

## Commentary

# Genetic epidemiology of acute lung injury: choosing the right candidate genes is the first step

Jason D Christie

Assistant Professor of Medicine and Assistant Professor of Epidemiology, Division of Pulmonary and Critical Care Medicine, Department of Medicine, and the Department of Biostatistics and Epidemiology, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, USA

Corresponding author: Jason D Christie, [jchristi@cceb.med.upenn.edu](mailto:jchristi@cceb.med.upenn.edu)

Published online: 19 August 2004

*Critical Care* 2004, **8**:411-413 (DOI 10.1186/cc2931)

This article is online at <http://ccforum.com/content/8/6/411>

© 2004 BioMed Central Ltd

See *Review*, page 440

## Abstract

In an innovative scientific review in this issue, Grigoryev and colleagues report a method for choosing candidate genes for acute lung injury (ALI) based on gene expression data derived from multiple animal models of mechanical ventilation and shear stress. The authors conclude there are five key biologic processes that warrant further investigation: inflammatory and immune responses, cell proliferation, chemotaxis, and blood coagulation. This review represents an important first step toward studying the genetic epidemiology of ventilator-induced lung injury and ALI. The application of these findings to future human studies of the genetic influence on ALI risks and outcomes is discussed here.

**Keywords** acute lung injury, acute respiratory distress syndrome, genetics, genomics, epidemiology, translational research

## Introduction

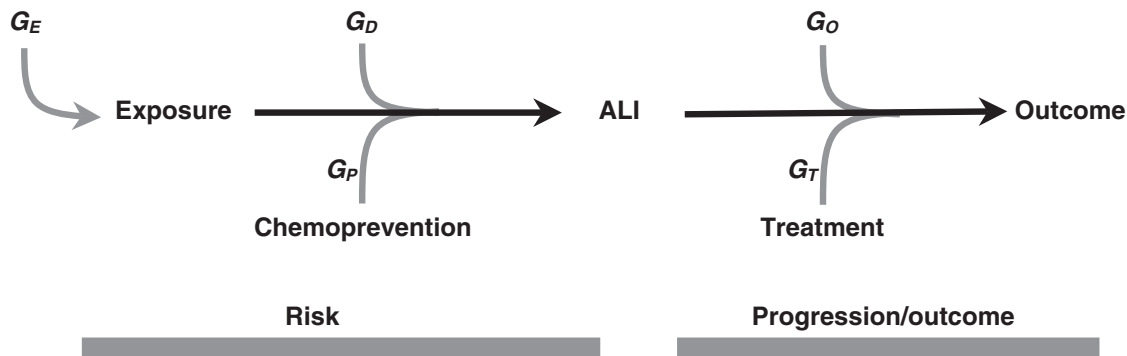
In recent years there has been growing interest in genetic susceptibility to acute lung injury (ALI) and in defining genetic determinants of outcomes in patients with established ALI [1,2]. The sporadic nature of ALI and the requirement for an extreme environmental predisposing insult (such as sepsis or trauma) make traditional family linkage studies of ALI unreasonable. Thus, the study of genetic influence on ALI incidence and outcomes will involve gene association studies, such as case-control and cohort studies. Choosing the right genes to study is the first step in designing such research.

## Choosing the right genes

There are several ways to choose appropriate genes for genetic epidemiologic study. Biologic inference on functionality of individual candidate single nucleotide polymorphisms (SNPs) within genes of interest may be limited and conflicting, and so choosing multiple genes in related pathways is attractive. When embarking on research such as this, the first question is which pathway(s) should we focus on?

In this issue of *Critical Care*, Grigoryev and colleagues [3] provide an important guide to this first step. In an innovative scientific review, those investigators report a method for choosing candidate genes for ALI based on gene expression data derived from multiple animal models of mechanical ventilation and shear stress. This is a strong approach, given that there are no underlying biologic hypotheses to steer the search other than grouping according to standard ontology. The authors conclude that there are five key biologic processes that warrant further investigation: inflammatory and immune responses, cell proliferation, chemotaxis, and blood coagulation. Errors in this approach could potentially arise because of the multiple cell types in the animal expression experiments or the arbitrary cutoffs utilized in the statistical filters. However, these potential limitations are minor and the report makes a significant contribution to the body of literature using expression data and genome-wide scans in different animal models of lung injury, such as ozone [4], nickel, and hyperoxia [5]. Furthermore, it achieves a new level of complexity by considering multiple pathways at once in a comprehensive manner.

Figure 1



Inherited susceptibility genes in the pathway of acute lung injury (ALI) risk and outcomes.  $G_E$  denotes genotypes affecting exposures (such as for developing sepsis or trauma);  $G_D$  denotes genotypes affecting disease risk for a given exposure;  $G_O$  denotes genotypes affecting outcomes of established ALI; and  $G_P$  and  $G_T$  denote genotypes affecting response to chemoprevention and treatments. Adapted from Rebbeck [11].

### Choosing the right question

Although generating key information for the investigation of ventilator-associated lung injury, the animal and cell models chosen to generate expression data in the review by Grigoryev and colleagues [3] may be limited in their generalizability to human ALI risk. Because ALI occurs in diverse populations and may be present before initiation of mechanical ventilation, application of the approach suggested by Grigoryev and coworkers to studies of ALI risk may not be as appropriate as application to studies investigating the influence of these genes on mortality risk or disease progression in ALI. The usefulness of these findings in extrapolating to human studies will largely be determined by designing the best clinical study to evaluate the appropriate genetic epidemiologic hypothesis. In general, genes that influence human disease states can be thought of as disease susceptibility genes – genes affecting outcomes and genes affecting chemoprevention and therapy responses (Fig. 1).

Given that they are derived from models of mechanical ventilation, the candidate genes advocated by Grigoryev and colleagues seem to be best applied to human studies of response to mechanical ventilation, perpetuation of lung injury, and/or outcomes in patients with ALI. However, as the authors point out, the same five fundamental processes they identified have been implicated in risk for ALI from other models, and thus it may be appropriate to study them as mediators of ALI risk in humans as well. Nonetheless, the unsupervised bioinformatics approach employed by the authors should serve as a model for choosing candidate genes derived from other basic science models of ALI (including susceptibility), such as sepsis, chemical aspiration, trauma, and endothelial injury.

### Choosing the right study design

Two general methods are available for studying candidate genes in gene association studies: approaches based on

functional inference of individual SNPs within a gene of interest and methods that evaluate the association of the entire gene by using haplotype-based analyses. Traditional ‘functional SNP’ investigating coding region or promoter SNPs studies may help identify mechanistically important variants that can serve as potential therapeutic targets. However, these studies are limited by current inconsistencies in functional inference, effects of locus heterogeneity, and they may not account for important variants in introns and/or untranslated regions [6]. Thus, if a selected SNP is not associated, this does not rule out involvement of the gene in the disease. Haplotype-based approaches use multiple genetic markers to test the association of the entire genetic locus with disease, but they do not necessarily identify the mechanistic underpinnings of the putative association [7]. These two approaches may be complementary and proceed in parallel.

Human gene association studies of ALI are susceptible to potential problems of all epidemiology in the intensive care unit [8]. They require rigorous definition of clinical variables (including ALI), consideration of confounding and causal pathway variables, careful attention to avoid biases introduced by the selection of the study population and controls, and biases introduced by underlying population architecture [9]. Of note, a recent review [10] concluded that the failure to replicate gene associations from published case-control and cohort studies was likely to be due to, ‘poor study design and execution’, including inadequate sample size.

### Conclusion

The introduction of genetic epidemiology as a tool to improve our understanding of mechanisms, prognosis, and therapeutic response of ALI and other critical illnesses is still in the early stages. The novel approach suggested by Grigoryev and colleagues [3] provides us with a model for

how to focus our future energies in this emerging line of investigation.

## Competing interests

The author(s) declare that they have no competing interests.

## References

1. Gong MN, Wei S, Xu LL, Miller DP, Thompson BT, Christiani DC, Sai Y, Zhou W: **Polymorphism in the surfactant protein-B gene, gender, and the risk of direct pulmonary injury and ARDS.** *Chest* 2004, **125**:203-211.
2. Marshall RP, Webb S, Bellingan GJ, Montgomery HE, Chaudhari B, McAnulty RJ, Humphries SE, Hill MR, Laurent GJ: **Angiotensin converting enzyme insertion/deletion polymorphism is associated with susceptibility and outcome in acute respiratory distress syndrome.** *Am J Respir Crit Care Med* 2002, **166**:646-650.
3. Grigoryev DN, Finigan JH, Hassoun P, Garcia JGN: **Science review: Searching for gene candidates in acute lung injury.** *Crit Care* 2004, **8**:440-447.
4. Cho HY, Zhang LY, Kleeberger SR: **Ozone-induced lung inflammation and hyperreactivity are mediated via tumor necrosis factor-alpha receptors.** *Am J Physiol Lung Cell Mol Physiol* 2001, **280**:L537-L546.
5. Cho HY, Jedlicka AE, Reddy SP, Zhang LY, Kensler TW, Kleeberger SR: **Linkage analysis of susceptibility to hyperoxia. Nrf2 is a candidate gene.** *Am J Respir Cell Mol Biol* 2002, **26**:42-51.
6. Pagani F, Baralle FE: **Genomic variants in exons and introns: identifying the splicing spoilers.** *Nat Rev Genet* 2004, **5**:389-396.
7. Crawford DC, Carlson CS, Rieder MJ, Carrington DP, Yi Q, Smith JD, Eberle MA, Kruglyak L, Nickerson DA: **Haplotype diversity across 100 candidate genes for inflammation, lipid metabolism, and blood pressure regulation in two populations.** *Am J Hum Genet* 2004, **74**:610-622.
8. Rubenfeld GD, Christie JD: **The epidemiologist in the intensive care unit.** *Intensive Care Med* 2004, **30**:4-6.
9. Ellsworth DL, Manolio TA: **The emerging importance of genetics in epidemiologic research II. Issues in study design and gene mapping.** *Ann Epidemiol* 1999, **9**:75-90.
10. Cardon LR, Palmer LJ: **Population stratification and spurious allelic association.** *Lancet* 2003, **361**:598-604.
11. Rebbeck TR: **Biomarkers of inherited susceptibility and cancer.** In *Mechanisms of Carcinogenesis: Contributions of Molecular Epidemiology*. IARC Scientific Publication 157. Edited by Buffler P, Rice J, Baan R, Bird M, Boffetta P. Lyon: IARC Press; 2004:91-103.