Chapter 9

General discussion and future perspectives
General discussion and future perspectives

The title of this thesis is ‘Genetics of different asthma phenotypes’. We have performed several genetic studies on different phenotypes to further unravel the genetics of asthma. The knowledge of genetics in general has been growing rapidly in the past decennia and we foresee that many new genetic markers will be identified in the coming time.

The importance of disease phenotypes in genetic studies

In the introduction of this thesis, we emphasized the importance of the definition of a disease in complex traits. Generally, asthma in GWAS studies is being defined as a doctor diagnosis of asthma. This is a rather loose definition, but allows studying large numbers of asthmatics. Another approach is to take the heterogeneity of asthma into account. For instance, the severe Asthma Research Program\(^1\) already identified five clusters within asthma. It would thus be of interest to genetically study asthmatics within their own cluster. A major issue with this approach is that by restricting the phenotype definition, study groups will become much smaller, making it more difficult to gain sufficient power to obtain reliable answers to the role of genes in the development of a particular asthma cluster. In Chapter 2, we applied a stricter definition of asthma by adding the presence of BHR to the definition of asthma. The advantage of adding BHR to the definition of asthma is that BHR can be clinically and objectively tested. Patients do not always remember if a doctor has told them that they really had asthma, or ‘just’ asthma-like complaints, like cough or wheeze, for instance after a viral upper respiratory tract infection. This recall bias can lead to treating subjects in a GWAS study as a case (asthmatic) even though they never had asthma. Even when using the strict definition of asthma, including objective presence of BHR in the past, we had approximately 1000 asthmatics, quite a reasonable number for genetic studies. Generally, 80-90% of asthmatics express BHR, therefore we anticipated that we would replicate earlier findings in the literature on GWAS of asthma and would find some additional novel genes. Indeed we confirmed the relevance of the 17q locus (ORMDL3 gene region) for asthma plus BHR and identified three novel loci that had not been linked to asthma before. However, a limitation of our study could be that we assessed replication of our findings in other asthma cohorts that did not perform a challenge test. Although approximately 80-90% of asthmatics express BHR, it could be that SNPs specifically associated to BHR in asthma did not replicate because of lack of BHR testing in the replication populations.
When a SNP is associated with a disease, mechanisms behind the association are often unknown. Even the possible function of the gene may be unknown. In this case, a subphenotype analysis can help with identifying the function of the gene. In Chapter 2, a subphenotype analysis was performed on SNPs that were replicated in four asthma cohorts. The replicated SNPs were subsequently associated to several different phenotypes within asthma. All replicated SNPs could be linked to one or more subphenotypes, i.e. SNPs in the genes \textit{ABI3BP}, \textit{NAF1}, \textit{MICA}, and \textit{ORMDL3/17q21} region. In case of \textit{ORMDL3} this resulted in a positive association with CD4\textsuperscript{+} cell numbers and eosinophil counts in airway wall biopsies of asthmatics. This finding was in line with previous literature that showed that Ormdl3 promoted eosinophil trafficking to the sites of inflammation in mice.\textsuperscript{2} One of the other genes, \textit{NAF1}, had never been associated with asthma before, but since it is also associated with the subphenotype atopy (based on skin prick test), it could very well be that this gene plays a role especially in atopic asthma. A requirement for performing a subphenotype analysis is that one needs a well phenotyped cohort. Population based cohorts often use a loose definition of asthma, like doctor diagnosed asthma, but do not have additional measurements for asthma like lung function or allergic phenotypes, let be airway wall biopsies. In our cohort we had data available from airway wall biopsies of the asthmatics to which we could associate our SNPs. Furthermore, the availability of gene expression data in lung tissue in another cohort, gave us the opportunity to study eQTLs of the replicated SNPs. This resulted in the identification of 35 gene transcripts that we could enter in a pathway analysis. This way we could associate \textit{ABI3BP} with a pathway involved in the (positive) regulation of cell adhesion.

**Replication of GWA studies**

Genetic studies need replication of the findings in other cohorts to validate the results found. Replication of genetic studies is needed because chances of finding false positives become much bigger when a high number of tests are being performed. My concern with respect to replication is if it the correct approach to always replicate your results in different populations. Genetic studies are often performed on a specific group of patients with a specific trait. As emphasized previously, asthma consists of different subphenotypes, which could be associated to different SNPs and genes. What does that mean for replication? It could be that the more general genetic mechanisms for asthma can be replicated in several cohorts, but that more specific mechanisms that underlie a subphenotype of asthma are being missed. Therefore, specifically characterized cohorts are required for replication. An example is \textit{ORMDL3} in asthma. As was already discussed in the general introduction, the association with asthma was only found in childhood onset asthma and not in adult onset asthma.\textsuperscript{3} In case replication was sought for in a predominantly adult onset asthma cohort, this would not have led to any significant replication.
In Chapters 2 and 3, we found novel loci that were associated with asthma and the severity of BHR. These loci were theoretically plausible for their relationship with asthma, but did not provide genome wide significant results after replication. In some cohorts the replication was not even nominally significant. Does this mean that the loci are irrelevant to the trait? I do not think they are irrelevant, but rather that the loci found are specific to a subphenotype in the cohort under study. Therefore, the replication cohorts should match to the identification cohort. In case of \textit{ORMDL3} only cohorts with childhood onset asthma should be included and not with adult onset asthma. The same can be true for other study characteristics, like allergy, severity of BHR, lung function results, airway wall biopsies etc. However, sometimes replication cohorts with the specific phenotype under study are extremely difficult to find, which makes matching impossible.

\textbf{Power and phenotype definition in GWA studies}

It is often hard to gain sufficient power in genetic studies. Although cohorts are big compared to most clinical observational studies, the number of patients needed for a genetic study is high. In Chapter 2, we had a fair number of cases and controls to study, 920 asthmatics and 980 controls, but in Chapters 3, 4 and 7, the number of asthmatics included in the studies, respectively 650, 415 and 790, is reasonable but not very high. There are several factors that can contribute to the size of a GWA study. Not only the amount of money and time that is needed for performing the study and genotyping of the patients is important, but also the access to patients. If a rare phenotype is taken, like for example complete remission in Chapter 7 that only occurred in 7\% of our total cohort, it is hard to include a sufficient number of patients with this phenotype. Hence, the power to detect genes associated with this rare phenotype becomes low. Asthma remission can be defined in several ways.\textsuperscript{4} When the phenotype is rare, like in complete remission of asthma, it is of importance to choose a strict definition of asthma remission. Again, the trade of is more subjects and less precise phenotyping, which could lead to more heterogeneity in the population under study.\textsuperscript{5} Chapter 7 is a good example of this strategy. Compared to the literature, we choose a strict definition of asthma and asthma remission that was clinically tested with a bronchial challenge test and spirometry in childhood and adulthood. Despite the low power we were able to replicate our findings on complete remission in other cohorts, and, of importance, they used the same asthma and asthma remission definition, i.e. the absence of asthma symptoms (no wheezing without having a cold in the last year and no asthma attacks in the last 3 years), no use of asthma medication in the last year, and no presence of BHR with a good lung function (\textit{FEV}_1 > 80\% of predicted).

\textbf{Design of genetic studies}

Nowadays, genetic studies are often based on more SNPs and a combination of several genetic markers. With prices of whole genome sequencing going down, it becomes even feasible in the near future to sequence the whole genome instead of using a chip. This way, rare variants which could be responsible for rare phenotypes, like complete remission of asthma, could be identified.
In most complex diseases a combination of different genomic markers plays a role in disease development. Thus, not one marker is responsible for the disease, but several. The identified SNPs are then investigated for their functional relevance, which can be approached by searching for their association with gene expression using eQTL analyses. However, several other genomic factors can play a role like for example epigenetic factors such as methylation, microRNA or histone modification. These markers often have a strong association with environmental exposures of the patient. Therefore, it is of importance to include the markers in studies instead of using data from another study.

If a genetic underlying disease mechanism is proposed after performing genetic studies, functional studies are needed to confirm the mechanism. If this mechanism is based on a specific allele, functional tests should be performed on tissues with the same genotype. If the functional study is performed on tissue with a different genotype, it is well possible that the proposed mechanism is not found, because the risk variant of the genotype is not available.

One way of using the correct genotype is by using tissue of the patients under study for genetic variants. In the study design of the Roorda cohort, culturing of epithelial cells of nasal epithelium was included for further functional studies. The collection of nasal brushes is a non-invasive way of obtaining epithelial cells from patients which could be used for functional testing. This functional testing could eventually lead to the confirmation of mechanisms associated to the trait, in this case asthma remission. If specific mechanisms are altered by one specific gene, it opens avenues for pharmaceutical companies to create drugs to treat asthmatic patients resulting into remission of asthma. In case of the Roorda cohort, we are still analysing all genetic data that is available and have not come to any functional testing yet. We envisage that novel techniques to immortalize cells, for example using organoid cell cultures, could lead to patient and genotype specific tools to test new drugs.

Highly correlated to this subject is the personalised medicine. It can be foreseen that within a decade or longer, it may be possible to prescribe drugs to an individual person based on his/her genotype. In chapter 4 we showed the interaction of ICS-by-SNP in the severity of BHR. This resulted in the identification of 6 loci with possible involvement in the interaction of SNP and ICS use. Minor alleles of 3 SNPs associated with a worsening of BHR severity when using ICS and minor alleles of 3 SNPs associated with an improvement of BHR severity when using ICS. One would not have prescribed ICS in the knowledge that that particular patient was losing control of asthma instead of stabilising. By identifying SNPs that alter the effect of drugs one thus could prescribe specific drugs based on genotype. If we want to be able prescribe drugs based on phenotype and genotype of a patient, large cohorts with discrete data are needed. In some diseases, like for example inflammatory bowel disease (IBD) and diabetes, this information is stored in electronic medical records. Patients who are being treated in the hospital for IBD or diabetes have signed consent to use their data for research.
By including all patients treated in the hospital, it becomes easier to include more patients in a study. This gives the opportunity to study interaction between the genome and different treatments. The next step would be to ask consent of all patients attending the UMCG, and not only for specific diseases, to using their data for research purposes. One necessary step further would be to also add their DNA to the database, thereby allowing studying SNP-drug interactions.

**Predictors of asthma remission**

It is hard to predict the outcome of asthma in childhood. Sometimes asthma is really driven by an allergen that can be avoided. This avoidance could result in less or no symptoms of asthma, which could lead to asthma remission. Whether or not this is clinical or complete remission remains to be investigated. Other phenotypes of asthma are not ‘cured’ by avoiding allergens, but need treatment. How can we predict the outcome of asthma over time? Fortunately, many childhood data were available in the Roorda cohort. Asthma remission predictors have been studied before, in complete as well as clinical remission; among other factors a good lung function and a low total IgE level are important contributors to remission. Since medical records were available for all patients in the Roorda cohort, we could run analyses on several potential predictors that were previously not investigated. We confirmed that a good lung function was associated with complete remission and could add new factors like the presence of leukaemia in the family and a positive skin prick test to mould. Especially the presence of leukaemia in the family is a striking finding. We identified two studies which linked atopic disease to acute lymphocytic leukaemia. This suggests that atopy may protect against developing leukaemia. These data were probably never studied, because asthmatics were never asked this question. Interestingly, our group studied if the protective role of atopic diseases in leukaemia was associated with innate immunity genes that were previously associated to atopy diseases. They identified two SNP, rs5743798 and rs6531666, located in TLR6 with a different genotype distribution in the two diseases. In children with an atopic disease, especially children with atopic eczema, the risk alleles were more often observed as compared to children with leukaemia. It could be of interest to study shared mechanisms underlying both diseases, by studying patients with both diseases.

Like in genetic studies, it is important to have a big dataset when analysing prediction factors for a disease. The same approach, as described before can be taken to include more patients in a cohort. If all medical record are stored discretely and patients sign consent to use their data, retrospective studies can be performed to identify factors involved in the prediction of disease outcome.
References


