Asthma Metabolomics: The Missing Step for Translating Bench Work into the Clinic

Amber Dahlin*, Michael J. McGeachie# and Jessica A. Lasky-Su*
Department of Medicine, Channing Division of Network Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA
*These authors contributed equally to this manuscript.

Abstract
Metabolite profiling, the systematic analysis of all metabolites, has been used successfully to identify new biomarkers for several complex diseases. In this review we describe how metabolomics data are generated, review the existing literature using metabolomics with asthma and related phenotypes, and discuss the need for more comprehensive asthma metabolomics research.

Keywords: Asthma; Metabolomics; Genetics; Mass spectrometry; Pharmacometabolomics

Introduction
Asthma is a widespread disease with over 23 million diagnosed cases in the United States, and is the most common health-related cause of lost school and work days [1], constituting a large public health concern. Patient genetics is a major driver in asthma etiology [2,3] and although multiple molecular determinants have been identified [4], the biological mechanisms by which these determinants impact disease is not known. Furthermore, it is likely that diverse genetic profiles among patients confer different asthma treatment responses. The metabolome constitutes the entire complement of low molecular weight molecules in a biological sample from an individual, where the sample may include diverse biospecimens such as plasma [5], serum [6-8], saliva [9], exhaled breath condensate [10,11] and urine [12]. Metabolomics, defined as the systematic analysis of all metabolites, provides an opportunity to connect genetic and molecular mechanisms with disease outcomes, as fluctuations in metabolite concentrations are likely to directly contribute to asthma pathology. Therefore, metabolomics offers a unique opportunity to link molecular determinants with asthma diagnoses and related outcomes. To date, metabolomics studies have been limited in size and scope for asthma [13,14], making the metabolome an untapped resource with the potential to transform the current understanding of asthma pathogenesis, and to more effectively personalize treatment approaches. In this review, we review key concepts in the generation and analysis of metabolomics data, summarize the current state of asthma metabolomics, and explain the potential for asthma pharmacometabolomics, and, finally, advocate for the increasing need for asthma metabolomics studies.

Metabolomics: A Missing but Crucial ‘omic’ to Identify Asthma Disease Pathways

Current technologies enable the assessment of a large number of metabolites that result from environmental, genomic, transcriptomic, and proteomic variability. As such, metabolomics data provides the most integrated profile of biological status reflect the complex interplay of genetic and environmental interactions [15,16]. Metabolomic data therefore are amenable to study disease predisposition, diagnosis, and progression. Endogenous metabolites span a variety of compound classes, with significant differences in size and polarity, across wide concentration ranges. Mass spectrometry (MS) coupled with separation techniques including liquid chromatography (LC) is currently the most advanced technology available [17,18]. It can be used both in a non-targeted, pattern-recognition manner, or a targeted manner, for confirmation. In most complex diseases, like asthma, the perturbations involve activation of multiple pathways. By using clinical, environmental, genetic, and genomic data in conjunction with descriptive metabolic profiles obtained by MS, it is possible to describe patterns of changes and biomarkers that discriminate between states of asthma severity and asthmatic cases and controls [19,20]. Investigators have already successfully identified biomarkers in type 2 diabetes, Alzheimer’s disease, and cardiovascular disease [18,21-23] that have led to the discovery of novel disease pathways. In contrast to those diseases, there has been limited work in asthma metabolomics to date.

Rapid technological advances have enabled the generation of vast ‘omics’ datasets comprising millions of biological measurements on genomic, transcriptomic, proteomic, and metabolomic data. The emerging field of systems biology uses these data to understand the complex interactions that result in the development of a disease. For a complex disease such as asthma, significant advances have been made in identifying genetic determinants of the disease [24-26]. While earlier linkage and candidate gene studies identified a small number of genetic variants for asthma [27-37], genome-wide association studies (GWAS) have confirmed a number of asthma loci including ORMDL3/GSDMB, HLA-DQ1, IL1RL1, and IL33, among others [4,38-42]. Despite the identification of these genetic contributors, little is known regarding how these variants impact asthma [43]. Furthermore, both genetic heterogeneity and host environment affect the diversity of biological pathways underlying diverse asthma phenotypes. Therefore, identifying the mechanisms by which genomic variations interact with the environment and lead to perturbations in biological pathways resulting in a disease state is crucial for both understanding and treating asthma.

Metabolomics, the systematic analysis of all metabolites (including sugars, amino acids, organic acids, nucleotides, and lipids), offers a snapshot of the integrated profile of a biological state. Compared with proteomic or genomic profiling, comprehensive metabolite profiling represents more proximate measures of underlying genetic and environmental exposures contributing to disease. Because the perturbations in biological pathways that occur with disease states
produce fluctuations in metabolite levels, metabolomic profiles can be more directly linked to a disease outcome, such as asthma [15,16]. Therefore, a hypothetical complete metabolomic profile would provide the most informative assessment of biological status, reflecting the net effect of genetic and environmental interactions [15,16], thereby making this approach a promising strategy to examine asthma pathogenesis [44].

The Generation and Analysis of Metabolomic Data

For metabolomic profiling, the metabolites are first extracted from a biospecimen using separation and quantification techniques. The analytical tools most commonly used in metabolomics include nuclear magnetic resonance spectroscopy (NMR), liquid and gas chromatography (LC, GC), and mass spectrometry (MS). A comprehensive measurement of the metabolome continues to be an analytical challenge because of the differences in physical properties among compounds that constitute the metabolome. For example, differences in polarity among metabolites require different extraction procedures for the preparation of analytical samples. As a result, multiple procedures must be combined in order to generate comprehensive measures of the existing metabolites. For metabolite quantitation, two main strategies exist: targeted and untargeted approaches. Targeted metabolomics approaches are applicable for hypothesis-driven studies, where a predefined set of select metabolites is quantified. In contrast, untargeted metabolomics approaches measure all endogenous metabolic signals in a biological sample, which results in a larger range of metabolites, but has reduced sensitivity to detect individual metabolites, or identify unknown metabolites. Here, we summarize commonly sampled biofluids, and routine analytical methods for analyzing metabolomics data.

Biofluids and sampling methods

Metabolomics data can be generated using a vast array of biologic specimens. While blood plasma and serum [8,17] are most common, metabolomics data is also being generated using urine [14], stool [45], exhaled breath condensate [11], and bronchoalveolar lavage (BAL) fluid [46].

Common analytic methods for metabolomics data

Several common metabolomics methods have been described in the literature that can also be applied to asthma metabolomics data. As with other 'omics' studies, in metabolomics, there are often many more variables to measure than there are samples available. To circumvent this issue, and to identify metabolites that are predictors of asthma or related outcomes, multivariate and machine-learning techniques can be applied. Standard parametric statistical methods such as regression analysis are inapplicable when the number of samples is far less than the number of parameters, making data-reducing methods such as principal component analysis [47], discriminant analysis, Gaussian graphical models, and Bayesian networks, attractive options. We review these common approaches below.

Principal component analysis (PCA) is one of the most common techniques used for analyzing metabolomics data [48]. The popularity of PCA in metabolomics is due to the fact that it is a simple, non-parametric method that can project the metabolites into lower-dimensional space, revealing inherent data structure, and greatly reducing the number of variables in the original data. In this case, each principal component is a summary of metabolite contributions that are orthogonal from the other principal components and can be used as independent variables; together these variables summarize the overall metabolic state of the individual samples. This is an effective way to summarize the contribution of many correlated metabolites to a phenotype; however, the principal components themselves are often difficult to interpret biologically.

Partial least square discriminant analysis (PLS-DA) is similar in some respects to PCA in that both reduce the dimensions of the analysis by collapsing correlated variables into fewer, uncorrelated variables. PLS-DA first transforms the metabolites into uncorrelated variables. These uncorrelated variables are then included in a regression analysis with the case/control status as the dependent variable of interest, and the set of metabolites that most accurately predicts the case/control status is selected. This approach identifies a metabolic signature for the condition in question. PLS-DA has been used to discriminate patients with and without various diseases, including coronary heart disease [49], cardiovascular disease [50], schizophrenia [7], inflammatory bowel disease [51], diabetes [21,52], low birth weight [53], ovarian cancer [54,55], and multiple sclerosis [56].

For broad-based, untargeted metabolomic profiling approaches, reconstructing metabolic pathways from the observed data is an attractive endeavor that also facilitates the identification of pathological disruptions or phenotypic differences between cases and controls. Gaussian graphical models (GGMs) use partial Pearson correlation of residuals for each pair of metabolites and identifies significantly associated pairs to identify molecules that are correlated after other effects are conditioned away, ideally resulting in connections that are representative of direct interactions [57]. In this manner, GGMs have been used to identify networks of interacting metabolites [58].

In some applications, it is desirable to obtain a metabolomic signature that predicts case/control status or differentiates cases from controls, for which a Bayesian network is appropriate [59]. A Bayesian network (BN) is a data structure that encodes conditional probability distributions among variables of interest by using a graph composed of nodes and directed edges [60]. Bayesian networks are an attractive modeling methodology since they can model complex interactions between many variables of interest [61], and are particularly appropriate in metabolomics where the number of predictors is relatively low and the emphasis is more on identifying nonlinear and conditional interactions that may have large effect sizes than on filtering through many noisy variables. These attributes have led to successfully replicated BN prediction from metabolomic profiling [62]. In combination with metabolomics, BNs are ideal for incorporating integrative 'omics' data, including gene, SNP, and methylation data [57]. Furthermore, as the cost of metabolomic profiling decreases, it is possible that dynamic Bayesian networks will be used in time-series metabolomic datasets to reconstruct concentration changes over time, just as they have been effective in the reconstruction of gene regulatory networks from time series gene expression experiments [63].

The application of these analytic approaches to investigate a comprehensive panel of metabolites in a large number of samples from well-characterized asthma cohorts has the potential to greatly inform asthma pathogenesis.

Review of Metabolomics Studies of Asthma

Asthma diagnosis and severity

As of early 2015, ten asthma metabolomics studies have been published, all of which focused on asthma diagnosis, control, and/or severity [11-14,64-69]. Recent studies focused on investigating the potential of metabolomics profiling of urine, serum and EBC to differentiate asthmatic from non-asthmatic subjects. Galhleitner et al.
identified a panel of eight candidate asthma-specific, volatile organic compounds from EBC that could differentiate asthmatic (N = 11) vs. healthy control (N = 12) samples using 2D-PCA [64]. More recently, Motta et al. applied PLS-DA to NMR-based metabolomics data from 35 asthmatic and 35 healthy EBC samples, conducted a subsequent multivariate statistical data analysis using projection methods to correctly differentiate asthmatics from healthy controls, and validated their model in 20 additional asthmatics and 20 controls [10]. Similarly, Jung et al. applied an NMR-based profiling approach of sera to distinguish 39 asthmatics and 26 controls, and identified metabolites related to hypermethylation, hypoxia response and immune responses [13].

While these studies demonstrated the applicability of metabolomics for noninvasive asthma diagnostics and therapeutic monitoring, others have revealed its potential in differentiating asthmatic phenotypes, and providing mechanistic insight. Furthermore, these studies have also applied metabolomics profiles to distinguish individuals with well-controlled asthma from those with poorly controlled asthma. An NMR-based approach combined with multivariate modeling produced a discriminatory model that not only successfully differentiated asthmatics from healthy subjects but also could classify asthmatic sub-phenotypes based on sputum eosinophilia, neutrophilia, asthma control and inhaled corticosteroid use [70]. Using an untargeted LC-MS approach to profile urine samples from 41 asthmatics and 12 healthy control subjects, Matturucchi et al. applied multivariate models to distinguish metabolic profiles that could characterize asthma from non-asthmatics, and asthma control [14]. In particular, the model differentiated poor vs. well controlled asthma in patients taking SABA (N = 14), and poor vs. well controlled asthma in patients using a daily controller (N = 11) [14]. In addition, the authors showed that metabolites could have an underlying role in the inflammatory mechanisms contributing to asthma severity. In general, these studies had good predictive accuracy (>80%) in differentiating asthmatics versus controls; however, due to the small sample sizes, statistical power was limited in all studies. Furthermore, only one of these studies included an independent replication cohort to validate their initial findings.

Metabolomics studies of asthma severity have focused on differentiating severe from non-severe asthma, identifying biomarkers that can predict asthma exacerbations, and profiling metabolites associated with oxidative stress resulting from exacerbations. Carraro et al. profiled EBC samples from 31 non-severe asthmatic children, 11 children with severe asthma, and 15 healthy control children and found that metabolites related to retinoic acid, adenosine and vitamin D could distinguish between severe vs. non-severe asthma, and severe asthma vs. healthy controls [68]. Loureiro et al. profiled urinary metabolic changes related to asthma exacerbation in a cohort of 10 adult asthmatics, and conducted PCA-based analyses of metabolites obtained by GC × GC-TOFMS and NMR based methods, finding that urinary metabolic profiles were markedly altered during exacerbations, with increased levels of aldehyde and alkane metabolites in particular [65]. As alkanes and aldehydes are products of oxidative metabolism and increased oxidative states, these results implicate greater oxidative stress during exacerbation vs. stable asthma. Saude et al. also identified five metabolites in the TCA cycle with a higher abundance in exacerbators [12]. This is consistent with the histamine release that occurs during mast cell activation in asthma exacerbations. Furthermore, Voraphani et al. identified a novel metabolome driving nitrative stress in human airway cells that was associated with severe refractory asthma [66]. Together, these findings and others are promising, identifying metabolites related to vitamin D, retinoic acid, TCA metabolism, hypoxic and nitrative stress, immune reaction and inflammation, all of which are biologically plausible metabolites for asthma.

Although encouraging, the studies to date are limited in size, complexity of phenotype(s), number of metabolites, and, in general, lack validation. Therefore, the use of large, well-characterized asthmatic cohorts that can validate initial findings using an independent population(s) is necessary. Additionally, leveraging longitudinal data on asthma severity together with metabolic profiling can establish temporality, thereby further informing prediction and enhancing clinical relevance.

Pathway-based asthma metabolomics

The heterogeneity in the biological pathways that underlie asthma development and progression has different implications for asthma treatment. Asthma is increasingly recognized as an inflammatory mediator-driven process. Leukotrienes and prostaglandins, two families of pro-inflammatory mediators that arise via arachidonic acid metabolism, have been implicated in the inflammatory cascade occurring in asthmatic airways [71,72]. Therapeutic agents such as the leukotriene pathway inhibitors montelukast, zafirlukast, and zileuton have become established medications for reducing asthma symptom severity, as they are known to modulate important cellular and physiological activities related to asthma symptoms, including neutrophil activation, chemotaxis, eosinophil migration and smooth muscle contraction [73,74]. These medications exert their leukotriene reducing effects through blocking leukotriene production (zileuton) or by interfering with cellular leukotriene responses via inhibition of leukotriene receptor binding (e.g. montelukast). However, the clinical response to anti-leukotriene medications is highly variable, but repeatable between individuals, suggesting that biological variation plays an important role in individual response [75-77]. Variation in multiple genes involved in the leukotriene pathway has been implicated in heterogeneous responses to anti-leukotriene medications, and increased leukotriene production is directly related to asthma symptom severity. Therefore, identifying specific metabolites that reflect the activity of the leukotriene pathway may be useful in guiding pathway-specific treatment in asthma. In addition to the leukotriene pathway, a wide body of literature exists to support the role of tryptophan metabolism in the development and treatment of asthma. A partial blockade in tryptophan metabolism has been identified in some asthmatic individuals [78]. In addition, a recent study suggests that tryptophan hydroxylase 1, an enzyme involved in the conversion of tryptophan to serotonin [79], represents a novel therapeutic target for asthma, as it markedly reduces allergic airway inflammation [80]. Serotonin has also been identified as a potent bronchoconstrictor associated with asthma; blockade of serotonin reuptake results in increased pulmonary function correlated with increased free serotonin plasma levels [81,82]. These examples illustrate the advantage that identifying pathway specific metabolites may help in more effectively connecting underlying genetics to pathway specific markers, to more effectively inform asthma treatment.

Limitations

Published studies on the metabolomics of asthma are limited in number and scope. These studies report good predictive accuracy in differentiating asthmatics from control individuals; however, they are limited in size, complexity of phenotype(s), and the number of metabolites investigated. In addition, all of these studies lack replication,
which is essential when both the sample size and number of metabolites evaluated are limited in the initial analysis. The use of large, well-characterized asthmatic cohorts for both discovering and validating asthma metabolomics studies is crucial, and is yet to be performed.

Clinical pharmacometabolomics—the future promise for asthma clinical care

The goal of pharmacogenomics is to reliably predict individual variation in drug responses in order to improve therapeutic outcomes for patients. However, due to the complexity of drug actions, which are mediated by diverse metabolic and pharmacological pathways, pharmacogenomic investigations have only reliably identified genetic variants with large effect sizes, which represent a minority of potential pharmacogenetic variants. To clarify this heterogeneity in patient drug responses, investigating the perturbations in the metabolic pathways of “good” and “poor” drug responders has been an emerging focus of personalized medicine. Using metabolomics to inform pharmacogenomic investigations, also known as “pharmacometabolomics”, assists in efforts to predict individual variation in drug response phenotypes, leading to the identification of novel diagnostic markers for drug responses in addition to clarifying the mechanisms underlying adverse drug events. Pharmacometabolomics has been applied to investigate multiple complex diseases; however, pharmacometabolomic studies are presently lacking for asthma. Despite this, the performance of such studies has the potential inform the current treatment of asthma. Below, we discuss the emerging application of pharmacometabolomics in asthma.

In a recent study that represents one of the first, if not the foremost, asthma pharmacometabolomics studies to date, we sought to identify novel predictors of asthma control based on albuterol inhaler use [69]. First, we generated lipidomic data from plasma samples obtained from a case-control cohort with poorly controlled and well-controlled asthma, using liquid chromatography tandem mass spectrometry (LC-MS). The outcome of interest was a binary indicator of asthma control defined by albuterol inhaler use in the preceding week. We then integrated metabolomic data with genotype, expression, and methylation data from this cohort to identify genomic and molecular indicators of asthma control. A Conditional Gaussian Bayesian Network (CGBN) model, based on four SNPs and two metabolites (arachidonic acid, prostaglandin E2 and S1P), were enriched by asthma control phenotype.

In this commentary, we provide a brief summary of metabolomics as a field of study, and discuss how this field has the potential to impact our current understanding of asthma pathogenesis. Herein, we defined metabolomics and summarized how metabolomics data are generated, reviewed the basic analytical strategies for metabolomics data, defined pharmacometabolomics, and discussed recent advances. Most importantly, we emphasized the need for more comprehensive asthma metabolomics studies, and the potential that these data may have on improving our understanding and treatment of the disease.

References


