Biomarker discovery for asthma phenotyping: From gene expression to the clinic
Wagener, A.H.

Citation for published version (APA):
Wagener, A. H. (2016). Biomarker discovery for asthma phenotyping: From gene expression to the clinic.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 7

Summary and General Discussion
SUMMARY

Background
Asthma is considered a complex respiratory disease of which various asthma phenotypes are being discovered. Clinical biomarkers have shown to be successful in the management of asthma phenotypes. However, to increase the understanding of this complex disease and discover new biomarkers, will require knowledge of the molecular mechanisms involved. High-throughput omics technologies are now available, such as transcriptomics, proteomics, lipidomics, and breathomics. Using these methods will allow further understanding of complex diseases such as asthma, and will potentially offer biomarker discovery (1).

In this thesis, I have reviewed the strengths and limitations of both transcriptomics analysis and breathomics analysis, and I have applied both technologies to increase understanding and discover surrogate biomarkers. Furthermore I have validated previously examined biomarkers.

Conclusions of the studies

Towards composite molecular signatures in the phenotyping of asthma
In chapter 2 I discussed the strengths and limitations of transcriptomics by microarrays and next-generation RNA sequencing. Next, I reviewed metabolomics in exhaled air (breathomics) as a non-invasive tool for the clinic.

Main conclusions:
• For signature discovery, transcriptomics analysis using microarrays is at present the-state-of-the-art method, because of available data for comparisons and validation, and its maturity of experimental design and analysis.
• When experience has increased and costs have decreased, RNA sequencing will become preferable since it allows an unbiased analysis and enables higher dynamic detection ranges.
• Breathomics has diagnostic potential as a non-invasive method, though sampling methods and devices still need to be standardised to enable validation.

Implications:
Composite molecular fingerprints have the best prospect as biomarkers in the phenotyping of patients with complex respiratory diseases such as asthma.
The impact of allergic rhinitis and asthma on human nasal and bronchial epithelial gene expression

In chapter 3 I studied the link between the upper and lower airways by analysing gene expression profiles of upper and lower airway epithelial cells in healthy individuals and examining the impact of allergic rhinitis and asthma on these expression profiles.

Main conclusions:

• There were substantial differences in gene expression between the upper and lower airway epithelium of healthy individuals but many of these differences disappeared in patients with allergic rhinitis with or without asthma.
• Genes that were influenced by allergic rhinitis and asthma were related to lung development, remodelling, regulation of peptidases and normal epithelial barrier function.
• Our unbiased approach identified genes, such as UDP-glucuronosyltransferase genes and RUNX2 that have not been previously described in relation to allergy, rhinitis or asthma.
• Allergic rhinitis affected the epithelial gene expression in both the upper and lower airway epithelium.

Implications:

Differences in epithelial gene expression between upper and lower airway epithelial cells of healthy individuals largely disappeared in patients with allergic rhinitis with or without asthma, with a main impact of allergic rhinitis. Several new genes and pathways were identified that might be potential targets for future drug development.

dsRNA-induced changes in gene expression profiles of primary nasal and bronchial epithelial cells from patients with asthma, rhinitis and controls

In chapter 4 I examined the responses of airway epithelium to double-stranded RNA (dsRNA) as a model of viral induced exacerbations. This was done by comparing dsRNA-induced gene expression profiles of primary nasal and bronchial epithelial cells and observing modulation of these expression profiles by the presence of allergic rhinitis and asthma.

Main conclusions:

• The dsRNA-induced transcriptional response was characterized by a strong induction of genes involved in the response to virus, apoptotic processes and antigen presentation.
• The airway epithelium of patients with asthma demonstrated significantly fewer induced genes, in particular with regard to impaired interferon expression and reduced down-regulation of mitochondrial genes.
• Several disease-specific genes were identified that are induced in patients with allergic rhinitis with or without asthma but not in healthy controls.

Implications:
The viral-induced differences in gene expression between upper and lower airways improved the understanding of mechanistic pathways of the mutual interaction between asthma and rhinitis. Furthermore, potential targets for drug-discovery studies were identified, related to mitochondrial dysfunction and interferon signalling.

External validation of blood eosinophils, \( FE_{NO} \) and serum periostin as surrogates for sputum eosinophils in asthma
In chapter 5 the mutual relationship between blood eosinophils, exhaled nitric oxide (\( FE_{NO} \)) and serum periostin with sputum eosinophils was quantified by external validation in two independent cohorts of patients with mild to severe asthma.

Main conclusions:
• Blood eosinophil count was an accurate surrogate marker for sputum eosinophils in patients with mild to moderate asthma and this relationship was replicated in patients with severe asthma.
• Serum periostin was not able to differentiate between eosinophilic and non-eosinophilic airway inflammation.
• When combining the three markers, neither \( FE_{NO} \) nor periostin showed any improvement to the diagnostic value of blood eosinophils.

Implications:
Blood eosinophil cell count represented an accurate biomarker for eosinophilic airway inflammation which can facilitate guidance of current and novel individualised asthma treatment.

Predicting eosinophilic airway inflammation in asthma using exhaled breath profiling
In chapter 6 the relationship of breathprints analysed by a composite electronic nose (eNose) platform with sputum eosinophils was validated in patients with mild to severe asthma.
Main conclusions:
• The eNose platform was able to differentiate between eosinophilic and non-eosinophilic airway inflammation in asthma, which is validated at a second visit.
• Blood eosinophil cell counts showed similar discriminant accuracies.

Implications:
ENoses have potential to assess eosinophilic airway inflammation in patients with asthma in a quick and non-invasive way, thereby potentially facilitating personalized asthma management.

GENERAL DISCUSSION

Transcriptomics

Transcriptomics analysis using microarrays
In chapters 3 and 4 I used transcriptomics analysis by microarrays to explore gene expression in airway epithelial cells with an unbiased approach. Unlike gene sequencing though, microarrays are limited to known RNA transcripts and therefore these analyses are not completely unbiased (2), as is discussed in chapter 2. Still, expression of more than 33,000 well-characterized RNA probes can be analysed by microarrays, which was shown to be promising in the biomarker discovery in pulmonary diseases.

Transcriptomics analysis in asthma
Recently, distinct phenotypes defined by Th2-high or Th2-low inflammation were identified by gene expression profiling of bronchial epithelial cells from patients with mild to moderate asthma using microarrays (3). Following these results, patients with severe asthma appeared to benefit from anti-interleukin-13 treatment if serum periostin levels were high, which corresponds to the Th2-high phenotype (4). Subsequently, anti-interleukin-4 receptor treatment seemed successful in patients with moderate to severe asthma with increased sputum or blood eosinophils (5), which appeared to identify patients with Th2-high asthma (6). Furthermore, induced sputum gene expression profiles identified three phenotypes related to both clinical asthma status and airway inflammation (7). This showed the potential of this technology to have additive value in clinical phenotyping. Additionally, analysing gene expression improved our understanding of disease mechanisms underlying the associations between airway inflammation and systemic inflammation. Gene profiling of induced sputum from asthmatics with systemic inflammation showed upregulation of signalling pathways involved in particularly neutrophilic inflammation (8). Also, mechanisms of treatment resistance in
asthma were studied by gene profiling bronchoalveolar lavage cells of patients with corticosteroid-resistant asthma and corticosteroid-sensitive asthma. This revealed involvement of endotoxin (LPS) and classical macrophage activation in corticosteroid resistance (9). Taken together, transcriptomics analysis in pulmonary disease was shown to be promising in biomarker discovery and in improving our understanding of asthma.

Sample method
In chapters 3 and 4 I explored gene expression profiles of both upper and lower airway epithelial cells from subjects with allergic asthma, allergic rhinitis, and healthy controls. I chose to isolate and culture epithelial cells only, without the influences of other cell types. Of course, extracted and cultured cells will not exactly represent the same conditions in tissue. Alternative procedures to obtain epithelial cells such as direct measurement after isolation or laser capture might mitigate these effects of cell culturing, but could introduce new biases introduced by contamination by other cell types or the isolation procedure itself. Air-liquid interface induces differentiation into a pseudostratified mucociliary epithelium resembling the in vivo appearance of the airway epithelium. Nevertheless, this still cannot ensure that the in vivo expression profile is fully preserved after differentiation ex vivo. Other studies used whole biopsies to study airway gene expression. However, by using entire biopsies it is not clear how the different cell types contribute to the gene expression since every biopsy contains different types and numbers of cells. Hence, every method has its strengths and limitations, and there is no consensus on the most desirable sample method. Standardisation of these method procedures would improve the ability to compare and validate results.

Transcriptomics analysis in this thesis
In both studies the unbiased analysis identified several new genes that were influenced by allergic rhinitis and asthma, and also confirmed the role of previously described genes. Since microarrays produce such large datasets I used different techniques in both studies to enable interpretation. In chapter 3 I used K-means clustering to reduce the data into subgroups and to be able to differentiate patterns of gene expression between the different cohorts. In this way we were able to see the impact of allergic disease, in particular of allergic rhinitis, on the gene expression differences between upper and lower airways. Furthermore, I made use of overrepresentation of gene ontology groups. A regulation interaction network discovery was performed to study the function and relationships between differential expressed genes. In chapter 4 I mainly focused on the overrepresentation of gene ontology groups because of so many differentially expressed genes after dsRNA-stimulation.

In chapter 3 one of the conclusions was that in a healthy state considerable differences exist in gene expression between nose and bronchus. However, in case of allergic rhinitis
many of these differences seemed to disappear, especially in case of allergic asthma with concomitant rhinitis. Using K-means clustering I demonstrated that gene expression in the lower airways of patients with allergic rhinitis was altered as well, which suggested a major impact of allergic disease. This also suggested interaction between the upper and lower airways in patients with allergic asthma and/or allergic rhinitis (10). Previously, gene expression of allergic nasal epithelial cells in response to house dust mite showed smaller changes as compared to healthy epithelial cells, which was explained by an activated state of allergic nasal epithelium before stimulation (11). Therefore, the diminished differences between gene expression of upper and lower airway epithelium in allergic disease could be explained by this activated state of genes in both upper and lower airways, and thereby reducing differences.

In chapter 4 a considerable loss in dsRNA-induced down-regulation of mitochondrial genes was observed in both upper and lower airways of patients with allergic asthma with concomitant rhinitis. However, I did not see this effect in patients with allergic rhinitis, suggesting changed host characteristics of the upper airways in patients with allergic rhinitis plus asthma as compared to those with allergic rhinitis alone. Again, this supported mutual interaction between upper and lower airways of patients with asthma and rhinitis. Furthermore, a considerable proportion of the most highly up-regulated genes in all three groups were interferon-related genes. Nevertheless, interferon-β1 and interferon-λ3 were induced in all subjects except for patients with asthma. This is in line with previous studies reporting reduced interferon-β and interferon-λ in primary bronchial epithelial cells from those patients with asthma following rhinovirus-infection (12;13). However, in a very recent report the same group failed to find any evidence for deficient type I or III interferon induction in rhinovirus-infected primary bronchial epithelial cells from patients with asthma (14). They speculated on the relation with asthma severity that these patients were too mild for a difference to be detected. However, in our study patients with asthma had even milder disease when comparing medication usage. The reason for this contrast in data is unknown. Recently, a study showed impaired interferon induction by bronchial epithelial cells following rhinovirus infection if these cells were pretreated with interleukin-4 (IL-4) and interleukin-13 (IL-13) prior to infection, suggesting that enhanced Th2 inflammation can dampen the antiviral response (15). This shows the need for further studies to validate and improve the understanding of the complex interactions in these signalling pathways.

**Biased surrogate markers for sputum eosinophils**

Asthma phenotyping started a long time ago using a biased approach, identifying allergic and nonallergic asthma (16). Subsequently, based on sputum differential cell counts, inflammatory subtypes were identified (17) with clinical consequences (18). Since sputum eosinophil count is considered an important clinical biomarker, many
studies have tried to find surrogate markers because sputum analysis is hindered by the duration of the analysis and the possible failure of producing an adequate sample. Furthermore, in patients with severe and uncontrolled asthma, who are exactly those of which the inflammatory profile is of interest for treatment adjustments, the procedure can cause unwanted hypertonic saline-induced airway narrowing (19). Several surrogate markers have been considered, such as FE\textsubscript{INO}, blood eosinophils, and serum periostin. However, blood eosinophils and FE\textsubscript{INO} have shown varying correlations with sputum eosinophils (20-24). FE\textsubscript{INO} demonstrated insufficient capacity to monitor steroid therapy (18), although this was challenged by a recent positive study in primary care and a different study using an improved algorithm to successfully reduce exacerbations by FE\textsubscript{INO}-guided therapy (25;26). In chapter 5 I concluded that blood eosinophil count is an adequate surrogate marker for sputum eosinophils in mild to moderate asthma, and these findings were replicated in a cohort with more severe asthma. However, a very recent meta-analysis concluded that blood eosinophil count has only moderate diagnostic accuracy as a single surrogate marker for airway eosinophilia (27). But perhaps more importantly, two recent trials using anti-IL-5 (mepolizumab) to target eosinophilic airway inflammation in patients with severe asthma, defined by blood eosinophils, resulted in a significant reduction in the daily requirement of oral glucocorticoids, reduced exacerbations and improved asthma symptoms (28;29). In a different large study using anti-IL-5, blood eosinophil count predicted the efficacy of reducing exacerbations in patients with severe eosinophilic asthma (30). These trials support the potential role of blood eosinophil count as a non-invasive marker to predict responsiveness to novel personalised therapies. Obviously, finding the best biomarkers to identify individual patients most appropriate for a targeted therapy is the ultimate goal, instead of determining the best surrogate marker for sputum eosinophilia. In other words, a moderate surrogate marker for sputum eosinophils might perform better as biomarker itself to predict treatment responsiveness.

Finally, serum periostin has showed high diagnostic accuracy to identify sputum eosinophils (31) but these results had not been replicated yet. Therefore, chapter 5 was the first to validate serum periostin as biomarker for eosinophilic airway inflammation, a relationship that we did not find in the mild to moderate cohort nor in the more severe patients with asthma. Periostin might be more related to IL-13 instead of sputum eosinophils, because the protein is partly regulated by IL-13 (32) and was shown to be predictive in the treatment response to anti-IL13 medication (4). As IL-13 is a Th2 cytokine, it is not surprising that periostin will correlate with sputum eosinophils in some patients. Taken together, it seems that variation exist within patients with a Th2-phenotype or high sputum eosinophils regarding levels of IL-13 and periostin. Therefore it seems that periostin may have complementary information to sputum or blood eosinophils in identifying patients with specific profiles based on a Th2-high cytokine profile.
Breathomics

Breathomics analysis using eNoses

In chapter 6 I have used eNoses to measure metabolites in the exhaled air that are potential non-invasive biomarkers of disease. These metabolites or volatile organic compounds (VOCs) originate from both local and systemic metabolic processes. The gold-standard for exhaled breath analysis is gas chromatography-mass spectrometry (GC-MS), which identifies individual molecular compounds by determining the composition and the concentration of the component (33). Unlike GC-MS, eNoses analyse VOCs using pattern recognition by arrays of nanosensors that capture various combinations of VOCs (34). Before implementation of eNoses in clinical practice can be accomplished, the pattern recognition algorithms require training and staged validation. There are different types of VOC-sensors, such as organic polymers, quartz crystals, metaloxide, and ion mobility spectrometry. The choice of sensors will influence the suitability of the eNose in the phenotyping or diagnosis of a specific disease. However, it is still unknown what type or combination of sensors is best in detecting which disease. Therefore, a composite eNose platform, used in chapter 6, was developed for the U-BIOPRED study group that integrates different types of eNose sensors to measure breathprints in parallel. Metal oxide semiconductor sensors and the field asymmetric ion mobility spectrometer were the most discriminative sensors to assess eosinophilic airway inflammation. To improve the discrimination performance of the eNose, a tailor made eNose should be developed to detect those VOCs that are specifically associated with eosinophilic airway inflammation, or eventually with treatment responsiveness. The eNose has already showed potential in predicting corticosteroid responsiveness in a small cohort of patients with asthma (35).

Breathomics analysis in asthma

Breathomics research in pulmonary diseases is rapidly expanding, especially concerning lung cancer, infectious and airway diseases (36). ENose studies in asthma have shown that eNoses can discriminate patients with asthma from healthy controls (37), and from those with COPD (38). Moreover, a subsequent external validation study showed that patients with fixed asthma and COPD can be discriminated using exhaled breath pattern recognition based on the previous training set (39). These previous studies mainly focus on disease diagnosis, though the results of chapter 6 indicate additive value of the eNose in asthma phenotyping. This is supported by GC-MS studies that were able to identify eosinophilic inflammation in asthma (40) and COPD (41;42), whilst a recent study showed potential in asthma clustering (43). The eNose platform was able to discriminate eosinophilic from non-eosinophilic asthma with an ROC area under the curve (AUC) of 73%, and validation of the model in the longitudinal cohort resulted in an
AUC of 78%. Blood eosinophils showed nearly similar discriminant values. Interestingly, when adding blood eosinophils to the predictive model of the eNose the AUC increased slightly to 83%. This difference in AUC is probably not significantly different but may suggest that VOCs, small molecules, and systemic markers all reflect different aspects of eosinophilic inflammatory pathways. This is in line with a recent study in which $\text{FE}_{\text{NO}}$ and blood eosinophils were independently associated with wheeze and asthma events (44). Therefore, composite marker signatures might perform better than single biomarkers in the phenotyping of asthma. A recent study identified the eNose as predictor for steroid responsiveness in asthma, being more accurate than sputum eosinophils or $\text{FE}_{\text{NO}}$ (35). Unfortunately these authors did not show the results of adding sputum eosinophils or $\text{FE}_{\text{NO}}$ to the predictive model of the eNose, which could have improved the predictive value. On the other hand, the eNose data themselves represent a composite molecular signature which supports the concept that composite biomarker fingerprints are more powerful than single biomarkers.

**Validation**

Regardless of the technique chosen and whether the aim is to find a biomarker or composite marker signature to diagnose or phenotype disease, validation of tools and results is required based on international guidelines on STAndards for the Reporting of Diagnostic accuracy studies (STARD) (45) and the very recent published Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) (46). First of all, internal validity of a test or new application should be determined and possible introduction of bias should be considered, after which external validation is necessary.

When we focus on the microarray analysis in chapter 3 and chapter 4, expression values can be validated using RT-qPCR. In this thesis, nine differentially expressed genes were identified for RT-qPCR validation in both chapters since these genes represented a complete range of fold change values. Since there were differences in sample sizes between the groups and in distribution of age and gender, I re-analysed the data randomly excluding one subject from each of the other groups and tested for associations with gender and age to exclude bias. Next, internationally accepted criteria were used for the diagnosis of allergic rhinitis and asthma. Unfortunately, further internal validation tests such as cross validation were impossible since the cohort was too small to split in training and validation sets. The ultimate test now will be external validation in large cohorts, such as U-BIOPRED (Unbiased BIOmarkers in PREDiction of respiratory disease outcomes) (www.ubiopred.eu).

In chapters 5 and 6 I tested surrogate markers for sputum eosinophils. In order to evaluate diagnostic accuracy the selection of subjects in whom the markers are being tested is crucial. Current guidelines recommend using the golden standard to discriminate
between patients with the targeted condition and those without it (45). Therefore the cohorts included in chapters 5 and 6 were well-characterised and stringent criteria for the diagnosis of asthma were used.

In chapter 5 blood eosinophils, $F_{E_{NO}}$ and serum periostin were externally validated in a large cohort of mild to moderate asthma and this was replicated in a smaller cohort of severe asthma. Blood eosinophils and $F_{E_{NO}}$ were measured using devices that meet the recommended technical specifications for measurements according to the STARD guidelines, whereas serum periostin and sputum eosinophils were measured by assays that do not meet this standard. Such a technically recommended assay for sputum eosinophils does not exist, and the quality of processing is especially dependent on the experiences of the technician. The processing of sputum and cell counts for the two cohorts from chapter 5 was done by the same experienced technicians. Moreover the correlation between sputum eosinophils and blood eosinophils was consistent and the fact that we were able to reproduce our results in a second independent cohort represents a major validation. Since the size of the cohort of patients with severe asthma was limited, further analysis in large multicentre severe asthma cohorts is required. However, the pitfall of such multicentre studies is bias because of potential differences in qualities and experience levels of technicians between centres. Multicenter studies therefore demand adequate training in advance. Finally, to measure serum periostin we used two different assays that showed similar results, one of which was used in the single previous study that demonstrated the highest AUC for serum periostin to diagnose airway eosinophilia (31).

The results presented in chapter 6 show the discriminative ability of the eNose platform to distinguish patients with eosinophilic airway inflammation from patients with non-eosinophilic airway inflammation. Previous studies using GC-MS breath analysis have already showed that exhaled VOCs can identify eosinophilic airway inflammation (40-42) which could be regarded as internal validation. Next, a largely independent second cohort was used to externally validate the retrieved model.

Taken together, I used several ways or techniques in order to improve the quality of results presented in this thesis and potential weaknesses were discussed in every chapter separately.

**Future opportunities for research**

This thesis has addressed several omics technologies to allow further understanding asthma as a complex disease, and to potentially offer biomarker discovery. During the discussion of this thesis I addressed the requirement of future studies for external validation in larger multicentre cohorts. A recent user-friendly, open-access microarray repository was constructed relevant to allergic airway inflammation (47). This database includes the microarray results from chapters 3 and 4, which will increase the use of
these large datasets and will allow validation of genes and further experiments. To increase the quality of transcriptomics analysis, we urged standardisation of techniques and further validation and development of next-generation RNA sequencing, as was discussed in chapter 2. Furthermore, I addressed the potential use of blood eosinophil count and composite molecular signatures to identify patients with asthma that are responsive to novel or existing therapies. First of all, large severe asthma cohorts, such as U-BIOPRED will be required to distinguish asthma profiles with targets per profile for novel treatments. Next, easy-to-measure biomarkers, such as blood eosinophil count and exhaled VOCs need to be discovered that rapidly identify every asthma profile that is sensitive to a certain therapy.
REFERENCES


(2) Malone JH, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol 2011;9:34.


CHAPTER 8
Nederlandse Samenvatting