REVIEWS

Discovering susceptibility genes for asthma and allergy

Donata Vercelli

Abstract | Asthma and asthma-related traits are complex diseases with strong genetic and environmental components. Rapid progress in asthma genetics has led to the identification of several candidate genes that are associated with asthma-related traits. Typically the phenotypic impact of each of these genes, including the ones most often replicated in association studies, is mild, but larger effects may occur when multiple variants synergize within a permissive environmental context. Despite the achievements made in asthma genetics formidable challenges remain. The development of novel, powerful tools for gene discovery, and a closer integration of genetics and biology, should help to overcome these challenges.

Although it has long been recognized that asthma and asthma-related traits run strongly in families and have a clear hereditary component[1], present-day genetics of these conditions is relatively young, having been inaugurated in the late-1980s and mid-1990s with the publication of the first genome-wide screens for allergy[2] and asthma[3] susceptibility loci. Rapid advances in the field of genetics have unveiled several pathways that are crucial for asthma pathogenesis and, perhaps more importantly, have shown that asthma and asthma-related traits behave as typical complex diseases. That is, conditions in which various genetic hits that are individually mild may be capable of major phenotypic effects when acting in concert within a permissive environmental context. This multifactorial pathogenesis differentiates complex diseases from Mendelian disorders, which are caused by single gene mutations, and renders their genetic components more elusive and difficult to trace.

Asthma genetics has made much progress over the past decade, and will advance even further in the near future, as increasingly powerful analytical tools are developed to unlock the complexities of genetic diseases. However, some formidable challenges still remain for asthma geneticists: the identification of all the genes involved in the disease, the mechanisms underlying the phenotypic heterogeneity of asthma, the difficulties in replicating associations between genotype and phenotype across populations and last, but not least, understanding how environmental and developmental factors interact with genetic determinants to affect disease susceptibility. This Review discusses the tools developed by asthma geneticists to understand how genes affect disease susceptibility, what has been learned from asthma genetics in the past decade, and what remains to be learned.

Functional Genomics Laboratory, Arizona Respiratory Center; Department of Cell Biology and The Bio5 Institute, University of Arizona, PO Box 245044, Tucson, Arizona 85724-5044, USA.
e-mail: donata@arc.arizona.edu
doi:10.1038/nri2257

The toolkit of asthma geneticists

Until a few years ago, scientists interested in discovering the genetic basis of inherited diseases relied primarily on two approaches: genome-wide linkage studies and candidate-gene association studies (BOX 1). Genome-wide linkage studies focus on families that comprise individuals affected by the disease of interest. Family members are genotyped with evenly spaced genetic markers covering all chromosomes, and a search is made for genetic regions containing a higher-than-expected number of shared alleles among the affected individuals. If such a region is discovered — which often spans 20–30 million base pairs (bp) of DNA and contains hundreds of genes — this suggests that somewhere within this genomic interval there is a disease-predisposing allele. This region is said to be linked (that is, co-inherited) with the disease in question. The genes in this region become positional candidates and are examined further by typing denser and denser collections of genetic variants until the underlying disease-associated gene or genes are identified.

The use of linkage studies is typically motivated by two considerations. This approach promises the identification of new genes and pathways, as it requires no a priori hypotheses and allows for studies of increasingly high resolution. Moreover, linkage analyses were deemed to be ideal for identifying alleles with large phenotypic effects, because only such alleles were expected to generate linkage signals in the small population samples that

Box 1 | **Candidate gene association studies and genome-wide linkage analyses**

An advantage of linkage analysis within a family is that, as closely related individuals tend to share large regions of the genome that are inherited from the same recent ancestor, genotyping fewer than 500 polymorphic markers across the genome is generally adequate to detect linked regions. In linkage studies, any polymorphism between a pair of linked markers will be associated with both markers. As a result, linkage maps have a logical hierarchical structure wherein an initial genome scan can be performed at low density (fewer than 500 markers) and additional markers can then be used to fine-map the boundaries of linked regions. However, the small number of recombination events within most families makes it difficult to narrow the region of interest to below several megabases, even though the availability of the complete sequence of the human genome has accelerated candidate gene analyses within linked regions[7].

An important practical difference between genome-wide linkage studies and gene association studies is that linkage disequilibrium assessed through association does not have a hierarchical structure[7]. Currently, association analyses are typically indirect, that is, a dense set of markers across a region is genotyped and tested for association. The assumption is that the genotype at polymorphisms in the region that are not genotyped directly will correlate, that is, will be in linkage disequilibrium, with one or more of the markers. If the degree of linkage disequilibrium between typed markers and adjacent unassayed polymorphisms is low, the power to detect disease associations at the unassayed polymorphism will be low even though the population sample may be large. In association analysis, a given single nucleotide polymorphism (SNP) located between two highly linked ones may not be strongly associated with either polymorphism, and therefore can easily go undetected. Therefore, association analyses can typically identify disease-risk variants only when such variants are strongly associated with a tag SNP.

are typically used for genetic studies[4]. Replication of linkage in multiple populations was thought to further support the results of linkage analysis. The reality, however, may be more complex. Many, and perhaps most, of the regions that show linkage to asthma or asthma-related traits in multiple populations may in fact house multiple susceptibility loci, each with small effects on disease risk. For instance, the well-replicated asthma and atopy linkage region in chromosome 5q31–33 actually includes 14 genes that are associated with asthma, atopy or related intermediate phenotypes, some of which — for example, interleukin 4 (*IL4*), *IL13*, *CD14*, the β_2 adrenergic receptor (*ADRB2*), and serine peptidase inhibitor, Kazal type 5 (*SPINK5*) — are among the most replicated in allergic inflammation[4]. The same is true of a broad region on chromosome 5p, which harbours at least two, and possibly more, susceptibility loci for asthma or bronchial hyper-responsiveness separated by >9 Mb[5]. Therefore, the only advantage to genome-wide linkage studies is the potential for discovery of new genes and pathways relevant to asthma, and even this advantage is rapidly fading away with the advent and increasing feasibility of genome-wide association (GWA) studies.

Candidate-gene association studies (BOX 1) focus on a selected number of genes that have been implicated as having a role in disease pathogenesis. Association studies between variants in candidate genes and relevant phenotypes are mostly conducted by comparing allele or genotype frequencies between groups of unrelated cases and unrelated controls. Association analysis is expected to be more powerful for the detection of common disease alleles that confer modest disease risk[6] in sample sizes that are comparable with those used in linkage studies. This reflects the fact that for modest-risk alleles the patterns of allele sharing among affected individuals within a family are less striking than patterns of allele sharing between unrelated affected individuals. Another advantage of association analysis is that it is easier to recruit large numbers of unrelated affected individuals than it is to collect large numbers of families, and there is increased statistical power in studying the equivalent number of individuals in a case-control association study than in family-based linkage studies. However, the region around a marker that is shared identically by descent in unrelated, affected individuals will be much smaller than the shared region for related individuals because of the higher number of generations from the most recent common ancestor. Therefore, association analysis requires higher marker densities than linkage analysis[7].

The main criticism aimed at case-control studies has been that an unknown population structure could result in differences in allele frequencies between cases and controls that are unrelated to their disease status. After the development of new statistical methods, which allow for the detection and correction of any imbalances between cases and controls, the advantages of family-based linkage analysis may not outweigh the power and logistical advantages provided by association studies[4]. Nevertheless, regardless of the method used to discover genes that putatively modify disease susceptibility, a causative role for such genes can only be established through further functional characterization of the genes and their variants, *in vivo* or *in vitro* studies with chemical or pharmacological inhibitors and, possibly, genetic manipulations in animal models.

Known susceptibility genes in asthma and allergy

Most of the genes currently known to modify asthma and allergy susceptibility have been identified through hypothesis-driven studies that sought to identify an association between variants such as single nucleotide polymorphisms (SNPs) in the main pathways that influence allergic inflammation and asthma or asthma-related (intermediate) phenotypes. Asthma-related phenotypes typically include respiratory traits (for example, wheezing, bronchial hyper-responsiveness, lung function parameters); immunological traits (for example, total and specific serum IgE levels, atopy); and clinical traits (for example, atopic dermatitis and/or eczema and rhinoconjunctivitis).

Several excellent articles have reviewed recent advances in the genetics of asthma and allergy for a general and

Atopy
The propensity of an individual to develop allergic diseases, such as asthma, atopic dermatitis, food allergy or hay fever. It is defined operationally by elevations in serum levels of IgE reactive with allergens or by skin-test reactivity to allergens.

Population structure
Any deviation from the ideal state of a single population in which every individual has the same chance of mating with every other.

Single nucleotide polymorphisms
(SNPs). Variations in DNA sequence in which one of the four nucleotides is substituted for another (for example, C for A). SNPs are the most frequent type of polymorphism in the genome.

Tag SNP
A single nucleotide polymorphism (SNP) that is correlated with a neighbouring variant, which serves as a proxy for that (not genotyped) variant.

Gene	Chromosome	Function and pathway	Common variants
GSTM1	1p13.3	Environmental and oxidative stress — detoxification	+/null
FLG	1q21.3	Epithelial barrier integrity	Arg510X, 2282del4
IL10	1q31-q32	Immunoregulation	–1082A/G, –571C/A
CTLA4	2q33	T-cell-response inhibition and immunoregulation	–318C/T, 49A/G
IL13	5q31	T_H2 effector functions	–1112C/T, Arg130Gln
IL4	5q31.1	T_H2 differentiation and IgE induction	–589C/T, +33C/T
CD14	5q31.1	Innate immunity — microbial recognition	–1721G/A, –260C/T
SPINK5	5q32	Epithelial serine protease inhibitor	Glu420Lys
ADRB2	5q31-q32	Bronchial smooth-muscle relaxation	Arg16Gly, Gln27Glu
HAVCR1	5q33.2	T-cell-response regulation — HAV receptor	5383_5397del
LTC4S	5q35	Cysteinyl leukotriene biosynthesis — inflammation	–444A/C
LTA	6p21.3	Inflammation	*Nco*I (intron 1)
TNF	6p21.3	Inflammation	–308G/A, –857C/T
HLA-DRB1	6p21	Antigen presentation	Multi-SNP alleles
HLA-DQB1	6p21	Antigen presentation	Multi-SNP alleles
HLA-DPB1	6p21	Antigen presentation	Multi-SNP alleles
GPRA	7p14.3	Regulation of cell growth and neural mechanisms	Haplotypes
NAT2	8p22	Detoxification of drugs and carcinogens	Slow acetylation SNPs
FCERIB	11q13	High-affinity Fc receptor for IgE	Ile181Leu, Gly237Glu
CC16	11q12.3-q13.1	Epithelium-derived anti-inflammatory protein	38A/G
GSTP1	11q13	Environmental and oxidative stress — detoxification	Ile105Val
IL18	11q22.2-q22.3	Induction of IFNγ and TNF	–656T/G, –137G/C
STAT6	12q13	IL-4 and IL-13 signalling	2964G/A, (GT)n exon 1
NOS1	12q24.2-q24.31	Nitric oxide synthesis — cell–cell communication	3391C/T, 5266C/T
CMA1	14q11.2	Mast-cell chymotryptic serine protease	*Bst*X1, –1903G/A
IL4R	16p12.1-p12.2	α-chain of the IL-4 and IL-13 receptors	Ile50Val, Glu551Arg
CCL11	17q21.1-q21.2	Epithelium-derived eosinophil chemoattractant	Ala23Thr, –1328G/A
CCL5	17q11.2-q12	Monocyte, T-cell and eosinophil chemoattractant	–403A/G, –28C/G
ACE	17q23.3	Inactivation of inflammatory mediators	In/del
TBXA2R	19p13.3	Smooth-muscle contraction, inflammation	924T/C, 795T/C
TGFB1	19q13.1	Immunoregulation, cell proliferation	–509C/T
ADAM33	20p13	Cell–cell and cell–matrix interactions	Multiple SNPs
GSTT1	22q11.23	Environmental and oxidative stress — detoxification	A/null

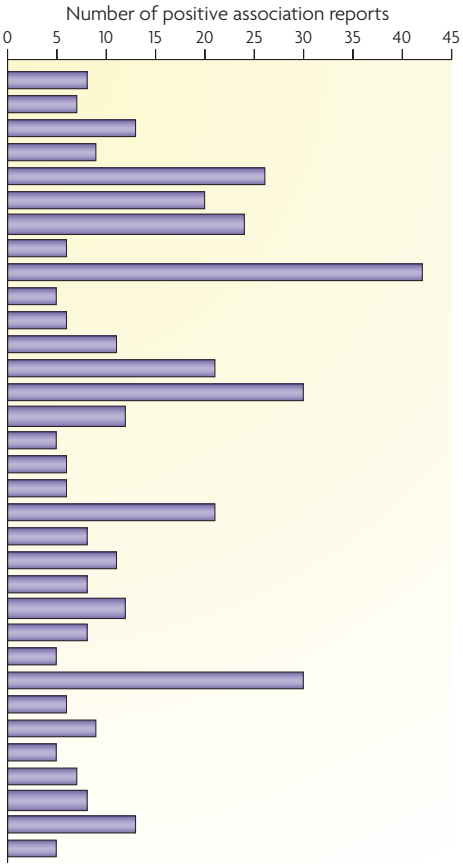

Figure 1 | **Susceptibility genes for asthma and asthma-related traits.** Summary of the genes that were found to be associated with asthma and/or asthma-related phenotypes in at least five independent reports of candidate-gene association or positional-cloning studies (these are underlined). Genes were identified by searching the public databases using the keyword 'association' together with each of the following terms: asthma, bronchial hyper-responsiveness, atopic dermatitis or IgE. Genes are ordered by their chromosomal location. Selected gene variants commonly associated with relevant phenotypes are shown. The information is updated to December 2007. References for studies published before 2006 are listed in REF. 4. The references for the studies published in 2006 and 2007 are provided as Supplementary information S1 (table). *ACE*, angiotensin I converting enzyme 1 (also known as peptidyl-dipeptidase A); *ADAM33*, a disintegrin and metalloproteinase domain 33; *ADRB2*, β_2 adrenergic receptor; *CC16*, Clara cell-specific 16 kD protein (also known as SCGB1A1); *CCL11*, CC-chemokine ligand 11 (also known as eotaxin-1); *CCL5*, CC-chemokine ligand 5 (also known as RANTES); *CD14*, monocyte differentiation antigen 14; *CMA1*, chymase 1, mast cell; *CTLA4*, cytytoxic T-lymphocyte antigen 4; *FCERIB*, high-affinity Fc receptor for IgE β-chain; *FLG*, filaggrin; *GPRA*, G-protein-coupled receptor for asthma susceptibility (also known as NPSR1, and GPRA154); *GSTM1*, glutathione *S*-transferase M1; *GSTP1*, glutathione *S*-transferase P1; *GSTT1*: glutathione *S*-transferase T1; *HAVCR1*, hepatitis A virus cellular receptor 1 (also known as TIM1); IL, interleukin; *IL4R*, interleukin-4 receptor (α-chain); *LTA*, lymphotoxin-α (also known as TNFβ); *LTC4S*, leukotriene C4 synthase; *NAT2*, *N*-acetyltransferase 2; *NOS1*, nitric oxide synthase 1 (neuronal); *SPINK5*, serine protease inhibitor, Kazal-type, 5; *STAT6*, signal transducer and activator of transcription 6; *TBXA2R*, thromboxane A_2 receptor; *TGFB1*, transforming growth factor-β1; *TNF*, tumour necrosis factor.

clinical readership[4,8–11]. Here, I will discuss the most robust asthma and allergy susceptibility genes within a functional and immunological perspective. As replication of results across studies remains the gold standard to assess the robustness of genetic associations, I will focus on the most robust, that is, the most replicated, asthma gene candidates. To provide a view of current asthma genetics that is both comprehensive and concise, I follow the recent suggestion to accept the gene as the unit of replication[12]. That is, I consider an association of any asthma and asthma-related phenotypes with any variant or combination of variants within the same gene to represent a replication, even when the association is not with the variant or phenotype reported in the initial study. Relying on this approach, FIG. 1 shows the genes that were found to be associated with asthma or asthma-related traits in at least five independent studies. Some limitations of this approach, and other definitions of replication, are discussed in BOX 2. SNP nomenclature, which is occasionally still confusing, is outlined in BOX 3.

Asthma susceptibility genes fall into four main groups: genes associated with innate immunity and immunoregulation; genes associated with T helper 2 (T_H2)-cell differentiation and effector functions; genes associated with epithelial biology and mucosal immunity; and genes associated with lung function, airway

T_H2 cells
CD4$^+$ T helper (T_H) cells differentiated along a pathway that leads to coordinated expression of IL-4, IL-13 and IL-5. T_H2 cells and their cytokine products are central mediators of allergic inflammation.

Box 2 | **Replication, strict and loose**

Independent replication of genotype–phenotype associations in distinct populations is generally thought to provide the most convincing evidence for the identification of a true disease susceptibility gene[99]. However, replication can be defined in different ways. Historically, the allele, genotyped at a single variant, was the basic unit of genetic association analysis. However, with the increasing availability of single nucleotide polymorphisms (SNPs) and the use of linkage-disequilibrium-based approaches to association analysis, the focus of association studies has shifted more towards haplotypes[10]. As a consequence, it was argued that the gene itself, rather than specific variants or haplotypes, should become the unit of replication[4,12]. Although genes remain consistent in terms of genomic location, sequence and function across human populations, this is not necessarily the case for specific SNPs or haplotypes.

There are now many examples of established associations with different functional variants within the same gene or with opposite alleles at the same SNP in different populations[4]. For example, IgE levels are associated with *IL13*-1112CT (*rs1800925*) in some populations[49,148], and with *IL13*+2044GA (*rs20541*)[46,47,49] or *IL13*-1512AC (*rs1881457*)[103] in others. Protection from allergy and/or asthma is associated with both the T^{13} and C^{112} allele of *CD14*-159CT (*rs2569190*). A disintegrin and metalloproteinase 33 (*ADAM33*) is also a good example of a gene that shows complex patterns of association, in that no single SNP or haplotype was associated across all populations tested, and the risk SNPs or haplotypes were different in different populations. Yet, convincing evidence supports the identification of all these genes as true susceptibility loci for asthma-related traits. The nature of these associations suggests that susceptibility is complex, raising the possibility that there may not be one susceptibility variant or haplotype in these genes in all populations[4].

Of note, using the gene as the unit of replication is considered by some to be too permissive, and the distinction between strict and loose replication has been emphasized[149]. In terms of phenotype, strict replication implies that the initial association between genotype and phenotype is replicated with exactly the same phenotype. By contrast, loose replication includes all other positive replications with other (related) phenotypes. In terms of genotype, strict replication implies that the same SNP(s) and same risk allele or haplotype are associated with the disease, whereas loose replication includes other SNPs in the gene or a different allele of the same SNP. The strongest evidence for the identification of a true disease susceptibility gene is provided by strict replication at the level of both genotype and phenotype, as is the case for the association between *IL13*+2044GA (*rs20541*) and total IgE levels[46,47,49,57,150]. Although also informative, loose replication raises the possibility that one gene may have multiple functional variants — for instance, *IL13*-1112CT (*rs1800925*) and *IL13*+2044GA (*rs20541*) both of which are functional[45,58] — and/or may exert pleiotropic effects, as suggested by the association of *IL13*+2044GA (*rs20541*) with IgE levels[46], asthma[50], atopic dermatitis[52] and lung function[151].

remodelling and disease severity. FIG. 2 places the genes from the first three groups — plus others identified by positional cloning — and their proposed interactions in a dynamic, cell-based framework.

Innate immunity and immunoregulation. The first group of genes is located at the interface between innate immune sensing and immunoregulation, and is involved in triggering the immune response. Allergic inflammation and IgE regulation are strongly influenced by polymorphisms in the genes encoding pattern recognition receptors and extracellular receptors such as CD14 (REF. 13), Toll-like receptor 2 (TLR2) (REF. 14), TLR4 (REFS 15,16), TLR6 (REF. 17) and TLR10 (REF. 18); and intracellular receptors such as nucleotide-binding oligomerization domain containing 1 (NOD1; also known as CARD4)[19], and NOD2 (also known as CARD15)[20]. Interestingly, variants in pattern recognition receptors are associated with complex diseases other than asthma and allergy, particularly inflammatory bowel disease[21,22] and cardiovascular diseases[23], highlighting the fundamental role of innate immunity in regulating the interface between the organism and the microbial environment. Allergy susceptibility is also influenced by variants in the immunoregulatory cytokines *IL10* (REF. 24) and transforming growth factor-β1 (*TGFB1*)[25] and the transcription factor signal transducer and activator of transcription 3 (*STAT3*)[26], which mediates the inhibitory effects of IL-6 on dendritic cells (DCs). That distinct innate immunity genes are associated with asthma-related traits in different populations suggests that distinct microbial products might have an impact on the innate immune interface in different environments — a recurrent theme in asthma genetics, as discussed below. The pre-eminent role of HLA class II molecules in antigen presentation is reiterated by many studies that show strong associations between specific HLA-DR, -DQ and -DP alleles, and allergen-specific IgE responses[27]. Interestingly, variants of the prostaglandin E_2 receptor (*PTGER2*) were found to be associated with increased susceptibility to asthma[28]. Consistent with this finding, pollen-associated phytoprostanes (bioactive molecules that resemble endogenous prostaglandin E_2 both structurally and functionally) seem to act directly on DCs, decreasing IL-12 expression and favouring T_H2-cell polarization[29].

T_H2-cell differentiation and effector function. The second group of asthma susceptibility genes includes genes that regulate the differentiation of naive $CD4^+$ T_H cells into a T_H2-cell polarized effector phenotype, a key process in allergic inflammation and asthma. Full T_H2-cell polarization requires GATA-binding protein 3 (GATA3) induction by signals delivered through IL-4–IL-4Rα interactions and mediated by STAT6. The activity of GATA3, the T_H2-cell master switch, appears to be inhibited by the transcription factor T-bet, the T_H1-cell master switch, through protein–protein interactions that interfere with GATA3 binding to its target DNA[30]. Common variants of *GATA3* (REF. 31); *TBX21* (which encodes T-bet)[32]; *IL4* (REFS 33,34); *IL4RA*[35]; *STAT6* (REF. 36); and *IL12B* (which encodes IL-12 p40)[37], were recently found to be associated with asthma and allergy, as well as with the response to corticosteroid treatment.

IL-13 is the central effector cytokine of allergic inflammation. Indeed, IL-13 is sufficient to mediate all of the cardinal features of T_H2-cell-mediated responses in the lungs of animal models (airway hyperresponsiveness, inflammatory-cell infiltration, mucus hypersecretion and airway fibrosis)[38,39]. Moreover IL-13, with IL-4, is the only cytokine that induces human IgE synthesis. IL-13 and its receptors are highly expressed in the respiratory tract of patients with asthma and rhinitis[40,41], and IL-13 is expressed in the placenta[42] and is actively secreted by T cells in the neonatal period[43], a time critical for susceptibility to allergic disease. The central role of T_H2-type cytokines in human asthma has been recently emphasized by pharmacological studies in individuals with asthma, which have shown that local administration of an IL-4 variant that inhibits the binding of IL-4 and IL-13 to IL-4Rα substantially diminished asthma symptoms in response to an inhaled allergen challenge[44].

Positional cloning
The process of systematically identifying mutations or susceptibility alleles by studying genetic markers in families or high-risk individuals.

Pattern recognition receptors
Proteins expressed by innate immune cells that detect molecules associated with microbial pathogens or cellular stress.

Box 3 | **The nomenclature of single nucleotide polymorphisms**

The nomenclature of genetic variants is still not standardized completely and is therefore occasionally confusing. For example, single nucleotide polymorphisms (SNPs), the variants most frequently assayed in genetic studies, are ideally defined by their position within the gene and the two alleles found at that position. However, the same SNP often receives different identifiers. For SNPs in regulatory regions, some groups use the ATG as a reference for numbering, whereas other groups count from the transcription start site. Thus, the same replacement of a C with a T in the CD14 promoter may be identified as *CD14*-260CT[118,119] or *CD14*-159CT[13]. For SNPs in coding regions, which are typically defined by the amino-acid change, numbering may or may not include the signal peptide. Thus, the same non-synonymous SNP in *IL13* may be defined as *IL13*Arg130Gln[45] or *IL13*Arg110Gln[152]. A solution to this problem may be found through concerted efforts such as the one led by the SNP database at the National Center for Biotechnology Information, which acts as a public-domain archive for a collection of genetic polymorphisms in various organisms. dbSNP maps each submitted SNP assay to the genome and assigns to each submitted SNP assay a RefSNP accession ID (rs number) that corresponds to the position in the idealized genome where the variation can be assayed. For example, the dbSNP notations of the CD14 and IL13 polymorphisms mentioned above are *rs2569190* and *rs20541*, respectively. The strength of this system is that submitted SNPs that map to the same location are clustered into the same RefSNP and have the same rs number. The SNPs discussed in this article are identified both by their traditional notation, which has the advantage of pointing readily to the gene, and by their rs number, which is less easy to visualize but unambiguous.

Not surprisingly, *IL13* is also one of the strongest and most studied of the candidate genes for asthma and allergy. *IL13*+2044GA (*IL13* Arg130Gln, *rs20541*) is a common coding SNP that results in the non-conservative replacement of Arg130 with Gln and the expression of an IL-13 variant that has increased biological activity[45]. This polymorphism is associated with increased total serum IgE[46–49], asthma[50], atopy[51] and atopic dermatitis[47,52,53]. *IL13*-1112CT (*rs1800925*), which is a promoter SNP associated with asthma, atopic dermatitis and increased risk of allergic sensitization[54–57], leads to increased *IL13* transcription in polarized T_H2 cells and enhanced IL-13 secretion by mitogen-stimulated mononuclear cells[58]. It is important to note that *IL13*+2044GA and *IL13*-1112CT are in high linkage disequilibrium. Therefore, individuals carrying both risk alleles are expected to express higher levels of an overactive IL-13 variant (FIG. 3).

Polymorphisms in the T_H2 signalling pathway may influence not only T-cell-dependent induction of IgE synthesis but also IgE amplification dependent on T_H2 cytokines expressed by basophils and mast cells. Crosslinking of the high-affinity IgE receptor FcεR1 is the main stimulus for T_H2 cytokine secretion by mast cells and basophils. Because of this, SNPs in membrane-spanning 4-domains, subfamily A, member 2 (*MS4A2*; also known as *FCER1B*) — which encodes the β-chain of FcεR1 and is one of the most replicated asthma susceptibility genes[59] — may affect receptor expression[60] and IgE-dependent release of inflammatory mediators by mast cells. SNPs in *MS4A2* may also affect T_H2 cytokine expression in non-T cells and IgE-dependent antigen presentation by DCs. The effector phase of allergic inflammation also relies on the interaction between T_H2-cell-derived IL-5 and IL-5RA on eosinophil progenitors. IL-5 drives the development of eosinophils, which can then be recruited to the airways by epithelial-cell-derived chemokines and in turn become an important source of T_H2 cytokines. SNPs in *IL5* and *IL5RA*[61,62] are associated with asthma-related traits, but not as strongly as polymorphisms in other genes of the T_H2 pathway.

Epithelial cells. The third group of asthma and allergy susceptibility genes is expressed in epithelial cells, at the interface between innate and adaptive immunity. The epithelium secretes chemokines such as CC-chemokine 5 (CCL5; also known as RANTES), which recruits T cells and eosinophils, and CCL11, CCL24 and CCL26 (known as eotaxins 1–3), which are powerful eosinophil attractants. The epithelium also secretes antimicrobial peptides (such as defensin-β1) and uteroglobin/Clara cell 16-kD protein (CC16), which inhibits T_H2-cell differentiation by acting on DCs[63]. All of these genes are robust asthma and allergy candidates[64–67]. The expression of SPINK5, which serves an important protective role against proteases released by mast cells or allergens, is restricted to the epithelium. Mutations in *SPINK5* lead to Netherton's syndrome, a rare autosomal recessive disorder that is characterized by defective epithelial function with features of eczema[68]. Also, a non-synonymous SNP (*SPINK5* Glu420Lys, *rs2303067*) has been reported to be strongly associated with asthma and eczema[69].

Recently, null mutations in another gene crucial for the integrity of the epithelial barrier, filaggrin (*FLG*), a member of the epidermal differentiation complex on chromosome 1q21, were reported to be strongly associated with atopic dermatitis and eczema[70,71], and to influence asthma, dependent[72] or independent[73] of atopic dermatitis. *FLG* deserves particular attention for several reasons. Filaggrin is expressed in the epidermis and in the oral and nasal mucosa, but not in the bronchial mucosa[74]. A direct association between asthma and *FLG* variants is rarely found in the absence of concomitant atopic dermatitis. These combined findings suggest that asthma in individuals with atopic dermatitis is secondary to allergic sensitization that occurs after the breakdown of the epidermal skin barrier[75]. Interestingly, according to a recent meta-analysis, the effect of *FLG* variants on the risk of atopic dermatitis exceeds that of any other candidate gene investigated so far[76]. Yet, the *FLG* mutations that predispose to atopic dermatitis are extremely rare. Indeed, their association with the disease was tested only because these mutations are known to cause ichthyosis vulgaris[77], a common recessive Mendelian disorder of skin keratinization, and atopic dermatitis was highly prevalent among patients with ichthyosis vulgaris who were null or heterozygous for *FLG*[70]. All of these findings underscore both the important role of an intact epithelial barrier in protecting against environmental agents[8] and the complexities of disease gene discovery.

Lung function. The final group of candidate genes is more heterogeneous and is comprised of genes associated with lung function, airway remodelling and disease severity. This group includes *ADRB2* (REF. 78)

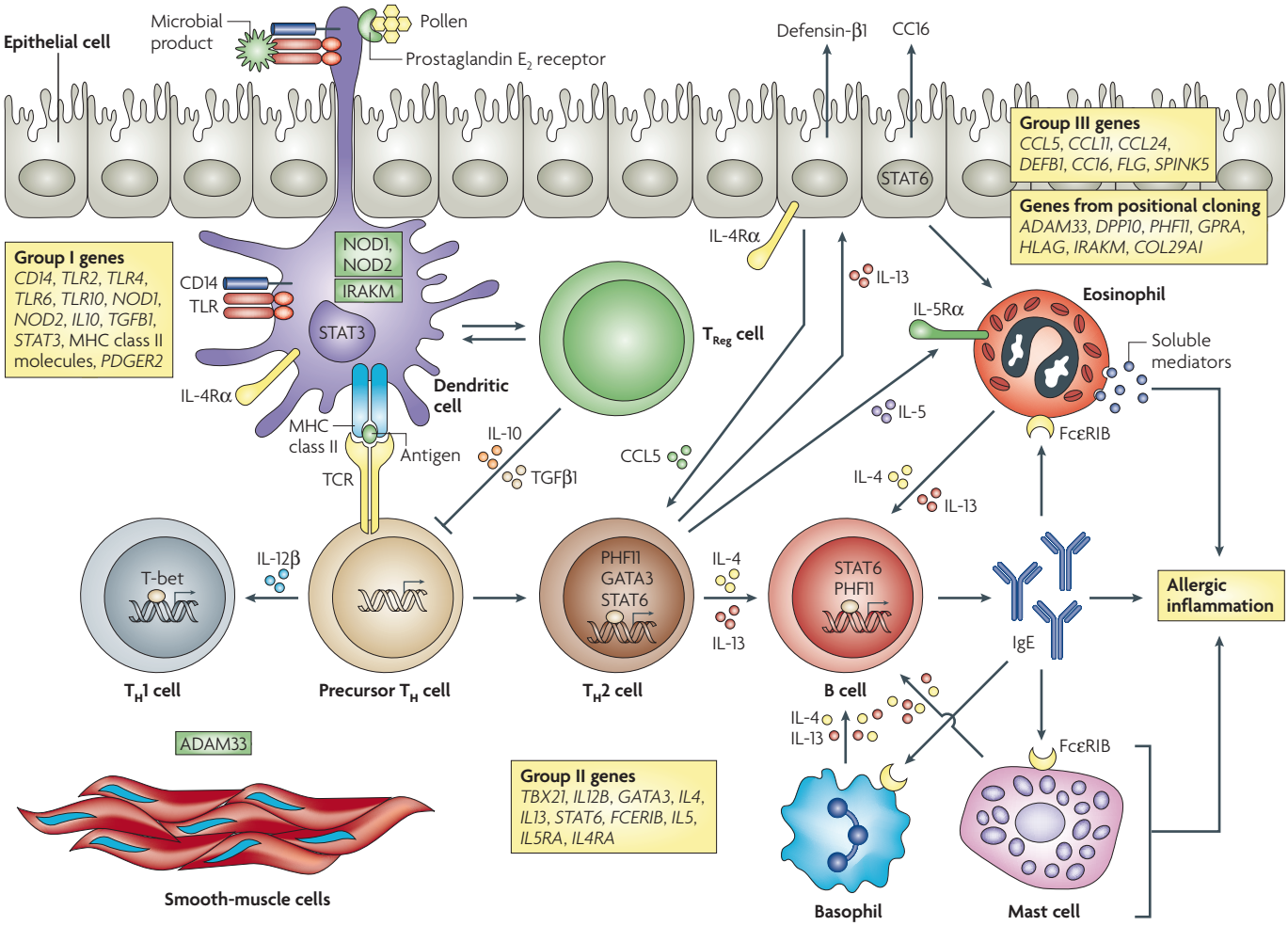

Figure 2 | **Susceptibility genes for asthma and asthma-related traits.** Representation of the most robust asthma candidate genes identified through association studies or positional cloning in a cell-based framework. Among the genes discovered through association studies, the first group is involved in triggering the immune response and directing $CD4^+$ T helper (T_H)-cell differentiation. This group includes genes encoding pattern recognition receptors (*CD14*, *TLR2*, *TLR4*, *TLR6*, *TLR10*, *NOD1* and *NOD2*), immunoregulatory cytokines (*IL10* and *TGFB1*), the transcription factor *STAT3*, molecules involved in antigen presentation (*HLA-DR*, *HLA-DQ* and *HLA-DP* alleles) and the prostaglandin receptor *PTGER2*. The second group of asthma susceptibility genes includes those that regulate T_H2-cell differentiation and T_H2-cell effector functions (*GATA3*, *TBX21*, *IL4*, *IL13*, *IL4RA*, *FCER1B*, *IL5*, *IL5RA*, *STAT6* and *IL12B*). The third group of genes is expressed in epithelial cells, and includes chemokines (*CCL5*, *CCL11*, *CCL24* and *CCL26*), antimicrobial peptides (*DEFB1*), CC16 and factors involved in maintaining the integrity of the epithelial-cell barrier (*SPINK5* and *FLG*). The group of asthma susceptibility genes discovered through positional-cloning approaches includes *ADAM33*, *DPP10*, *PHF11*, *GPRA*, *HLA-G*, *IRAKM* and *COL29A1* (which encodes collagen XXIX). These genes are expressed in the epithelium and/or smooth muscle. ADAM33, a disintegrin and metalloproteinase domain 33; CC16, Clara cell-specific 16 kD protein (also known as SCGB1A1); CCL, CC-chemokine ligand; CD14, monocyte differentiation antigen 14; COL29A1, collagen XXIX; DEFB1, defensin β1; DPP10, dipeptidyl peptidase 10; FCERIB, high-affinity Fc receptor for IgE β-chain; FLG, filaggrin; *GATA3*, GATA-binding protein 3; GPRA, G-protein-coupled receptor for asthma susceptibility (also known as NPSR1 and GPRA154); IL, interleukin; IL4RA, interleukin-4 receptor (α-chain); *IL5RA*, interleukin 5 receptor (α-chain); IRAKM, interleukin-1 receptor-associated kinase 1; NOD, nucleotide-binding, oligomerization-domain-containing; PDGER2, prostaglandin E_2 receptor; PHF11, plant homeodomain finger protein 11; SPINK5, serine protease inhibitor, Kazal-type, 5; STAT, signal transducer and activator of transcription; *TBX21*, T-box 21 (also known as T-bet); TCR, T-cell receptor; TGFβ1, transforming growth factor-β1; T_H, T helper cell; TLR, Toll-like receptor; T_{Reg}, regulatory T cell.

and tumour necrosis factor (*TNF*)[79], which are among the most robust asthma gene candidates. Other genes include *TGFB1* (REF. 25); leukotriene C_4 synthase (*LTC4S*)[80]; glutathione-*S*-transferases (*GSTP1* and *GSTM1*)[81,82]; thromboxane A2 receptor (*TBXA2R*)[83]; arachidonate 5-lipoxygenase (*ALOX5*)[84]; lymphotoxin-α (*LTA*; also known as *TNFB*)[79]; tenascin-C (*TNC*)[85]; and nitric oxide synthase 1 (*NOS1*)[86]. Several of these genes also seem to influence treatment requirements and are systematically investigated in pharmacogenomics studies aimed at identifying variants that determine patterns of pharmacological response[87].

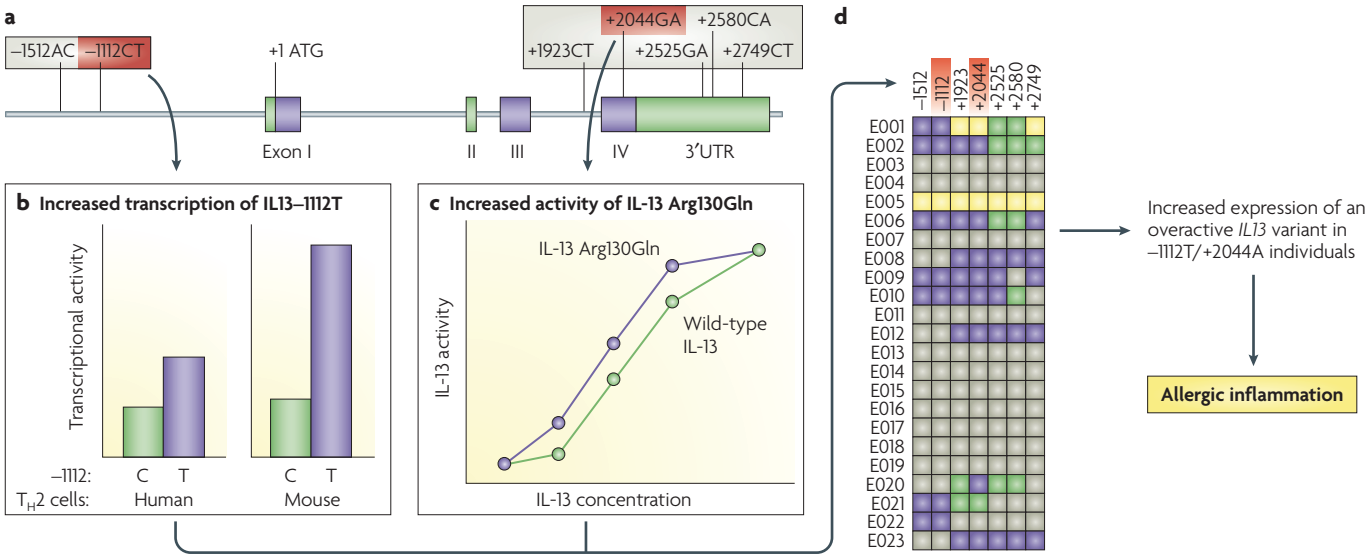

Figure 3 | **Distinct variants of *IL13* may synergize and increase susceptibility to allergic inflammation. a** | Analysis of genetic variation across the interleukin-13 (*IL13*) locus in European ancestry subjects[46] identified a block of common single nucleotide polymorphisms (SNPs) at the 5′ end of the gene, that includes two promoter polymorphisms, *IL13*-1512AC (*rs1881457*) and *IL13*-1112CT (*rs1800925*). Another SNP block at the 3′end of the gene extends from *IL13*+1923CT (*rs1295686*) in the third intron to *IL13*+2749CT (*rs847*) in the 3′ untranslated region. SNP nomenclature is discussed in BOX 3. **b** | *IL13*-1112CT (*rs1800925*) results in increased *IL13* transcription in human and mouse T_H2 cells[58]. **c** | *IL13*+2044GA (*rs20541*) in the 3′ block is a non-synonymous SNP that leads to the expression of a gain-of-function variant, IL-13 Arg130Gln[45]. **d** | The 5′ and 3′ blocks are frequently, albeit not invariably, found in the same individuals, as shown by the Visual Genotype for *IL13* provided by Seattle SNPs. In a Visual Genotype, polymorphic sites are consecutively ordered (5′ to 3′) on the horizontal axis and individual reference DNA samples (here, from 23 individuals of European descent) are listed on the vertical axis of the array. The genotype for each sample at each polymorphic site is coded as: grey = homozygous (common allele), yellow = homozygous (rare allele), purple = heterozygous (both alleles), green = genotype not determined. Co-occurrence of the rare alleles at positions *IL13*-1112T and *IL13*+2044A would be expected to result in increased expression of an overactive IL-13 variant, which may contribute to enhanced susceptibility to allergic inflammation. Figure 3b is modified with permission from REF. 58 © (2006) The American Association of Immunologists, Inc. Figure 3c is modified with permission from REF. 45 © (2005) The American Society for Clinical Investigation.

Asthma genes identified by positional cloning. The asthma susceptibility genes identified through linkage studies followed by positional cloning deserve a separate discussion. The first positionally cloned asthma gene, a disintegrin and metalloproteinase 33 (*ADAM33*)[88], ranks among the genes most replicated so far. The linkage analysis that led to the identification of *ADAM33* as an asthma susceptibility gene was conducted on 460 affected Caucasian families from the UK and USA. Significant linkage to asthma and bronchial hyper-responsiveness was identified in close proximity to the tip of the p-arm of chromosome 20. Physical mapping and direct cDNA selection identified 40 genes under the linkage peak, and further high-resolution SNP analyses identified *ADAM33* as the source of the linkage signal. The association of *ADAM33* with asthma and lung function has been confirmed in multiple populations of distinct ethnic backgrounds[89]. The identification of *ADAM33* represented a breakthrough because this gene pointed to potentially new pathogenetic pathways for asthma. Indeed, unlike most of the previously known candidates, *ADAM33* is expressed by lung fibroblasts and bronchial smooth-muscle cells, but not by bronchial epithelial cells or immune cells. *ADAM33* is also preferentially expressed during branching morphogenesis in mouse and human lungs, suggesting a function linked to the role of the epithelial–mesenchymal trophic unit in lung development[89].

The year 2003 witnessed the positional cloning of two other genes that modify the risk of developing asthma: plant homeodomain finger protein 11 (*PHF11*) and dipeptidyl-peptidase 10 (*DPP10*). *PHF11* is located within a region on chromosome 13q14 that had shown consistent linkage to atopy and total serum IgE concentration. The quantitative-trait locus associated with IgE levels was identified in a comprehensive SNP map that led to several alleles in a single gene, *PHF11*. These variants were also associated with severe clinical asthma[90]. *PHF11* is a member of the plant homeodomain (PHD) family, a collection of proteins named for the presence of one or more PHD-type zinc fingers. Proteins with a PHD finger are commonly found in the nucleus and are implicated in transcriptional regulation through protein–protein interactions. Consistent with a role in transcription, PHD domains can interact with the trimethylated lysine residue 4 on histone 3, a modification that marks transcriptionally active regions of the human genome[91]. Expression of PHF11 in both

Quantitative-trait locus
A polymorphic locus that contains alleles that differentially affect the expression of a continuously distributed phenotypic trait (for example, total serum IgE levels).

T and B cells[90] raises the possibility that this protein may regulate the transcription of lymphocyte genes involved in allergic inflammation.

The search that led to the discovery of *DPP10* was stimulated by reports linking asthma to an ill-defined interval in chromosome 2 (2q14–32), and by two mouse genome screens that had linked bronchial hyper-responsiveness to the region homologous to 2q14 (REF. 92). Association studies with a relevant microsatellite, followed by sequencing of the surrounding region and generation of a high-density SNP and linkage disequilibrium map, showed that the association was limited to the initial exons of a solitary 3.6 kb gene, *DPP10*, which encodes a homologue of dipeptidyl peptidases involved in cleavage of terminal dipeptides from cytokines and chemokines[92].

G-protein-coupled receptor for asthma susceptibility (*GPRA*; also known as *NPSR1* and *GPRA154),* an asthma susceptibility gene on chromosome 7p, was positionally cloned in 2004 using a hierarchical genotyping design that led to the identification of a 133 kb risk-conferring segment that contained two genes. One of these encoded *GPRA*, an orphan G-protein-coupled receptor that showed a distinct distribution of protein isoforms in bronchial biopsies from healthy and asthmatic individuals. In three cohorts from Finland and Canada, SNP-tagged haplotypes were associated with high serum IgE levels or asthma. Moreover, the mouse orthologue of *GPRA* was upregulated in a model of ovalbumin-induced inflammation, thus supporting a role for this gene in the pathogenesis of atopy and asthma[93].

Two more asthma susceptibility genes were identified in 2005 through linkage analysis. *HLA-G*, an HLA class Ib gene on chromosome 6p21, was identified in white families participating in a study in Chicago, USA[94]. This gene is primarily expressed by fetal cells at the maternal–fetal interface, where it is involved in immunosuppression and is thought to contribute to the maternal tolerance of the genetically foreign fetus. Expression of soluble HLA-G was detected in bronchial epithelial cells in two individuals with asthma but not in one unaffected individual, suggesting that expression of this gene may be upregulated in asthmatic lungs[95]. Interestingly, the risk genotype at *HLA-G* varied depending on whether the mother had asthma or bronchial hyper-responsiveness or was unaffected. More recently, a SNP in the 3′ untranslated region of *HLA-G* was shown to influence the targeting of three microRNAs to this gene, raising the interesting possibility that microRNA-mediated mechanisms may contribute to the impact of *HLA-G* on asthma risk[95].

The other asthma gene published in 2005 was cytoplasmic fragile X mental retardation protein (FMR) interacting protein 2 (*CYFIP2*), which was identified in a region of chromosome 5q33 linked to mite-sensitive asthma. Mutation screening and association analyses in 155 families with asthma revealed that six polymorphisms in *CYFIP2* were significantly associated with the development of asthma. *CYFIP2* expression was significantly increased in lymphocytes from subjects homozygous for the disease-associated haplotype[96].

The latest additions to the group of positionally cloned genes for asthma and asthma-related traits are IL-1 receptor-associated kinase-M (*IRAKM*; also known as *IRAK3*), and collagen XXIX (*COL29A1*), both reported in 2007. *IRAKM*, a negative regulator of the TLR and IL-1R pathway located in the candidate region 12q13–24, was found to be associated with early onset persistent asthma both in a Sardinian founder population and in an outbred Italian population. Sequence analysis in patients found inactivating mutations in the *IRAKM* coding region. Immunohistochemistry of lung biopsies showed that IRAKM is highly expressed in epithelial cells, which suggests a mechanistic link between hyperactivation of the innate immune system and chronic airway inflammation[97].

COL29A1 was identified in a susceptibility locus for atopic dermatitis on chromosome 3q21 through dense mapping of microsatellites and SNPs[98]. The region strongly associated with atopic dermatitis contained a single gene, *COL29A1*, which encodes a recently discovered epidermal collagen highly expressed in the skin, lungs and gastrointestinal tract, sites that are highly relevant to allergic inflammation. Lack of *COL29A1* expression in the outer epidermis of patients with atopic dermatitis suggests that this protein contributes to epidermal integrity and function.

Overall, genome-wide linkage analyses have kept their promise. Indeed, their results have renewed the interest of asthma researchers in the epithelium and the smooth muscle, which are at the core of the organ-specific component of asthma pathogenesis, but had been neglected by association studies, most of which had focused on innate and adaptive immune genes.

Discrepancies, controversies, complexities

Despite its past, current and expected achievements, asthma genetics is still burdened with puzzling inconsistencies in the genes that have been associated with disease phenotypes, the alleles specifically involved and the direction of their effects. A recent Editorial in *Nature Genetics* expressed alarm for the more and more frequent inability of association studies to be replicated[99]. Likewise, an Editorial published in *Thorax* lamented that genetic association studies of complex diseases have "acquired a bad reputation…because of problems with poor design and variable replication of findings"[100]. Problems of consistency and replication are not unique to asthma, but extend to other complex diseases[101]. Moreover, even the most replicated genes, such as *IL13*, *IL4*, *IL4RA*, *CD14*, *ADRB2*, *FCER1B*, *TNF* and *ADAM33*, failed to show an association with asthma or allergy in a substantial number of studies[4,102], and when successfully replicated, they typically have small effects and explain a negligible proportion of the phenotypic variance[103]. Why is there so much controversy and complexity? As extensively discussed elsewhere[9,11,101], inconsistent associations may reflect flaws in study design, size or analysis. However, a role for factors and processes rooted in the biology of complex diseases in general, and asthma and allergy in particular, should also be considered.

Microsatellites
Polymorphic DNA loci that consist of repeating units of 1–4 bp in length.

Linkage disequilibrium map
Map of non-random associations between alleles at two or more loci.

Haplotype
A combination of alleles at different markers located on the same chromosome in a specific genomic region.

MicroRNAs
Single-stranded RNA molecules of approximately 21–23 nucleotides in length that are thought to regulate the expression of other genes.

Phenotypic heterogeneity. Similar to other complex diseases, asthma is phenotypically heterogeneous. It can have its onset at any age; it is triggered by numerous stimuli, both specific and nonspecific; it can present as a mild, transient condition or as a severe, persistent, life-threatening illness; and it is associated with several intermediate phenotypes, including atopy, IgE levels, bronchial hyperresponsiveness, eosinophilia, atopic dermatitis and others, which can be assorted in any combination[11]. By hindering phenotypic characterization and standardization, this heterogeneity can greatly contribute to inconsistencies in replication, particularly if distinct phenotypes to some extent reflect the involvement of distinct genetic pathways[11,101].

Gene–environment interactions in the population. Gene–environment interactions critically modify the impact of a given gene on complex phenotypes. This notion is consistent with the results of multiple epidemiological studies demonstrating how strongly allergic inflammation phenotypes are influenced by the environment (reviewed in REF. 104). Moreover, rigorous studies in mice conclusively showed that gene–environment interactions explain a proportion of the phenotypic variance that may be higher than that explained by either genetic or environmental effects considered separately[105]. The problem is, the influence of variation on gene expression patterns is difficult to decipher unless the appropriate exposures are taken into consideration, but what these exposures are is not always clear.

Perhaps not surprisingly, interactions involving genes associated with innate immunity and exposure to pathogen products offer eloquent examples of the inconsistencies that plague complex disease genetics[106]. CD14 is a component of the TLR signalling complex that facilitates endotoxin responsiveness through the TLR4–MD2 complex, but it can also associate with various other bacterial TLR ligands[107] and even with TLR3, the receptor for double-stranded RNA[108]. Many association studies have explored the relationship between *CD14*-159CT (also known as *CD14*-260CT, *rs2569190*), which is a functional promoter polymorphism[109], and allergic inflammation. The effects revealed by these studies are diverse, ranging from protective effects[13,110,111] to disease-promoting effects[112,113], with some reports failing to find an association[114–116]. Moreover, the *CD14*-159 alleles have been found to have opposite effects even within the same population (reviewed in REF. 117). Gene–environment interactions provide a framework to reconcile these discrepancies (FIG. 4). The influence that the *CD14*-159CT polymorphism has on IgE levels and asthma seems to depend on the relevant environmental exposures: cat and dog versus stable animals in a European population[118], and low versus high endotoxin in subjects of African descent[119]. In another study, the *CD14*-159TT genotype protected against atopic dermatitis, but only in children with a dog at home[120]. Gene–environment interactions have been implicated even more explicitly by evidence of the *CD14* genotype-dependent effect of domestic endotoxin exposure on allergic sensitization, eczema and wheezing. Children who were homozygous for the *CD14*-159C allele showed a dose-dependent response to endotoxin — high exposures were associated with a decreased risk of allergic sensitization and eczema, and low exposures with an increased risk, relative to the rest of the population. However, in the same children, high endotoxin exposure was associated with increased non-atopic wheezing. Among the remaining children, there was no effect of endotoxin exposure on sensitization, wheeze or eczema[121].

The conclusion reached by all these studies is that the same polymorphism can be associated with either disease or protection depending on the environment a subject is exposed to. If one considers how complex the definition of environmental exposure is (smoking is a factor involved in gene–environment interactions[122], but sex is too[123]), it is not unreasonable to surmise that other, as yet undetected, gene–environment interactions may contribute to the problems of replication in genetic association studies. An intriguing, still unanswered, mechanistic question raised by these data pertains to the nature of the biological switches that turn distinct environmental exposures into distinct immune response patterns[124].

Gene–environment interactions in the nucleus. A special case of gene–environment interactions are those that occur within the nucleus and are best revealed by functional studies of polymorphisms in regulatory regions. For instance, the relative transcriptional activities of the C and T alleles of *CD14*-159 (*rs2569190*) differ in monocytes and hepatocytes, depending on the ratio between SP1 and SP3, which are the transcription factors that

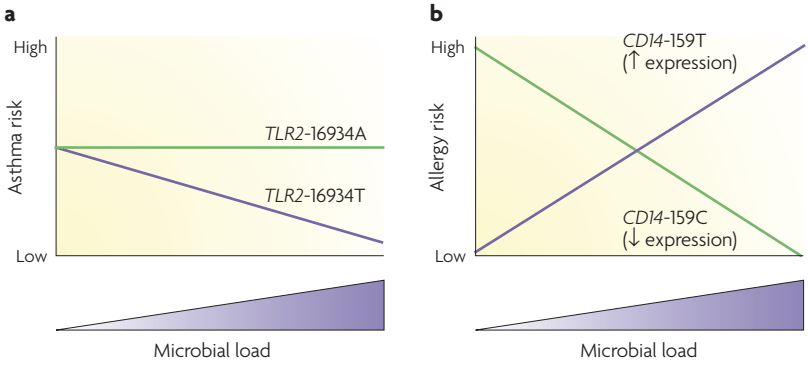

Figure 4 | **Gene–environment interactions in human populations.** The impact of genetic variants on disease susceptibility often depends on the environment subjects are exposed to. **a** | Among farmers' children, those carrying a T allele at Toll-like receptor 2 (*TLR2*)-16934 (*rs4696480*) were significantly less likely to have a diagnosis of asthma or asthma-related phenotypes compared with children carrying the *TLR2*-16934A genotype. No such association was found among children from the same rural communities but not living on farms, suggesting that genetic variation in *TLR2* is a major determinant of the susceptibility to asthma and allergies, but only in children of farmers[14]. **b** | In rural populations from the Alpine regions of Europe, the T allele of *CD14*-159CT (*rs2569190*) was neutral when tested for association with serum IgE levels in the totality of the population, was protective in children exposed to cats and dogs, and was associated with high IgE levels in children exposed to stable animals[118]. In these examples, the impact of genetic variants on disease susceptibility appeared to be modified by quantitative factors (levels of exposure), but a qualitative component — that is, the kind of microbes present in the environment — might also have an important role in affecting risk.

bind the polymorphic promoter region[109]. Moreover, the minor T allele of *IL13*-1112CT (*rs1800925*), a common promoter SNP, was more active than the –1112C allele in primary human T_H2 cells, but was less active than the –1112C allele in non-polarized $CD4^+$ T cells. Distinct patterns of transcription factor binding to the *IL13* promoter appeared to determine differential promoter activity[58]. In this case, differential gene expression was detected not in distinct cell types but in distinct phenotypes of the same cell, demonstrating the subtle, yet critical, impact of the nuclear environment on genetic variants.

Gene–gene interactions. Gene–gene interactions, that is, the functional interplay between genetic variants within a pathway, are also likely to contribute to the complexity of genetic diseases. Each variant typically has modest effects in isolation, but synergizes effectively with other variants to magnify the impact on disease risk. Analyses of associations between allergy and variants in the T_H2-cell differentiation and signalling pathway provide good examples of gene–gene interactions. The analysis of genotyping data from a large population of German children showed that when polymorphisms in *IL4*, *IL13*, *IL4RA* and *STAT6* were combined in a stepwise procedure, the risk for high serum IgE levels increased by 10.8-fold and the risk for the development of asthma increased by 16.8-fold compared with the maximum effect of any individual SNP[125] (FIG. 5). In other studies, significant interactions were detected between *IL4RA* Pro478Ser (*rs1805015*) and *IL13*-1112CT (*rs1800925*). Individuals with the risk genotype for both genes were at almost five-times greater risk for the development of asthma compared with individuals with both non-risk genotypes[126]. The effects of the *IL13*-1112TT genotype on risk of food sensitization were modified by *IL4RA* Pro478Ser and *IL4RA* Gln551Arg (*rs1801275*)[57], and significant interactions were found between *IL13* and *IL4RA* for asthma, and *IL13* and *CCL17* (also known as *TARC*) for total plasma IgE[127]. As interaction analyses require large population samples, all of these studies suffer from limited statistical power. However, the biological plausibility of the underlying hypothesis lends credibility to their results.

The complexity of the scenario that is emerging from the studies discussed in this section highlights the need for novel, more incisive approaches that are capable of better characterizing the interface between genes and the environment in both genetic and functional terms.

From single gene to GWA studies

The year 2007 has witnessed a quantum leap in genotype and phenotype association analyses with the publication of the first GWA study for an asthma trait (childhood asthma)[128]. GWA studies rely on dense sets of SNPs across the genome to survey the most common genetic variants for a role in disease or to identify the heritable quantitative traits that are risk factors for disease[129]. The advent of GWA studies has been fostered by three recent advances. First, the International HapMap resource[130], which documents patterns of genome-wide variation and linkage disequilibrium in four population samples,

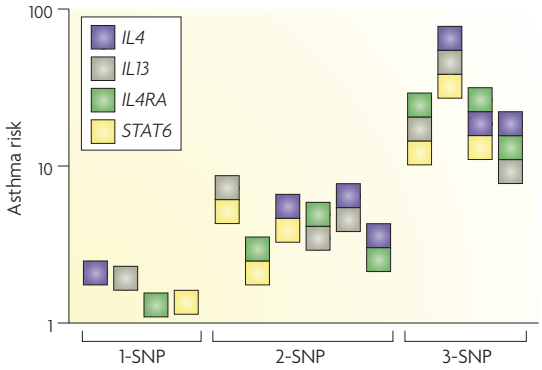

Figure 5 | **Effects of gene–gene interactions on asthma susceptibility.** The risk of developing asthma is synergistically increased by combinations of single nucleotide polymorphisms (SNPs) in individual genes of the T helper 2 (T_H2)-cell-associated signalling pathway. Individual polymorphisms in interleukin-4 (*IL4*) (*rs2243250*), *IL13* (*rs1800925*), *IL4RA* (*rs1805010*) and signal transducer and activator of transcription 6 (*STAT6*) (*rs324011*), and their combinations are shown as colour-coded squares. Asthma risk is indicated on the *y*-axis. In a group of 1,120 German children, 9–11 years old, each polymorphism in isolation had a modest effect on asthma susceptibility (left), but risk was synergistically enhanced by carrying SNPs in two (centre) or three (right) genes within the T_H2-cell-associated pathway. This figure is modified with permission from REF. 125 © (2006) Elsevier Science.

greatly facilitates both the design and the analysis of association studies. Second, the availability of dense genotyping chips, containing sets of hundreds of thousands of SNPs that provide good coverage of much of the human genome, means that for the first time GWA studies for thousands of cases and controls are technically and financially feasible. Third, large and well-characterized clinical samples have been assembled for many common diseases, including asthma and asthma-related traits[131].

The main strength of GWA studies is expected to lie in their ability to discover truly novel disease candidate genes, especially those associated with moderate risks[6]. However, as with any new method, it is important not to overstate what GWA studies can do. First, the study populations must be carefully characterized to allow the selection of patients who are likely to share a genetic cause of disease. Second, thousands of cases and controls may be needed if a study is to have sufficient statistical power to identify the alleles of interest, and some relevant statistical strategies are still being developed. The need for extremely large populations may also lead to heterogeneity in environmental exposures[132], a variable critical for the outcome of any genetic study. Third, the need to test hundreds of thousands of DNA variants in thousands of subjects creates challenges in bioinformatics and raises questions about how to identify true positive signals in a sea of false positives. Moreover, a few technological challenges remain, as do a few gaps in the genetic map. Last, even when research involving GWA identifies a physical region of interest, finding the specific mutation may not be straightforward without in-depth functional studies[133]. More generally, even at this early

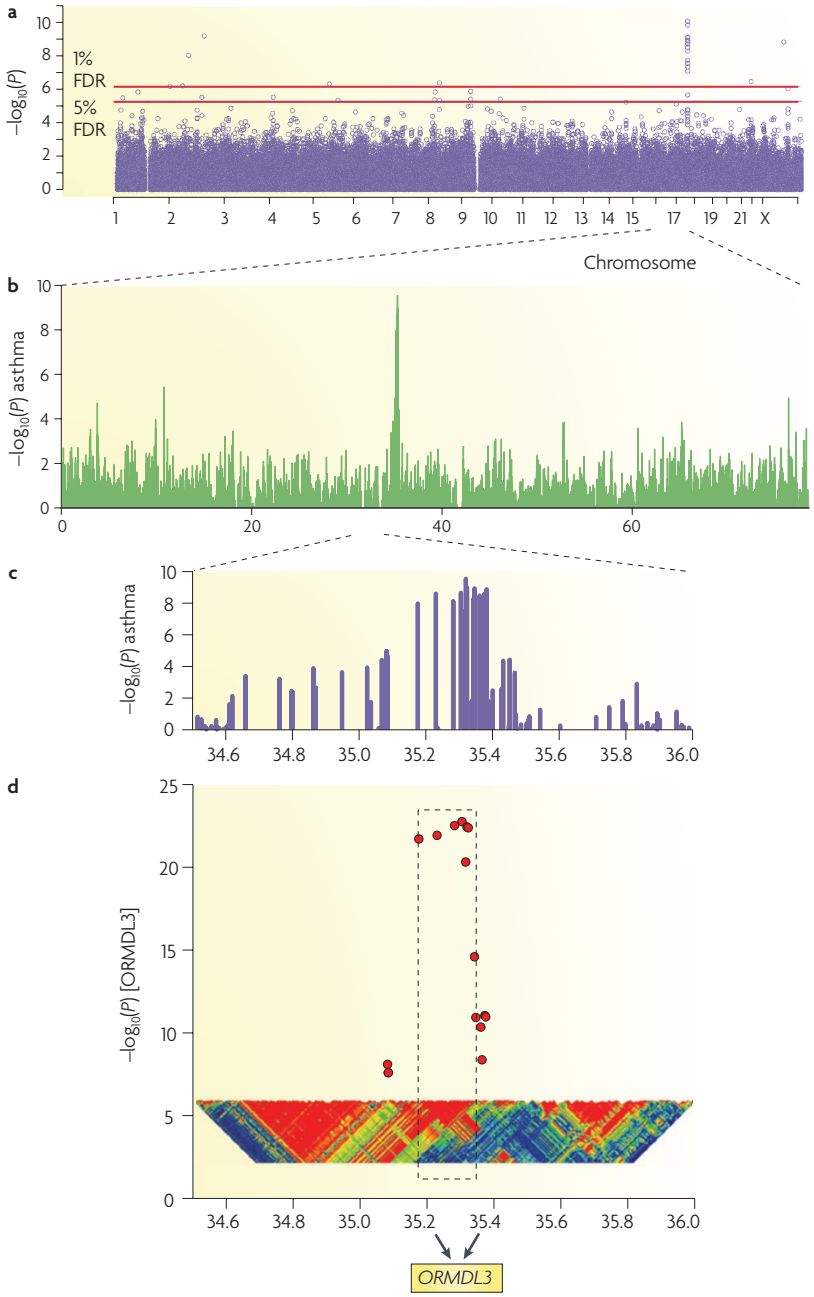

Figure 6 | **Steps in the genome-wide association study that led to the identification of ORMDL3 as an asthma gene. a** | Results of a genome-wide association (GWA) study of 317,447 single nucleotide polymorphisms (SNPs) and asthma in 994 asthmatic children and 1,243 non-asthmatic children. Position in the genome, divided by chromosome, is depicted along the *x*-axis. Strength of association is shown on the *y*-axis. The result for each individual marker is depicted as a circle. The genome-wide thresholds for 1% and 5% false discovery rates (FDR) are shown as horizontal red lines. Numerous markers on chromosome 17q21 showed association to asthma above the 1% FDR threshold in the region of maximum association. **b** | Mapping of association to asthma on chromosome 17. **c** | Detail of association to SNPs on chromosome 17q21. **d** | Association to *ORMDL3* transcript abundance with the same markers. A linkage disequilibrium plot between markers is also shown, with red indicating high linkage disequilibrium and blue denoting low linkage disequilibrium. The central island of linkage disequilibrium, which contains maximum association to *ORMDL3* and asthma, is contained within the grey rectangle. **e** | Identification of *ORMDL3* as the gene in the region of maximum association for which transcript levels were strongly and consistently associated with the same 17q21 SNPs associated with childhood asthma. This figure is modified with permission from *Nature* REF. 128 © (2007) Macmillan Publishers Ltd.

stage it is clear that not all GWA results replicate consistently. In this respect, GWA studies are reminiscent of earlier candidate gene studies, and conclusions should be similarly cautiously made[134,135].

The results of the first, and so far only, GWA study for asthma were published in the summer of 2007 by a network including several European groups. In the initial phase of the study (FIG. 6), more than 317,000 SNPs were characterized in DNA from 994 patients with childhood onset asthma and 1,243 non-asthmatics, using family and case-control panels. Multiple markers on chromosome 17q21.1 were found to be strongly and reproducibly associated with the disease phenotype of interest. The association between the 17q21 locus and diagnosis of childhood asthma was independently replicated in 2,320 subjects from a cohort of German children and in 3,301 subjects from the British 1958 birth cohort[128].

As the region of association on chromosome 17q21.1 spanned 206 kb and included 19 annotated genes, subsequent efforts focused on the identification of the gene(s) responsible for the association, which was pursued by profiling transcription of the genes located within the relevant genomic interval. The rationale underlying this approach was that variation in gene transcription may mediate disease susceptibility, and transcript abundance may be directly modified by polymorphisms in regulatory elements. Analysis of the relationships between markers in the 17q21 locus and transcript levels in Epstein–Barr virus-transformed lymphoblastoid cell lines from children in the asthma family panel used in the association study showed that the SNPs associated with childhood asthma were consistently and strongly associated in *cis* with transcript levels of *ORMDL3*. This result suggests that genetic variants regulating *ORMDL3* expression are determinants of susceptibility to childhood asthma. No other markers were significantly correlated to *ORMDL3* expression.

ORMDL3 is the third member of a new class of genes of unknown function and almost ubiquitous expression that encode transmembrane proteins anchored in the endoplasmic reticulum[136]. Its identification as an asthma susceptibility gene epitomizes the heuristic power of GWA studies, particularly their ability to identify novel genes and pathways that potentially affect disease susceptibility. Several known asthma candidate genes (*STAT3*, *CRHR1*, *ITGB3* and *TBX21*) map to chromosome 17q21 (REF. 4), but *ORMDL3* had never emerged as a candidate. Genes similar to *ORMDL3* are found in yeast and other primitive organisms, suggesting that they may be components of an ancient and conserved immune mechanism.

Future research will need to explore the mechanisms underlying the impact of *ORMDL3* on asthma susceptibility. Also, despite the statistical strength of the association data, the presence of several SNPs independently associated with asthma makes it possible that *ORMDL3* may not be the exclusive determinant of disease susceptibility at the 17q21 locus. Thus, further analyses and replication studies are warranted. The strength of the effects unveiled by this study also deserves comment. Although the association signal from chromosome 17

Epigenetics
The study of heritable changes in gene function that occur without a change in the DNA sequence.

was the strongest genome-wide signal, the marker most robustly associated with disease had a frequency of 62% among asthmatics and 52% in non-asthmatics. The large size of the study population rendered this difference highly significant in statistical terms, but this distribution reiterates that variants in single genes, even those discovered through GWA searches, are unlikely to be sufficient for disease causation[135].

Looking forward

At first glance, the achievements in asthma genetics may appear both impressive and confusing, and for the same reasons. In a little over 10 years, many susceptibility genes have been identified as robust candidates, and the list keeps growing longer. A similar situation exists for diabetes, inflammatory bowel disease, obesity, Alzheimer's disease, coronary artery disease and schizophrenia, among others. There is now consensus that all these common conditions with familial clustering represent true complex diseases and are due to complicated interactions between an unknown number of genes and environmental (and possibly developmental) factors[8]. What remains unclear is the extent to which complex diseases are influenced by a few polymorphisms with large effects[137] or by the synergism among multiple variants individually associated with low risk[7]. It is also possible that common, low-risk polymorphisms set the threshold for disease susceptibility, whereas less frequent, more penetrant polymorphisms associate with extreme phenotypes. There is hope that GWA studies, if adequately designed and powered to detect gene–environment and gene–gene interactions, may prove to be effective tools for discovery and shed light on these issues, taking us closer to a truly comprehensive set of asthma genes.

What comes next? The necessity of addressing the interplay between genetic variants, developmental processes and epigenetic mechanisms is recognized more and more widely. Epidemiological studies consistently show that most of the factors affecting the initiation and course of allergic inflammation later in life appear to act within a narrow window of opportunity in early life[138] or even prenatally[139]. This evidence implies that developmental processes might critically modify the impact of genetic variation on asthma and allergy susceptibility. However, the interactions between development and genetic variants have been explored in only a few studies, such as a recent one which showed an age-specific influence of *CD14*-159CT (*rs2569190*) on the atopic phenotype[140]. The role of epigenetics in asthma and allergy is also still unexplored, but an argument, supported by evidence in plants, animals and humans[141–144], can be made for epigenetic mechanisms as the most plausible mediators of the plasticity of gene–environment interactions[117]. Moreover, epigenetic chromatin modifications are critical for the regulation of some of the strongest asthma susceptibility candidates, such as *IL13* (REF. 145) and *IL4* (REF. 146), and the activity and expression of enzymes involved in histone modifications appear to be altered in the airways of asthmatic patients[147]. Finally, developing approaches capable of identifying functional variants and the relevant mechanisms is undoubtedly a priority, particularly if these variants are to become therapeutic targets. Indeed, the ultimate promise of asthma genetics is not only to help unravel disease pathogenesis, but also, and perhaps more importantly, to lead to more effective prevention and cure.

1. Los, H., Postmus, P. E. & Boomsma, D. I. Asthma genetics and intermediate phenotypes: a review from twin studies. *Twin Res.* **4**, 81–93 (2001).
2. Cookson, W. O., Sharp, P. A., Faux, J. A. & Hopkin, J. M. Linkage between immunoglobulin E responses underlying asthma and rhinitis and chromosome 11q. *Lancet* **333**, 1292–1295 (1989).
3. Daniels, S. E. *et al*. A genome-wide search for quantitative trait loci underlying asthma. *Nature* **383**, 247–250 (1996).
4. Ober, C. & Hoffjan, S. Asthma genetics 2006: the long and winding road to gene discovery. *Genes Immun.* **7**, 95–100 (2006).
An excellent review of 10 years of genetics of asthma and asthma-related traits.
5. Kurz, T. *et al*. Fine mapping and positional candidate studies on chromosome 5p13 identify multiple asthma susceptibility loci. *J. Allergy Clin. Immunol.* **118**, 396–402 (2006).
6. Risch, N. & Merikangas, K. The future of genetic studies of complex human diseases. *Science* **273**, 1516–1517 (1996).
7. Carlson, C. S., Eberle, M. A., Kruglyak, L. & Nickerson, D. A. Mapping complex disease loci in whole-genome association studies. *Nature* **429**, 446–452 (2004).
A balanced discussion of the toolkit of asthma genetics and the strengths and weaknesses of each major approach.
8. Cookson, W. The immunogenetics of asthma and eczema: a new focus on the epithelium. *Nature Rev. Immunol.* **4**, 978–988 (2004).
9. Wills-Karp, M. & Ewart, S. L. Time to draw breath: asthma-susceptibility genes are identified. *Nature Rev. Genet.* **5**, 376–387 (2004).
10. Bosse, Y. & Hudson, T. J. Toward a comprehensive set of asthma susceptibility genes. *Annu. Rev. Med.* **58**, 171–184 (2007).
11. Guerra, S. & Martinez, F. D. Asthma genetics: from linear to multifactorial approaches. *Annu. Rev. Med.* **59**, 199–213 (2008).
A thoughtful discussion of phenotypic heterogeneity and gene–environment interactions in asthma.
12. Neale, B. M. & Sham, P. C. The future of association studies: gene-based analysis and replication. *Am. J. Hum. Genet.* **75**, 353–362 (2004).
13. Baldini, M. *et al*. 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.* **20**, 976–983 (1999).
14. Eder, W. *et al*. Toll-like receptor 2 as a major gene for asthma in children of European farmers. *J. Allergy Clin. Immunol.* **113**, 482–488 (2004).
Together with references104–107, this paper provides strong evidence for gene–environment interactions in human populations.
15. Werner, M. *et al*. *TLR4* gene variants modify endotoxin effects on asthma. *J. Allergy Clin. Immunol.* **112**, 323–330 (2003).
16. Fageras Bottcher, M. *et al*. A TLR4 polymorphism is associated with asthma and reduced lipopolysaccharide-induced interleukin-12(p70) responses in Swedish children. *J. Allergy Clin. Immunol.* **114**, 561–567 (2004).
17. Tantisira, K. *et al*. Toll-like receptor 6 gene (*TLR6*): single-nucleotide polymorphism frequencies and preliminary association with the diagnosis of asthma. *Genes Immun.* **5**, 343–346 (2004).
18. Lazarus, R. *et al*. Single nucleotide polymorphisms in innate immunity genes: abundant variation and potential role in complex human disease. *Immunol. Rev.* **190**, 9–25 (2002).
19. Hysi, P. *et al*. NOD1 variation, immunoglobulin E and asthma. *Hum. Mol. Genet.* **14**, 935–941 (2005).
20. Kabesch, M. *et al*. Association between polymorphisms in caspase recruitment domain containing protein 15 and allergy in two German populations. *J. Allergy Clin. Immunol.* **111**, 813–817 (2003).
21. Rosenstiel, P., Till, A. & Schreiber, S. NOD-like receptors and human diseases. *Microbes Infect.* **9**, 648–657 (2007).
22. Hong, J. *et al*. TLR2, TLR4 and TLR9 polymorphisms and Crohn's disease in a New Zealand Caucasian cohort. *J. Gastroenterol. Hepatol.* **22**, 1760–1766 (2007).
23. Arroyo-Espliguero, R., Avanzas, P., Jeffery, S. & Kaski, J. C. CD14 and toll-like receptor 4: a link between infection and acute coronary events? *Heart* **90**, 983–988 (2004).
24. Hobbs, K., Negri, J., Klinnert, M., Rosenwasser, L. J. & Borish, L. Interleukin-10 and transforming growth factor-β promoter polymorphisms in allergies and asthma. *Am. J. Respir. Crit. Care Med.* **158**, 1958–1962 (1998).
25. Silverman, E. S. *et al*. Transforming growth factor-β1 promoter polymorphism C-509T is associated with asthma. *Am. J. Respir. Crit. Care Med.* **169**, 214–219 (2004).
26. Litonjua, A. A. *et al*. Polymorphisms in signal transducer and activator of transcription 3 and lung function in asthma. *Respir. Res.* **6**, 52 (2005).
27. Shiina, T., Inoko, H. & Kulski, J. K. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens* **64**, 631–649 (2004).

28. Jinnai, N. *et al*. Polymorphisms in the prostaglandin E2 receptor subtype 2 gene confer susceptibility to aspirin-intolerant asthma: a candidate gene approach. *Hum. Mol. Genet.* **13**, 3203–3217 (2004).
29. Traidl-Hoffmann, C. *et al*. Pollen-associated phytoprostanes inhibit dendritic cell interleukin-12 production and augment T helper type 2 cell polarization. *J. Exp. Med.* **201**, 627–636 (2005).
30. Hwang, E. S., Szabo, S. J., Schwartzberg, P. L. & Glimcher, L. H. T helper cell fate specified by kinase-mediated interaction of T-bet with GATA-3. *Science* **307**, 430–433 (2005).
31. Pykalainen, M. *et al*. Association analysis of common variants of STAT6, GATA3, and STAT4 to asthma and high serum IgE phenotypes. *J. Allergy Clin. Immunol.* **115**, 80–87 (2005).
32. Tantisira, K. G. *et al*. TBX21: a functional variant predicts improvement in asthma with the use of inhaled corticosteroids. *Proc. Natl. Acad. Sci. USA* **101**, 18099–18104 (2004).
33. Kabesch, M. *et al*. A complete screening of the *IL4* gene: novel polymorphisms and their association with asthma and IgE in childhood. *J. Allergy Clin. Immunol.* **112**, 893–898 (2003).
34. Basehore, M. J. *et al*. A comprehensive evaluation of IL4 variants in ethnically diverse populations: association of total serum IgE levels and asthma in white subjects. *J. Allergy Clin. Immunol.* **114**, 80–87 (2004).
35. Loza, M. J. & Chang, B. L. Association between Q551R *IL4R* genetic variants and atopic asthma risk demonstrated by meta-analysis. *J. Allergy Clin. Immunol.* **120**, 578–585 (2007).
36. Schedel, M. *et al*. A signal transducer and activator of transcription 6 haplotype influences the regulation of serum IgE levels. *J. Allergy Clin. Immunol.* **114**, 1100–1105 (2004).
37. Morahan, G. *et al*. Association of IL12B promoter polymorphism with severity of atopic and non-atopic asthma in children. *Lancet* **360**, 455–459 (2002).
38. Wills-Karp, M. *et al*. Interleukin-13: central mediator of allergic asthma. *Science* **282**, 2258–2261 (1998).
39. Grunig, G. *et al*. Requirement for IL-13 independently of IL-4 in experimental asthma. *Science* **282**, 2261–2263 (1998).
40. Ghaffar, O. *et al*. IL-13 mRNA and immunoreactivity in allergen-induced rhinitis: comparison with IL-4 expression and modulation by topical glucocorticoid therapy. *Am. J. Respir. Cell. Mol. Biol.* **17**, 17–24 (1997).
41. Lordan, J. L. *et al*. Cooperative effects of T_H2 cytokines and allergen on normal and asthmatic bronchial epithelial cells. *J. Immunol.* **169**, 407–414 (2002).
42. Dealtry, G. B., Clark, D. E., Sharkey, A., Charnock-Jones, D. S. & Smith, S. K. Expression and localization of the T_H2-type cytokine interleukin-13 and its receptor in the placenta during human pregnancy. *Am. J. Reprod. Immunol.* **40**, 283–290 (1998).
43. Ribeiro-do-Couto, L. M. *et al*. High IL-13 production by human neonatal T cells: neonate immune system regulator? *Eur. J. Immunol.* **31**, 3394–3402 (2001).
44. Wenzel, S., Wilbraham, D., Fuller, R., Getz, E. B. & Longphre, M. Effect of an interleukin-4 variant on late phase asthmatic response to allergen challenge in asthmatic patients: results of two phase 2a studies. *Lancet* **370**, 1422–1431 (2007).
45. Vladich, F. D. *et al*. IL-13 R130Q, a common variant associated with allergy and asthma, enhances effector mechanisms essential for human allergic inflammation. *J. Clin. Invest.* **115**, 747–754 (2005).
This paper demonstrates how functional studies can elucidate the contribution of genetic variants to disease susceptibility.
46. Graves, P. E. *et al*. A cluster of seven tightly linked polymorphisms in the *IL-13* gene is associated with total serum IgE levels in three populations of white children. *J. Allergy Clin. Immunol.* **105**, 506–513 (2000).
47. Liu, X. *et al*. An *IL13* coding region variant is associated with a high total serum IgE level and atopic dermatitis in the German multicenter atopy study (MAS-90). *J. Allergy Clin. Immunol.* **106**, 167–170 (2000).
48. Wang, M. *et al*. A common IL-13 Arg130Gln single nucleotide polymorphism among Chinese atopy patients with allergic rhinitis. *Hum. Genet.* **113**, 387–390 (2003).
49. Heinzmann, A. *et al*. Association study of the IL13 variant Arg110Gln in atopic diseases and juvenile idiopathic arthritis. *J. Allergy Clin. Immunol.* **112**, 735–739 (2003).
50. Heinzmann, A. *et al*. Genetic variants of IL-13 signalling and human asthma and atopy. *Hum. Mol. Genet.* **9**, 549–559 (2000).
51. He, J. Q. *et al*. Genetic variants of the *IL13* and *IL4* genes and atopic diseases in at-risk children. *Genes Immun.* **4**, 385–389 (2003).
52. Tsunemi, Y. *et al*. Interleukin-13 gene polymorphism G4257A is associated with atopic dermatitis in Japanese patients. *J. Dermatol. Sci.* **30**, 100–107 (2002).
53. DeMeo, D. *et al*. Univariate and multivariate family-based association analysis of the IL-13 ARG130GLN polymorphism in the Childhood Asthma Management Program. *Genet. Epidemiol.* **23**, 335–348 (2002).
54. van der Pouw Kraan, T. C. *et al*. An IL-13 promoter polymorphism associated with increased risk of allergic asthma. *Genes Immun.* **1**, 61–65 (1999).
55. Howard, T. D. *et al*. Identification and association of polymorphisms in the Interleukin-13 gene with asthma and atopy in a Dutch population. *Am. J. Respir. Cell. Mol. Biol.* **25**, 377–384 (2001).
56. Hummelshoj, T. *et al*. Association between an interleukin-13 promoter polymorphism and atopy. *Eur. J. Immunogenet.* **30**, 355–359 (2003).
57. Liu, X. *et al*. Associations between specific serum IgE response and 6 variants within the genes *IL4*, *IL13*, and *IL4RA* in German children: the German Multicenter Atopy Study. *J. Allergy Clin. Immunol.* **113**, 489–495 (2004).
58. Cameron, L. *et al*. T_H2-selective enhancement of human *IL13* transcription by *IL13*-1112C > T, a polymorphism associated with allergic inflammation. *J. Immunol.* **177**, 8633–8642 (2006).
This paper provides strong evidence for the existence of gene–environment interactions in the nucleus.
59. Shirakawa, T. *et al*. Association between atopy and variants of the β subunit of the high-affinity immunoglobulin E receptor. *Nature Genet.* **7**, 125–129 (1994).
60. Donnadieu, E. *et al*. Competing functions encoded in the allergy-associated *FcεRI*β gene. *Immunity* **18**, 665–674 (2003).
61. Kabesch, M. *et al*. Polymorphisms in eosinophil pathway genes, asthma and atopy. *Allergy* **62**, 423–428 (2007).
62. Namkung, J. H. *et al*. IL-5 and IL-5 receptor α polymorphisms are associated with atopic dermatitis in Koreans. *Allergy* **62**, 934–942 (2007).
63. Johansson, S., Wennergren, G., Aberg, N. & Rudin, A. Clara cell 16-kd protein downregulates T_H2 differentiation of human naive neonatal T cells. *J. Allergy Clin. Immunol.* **120**, 308–314 (2007).
64. Nickel, R. G. *et al*. Atopic dermatitis is associated with a functional mutation in the promoter of the C-C chemokine RANTES. *J. Immunol.* **164**, 1612–1616 (2000).
65. Levy, H. *et al*. Association of defensin β-1 gene polymorphisms with asthma. *J. Allergy Clin. Immunol.* **115**, 252–258 (2005).
66. Sengler, C. *et al*. Clara cell protein 16 (*CC16*) gene polymorphism influences the degree of airway responsiveness in asthmatic children. *J. Allergy Clin. Immunol.* **111**, 515–519 (2003).
67. Nakamura, H. *et al*. Variant eotaxin: its effects on the asthma phenotype. *J. Allergy Clin. Immunol.* **108**, 946–953 (2001).
68. Chavanas, S. *et al*. Mutations in *SPINK5*, encoding a serine protease inhibitor, cause Netherton syndrome. *Nature Genet.* **25**, 141–142 (2000).
69. Walley, A. J. *et al*. Gene polymorphism in Netherton and common atopic disease. *Nature Genet.* **29**, 175–178 (2001).
70. Palmer, C. N. *et al*. Common loss-of-function variants of the epidermal barrier protein filaggrin are a major predisposing factor for atopic dermatitis. *Nature Genet.* **38**, 441–446 (2006).
71. Marenholz, I. *et al*. Filaggrin loss-of-function mutations predispose to phenotypes involved in the atopic march. *J. Allergy Clin. Immunol.* **118**, 866–871 (2006).
72. Palmer, C. N. *et al*. Filaggrin null mutations are associated with increased asthma severity in children and young adults. *J. Allergy Clin. Immunol.* **120**, 64–68 (2007).
73. Morar, N., Cookson, W. O., Harper, J. I. & Moffatt, M. F. Filaggrin mutations in children with severe atopic dermatitis. *J. Invest. Dermatol.* **127**, 1667–1672 (2007).
74. Ying, S., Meng, Q., Corrigan, C. J. & Lee, T. H. Lack of filaggrin expression in the human bronchial mucosa. *J. Allergy Clin. Immunol.* **118**, 1386–1388 (2006).
75. Hudson, T. J. Skin barrier function and allergic risk. *Nature Genet.* **38**, 399–400 (2006).
76. Baurecht, H. *et al*. Toward a major risk factor for atopic eczema: meta-analysis of filaggrin polymorphism data. *J. Allergy Clin. Immunol.* **120**, 1406–1412 (2007).
77. Smith, F. J. *et al*. Loss-of-function mutations in the gene encoding filaggrin cause ichthyosis vulgaris. *Nature Genet.* **38**, 337–342 (2006).
78. Drysdale, C. M. *et al*. Complex promoter and coding region β 2-adrenergic receptor haplotypes alter receptor expression and predict in vivo responsiveness. *Proc. Natl. Acad. Sci. USA* **97**, 10483–10488 (2000).
79. Moffatt, M. F. & Cookson, W. O. Tumour necrosis factor haplotypes and asthma. *Hum. Mol. Genet.* **6**, 551–554 (1997).
80. Sayers, I. *et al*. Allelic association and functional studies of promoter polymorphism in the leukotriene C4 synthase gene (*LTC4S*) in asthma. *Thorax* **58**, 417–424 (2003).
81. Fryer, A. A. *et al*. Polymorphism at the glutathione *S*-transferase *GSTP1* locus. A new marker for bronchial hyperresponsiveness and asthma. *Am. J. Respir. Crit. Care Med.* **161**, 1437–1442 (2000).
82. Kabesch, M. *et al*. Glutathione *S* transferase deficiency and passive smoking increase childhood asthma. *Thorax* **59**, 569–573 (2004).
83. Shin, H. D. *et al*. Association of thromboxane A_2 receptor (TBXA2R) with atopy and asthma. *J. Allergy Clin. Immunol.* **112**, 454–457 (2003).
84. Drazen, J. M. *et al*. Pharmacogenetic association between *ALOX5* promoter genotype and the response to anti-asthma treatment. *Nature Genet.* **22**, 168–170 (1999).
85. Matsuda, A. *et al*. Coding SNP in tenascin-C Fn-III-D domain associates with adult asthma. *Hum. Mol. Genet.* **14**, 2779–2786 (2005).
86. Gao, P. S. *et al*. Variants of *NOS1*, *NOS2*, and *NOS3* genes in asthmatics. *Biochem. Biophys. Res. Commun.* **267**, 761–763 (2000).
87. Hall, I. P. & Sayers, I. Pharmacogenetics and asthma: false hope or new dawn? *Eur. Respir. J.* **29**, 1239–1245 (2007).
88. Van Eerdewegh, P. *et al*. Association of the *ADAM33* gene with asthma and bronchial hyperresponsiveness. *Nature* **418**, 426–430 (2002).
This paper reported the identification of *ADAM33*, the first asthma susceptibility gene discovered by positional cloning.
89. Holgate, S. T. *et al*. Local genetic and environmental factors in asthma disease pathogenesis: chronicity and persistence mechanisms. *Eur. Respir. J.* **29**, 793–803 (2007).
90. Zhang, Y. *et al*. Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. *Nature Genet.* **34**, 181–186 (2003).
91. Mellor, J. It takes a PHD to read the histone code. *Cell* **126**, 22–24 (2006).
92. Allen, M. *et al*. Positional cloning of a novel gene influencing asthma from chromosome 2q14. *Nature Genet.* **35**, 258–263 (2003).
93. Laitinen, T. *et al*. Characterization of a common susceptibility locus for asthma-related traits. *Science* **304**, 300–304 (2004).
94. Nicolae, D. *et al*. Fine mapping and positional candidate studies identify *HLA-G*. as an asthma susceptibility gene on chromosome 6p21. *Am. J. Hum. Genet.* **76**, 349–357 (2005).
95. Tan, Z. *et al*. Allele-specific targeting of microRNAs to HLA-G. and risk of asthma. *Am. J. Hum. Genet.* **81**, 829–834 (2007).
96. Noguchi, E. *et al*. Positional identification of an asthma susceptibility gene on human chromosome 5q33. *Am. J. Respir. Crit. Care Med.* **172**, 183–188 (2005).
97. Balaci, L. *et al*. IRAK-M is involved in the pathogenesis of early-onset persistent asthma. *Am. J. Hum. Genet.* **80**, 1103–1114 (2007).
98. Soderhall, C. *et al*. Variants in a novel epidermal collagen gene (*COL29A1*) are associated with atopic dermatitis. *PLoS Biol.* **5**, e242 (2007).
99. Editorial. Framework for a fully powered risk engine. *Nature Genet.* **37**, 1153 (2005).
100. Hall, I. P. & Blakey, J. D. Genetic association studies in Thorax. *Thorax* **60**, 357–359 (2005).
101. Chanock, S. J. *et al*. Replicating genotype-phenotype associations. *Nature* **447**, 655–660 (2007).

This paper provides a thoughtful and in-depth discussion of how the results of association studies and their replication, or lack thereof, should be interpreted.
102. Hersh, C. P. *et al*. Comprehensive testing of positionally cloned asthma genes in two populations. *Am. J. Respir. Crit. Care Med.* **176**, 849–857 (2007).
103. Maier, L. M. *et al*. Association of IL13 with total IgE: Evidence against an inverse association of atopy and diabetes. *J. Allergy Clin. Immunol.* **117**, 1306–1313 (2006).
104. von Mutius, E. Influences in allergy: epidemiology and the environment. *J. Allergy Clin. Immunol.* **113**, 373–379 (2004).
105. Valdar, W. *et al*. Genetic and environmental effects on complex traits in mice. *Genetics* **174**, 959–984 (2006).
106. Martinez, F. D. Gene–environment interactions in asthma and allergies: a new paradigm to understand disease causation. *Immunol. Allergy Clin. North Am.* **25**, 709–721 (2005).
107. Finberg, R. W. & Kurt-Jones, E. A. CD14: chaperone or matchmaker? *Immunity* **24**, 127–129 (2006).
108. Lee, H. K., Dunzendorfer, S., Soldau, K. & Tobias, P. S. Double-stranded RNA-mediated TLR3 activation is enhanced by CD14. *Immunity* **24**, 153–163 (2006).
109. LeVan, T. D. *et al*. A common single nucleotide polymorphism in the CD14 promoter decreases the affinity of Sp protein binding and enhances transcriptional activity. *J. Immunol.* **167**, 5838–5844 (2001).
110. Jones, C. A. *et al*. Reduced soluble CD14 levels in amniotic fluid and breast milk are associated with the subsequent development of atopy, eczema, or both. *J. Allergy Clin. Immunol.* **109**, 858–866 (2002).
111. Leynaert, B. *et al*. Association between farm exposure and atopy, according to the CD14 C-159T polymorphism. *J. Allergy Clin. Immunol.* **118**, 658–665 (2006).
112. Ober, C., Tsalenko, A., Parry, R. & Cox, N. J. A second-generation genomewide screen for asthma-susceptibility alleles in a founder population. *Am. J. Hum. Genet.* **67**, 1154–1162 (2000).
113. Woo, J., Assa'ad, A., Heizer, A., Bernstein, J. & Hershey, G. The–59 C→T polymorphism of CD14 is associated with nonatopic asthma and food allergy. *J. Allergy Clin. Immunol.* **112**, 438–444 (2003).
114. Sengler, C. *et al*. Evaluation of the CD14 C-159T polymorphism in the German MAS cohort. *Clin. Exp. Allergy* **33**, 166–169 (2003).
115. Kabesch, M. *et al*. A promoter polymorphism in the *CD14* gene is associated with elevated levels of soluble CD14 but not with IgE or atopic diseases. *Allergy* **59**, 520–525 (2004).
116. Kedda, M. A. *et al*. The CD14 C-159T polymorphism is not associated with asthma or asthma severity in an Australian adult population. *Thorax* **60**, 211–214 (2005).
117. Vercelli, D. Genetics, epigenetics and the environment: switching, buffering, releasing. *J. Allergy Clin. Immunol.* **113**, 381–386 (2004).
118. Eder, W. *et al*. Opposite effects of CD14/-260 on serum IgE levels in children raised in different environments. *J. Allergy Clin. Immunol.* **116**, 601–607 (2005).
119. Zambelli-Weiner, A. *et al*. Evaluation of the CD14/-260 polymorphism and house dust endotoxin exposure in the Barbados Asthma Genetics Study. *J. Allergy Clin. Immunol.* **115**, 1203–1209 (2005).
120. Gern, J. E. *et al*. Effects of dog ownership and genotype on immune development and atopy in infancy. *J. Allergy Clin. Immunol.* **113**, 307–314 (2004).
121. Simpson, A. *et al*. Endotoxin exposure, CD14 and allergic disease: an interaction between genes and the environment. *Am. J. Respir. Crit. Care Med.* **174**, 386–392 (2006).
122. Meyers, D. A. *et al*. Genome screen for asthma and bronchial hyperresponsiveness: interactions with passive smoke exposure. *J. Allergy Clin. Immunol.* **115**, 1169–1175 (2005).
123. Weiss, L. A., Pan, L., Abney, M. & Ober, C. The sex-specific genetic architecture of quantitative traits in humans. *Nature Genet.* **38**, 218–222 (2006).
124. Vercelli, D. Learning from discrepancies: CD14 polymorphisms, atopy and the endotoxin switch. *Clin. Exp. Allergy* **33**, 153–155 (2003).
125. Kabesch, M. *et al*. IL-4/IL-13 pathway genetics strongly influence serum IgE levels and childhood asthma. *J. Allergy Clin. Immunol.* **117**, 269–274 (2006).
This work highlighted gene–gene interactions along the T_H2-cell-associated signalling pathway that might be important for allergy and asthma risk.
126. Howard, T. D. *et al*. Gene–gene interaction in asthma: IL4RA and IL13 in a Dutch population with asthma. *Am. J. Hum. Genet.* **70**, 230–236 (2002).
127. Chan, I. H. *et al*. Gene–gene interactions for asthma and plasma total IgE concentration in Chinese children. *J. Allergy Clin. Immunol.* **117**, 127–133 (2006).
128. Moffatt, M. F. *et al*. Genetic variants regulating *ORMDL3* expression contribute to the risk of childhood asthma. *Nature* **448**, 470–473 (2007).
This paper reported the results of the first GWA study for an asthma phenotype.
129. Hirschhorn, J. N. & Daly, M. J. Genome-wide association studies for common diseases and complex traits. *Nature Rev. Genet.* **6**, 95–108 (2005).
130. The International HapMap Consortium. A haplotype map of the human genome. *Nature* **437**, 1299–1320 (2005).
131. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
This paper provides an eloquent demonstration of the power of GWA studies and an in-depth discussion of issues related to study design, sample size and phenotyping.
132. Vercelli, D. & Martinez, F. D. The Faustian bargain of genetic association studies: bigger might not be better, or at least it might not be good enough. *J. Allergy Clin. Immunol.* **117**, 1303–1305 (2006).
133. Christensen, K. & Murray, J. C. What genome-wide association studies can do for medicine. *N. Engl. J. Med.* **356**, 1094–1097 (2007).
134. Duerr, R. H. *et al*. A genome-wide association study identifies *IL23R* as an inflammatory bowel disease gene. *Science* **314**, 1461–1463 (2006).
135. Williams, S. M. *et al*. Problems with genome-wide association studies. *Science* **316**, 1840–1842 (2007).
136. Hjelmqvist, L. *et al*. ORMDL proteins are a conserved new family of endoplasmic reticulum membrane proteins. *Genome Biol.* **3**, research0027.1–0027.16 (2002).
137. Farrall, M. Quantitative genetic variation: a post-modern view. *Hum. Mol. Genet.* **13** (Spec. No. 1), R1–R7 (2004).
138. Martinez, F. D. Toward asthma prevention — does all that really matters happen before we learn to read? *N. Engl. J. Med.* **349**, 1473–1475 (2003).
139. Ege, M. J. *et al*. Prenatal farm exposure is related to the expression of receptors of the innate immunity and to atopic sensitization in school-age children. *J. Allergy Clin. Immunol.* **117**, 817–823 (2006).
140. O'Donnell, A. R. *et al*. Age-specific relationship between CD14 and atopy in a cohort assessed from age 8 to 25 years. *Am. J. Respir. Crit. Care Med.* **169**, 615–622 (2004).
141. Cropley, J., Suter, C., Beckman, K. & Martin, D. Germline epigenetic modification of the murine A vy allele by nutritional supplementation. *Proc. Natl. Acad. Sci. USA* **103**, 17308–17312 (2006).
142. Rutherford, S. & Lindquist, S. Hsp90 as a capacitor for morphological evolution. *Nature* **396**, 336–342 (1998).
143. Queitsch, C., Sangster, T. & Lindquist, S. Hsp90 as a capacitor of phenotypic variation. *Nature* **417**, 618–624 (2002).
144. Kaati, G., Bygren, L. O., Pembrey, M. & Sjostrom, M. Transgenerational response to nutrition, early life circumstances and longevity. *Eur. J. Hum. Genet.* **15**, 784–790 (2007).
145. Webster, R. B., Rodriguez, Y., Klimecki, W. T. & Vercelli, D. The human IL-13 locus in neonatal $CD4^+$ T cells is refractory to the acquisition of a repressive chromatin architecture. *J. Biol. Chem.* **282**, 700–709 (2007).
146. Santangelo, S., Cousins, D. J., Winkelmann, N. E. & Staynov, D. Z. DNA methylation changes at human T_H2 cytokine genes coincide with DNase I hypersensitive site formation during $CD4^+$ T cell differentiation. *J. Immunol.* **169**, 1893–1903 (2002).
147. Barnes, P. J., Adcock, I. M. & Ito, K. Histone acetylation and deacetylation: importance in inflammatory lung diseases. *Eur. Respir. J.* **25**, 552–563 (2005).
148. Hoffjan, S. *et al*. Genetic variation in immunoregulatory pathways and atopic phenotypes in infancy. *J. Allergy Clin. Immunol.* **113**, 511–518 (2004).
149. Holloway, J. W. & Koppelman, G. H. Identifying novel genes contributing to asthma pathogenesis. *Curr. Opin. Allergy Clin. Immunol.* **7**, 69–74 (2007).
150. Leung, T. F. *et al*. A polymorphism in the coding region of interleukin-13 gene is associated with atopy but not asthma in Chinese children. *Clin. Exp. Allergy* **31**, 1515–1521 (2001).
151. Leung, T. F. *et al*. Association between candidate genes and lung function growth in Chinese asthmatic children. *Clin. Exp. Allergy* **37**, 1480–1486 (2007).
152. Chen, W., Ericksen, M. B., Levin, L. S. & Khurana Hershey, G. K. Functional effect of the R110Q *IL13* genetic variant alone and in combination with *IL4RA* genetic variants. *J. Allergy Clin. Immunol.* **114**, 553–560 (2004).

DATABASES

OMIM: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM
Asthma | atopic dermatits | inflammatory bowel disease | ichthyosis vulgaris | Netherton's syndrome
dbSNP: http://www.ncbi.nlm.nih.gov/projects/SNP/
rs1800925 | *rs1801275* | *rs1805010* | *rs1805015* | *rs1881457* | *rs1295686* | *rs20541* | *rs2243250* | *rs2303067* | *rs2569190* | *rs324011* | *rs847* | *rs4696480*

FURTHER INFORMATION

Donata Vercelli's homepage: http://bio5.arizona.edu/bio5/database.php?cmd=fac&faculty_id=2803
International HapMap Project: http://www.hapmap.org/
Seattle SNPs: http://pga.mbt.washington.edu/
The Single Nucleotide Polymorphism database: http://www.ncbi.nlm.nih.gov/projects/SNP/

SUPPLEMENTARY INFORMATION

See online article: S1 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF.