Portelli, Michael A. and Sayers, Ian (2016) Genome-wide association studies in asthma. eLS . pp. 1-10.

eLS

Genome-wide Association Studies in Asthma

A24639

Michael Portelli and Ian Sayers

Division of Respiratory Medicine, Queen's Medical Centre, University of Nottingham, Nottingham, NG7 2UH.

Abstract:

Asthma is a complex respiratory disease, with both genetic and environmental factors contributing to disease susceptibility. Genome-wide Association Studies (GWAS) have now identified novel risk alleles and loci associated with asthma diagnosis and more recently clinical sub-groups of disease. However, while providing insight into potential disease mechanisms these risk alleles have modest effect sizes and account for a small proportion of the anticipated heritability of asthma. In this review we provide an overview of GWAS in asthma to date including reproducible associations and advances in our understanding of the biology of asthma. In addition we discuss ancestry specific findings and how genetics may contribute to the development of multiple allergic conditions known as the 'atopic march'. Finally, we outline the strengths and weaknesses of GWAS and look to future approaches including a greater focus to functional variation and assessment of gene-gene and gene-environment interactions.

Key words: Asthma, genome wide association study, single nucleotide polymorphism, heritability

Key Concepts:

- Asthma is a common respiratory disease that is heterogeneous with respect to its underlying pathology and clinical presentation.
- Susceptibility to develop asthma involves both genetic and environmental risk factors.
- Genome wide association is the current method of choice to identify genetic factors underlying complex, multifactorial diseases such as asthma with sufficient confidence.
- GWAS of asthma have identified several genetic risk factors and genes with confidence, *e.g. IL33, IL1RL1, ORMDL3* loci. Replication of findings remains the gold standard.
- GWAS of sub-groups of asthma e.g. childhood onset disease have identified specific genetic risk factors.
- The effect sizes identified in GWAS are typically modest for single variants.
- GWAS have provided a unique insight into the altered biology of asthma including changes in innate and adaptive immune responses, altered airway smooth muscle function and epithelial barrier/function abnormalities.
- While GWAS of asthma have been successful there remains a large missing heritability.
- Future approaches include better clinical definition of asthma and greater interrogation of genetic factors not currently addressed e.g. regulatory variants, rare variants, copy number variants and greater attention to gene-environment interactions, gene-gene interactions and epigenetic mechanisms.

Introduction:

Asthma is a common respiratory disease characterised by acute episodes of breathlessness, chronic inflammation of the airways, reversible airflow obstruction and increased airway hyper-responsiveness to a variety of environmental stimuli and allergens (1). Asthma is a complex disease with a large degree of heterogeneity in the age of onset, the nature of triggers, the severity of symptoms and the contribution of atopy. There is now compelling evidence that asthma is effected by the joint action of both genetic and environmental risk factors in addition to their main effects (2). It affects both children and adults and commonly exists with comorbidities including other allergic diseases such as Allergic Rhinitis (AR) and Atopic Dermatitis (AD), which also have substantial heritability (3). Genome-wide association studies (GWAS), which involve the testing of typically 500,000+ genetic variants for association with the disease, are currently the preferred method for studying complex multifactorial diseases such as asthma.

In this review, we discuss recent advances in our understanding of the genetic basis of asthma that have come from GWAS, including the strengths and limitations of these genetic approaches. Additionally, we discuss the new insights into the biology of asthma provided to date. Finally, we outline future directions in this area including improved phenotype definition and additional genetic approaches to identify causative variants.

Asthma is a complex genetic disorder

It has been known for over 100 years that asthma and atopic diseases asthma run in families. Using 621 atopic probands and 76 non-atopic controls and their families, it was shown in 1916 that 48.4% of atopic probands had a family history of atopy, compared with just 14.5% in the control population (4). Similarly, a very high concordance of asthma, AR and AD in parents and children was established in the 1970s in a study of 176 families (5). Twin studies have been instrumental in identifying a significant concordance of asthma that is higher in monozygotic twins (identical genotype) than in dizygotic twins (on average sharing half of their genes). A recent study using 25,306 twins aged 9 or 12 years identified the heritability of childhood asthma to be 82%(6). Overall genetic factors are thought to account for 60-80% of the susceptibility to develop asthma with a smaller effect attributable to environmental factors, however this does not preclude that the environment is important.

Therefore, asthma is considered a complex genetic disorder and, in contrast to single-gene disorders (e.g. cystic fibrosis), involves multiple genes with expression influenced by both genetic and environmental factors. Several environmental factors are important in asthma development including tobacco smoke exposure, respiratory viral infections, antibiotic use, diet, and allergen exposure. In particular, early-life exposures play an important role. Gender and ethnic background also have a significant contribution. Environmental contributions to asthma risk is nicely demonstrated by two key observations; i) the increase in asthma prevalence in developed countries over the last few decades and ii) the differences in asthma prevalence between rural/farming and city/non-farming children which cannot be driven by genetic factors alone (7). This complex mode of inheritance, combined with the heterogeneity in the presentation of the disease and differing environmental influences has made gene discovery in asthma a challenge. See also DOI: 10.1002/9780470015902.a0005565.pub2

Methods for gene identification: The move to Genome Wide Association Studies (GWAS) for complex disorders

Early studies of the genetics of asthma investigated inheritance through families containing multiple affected children, using linkage analyses and candidate gene approaches based on biology or location in the genome. However, the reproducibility of these findings was limited primarily because of inadequate power, subject heterogeneity (different phenotype definition), population stratification, and multiple testing without correction. This aside, several genes/loci where identified with confidence including; *DPP10, PCDH1, HLAG, NPSR1, PHF11, PLAUR, ADAM33, IL10, CD14, IL4, IL13, ADRB2, HLA-DRB1, HLA-DQB1, TNFA, FCER1B, INPP4A, STAT6* and *IL4RA* providing a novel insight into asthma biology (For excellent reviews see (8, 9))**.**

Our understanding of the complexity of genetic variation present in the human genome has improved dramatically with sequencing initiatives such as the HapMap project, 1000 genomes and most recently the 100,000 genomes projects. Recent figures suggest > 60 million single nucleotide polymorphisms (SNP) or single base pair changes exist in humans. Similarly, there is a growing realization that deletions, insertions, and expansions of tandem repeats also represent significant variation. Technological advances enabling the simultaneous genotyping of >1 million SNPs allows for the investigation of the role of polymorphisms spanning the entire genome in cases and controls with very stringent statistical thresholds, *e.g.* $P<5 \times 10^{-8}$ to account for the large number of tests completed. See Figure 1 for an overview of

approaches used to identify genetic loci associated with asthma diagnosis. See also DOI: 10.1002/9780470015902.a0021458 and DOI: 10.1002/9780470015902.a0021995

Insert Figure 1 here

In the following sections we provide an overview of current findings of recent GWAS for i) self-reported/doctor diagnosed asthma and ii) asthma that has been refined clinically to a specific sub-population of asthma patients, namely; childhood onset asthma, severe asthma, asthma with frequent exacerbation, asthma with co-morbidities, including allergic rhinitis, atopy, COPD and gender specific analyses.

Genetic associations identified in GWAS of asthma diagnosis

The first asthma GWAS was completed in 2007 and utilised a discovery cohort of 994 patients who presented with childhood onset asthma in comparison to 1,243 non-asthma controls (10). This GWAS identified a significant association to a locus on chromosome 17q21 that included multiple genes of interest, including genes for a) *zona pellucida* binding protein 2 (*ZPBP2*), b) gasdermin B (*GSDMB*) and c) orm1 like protein 3 (*ORMDL3*). Over time, this 17q21 locus has been confirmed as an association locus in independent studies with asthma (11), severe asthma (12) and asthma with severe exacerbations (13) as phenotypic end-points. Further evidence of the importance of this first GWAS defined locus has come through associations with several asthma-relevant clinical measures in independent cohorts such as lung function, bronchial hyper-responsiveness (BHR) and disease severity for the key 17q21 GWAS SNPs (14). However, the specific underlying gene(s) that explain the genetic association remains to be resolved and it is likely that multiple genes are underlying the signal(s). *ZPBP2, GSDMB* and *ORMDL3* have been reported to have a role in gene transcription, cell apoptosis and sphingolipid synthesis respectively. Recently a role for *ORMDL3* in eosinophil trafficking and degranulation, mechanisms thought to be important in asthma, has been identified (15).

There are now more than 50 studies registered with the NHGRI-EBI catalog of published GWAS for asthma and related traits (16). Typically, these studies include between 300-2,000 asthma subjects and therefore may be anticipated to identify ~50% of associations for common variants (minor allele frequency 10-50%) (17). A summary of the main findings from these asthma GWAS is shown in Table 1, focussed to loci that have been identified in the Caucasian population and have been verified by replication.

It was realised early on that very large numbers of subjects would be needed to identify genetic variants associated with asthma diagnosis with sufficient confidence to overcome the issues of differential asthma definition, ancestry diversity and the large number of known environmental factors contributing to susceptibility. The largest of these meta-analyses to date is the study carried out by the European GABRIEL (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community) consortium involving 10,365 cases and 16,110 controls (11). This study used primarily doctor diagnosed asthma as an end point identifying association to loci spanning multiple genes, including: *IL33, IL1RL1/IL18R1, HLA-DQ, SMAD3, IL2RB* and the 17q21 locus (11). Similarly, there has been a US led meta-analyses, the EVE consortium, consisting of 3,246 asthma cases and 3,385 controls and additional cohorts (1,702 case-parent trios, 355 family based cases and 468 family based control) including subjects from European, Latino and African ancestry (18). Four previously described loci associated in Caucasian subjects were identified; 17q21, *IL1RL1, IL33* and *TSLP* (18). Therefore, to date 12 asthma susceptibility loci have been identified using asthma diagnosis as and end point (Table 1) however it is important to note that the effect sizes of any single variant is modest, odds ratio (OR) 1.1-1.4. Overall the susceptibility genes identified to date are consistent with the hypothesis that asthma is caused by epithelial barrier/function abnormalities and altered innate and adaptive immune responses. It was reported by the GABRIEL consortium that ~49% of the lifetime risk of asthma could be explained by the loci identified in this study (11).

Insert Table 1 here.

Clinical refinement of asthma for GWAS

There is accumulating evidence that asthma is a heterogeneous condition involving multiple sub-groups with potentially different underlying causation, clinical presentation and therefore genetic basis. These groups have been identified through approaches such as cluster analyses that examine clinical (e.g. lung function), immunological (e.g. blood inflammatory cells) and epidemiological data (gender, age of onset) (19-21). A recent study combined this clustering of phenotypic information in 3,001 asthma subjects to identify four asthma groups and then completed a GWAS which identified novel genetic associations for i) active adult-onset non-allergic asthma and *CD200* and ii) inactive/mild non-allergic asthma with *GRIK2* (22).

Of these sub-groups, asthma age of onset has emerged as an important phenotype for asthma development. From a genetic perspective, heritability estimates have been shown to be inversely correlated with age of onset , suggesting that in childhood onset disease genetic factors are more important (23). This same study also demonstrated that genetic factors explained 34% of the variation in the age at onset of and environmental factors 66% (23). Also, as a sub-analyses of the GABRIEL study, chromosome 17q21 was identified as a specific locus for childhood onset asthma (11), similarly GWAS of mild-moderate childhood asthma with methacholine sensitivity and moderate-severe childhood asthma identified *PDE4D* and *DENND1B* loci respectively ((24, 25) Table 1).

Multiple recent studies have now started to investigate different sub-phenotypes of asthma to try and identify genetic drivers of specific asthma phenotypes. Such phenotypes have included: i) increased asthma exacerbations in patients taking inhaled corticosteroids (26), ii) early childhood asthma with exacerbations (13), iii) in never/low smoking asthma subjects (27) and moderate-severe asthma (12). These studies have identified several variants distinct to those previously reported for asthma diagnosis such as *CDHR3* for early childhood asthma with exacerbations (13). Interestingly, for several previously identified regions of interest, the median effect size reported was higher in studies with a refined clinical phenotype, suggesting that interpretation of doctor diagnosed asthma GWAS requires caution. However, this is not all together surprising as some loci will be of greater importance for different subsets of asthma patients. This is exemplified in the GWAS for severe asthma with exacerbations were the number of hospitalisations reported positively correlated to the SNP effect sizes, including those in the *IL33* locus (*e.g.* OR 1.32, 1.22, 1.47, 1.91 for 2, 3, 4/5 and 6 or more hospitalisations respectively) (13). Focussing to moderate-severe asthma, a recent UK study did not identify any novel locus meeting genome wide significance in the discovery analyses with suggestive data for novel loci *e.g. C5orf56, CD83* however this study confirmed previous signals at 17q21 and the *IL1RL1* loci in the combined analysis (12). As illustration we have included the Manhattan plot and 17q21 region plot from this study of moderate-severe asthma to demonstrate typical findings from a GWAS (Figure 2). In a similar approach focussed to severe, difficult to treat asthma subjects as part of the The Epidemiology and Natural History of Asthma: Outcomes and Treatment Regimens

(TENOR) study did not identify novel loci however association to the *IL13/RAD50* locus was confirmed (28).

Insert Figure 2 here.

Stratification has also included investigations into co-morbidities commonly associated with asthma such as asthma-COPD (29) and asthma-AR (30) overlap. The presence or absence of these co-morbidities with asthma may be triggered by a distinct and overlapping genetic profile. Loci associated with the combined phenotype were attenuated or absent when each phenotype was investigated in isolation when compared to the combined phenotype, again highlighting the need to consider asthma as a more complex and multi-stratified disease. Potential genetic associations included SNPs in genes: *CSMD1*, *SOX5*, for asthma/COPD (29) and *ZBTB10, IL33*, *IL1RL1*, *SMAD3*, *TSLP*, *c11orf30, ORMDL3* and *CLEC16A* for asthma/AR (30) (see Table 1).

Interestingly, stratification of asthma patients based on gender has recently led to the discovery of novel genetic determinants (31), suggesting that gender-stratification of asthma GWA has an important role to play in dissecting the genetic architecture of asthma. Another study confirmed the importance of gender-linked association by comparing three groups consisting of 2566 female cases, 2653 male cases and 3830 controls identifying four female specific association loci (Rap1GAP2/17p13.3, C6orf118/6q27, ERBB4/2q34, AK057517/2q23.3) and two male loci (IRF1/5q31.1, RAB11FTP2/10q26.11) in multiple ancestry groups (32).

Therefore while data generated by GWAS of asthma focussed to specific sub-groups of patients is only just emerging, it is clear that these analyses have identified overlapping and distinct genetic loci from asthma diagnosis adding further to the concept that genetics may contribute to the differential expression of asthma (see Figure 3).

Insert Figure 3 here.

Results of GWAS in other ethnic populations

Although the majority of GWAS to date have focussed to populations of European descent, recent studies have also considered other ancestries including African American, Mexican, Korean and Japanese cohorts. Torgerson *et al.* have used diverse North American populations including 5,416 individuals with asthma of

European American, African American or African Caribbean, and Latino ancestry with replication in 12,649 individuals from the same ethnic groups (18). Four previously described loci associated in Caucasian studies were identified; 17q21, *IL1RL1, IL33* and *TSLP*. However, importantly there appears to be some ancestry specific loci *e.g*. the 17q21 loci was particularly relevant to the Caucasian and Latino populations and a novel locus, *PYHIN1* was identified in populations of African ancestry only (18). *PYHIN1* encodes pyrin and HIN-domain family, member 1 and is an interferon inducible protein shown to regulate IFN-β and NO production in macrophages. Ancestry specific associations have also been identified in other populations. In the largest GWAS of asthma in the Japanese population to date, 7,171 cases and 27,912 controls were used to identify five loci; 4q31 (*USP38-GAB1*), 5q22 (*TSLP*), 6p21 (*HLA*), 10p14 (intergenic) and 12q13 (*IKZF4*) (33).

Overlap with other allergic diseases – the atopic march?

There is accumulating evidence that allergic diseases *e.g.* asthma, AR, AD and traits e.g. serum IgE, blood eosinophil counts share a large number of genetic susceptibility loci (3). Of note genetic polymorphisms within the *IL33* and *IL1RL1* (IL33 receptor) loci are thought to be of relevance for asthma, AD, allergic sensitisation and blood eosinophil counts suggesting the IL33 pathway may represent an underlying mechanism and therapeutic opportunity. Polymorphisms spanning *C11orf30/LRRC32* also show association with these traits. These overlapping loci may at least in part explain the concept of the “atopic march” *e.g.* childhood AD leads to an increased risk of developing asthma, as there is overlapping genetic susceptibility to both conditions. It is important to note that there is also clear trait specificity for many loci, most evident for *e.g.* the *FLG* locus for AD. Due to the common occurrence of comorbidities it is difficult to define which susceptibility loci are shared or specific. As discussed earlier, there is move to a more comprehensive phenotype definition in asthma genetics including stratification based on co-morbidities. One recent study aimed to address this for asthma and AD by stratifying patients based on AD (all), AD and asthma, and AD (no asthma) (34). Using a cohort of 1,563 childhood onset AD cases and 4,054 controls, five loci were identified as genome wide significant in all subjects; of interest the 1p21 (*FLG*) and 5q31 (*RAD50/IL13*) loci achieved markedly greater significance in the AD plus

asthma compared to the AD (no asthma) group (34). More recently, a GWAS in infantile AD followed by childhood asthma using 2,428 cases and 17,034 controls identified both novel loci (*EFHC1* on 6p12.3 and *TMTC2/SLC6A15* on12q21.3) and loci previously associated with multiple allergic traits (*FLG* (1q21.3), *IL4/KIF3A* (5q31.1), *AP5B1/OVOL1* (11q13.1), *C11orf30/LRRC32* (11q13.5) and *IKZF3* (17q21)) (35). This study provides further evidence for a genetic contribution to the atopic march. See also DOI: 10.1002/9780470015902.a0001887.pub3

What have we learnt so far – biology

Genetic findings from asthma GWAS have provided novel insights into the potential molecular mechanisms that underlie asthma development. For example *HLA* is anticipated to be important for T-cell-mediated inflammatory responses as is *IL2RB,* which is an intermediate molecule for T-cell survival. The signalling molecule, *SMAD3*, is known to be involved in fibrosis. Asthma GWAS results identifying loci related to interleukin 33 (*IL33*) and interleukin 33 receptor (*IL1RL1* or ST2) genes have identified mechanisms of relevance related to allergic sensitisation and blood eosinophilia (through other GWAS). These associations to *IL33* and *IL1RL1* were replicated in subsequent GWAS for asthma and severe asthma phenotypes (12, 13, 36) confirming their role. IL33 has been shown to be elevated in the airways of asthma patients; particularly in the airway structural cells including the bronchial epithelium, while the soluble form of its receptor ST2 encoded by *IL1RL1* was shown to be elevated during asthma exacerbation. Functional genetics, following GWAS results have allowed for the determination of putative mechanisms for GWAS identified polymorphisms, specifically those known to alter amino acid residues. An example of such is the functional genetic study focussing to the *IL1RL1* locus which has shown that the GWAS tagged SNPs may influence IL33 and sST2 production (37).

Overall to date genes identified may be involved in diverse roles such as the function and activation of inflammatory cells (*IL13, IL6R, DENND1B, LRRC32, IL2RB,* and *IL1RL1*), airway smooth muscle contraction (*PDE4D*), and cell apoptosis and

differentiation (*GSDMB*). This once more provides insight into the potential mechanisms of action that are involved in the development and modulation of asthma. Of special note is that a significant number of genes (*e.g. IL33, IL1RL1, C11orf30* and *TSLP*) that are known to be associated with epithelial cell functions and homeostasis. This further supports the hypothesis that the bronchial epithelium is altered in asthma (38). Further evidence for this concept is the recent finding that polymorphisms spanning *CDHR3* are associated with severe asthma with exacerbation (13). *CDHR3* encodes cadherin-related family member 3, with other family members being involved in epithelial polarity and cell-cell interactions. Recent data suggest that CDHR3 is the receptor for Rhinovirus, the most common respiratory virus associated with exacerbations in asthma, and that the key variant identified in GWAS modulates levels of CDHR3 providing a putative mechanism (39).

Of note is that, for the majority of asthma susceptibility loci identified to date it is unclear what are the key causative variants and genes(s) underlying the association. There are intensive efforts to close this gap in knowledge using approaches such as linkage disequilibrium mapping and eQTL analyses in lung tissue and lung relevant cells.

GWAS strengths and weaknesses: Reproducibility between approaches and missing heritability

While GWAS have many design/technological strengths including the ability to interrogate the genome on an unprecedented scale, the hypothesis free nature of the approach and the potential to identify causative variants it is important to note that GWAS to date in asthma have only been able to identify common variants with modest effect sizes (OR: 1.1-1.4) and have shown limited concordance with previous work. The lack of concordance between approaches (e.g., linkage versus GWAS) can be explained by the fact that the methodologies are designed to detect different types of variants (e.g., linkage analysis has good power to detect high-risk disease-causing alleles but is not effective at identifying common alleles of modest effect size as GWAS does). It is reassuring that many of the genes identified in candidate gene approaches have been reproduced in GWAS (e.g., the *IL13/IL4* locus on chromosome 5q31). Another key limitation of current GWAS in asthma is that SNPs

chosen for array design were not selected specifically for function. This means that results reported to date using very stringent statistical approaches represent “the low hanging fruit” and it is likely that causative SNPs exist in the statistical significance range $<10^{-4}$ or have simply not been interrogated yet (40). These considerations underlie the observation that variants identified by GWAS in asthma account for only a small fraction of the heritability, a concept that is called the “missing heritability” (41). Additional possible explanations to account for missing heritability include, i) rare variants with larger effect size not measured on existing platforms, ii) structural variation e.g. copy number variation, iii) gene-environment contributions, iv) gene-gene interactions v) epigenetic mechanisms and vi) overestimation of initial heritability.

Future Directions

It is beyond doubt GWAS in asthma have significantly increased our understanding of the genetic architecture of this complex respiratory disease and provided a novel understanding of potentially altered biology in the disease. These genetic findings highlight alterations in innate and adaptive immune responses, airway smooth muscle function and epithelial barrier/function. The future holds great promise to extend these studies particularly beyond asthma diagnosis to further define asthma sub-phenotypes with recent success including childhood onset asthma (17q21), childhood severe asthma with exacerbation (*CDHR3)* and identifying novel genetic determinants underlying the atopic march from AD to asthma (*EFHC1* and *TMTC2/SLC6A15*). Therefore, in addition to larger International Consortia involving tens of thousands of subjects investigating asthma diagnosis with improved power we also anticipate a drive to GWAS in refined studies of carefully characterised patients. This shift in focus is at least in part driven by the greater appreciation that asthma is heterogeneous and while very large numbers have been able to identify the “low hanging fruit” future approaches need to be focussed to asthma patients with more thorough clinical characterisation.

Data from GWAS of several human traits/diseases including asthma suggest that the majority of associated common SNPs are found in regulatory regions not in the

coding regions of genes and that these regions are enriched for *e.g.* DNase I hypersensitivity sites. Therefore the design of current platforms for GWAS is also an area of intense focus with greater emphasis on validated functional variation identified in initiatives such as Encyclopaedia of DNA Elements Consortium (ENCODE) (42) being a priority. Significant advances in our understanding of expression trait quantitative loci (eQTL) importantly in airway relevant cells and in lung tissue (43, 44) have helped identify potentially functional SNPs driving mRNA levels in both a cis and trans mechanism. These initiatives and improved sequencing information on rare variants were fundamental in the design of arrays used in UK initiatives such as the custom Affymetrix® array for UK Biobank, a study of 502,682 participants in the UK.

In addition to GWAS additional approaches are being used including exome sequencing and candidate gene resequencing which suggested an increased heterogeneity in asthma and the importance of rare variants. As costs for targeted resequencing and whole genome sequencing continue to decrease this makes approaches to investigate variation per se on a large scale a real possibility. The integration of environmental factors, known to be an important contributing factor in asthma will be a focus for research efforts allowing gene-environmental interaction to be identified beyond those identified for single genes e.g. interaction between CD14 rs2569190 and endotoxin exposure determining disease risk (2). The environment is particularly important for epigenetic changes driving disease, with accumulating evidence that the epigenome may be important in allergic diseases such as asthma. Recently, a genome-wide methylation association study identified a significant contribution of CpG islands in determining serum IgE levels, a major driver of multiple allergic diseases including allergic asthma (45).

In summary, future approaches to asthma gene discovery and translation will include: improved clinical definition, integrated models that include interactions with environmental factors, GWAS data from custom/functional arrays, epigenetic data, eQTL analyses, emerging sequencing approaches leading to pathways analyses and biological approaches. Overall a greater understanding of genetic variation in specific pathways which results in increased risk of developing asthma will generate greater understanding of the biology of this complex disease. This represents the

first stage to clinical translation and the development of new more effective treatments for asthma.

Acknowledgements

Research in the authors' laboratory is funded by Asthma UK, British Lung Foundation, Medical Research Council and Biotechnology and Biological Sciences Research Council.

1 Figures and Table

Reported Gene(s)	Locus	Biology	Associated end-point	Study (reference)
IL6R	1q21	Regulatory T-cell function, T-cell differentiation	A	(46)
DENND1B	1q31	Memory T-cell function	B	(25)
IL1RL1/IL18R1, SLC9A4	2q12	IL-33 receptor/sodium-hydrogen exchanger	A, B, C, D	(11, 12, 18, 30)
CD200	3q13	T-cell proliferation	E	(22)
TLR4	4p14	Pathogen recognition and activation of innate immunity	D	(30)
PDE4D	5q12	Cell signalling, inflammation, ASM function	B	(24)
TSLP	5q22	Activates dendritic cells, Th2 immune responses	A, D	(18, 30)
SLC22A4/RAD50/IL13/KIF3A	5q31	Organic cationic transporter/DNA repair/Th2 cytokine/cilia protein	A, C, F	(11, 13, 28)
IRF1	5q31	Involved in B lymphocyte expression	G	(32)
HLA-DRA/DRQ	6p21	T-cell responses/many additional genes in region	A, C	(11, 28, 36, 47)
GRIK2	6q16	Excitatory neurotransmitter	H	(22)
CDHR3	7q22	Epithelial polarity, cell–cell contact and differentiation	F	(13)
CSMD1	8p23	Regulator of complement activation and inflammation in the developing central nervous system	I	(29)
ZBTB10	8q21	May be involved in transcriptional regulation	D	(30)
IL33	9p24	Recruitment/activation of inflammatory cells	A, D, F	(11, 13, 30, 48)
C11orf30/LRRC32	11q13	Regulates gene expression, epithelial barrier/regulatory T-cell function.	A, D	(30, 46)
SOX5	12p12	Controls expression of extracellular matrix genes and cell proliferation	I	(29)
SMAD3	15q22	TGF-β signalling intermediate, fibrosis	A, D	(11, 30)
CLEC16A	16p13	Inflammatory cell function (ITAM receptor), Regulator of mitophagy	D	(30)
ORMDL3/GSDMB/ZPB2	17q21	Sphingolipid synthesis/cell apoptosis	A, B, C, D	(10-12, 30)
IL2RB	22q12	Binds IL-2/IL-15, lymphoid cell differentiation	A	(11)

2

3 **Table 1: Susceptibility genes for asthma diagnosis or asthma stratified into specific sub-**
4 **groups identified by Genome Wide Association Studies (GWAS).** A: Asthma Diagnosis, B:
5 Childhood asthma, C: Severe Asthma, D: Asthma with a diagnosis of AR, E: Active adult-
6 onset non-allergic asthma, F: Childhood severe asthma with exacerbation, G: Asthma
7 associated with the male gender, H: inactive/mild non-allergic asthma, I: Asthma with a
8 diagnosis of COPD. Data is focused to studies using individuals with Caucasian ancestry.

9

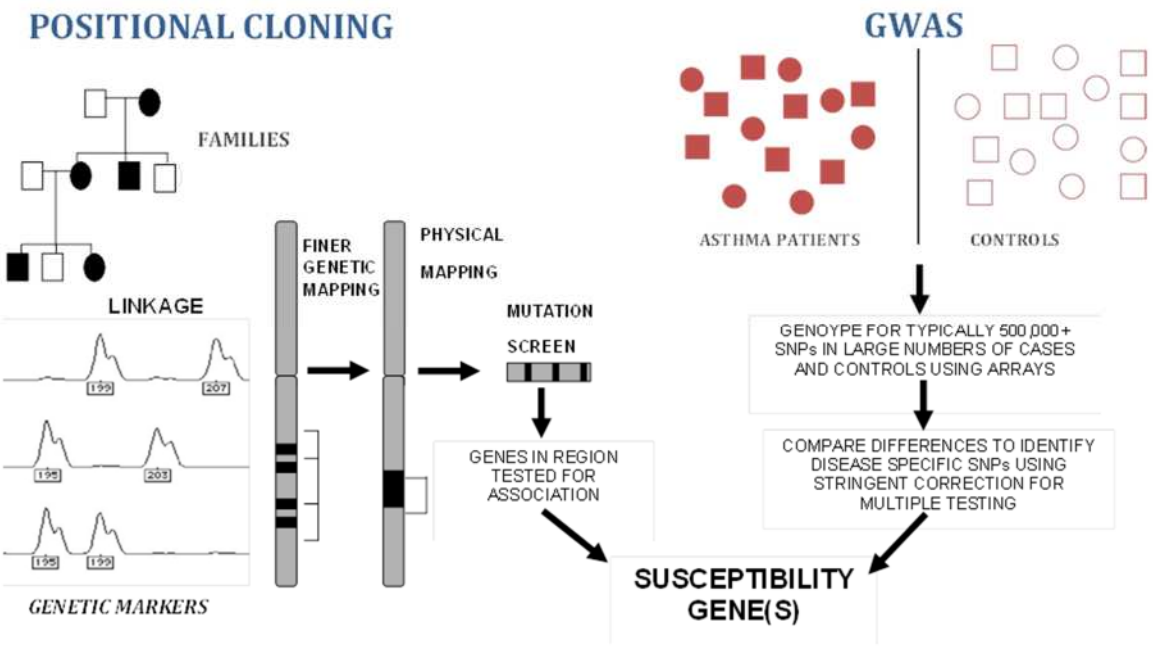
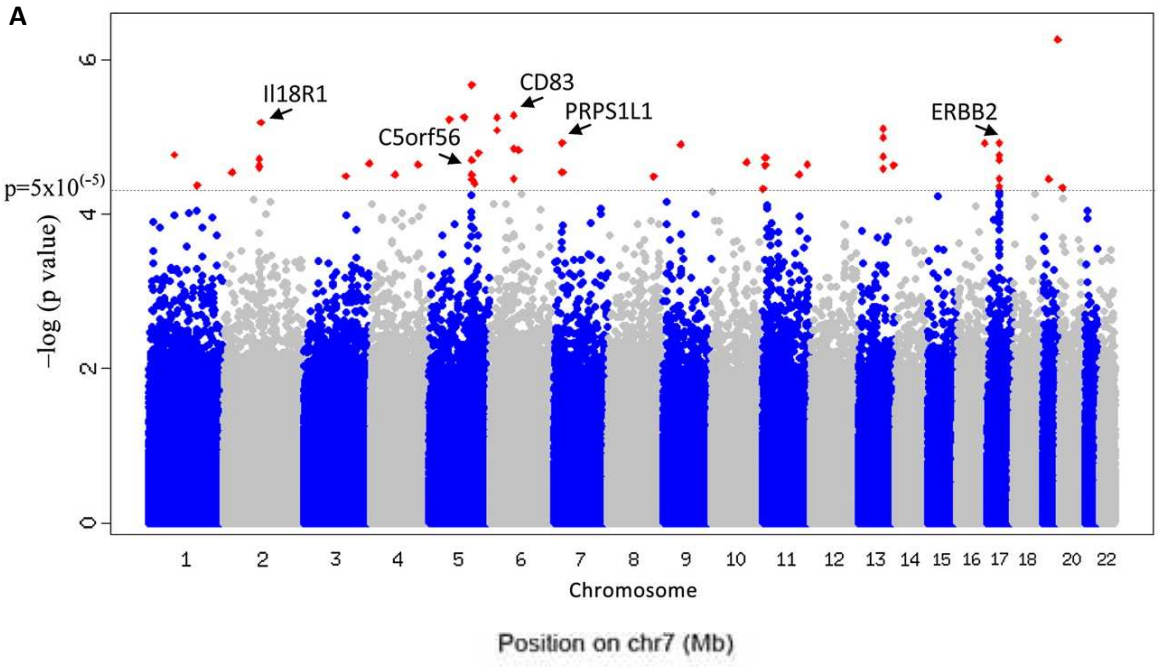


Figure 1: Asthma Gene Discovery Methods. Positional cloning involves linkage analyses which follows the transmission of genetic information through families with multiple affected children followed by fine association mapping. Genome Wide Association Studies (GWAS) looks at the frequency of a large number of common variants between cases and controls. Both approaches lead to novel gene discovery. Reproduced with permission from (1).



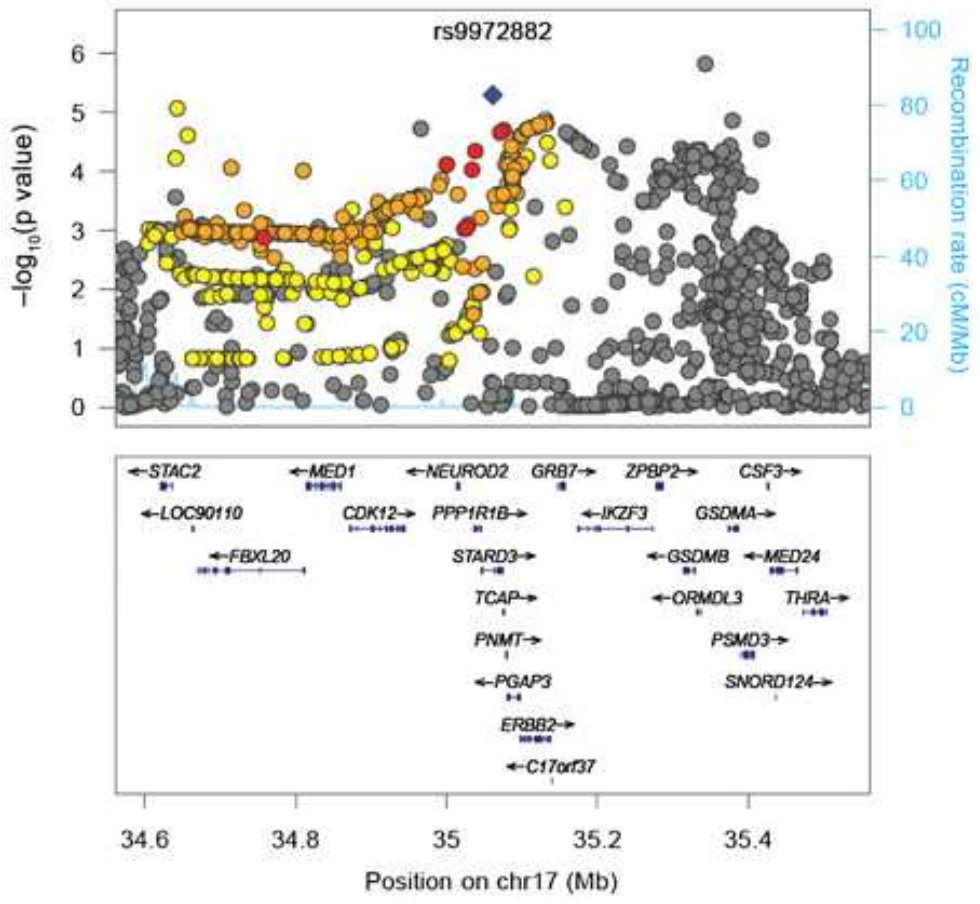


Figure 2: Manhattan (A) and chromosome 17q21 region (B) plots from the GWAS of moderate-severe asthma. Multiple suggestive signals (P<10-5) are apparent (red) and closer examination of 17q21 (B) illustrates the complexity of the signal demonstrating how

identification of the causative variant and gene can be a challenge. Reproduced with permission from (12).

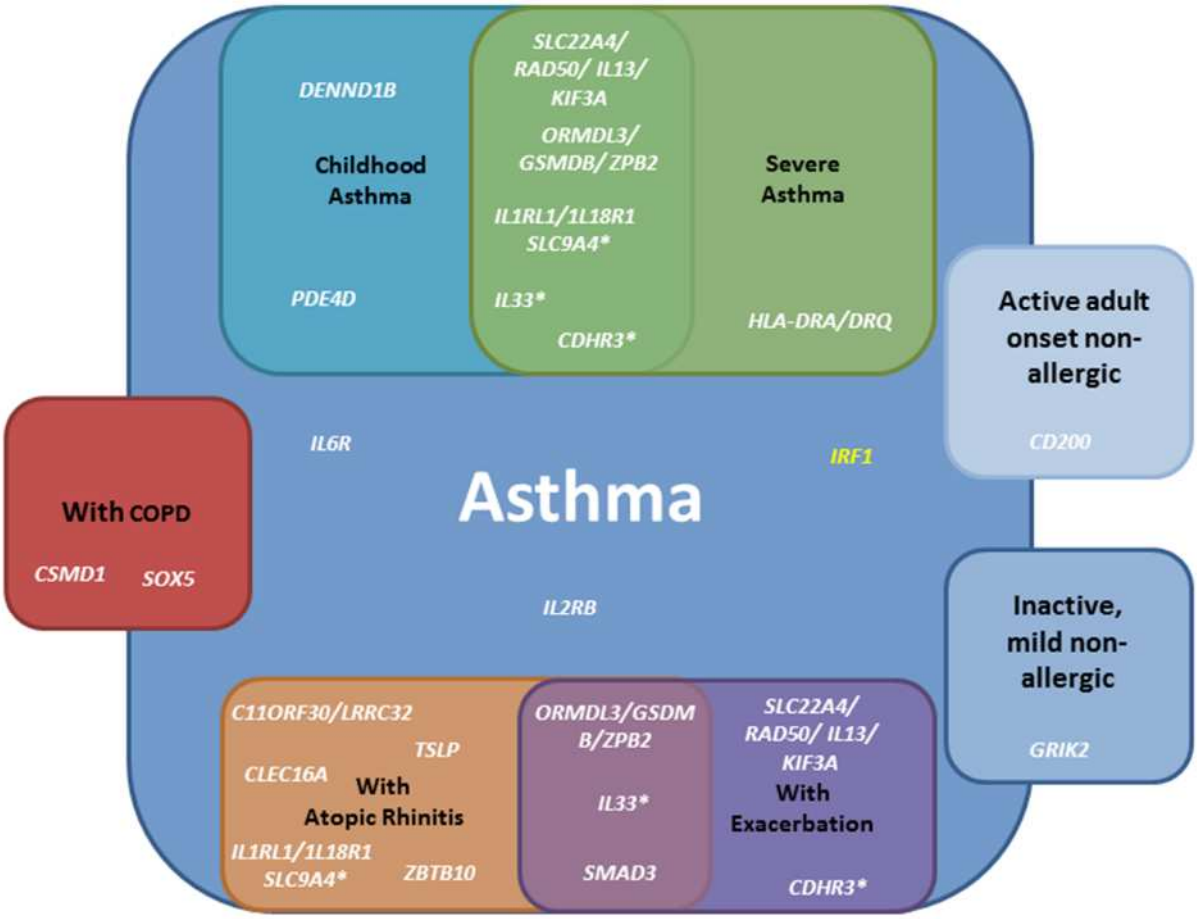


Figure 3: Schematic illustrating genetic loci identified in GWAS for asthma diagnosis or asthma stratified into specific sub-groups. Multiple signals identified in different populations are highlighted in the main blue box. Signals specific to the male gender are highlighted in yellow, while genes associated to asthma with co-morbidities are highlighted in their respective boxes. Loci associated with a specific sub-set of asthma are listed in their respective groups a top the main box. Genes that are relevant to different groups presented in box overlaps. Where overlap was not possible, genes presented multiple times in the diagram are highlighted with an asterisk.

References

1. Shaw DE, Portelli MP, Sayers I. Asthma. In: Padmanabhan S, editor. Handbook of Pharmacogenomics and Stratified Medicine. London: Elsevier; 2014. p. 617-51.
2. Custovic A, Marinho S, Simpson A. Gene-environment interactions in the development of asthma and atopy. Expert Rev Respir Med. 2012 Jun;6(3):301-8. PubMed PMID: 22788944.
3. Portelli MA, Hodge E, Sayers I. Genetic risk factors for the development of allergic disease identified by genome-wide association. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology. 2015 Jan;45(1):21-31. PubMed PMID: 24766371. Pubmed Central PMCID: PMC4298800.
4. Cooke RA, Vander Veer A. Human Sensitisation. Journal of immunology. 1916;1:201-305.
5. Gerrard JW, Vickers P, Gerrard CD. The familial incidence of allergic disease. Annals of allergy. 1976 Jan;36(1):10-5. PubMed PMID: 1247187. Epub 1976/01/01. eng.
6. Ullemar V, Magnusson PK, Lundholm C, Zettergren A, Melen E, Lichtenstein P, et al. Heritability and confirmation of genetic association studies for childhood asthma in twins. Allergy. 2016 Feb;71(2):230-8. PubMed PMID: 26786172.
7. von Mutius E. The microbial environment and its influence on asthma prevention in early life. The Journal of allergy and clinical immunology. 2016 Mar;137(3):680-9. PubMed PMID: 26806048.
8. Ober C, Yao TC. The genetics of asthma and allergic disease: a 21st century perspective. Immunological reviews. 2011 Jul;242(1):10-30. PubMed PMID: 21682736. Pubmed Central PMCID: 3151648.
9. Undarmaa S, Mashimo Y, Hattori S, Shimojo N, Fujita K, Miyatake A, et al. Replication of genetic association studies in asthma and related phenotypes. Journal of human genetics. 2010 Jun;55(6):342-9. PubMed PMID: 20395963. Epub 2010/04/17. eng.
10. Moffatt MF, Kabesch M, Liang L, Dixon AL, Strachan D, Heath S, et al. Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. Nature. 2007 Jul 26;448(7152):470-3. PubMed PMID: 17611496.
11. Moffatt MF, Gut IG, Demenais F, Strachan DP, Bouzigon E, Heath S, et al. A large-scale, consortium-based genomewide association study of asthma. The New England journal of medicine. 2010 Sep 23;363(13):1211-21. PubMed PMID: 20860503. Pubmed Central PMCID: 4260321.
12. Wan YI, Shrine NR, Soler Artigas M, Wain LV, Blakey JD, Moffatt MF, et al. Genome-wide association study to identify genetic determinants of severe asthma. Thorax. 2012 Sep;67(9):762-8. PubMed PMID: 22561531.
13. Bonnelykke K, Sleiman P, Nielsen K, Kreiner-Moller E, Mercader JM, Belgrave D, et al. A genome-wide association study identifies CDHR3 as a susceptibility locus for early childhood asthma with severe exacerbations. Nature genetics. 2014 Jan;46(1):51-5. PubMed PMID: 24241537.
14. Tulah AS, Holloway JW, Sayers I. Defining the contribution of SNPs identified in asthma GWAS to clinical variables in asthmatic children. BMC medical genetics. 2013;14:100. PubMed PMID: 24066901. Pubmed Central PMCID: 3849932.
15. Ha SG, Ge XN, Bahaie NS, Kang BN, Rao A, Rao SP, et al. ORMDL3 promotes eosinophil trafficking and activation via regulation of integrins and CD48. Nature communications. 2013;4:2479. PubMed PMID: 24056518. Pubmed Central PMCID: 3940275.
16. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic acids research. 2014 Jan;42(Database issue):D1001-6. PubMed PMID: 24316577. Pubmed Central PMCID: 3965119.
17. March ME, P.M.A. S, Hakonarson H. Genome-wide association studies in asthma: progress and pitfalls. Advances in Genomics and Genetics. 2015;5:1-7-109.
18. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American

populations. Nature genetics. 2011 Sep;43(9):887-92. PubMed PMID: 21804549. Pubmed Central PMCID: 3445408.
19. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, et al. Cluster analysis and clinical asthma phenotypes. American journal of respiratory and critical care medicine. 2008 Aug 1;178(3):218-24. PubMed PMID: 18480428. Pubmed Central PMCID: PMC3992366.
20. Jarjour NN, Erzurum SC, Bleecker ER, Calhoun WJ, Castro M, Comhair SA, et al. Severe asthma: lessons learned from the National Heart, Lung, and Blood Institute Severe Asthma Research Program. American journal of respiratory and critical care medicine. 2012 Feb 15;185(4):356-62. PubMed PMID: 22095547. Pubmed Central PMCID: PMC3297096.
21. Newby C, Heaney LG, Menzies-Gow A, Niven RM, Mansur A, Bucknall C, et al. Statistical cluster analysis of the British Thoracic Society Severe refractory Asthma Registry: clinical outcomes and phenotype stability. PloS one. 2014;9(7):e102987. PubMed PMID: 25058007. Pubmed Central PMCID: PMC4109965.
22. Siroux V, Gonzalez JR, Bouzigon E, Curjuric I, Boudier A, Imboden M, et al. Genetic heterogeneity of asthma phenotypes identified by a clustering approach. The European respiratory journal. 2014 Feb;43(2):439-52. PubMed PMID: 24311777.
23. Thomsen SF, Duffy DL, Kyvik KO, Backer V. Genetic influence on the age at onset of asthma: a twin study. The Journal of allergy and clinical immunology. 2010 Sep;126(3):626-30. PubMed PMID: 20673982.
24. Himes BE, Hunninghake GM, Baurley JW, Rafaels NM, Sleiman P, Strachan DP, et al. Genome-wide association analysis identifies PDE4D as an asthma-susceptibility gene. American journal of human genetics. 2009 May;84(5):581-93. PubMed PMID: 19426955. Pubmed Central PMCID: 2681010.
25. Sleiman PM, Flory J, Imielinski M, Bradfield JP, Annaiah K, Willis-Owen SA, et al. Variants of DENND1B associated with asthma in children. The New England journal of medicine. 2010 Jan 7;362(1):36-44. PubMed PMID: 20032318. Epub 2009/12/25. eng.
26. Dahlin A, Denny J, Roden DM, Brilliant MH, Ingram C, Kitchner TE, et al. CMTR1 is associated with increased asthma exacerbations in patients taking inhaled corticosteroids. Immunity, inflammation and disease. 2015 Dec;3(4):350-9. PubMed PMID: 26734457. Pubmed Central PMCID: 4693729.
27. Yatagai Y, Sakamoto T, Yamada H, Masuko H, Kaneko Y, Iijima H, et al. Genomewide association study identifies HAS2 as a novel susceptibility gene for adult asthma in a Japanese population. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology. 2014 Nov;44(11):1327-34. PubMed PMID: 25251750.
28. Li X, Howard TD, Zheng SL, Haselkorn T, Peters SP, Meyers DA, et al. Genome-wide association study of asthma identifies RAD50-IL13 and HLA-DR/DQ regions. The Journal of allergy and clinical immunology. 2010 Feb;125(2):328-35 e11. PubMed PMID: 20159242. Pubmed Central PMCID: 2824608. Epub 2010/02/18. eng.
29. Hardin M, Cho M, McDonald ML, Beaty T, Ramsdell J, Bhatt S, et al. The clinical and genetic features of COPD-asthma overlap syndrome. The European respiratory journal. 2014 Aug;44(2):341-50. PubMed PMID: 24876173. Pubmed Central PMCID: 4154588.
30. Ferreira MA, Matheson MC, Tang CS, Granell R, Ang W, Hui J, et al. Genome-wide association analysis identifies 11 risk variants associated with the asthma with hay fever phenotype. The Journal of allergy and clinical immunology. 2014 Jun;133(6):1564-71. PubMed PMID: 24388013. Pubmed Central PMCID: 4280183.
31. Mersha TB, Martin LJ, Biagini Myers JM, Kovacic MB, He H, Lindsey M, et al. Genomic architecture of asthma differs by sex. Genomics. 2015 Jul;106(1):15-22. PubMed PMID: 25817197. Pubmed Central PMCID: 4458428.
32. Myers RA, Scott NM, Gauderman WJ, Qiu W, Mathias RA, Romieu I, et al. Genome-wide interaction studies reveal sex-specific asthma risk alleles. Human molecular genetics. 2014 Oct 1;23(19):5251-9. PubMed PMID: 24824216. Pubmed Central PMCID: 4159149.

33. Hirota T, Takahashi A, Kubo M, Tsunoda T, Tomita K, Sakashita M, et al. Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. Nature genetics. 2012 Nov;44(11):1222-6. PubMed PMID: 23042114. Epub 2012/10/09. eng.
34. Weidinger S, Willis-Owen SA, Kamatani Y, Baurecht H, Morar N, Liang L, et al. A genome-wide association study of atopic dermatitis identifies loci with overlapping effects on asthma and psoriasis. Human molecular genetics. 2013 Dec 1;22(23):4841-56. PubMed PMID: 23886662. Pubmed Central PMCID: 3820131. Epub 2013/07/28. eng.
35. Marenholz I, Esparza-Gordillo J, Ruschendorf F, Bauerfeind A, Strachan DP, Spycher BD, et al. Meta-analysis identifies seven susceptibility loci involved in the atopic march. Nature communications. 2015;6:8804. PubMed PMID: 26542096. Pubmed Central PMCID: PMC4667629.
36. Ramasamy A, Kuokkanen M, Vedantam S, Gajdos ZK, Couto Alves A, Lyon HN, et al. Genome-wide association studies of asthma in population-based cohorts confirm known and suggested loci and identify an additional association near HLA. PloS one. 2012;7(9):e44008. PubMed PMID: 23028483. Pubmed Central PMCID: 3461045.
37. Ho JE, Chen WY, Chen MH, Larson MG, McCabe EL, Cheng S, et al. Common genetic variation at the IL1RL1 locus regulates IL-33/ST2 signaling. The Journal of clinical investigation. 2013 Oct;123(10):4208-18. PubMed PMID: 23999434. Pubmed Central PMCID: 3784527.
38. Holgate ST, Arshad HS, Roberts GC, Howarth PH, Thurner P, Davies DE. A new look at the pathogenesis of asthma. Clinical science. 2010 Apr;118(7):439-50. PubMed PMID: 20025610. Pubmed Central PMCID: 2805922.
39. Bochkov YA, Watters K, Ashraf S, Griggs TF, Devries MK, Jackson DJ, et al. Cadherin-related family member 3, a childhood asthma susceptibility gene product, mediates rhinovirus C binding and replication. Proceedings of the National Academy of Sciences of the United States of America. 2015 Apr 28;112(17):5485-90. PubMed PMID: 25848009. Pubmed Central PMCID: PMC4418890.
40. Ober C. Asthma Genetics in the Post-GWAS Era. Ann Am Thorac Soc. 2016 Mar;13 Suppl 1:S85-90. PubMed PMID: 27027959.
41. Maher B. Personal genomes: The case of the missing heritability. Nature. 2008 Nov 6;456(7218):18-21. PubMed PMID: 18987709.
42. Consortium EP. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012 Sep 6;489(7414):57-74. PubMed PMID: 22955616. Pubmed Central PMCID: PMC3439153.
43. Luo W, Obeidat M, Di Narzo AF, Chen R, Sin DD, Pare PD, et al. Airway Epithelial Expression Quantitative Trait Loci Reveal Genes Underlying Asthma and Other Airway Diseases. American journal of respiratory cell and molecular biology. 2016 Feb;54(2):177-87. PubMed PMID: 26102239. Pubmed Central PMCID: PMC4821039.
44. Obeidat M, Hao K, Bosse Y, Nickle DC, Nie Y, Postma DS, et al. Molecular mechanisms underlying variations in lung function: a systems genetics analysis. The Lancet Respiratory medicine. 2015 Oct;3(10):782-95. PubMed PMID: 26404118.
45. Liang L, Willis-Owen SA, Laprise C, Wong KC, Davies GA, Hudson TJ, et al. An epigenome-wide association study of total serum immunoglobulin E concentration. Nature. 2015 Apr 30;520(7549):670-4. PubMed PMID: 25707804. Pubmed Central PMCID: PMC4416961.
46. Ferreira MA, Matheson MC, Duffy DL, Marks GB, Hui J, Le Souef P, et al. Identification of IL6R and chromosome 11q13.5 as risk loci for asthma. Lancet. 2011 Sep 10;378(9795):1006-14. PubMed PMID: 21907864. Pubmed Central PMCID: 3517659.
47. Lasky-Su J, Himes BE, Raby BA, Klanderman BJ, Sylvia JS, Lange C, et al. HLA-DQ strikes again: genome-wide association study further confirms HLA-DQ in the diagnosis of asthma among adults. Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology. 2012 Dec;42(12):1724-33. PubMed PMID: 23181788. Epub 2012/11/28. eng.
48. Torgerson DG, Ampleford EJ, Chiu GY, Gauderman WJ, Gignoux CR, Graves PE, et al. Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American

populations. Nature genetics. 2011 Sep;43(9):887-92. PubMed PMID: 21804549. Pubmed Central PMCID: 3445408. Epub 2011/08/02. eng.

Further Reading

1. Bonnelykke K, Sparks R, Waage J, Milner JD. Genetics of allergy and allergic sensitization: common variants, rare mutations. Current opinion in immunology. 2015 Oct;36:115-26. PubMed PMID: 26386198. Pubmed Central PMCID: 4610744.
2. Meyers DA, Bleecker ER, Holloway JW, Holgate ST. Asthma genetics and personalised medicine. The Lancet Respiratory medicine. 2014 May;2(5):405-15. PubMed PMID: 24794577. Pubmed Central PMCID: 4768462.
3. Rava M, Smit LA, Nadif R. Gene-environment interactions in the study of asthma in the postgenomewide association studies era. Curr Opin Allergy Clin Immunol. 2015 Feb;15(1):70-8. PubMed PMID: 25479314.