

Anil Sapru
Joseph L. Wiemels
John S. Witte
Lorraine B. Ware
Michael A. Matthay

Acute lung injury and the coagulation pathway: potential role of gene polymorphisms in the protein C and fibrinolytic pathways

Received: 14 October 2005
Accepted: 3 May 2006
Published online: 13 June 2006
© Springer-Verlag 2006

A. Sapru (✉)
University of California, Pediatric Critical
Care Medicine,
505 Parnassus Avenue, Moffitt 680-B,
San Francisco CA 94143-0106, USA
e-mail: anil@itsa.ucsf.edu
Tel.: +1-415-4760963
Fax: +1-415-5024186

J. L. Wiemels · J. S. Witte
University of California, Department of
Epidemiology and Biostatistics,
San Francisco California, USA

L. B. Ware
Vanderbilt University School of Medicine,
Division of Allergy, Pulmonary and Critical
Care Medicine, Department of Medicine,
Nashville Tennessee, USA

M. A. Matthay
University of California, Department of
Medicine and Anesthesia and
Cardiovascular Research Institute,
San Francisco California, USA

Abstract There is evidence that dysregulation of coagulation and fibrinolysis may participate in the pathogenesis of acute lung injury (ALI) and the acute respiratory distress syndrome (ARDS). Altered concentrations of several proteins of the coagulation and fibrinolytic pathways in plasma and pulmonary edema fluid from patients with acute lung injury have been related to the severity of lung injury and clinical outcomes. Polymorphisms in the genes encoding for proteins of the protein C and fibrinolysis pathways are known to regulate the production of the respective proteins. It is plausible that these polymorphisms may be associated with the susceptibility to and severity of illness in ALI and ARDS. Well-designed studies that examine the association of these polymorphisms with susceptibility and severity of ALI and ARDS are needed to test the influence of both genetic and environmental factors on

the clinical outcomes in patients with ALI and ARDS. There are several important considerations in the design of these genetic association studies, including selection of candidate genes with the most biological plausibility, definition of the phenotype, selection of appropriate controls, determination of the appropriate sample size and assessment of Hardy–Weinberg equilibrium among controls as a measure of internal validity.

Introduction

Acute lung injury (ALI) and its more severe form, the acute respiratory distress syndrome (ARDS), are syndromes of acute respiratory failure that are characterized by acute pulmonary edema and lung inflammation. ALI and ARDS include a spectrum of increasing severity of lung injury defined by physiologic and radiographic criteria. The American–European consensus committee has defined the illness as acute onset with a $\text{PaO}_2/\text{FiO}_2 \leq 300$ for ALI and ≤ 200 for ARDS, bilateral infiltrates on chest

radiograph, and a pulmonary artery occlusion pressure < 18 mmHg or the absence of clinical evidence of left atrial hypertension [1]. ALI has an incidence of approximately 200,000 patients per year in the United States [2]. Estimates of mortality rates for ARDS vary from 26% to 58%, and epidemiologic data indicate that ARDS may account for up to 75,000 deaths per year in the USA [2, 3, 4].

Alterations in the levels of mediators of coagulation such as protein C, thrombomodulin, and plasminogen activator inhibitor-1 (PAI-1) have been associated with worse clinical outcomes in patients with ALI. It is not

clear whether the alterations in these protein levels are determined exclusively by environmental factors that precipitate ALI, such as the virulence of the infection, the extent of aspiration, and the severity of shock or whether the alterations in coagulation protein concentrations are determined by genetic factors as well. Clinical and epidemiological observations in patients with ALI and data from animal experiments support the hypothesis that genetic factors may play a part in the susceptibility to development of ALI. Polymorphisms in genes encoding for proteins of the protein C and fibrinolytic pathways are associated with altered levels of their respective mediators, and phenotypes of disordered coagulation. It is possible that these genetic polymorphisms are associated with the clinical outcomes in patients with ALI because they can regulate production and activity of proteins of the coagulation and fibrinolytic pathways. An association between polymorphisms of the genes encoding for these proteins and clinical outcomes would (a) strengthen the importance of the association between the proteins and clinical outcomes, (b) provide further evidence for a cause and effect relationship of this association, (c) lead to a better understanding of the biological role of different genes and their products and in the pathogenesis of ALI and (d) help develop profiling of clinical risk, so that specific therapeutic and preventive measures might be directed towards more susceptible patients [5].

This article summarizes the coagulation pathway abnormalities in ALI, presents evidence for a possible contribution of genetic susceptibility to ALI/ARDS, and provides an overview of important considerations in the design and interpretation of investigations into the genetic basis of ALI focused primarily on the protein C and fibrinolytic pathways.

Role of coagulation and fibrinolysis in ALI/ARDS

Activation of the coagulation cascade, specifically thrombin formation, can lead to inflammatory events [6]. Sev-

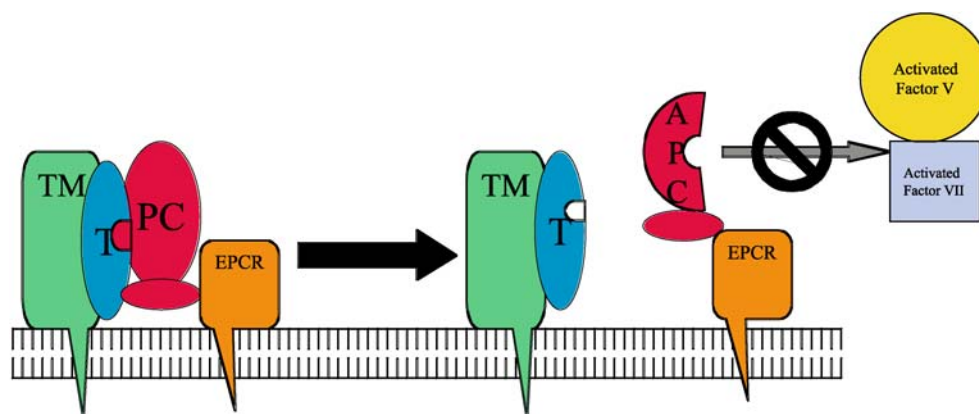
eral inflammatory mediators are involved in the activation of coagulation and, in turn, coagulation proteins are themselves actively involved in the inflammatory process. Interaction of coagulation and inflammatory pathways are well described in related clinical conditions like sepsis and cardiopulmonary bypass [7, 8].

Intra-alveolar fibrin deposition is a hallmark of several acute inflammatory lung diseases, and enhanced alveolar procoagulant activity is reported in patients with ventilator-associated pneumonia, severe pneumonia, and ARDS. Fibrin deposition may exert beneficial effects on gas exchange by sealing leakage sites if the lung capillary endothelial and epithelial barriers are compromised. On the other hand fibrin deposition in the alveoli may be harmful since it can lead to activation of neutrophils and fibroblasts, compromise endothelial integrity, contribute to loss of surfactant activity favoring alveolar collapse, decrease alveolar fluid clearance and induce thrombotic obstruction of the microcirculation [9, 10, 11, 12, 13, 14]. Injury to the pulmonary microcirculation from inflammatory and thrombotic mechanisms probably contributes to increase in dead space fraction that may be an independent predictor of mortality in ARDS [15]. Elevated plasma tissue factor levels in patients with ARDS have been associated with poor clinical outcomes, suggesting that activation of systemic coagulation may be detrimental in patients with ARDS [16].

ALI and the protein C and fibrinolytic pathways

The protein C pathway plays an important role in both the regulation of blood coagulation and the prevention of lethal effects of sepsis. Protein C is an endogenous anticoagulant that is activated by thrombin in the presence of thrombomodulin. The membrane-bound endothelial protein C receptor (EPCR) potentiates this activation (Fig. 1). Activated protein C (APC) in turn inactivates the activated factors V and VII, thereby inhibiting the coagulation cascade [17]. In addition, neutrophils express EPCR on their

Fig. 1 Protein C (PC) is converted to activated protein C (APC) while bound to endothelial protein C receptor (EPCR) in the presence of thrombin (T) and thrombomodulin (TM). APC then inactivates factors V and VII, thereby inhibiting the coagulation cascade



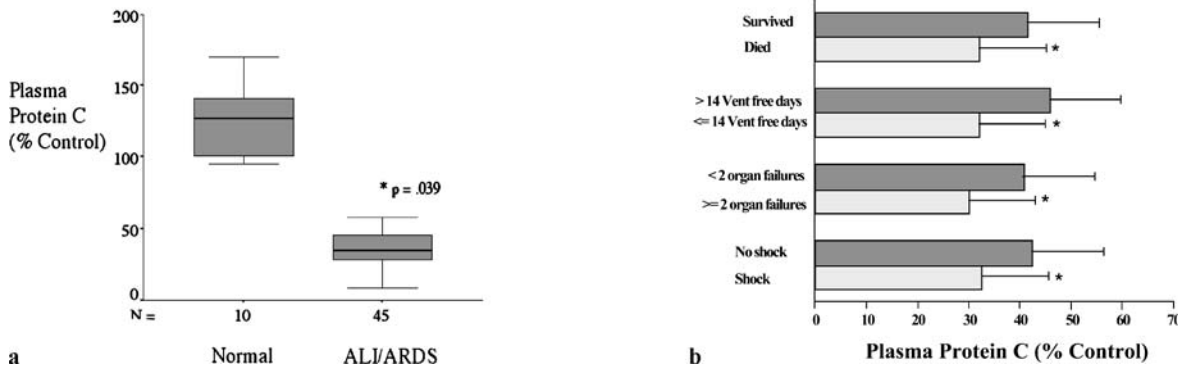


Fig. 2 **a** Boxplot summary of plasma protein C levels in 10 normal controls and 45 patients with ALI. *Box* encompasses 25th to 75th percentile, *error bars* encompass 10th to 90th percentile, and *horizontal bar* shows median. **b** Plasma protein C levels in 45 patients with ALI were consistently higher in those with favorable clinical outcomes, including hospital survival, greater than 14 days of un-

sisted ventilation over a 28-day period, less than two organ failures or the absence of circulatory shock. Data as mean \pm SD, [Ware LB, Fang X, Matthay MA (2003) Protein C and thrombomodulin in human acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 285:L516, Fig. 1 and Fig. 2. Reprinted with permission of The American Physiological Society]

surface and APC inhibits neutrophil chemotaxis in experimental models of pulmonary inflammation [18, 19].

Pulmonary edema fluid obtained early in the course of ALI has lower concentrations of protein C than do controls with hydrostatic pulmonary edema, and reduced plasma

protein C levels are associated with higher mortality and more non-pulmonary organ system dysfunction (Fig. 2). In addition, markedly elevated levels of thrombomodulin in the pulmonary edema fluid from patients with ALI suggest that release of thrombomodulin probably occurs from both epithelial and endothelial sources in the lung [20].

Fibrin is degraded by plasmin, a proteolytic enzyme present in the tissues in the form of an inactive precursor, plasminogen. Plasminogen activator inhibitor-1 is the primary inhibitor of plasminogen activation and thus is a major inhibitor of fibrinolysis. Fibrinolysis is impaired in ALI, in part from elevated levels of PAI-1. Pulmonary edema fluid obtained early in the course of ALI has elevated concentrations of PAI-1 (Fig. 3). The elevated PAI-1 levels have predictive value for identifying patients who are more likely to die among all patients with ALI, regardless of the clinical risk factor that predisposes to the development of ALI [21].

An analysis of 779 patients from the ARDS network clinical trial of low tidal volume versus high tidal volume has confirmed these findings [22, 23]. Lower levels of plasma protein C and higher levels of plasma PAI-1 were independent predictors of mortality, and the combination of low protein C and high PAI-1 carried an even higher risk of mortality. The prognostic value of protein C and PAI-1 was not altered by exclusion of patients with co-existing sepsis. Thus, abnormalities of coagulation and fibrinolysis are not simply a manifestation of sepsis, but are associated with ALI, regardless of the underlying cause. Interestingly, a protective ventilator strategy did not have a significant effect on plasma levels of protein C and PAI-1. The alterations in the levels of these coagulation and fibrinolysis proteins may be in part determined by genetic factors independent of the degree of direct lung injury.

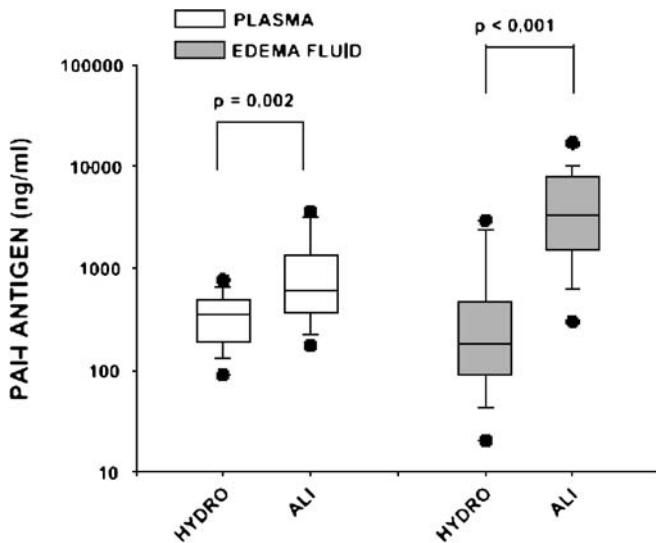


Fig. 3 Plasminogen activator inhibitor-1 (PAI-1) antigen levels in patients with acute lung injury (ALI) and hydrostatic pulmonary edema (HYDRO). PAI-1 antigen level was measured in plasma (*open boxes*) and edema fluid (*filled boxes*) from patients with either clinical ALI or HYDRO by ELISA. Data are plotted in box plot format (median, 25–75%) and compared using the Mann–Whitney U-test. *P* values are as indicated. [Prabhakaran P, Ware LB, White KE, Cross MT, Matthay MA, Olman MA (2003) Elevated levels of plasminogen activator inhibitor-1 in pulmonary edema fluid are associated with mortality in acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 285:L22, Fig. 1. Reprinted with permission of The American Physiological Society]

Evidence for a possible genetic contribution to ALI

Variability in susceptibility and outcome

Both adult and pediatric patients develop ALI [4]. A number of clinical disorders that directly or indirectly injure the lung can lead to ALI, but of the apparently large population at risk of developing ALI, only a small fraction develop the clinical syndrome. Among patients who have identical risk factors for lung injury, such as sepsis and aspiration, fewer than 50% develop ALI [24, 25]. Likewise, it is difficult to predict who will recover from ALI. Current clinical and physiological measurements cannot account for outcome variability. Individual susceptibility and variation in the clinical course and outcome may be related in part to the underlying genetic characteristics of the individual patient.

Positional cloning linkage studies

Genome-wide searches in animal models have identified a number of quantitative trait loci (QTL) that associate with susceptibility to experimental lung injury. A QTL on chromosome 4 that contains the gene for the toll-like receptor 4 is associated with ozone-induced injury in animal models [26]. An association between ozone and

nickel induces lung injury, and a region of chromosome 6 that contains genes for pro-inflammatory cytokines, as well as HLA, has also been reported in animal models. Similarly, nickel sulfate (NiSO₄)-induced ALI in rat models is controlled by at least five genes, including aquaporin 1, transforming growth factor- α (TGF- α), and surfactant protein B [27].

Gene expression studies

Researchers at the Hopgene Programs for Genomic Applications and other investigators have used micro-arrays to study gene expression profiles in animal models subject to mechanical stress and in human endothelial cells subject to stretch in vitro. These investigators reported up-regulation in the expression of multiple genes involved in the immune response, including chemotaxis, cell proliferation, inflammatory responses, and coagulation genes for PAI-1, tissue factor precursor, and fibrinogen- α [28].

Candidate gene association studies

Much of the genetic variation between individuals lies in differences known as single-nucleotide polymorphisms

Table 1 Association studies of genetic polymorphisms and ALI

Author and year	Gene	Polymorphism	Associated phenotype
Lin (2000)	Surfactant protein B	1580C/T	C allele associated with ARDS [36]
Gong (2004)	Surfactant protein B	Tandem repeats in intron4	Variant allele associated with susceptibility to ARDS in women [33]
Quasney (2004)	Surfactant protein B	1580 C/T polymorphism	C allele associated with ARDS and mechanical ventilation in patients with community-acquired pneumonia [39]
Currier (2005)	Surfactant protein B	Tandem repeats in intron4	Variant allele associated with increased 60-day mortality in Caucasians [29]
Marshall (2002)	Angiotensin-converting enzyme	Insertion/deletion (I/D) of a 287BP repeat sequence in intron 16	DD genotype had increased frequency in patients with ARDS [37]
Gong MN (2005)	Mannose-binding lectin	Codon 54 (G/A) polymorphism	Variant allele associated with increased risk of ARDS, higher APACHE III score and increased ICU mortality [35]
Medford (2005)	Vascular endothelial growth factor	+936 C/T polymorphism	T allele associated with ARDS and higher mean APACHE III scores [38]
Ye (2005)	Pre-B cell colony enhancing factor	1001T/G and 1543C/T SNPs	Haplotype GC associated with higher risk of ALI [40]
Gong MN (2005)	Tumor necrosis factor	-308GA polymorphism and TNFB1/2 polymorphisms	-308 A allele associated with increased chance of developing ARDS in direct lung injury and increased 60-day mortality [34].
Gong MN (2006)	Interleukin-10	IL-10 promoter polymorphism (-10822G/A)	IL-10 (-10822) GG genotype was associated with susceptibility to ARDS with advancing age but decreased severity of illness and better outcomes overall [82]
Gao (2006)	Myosin light chain kinase	Multiple SNPs and haplotypes	A single SNP (rs 820336) and a haplotype on 5' end associated with development of ALI; an additional haplotype on 3' end associated with ALI in African Americans only [31]

(SNPs). Polymorphisms are variant forms of genes that occur in at least 1% of the population. Associations between susceptibility to and severity of ALI and polymorphisms encoding cytokines and other mediators of inflammation have been reported in a number of association studies [29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41] (Table 1). Results of these studies suggest that it is unlikely that a single gene or SNP will explain the diversity in risk, clinical manifestations (diverse phenotypes) and outcomes of ALI. ALI is probably a disease with genetic heterogeneity involving a number of genes at different loci with interactions both within the multiple genes and with environmental factors that predispose to the development of the clinical disorder. Based on our prior work and that of other investigators [16, 20, 21, 22, 23, 42], we hypothesize that genes encoding for the mediators of the coagulation pathways have a strong biological plausibility for a significant role in the susceptibility to and clinical severity of ALI.

ALI tends to be individual-specific, it is not suitable for family-based study designs.

More recently, genome-wide association studies of unrelated individuals have been used as well. However, these studies are hypothesis generating because they lead to the identification of a limited set of genes, which can then be further studied. It is also possible to begin with a candidate gene approach, evaluating the association between genetic variants and disease directly. The first step is to identify genotypes a priori, based on the results of previous positional cloning studies or biological plausibility, and then study the association of specific polymorphisms with the phenotype of interest (mild versus severe disease or the presence or absence of disease). Some of the important considerations (challenges) encountered in the design and interpretation of genetic association studies are detailed in the next section.

The study of genetic factors

Genetic factors traditionally have been studied using either positional cloning or focused candidate gene approaches. Positional cloning approaches have most often entailed genome-wide linkage studies of families with multiple affected individuals. Since the precipitating event to

Considerations in the design of association studies

Candidate genes
The first step in the design of association studies is the selection of the most appropriate candidate genes. Knowledge of the gene and SNP function is mandatory for the appropriate design and interpretation of candidate gene association studies. Investigators usually rely on making an

Table 2 Well-characterized coagulation pathway polymorphisms that may have a role in ALI

Gene	Function of gene product	Rationale in ALI	SNP/haplotype (minor allele frequency)	Known phenotype of the SNP/haplotype
Protein C	Anti-inflammatory and anti-thrombotic	Lower levels predict worse outcome	Haplotypes based on promoter region SNPs CG (0.48) CA (0.20) TA (0.31)	Decrease in protein C and increased incidence of venous thromboembolism [44]
Endothelial protein C receptor (EPCR)	Activation of protein C	Activated protein C protective in ALI	Haplotypes based on promoter region SNPs A1 (0.39) A2(0.43) A3 (0.14)	Increase in plasma levels of sEPCR and increased risk of thrombosis [49]
Plasminogen activator inhibitor-1 (PAI-1)	Inhibitor of plasminogen activator, net effect inhibits fibrinolysis	Increased levels associated with worse outcome in ALI	Promoter region polymorphism 4G/5G (0.45)	Increase in plasma levels of [46] PAI-1 and worse outcome in children with sepsis
Thrombomodulin (TMBD)	Activation of protein C	Higher levels associated with worse outcome in ALI	Ala 455 Val (0.28)	Associated with coronary heart disease [48]
Tissue factor (TF)	Activation of coagulation cascade	Higher plasma levels associated with mortality in ALI	Haplotype 1208 D (0.5) Haplotype 1208 I	Decrease in plasma levels of tissue factor and lower risk of thrombosis [45]
Factor V	Target of activated protein C action	Lower levels of protein C associated with poor outcome in ALI	Factor V 1691 G/A (Factor V Leiden)	Factor V Leiden is resistant to cleavage by activated protein C and is a known cause of thrombophilia [51, 53]

assessment of the function of a candidate gene or SNP from previous work and published literature [43].

A number of SNPs and haplotypes in the genes encoding for mediators of protein C and fibrinolytic components of the coagulation pathway alter levels of their respective mediators. Some of these polymorphisms and SNPs are relatively common in the general population (MAF > 5%) and have been associated with specific disease phenotypes of disordered coagulation (Table 2) [44, 45, 46, 47, 48, 49, 50]. However, there are other less frequent but well-described mutations and polymorphisms in the coagulation pathway, which may also have a role in ALI due to their effect on the coagulation cascade [51, 52, 53].

Biological plausibility

Finding an association does not provide proof of causality. Also, defining the biological mechanism underlying the associations should be an important part of a genetic association study. The gene product must be known to alter the risk or course of the disease, and the proposed mechanisms by which the gene product may influence the course of the disease should either be known or be a part of such a study. Whenever the biological plausibility for the choice of a candidate gene is based on a gene product that can be measured in plasma or other body fluids, it is preferable to include these measurements in both the plasma and broncho-alveolar lavage fluid. This should be done regardless of what is known about the effect of the particular polymorphism on the baseline levels of the gene product. For example, the 4G/5G genotype accounts for only a small part of variance in PAI-1 levels at baseline, but the difference in the levels of PAI-1 between 4G and 5G becomes more obvious in the presence of environmental and disease factors that stimulate PAI-1 expression [54, 55, 56]. Associations with these intermediate phenotypes, whenever present, will further strengthen the significance of the overall association with the clinical phenotype.

Defining the phenotype

Special attention is required in the design of association studies for correct classification of cases and controls. Improper classification may lead to a bias in the association and reduced power to detect an effect. Definition of the phenotype in ALI is problematic because there is no definitive diagnostic test for ALI. In addition, ALI has a complex etiology and there may be multiple subtypes within the ALI phenotype [57]. One option is to study ALI from a specific etiology to decrease the heterogeneity of the cases, e.g., ALI due to direct pulmonary injury, ALI due to trauma or ALI due to pancreatitis. However, the potential disadvantages of this approach include the possibility that there may be further subtypes within each

of these groups and that sufficient number of cases with a homogenous phenotype may be difficult to ascertain, especially in single-center studies [57]. On the other hand, even though the etiology of ARDS is complicated, a common injury pathway may contribute to the damage of lung tissue in these patients [58]. Though there are limitations to the existing definitions of ALI and sepsis, syndromes can be studied with a reasonable degree of validity and reliability as has been done in other disciplines [59]. The American European Consensus definition serves as a uniformly accepted guideline for defining lung injury and is the most universally accepted criterion for ALI/ARDS. However, in order to avoid misclassification that may result from lack of reliability among the ALI diagnostic criteria, more than one investigator should review the entire primary diagnostic data, including the chest radiographs. Only subjects with a high degree of agreement among the different investigators should be included in such studies.

Sample size and power

Despite the fact that the power of association analysis to detect genetic contribution to complex disease is greater than that of linkage studies, many association studies suffer from a lack of adequate power because of a small sample size [60]. The number of subjects needed to achieve adequate power is dependent on the frequency of the polymorphism, the disease model (dominant, recessive, co-dominant), and the genetic risk ratio, a measure of gene effect on the scale of risk [61, 62]. Therefore, studies based on polymorphisms with a higher minor allele frequency require a smaller number of subjects for adequate power than studies based on rare polymorphisms with similar genetic effect. Also, rare variants (< 5%) are more likely to be limited to specific population groups, whereas common alleles (> 10%) are more likely to be found globally [61]. The common disease–common variant hypothesis proposes that most of the genetic risk for common complex diseases is due to a small number of relatively common variants [63]. However, it is possible that a rare allele with a strong penetrance may have a more profound effect on the phenotype.

Added power and precision is provided in epidemiological studies by continuous variables relative to categorical variables [64]. There are advantages to the use of quantitative traits as outcomes, especially in studies of modifier genes in complex diseases [65]. Association studies using quantitative traits as outcome in complex diseases have been previously reported in asthma, Alzheimer's disease, diabetes and obesity [66, 67, 68, 69, 70]. Using a continuous phenotype is less ambiguous and more informative, objective and statistically powerful than binary categorizations of disease status [71]. Although, similar to the binary outcomes (e.g., mortality), continuous phenotypes

Table 3 Sample size estimates for continuous outcomes (assumptions: allele frequency = 0.25, $\alpha = 0.05$, $\beta = 0.2$)

Effect size (% standard deviation)	Sample size
30%	350
40%	195
50%	124
60%	85

(e.g., ventilator-free days, pulmonary dead space fraction, plasma levels of the mediator) as the outcome in association studies of ALI are susceptible to ventilator strategies and interventions such as fluid loading or diuresis, they may yield greater power than dichotomous outcomes such as mortality, depending on their effect size and variability. Using standard power calculations [72] we estimated the sample size required for $\alpha = 0.05$ and $\beta = 0.2$ (power of 80% at a significance of 95%) in a study with a minor allele frequency of 25%, dominant mode of inheritance, and a continuous phenotype as the outcome (Table 3). If there is a sufficiently large effect, even a modest number of patients may yield satisfactory power in such study designs. If the effect is smaller, substantially more subjects will be required. In contrast, a study using a dichotomous endpoint such as mortality (estimated 25%) would need a sample size of approximately 500 patients to detect an OR of 1.8 or more with the assumptions described above.

Selection of controls

Controls should represent the reference population at risk for development of the disease that have not yet developed the disease. There is no agreement on ideal controls that best match the reference population in ALI or sepsis [57, 73]. Studies of ALI in the past have utilized either normal healthy individuals or critically ill patients at risk for development of ALI as controls. Studies that utilize the

healthy population as controls may result in identification of genes that predispose to development of critical illness rather than to the identification of genes that predispose to ALI itself. Studies utilizing just the hospital-based at-risk population may be confounded by a selection bias as well. They tend to eliminate the population of controls that contributes to cases with rapidly evolving ALI presenting directly with ALI/ARDS without passing through an intermediate phenotype. In order to overcome these issues, one may select two groups of controls, i.e., healthy age- and ethnicity-matched controls from the community and another group of hospital-based at-risk controls. Another approach is to use the case-cohort study design, where a cohort of at-risk patients is followed for the development of ALI, thereby eliminating bias that may occur due to inappropriate selection of a control population. This approach requires significantly more resources and patients to allow a sufficient number of cases to emerge before the association study can proceed [60].

Linkage disequilibrium

Linkage disequilibrium refers to the observation that specific neighboring alleles tend to occur together on the same chromosome more often than would be expected by chance alone. The combination of alleles along a chromosome that are all in high linkage disequilibrium with each other is referred to as a haplotype. A chromosomal region with high linkage disequilibrium and low haplotype diversity is referred to as a haplotype block. Each time a mutation occurs, it results in the creation of a new haplotype. Over generations this linkage is broken into several further haplotypes by recombination, new mutations and genetic drift. Haplotypes grouped on the basis of their evolutionary origin are referred to as clades.

Both haplotype-based and single SNP approaches (Table 4) can be applied concurrently to candidate genes that are implicated in disease pathogenesis, and the two approaches are not mutually exclusive. Combining infor-

Table 4 Comparison of a haplotype versus single SNP approach

Haplotype-based association	Single SNP-based association
<p>Advantages: Haplotypes represented by tag SNPs have greater power to comprehensively characterize all the common genetic variation in a particular gene or region and can localize a disease-conferring gene or locus when the causal SNP is unknown. It is possible to test the overall effect of multiple SNPs on the trait of interest</p> <p>Disadvantages: It does not identify the injurious SNP but only narrows the possibility. Marked haplotype diversity among populations leads to requirement of multiple ethnicity-based markers.</p>	<p>Advantage: Causal SNP can be the basis of further study of mechanistic pathways and therapeutic interventions.</p> <p>Disadvantage: An individual polymorphism may not be the disease locus but rather may be in strong linkage disequilibrium with the causal allele.</p>

mation on re-sequencing of coagulation genes available at the Seattle SNPs website (<http://pga.mbt.washington.edu>) with one of the available software programs to identify areas of linkage disequilibrium and tag SNPs representing these blocks of linkage equilibrium, association studies can be designed that can account for most of the common genetic variation in these genes [74, 75, 76]. This approach can be done in addition to testing for specific SNPs that have been shown to be associated with altered levels of gene product in previous studies.

Multiple comparisons

The increasing number of candidate genes implies that multiple association scans will be conducted on the same patient samples. This approach leads to testing of multiple hypotheses, which can lead to false-positive results. However, the increased false-positive rate from multiple comparisons in genetic association studies is less than the expected increase if the markers are entirely independent. Therefore, Bonferroni correction of these data markedly overcorrects for the inflated false-positive rate and thereby discards valid information in the sample [60]. Use of computational techniques based on multiple permutations of the data set to obtain a distribution for the test statistic under the null or the use of the false discovery rate method are accepted methods of adjusting for multiple comparisons in genetic association studies [77].

Population stratification

Confounding by population stratification is another important issue in population-based association studies. False associations may be reported if there is variation in the incidence or severity of a particular disease among different ethnic groups, which may also happen to have a variation in the allele frequencies of the studied SNP. This issue can be managed at the design stage by choosing to study only a single ethnic group or by stratifying by ethnicity. An alternative approach involves testing for other unrelated SNPs at the so-called null loci and then adjusting for population stratification using either structured association or genomic control strategy. The structured association method reconstructs the population substructure using the data from the many unlinked markers, splits the data accordingly, and then tests for allele frequency differences between those sub-populations [78]. Modifications to improve the robustness of these procedures have also been described [79]. On the other hand, the genomic control method estimates background association between populations by using the data from SNPs not related to the disease of interest. The individual test statistic of interest is

corrected by dividing it by a correction factor, which is equal to the median of the entire test statistics generated for the null loci. This method adjusts for the overdispersion of the test statistics generated due to population substructure [80]. However, there can be loss of power in the face of increasing population substructure, and therefore experiments should preferably be designed to minimize population substructure [81].

Hardy–Weinberg equilibrium

Hardy–Weinberg equilibrium indicates that in a freely breeding population genotype frequencies can be determined directly from allele frequencies. It can be readily assessed with the use of a goodness-of-fit chi-square test for bi-allelic markers. Testing for Hardy–Weinberg equilibrium is frequently used in genetic association studies to test for internal validity. Failure to demonstrate Hardy–Weinberg equilibrium among controls in an association study can result from genotyping errors, inbreeding, population substructure or selection bias [71]

Conclusions

There is considerable interest in testing candidate genes that may increase the susceptibility to or severity of ALI. There is some evidence that dysregulated coagulation and impaired fibrinolysis may contribute to the pathogenesis of ALI. Altered levels of the plasma and pulmonary edema fluid protein concentrations of protein C, thrombomodulin, tissue factor, and PAI-1 are biological markers of worse clinical outcomes in patients with ALI. However, it is not known whether the altered concentrations in these mediators are dependent only on environmental factors, such as the severity of pneumonia, aspiration or shock, or also on underlying genetic factors. Several polymorphisms in genes encoding for these coagulation proteins are present in the general population with high allele frequencies and are associated with altered levels of the gene products and clinical phenotypes of disordered coagulation. Therefore, clinical studies are needed to test the possible contributions of these polymorphisms to the severity and susceptibility to ALI. These studies should include determination of the polymorphism at the genetic level as well as measurement of levels of the respective gene product in the plasma and broncho-alveolar lavage fluid of patients with ALI. Several issues need to be carefully managed during the design of such studies, including definition of the phenotype, selection of controls, choice of candidate genes, attention to linkage disequilibrium and adjustment for multiple comparisons and population stratification.

References

- Matthay MA, Zimmerman GA, Esmon C, Bhattacharya J, Collier B, Doerschuk CM, Floros J, Gimbrone MA Jr, Hoffman E, Hubmayr RD, Leppert M, Matalon S, Munford R, Parsons P, Slutsky AS, Tracey KJ, Ward P, Gail DB, Harabin AL (2003) Future research directions in acute lung injury: summary of a National Heart, Lung, and Blood Institute working group. *Am J Respir Crit Care Med* 167:1027–1035
- Rubinfeld GD, Caldwell E, Peabody E, Weaver J, Martin DP, Neff M, Stern EJ, Hudson LD (2005) Incidence and outcomes of acute lung injury. *N Engl J Med* 353:1685–1693
- Brun-Buisson C, Minelli C, Bertolini G, Brazzi L, Pimentel J, Lewandowski K, Bion J, Romand JA, Villar J, Thorsteinsson A, Damas P, Armaganidis A, Lemaire F (2004) Epidemiology and outcome of acute lung injury in European intensive care units. Results from the ALIVE study. *Intensive Care Med* 30:51–61
- Flori HR, Glidden DV, Rutherford GW, Matthay MA (2005) Pediatric acute lung injury: prospective evaluation of risk factors associated with mortality. *Am J Respir Crit Care Med* 171:995–1001
- Villar J (2002) Genetics and the pathogenesis of adult respiratory distress syndrome. *Curr Opin Crit Care* 8:1–5
- Coughlin SR (2000) Thrombin signalling and protease-activated receptors. *Nature* 407:258–264
- Paparella D, Brister SJ, Buchanan MR (2004) Coagulation disorders of cardiopulmonary bypass: a review. *Intensive Care Med* 30:1873–1881
- Amaral A, Opal SM, Vincent JL (2004) Coagulation in sepsis. *Intensive Care Med* 30:1032–1040
- Idell S (2002) Endothelium and disordered fibrin turnover in the injured lung: newly recognized pathways. *Crit Care Med* 30:S274–280
- Bertozzi P, Astedt B, Zenzius L, Lynch K, LeMaire F, Zapal W, Chapman HA Jr (1990) Depressed bronchoalveolar urokinase activity in patients with adult respiratory distress syndrome. *N Engl J Med* 322:890–897
- Su X, Camerer E, Hamilton JR, Coughlin SR, Matthay MA (2005) Protease-activated receptor-2 activation induces acute lung inflammation by neuropeptide-dependent mechanisms. *J Immunol* 175:2598–2605
- Schultz MJ, Millo J, Levi M, Hack CE, Wewerling GJ, Garrard CS, van der Poll T (2004) Local activation of coagulation and inhibition of fibrinolysis in the lung during ventilator associated pneumonia. *Thorax* 59:130–135
- Gunther A, Mosavi P, Heinemann S, Ruppert C, Muth H, Markart P, Grimminger F, Walrath D, Temmesfeld-Wollbrück B, Seeger W (2000) Alveolar fibrin formation caused by enhanced procoagulant and depressed fibrinolytic capacities in severe pneumonia. Comparison with the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 161:454–462
- Vadasz I, Morty RE, Olschewski A, Königshoff M, Kohstall MG, Ghofrani HA, Grimminger F, Seeger W (2005) Thrombin impairs alveolar fluid clearance by promoting endocytosis of Na⁺,K⁺-ATPase. *Am J Respir Cell Mol Biol* 33:343–354
- Nuckton TJ, Alonso JA, Kallet RH, Daniel BM, Pittet JF, Eisner MD, Matthay MA (2002) Pulmonary dead-space fraction as a risk factor for death in the acute respiratory distress syndrome. *N Engl J Med* 346:1281–1286
- Prudhomme J.B, Wang L, Wickersham N, Matthay M.A, Ware L.B (2005) Tissue factor levels in ARDS are 100-fold higher in edema fluid compared to plasma and correlate with mortality. *Proc Am Thoracic Soc* 2:A826
- Van de Wouwer M, Collen D, Conway EM (2004) Thrombomodulin-protein C-EPCR system: integrated to regulate coagulation and inflammation. *Arterioscler Thromb Vasc Biol* 24:1374–1383
- Sturm DH, Kaneider NC, Feistritz C, Djanani A, Fukudome K, Wiedermann CJ (2003) Expression and function of the endothelial protein C receptor in human neutrophils. *Blood* 102:1499–1505
- Nick JA, Coldren CD, Geraci MW, Poch KR, Fouty BW, O'Brien J, Gruber M, Zarini S, Murphy RC, Kuhn K, Richter D, Kast KR, Abraham E (2004) Recombinant human activated protein C reduces human endotoxin-induced pulmonary inflammation via inhibition of neutrophil chemotaxis. *Blood* 104:3878–3885
- Ware LB, Fang X, Matthay MA (2003) Protein C and thrombomodulin in human acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 285:L514–521
- Prabhakaran P, Ware LB, White KE, Cross MT, Matthay MA, Olman MA (2003) Elevated levels of plasminogen activator inhibitor-1 in pulmonary edema fluid are associated with mortality in acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 285:L20–28
- Ware LB, Eisner M, Wickersham N, Olman MA, Thompson B, Parsons P, Matthay MA (2004) Plasminogen activator inhibitor-1 (PAI-1), a marker of impaired fibrinolysis, is associated with higher mortality in patients with ALI/ARDS. *Am J Respir Crit Care Med* 169:A115
- Ware LB, Ancukiewicz M, Parsons PI, Matthay MA (2003) Prognostic value of plasma protein C in patients with acute lung injury. *Am J Respir Crit Care Med* 167:A661
- Pepe PE, Potkin RT, Reus DH, Hudson LD, Carrico CJ (1982) Clinical predictors of the adult respiratory distress syndrome. *Am J Surg* 144:124–130
- Hudson LD, Milberg JA, Anardi D, Maunder RJ (1995) Clinical risks for development of the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 151:293–301
- Kleeberger SR, Reddy S, Zhang LY, Jedlicka AE (2000) Genetic susceptibility to ozone-induced lung hyperpermeability: role of toll-like receptor 4. *Am J Respir Cell Mol Biol* 22:620–627
- Leikauf GD, McDowell SA, Wesselkamper SC, Hardie WD, Leikauf JE, Korfhagen TR, Prows DR (2002) Acute lung injury: functional genomics and genetic susceptibility. *Chest* 121:70S–75S
- Grigoryev DN, Finigan JH, Hassoun P, Garcia JG (2004) Science review: searching for gene candidates in acute lung injury. *Crit Care* 8:440–447
- Currier PF GM, Pothier LJ, Boyce PD, Thompson BT, Christiani DC (2005) Surfactant protein-B polymorphism is associated with increased mortality at 28 days in ARDS patients. *Proc Am Thoracic Soc* 2:A243
- Wattanathum A, Manocha S, Groshaus H, Russell JA, Walley KR (2005) Interleukin-10 haplotype associated with increased mortality in critically ill patients with sepsis from pneumonia but not in patients with extrapulmonary sepsis. *Chest* 128:1690–1698

31. Gao L, Grant A, Halder I, Brower R, Sevransky J, Maloney JP, Moss M, Shanholtz C, Yates CR, Umberto Meduri G, Shriver MD, Ingersoll R, Scott AF, Beaty TH, Moitra J, Ma SF, Ye SQ, Barnes KC, Garcia JG (2006) Novel polymorphisms in the myosin light chain kinase gene confer risk for acute lung injury. *Am J Respir Cell Mol Biol* 1044–1549
32. Gomi K, Tokue Y, Kobayashi T, Takahashi H, Watanabe A, Fujita T, Nukiwa T (2004) Mannose-binding lectin gene polymorphism is a modulating factor in repeated respiratory infections. *Chest* 126:95–99
33. Gong MN, Wei Z, Xu LL, Miller DP, Thompson BT, Christiani DC (2004) Polymorphism in the surfactant protein-B gene, gender, and the risk of direct pulmonary injury and ARDS. *Chest* 125:203–211
34. Gong MN, Zhou W, Williams PL, Thompson BT, Pothier L, Boyce P, Christiani DC (2005) -308GA and TNFB polymorphisms in acute respiratory distress syndrome. *Eur Respir J* 26:382–389
35. Gong MN ZW, Pothier L, Thompson BT, Christiani DC (2005) Mannose-binding lectin polymorphisms and ARDS. *Proc Am Thoracic Soc* 2:A243
36. Lin Z, Pearson C, Chinchilli V, Pietschmann SM, Luo J, Pison U, Floros J (2000) Polymorphisms of human SP-A, SP-B, and SP-D genes: association of SP-B Thr131Ile with ARDS. *Clin Genet* 58:181–191
37. Marshall RP, Webb S, Bellingan GJ, Montgomery HE, Chaudhari B, McAnulty RJ, Humphries SE, Hill MR, Laurent GJ (2002) Angiotensin converting enzyme insertion/deletion polymorphism is associated with susceptibility and outcome in acute respiratory distress syndrome. *Am J Respir Crit Care Med* 166:646–650
38. Medford AR, Keen LJ, Bidwell JL, Millar AB (2005) Vascular endothelial growth factor gene polymorphism and acute respiratory distress syndrome. *Thorax* 60:244–248
39. Quasney MW, Waterer GW, Dahmer MK, Kron GK, Zhang Q, Kessler LA, Wunderink RG (2004) Association between surfactant protein B + 1580 polymorphism and the risk of respiratory failure in adults with community-acquired pneumonia. *Crit Care Med* 32:1115–1119
40. Ye SQ, Simon BA, Maloney JP, Zambelli-Weiner A, Gao L, Grant A, Easley RB, McVerry BJ, Tuder RM, Standiford T, Brower RG, Barnes KC, Garcia JG (2005) Pre-B-cell colony-enhancing factor as a potential novel biomarker in acute lung injury. *Am J Respir Crit Care Med* 171:361–370
41. Wunderink RG, Waterer GW, Cantor RM, Quasney MW (2002) Tumor necrosis factor gene polymorphisms and the variable presentation and outcome of community-acquired pneumonia. *Chest* 121:87S
42. Russell JA (2003) Genetics of coagulation factors in acute lung injury. *Crit Care Med* 31:S243–247
43. Rebbeck TR, Spitz M, Wu X (2004) Assessing the function of genetic variants in candidate gene association studies. *Nat Rev Genet* 5:589–597
44. Aiach M, Nicaud V, Alhenc-Gelas M, Gandrille S, Arnaud E, Amiral J, Guize L, Fiessinger JN, Emmerich J (1999) Complex association of protein C gene promoter polymorphism with circulating protein C levels and thrombotic risk. *Arterioscler Thromb Vasc Biol* 19:1573–1576
45. Arnaud E, Barbalat V, Nicaud V, Cambien F, Evans A, Morrison C, Arveiler D, Luc G, Ruidavets JB, Emmerich J, Fiessinger JN, Aiach M (2000) Polymorphisms in the 5' regulatory region of the tissue factor gene and the risk of myocardial infarction and venous thromboembolism: the ECTIM and PATHROS studies. *Etude Cas-Temoins de l'Infarctus du Myocarde. Paris Thrombosis case-control Study. Arterioscler Thromb Vasc Biol* 20:892–898
46. Hermans PW, Hibberd ML, Booy R, Daramola O, Hazelzet JA, de Groot R, Levin M (1999) 4G/5G promoter polymorphism in the plasminogen-activator-inhibitor-1 gene and outcome of meningococcal disease. *Meningococcal Research Group. Lancet* 354:556–560
47. Hoekstra T, Geleijnse JM, Schouten EG, Kluft C (2004) Plasminogen activator inhibitor-type 1: its plasma determinants and relation with cardiovascular risk. *Thromb Haemost* 91:861–872
48. Norlund L, Holm J, Zoller B, Ohlin AK (1997) A common thrombomodulin amino acid dimorphism is associated with myocardial infarction. *Thromb Haemost* 77:248–251
49. Saposnik B, Reny JL, Gaussem P, Emmerich J, Aiach M, Gandrille S (2004) A haplotype of the EPCR gene is associated with increased plasma levels of sEPCR and is a candidate risk factor for thrombosis. *Blood* 103:1311–1318
50. Uitte de Willige S, Van Marion V, Rosendaal FR, Vos HL, de Visser MC, Bertina RM (2004) Haplotypes of the EPCR gene, plasma sEPCR levels and the risk of deep venous thrombosis. *J Thromb Haemost* 2:1305–1310
51. Endler G, Mannhalter C (2003) Polymorphisms in coagulation factor genes and their impact on arterial and venous thrombosis. *Clin Chim Acta* 330:31–55
52. Kottke-Marchant K (2002) Genetic polymorphisms associated with venous and arterial thrombosis: an overview. *Arch Pathol Lab Med* 126:295–304
53. Lane DA, Grant PJ (2000) Role of hemostatic gene polymorphisms in venous and arterial thrombotic disease. *Blood* 95:1517–1532
54. Eriksson P, Kallin B, van 't Hooft FM, Bavenholm P, Hamsten A (1995) Allele-specific increase in basal transcription of the plasminogen-activator inhibitor 1 gene is associated with myocardial infarction. *Proc Natl Acad Sci USA* 92:1851–1855
55. Henry M, Tregouet DA, Alessi MC, Aillaud MF, Visvikis S, Siest G, Tiret L, Juhan-Vague I (1998) Metabolic determinants are much more important than genetic polymorphisms in determining the PAI-1 activity and antigen plasma concentrations: a family study with part of the Stanislas Cohort. *Arterioscler Thromb Vasc Biol* 18:84–91
56. McCormack LJ, Semple JI, Stickland MH, Barrett JH, Grant PJ (1998) The effect of number of days in culture and plasminogen activator inhibitor-1 (PAI-1) 4G/5G genotype on PAI-1 antigen release by cultured human umbilical vein endothelial cells. *Thromb Res* 92:199–206
57. Barnes KC (2005) Genetic determinants and ethnic disparities in sepsis-associated acute lung injury. *Proc Am Thorac Soc* 2:195–201
58. Khadaroo RG, Marshall JC (2002) ARDS and the multiple organ dysfunction syndrome. Common mechanisms of a common systemic process. *Crit Care Clin* 18:127–141
59. Rubenfeld GD, Christie JD (2004) The epidemiologist in the intensive care unit. *Intensive Care Med* 30:4–6

60. Cardon LR, Bell JI (2001) Association study designs for complex diseases. *Nat Rev Genet* 2:91–99
61. Risch NJ (2000) Searching for genetic determinants in the new millennium. *Nature* 405:847–856
62. Romero R, Kuivaniemi H, Tromp G, Olson J (2002) The design, execution, and interpretation of genetic association studies to decipher complex diseases. *Am J Obstet Gynecol* 187:1299–1312
63. Lander ES (1996) The new genomics: global views of biology. *Science* 274:536–539
64. Browner W (2001) Estimating sample size. In Browner W, Newman T, Cummings S, Hulley S (eds) *Designing clinical research: an epidemiologic approach*. Lippincott Williams & Wilkins, Philadelphia, p 75
65. Newton-Cheh C, Hirschhorn JN (2005) Genetic association studies of complex traits: design and analysis issues. *Mutat Res* 573:54–69
66. Lyon H, Lange C, Lake S, Silverman EK, Randolph AG, Kwiatkowski D, Raby BA, Lazarus R, Weiland KM, Laird N, Weiss ST (2004) IL10 gene polymorphisms are associated with asthma phenotypes in children. *Genet Epidemiol* 26:155–165
67. Ozturk A, Dekosky ST, Kamboh MI (2005) Genetic variation in the choline acetyltransferase (CHAT) gene may be associated with the risk of Alzheimer's disease. *Neurobiol Aging* DOI: 10.1016/j.neurobiolaging.2005.08.02.4
68. Prince JA, Feuk L, Gu HF, Johansson B, Gatz M, Blennow K, Brookes AJ (2003) Genetic variation in a haplotype block spanning IDE influences Alzheimer disease. *Hum Mutat* 22:363–371
69. Wanic K, Malecki MT, Wolkow PP, Klupa T, Skupien J, Bobrek J, Kozek E, Krolewski AS, Sieradzki J (2006) Polymorphisms in the gene encoding hepatocyte nuclear factor-4alpha and susceptibility to type 2 diabetes in a Polish population. *Diabetes Metab* 32:86–88
70. Mattevi VS, Zembruski VM, Hutz MH (2006) Impact of variation in ADRB2, ADRB3, and GNB3 genes on body mass index and waist circumference in a Brazilian population. *Am J Hum Biol* 18:182–186
71. Silverman EK, Palmer LJ (2000) Case-control association studies for the genetics of complex respiratory diseases. *Am J Respir Cell Mol Biol* 22:645–648
72. Morrison J, Gauderman J (2004) Quanto, accessed at <http://hydra.usc.edu/GxE/>
73. Tracey KJ, Warren HS (2004) Human genetics: an inflammatory issue. *Nature* 429:35–37
74. Carlson CS, Eberle MA, Kruglyak L, Nickerson DA (2004) Mapping complex disease loci in whole-genome association studies. *Nature* 429:446–452
75. Crawford DC, Carlson CS, Rieder MJ, Carrington DP, Yi Q, Smith JD, Eberle MA, Kruglyak L, Nickerson DA (2004) Haplotype diversity across 100 candidate genes for inflammation, lipid metabolism, and blood pressure regulation in two populations. *Am J Hum Genet* 74:610–622
76. Crawford DC, Nickerson DA (2005) Definition and clinical importance of haplotypes. *Annu Rev Med* 56:303–320
77. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I (2001) Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* 125:279–284
78. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
79. Kohler K, Bickeboller H (2006) Case-control association tests correcting for population stratification. *Ann Hum Genet* 70:98–115
80. Devlin B, Roeder K, Wasserman L (2001) Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol* 60:155–166
81. Bacanu SA, Devlin B, Roeder K (2000) The power of genomic control. *Am J Hum Genet* 66:1933–1944
82. Gong MN, Thompson BT, Williams PL, Zhou W, Wang MZ, Pothier L, Christiani DC (2006) Interleukin-10 polymorphism in position -1082 and acute respiratory distress syndrome. *Eur Respir J* 27:674–681