Genomic DNA from patients in the discovery population was extracted from whole blood sample, using the Puregene DNA Isolation Kit (Gentra Systems, Minneapolis, MN) or the Autopure LS workstation and Autopure reagents (Qiagen, Valencia, CA) and normalized to a

concentration of 50-100 ng/µl. For the replication populations (Stage II and III) DNA was isolated using the Qiagen QIAamp DNA Blood Midi Kit or the Qiagen Qiamp 96 blood kit (Qiagen™, Valencia, CA).

2.1 Stage I (MGH, IBC Chip): Genotyping was attempted on 2395 patients with available high-quality DNA samples, as determined by optical density spectrophotometry and agarose gel electrophoresis. Samples were genotyped at the Center for Applied Genomics, Children's Hospital of Philadelphia (Philadelphia, PA), using a custom SNP array designed by the Institute of Translational Medicine and Therapeutics, the Broad Institute, and the National Heart Lung and Blood Institute supported Candidate gene Association Resource Consortium (**I**TMAT-**B**road_**C**ARe (IBC) array; Illumina®, San Diego, CA).(11) Non-synonymous SNPs with a MAF > 0.01 and tagging SNPs with MAF > 0.02 located in regions of functional significance were selected to be included in the array. A list of genes and SNPs on the array is available at http://bmic.upenn.edu/cvdsnp. The chip included 48,742 markers in ~2,100 potentially relevant loci related to cardiovascular, metabolic, and inflammatory syndromes. As a candidate gene chip designed to capture variation in loci important to inflammatory and vascular phenotypes the IBC chip also includes many genes with plausible role in ALI development. This platform was chosen based on its design that allows saturation of these regions with a density greater than afforded by genome-wide genotyping platforms. Most of candidate genes previously associated with ALI/ARDS(7, 12-15) are present on the array and recent associations with ALI phenotype has been reported using this platform.(7) Genotyping was carried out by laboratory personnel without the knowledge of case-control status. Quality control measures were conducted using the software package PLINK version 1.06 (http://pngu.mgh.harvard.edu/~purcell/plink/).(16) For samples quality control samples with call rate ≤ 95%, with missing clinical information or with

previous enrollment or history of ARDS, and non-Caucasian were excluded from analysis. For SNP quality control we removed markers with genotyping call rate $< 95\%$, those that were non-autosomal, had a MAF < 0.05, or were deviated from Hardy-Weinberg equilibrium (HWE) in the control sample ($p < 0.001$) (supplementary Figure 2 and supplementary Table 1).

2.2 Stage II (Population I: UW trauma cohort and Population II: ARDSnet/MGH: pneumonia/pulmonary sepsis) (I6Q Genome Wide Platform)**:* All DNA was shipped to Center for Applied Genomics (CAG) at CHOP and genotyped using the Infinium™ II HumanHap610K-quad BeadChip (I6Q) Illumina™, San Diego, CA)(17, 18) by lab personnel unaware of the case status of each sample. For quality control only samples with genotyping call rate $\geq 95\%$ and SNPs with genotyping call rate $\geq 95\%$; HWE p-value $\geq$ 10E-04; and MAF ≥ 0.01 were included in the analyses (supplementary Table 1). For those IBC SNPs not directly typed on the genome wide array, imputation was carried out using MACH v 3.0.(19) run with default settings and 50 iterations of the Markov sampler, using the haplotypes and the snp files from and 1000 Genomes European ancestry samples as reference panel. To make the computation feasible, we limited the imputation to reference haplotypes chopped the haplotypes ~500kb up and downstream of the genes of interest. ***2.3 Stage III (HUP trauma cohort, IBC Chip): Subjects in Stage III were genotyped with the IBC Chip. Analysis was restricted to those samples with genotyping call rate $\geq 95\%$ and only to autosomal SNPs with genotyping call rate $\geq 95\%$; HWE p-value $\geq$ 10E-04; and MAF ≥ 0.01 (supplementary Table 1).

3. Statistical analysis

All statistical analyses were performed using PLINK (http://pngu.mgh.harvard.edu/purcell/plink/)(16) v1.07, R version 2.12 (http://www.r-project.org), and SAS® v9.1.3. statistical software package (SAS Institute, Inc., Cary, NC).

Demographic variables were compared between ARDS/ALI patients and controls using Fisher's exact test for categorical variables and Student's *t*-test for continuous variables. We used multivariate logistic regression to estimate the genotype specific odds ratio (OR) and 95% confidence interval (CI) for ARDS/ALI susceptibility, as implemented in PLINK. Significance of odds ratios was determined using the χ^2 test. The genotype associations were analyzed in additive models adjusting for clinical covariates available at each stage.

3.1 Population Stratification and Genetic Determination of Ancestry

To minimize the risk of confounding due to ethnic differences, analyses in each stage were restricted to subjects of European ancestry. In Stage I population stratification was analyzed using the quantile-quantile (Q-Q) plot. The Q-Q plots were used to validate the observed associations, compared with the expectations under the null distribution that assumes no association, potential population stratification or genotyping errors. We also computed the genomic control (λ), defined as the median of the observed 1-df chi-squared association statistics divided by its theoretical median under null distribution.(20) Values of $\lambda < 1.05$ are considered benign.(21) In Stage II (Population I and II), population stratification was assessed by using principal components analysis (PCA). In Population I, reported ethnicity was screened using the STRUCTURE package(22) and over 200 ancestry informative markers (AIMs) to cluster the TASC submissions with 90 HapMap individuals (CEU, Yoruban, and Chinese/Japanese). Samples with an inferred proportion of CEU ancestry < 90% were determined to be non-European American and excluded from Stage II. In population II, PC were calculated by EIGENSOFT 4.2 using the SNPs on the Illumina 610 chip (http://www.hsph.harvard.edu/faculty/alkes-price/software/). Procedures follow those described in.(23) PCA analysis was carried out on the 838 subjects and the genome-wide SNPs on the

Illumina 610 chip, and top 6 principal components were chosen to be used in following association analyses. In Stage III ancestry was inferred by multidimensional scaling (MDS) using all markers on the IBC chip as enacted in PLINK and results were adjusted for 2 principal components from MDS as described previously.(7, 16, 24)

3.2 Selection of threshold of significance for the selection of SNPs at each stage.

In Stage I, SNPs were selected according to a pre-specified p-value $\leq 5x10^{-4}$. A significance threshold of 10^{-6} resulting after a strict Bonferroni correction for multiple testing (0.05/50,000 SNPs on the array) would be overly conservative considering that association tests are not completely independent due to the high degree of LD among the SNPs. Furthermore, SNPs on the array were selected to densely cover loci of established relevance to disease.(11) The statistical threshold to declare significance when using a dense, hypothesis-driven candidate gene SNP array is uncertain.(24) Previous studies using this platform have adopted a cut-off thresholds ranging from 10^{-4} to 10^{-6} .(7, 25-27) We adopted a pre-specified p-value $\leq 5x10^{-4}$ in order to reach satisfactory power for our cohort (see below).The significance of the associations observed in Stage I was then established by independent replication of our findings in the Stage II and III of the study. To be considered as replicated the direction of the association was required to be the same as in Stage I and with a p < (0.05/ number of SNPs tested).

3.3 Power Calculations

In Stage I (discovery phase), our extrapulmonary subpopulation with 290 cases and 549 controls yielded greater than 80% power to detect a minimum relative risk (RR) of 1.65 at an alpha level of 0.0005 for variants with a MAF $\geq$ 0.2. Variants with a MAF $\geq$ 0.1 would be detected at the same significant level only with a RR $\geq$ 1.9. For the association with pulmonary injury-related ARDS, based on a sample size of 127 cases and 751 controls we determined that

we would have greater than 80% power to detect RR $\geq$ 1.95 at an alpha level of 0.0005 for MAF $\geq$ 0.2. Variants with a MAF $\geq$ 0.1 would be detected with a RR $\geq$ 2.25. In Stage II, Population I (Harborview Trauma Cohort) with 597 cases and 168 controls yielded greater than 80% power to detect a minimum relative risk (RR) of 1.65 at an alpha level of 0.006 (0.05/8 SNPs tested) for variants with a MAF $\geq$ 0.2. Variants with a MAF $\geq$ 0.1 would be detected at the same significant level only with a RR $\geq$ 1.9.Population II (MGH/ARDS net, pneumonia/pulmonary sepsis Cohort) yielded greater than 80% power to detect a minimum relative risk (RR) of 1.6 at an alpha level of 0.003 (0.05/17 SNPs tested) for variants with a MAF $\geq$ 0.2. Variants with a MAF $\geq$ 0.1 would be detected at the same significant level only with a RR $\geq$ 1.95. In Stage III (Penn Trauma Cohort) with 74 cases and 150 controls yielded greater than 80% power to detect a minimum relative risk (RR) of 1.9 at an alpha level of 0.05 for variants with a MAF $\geq$ 0.2.(28)

3.4 Sensitivity Analysis

A sensitivity analysis was performed to assess the robustness of the replication results to the reclassification of ALI subjects into ARDS cases in Population I (Stage II). Approximately 70% (413 of 597) of ALI cases in the replication Population I also met the criteria for ARDS. To see whether there were differences in the magnitude and direction of the association of rs324420 (*FAAH* gene) with clinical phenotype, the association analyses were repeated after recategorizing ALI cases (defined as $PaO_2:FiO_2 < 300$ mmHg) into ARDS ($PaO_2:FiO_2 < 200$ mmHg).(5) The association of rs324420 with ARDS observed in Stage I and replicated in Stage II for ALI phenotype remained stable after reclassification of ALI subjects into ARDS cases (supplementary Table 3).

3.5 Multi-SNP genotypic risk score estimation

In order to know if the genetic factors associated with extrapulmonary injury-related

ARDS also influences the development of ARDS from pulmonary sources we assigned a multi-SNP genotypic risk score (MSRS) based on the SNPs significantly associated with risk of developing ARDS ($p \leq 0.0005$) from direct or indirect lung injury in Stage I of our study. Two different methods can be used in the estimation of MSRS. The first method known as the "count method", sums the total number of risk alleles each individual carries. The second method referred as the "log odds method" sums together the natural logarithm of the allelic odds ratio for each risk allele.(29) Little differences in discriminative accuracy has been found when MSRS is constructed counting the number of risk genotypes or by calculating the associated disease risk .(29,30) For this reason, we decided to adopt the count method in the calculation of our MSRSs. The allele-counting method assumed equal and additive effects of the individual variants (29). For each patient, we first calculated the sum across SNPs of the number of risk alleles at each SNP (genotypes were coded as 0, 1, and 2). For SNPs in high LD, we keep the one with smaller p-value.

The risk scores for the SNPs associated with the development of ARDS from extrapulmonary and pulmonary sources (RSEXP and RSP, respectively) were calculated as showed in (1) and (2):

(1) RSEXP = rs198977+rs9645765+rs2889490+rs3128126+rs16980496+rs10490072+rs324420. The mean for the RSEXP was 3.37 (SD = 1.58).

(2) RSP = rs7807769+2-rs1190286+rs9960450+2-rs2254358+rs732821+rs3887893+2-rs2671222+rs12080701+rs8112223+rs6451620+rs17419964+rs802440+2-rs4075731. For the RSP mean was 11.26 (SD = 2.36)

Scores were normalized by subtracting the mean and dividing by the standard error. The association between normalized PRS and the development of ARDS was next analyzed in the

extrapulmonary and pulmonary subpopulations by standard logistic regression including age, gender and APACHE III score as covariates.

In the extrapulmonary subpopulation, the RSEXP was significantly associated with the development of ARDS (OR: 1.81(1.58~2.07), P<1E-15), but not in the pulmonary subpopulation (OR: 0.99(0.90~1.09),P=0.8613). On the other hand, RSP was significantly associated with ARDS in the pulmonary subpopulation (OR: 1.57(1.45~1.71), P<1E-15), but not in the in the extrapulmonary subgroup (OR: 1.06(0.97~1.15), P=0.2345).

Our analyses suggest nonexistence of shared risk factors contributing to the development of ARDS from extrapulmonary or pulmonary sources. However, it is possible that genetic variants with *smaller* effects might modulate the risk of developing ARDS from both pulmonary and extrapulmonary sources. Such variants were not detected since our study was not designed to evaluate more modest effects size or rarer variants. This fact, however, does not invalidate our findings.

4. *Validation of prior genetic associations with ARDS risk*

Earlier findings by our group reporting genetic association with ARDS susceptibility(31-37) were not validated in Stage I of our study. There are several reasons for these conflicting results. First, due to the IBC chip's design, there is important genetic variation as microsatellites and in/del, previously associated with ARDS(31, 32) that we did not detect. Second, the chip also has a very limited coverage (50K SNPs), and did not include some of the genes(33) or SNPs (34-37) previously associated with the development of ARDS. Finally, unlike our prior studies, in the current study the genetic risk of ARDS is analyzed by carrying out a stratified analysis by the type of lung injury. This approach might reduce the statistical power of our study to detect previous associations.

REFERENCES

1. **Satagopan JM**, Elston RC. Optimal two-stage genotyping in population-based association studies. *Genet Epidemiol* 2003;**25**(2):149-57.

2. **Hirschhorn JN**, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005;**6**(2):95-108.

3. **Gong MN**, Thompson BT, Williams P, Pothier L, Boyce PD, Christiani DC. Clinical predictors of and mortality in acute respiratory distress syndrome: potential role of red cell transfusion. *Crit Care Med* 2005;**33**(6):1191-8.

4. **Tejera P**, Wang Z, Zhai R, Su L, Sheu CC, Taylor DM, Chen F, Gong MN, Thompson BT, Christiani DC. Genetic polymorphisms of peptidase inhibitor 3 (elafin) are associated with acute respiratory distress syndrome. *Am J Respir Cell Mol Biol*. 2009;**41**(6):696-704.

5. **Bernard GR**, Artigas A, Brigham KL, Carlet J, Falke K, Hudson L, Lamy M, Legall JR, Morris A, Spragg R. The American-European Consensus Conference on ARDS. Definitions, mechanisms, relevant outcomes, and clinical trial coordination. *Am J Respir Crit Care Med* 1994 ;**149**(3 Pt 1):818-24.

6. **Christie JD**, Wurfel MM, Feng R, O'Keefe GE, Bradfield J, Ware LB, Christiani DC, Calfee CS, Cohen MJ, Matthay M, Meyer NJ, Kim C, Li M, Akey J, Barnes KC, Sevransky J, Lanken PN, May AK, Aplenc R, Maloney JP, Hakonarson H; Trauma ALI SNP Consortium (TASC) investigators. Genome wide association identifies PPFIA1 as a candidate gene for acute lung injury risk following major trauma. *PLoS One* 2012;**7**(1):e28268.

7. **Meyer NJ**, Li M, Feng R, Bradfield J, Gallop R, Bellamy S, Fuchs BD, Lanken PN, Albelda SM, Rushefski M, Aplenc R, Abramova H, Atochina-Vasserman EN, Beers MF, Calfee CS, Cohen MJ, Pittet JF, Christiani DC, O'Keefe GE, Ware LB, May AK, Wurfel MM,

Hakonarson H, Christie JD. ANGPT2 Genetic Variant Is Associated with Trauma-associated Acute Lung Injury and Altered Plasma Angiopoietin-2 Isoform Ratio. *Am J Respir Crit Care Med* 2011;**183**(10):1344-53.

8. **Civil ID**, Schwab CW. The Abbreviated Injury Scale, 1985 revision: a condensed chart for clinical use. *J Trauma* 1988;**28**(1):87-90.

9. **Marzec JM**, Christie JD, Reddy SP, Jedlicka AE, Vuong H, Lanken PN, Aplenc R, Yamamoto T, Yamamoto M, Cho HY, Kleeberger SR. Functional polymorphisms in the transcription factor NRF2 in humans increase the risk of acute lung injury. *FASEB J* 2007 ;**21**(9):2237-46.

10. **Christie JD**, Ma SF, Aplenc R, Li M, Lanken PN, Shah CV, Fuchs B, Albelda SM, Flores C, Garcia JG. Variation in the myosin light chain kinase gene is associated with development of acute lung injury after major trauma. *Crit Care Med* 2008;**36**(10):2794-800.

11. **Keating BJ**, Tischfield S, Murray SS, Bhangale T, Price TS, Glessner JT, Galver L, Barrett JC, Grant SF, Farlow DN, Chandrupatla HR, Hansen M, Ajmal S, Papanicolaou GJ, Guo Y, Li M, Derohannessian S, de Bakker PI, Bailey SD, Montpetit A, Edmondson AC, Taylor K, Gai X, Wang SS, Fornage M, Shaikh T, Groop L, Boehnke M, Hall AS, Hattersley AT, Frackelton E, Patterson N, Chiang CW, Kim CE, Fabsitz RR, Ouwehand W, Price AL, Munroe P, Caulfield M, Drake T, Boerwinkle E, Reich D, Whitehead AS, Cappola TP, Samani NJ, Lusis AJ, Schadt E, Wilson JG, Koenig W, McCarthy MI, Kathiresan S, Gabriel SB, Hakonarson H, Anand SS, Reilly M, Engert JC, Nickerson DA, Rader DJ, Hirschhorn JN, Fitzgerald GA. Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS One*. 2008;**3**(10):e3583.

12. **Gao L**, Barnes KC. Recent advances in genetic predisposition to clinical acute lung

injury. *Am J Physiol Lung Cell Mol Physiol* 2009;**296**(5):L713-25.

13. **Flores C**, Pino-Yanes MM, Casula M, Villar J. Genetics of acute lung injury: past, present and future. *Minerva Anestesiol*. 2010;**76**(10):860-4.

14. **Glavan BJ**, Holden TD, Goss CH, Black RA, Neff MJ, Nathens AB, Martin TR, Wurfel MM. Genetic variation in the FAS gene and associations with acute lung injury. *Am J Respir Crit Care Med*. 2011;**183**(3):356-63.

15. **Wurfel MM**, Gordon AC, Holden TD, Radella F, Strout J, Kajikawa O, Ruzinski JT, Rona G, Black RA, Stratton S, Jarvik GP, Hajjar AM, Nickerson DA, Rieder M, Sevransky J, Maloney JP, Moss M, Martin G, Shanholtz C, Garcia JG, Gao L, Brower R, Barnes KC, Walley KR, Russell JA, Martin TR. Toll-like receptor 1 polymorphisms affect innate immune responses and outcomes in sepsis. *Am J Respir Crit Care Med* 2008;**178**(7):710-20.

16. **Purcell S**, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;**81**(3):559-75.

17. **Steemers FJ**, Chang W, Lee G, Barker DL, Shen R, Gunderson KL. Whole-genome genotyping with the single-base extension assay. *Nat Methods* 2006;**3**(1):31-3.

18. **Zhang FR**, Huang W, Chen SM, Sun LD, Liu H, Li Y, Cui Y, Yan XX, Yang HT, Yang RD, Chu TS, Zhang C, Zhang L, Han JW, Yu GQ, Quan C, Yu YX, Zhang Z, Shi BQ, Zhang LH, Cheng H, Wang CY, Lin Y, Zheng HF, Fu XA, Zuo XB, Wang Q, Long H, Sun YP, Cheng YL, Tian HQ, Zhou FS, Liu HX, Lu WS, He SM, Du WL, Shen M, Jin QY, Wang Y, Low HQ, Erwin T, Yang NH, Li JY, Zhao X, Jiao YL, Mao LG, Yin G, Jiang ZX, Wang XD, Yu JP, Hu ZH, Gong CH, Liu YQ, Liu RY, Wang DM, Wei D, Liu JX, Cao WK, Cao HZ, Li YP, Yan WG, Wei SY, Wang KJ, Hibberd ML, Yang S, Zhang XJ, Liu JJ. Genomewide association study of leprosy. N

Engl J Med 2009;**361**(27):2609-18.

19. **Li Y**, Willer C, Sanna S, Abecasis G. Genotype imputation. *Annu Rev Genomics Hum Genet* 2009;**10:**387-406.

20. **Reich DE**, Goldstein DB. Detecting association in a case-control study while correcting for population stratification. *Genet Epidemiol* 2001;**20**(1):4-16.

21. **Price AL**, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet* 2010;**11**(7):459-63.

22. **Falush D**, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Mol Ecol Notes* 2007;**7**(4):574-8.

23. **Anderson CA**, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. Data quality control in genetic case-control association studies. *Nat Protoc* 2010;**5**(9):1564-73.

24. **Cappola TP**, Li M, He J, Ky B, Gilmore J, Qu L, Keating B, Reilly M, Kim CE, Glessner J, Frackelton E, Hakonarson H, Syed F, Hindes A, Matkovich SJ, Cresci S, Dorn GW, 2nd. Common variants in HSPB7 and FRMD4B associated with advanced heart failure. *Circ Cardiovasc Genet* 2010;**3**(2):147-54.

25. **Lanktree MB**, Guo Y, Murtaza M, Glessner JT, Bailey SD, Onland-Moret NC, Lettre G, Ongen H, Rajagopalan R, Johnson T, Shen H, Nelson CP, Klopp N, Baumert J, Padmanabhan S, Pankratz N, Pankow JS, Shah S, Taylor K, Barnard J, Peters BJ, Maloney CM, Lobmeyer MT, Stanton A, Zafarmand MH, Romaine SP, Mehta A, van Iperen EP, Gong Y, Price TS, Smith EN, Kim CE, Li YR, Asselbergs FW, Atwood LD, Bailey KM, Bhatt D, Bauer F, Behr ER, Bhangale T, Boer JM, Boehm BO, Bradfield JP, Brown M, Braund PS, Burton PR, Carty C, Chandrupatla HR, Chen W, Connell J, Dalgeorgou C, Boer A, Drenos F, Elbers CC, Fang JC, Fox CS, Frackelton EC, Fuchs B, Furlong CE, Gibson Q, Gieger C, Goel A, Grobbee DE, Hastie C,

Howard PJ, Huang GH, Johnson WC, Li Q, Kleber ME, Klein BE, Klein R, Kooperberg C, Ky B, Lacroix A, Lanken P, Lathrop M, Li M, Marshall V, Melander O, Mentch FD, Meyer NJ, Monda KL, Montpetit A, Murugesan G, Nakayama K, Nondahl D, Onipinla A, Rafelt S, Newhouse SJ, Otieno FG, Patel SR, Putt ME, Rodriguez S, Safa RN, Sawyer DB, Schreiner PJ, Simpson C, Sivapalaratnam S, Srinivasan SR, Suver C, Swergold G, Sweitzer NK, Thomas KA, Thorand B, Timpson NJ, Tischfield S, Tobin M, Tomaszweski M, Verschuren WM, Wallace C, Winkelmann B, Zhang H, Zheng D, Zhang L, Zmuda JM, Clarke R, Balmforth AJ, Danesh J, Day IN, Schork NJ, de Bakker PI, Delles C, Duggan D, Hingorani AD, Hirschhorn JN, Hofker MH, Humphries SE, Kivimaki M, Lawlor DA, Kottke-Marchant K, Mega JL, Mitchell BD, Morrow DA, Palmen J, Redline S, Shields DC, Shuldiner AR, Sleiman PM, Smith GD, Farrall M, Jamshidi Y, Christiani DC, Casas JP, Hall AS, Doevendans PA, Christie JD, Berenson GS, Murray SS, Illig T, Dorn GW, 2nd, Cappola TP, Boerwinkle E, Sever P, Rader DJ, Reilly MP, Caulfield M, Talmud PJ, Topol E, Engert JC, Wang K, Dominiczak A, Hamsten A, Curtis SP, Silverstein RL, Lange LA, Sabatine MS, Trip M, Saleheen D, Peden JF, Cruickshanks KJ, Marz W, O'Connell JR, Klungel OH, Wijmenga C, Maitland-van der Zee AH, Schadt EE, Johnson JA, Jarvik GP, Papanicolaou GJ, Grant SF, Munroe PB, North KE, Samani NJ, Koenig W, Gaunt TR, Anand SS, van der Schouw YT, Soranzo N, Fitzgerald GA, Reiner A, Hegele RA, Hakonarson H, Keating BJ. Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height. *Am J Hum Genet* 2011;**88**(1):6-18.

26. **Talmud PJ**, Drenos F, Shah S, Shah T, Palmen J, Verzilli C, Gaunt TR, Pallas J, Lovering R, Li K, Casas JP, Sofat R, Kumari M, Rodriguez S, Johnson T, Newhouse SJ, Dominiczak A, Samani NJ, Caulfield M, Sever P, Stanton A, Shields DC, Padmanabhan S, Melander O, Hastie C, Delles C, Ebrahim S, Marmot MG, Smith GD, Lawlor DA, Munroe PB, Day IN, Kivimaki M,

Whittaker J, Humphries SE, Hingorani AD. Gene-centric association signals for lipids and apolipoproteins identified via the HumanCVD BeadChip. *Am J Hum Genet* 2009;**85**(5):628-42.

27. **Clarke R**, Peden JF, Hopewell JC, Kyriakou T, Goel A, Heath SC, Parish S, Barlera S, Franzosi MG, Rust S, Bennett D, Silveira A, Malarstig A, Green FR, Lathrop M, Gigante B, Leander K, de Faire U, Seedorf U, Hamsten A, Collins R, Watkins H, Farrall M. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med* 2009;**361**(26):2518-28.

28. **Menashe I**, Rosenberg PS, Chen BE. PGA: power calculator for case-control genetic association analyses. *BMC Genet* 2008;**9**:36.

29. **Evans DM**, Visscher PM, Wray NR. Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. *Hum Mol Genet* 2009;**18**(18):3525-31.

30. **Janssens AC**, Moonesinghe R, Yang Q, Steyerberg EW, van Duijn CM, Khoury MJ. The impact of genotype frequencies on the clinical validity of genomic profiling for predicting common chronic diseases. *Genet Med* 200;**9**:258-35.

31. **Bajwa EK**, Cremer PC, Gong MN, Zhai R, Su L, Thompson BT, Christiani DC. An NFKB1 promoter insertion/deletion polymorphism influences risk and outcome in acute respiratory distress syndrome among Caucasians. *PLoS One* 2011;**6**(5):e19469.

32. **Sheu CC**, Zhai R, Wang Z, Gong MN, Tejera P, Chen F, Su L, Thompson BT, Christiani DC. Heme oxygenase-1 microsatellite polymorphism and haplotypes are associated with the development of acute respiratory distress syndrome. *Intensive Care Med* 2009;**35**(8):1343-51.

33. **Tejera P**, Wang Z, Zhai R, Su L, Sheu CC, Taylor DM, Chen F, Gong MN, Thompson BT, Christiani DC. Genetic polymorphisms of peptidase inhibitor 3 (elafin) are associated with acute respiratory distress syndrome. Am J Respir Cell Mol Biol. 2009;**41**(6):696-704.

34. **Bajwa EK**, Yu CL, Gong MN, Thompson BT, Christiani DC. Pre-B-cell colony-enhancing factor gene polymorphisms and risk of acute respiratory distress syndrome. *Crit Care Med* 2007;**35**(5):1290-5.

35. **Zhai R**, Gong MN, Zhou W, Thompson TB, Kraft P, Su L, Christiani DC. Genotypes and haplotypes of the VEGF gene are associated with higher mortality and lower VEGF plasma levels in patients with ARDS. *Thorax* 2007;**62**(8):718-22.

36. **Gong MN**, Zhou W, Williams PL, Thompson BT, Pothier L, Christiani DC. Polymorphisms in the mannose binding lectin-2 gene and acute respiratory distress syndrome. *Crit Care Med* 2007;**35**(1):48-56.

37. **Gong MN**, Thompson BT, Williams PL, Zhou W, Wang MZ, Pothier L, Christiani DC. Interleukin-10 polymorphism in position -1082 and acute respiratory distress syndrome. *Eur Respir J* 2006;**27**(4):674-81.

38. **Johnson AD**, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;**24**(24):2938-9.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1. Study Overview. We implemented a multi-stage design in our study. In Stage I, 1,717 critical ill patients (1,300 controls and 417 cases) in the discovery set were classified as having pulmonary or extrapulmonary-injury related ARDS and genotyped using the ITMAT-Broad_CARe (IBC) array (Illumina®, San Diego, CA). SNPs associated with development of pulmonary and extrapulmonary injury-related ARDS at $p \leq$ 5E-04 were tested for validation in Stage II of the study using two separated populations, with extrapulmonary (Population I: UW trauma cohort) or pulmonary (Population II, ARDSnet/MGH: pneumonia/pulmonary sepsis) injury as causes for ARDS/ALI. Results were adjusted for clinical covariates available at each site. SNP rs1190286 in *POPODC3* gene was the only SNP replicating its association with the development of ARDS from pulmonary causes in Stage II of our study. Among top SNPs associated with risk of ARDS from extrapulmonary origin, SNP rs324420 replicated its association with development of trauma related-ALI cohort in Stage II and StageIII.

Supplementary Figure 2. Study Design and Patient Selection, Stage I. Among the 4148 patients eligible for enrollment, 2786 patients with informed consent were enrolled into a prospective cohort. Genotyping was attempted on 2395 patients with available high-quality DNA. Samples with call rate ≤ 95% (n = 101), with missing clinical information or with previous enrollment or history of ARDS (n = 22), and non-Caucasian (n = 209), were excluded from analysis. The total genotyping rate in remaining individuals was 99.84 %. After quality control 2,063 subjects were left for further analyses. 1,717 Caucasian critically

ill patients at risk for ARDS and with only a type of lung injury were included in the first stage of the study. Among them, 417 were ARDS cases (127 with pulmonary injury and 290 with extrapulmonary injury) and 1,300 controls (751 with pulmonary injury and 549 with extrapulmonary injury).

Supplementary Figure 3. Manhattan plot for Stage I association study showing 29,483 markers distributed across chromosomes. Data are –log transformed p-values from association analyses with pulmonary (**A**) and extrapulmonary (**B**) injury-related ARDS, under an additive model.

Supplementary Figure 4. Quantile-Quantile (Q-Q) plots of Stage I genetic association results with ARDS from pulmonary (A) and extrapulmonary (B) origin. Observed p-values are plotted against expected p-values. The slope line shows the distribution of p-values under the null hypothesis of no association at any locus. For both the pulmonary (left) and extrapulmonary (right) groups that compose the discovery set, the observed p values match reasonably well with the expected values, suggesting that our associations with pulmonary/extrapulmonary injury-related ARDS in the discovery population were more likely due to true genetic variation than other reason like population stratification or genotyping bias. The calculated λ values for the pulmonary and extrapulmonary groups (λ = 1.000 and λ = 1.018, respectively) also suggested that our association results were not confounded by population stratification.

Supplementary Figure 5. Regional association plot of *FAAH* for its association with trauma associated ALI in Stage II. The scatter graph indicates the negative logarithm of p-value (additive model) for each SNP. Markers above the dashed blue line were significantly associated with ALI with an alpha the significance, $p < 0.0005/8= 0.006$. The color scheme of a white to red gradient reflects lower to higher linkage disequilibrium (LD) values (r^2) with rs324420. Blue line reflects the global genome recombination rate for this region of the genome which was downloaded from HapMap. SNP rs324420 in exon 3 of FAAH was the most significant SNP associated with ALI. SNP rs324418 (intronic) in high LD with rs324420, also showed a significant association (impute $p = 0.00552$) with increased risk of trauma-related ALI.

Supplementary Figure 6. Regional association plot of *POPDC3* for its association with trauma associated ARDS in Stage II. The y axis is the negative logarithm of p-value (additive model) for each SNP for and additive model of association with ARDS. Markers above the dashed blue line (rs1190267, rs1190272, rs1190293, rs1190294, rs1190295, rs1190297, rs1190298, rs4946659, rs9399904, rs9373786, rs4945717, rs1190271, rs1190286, rs1190290, rs1190292, rs7749438 and rs6571221), were significantly associated with ALI with an alpha the significance, $p < 0.0005/17= 0.003$. Each diamond show an SNP with a color scale relating the r2 value for that SNP and the top SNP, with increase red intensity reflecting greater degree of linkage disequilibrium. Blue lines indicated estimated recombination rate from HapMap.

Supplementary Table 1 *Filtering Criterion Used for SNPs Quality Control in Stage I, II and III*

Stage I		
Filtering Criterion	**SNPs**	**Remaining SNPs of 48,742 on IBC Chip**
Non-autosomal SNPs	1121	47,621
HWE p-value < 10E-03	2403	45,218
Genotype call rate < 95%	505	44,713
MAF < 0.05	15,228	29,485
Duplicate SNPs	2	29,483
Stage II		
Filtering Criterion	**SNPs**	**Remaining SNPs of 620,091 on 610-quad**
SNP failure, monomorphic, genotype call rate < 95%, or MAF < 0.01	159,098	530,993
HWE p-value < 10E-04	534	530,459
Stage III		
Filtering Criterion	**SNPs**	**Remaining SNPs of 48,742 on IBC Chip**
Monomorphic	1820	46,922
Genotype call rate < 95%	850	46,072
HWE p-value < 10E-04	39	46,033
MAF < 0.01	585	45,448

***Definition of abbreviations*;** SNPs: Single Nucleotide Polymorphisms; HWE = Hardy-Weinberg Equilibrium; MAF = Minor Allele Frequency.

Supplementary Table 2 ***SNPs associated with extrapulmonary and pulmonary injury-related ARDS in Stage I***

							Association with extrapulmonary injury-related ARDS		**Association with pulmonary injury-related ARDS**	
Chr	**SNP**	**Gene**	**Location**	**Allele**	**MAF Case/Ctrl**	**HWE**	**OR (95% CI)**	**P^a (additive)**	**OR (95% CI)**	**P^a (additive)**
19	rs198977	KLK2	Exon	C>T	0.38/0.22	0.25	**1.74 (1.23-2.32)**	**0.00021**	0.96 (0.75-1.22)	0.7162
12	rs9645765	VWF	Intron	A>G	0.14/0.07	1	**2.17 (1.43-3.28)**	**0.000276**	1.20 (0.82-1.74)	0.3593
19	rs2889490	SFRS16	Intron	A>G	0.58/0.46	0.71	**1.67 (1.26-2.19)**	**0.000276**	0.99 (0.81-1.22)	0.9636
1	rs3128126	ISG15	Intron	A>G	0.48/0.36	0.032	**1.70 (1.28-2.26)**	**0.000278**	1.04 (0.84-1.28)	0.7467
22	rs16980496	ADRBK2	Intron	A>G	0.13/0.06	0.75	**2.23 (1.44-3.45)**	**0.00034**	0.91 (0.62-1.35)	0.6545
12	rs2070887	VWF	Intron	A>G	0.15/0.08	0.44	**2.07 (1.38-3.10)**	**0.0004**	1.14 (0.79-1.66)	0.4804
2	rs10490072	BCL11A	3' near gene	C>T	0.30/0.21	0.65	**1.72 (1.27-2.33)**	**0.000476**	0.99 (0.77-1.27)	0.9656
1	rs324420	FAAH	Exon	C>A	0.29/0.19	0.13	**1.74 (1.27-2.39)**	**0.000503**	0.92 (0.70-1.20)	0.5669
7	rs7807769	PRKAG2	Intron	C>A	0.48/0.39	0.37	1.15 (0.87-1.51)	0.3142	**1.58 (1.28-1.94)**	**1.61E-05**
7	rs7801616	PRKAG2	Intron	C>T	0.48/0.39	0.33	1.18 (0.90-1.56)	0.2302	**1.54 (1.25-1.89)**	**4.37E-05**
6	rs1190286	POPDC3	Intron	C>T	0.13/0.20	0.29	1.28 (0.91-1.80)	0.1548	**0.53 (0.39-0.72)**	**5.30E-05**
18	rs9960450	TNFRSF11A	Intron	C>T	0.08/0.03	0.38	0.99 (0.54-1.79)	0.9634	**2.48 (1.56-3.93)**	**0.000114**
1	rs2254358	HSPG2	Exon	A>C	0.25/0.33	0.63	0.88 (0.65-1.19)	0.401	**0.63 (0.50-080)**	**0.000129**
13	rs732821	HTR2A	5'near gene	G>A	0.54/0.45	1	1.17 (0.89-1.55)	0.2571	**1.52 (1.22-1.88)**	**0.000136**
7	rs6970522	PRKAG2	Intron	A>G	0.51/0.44	0.26	1.06 (0.81-1.39)	0.6365	**1.49 (1.21-1.82)**	**0.000167**
16	rs3887893	ABCC1	Intron	A>G	0.45/0.36	0.07	1.20 (0.91-1.58)	0.1956	**1.48 (1.20-1.83)**	**0.000287**
18	rs17069902	TNFRSF11A	Intron	C>T	0.09/0.04	0.05	1.13 (0.65-1.96)	0.6564	**2.12 (1.41-3.20)**	**0.000312**
2	rs2671222	IL8RA	5'near gene	A>G	0.02/0.06	1	0.72 (0.35-1.50)	0.3834	**0.34 (0.19-0.61)**	**0.000325**
1	rs12080701	PDE4B	Intron	G>A	0.14/0.09	0.30	0.88 (0.54-1.43)	0.6109	**1.85 (1.32-2.60)**	**0.000362**
19	rs8112223	HAS1	5'near gene	G>A	0.45/0.36	0.78	0.92 (0.70-1.22)	0.5037	**1.48 (1.19-1.84)**	**0.00037**
5	rs6451620	GHR	Intron	G>A	0.09/0.04	0.32	1.28 (0.75-2.17)	0.3601	**2.15 (1.41-3.29)**	**0.000372**
1	rs17419964	PDE4B	Intron	A>G	0.34/0.26	0.83	0.91 (0.68-1.25)	0.5877	**1.50 (1.20-1.88)**	**0.000433**
7	rs802440	GRM3	Intron	C>T	0.40/0.30	0.11	1.05 (0.79-1.41)	0.7087	**1.47 (1.19-1.83)**	**0.000452**
1	rs4075731	MAP3K6	Intron	A>C	0.32/0.40	1	0.97 (0.74-1.28)	0.855	**0.67 (0.54-0.84)**	**0.000468**
2	rs2854386	IL8RA	3'near gene	C>G	0.03/0.07	1	0.98 (0.51-1.89)	0.9589	**0.36 (0.20-0.64)**	**0.000476**

Definitions of abbreviations: ARDS: Acute Respiratory Distress Syndrome; Chr: Chromosome; SNP: Single Nucleotide Polymorphism; MAF: Minor Allele Frequency; HWE: Hardy-Weinberg Equilibrium; OR: Odd Ratio. *P values were adjusted for age, gender and APACHE III score in Stage I population. Shaded area shows SNPs associated with ARDS from direct or indirect lung injury in Stage I ($p \leq 0.0005$). No variant exhibited even a marginal association in both types of lung injury.

Supplementary Table 3. Sensitivity analysis of case definition on the association of SNP 324420 with ARDS in Stage II (Population I)

SNP	Gene	Odds Ratio (95% CI)	†P(additive)
rs324420[a]	FAAH	1.59 (1.10-2.31)[a]	0.0131[a]
rs324420[b]	FAAH	1.58 (1.07-2.34)[b]	0.0215[b]

Definitions of abbreviations; SNP: Single Nucleotide Polymorphism; MAF: Minor Allele Frequency; Ctrl: Control. [a]SNP was tested for the association with ALI. [b]SNP was tested for the association with ARDS. †P values were adjusted for age, ISS and APACHE II score. SNP rs324420 remains associated with ARDS in Stage II after reclassification of ALI subjects into ARDS cases.

Supplementary Ttable 4. Meta-analysis results for SNPs significantly associated with extrapulmonary injury-related ARDS in Stage I

				Discovery phase (Stage I)				***Replication phase I (Stage II)***				***Replication phase II (Stage III)***				***Meta-analysis***		
SNP	Gene	Chr	Minor allele	MAF Case Ctrl	OR	95%CI	*P*	MAF Case Ctrl	OR	95% CI	*P*	MAF Case Ctrl	OR	95% CI	*P*	OR	*P-meta*	*Q**
rs3128126	ISG15	1	G	0.48 0.36	1.70	(1.28-2.26)	0.000278	0.33 0.36	0.77	(0.52/1.16)	0.2112	--	--	--	--	1.16	0.7088	0.002
rs324420	FAAH	1	A	0.29 0.19	1.74	(1.27-2.39)	0.000503	0.23 0.16	1.59	(1.10-2.31)	0.0131	0.24 0.17	1.85	(1.08-3.19)	0.026	1.70	2×10^{-6}	0.89
rs10490072	BCL11A	2	C	0.30 0.21	1.72	(1.27-2.33)	0.000476	0.24 0.23	1.09	(0.78/1.52)	0.6125	--	--	--	--	1.38	0.1608	0.05
rs2070887	VWF	12	G	0.15 0.08	2.07	(1.38-3.10)	0.0004	0.08 0.08	1.14	(0.69-1.88)	0.5928	--	--	--	--	1.56	0.132	0.07
rs9645765	VWF	12	G	0.14 0.07	2.17	(1.43-3.28)	0.000276	0.08 0.08	0.98	(0.60-1.61)	0.942	--	--	--	--	1.47	0.3276	0.016
rs2889490	SFRS16	19	G	0.58 0.46	1.67	(1.26-2.19)	0.000276	0.48 0.49	1.03	(0.77-1.37)	0.8319	--	--	--	--	1.31	0.2586	0.017
rs198977	KLK2	19	T	0.34 0.22	1.74	(1.23-2.32)	0.00021	0.38 0.22	1.23	(0.89-1.70)	0.2019	--	--	--	--	1.46		0.02763
rs16980496	ADRBK2	22	A	0.13 0.06	2.23	(1.44-3.45)	0.00034	0.07 0.09	1.05	(0.60-1.86)	0.8509	--	--	--	--	1.56	0.2337	0.04

Definitions of abbreviations: ARDS: Acute Respiratory Distress Syndrome; Chr: Chromosome; SNP: Single Nucleotide Polymorphism; MAF: Minor Allele Frequency; OR: Odd Ratio; CI: confidence interval; Case/Ctrl: Case/Controls. *The meta-analysis was performed using a fixed ($P<0.1$ for Cochran's Q test) or random effects-model ($P>0.1$ for Cochran's Q test).

Supplementary Table 5. Meta-analysis results for SNPs significantly associated with ARDS from pulmonary injury-related ARDS in Stage I																		
				Discovery phase (Stage I)				***Replication phase I (Stage II)***				***Replication phase II (Stage III)***				***Meta-analysis***		
SNP	Gene	Chr	Minor allele	MAF Case Ctrl	OR	95%CI	*P*	MAF Case Ctrl	OR	95% CI	*P*	MAF Case Ctrl	OR	95% CI	*P*	OR	*P-meta*	*Q*
rs2254358	HSPG2	1	C	0.25 0.33	0.63	(0.50-080)	0.000129	0.33 0.31	0.98	(0.75-1.28)	0.8977	--	--	--	--	0.78	0.2656	0.015
rs4075731	MAP3K6	1	A	0.32 0.40	0.67	(0.54-0.84)	0.000468	0.37 0.41	0.93	(0.73-1.18)	0.547	--	--	--	--	0.79	0.1432	0.05
rs12080701	PDE4B	1	G	0.14 0.09	1.85	(1.32-2.60)	0.000362	0.09 0.09	1.27	(0.81-1.98)	0.2997	--	--	--	--	1.61	0.0005	0.19
rs17419964	PDE4B	1	G	0.34 0.26	1.50	(1.20-1.88)	0.000433	0.25 0.27	1.24	(0.94-1.62)	0.12	--	--	--	--	1.39	0.0002	0.29
rs2854386	IL8RA	2	C	0.03 0.07	0.36	(0.20-0.64)	0.000476	0.06 0.07	0.96	(0.59-1.58)	0.8806	--	--	--	--	0.59	0.2902	0.01
rs2671222	IL8RA	2	A	0.02 0.06	0.34	(0.19-0.61)	0.000325	0.06 0.07	0.96	(0.59-1.58)	0.889	--	--	--	--	0.57	0.2913	0.008
rs6451620	GHR	5	A	0.090 .04	2.15	(1.41-3.29)	0.000372	0.050 .04	0.98	(0.54-1.77)	0.9416	--	--	--	--	1.49	0.3057	0.04
rs1190286	POPDC3	6	C	0.13 0.20	0.53	(0.39-0.72)	5.30E-05	0.14 0.20	0.65	(0.46-0.90)	0.0094	--	--	--	--	0.58	2.7×10^{-6}	0.38
rs802440	GRM3	7	T	0.40 0.30	1.47	(1.19-1.83)	0.000452	0.32 0.31	0.96	(0.75-1.25)	0.7879	--	--	--	--	1.19	0.4032	0.01
rs6970522	PRKAG2	7	G	0.51 0.44	1.49	(1.21-1.82)	0.000167	0.47 0.44	1.07	(0.84-1.35)	0.5847	--	--	--	--	1.27	0.1487	0.04
rs7807769	PRKAG2	7	A	0.48 0.39	1.58	(1.28-1.94)	1.61E-05	0.42 0.39	1.08	(0.85-1.38)	0.5082	--	--	--	--	1.31	0.1528	0.018
rs7801616	PRKAG2	7	T	0.48 0.39	1.54	(1.25-1.89)	4.37E-05	0.42 0.40	1.08	(0.85-1.37)	0.5195	--	--	--	--	1.29	0.1432	0.03
rs732821	HTR2A	13	A	0.54 0.45	1.52	(1.22-1.88)	0.000136	0.46 0.47	0.99	(0.78-1.25)	0.9171	--	--	--	--	1.23	0.3342	0.009
rs3887893	ABCC1	16	G	0.45 0.36	1.48	(1.20-1.83)	0.000287	0.39 0.36	1.23	(0.97-1.58)	0.0912	--	--	--	--	1.37	0.0001	0.16
rs9960450	TNFRSF11A	18	C	0.08 0.03	2.48	(1.56-3.93)	0.000114	0.05 0.03	1.64	(0.91-3.00)	0.1009	--	--	--	--	2.12	5.3×10^{-5}	0.28
rs17069902	TNFRSF11A	18	T	0.09 0.04	2.12	(1.41-3.20)	0.000312	0.06 0.04	1.51	(0.91-2.50)	0.1098	--	--	--	--	1.85	0.0001	0.31
rs8112223	HAS1	19	A	0.45 0.36	1.48	(1.19-1.84)	0.00037	0.40 0.35	1.06	(0.82-1.35)	0.6609	--	--	--	--	1.26	0.1663	0.05

Definitions of abbreviations: ARDS: Acute Respiratory Distress Syndrome; Chr: Chromosome; SNP: Single Nucleotide Polymorphism; MAF: Minor Allele Frequency; OR: Odd Ratio; CI: confidence interval; Case/Ctrl: Case/Controls. *The meta-analysis was performed using a fixed (P<0.1 for Cochran's Q test) or random effects-model (P>0.1 for Cochran's Q test).

Supplementary Table 6 ***SNPs associated with pulmonary injury-related ARDS in Stage I and tested for validation in Stage II (Population I: trauma)***

SNP	Gene	Human 610-quad*	Best proxy (LD:r^2)†	MAF Case/Ctrl.	OR (95% CI)	P‡ (additive)
rs7807769	PRKAG2	Not typed	rs7801616 (1)	0.43/0.43	1.11(0.85-1.47)	0.4476
rs7801616	PRKAG2	Typed	--	0.43/0.43	1.11(0.85-1.47)	0.4476
rs1190286	POPDC3	Not typed	rs1190298 (1)	0.17/0.16	1.32(0.89-1.95)	0.1672
rs9960450	TNFRSF11A	Typed	--	0.06/0.07	0.70(0.41-1.19)	0.1894
rs2254358	HSPG2	Not typed	rs3767141 (1)	0.32/0.29	0.99(0.73-1.34)	0.9432
rs732821	HTR2A	Not typed	rs4142900 (1)	0.49/0.49	0.87(0.66-1.16)	0.3544
rs6970522	PRKAG2	Typed	rs6970522 (1)	0.48/0.46	1.21(0.92-1.59)	0.1689
rs3887893	ABCC1	Typed	--	0.39/0.41	0.83(1.62-1.10)	0.1973
rs17069902	TNFRSF11A	Typed	--	0.06/0.09	0.67(0.40-1.12)	0.1309
rs2671222	IL8RA	Not typed	rs4672875 (1)	0.06/0.09	0.82(0.47-1.41)	0.4736
rs12080701	PDE4B	Not typed	rs11208772 (1)	0.10/0.13	0.79(0.51-1.23)	0.3055
rs8112223	HAS1	Typed	--	0.39/0.38	1.19(0.89-1.58)	0.2958
rs6451620	GHR	Typed	--	0.06/0.06	0.96(0.52-1.78)	0.9034
rs17419964	PDE4B	Not typed	rs12757542 (0.960)	0.28/0.27	1.06(0.77-1.47)	0.7139
rs802440	GRM3	Not typed	rs802443 (1)	0.35/0.33	1.17(0.87-1.57)	0.2958
rs4075731	MAP3K6	Not typed	rs12727507 (1)	0.36/0.34	1.23(0.91-1.67)	0.1809

SNPs associated with ARDS from direct lung injury in Stage I were tested in Stage II using Population I (a trauma-related ALI) in order to validate blunt trauma as an extrapulmonary insult. Proxies SNPs were identified for those SNPs not directly typed on the 610-Quad platform (38). No replications were observed for any of the SNPs previously associated with pulmonary injury-related ARDS. *Definitions of abbreviations*: ARDS: Acute Respiratory Distress Syndrome; SNP: Single Nucleotide Polymorphism; MAF: Minor Allele Frequency; Ctrl: Control; OR: Odd Ratio. *Indicates whether the SNP was directly genotyped by the Human 610-quad. †For those non-genotyped SNPs, markers in high linkage disequilibrium (LD) and typed in the genome-wide platform were identified (38) (best proxy, $\mathbf{r^2} \geq 0.8$) and used to infer the association with ARDS/ALI. ‡P values were adjusted for age, ISS and APACHE II.

Supplementary Table 7 ***SNPs associated with extrapulmonary injury-related ARDS in Stage I and tested for validation in Stage II (Population II: pneumonia/pulmonary sepsis)***

SNP	Gene	Human 610-quad*	MAF Case/Ctrl.	OR (95% CI)	P‡ (additive)
rs198977	KLK2	Typed	0.23/0.26	0.96 (0.73-1.27)	0.788
rs9645765	VWF	Imputed	0.08/0.08	0.90 (0.57-1.42)	0.6527
rs2889490	SFRS16	Imputed	0.49/0.48	1.06 (0.83-1.35)	0.6473
rs3128126	ISG15	Imputed**	--	--	--
rs16980496	ADRBK2	Imputed	0.06/0.08	0.83 (0.52-1.3)	0.4282
rs2070887	VWF	Typed	0.08/0.08	1.10 (0.70-1.72)	0.6799
rs10490072	BCL11A	Imputed	0.26/0.22	1.04 (0.79-1.38)	0.7638
rs324420	FAAH	Typed	0.22/0.22	0.92 (0.69-1.22)	0.5631

SNPs associated with ARDS from indirect lung injury in Stage I were tested in Stage II using Population II (pneumonia/pulmonary sepsis) in order to validate this population as a replication population for direct-cause ARDS associations. Imputation was carried out for those SNPs not directly typed on the 610-Quad platform. No replications were observed for any of the SNPs previously associated with extrapulmonary injury-related ARDS. ***Definitions of abbreviations*****:** ARDS: Acute Respiratory Distress Syndrome; SNP: Single Nucleotide polymorphism; MAF: Minor Allele Frequency; Ctrl: Control; OR: Odd Ratio. *Indicates whether the SNP was directly genotyped by the Human 610-quad. ‡P values were adjusted for age, gender and top 6 principal components **SNP rs3128126 (ISG15) was inadequately imputed ($r^2 < 0.1$) and it is not considered here.