Since the recognition that asthma is genetically determined, enormous progress has been made to understand which genes determine disease development in interaction with each other and/or with different environmental factors. This is the result of rapid developments in techniques for genotyping and statistical analyses. However, we are only at the beginning of understanding the complex nature of asthma. Follow-up of both clinical and environmental measures in a standardized way in numerous prospective cohorts is needed to establish which sets of genes and environmental factors determine the different phenotypes of asthma. This includes differences between sex, asthma severity, asthma remission, and asthma progression. Investigation of genetic profiling and gene expression profiling may further help to better understand the underlying pathophyslogic mechanisms, taking into account the insight in the biology currently present. Only by collaborative efforts of many groups of researchers that join forces in DNA analyses will it be possible to help to develop preventive strategies for asthma.

Keywords: asthma; genetics; gene–gene; environment; association

Since Mendel discovered the law of segregation in 1865 and, almost a century after this, Watson and Crick made public the structure of DNA, we have made progress in better understanding the role of genetics in disease development. Unlike monogenic diseases like cystic fibrosis, in which alterations in a single gene explain all or nearly all development of disease, genes underlying common diseases like asthma are likely to be multiple. Each of these genes acts in concert, or act in interaction with environmental stimuli leading to clinical disease (Figure 1). Asthma prevalence and severity vary significantly across populations, a variance that may be due to genetic and/or environmental differences. Since many genes and many environmental factors play a role, it is clear that intricate networks have to be elucidated before we know all possible combinations in the causal origins of asthma. This may give more insight into different types of asthma that may have similar symptoms and/or disease severity, but different underlying genes and/or environmental interaction of these genes.

WHERE WERE WE?

Segregation

It has been known for centuries that asthma clusters in families. By two-locus segregation of immunoglobulin (Ig)E levels in families with an asthma proband, Meyers and coworkers showed that the first locus explained 65% of the variation in IgE, and the second locus explained 19%, whereas their combination explained 87.4% (1). This seemed an easy way to go. However, it has now become clear that no single gene so far explains 65% of the variation in IgE.

Positional Cloning Approach

Positional cloning starts with the investigation of families without a predetermined hypothesis regarding the location or identity of the underlying susceptibility gene or genes. Markers are randomly spaced throughout the entire genome and tested for linkage (i.e., coinheritance) with a disease phenotype. This is then followed by finely spaced typing of genetic markers (fine mapping) to pinpoint the exact gene causing the linkage. The approach is time consuming, since in-depth analysis of a particular region of linkage that still can cover a large part of a chromosome requires considerable molecular analysis. So far, seven genes have been found by positional cloning (2–4): DPP10, CYFIP2, HLAG, GPRA, SFRS8, PHF11, and ADAM33. Table 1 presents their chromosomal location, replication, and putative function. Several of these genes have been replicated, including at the level of a single nucleotide polymorphism (SNP). Many genes are still under scrutiny for their function. Of interest, the table shows that this unbiased approach finds genes that are important modulators of epithelial barrier and lymphocyte function in asthma, functions that are important in the host defense.

Meta-Analysis of Linkage Studies

The lack of replication in linkage studies is probably due to low statistical power resulting from insufficient sample sizes and the potential of type I and Type II error. A recent meta-analysis combined data of 11 linkage studies to investigate consistency of linkage across studies (5). It replicated suggestively five genes found by positional cloning. The chromosomal regions that contained susceptibility genes using the most stringent genomewide statistical criteria were 6p22.3-p21.1 (hyperresponsiveness), 5q11.2-q14.3 and 6pter-p22.3 (total IgE), 3p22.1-q22.1, and 17p12-q24.3 (positive skin test). Interestingly, no significant association was found with asthma, either due to heterogeneity of the disease, variable diagnosis made by doctors in the different countries of the populations under study in combination with the difficulty of the diagnosis in childhood, or the fact that binary data (affected or not) are less powerful than continuous traits (e.g., IgE) to detect linkage. This suggests that objective markers of disease add homogeneity to the data and improve results, an observation that is important for all genetic research in complex, heterogeneous diseases. The meta-analysis demonstrated that overlap between genomic regions for the different traits under study (i.e., 6p22.3-p21.1 [HLA region], chromosome 5q23.2-q34 [HLA, IL13, IL5, IL12B], and
Asthma and atopy development.

Several studies, suggesting that they indeed are involved in the pathogenesis of atopy and allergic disease phenotypes (6). Several genes (e.g., \( IL13 \), \( HLA-DR \), \( IL4RA \)) have been replicated over 10 times in various studies, suggesting that they indeed are involved in asthma and atopy development.

**Gene–Environment Interaction**

The heritability of asthma has been reported to vary between 60 and 80%, leaving a remaining 20 to 40% for environmental contributions (11). Both linkage and association studies indicate that many early life exposures influence the risk for asthma and its related phenotypes in a genotype-specific manner. These early life exposures include exposure to endotoxins, viruses, pets, daycare environment, and environmental tobacco smoke (13). For instance, linkage of hyperresponsiveness to chromosome 5q was conditioned to families in whom environmental tobacco smoke exposure was present in the offspring during pregnancy or early childhood (12). An important example of gene–environmental interaction has been provided by studies showing that children on farms had a lower prevalence of asthma and allergy than those raised outside a farm. It has been hypothesized that contact with microbial products, such as endotoxin, modifies the immune response and asthma development (14). Eder and coworkers showed that children raised on a farm carrying an SNP in the gene encoding TLR2, a receptor for microbial products like endotoxin, had a significantly less frequent diagnosis of asthma and current asthma symptoms than farmers’ children without this SNP (15). Thus the environment (i.e., living on farms) modified the association between the \( TLR2 \) gene and asthma.

** Advances in Technology**

Genome-wide association (GWA) scans are emerging as powerful tools to identify genes involved in complex diseases, and this is currently the mapping strategy of choice. On the basis of phase I HapMap data, a catalog of all human genetic variation, it was shown that approximately 250,000 to 500,000 SNPs are required to capture all common SNPs in human populations. Although these numbers appear impressive, current technologies can evaluate 1,000,000 SNPs simultaneously. Improved cost efficiency of genotyping platforms has allowed the performance of many GWA studies, since large sample sizes are required to detect genetic variants of modest effects. To find a true association, larger sample sizes are required with:

- A greater number of genotypes and association tests (greater type I error)
- Phenotyping misclassification
- A lower frequency of the risk allele
- A lower size of the genetic effect
- A lower \( r^2 \) between disease associated SNPs and tag SNPs on the platform
- Heterogeneity of the association, that is, ancestry difference across population subsets, gene–environment and gene–gene interaction

The conventional \( P \) value of less than 0.05 is divided by the number of tests performed to assess whether true association is found with GWA, thus requiring \( P \) values of \( 5 \times 10^{-7} \) to \( 10^{-8} \),
a stringent level of significance. GWA has resulted in the identification of two novel asthma susceptibility genes, ORMDL3 and CHI3L1 (16, 17). These genes have been replicated and functional studies have supported their relevance in asthma (16, 17). Even though multiple, frequently replicated genes have now been put forward, each of these genes only explains a small part of the asthma risk in the population. This suggests that gene–gene interaction and interaction with various environmental factors will have to be solved before the development of asthma can be fully understood.

WHERE DO WE GO?

It is clear that GWAs have brought a lot of new information on complex diseases, and we eagerly await results of GWAs that are currently underway. There are additional variants in the genome that are not easily captured by GWA (i.e., deletions and insertions, inversions, and copy number variants [CNVs]). These variants have been largely ignored so far, yet it is clear that they are important in asthma, as exemplified by the relevance of the GSTM1-null variant that has been associated with asthma in many studies (2). Expansions and refinements of current technologies already contribute to the elucidation of these variants and some successes have been made to date (18). Structural variants (particularly CNVs) can now be studied at a genome-wide scale, using either SNP data or array CGH technology, and will provide new insights into asthma genetics as well.

Epigenetics

Epigenomics, the merged science of epigenetics and genomics, has arisen as a new discipline with the aim of understanding genetic regulation and its contribution to cellular growth and differentiation, disease, and aging. Epigenetics is the study of heritable changes other than those in the DNA sequence and encompasses two major modifications of DNA or chromatin: DNA methylation, the covalent modification of cytosine, and post-translational modification of histones including methylation, acetylation, phosphorylation, and sumoylation. An example of an early step in approaching the epigenome comes from a recent study by Fraga and coworkers (19). They used samples from monozygotic (MZ) twins, a tactic commonly used to quantify the environmental component of complex human diseases. MZ twins, despite being genetically identical, have been observed to frequently display phenotypic discordance for complex diseases. Fraga and colleagues determined that epigenetic disparities are present in the identical genetic background of MZ twins and that the number of differences increases

---

**TABLE 1. CHROMOSOMAL LOCATION, PHENOTYPES OF ASSOCIATION, REPLICATION, PROPOSED FUNCTION AND RELEVANT TISSUE AND CELL EXPRESSION FOR ASTHMA AND ATOPY GENES IDENTIFIED BY POSITIONALLY CLONING AND GENOME-WIDE ASSOCIATION**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Phenotype</th>
<th>Replication*</th>
<th>Possible Function</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positionally cloned genes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPP10</td>
<td>2q14</td>
<td>Asthma</td>
<td>11/2</td>
<td>Binding voltage-gated potassium channels</td>
<td>Epithelium asthma</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cleaving peptides from some chemokines</td>
<td></td>
</tr>
<tr>
<td>CYFIP2</td>
<td>5q33</td>
<td>Asthma</td>
<td>11/2</td>
<td>Th1/Th2 balance</td>
<td>Lymphocytes, undifferentiated</td>
</tr>
<tr>
<td>HLAG</td>
<td>6p21</td>
<td>Asthma</td>
<td>11/2</td>
<td>Modulation immune relationship mother and fetus</td>
<td>Epithelial cells PBMC</td>
</tr>
<tr>
<td>GPR4</td>
<td>7p14</td>
<td>Asthma</td>
<td>11/2</td>
<td>Smooth muscle contractility</td>
<td>B isomor stronger in epithelium and smooth muscle in asthma</td>
</tr>
<tr>
<td>SFRS8</td>
<td>12q24</td>
<td>Asthma</td>
<td>11/2</td>
<td>T cell adhesion</td>
<td>T cells</td>
</tr>
<tr>
<td>PHF11</td>
<td>13q14</td>
<td>Asthma</td>
<td>11/2</td>
<td>Th1-type cytokine gene expression</td>
<td>Lymphocytes</td>
</tr>
<tr>
<td>ADAM33</td>
<td>20p13</td>
<td>Asthma</td>
<td>11/2</td>
<td>Epithelial repair</td>
<td>Mesenchymal cells, Smooth muscle</td>
</tr>
<tr>
<td>GWA Genes</td>
<td></td>
<td></td>
<td></td>
<td>Many consistent replications of SNPs</td>
<td>Bronchial epithelium</td>
</tr>
<tr>
<td>CHI3L1</td>
<td>1q32</td>
<td>Asthma</td>
<td>11/2</td>
<td>Inflammatory processes</td>
<td>Epithelium reticulum</td>
</tr>
<tr>
<td>ORMDL</td>
<td>17q21</td>
<td>Asthma</td>
<td>11/2</td>
<td>Not yet known</td>
<td>Ubiquitous in endoplasmatic reticulum</td>
</tr>
</tbody>
</table>

Definition of abbreviations: AD = atopic dermatitis; BHR = bronchial hyperresponsiveness; IgE = immunoglobulin E; PBMC = peripheral blood mononuclear leukocytes; SNP = single-nucleotide polymorphism.
significantly with age (19). This study serves to highlight the important contribution of an individual’s epigenotype to the phenotypic manifestation of the inherited genotype. By comparing the epigenotype of discordant MZ twins in disease studies, researchers may eventually determine the source of many complex common diseases. An example of methylation of Cpg islands that may contribute to genetic variability between individuals has recently been described to be important for ADAM33 expression in epithelial cells in asthma (20). It has yet to be established whether the methylation status is a result of environmental and genetic factors, or by itself independently contributes to disease development (Figure 2). Interestingly, a maternal diet supplemented with methyl donors in C57BL/6J mice enhanced the severity of allergic airway disease in the offspring, and the trait was inherited transgenerationally (21). The 82 gene-associated loci that were differentially methylated with the diet of the mother were associated with decreased transcriptional activity and increased disease severity, showing that in utero dietary exposure may alter development of allergic phenotypes.

Overlapping Regions between Diseases

Of interest to a better understanding of the pathophysiology of various diseases is the observation that gene variants or regions in the genome have been implicated in multiple diseases (e.g., the 8q24 region in prostate, breast, and colorectal cancer [22–24] and the PTPN2 in gene in type 1 diabetes and Crohn’s disease [25]). In asthma research, we may thus learn from comparing GWAs in asthma and COPD, two different diseases with similar pulmonary symptoms and common clinical phenotypes like hyperresponsiveness and airway obstruction. Different factors may contribute to hyperresponsiveness in asthma and COPD, and comparing GWAs conditioning on this phenotype may help to unravel the different underlying pathophysiologies in asthma and COPD. Finally, meta-analysis of whole genome scans is perceived as the most reliable path to identification of genes in complex inheritance, and statistical tools to analyze these are under scrutiny and further development.

The Genetics of Gene Expression

Another approach to better identification of causal genes and consequent protein changes is genetics of gene expression (26). By comparing SNPs in genes and their expression in different tissues, one can dissect whether the SNP is causatively related to disease, and the protein expression is causative or the consequence of the SNP or possibly of other environmental determinants (e.g., smoking, or treatment like inhaled steroids). This approach requires a smaller number of samples than linkage and GWA studies, since the genetic effect of an SNP on gene expression is larger than on clinical phenotypes. This approach has been successfully applied to the identification of ORM DL in the first GWA study of asthma (16).

Different Phenotypes in Asthma

Though techniques to dissect the genetics of asthma have improved, proper phenotyping remains crucial given the heterogeneity of asthma and atopy. It may well be that by defining subtypes of asthma, like asthma in remission, progressive asthma, severe corticosteroid insensitive asthma, and asthma with systemic inflammation, we will have better insight into which genetic and environmental risk factors contribute to the course of asthma over a life span. Crucial to the unraveling of the disease is similar phenotyping of different cohorts when combining data to enhance the required sample sizes.

It has been known for a long time that asthma may remit in boys and develop in girls during puberty (27), implicating hormonal factors. However, since not all girls develop asthma, genetic variants may play a role. SNPs in ESRI, the estrogen receptor, have already been associated with lung function decline in females with asthma in particular (28). Thus, it is important to study girls and boys around puberty and measure hormonal changes in connection with other genetic and environmental risk factors for the development of asthma and its phenotypes.

Will it be possible in the future to predict asthma based on genetic testing? The predictive value of a single gene test in a complex disease is very limited for diagnostic or preventive purposes. In the future, asthma prediction may be possible, based on a prediction model that incorporates genes, personal factors, and environmental risk factors. Studies in general and at-risk populations are needed to investigate and validate this approach (29).

CONCLUSIONS

The current enormous improvements in the application of GWAs, genome-wide expression, microRNA studies, and new epigenetic insights will undoubtedly result in a substantial step forward in the next decade to understand the mechanisms underlying asthma. Collaboration between many groups following cohorts of children with asthma from birth till adulthood, including follow-up of measures of environmentally relevant factors, is mandatory. Huge numbers of participants in such studies are required to identify sets of genes and sets of environmental factors that together contribute to development, persistence, remission, and progression of asthma. This will eventually allow us to start preventive measures and therapies that are tailored to an individual’s needs.

Conflict of Interest Statement: Neither author has a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

References


25. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678.


29. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678.