

Genetic polymorphism in allergy and asthma

Donata Vercelli

The impact of genetic factors on the pathogenesis of allergy and asthma is currently an area of intense investigation.

Although epidemiological studies find more and more genes that are likely to contribute to allergic inflammation, functional studies of the mechanisms that link genetic variation with dysregulation of gene expression and function are hindered by the high frequency of natural variation. The interplay between genes and environment adds another layer of complexity to the determination of allergic phenotypes. The complexity emerging from these studies warrants novel experimental strategies and bolder conceptual paradigms.

Addresses

Functional Genomics Laboratory, Arizona Respiratory Center,
University of Arizona Health Sciences Center, Room 2349, 1501 North
Campbell Avenue, Tucson, AZ 85724, USA
e-mail: donata@resp-sci.arizona.edu

Current Opinion in Immunology 2003, 15:609–613

This review comes from a themed issue on
Allergy and hypersensitivity
Edited by Raif S Geha

0952-7915/\$ – see front matter
© 2003 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.coi.2003.09.005

Abbreviations

CRS Children Respiratory Study
LPS lipopolysaccharide
MAS Multicenter Allergy Study
SNP single nucleotide polymorphism
Th T helper
TLR Toll-like receptor

Introduction

The recent advent of high-throughput methods for gene sequencing and genotyping, coupled with a growing interest in functional studies, has resulted in the generation of an impressive amount of data on the genetics of allergy and asthma. Overall, these studies have reiterated the essential role played by genetic factors in the pathogenesis of allergic disease.

Such analyses have resulted in the emergence of three interrelated themes: the heuristic value of genetic studies, which have the power to identify new candidate genes; the remarkable frequency of natural variation even in highly conserved genes; and the interplay between genes and environment. In this review, I discuss each of these three themes.

Uncovering the unexpected: the heuristic value of genetic studies

These days, a common trend in biology is to complement hypothesis-driven research with hypothesis-independent, global types of analysis. In studies of the genetics of asthma and allergy, this means pursuing the identification of putative allergy or asthma susceptibility genes by positional cloning in outbred populations, and then characterizing the associations between patterns of variation in those candidate genes and specific disease phenotypes.

To date, linkage studies have identified more than a dozen genomic regions that are associated with asthma and/or increased serum concentrations of IgE [1]. For certain regions (e.g. the Th2 cytokine locus on chromosome 5q31), identifying some of the genes that are likely to contribute to the linkage signal (e.g. IL-13, IL-4, CD14) was relatively straightforward because the pathogenetic role of these genes was already known, or at least predicted. In other studies, deciphering linkage revealed the existence of novel genes.

The identification of *ADAM33* is a good example of gene discovery driven by genetic approaches. At the end of 2002, a genome-wide scan of 460 Caucasian families identified a locus on chromosome 20p13 that was linked to asthma and bronchial hyperresponsiveness. A survey of several single nucleotide polymorphisms (SNPs) in 23 genes showed that *ADAM33* was significantly associated with asthma in case-control, transmission disequilibrium and haplotype studies [2•].

ADAM proteins are membrane-anchored metalloproteases with diverse functions [3], including a role in the shedding of cell-surface proteins such as cytokines and cytokine receptors [4]. Notably, ADAM33 is expressed in lung fibroblasts and bronchial smooth muscle, but not in bronchial epithelial or immune cells [2•]. This expression pattern, combined with the known functions of ADAM family members in promoting myogenic fusion and the release of proliferative growth factors [3], supports the idea that ADAM33 has a role in airway remodeling rather than in immune-mediated inflammation. Epithelial damage, smooth muscle hyperplasia and increased matrix deposition are important characteristics of asthmatic airways that are thought to contribute to bronchial hyperresponsiveness and may be controlled in part by ADAM33.

The strong heuristic value of positional cloning studies has been reiterated in the recent identification of a quantitative trait locus on chromosome 13q14 that influences serum concentrations of IgE and asthma [5•].

Chromosome 13q14 had shown consistent linkage to atopy (a state of IgE-mediated hypersensitivity towards environmental allergens) and total serum IgE concentrations [6,7]. Further studies narrowed the locus associated with atopy down to a region containing a novel micro-satellite, USAT24G1 [8]. The underlying quantitative trait locus was localized in a comprehensive SNP map, and several alleles in a single gene, *PHF11*, were found to be associated with IgE and severe asthma [5**].

PHF11 is an interesting candidate gene for allergy susceptibility. Indeed, the *PHF11* gene product contains two plant homeodomain zinc-finger domains, which suggests that it may be involved in chromatin-mediated transcriptional regulation [9]. Furthermore, distinct splice variants of *PHF11* have been found to be differentially expressed in T and B lymphocytes in different states of activation [5**]. Last, but not least, *PHF11* is closely flanked by two other genes, *CLLD7* and *CLLD8*.

Preliminary evidence suggests that all three genes may be coordinately regulated. Of note, *CLLD8* probably encodes a histone methyltransferase, whereas *CLLD7* shows sequence similarity to genes known to be involved in the regulation of chromatin condensation. The narrow segment of chromosome 13q14 that contains *PHF11*, *CLLD8* and *CLLD7* is within the minimum region commonly deleted in B cell chronic lymphocytic leukemias. Furthermore, one of the *PHF11* alleles that is not associated with serum IgE is associated with high concentrations of serum IgM. Thus, although the variation in IgE seems to be attributable to variation in *PHF11* and not *CLLD8* and *CLLD7*, this whole locus is an interesting target for studying the regulation of B cell clonal expansion and immunoglobulin expression.

So are *ADAM33* or *PHF11* 'the' asthma genes? Probably not, because it is clear that asthma, although often defined as a single disease, in reality includes several different phenotypes, each of which is likely to reflect different pathogenetic mechanisms. Indeed, an important reason why the analysis of the genetic components of asthma has proved so difficult probably lies in our inability to define phenotypes strictly enough to identify genetically homogeneous populations.

Genetic variation: so frequent, so deep

Basic immunology is making wider and wider use of animal models based on inbred strains. The genetic manipulations and permutations possible in such models enables a level of experimental elegance that is unattainable in human studies, which accounts for the popularity of this approach. Needless to say, using inbred strains of laboratory animals to model the function and behavior of human genes presupposes fundamental genetic similarities between these animals and humans, as well as within the human species. However,

the data currently emerging from studies of human genetic variation highlight a situation that is far more complex and warrant caution in generalizing to humans results that have been obtained in genetically homogeneous animal strains.

High-throughput sequencing methods can facilitate the fast in-depth analysis of a large number of genes. These methods, coupled with aggressive bioinformatics approaches for extensive sequence comparison, have led to the realization that variation in the human genome is incredibly pervasive. Millions of SNPs have been already identified [10], and the count is continuously rising.

As an example, Table 1 shows the number of SNPs in innate immunity genes belonging to the Toll-like receptor (TLR) response pathway detected by researchers of the National Heart Lung and Blood Institute (NHLBI)-sponsored Innate Immunity Program for Genomic Applications. All of these genes are highly conserved among species (in fact, the gene encoding human TLR4 was cloned on the basis of its homology with *Drosophila* Toll [11]), and yet they are all highly polymorphic. The high frequency at which non-conservative SNPs occur in the coding regions of these genes is particularly impressive. The genes encoding TLR2, TLR4, TLR6 and TLR9 contain four, six, seven and four non-conservative coding region SNPs, respectively. These SNPs are expected to affect the structure of the receptor protein and thus the response that the receptor mediates. Of note, these SNPs are not particularly rare: their cumulative frequency in the test population was 13.3% (TLR2), 25.5% (TLR4), 52.2% (TLR6) and 6.6% (TLR9).

Table 1

SNPs in innate immunity genes.

Gene	Number of SNPs
<i>TLR1</i>	77
<i>TLR2</i>	28
<i>TLR3</i>	23
<i>TLR4</i>	44
<i>TLR5</i>	143
<i>TLR6</i>	53
<i>TLR7</i>	99
<i>TLR8</i>	72
<i>TLR9</i>	22
<i>TLR10</i>	76
<i>MD-1</i>	32
<i>MD-2</i>	15
<i>CD14</i>	37
<i>MYD88</i>	8
<i>RP105</i>	88
<i>TOLLIP</i>	199
<i>LBP</i> (lipopolysaccharide-binding protein)	77
<i>BPI</i> (bactericidal/permeability-increasing protein)	86

Please visit the Innate Immunity Program for Genomic Application Website (<http://innateimmunity.net>) for further information.

The high frequency at which natural variation occurs in humans has important implications for studies aimed at understanding the genetic basis of disease. The realization that variation is at the same time both frequent and complex is prompting changes in genotyping strategies. Only a couple of years ago, most published association studies involved isolated SNPs. Now, more and more often, haplotypes are determined in an effort to capture all of the genetic diversity of the test population.

Of course, an immediate consequence of haplotype-based approaches is that the populations originally recruited to analyze one SNP at a time are too small to allow analyses of high statistical power, given the low frequency at which individual haplotypes appear in the population. Theoretically, this problem may be circumvented by meta-analyses, that is, by merging data from different populations. Unfortunately, this approach may often be doomed by the profound differences in allele frequency among populations of different ethnic background and by the need to take into account gene–environment interactions (see below). Thus, future association studies will require recruiting new populations of significantly larger size — a task that will not be easy.

The complexity of genetic variation patterns also has important consequences for functional genomic studies aimed at determining the mechanisms by which genetic variation affects disease. The crucial question nowadays is how to choose candidates for further functional analysis in a complex haplotype that includes several SNPs. Functional studies are in their infancy, are still low-throughput, and thus are difficult and time-consuming. Rational criteria to prioritize the choice of SNPs for functional analysis are not readily available, but are badly needed. Recent advances in comparative genomics and phylogenetic analysis may help to circumvent this stumbling block. Phylogenetic shadowing has been proposed as a tool to identify genomic regions of likely functional significance through sequence comparisons across primates [12^{••}]. This approach is based on sequence comparison of an extensive set of ‘old world’ and ‘new world’ monkeys and hominoids, and was able to identify primate-specific as well as common mammalian gene regulatory elements. Of note, several of these elements would have gone undetected when evaluating more evolutionarily distant species [12^{••}]. SNPs in regions identified through phylogenetic shadowing may represent good candidates for further functional studies.

Gene–environment interactions: the endotoxin switch

In 1999, a SNP in the CD14 promoter, CD14-159C/T, was found to be associated with increased amounts of soluble CD14 and decreased total serum IgE in the Children Respiratory Study (CRS), a study of an unse-

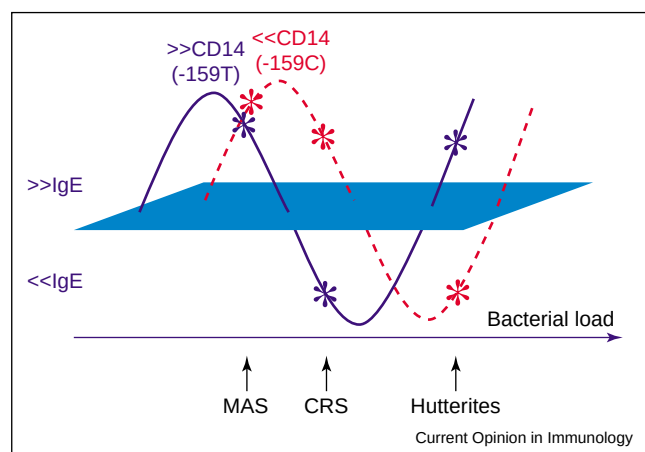
lected population enrolled at birth in Tucson, Arizona [13]. This work received wide attention, because the role of innate immunity in sensing the environment and regulating adaptive responses was becoming increasingly clear at that time. At the beginning of 2003, another group did not find such an association in the German Multicenter Allergy Study (MAS) cohort [14]. To complicate matters even further, an association between atopy and CD14-159T, rather than CD14-159C, was reported in members of the Hutterite community of South Dakota [15].

Inconsistencies among molecular epidemiology findings in different populations are not isolated episodes and often signal the existence of complex gene–environment interactions. In this context, CD14 represents a model case. CD14 is part of the TLR–lipopolysaccharide (LPS) signaling pathway, and as such is an essential component of the genetic interface between the host and the environment. Although the role of innate immunity in shaping allergic responses has been explored in various animal models, data on the role of LPS in the regulation of Th2 responses and in IgE production are apparently conflicting [16]. Both protection and exacerbation have been reported.

Part of this discrepancy depends on the time at which pathogen exposure occurs. In a rat model, administration of LPS before allergen challenge was strongly protective against sensitization, whereas administration after sensitization resulted in exacerbation of inflammation [17]. By contrast, mice carrying a mutation in TLR4 that caused LPS unresponsiveness were incapable of mounting strong antigen-specific Th2 responses owing to deficient dendritic cell maturation following pathogen exposure [18]. Even more intriguingly, when mice were immunized intra-nasally with an ovalbumin preparation that had been rigorously purified and lacked LPS, no response was elicited. When a low or high dose of LPS was added to ovalbumin at the time of immunization, however, mice given low-dose LPS developed Th2 responses, whereas those given high-dose LPS developed strong Th1 responses. The cells recruited to the lung were eosinophils in the first case, and neutrophils in the second case [19^{••}].

These experiments uncovered a complex, dose-dependent, non-linear relationship between environmental exposure to bacterial products and the outcome of immune responses. We have named this relationship the ‘endotoxin switch’ [20]. The mechanisms underlying the switch may involve differential induction of dendritic cell maturation and migration, which in turn may result in a differential effect on T regulatory cell function aimed at controlling Th2 responses [21]. A functional implication of the endotoxin switch is that antigen exposure may result in opposite responses depending on the environmental context.

Figure 1



The endotoxin switch. A functional implication of the endotoxin switch is that antigen exposure may result in different responses depending on the environmental context. Here, in the presence of a low bacterial load, such as that typically encountered in the West, the response may be shifted to the leftmost part of the curve, where Th2 responses are vigorous and IgE is high. Polymorphisms in genes at the host–environment interface (CD14 being a prototypical gene) will modulate immune responses and will reposition the curve along the bacterial load axis. Each haplotype will have a distinct dose–response curve and a distinct pathogen concentration required for the switch. Thus, carriers of CD14-159T with higher expression of CD14 will require fewer pathogen products to switch to low IgE responses, and will therefore be more protected from allergic inflammation than will carriers of CD14-159C. Conversely, extremely high endotoxin loads may bring about a second switch, again towards strong Th2 responses, which will now preferentially affect individuals with higher CD14 expression (CD14-159T carriers). For very low or very high endotoxin loads, responses will be similar across the population. Thus, the different impact of CD14-159C/T on the Children Respiratory Study (CRS), Multicenter Allergy Study (MAS) and Hutterite populations (arrowheads) may reflect low, intermediate and high levels of exposure to pathogen products. Asterisks indicate the expected level of exposure to bacterial products for each population.

Evolution may have shaped our immune system so that it recognizes antigens in a pathogen-rich environment [22]. As shown in Figure 1, in the presence of a low bacterial load, such as that typically encountered in Western countries and lifestyles, the response may be shifted to the leftmost part of the curve, where Th2 responses are vigorous and IgE levels are high. In this scheme, polymorphisms in genes functioning at the host–environment interface (CD14 being a prototypical one) would modulate immune responses and would reposition the curve along the bacterial load axis. Each haplotype would have a distinct dose–response curve and a distinct pathogen concentration required for the switch. Thus, carriers of CD14-159T with higher CD14 expression would require fewer pathogen products to switch to low IgE responses and would therefore be more protected from allergic inflammation than would CD14-159C individuals. Conversely, extremely high endotoxin loads may bring about

a second switch, again towards strong Th2 responses, which would now preferentially affect individuals with higher expression of CD14 (i.e. CD14-159T individuals). For very low or very high endotoxin loads, the responses would be similar across the population. Thus, we propose that the different impact of CD14-159C/T on the CRS, MAS and Hutterite populations may reflect low, intermediate and high levels of exposure to pathogen products.

A final, crucial corollary to the impact of gene–environment interactions on the regulation of allergic responses is that association studies in populations with different levels of the relevant exposure may highlight different genes. Furthermore, and perhaps even more importantly, the phenotypic impact of exposure may be detectable only in individuals carrying specific haplotypes or SNPs. Data confirming this hypothesis are now emerging. In the ALEX population [23,24], which is currently the prototype of a rural lifestyle, farm milk consumption in the first year of life was found to be protective against asthma, hay fever and pollen sensitization only in individuals carrying at least one CD14-1720A allele [25].

Conclusions

Far from simplifying our understanding of the pathogenesis of allergic disease, the advances in the genetic studies of asthma and allergy discussed above highlight a situation of remarkable complexity that warrants novel approaches. Haplotypes, rather than single SNPs, will have to be evaluated. New strategies will be needed to tease out the mechanisms through which natural variation has an impact on gene expression and/or function. Animal models will have to be improved to recapitulate better the regulation of allergic responses in outbred populations. Finally, gene–environment interactions will have to be taken into account when interpreting the results of genetic association studies. None of this will be easy, but it is all tremendously exciting.

Acknowledgements

The author was supported by grants PO1HL67672 and RO1HL66391 from the National Institutes of Health, and by the Program for Genomic Applications Innate Immunity in Heart, Lung and Blood Disease (U01-HL66803) from the National Heart, Lung and Blood Institute.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
 - of outstanding interest
1. Wjst M, Immervoll T: **An internet linkage and mutation database for the complex phenotype asthma.** *Bioinformatics* 1998, 14:827–828.
 2. Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J, Torrey D, Pandit S, McKenny J, Braunschweiler K et al.: **Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness.** *Nature* 2002, 418:426–430.

An important result of positional cloning studies, and strong evidence for the role of genetic dysregulation of non-immune pathways in the pathogenesis of asthma.

3. Yoshinaka T: **Identification and characterization of novel mouse and human ADAM33s with potential metalloprotease activity.** *Gene* 2002, **282**:227-236.
 4. Mullberg J, Althoff K, Jostock T, Rose-John S: **The importance of shedding of membrane proteins for cytokine biology.** *Eur Cytokine Netw* 2000, **11**:27-38.
 5. Zhang Y, Leaves NI, Anderson GG, Ponting CP, Broxholme J, ●● Holt R, Edser P, Bhattacharyya S, Dunham A, Adcock IM *et al.*: **Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma.** *Nat Genet* 2003, **34**:181-186.
- By using positional cloning, this group identify a novel gene that seems to be important for regulating serum IgE.
6. Daniels SE: **A genome-wide search for quantitative trait loci underlying asthma.** *Nature* 1996, **383**:247-250.
 7. Kimura K: **Linkage and association of atopic asthma to markers on chromosome 13 in the Japanese population.** *Hum Mol Genet* 1999, **8**:1487-1490.
 8. Anderson GG, Leaves NI, Bhattacharyya S, Zhang Y, Walshe V, Broxholme J, Abecasis G, Levy E, Zimmer M, Cox R, Cookson WO: **Positive association to IgE levels and a physical map of the 13q14 atopy locus.** *Eur J Hum Genet* 2002, **10**:266-270.
 9. Aasland R, Gibson TJ, Stewart AF: **The PHD finger: implications for chromatin-mediated transcriptional regulation.** *Trends Biochem Sci* 1995, **20**:56-59.
 10. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL *et al.*: **A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms.** *Nature* 2001, **409**:928-933.
 11. Medzhitov R, Preston-Hurlburt P, Janeway CA: **A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity.** *Nature* 1997, **388**:394-397.
 12. Boffelli D, McAuliffe J, Ovcharenko D, Lewis KD, Ovcharenko I, ●● Pachter L, Rubin EM: **Phylogenetic shadowing of primate sequences to find functional regions of the human genome.** *Science* 2003, **299**:1391-1394.
- A brilliant study that proposes an original perspective on comparative genomics and may facilitate the identification of novel functional elements in genes of interest.
13. Baldini M, Lohman IC, Halonen M, Erickson RP, Holt PG, Martinez FD: **A polymorphism in the 5' flanking region of the CD14 gene is associated with circulating soluble CD14 levels and with total serum IgE.** *Am J Respir Cell Mol Biol* 1999, **20**:976-983.
 14. Sengler C, Haider A, Sommerfeld C, Lau S, Baldini M, Martinez F, Wahn U, Nickel R: **Evaluation of the CD14C-159T polymorphism in the German MAS cohort.** *Clin Exp Allergy* 2003, **33**:166-169.
 15. Ober C, Tsalenko A, Parry R, Cox NJ: **A second-generation genomewide screen for asthma-susceptibility alleles in a founder population.** *Am J Hum Genet* 2000, **67**:1154-1162.
 16. Liu AH: **Endotoxin exposure in allergy and asthma: reconciling a paradox.** *J Allergy Clin Immunol* 2002, **109**:379-392.
 17. Tulic MK, Wale JL, Holt PG, Sly PD: **Modification of the inflammatory response to allergen challenge after exposure to bacterial lipopolysaccharide.** *Am J Respir Cell Mol Biol* 2000, **22**:604-612.
 18. Dabbagh K, Dahl ME, Stepick-Biek P, Lewis DB: **Toll-like receptor 4 is required for optimal development of Th2 immune responses: role of dendritic cells.** *J Immunol* 2002, **168**:4524-4530.
 19. Eisenbarth SC, Piggott DA, Huleatt JW, Visintin I, Herrick CA, ●● Bottomly K: **Lipopolysaccharide-enhanced, Toll-like receptor 4-dependent T helper cell 2 responses to inhaled antigen.** *J Exp Med* 2002, **196**:1645-1651.
- A seminal work that, for the first time, addresses directly and in quantitative terms the impact of endotoxin exposure on T helper cell responses.
20. Vercelli D: **Learning from discrepancies: CD14 polymorphisms, atopy and the endotoxin switch.** *Clin Exp Allergy* 2003, **33**:153-155.
 21. Herrick CA, Bottomly K: **To respond or not to respond: T cells in allergic asthma.** *Nat Rev Immunol* 2003, **3**:405-412.
 22. Medzhitov R, Janeway CA Jr: **Innate immunity: the virtues of a nonclonal system of recognition.** *Cell* 1997, **91**:295-298.
 23. Riedler J, Braun-Fahrlander C, Eder W, Waser M, Maisch S, Carr D, Schierl R, Nowak D, Von Mutius E: **Early life exposure to farming environment is essential for protection against the development of asthma and allergy: a cross-sectional survey.** *Lancet* 2001, **358**:1129-1133.
 24. Braun-Fahrlander C, Riedler J, Herz U, Eder W, Waser M, Grize L, ●● Maisch S, Carr D, Gerlach F, Bufe A *et al.*: **Environmental exposure to endotoxin and its relation to asthma in school-age children.** *N Engl J Med* 2002, **347**:869-877.
- A crucial demonstration of the *in vivo* protection conferred by environmental endotoxin against asthma and allergy.
25. Eder W, Klimecki W, Yu L, von Mutius E, Riedler J, Braun-Fahrlander C, Nowak D, Martinez FD: **A polymorphism in the CD14 gene is associated with asthma and allergies in rural children exposed to farm milk in early life.** *Am J Respir Crit Care Med* 2003, **167**:A578.