British Journal of Pharmacology

DOI:10.1111/j.1476-5381.2011.01222.x
www.brjpharmacol.org

Themed Issue: Respiratory Pharmacology

REVIEW

Genetics of complex respiratory diseases: implications for pathophysiology and pharmacology studies

Ma'en Obeidat and Ian P. Hall

Division of Therapeutics and Molecular Medicine, Nottingham Respiratory Biomedical Research Unit, University of Nottingham, Nottingham, UK

Correspondence
Ian P Hall, Division of Therapeutics and Molecular Medicine, Nottingham Respiratory Biomedical Research Unit, University Hospital of Nottingham, Nottingham NG7 2UH, UK. E-mail: ian.hall@nottingham.ac.uk

Keywords
asthma; COPD; FEV_1; genome wide association; meta-analysis

Received
23 November 2010
Revised
20 December 2010
Accepted
22 December 2010

There has been a huge influx of data on the genetics and genomics of respiratory diseases in the last few years. Powered by large sample sizes from collaborations worldwide, recent genome-wide association studies have convincingly implicated variants in different regions in the genome for association with complex respiratory traits. These new associations have the potential to offer invaluable insight into the pathophysiology of the normal and diseased respiratory system. The functional mechanisms underlying effects of both identified and novel variants will be the focus of research over the next few years. The identification of these mechanisms will not only increase our understanding of disease but may allow the development of new therapies to alleviate respiratory conditions. The implications of these approaches for studies of asthma and Chronic Obstructive Pulmonary Disease are covered in this review.

LINKED ARTICLES
This article is part of a themed issue on Respiratory Pharmacology. To view the other articles in this issue visit http://dx.doi.org/10.1111/bph.2011.163.issue-1

Background

It is well established that the risk of developing common respiratory diseases such as asthma and Chronic Obstructive Pulmonary Disease (COPD) is influenced by both genetic and environmental factors (Wang *et al.*, 2005). These influences are also important in quantitative traits such as lung function, which are important in such diseases. While the environmental factors underlying the development of asthma and COPD are reasonably well understood, with smoking for example as the main risk factor for COPD, until recently, less has been known about genetic factors underlying these conditions. It is clear from heritability studies that genetic factors play a major role in their development. For example, in asthma, twin studies have shown higher concordance rates in monozygotic than dizygotic twins (Duffy *et al.*, 1990), and there is a fourfold to fivefold increased prevalence in first-degree relatives. Heritability estimates range from 40% to 60% (Bosse and Hudson, 2007). Lung function measures are also highly heritable, with estimates for heritability reaching as high as 77% for forced expiratory volume in 1 second (FEV_1) (Hubert *et al.*, 1982; McClearn *et al.*, 1994). Airflow obstruction (defined by a reduced FEV_1 value and a reduced FEV_1/forced vital capacity (FVC) ratio) is a feature of both asthma and COPD, with fixed (i.e. nonreversible) airflow obstruction being a key diagnostic criterion for COPD (Rabe, Hurd *et al.*, 2007).

Identifying the genes underlying respiratory diseases is of major importance for a number of reasons. First, this will help us to understand more fully the pathophysiology underlying the development of disease and the normal functioning of the respiratory tract. Second, it may facilitate the development of novel treatment strategies based on newly identified drug targets. Third, by identifying a set of risk and safety genetic variants, we may be able to identify better ways to either prevent disease by improving risk assessment, or to make an earlier, more accurate diagnosis. In addition, we may be able to use genetic information to substratify disease into specific phenotypes which may respond differently. Finally, using genetic data, we hopefully will be able to tailor medicines to individuals who are more likely to benefit and less likely to develop adverse events (a subject area known as pharmacogenetics). The aims of this review are to describe recent advances in the genetics of airway disease, focusing on

asthma and COPD, and to discuss the implications of these advances for pharmacologists.

Over the last 10 years, there has been a revolution in the ability to identify underlying genetic factors responsible for the development of common complex diseases. This has been driven by the completion of the human genome project and the advent of novel high throughput platforms to aid extensive genotyping studies. Before the completion of the human genome project, two main methods were used for disease susceptibility gene identification: the genome-wide linkage approach and the candidate gene approach.

Linkage and candidate gene approaches

In genome-wide linkage scans, family members are genotyped with evenly spaced genetic markers covering all chromosomes. These are typically microsatellites: polymorphic DNA loci that consist of repeating units of 1–4 base pairs in length. A search is then made for genetic regions containing a higher-than-expected number of shared alleles among the affected individuals. If such a region is discovered, genes in the region become candidates for positional cloning and fine mapping (Baron 2001; Vercelli 2008), in which the region is examined by typing denser collections of single nucleotide polymorphisms (SNPs), the most common genetic variation in humans, in which one nucleotide base is substituted with another) (Bosse and Hudson, 2007). The regions identified by this approach were usually large, and a practical difficulty was that large family cohorts were hard to recruit, especially in late onset diseases like COPD. This approach has also proven generally underpowered to detect linkage where the underlying genetic risk factors were of modest magnitude (Risch and Merikangas 1996; Risch 2000). Nevertheless, there have been some success stories in asthma including the identification of *ADAM33* (Van Eerdewegh *et al.*, 2002), *DPP10* D (Allen *et al.*, 2003), *PHF11* (Zhang *et al.*, 2003), *GPRA* (Laitinen *et al.*, 2004), *HLA-G* (Nicolae *et al.*, 2005), *CYFIP2* (Noguchi *et al.*, 2005), *IRAK-M* (Balaci *et al.*, 2007), *OPN3*(White *et al.*, 2008) and *PLAUR* (Barton *et al.*, 2009). Full gene names are listed in Table 1.

The candidate gene approach, on the other hand, mostly utilizes population-based cohorts using a case control design. The underlying principle is to look for a significant difference in the frequency of genetic markers in a gene of interest between the two groups. If association is observed, then it suggests that the marker identified is either causally related to the disease or phenotype of interest, or is in Linkage Disequilibrium (LD) with a causative polymorphism (Rothman *et al.*, 2001). LD occurs when genotypes at two adjacent loci are not independent of each other because of the low probability of recombination events occurring within small genetic distances (Slatkin 2008). The gene choice for these studies is usually based on our knowledge of the function or pathway of that gene and its relevance to disease, and hence does not generally directly lead to discoveries of new biological pathways. Many candidate gene study findings have been hard to replicate (for example, see references Hersh *et al.*, 2005; Smolonska *et al.*, 2009 for reviews of the COPD literature). This is probably a reflection of the modest sample sizes used in many studies which makes them underpowered to detect true associations of modest magnitude. For a comprehensive list of genes identified using this approach in asthma and COPD, the reader is directed to a number of recent reviews (Hersh *et al.*, 2005; Ober and Hoffjan 2006; Vercelli 2008; Zhang *et al.*, 2008; Smolonska *et al.*, 2009; Weiss *et al.*, 2009). Table 2, adapted from reference Weiss *et al.* (2009), presents a list of candidate genes that have been associated with the asthma phenotype in at least three studies of sample sizes greater than a total of 300 subjects (150 cases and 150 controls).

In October 2010, the Human Genetic Epidemiology Navigator database (HuGE Navigator, Yu *et al.*, 2008) listed 674 and 519 genes reported to be associated with asthma and COPD, respectively, and their related traits. However, as discussed above, there has been a major problem in the replication of many of these findings.

Genome-Wide Association Studies (GWAS) approaches

The completion of the human genome sequencing led to the identification of a vastly expanded list of SNPs and also allowed the documention of the extent of linkage disequilibrium across the human genome in four populations from different ethnic backgrounds (the HapMap project, International HapMap 2005). Using this information, and taking advantage of the technological developments in dense SNP genotyping chips, it became feasible to conduct GWAS (Wellcome Trust Case Control 2007). The GWAS approach relies on the use of a dense set of SNPs giving coverage across the majority of the human genome to survey common genetic variation for a possible role in disease or to identify the heritable quantitative traits that underlie disease (Hirschhorn and Hirschhorn, 2005). By definition, this is a hypothesis-free approach that enables the discovery of novel disease associated genes and molecular pathways.

The era of GWAS in respiratory disease began in 2007 (Table 3), when the first asthma GWAS was published (Moffatt *et al.*, 2007). This study reported association of childhood asthma with *ORMDL3*, a gene of unknown function at the time. Several studies in different asthmatic populations have followed, mostly in children, replicating the findings (Bouzigon *et al.*, 2008; Galanter *et al.*, 2008; Hirota *et al.*, 2008; Tavendale *et al.*, 2008; Bisgaard *et al.*, 2009). The first GWAS to investigate association with quantitative pulmonary function measures was also reported in 2007. The study proposed a potential role for *GSTO2* and *IL6R* (Wilk *et al.*, 2007). A second GWA study for asthma published in 2008, reported findings of a genetic influence of variants in *CHI3L1* on asthma and a chitinase-like protein known as YKL-40 (Ober *et al.*, 2008).

A number of additional GWAS papers looking at asthma were published in 2009. One focused on the investigation of association with a specific disease subphenotype, eosinophil counts, in the blood of 9392 Icelanders (Gudbjartsson *et al.*, 2009), and showed associations with variants in five genes: additional analyses were presented looking at asthma in these

Table 1

Gene symbols reported in the review and their full names

Gene symbol	Gene name
ADAM33	ADAM metallopeptidase domain 33
DPP10	dipeptidyl-peptidase 10 (non-functional)
PHF11	PHD finger protein 11
GPRA (NPSR1)	neuropeptide S receptor 1
HLA-G	major histocompatibility complex, class I, G
CYFIP2	cytoplasmic FMR1 interacting protein 2
IRAK-M	interleukin-1 receptor-associated kinase 3
OPN3	opsin 3
PLAUR	plasminogen activator, urokinase receptor
ADRB2	adrenergic, beta-2-, receptor, surface
CCL11	chemokine (C-C motif) ligand 11
CCL5	chemokine (C-C motif) ligand 5
CD14	CD14 molecule
CYSLTR2	cysteinyl leukotriene receptor 2
EDN1	endothelin 1
FCER1B(MS4A2)	membrane-spanning 4-domains, subfamily A, member 2 (Fc fragment of IgE, high affinity I, receptor for; beta polypeptide)
GSTP1	glutathione S-transferase pi 1
IL10	interleukin 10
IL13	interleukin 13
IL4	interleukin 4
IL4 R	interleukin 4 receptor
ITGB3	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61
LTA	lymphotoxin alpha (TNF superfamily, member 1)
NAT2	N-acetyltransferase 2 (arylamine N-acetyltransferase)
NOD1	nucleotide-binding oligomerization domain containing 1
PAFAH (PAFAH1B1)	platelet-activating factor acetylhydrolase 1b, regulatory subunit 1 (45 kDa)
PTGDR	prostaglandin D2 receptor (DP
TLR9	toll-like receptor 9
TNF	tumor necrosis factor
UGB (SCGB1A1)	secretoglobin, family 2A, member 1
VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor
ORMDL3	ORM1-like 3 (S. cerevisiae)
GSTO2	glutathione S-transferase omega 2
IL6R	interleukin 6 receptor
CHI3L1	chitinase 3-like 1 (cartilage glycoprotein-39)
PDE4D	phosphodiesterase 4D, cAMP-specific
ADRA1B	adrenergic, alpha-1B-, receptor
PRNP	prion protein
TLE4	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila)
CHCHD9	coiled-coil-helix-coiled-coil-helix domain containing 9
IL1RL1	interleukin 1 receptor-like 1
IKZF2	IKAROS family zinc finger 2 (Helios)
GATA2	GATA binding protein 2
IL5	interleukin 5 (colony-stimulating factor, eosinophil)
SH2B3	SH2B adaptor protein 3

Table 1

Continued.

Gene symbol	Gene name
DENND1B	DENN/MADD domain containing 1B
CRB1	crumbs homolog 1 (Drosophila)
RAD50	RAD50 homolog (S. cerevisiae)
HLA-DQB1	major histocompatibility complex, class II, DQ beta 1
HHIP	hedgehog interacting protein
CHRNA 3	cholinergic receptor, nicotinic, alpha 3
CHRNA 5	cholinergic receptor, nicotinic, alpha 5
GSTCD	glutathione S-transferase, C-terminal domain containing
TNS1	tensin 1
HTR4	5-hydroxytryptamine (serotonin) receptor 4
AGER	advanced glycosylation end product-specific recepto
THSD4	thrombospondin, type I, domain containing 4
GPR126	G protein-coupled receptor 126
ADAM19	ADAM metallopeptidase domain 19
FAM13A	family with sequence similarity 13, member A
PTCH1	family with sequence similarity 13, member A
PID1	patched 1
IREB2	iron-responsive element binding protein 2
HLA-DQB1	major histocompatibility complex, class II, DQ beta 1
IL1RL1	interleukin 1 receptor-like 1
IL18R1	interleukin 18 receptor 1
IL33	interleukin 33
SMAD3	SMAD family member 3
IL2RB	interleukin 2 receptor, beta
GSDMA	gasdermin A
GSDMB	gasdermin B
BICD1	bicaudal D homolog 1 (Drosophila)

Nonstandard abbreviations used in the paper: Genes reported in the review are listed along with their full name in Table 1.

individuals. A second GWAS in Caucasian subjects reported association with variants in *PDE4D* (Himes *et al.*, 2009). Two further GWAS papers reported association with asthma in different populations. The first investigated susceptibility for asthma in children from the Mexican population and suggested a contribution of *TLE4* and *CHCHD9* (Hancock *et al.*, 2009). The second studied two independent populations of African ancestry and suggested association with SNPs in *ADRA1B, PRNP* and *DPP10*. (Mathias *et al.*, 2010). Momentum has gathered in 2010. A GWAS for childhood asthma suggested a role for variants in *DENND1B*. Another GWAS for asthma suggests SNPs in the *RAD50-IL13* and *HLA-DR/DQ* regions were associated with asthma (Li *et al.*, 2010).

The first GWAS for COPD was published in 2009 (Pillai *et al.*, 2009). This study identified risk SNPs in two regions. The first was at the alpha subunit of the nicotinic acetylcholine receptor (*CHRNA 3/5*) locus, a region previously linked to nicotine dependence and lung cancer (Saccone *et al.*, 2007; Berrettini *et al.*, 2008; Thorgeirsson *et al.*, 2008). The second region contains the gene for hedgehog-interacting protein (*HHIP*). An accompanying manuscript in the same journal reported the second GWAS for lung function measures in the Framingham Heart Study (Wilk *et al.*, 2009), and identified SNPs near *HHIP* to be associated with the percent predicted FEV_1/FVC ratio (Wilk *et al.*, 2009).

More recently, a third GWAS investigating associations with COPD has identified variants in *FAM13A* (Cho *et al.*, 2010). A recent GWAS for emphysema, assessed through high-resolution chest computed tomography in individuals with COPD has also implicated variants in *BICD1* (Kong *et al.*, 2010).

From the first wave of GWAS published for asthma and COPD, it became clear that the effect size estimates of the identified variants were typically modest (e.g. odds ratio <1.5). This means that very large sample sizes are needed to identify genetic variants of small magnitude with confidence. Many of these early papers published in the field reported associations which, while of potential interest, were not genome-wide significant using conventional cut offs (multiple testing corrections depending on the number of SNPs on the relevant

Table 2

Susceptibility genes for asthma and related traits using candidate gene approach. The genes that have been associated with the asthma phenotype and reported in at least three independent studies with sample sizes greater than 150 cases and 150 controls, and replication with the same single nucleotide polymorphism (SNP)

Gene	Reference sequence	Total populations showing SNP association with asthma
ADRB2	chr5	5
CCL11	chr17	3
CCL5	chr17	3
CD14	chr5	4
CYSLTR2	chr13	3
EDN1	chr6	3
FCER1B	chr11	9
GSTP1	chr11	8
IL10	chr1	4
IL13	chr5	8
IL4	chr5	11
IL4 R	chr16	7
ITGB3	chr17	3
LTA	chr6	3
NAT2	chr8	3
NOD1	chr7	4
PAFAH	chr6	3
PTGDR	chr14	5
TLR9	chr3	3
TNF	chr6	17
UGB (CC10)	chr11	4
VDR	chr12	3

Adapted from reference (Weiss, Raby *et al.*, 2009).

platforms used for genotyping). This has led to the establishment of consortia comprising multiple independent studies combined to allow pooled analyses to be undertaken (Herbert *et al.*, 2006; Rothman *et al.*, 2006; Zeggini *et al.*, 2008).

Genome-wide meta analyses

The SpiroMeta consortium was established to facilitate large-scale meta-analysis of GWAS of lung function from 14 cohorts of European ancestry (Repapi, Sayers *et al.*, 2010). The Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium also undertook a similar analysis of lung function associations (Hancock, Eijgelsheim *et al.*, 2010). The SpiroMeta consortium analysed associations with FEV_1 and FEV_1/FVC ($n = 20\,288$) using >2.5 million genotyped and imputed SNPs [imputation being the process of predicting genotypes that are not directly assayed in a sample of individuals (Li and Abecasis 2006; Marchini, Howie *et al.*, 2007)], followed by meta-analysis of top signals with data both from direct genotyping (32 184 additional individuals) and *in silico* summary association data from the CHARGE Consortium ($n = 21\,209$) and the Health 2000 survey ($n = 883$). SpiroMeta results confirmed the reported locus at 4q31 1 (the *HHIP* locus) and identified associations with FEV_1 or FEV_1/FVC and common variants at five additional loci: *TNS1*, *GSTCD*, *HTR4*, *AGER* and *THSD*4. The CHARGE consortium also analysed associations with FEV_1 and FEV_1/FVC ratio ($n = 20\,890$) and evaluated 30 high signal SNPs in 16 178 SpiroMeta participants. The CHARGE study confirmed the *HHIP* locus and also identified *AGER, HTR4* and *GSTCD*, and in addition, suggested a potential role of five additional genes: *GPR126, ADAM19, FAM13A PTCH1* and *PID1*.

A similar collaborative approach has also recently been undertaken for asthma. A recent large-scale collaborative GWAS investigated SNPs association in 10 365 persons with physician-diagnosed asthma and 16 110 unaffected persons (the GABRIEL consortium: a multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community) (Moffatt *et al.*, 2010). This study identified *IL1RL1/IL18R1*, *HLA-DQ*, *IL33*, *SMAD3* and *IL2RB* in addition to confirming the previously reported *ORMDL3/GSDMB* region. The association with the *ORMDL3/GSDMB* locus on chromosome 17q21 was specific to childhood-onset disease. In this study, only *HLA-DR* showed a significant genome-wide association with the total serum IgE concentration (Moffatt *et al.*, 2010).

A summary of the key findings from these GWAS approaches is shown in Table 3.

Lessons from genetic studies

Given all the recent papers reporting potentially novel genes important in the development of lung disease and relevant subphenotypes, what have we learned from GWAS approaches so far? Many of the genes/loci identified are novel in that they have not previously been linked to the traits investigated and hence, would have been missed using a traditional candidate gene approach. Interestingly, an evaluation in the SpiroMeta general population sample of genes previously reported in candidate gene studies to alter lung function did not suggest a strong contribution of these genes to FEV_1 or FEV_1/FVC ratio (unpublished data) suggesting that many previously reported candidate gene studies may prove to be false positives.

The statistically convincing novel associations seen in some of these large studies not only provide invaluable insights into the pathophysiology of lung disease but also provide insight into the genetic architecture of complex human diseases. Many of the SNPs reported map to introns or to intergenic regions, with no apparent connection to functionality. In fact, a survey of published GWAS papers found associated SNPs to be 45% intronic and 43% intergenic. Nine percent were nonsynonymous, 2% were in a 5′ or 3′ untranslated region, and 2% were synonymous (Hindorff *et al.*, 2009). This, however, is partly a reflection on the choice of SNPs used for genotyping on genome wide platforms. The important question now is where do we go next?

Table 3

Summery of published genome-wide association study in respiratory diseases and traits

Gene	Phenotype	Sample size and ethnicity	Year	Genome-wide significant findings (Yes/No)	Comments	Ref
ORMDL3	Childhood asthma	Discovery: 994 /1,243 * (Caucasians) Replication: 2320 subjects from Germany (Caucasians) 3301 subjects from the British 1958 Birth Cohort (Caucasians)	2007	Yes	SNPs in the 17q21 region showed a strong association with childhood asthma in both a UK family cohort and German case-control samples. SNPs in this region were also associated with increased *ORMDL3* mRNA expression in lymphoblastoid cell lines from asthmatic children.	(Moffatt, Kabesch *et al.*, 2007)
GSTO2 *IL6R*	Quantitative pulmonary function measures	1097–1222 (depending on phenotype) individuals from the Framingham Heart Study population. (Caucasians)	2007	No	The study utilized data on 71 000 SNPs. Two genes where proposed as potential candidate genes: *GSTO2* and *IL6R*.	(Wilk, Walter *et al.*, 2007)
CHI3L1	Asthma, bronchial hyperresponsiveness and measures of pulmonary function	632 Hutterites (Caucasian)	2008	Yes	Variants associated with elevated serum YKL-40 levels. YKL-40 was previously reported to be increased in the lungs and circulation of patients with severe asthma.	(Ober, Tan *et al.*, 2008)
PDE4D	Asthma	Discovery: 359 /846 * (Caucasians) Replication: Ten independent populations with different ethnicities totalling 18 891 individuals (4342 cases)	2009	No	Cases from the Childhood Asthma Management Program (CAMP) and genetically matched controls from the Illumina ICONdb public resource. The strongest region of association seen was on chromosome 5q12 in *PDE4D*	(Himes, Hunninghake *et al.*, 2009)
ADRA1B *PRNP* *DPP10*	Asthma	Discovery: 464 /471 * (African American) 929 asthmatics and their family members (African Caribbean) Replication: 994 / 1243* and 207 families (Caucasians) 1456/1973*, 200/200*, 264 /186*, 208/179 * (African Americans)	2009	No	None of the SNPs implicated in the discovery population were replicated in two European cohorts and four additional case-control studies of African Americans.	(Mathias, Grant *et al.*, 2010)
TLE4 *CHCHD9*	Childhood asthma	Discovery: 492 Trios (Mexicans) Replication: 177 Trios (Mexicans)	2009	No	Cases were children with asthma, predominantly atopic by skin prick test, and their parents using the Illumina HumanHap 550 K BeadChip.	(Hancock, Romieu *et al.*, 2009)
IL1RL1 *IKZF2* *GATA2* *IL5* *SH2B3*	Plasma eosinophil count	Discovery: 9392 (Icelanders) Replication: 12 118 (Europeans) 5212 (East Asians)	2009	Yes	Variants in *IL1RL1*, *IKZF2*, *GATA2*, *IL5*, and *SH2B3* showed association with eosinophil count. Three SNPs in *IL1RL1*, *IL33* and *WDR36* have also showed association with atopic asthma, with the *IL1RL1* SNP showing association with asthma as well.	(Gudbjartsson, Bjornsdottir *et al.*, 2009)
DENND1B *CRB1*	Asthma	Discovery: 793/1988* (Europeans) Replication: 917/1546* (Europeans) 1667/2045* (African Americans)	2010	Yes	The study also implicated the 17q21 locus harbouring *ORMDL3*	(Sleiman, Flory *et al.*, 2010)
RAD50 *IL13* *HLA-DQB1*	Asthma	473/1892* (Caucasians)	2010	No	Cases from The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens (TENOR) and 1892 Illumina general population controls.	(Li, Howard *et al.*, 2010)

Table 3

Continued.

Gene	Phenotype	Sample size and ethnicity	Year	Genome-wide significant findings (Yes/No)	Comments	Ref
HHIP	FEV_1/FVC ratio	Discovery: 7691 (Caucasians) Replication: 835 (Caucasians)	2009	Yes	Individuals from the Framingham Heart Study. This time, by genotyping more SNPs (550 000 SNPs) in more individuals (7691), the study identified associations of SNPs on chromosome 4q31 near *HHIP* with the percent predicted FEV_1/FVC ratio	(Wilk, Chen *et al.*, 2009)
CHRNA 3 *CHRNA 5* *HHIP*	COPD	Discovery: 823/810* (Caucasians) Replication: 389/472* (Caucasians) 2840 Caucasians family members	2009	Yes	The first GWAS for COPD. The study investigated association with ~500 000 SNPs. The *HHIP* locus did not reach genome-wide significance.	(Pillai, Ge *et al.*, 2009)
GSTCD *TNS1* *HHIP* *HTR4* *AGER* *THSD4*	Lung function measures (FEV_1 and FEV_1/FVC)	Discovery: 20 288 (Caucasians) Replication: 32 184 direct genotyping (Caucasians) 22 092 *in silico* replication (Caucasians) 54 276 total	2010	Yes	SpiroMeta consortium meta-analysis. >2.5 million genotyped and imputed SNPs tested. mRNA expression analysis showed all variants to be expressed in lung tissue.	(Repapi, Sayers *et al.*, 2010)
GSTCD *HHIP* *HTR4* *AGER* *GPR126* *ADAM19* *FAM13A* *PTCH1* *PID1*	Lung function measures (FEV_1 and FEV_1/FVC)	Discovery: 20 890 (Caucasians) Replication: 16 178 *in silico* (Caucasians)	2010	Yes	CHARGE consortium meta-analysis.	(Hancock, Eijgelsheim *et al.*, 2010)
FAM13A *HHIP* *CHRNA3* *CHRNA5* *IREB2*	COPD	Discovery: 2940 /1380* (Caucasians) Replication: 502/504* and two family-based cohorts (n = 3808) (Caucasians)	2010	Yes		(Cho, Boutaoui *et al.*, 2010)
HLA-DQ *IL1RL1* *IL18R1* *IL33* *SMAD3* *IL2RB* *GSDMA* *GSDMB*	Asthma	Discovery: 10 365/16 110 * (Caucasians)	2010	Yes	The GABRIEL (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community) Consortium. Association with the *ORMDL3*/*GSDMB* locus on chromosome 17q21 was specific to childhood-onset disease. Only *HLA-DR* showed a significant genome wide association with the total serum IgE concentration	(Moffatt, Gut *et al.*, 2010)
BICD1	Emphysema (Both qualitative and quantitative)	Caucasian subjects from three cohorts (n = 1586, 435 and 362)	2010	Yes	*BICD1* gene were previously associated with telomere length in leukocytes	(Kong, Cho *et al.*, 2010)

*Number of cases/number of controls.

Resequencing of loci showing strong and reproducible associations, to identify the rare (<1% frequency) and (hopefully) functional variants is the obvious next step (Mardis 2008; Metzker 2010). In addition, other relevant polymorphisms need to be identified and characterized (Weiss *et al.*, 2009). These include copy number variations (Wainr *et al.*, 2009), and other polymorphic variants such as insertions and deletions. Then, we need to investigate interactions (epistasis) between variants, and construct the pathways and networks which may underlie functional effects to better explain the observed genotype phenotype associations and answer the important question of why the identified variants explain only a small proportion of the trait heritability, the appropriately called 'missing heritability' issue (Altshuler *et al.*, 2008; Frazer *et al.*, 2009; Goldstein 2009; Manolio *et al.*, 2009).

All of this will only form the first step towards maximizing the biological and clinical applications of these findings. The identified variants need to be functionally characterized using integrated approaches. One example of this is the use of comparisons of genetic expression data and genotyping data to look for SNP associations with expression levels of particular genes of interest [the expression quantitative trait locus (eQTL) approach]. This has been used, for example, to try and determine the major genetic influences on the expression of genes identified to be relevant in asthma such as *ORMDL3* (mRNA levels) (Moffatt *et al.*, 2007) and *CHI3L1* (YKL-40 protein levels) (Ober *et al.*, 2008).

Nonsynonymous coding SNPs (change amino acid) should be considered a priority for future functional work, as their potential effects are easier to interpret. In addition, *in vitro* assays need to be developed to evaluate the effect of intronic and intergenic variants on gene expression and to help understand short- and long-range genetic control. Developing animal models to understand how these genes function in complex biological systems will no doubt be valuable (Ober *et al.*, 2010). Finally, integrated databases and bioinformatics tools will be needed to make the best use of data from multiple resources. These considerable challenges should lead to a clearer understanding of how the novel genes and pathways identified contribute to the development of respiratory diseases.

Finally, how relevant will this work prove to the pharmacologist? We believe the findings emerging from genetic approaches will form the basis for pharmacological studies for many years to come, given the probable key role for many of the genes identified in the pathophysiology of lung disease. With this in mind, some of the genes already identified are reasonably well understood: an example would be *HTR4*, which shows association with lung function in both the SpiroMeta and CHARGE studies described above. Not only has *HTR4* signalling and function already been studied (Repapi *et al.*, 2010), but there already exists a range of selective agents active at this receptor which will permit early functional studies to be pursued. However, other genes are much more challenging. For example, *GST-CD* was identified in both the CHARGE and SpiroMeta studies as being important in determining lung function parameters. Little is known regarding the function of this gene, and no selective agents exist which are known to inhibit or activate protein function. The potential challenges of gaining a full understanding of *GST-CD* function, and similar poorly characterized genes will be a mainstay of respiratory research over the next few years.

Acknowledgements

Work in the authors' laboratory is funded by MRC, Asthma UK, the EDRF and NIHR.

Conflict of interest

We have no conflict of interest to declare.

References

Allen M, Heinzmann A, Noguchi E, Abecasis G, Broxholme J, Ponting CP *et al.* (2003). Positional cloning of a novel gene influencing asthma from chromosome 2q14. Nat Genet 35: 258–263.

Altshuler D, Daly MJ, Lander ES (2008). Genetic mapping in human disease. Science 322: 881–888.

Balaci L, Spada MC, Olla N, Sole G, Loddo L, Anedda F *et al.* (2007). IRAK-M is involved in the pathogenesis of early-onset persistent asthma. Am J Hum Genet 80: 1103–1114.

Baron M (2001). The search for complex disease genes: fault by linkage or fault by association? Mol Psychiatry 6: 143–149.

Barton SJ, Koppelman GH, Vonk JM, Browning CA, Nolte IM, Stewart CE *et al.* (2009). PLAUR polymorphisms are associated with asthma, PLAUR levels, and lung function decline. J Allergy Clin Immunol 123: 1391–1400. e1317.

Berrettini W, Yuan X, Tozzi F, Song K, Francks C, Chilcoat H *et al.* (2008). Alpha-5/alpha-3 nicotinic receptor subunit alleles increase risk for heavy smoking. Mol Psychiatry 13: 368–373.

Bisgaard H, Bonnelykke K, Sleiman PM, Brasholt M, Chawes B, Kreiner-Moller E *et al.* (2009). Chromosome 17q21 gene variants are associated with asthma and exacerbations but not atopy in early childhood. Am J Respir Crit Care Med 179: 179–185.

Bosse Y, Hudson TJ (2007). Toward a comprehensive set of asthma susceptibility genes. Annu Rev Med 58: 171–184.

Bouzigon E, Corda E, Aschard H, Dizier MH, Boland A, Bousquet J *et al.* (2008). Effect of 17q21 variants and smoking exposure in early-onset asthma.[see comment]. N Engl J Med 359: 1985–1994.

Cho MH, Boutaoui N, Klanderman BJ, Sylvia JS, Ziniti JP, Hersh CP *et al.* (2010). Variants in FAM13A are associated with chronic obstructive pulmonary disease. Nat Genet 42: 200–202.

Duffy DL, Martin NG, Battistutta D, Hopper JL and Mathews JD (1990). Genetics of asthma and hay fever in Australian twins. Am Rev Respir Dis 142: 1351–1358.

Frazer KA, Murray SS, Schork NJ, Topol EJ (2009). Human genetic variation and its contribution to complex traits. Nat Rev Genetics 10: 241–251.

Galanter J, Choudhry S, Eng C, Nazario S, Rodriguez-Santana JR, Casal J *et al.* (2008). ORMDL3 gene is associated with asthma in three ethnically diverse populations. Am J Respir Crit Care Med 177: 1194–1200.

Goldstein DB (2009). Common genetic variation and human traits. N Engl J Med 360: 1696–1698.

Gudbjartsson DF, Bjornsdottir US, Halapi E, Helgadottir A, Sulem P, Jonsdottir GM *et al.* (2009). Sequence variants affecting eosinophil numbers associate with asthma and myocardial infarction. Nat Genet 41: 342–347.

Hancock DB, Eijgelsheim M, Wilk JB, S. Gharib SA, Loehr LR, Marciante KD *et al.* (2010). Meta-analyses of genome-wide association studies identify multiple loci associated with pulmonary function. Nat Genet 42: 45–52.

Hancock DB, Romieu I, Shi M, Sienra-Monge J-J, Wu H, Chiu GY *et al.* (2009). Genome-wide association study implicates chromosome 9q21.31 as a susceptibility locus for asthma in mexican children. Plos Genet 5: e1000623.

Herbert A, Gerry NP, McQueen MB, Heid IM, Pfeufer A, Illig T *et al.* (2006). A common genetic variant is associated with adult and childhood obesity. Science 312: 279–283.

Hersh CP, Demeo DL, Lange C, Litonjua AA, Reilly JJ, Kwiatkowski D *et al.* (2005). Attempted replication of reported chronic obstructive pulmonary disease candidate gene associations. Am J Respir Cell Mol Biol 33: 71–78.

Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP *et al.* (2009). Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. Am J Hum Genet 84: 581–593.

Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS *et al.* (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci USA 106: 9362–9367.

Hirota T, Harada M, Sakashita M, Doi S, Miyatake A, Fujita K *et al.* (2008). Genetic polymorphism regulating ORM1-like 3 (Saccharomyces cerevisiae) expression is associated with childhood atopic asthma in a Japanese population. J Allergy Clin Immunol 121: 769–770.

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

Hubert HB, Fabsitz RR, Feinleib M, Gwinn C (1982). Genetic and environmental influences on pulmonary function in adult twins. Am Rev Respir Dis 125: 409–415.

International HapMap C (2005). A haplotype map of the human genome.[see comment]. Nature 437: 1299–1320.

Kong X, Cho MH, Anderson W, Coxson HO, Muller N, Washko G *et al.* (2010). Genome-wide Association Study Identifies BICD1 as a Susceptibility Gene for Emphysema. Am J Respir Crit Care Med DOI: 10.1164/rccm.201004-05410C [Epub ahead of print].

Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P *et al.* (2004). Characterization of a common susceptibility locus for asthma-related traits.[see comment]. Science 304: 300–304.

Li X, Howard TD, Zheng RSL, Haselkorn T, Peters SP, Meyers DA *et al.* (2010). Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. J Allergy Clin Immunol 125: 328–335. e311.

Li Y, Abecasis GR (2006). Mach 1.0: rapid haplotype reconstruction and missing genotype inference. Am J Hum Genet 2290.

McClearn GE, Svartengren M, Pedersen NL, Heller DA, Plomin R (1994). Genetic and environmental influences on pulmonary function in aging Swedish twins. J Gerontol 49: 264–268.

Manolio TA, Collins FS *et al.* (2009). Finding the missing heritability of complex diseases. Nature 461: 747–753.

Marchini J, Howie B, Myers S, McVean G, Donnelly P (2007). A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39: 906–913.

Mardis ER (2008). The impact of next-generation sequencing technology on genetics. Trends Genet 24: 133–141.

Mathias RA, Grant AV, Rafaels N, Hand T, Gao L, Vergara C *et al.* (2010). A genome-wide association study on African-ancestry populations for asthma. J Allergy Clin Immunol 125: 336–346. e334.

Metzker ML (2010). Sequencing technologies – the next generation. Nat Rev Genetics 11: 31–46.

Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S *et al.* (2010). A large-scale, consortium-based genomewide association study of asthma. N Engl J Med 363: 1211–1221.

Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S *et al.* (2007). Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature 448: 470–473.

Nicolae D, Cox NJ, Lester LA, Schneider D, Tan Z, Billstrand C *et al.* (2005). Fine mapping and positional candidate studies identify HLA-G as an asthma susceptibility gene on chromosome 6p21. Am J Hum Genet 76: 349–357.

Noguchi E, Yokouchi Y, Zhang J, Shibuya K, Shibuya A, Bannai M *et al.* (2005). Positional identification of an asthma susceptibility gene on human chromosome 5q33. Am J Respir Crit Care Med 172: 183–188.

Ober C, Butte AJ, Elias JA, Lusis AJ, Gan W, Banks-Schlegel S *et al.* (2010). Getting from genes to function in lung disease: a National Heart, Lung, and Blood Institute workshop report. Am J Respir Crit Care Med 182: 732–737.

Ober C, Hoffjan S (2006). Asthma genetics 2006: the long and winding road to gene discovery. Genes Immun 7: 95–100.

Ober C, Tan Z, Sun Y, Possick JD, Pan L, Nicolae R *et al.* (2008). Effect of variation in CHI3L1 on serum YKL-40 level, risk of asthma, and lung function. N Engl J Med 358: 1682–1691.

Pillai SG, Ge D, Zhu G, Kong X, Shianna KV, Need AC *et al.* (2009). A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. Plos Genet 5: e1000421.

Rabe KF, Hurd S, Anzueto A, Barnes PJ, Buist SA, Calverley P *et al.* (2007). Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. Am J Respir Crit Care Med 176: 532–555.

Repapi E, Sayers I, Wain LV, Burton PR, Johnson T, Obeidat M *et al.* (2010). Genome-wide association study identifies five loci associated with lung function. Nat Genet 42: 36–44.

Risch NJ (2000). Searching for genetic determinants in the new millennium. Nature 405: 847–856.

Risch N, Merikangas K (1996). The future of genetic studies of complex human diseases. Science 273: 1516–1517.

Rothman N, Skibola CF, Wang SS, Morgan G, Lan Q, Smith MT *et al.* (2006). Genetic variation in TNF and IL10 and risk of non-Hodgkin lymphoma: a report from the InterLymph Consortium. Lancet Oncol 7: 27–38.

Rothman N, Wacholder S, Caporaso NE, Garcia-Closas M, Buetow K, Fraumeni JF Jr (2001). The use of common genetic polymorphisms to enhance the epidemiologic study of environmental carcinogens. Biochim Biophys Acta 1471: C1–10.

Saccone SF, Hinrichs AL, Saccone NL, Chase GA, Konvicka K, Madden PAF *et al*. (2007). Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Hum Mol Genet 16: 36–49.

Slatkin M (2008). Linkage disequilibrium – understanding the evolutionary past and mapping the medical future. Nat Rev Genetics 9: 477–485.

Sleiman PMA, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SAG *et al*. (2010). Variants of DENND1B associated with asthma in children. N Engl J Med 362: 36–44.

Smolonska J, Wijmenga C, Postma DS, Boezen HM (2009). Meta-analyses on suspected chronic obstructive pulmonary disease genes: a summary of 20 years' research. Am J Respir Crit Care Med 180: 618–631.

Tavendale R, Macgregor DF, Mukhopadhyay S, Palmer CN, Tavendale R, Macgregor DF *et al*. (2008). A polymorphism controlling ORMDL3 expression is associated with asthma that is poorly controlled by current medications. J Allergy Clin Immunol 121: 860–863.

Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP *et al*. (2008). A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 452: 638–642.

Van Eerdewegh P, Little RD, Dupuis J, Del Mastro RG, Falls K, Simon J *et al*. (2002). Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness.[see comment]. Nature 418: 426–430.

Vercelli D (2008). Discovering susceptibility genes for asthma and allergy. Nat Rev Immunol 8: 169–182.

Wain LV, Armour JAL, Tobin MD (2009). Genomic copy number variation, human health, and disease. Lancet 374: 340–350.

Wang WYS, Barratt BJ, Clayton DG, Todd JA (2005). Genome-wide association studies: theoretical and practical concerns. Nat Rev Genetics 6: 109–118.

Weiss ST, Raby BA, Rogers A (2009). Asthma genetics and genomics 2009. Curr Opin Genet Dev 19: 279–282.

Wellcome Trust Case Control C (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls.[see comment]. Nature 447: 661–678.

White JH, Chiano M, Wigglesworth M, Geske R, Riley J, White N *et al*. (2008). Identification of a novel asthma susceptibility gene on chromosome 1qter and its functional evaluation. Hum Mol Genet 17: 1890–1903.

Wilk JB, Chen T-H, Gottlieb DJ, Walter RE, Nagle MW, Brandler BJ *et al*. (2009). A genome-wide association study of pulmonary function measures in the Framingham Heart Study. Plos Genet 5: e1000429.

Wilk JB, Walter RE, Laramie JM, Gottlieb DJ, O'Connor GT (2007). Framingham Heart Study genome-wide association: results for pulmonary function measures. BMC Med Genet 8: S8.

Yu W, Gwinn M, Clyne M, Yesupriya A, Khoury MJ (2008). A navigator for human genome epidemiology. Nat Genet 40: 124–125.

Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T *et al*. (2008). Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 40: 638–645.

Zhang J, Pare PD, Sandford AJ (2008). Recent advances in asthma genetics. Respir Res 9: 4.

Zhang Y, Leaves NI, Anderson GG, Ponting CP, Broxholme J, Holt R *et al*. (2003). Positional cloning of a quantitative trait locus on chromosome 13q14 that influences immunoglobulin E levels and asthma. Nat Genet 34: 181–186.