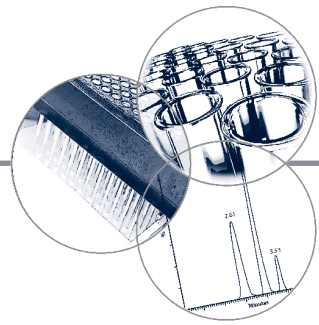
Review
Special Focus Issue: Metabolomics



Application of metabolomics approaches to the study of respiratory diseases

Metabolomics is the global unbiased analysis of all the small-molecule metabolites within a biological system, under a given set of conditions. These methods offer the potential for a holistic approach to clinical medicine, as well as improving disease diagnosis and understanding of pathological mechanisms. Respiratory diseases including asthma and chronic obstructive pulmonary disorder are increasing globally, with the latter predicted to become the third leading cause of global mortality by 2020. The root causes for disease onset remain poorly understood and no cures are available. This review presents an overview of metabolomics followed by in-depth discussion of its application to the study of respiratory diseases, including the design of metabolomics experiments, choice of clinical material collected and potentially confounding experimental factors. Particular challenges in the field are presented and placed within the context of the future of the applications of metabolomics approaches to the study of respiratory diseases.

Stuart Snowden[1,2], Sven-Erik Dahlén[2,3] & Craig E Wheelock*[1,2]

[1]Department of Medical Biochemistry & Biophysics, Division of Physiological Chemistry II, Karolinska Institutet, Stockholm, Sweden
[2]Centre for Allergy Research, Karolinska Institutet, Stockholm, Sweden
[3]Unit for Experimental Asthma & Allergy Research, Karolinska Institutet, Stockholm, Sweden
*Author for correspondence:
Tel.: +46 8 5248 7630
E-mail: craig.wheelock@ki.se

Respiratory diseases: the unmet need for new biomarkers & increased understanding of pathobiology

Respiratory diseases are a major cause of global morbidity and mortality, affecting all age groups of the population [1–3]. The principal lung diseases are **asthma**, **chronic obstructive pulmonary disorder** (COPD), pulmonary arterial hypertension, **cystic fibrosis** and, in particular, a range of infectious diseases including TB. Sarcoidosis is an example of a rare immunological disease that shares certain features both with TB and autoimmune reactions. There are in fact many different immunological and inflammatory lung diseases and they are often severe and progressive, as in the case of interstitial pulmonary fibrosis and several occupational diseases, including asbestosis. There is a wealth of published work considering these diseases in detail, and accordingly, the pathology of individual respiratory diseases will not be discussed in this review. Interested readers are directed to a number of review articles that deal with these topics in detail [4–7].

Lung diseases are very prevalent as demonstrated by the following two examples. First, asthma is in fact the greatest cause of handicap among children. The lifetime risk of developing asthma is as high as for cancer and diabetes, but in contrast to the latter diseases, asthma is a burden to the health system throughout adult life, making it a major reason for disease-related work absenteeism among adults [8]. Second, COPD alone is estimated to have an annual global mortality of approximately 2.7 million people, cause significant morbidity in >200 million people [9], and is on target to be the third largest cause of global mortality by 2020 [10].

While the exact mechanism(s) of disease onset are still unknown, environmental exposures are directly linked to the development of lung diseases. In addition to triggering of allergic asthma by a range of environmental allergens, including grass pollen [11], birch pollen [12,13] dust mites [14] or cat allergens [14], there are many other environmental factors that can exacerbate existing pulmonary conditions; for example, exposure to air pollution in urban areas [15–18], diesel exhaust [19], smoke from heating fires in rural areas [20], and smoke from indoor cooking fires [21,22]. Air pollution in urban areas has been shown to have a detrimental effect on the health of individuals with asthma, especially children [23–25]. Long- and short-term exposures have been shown to induce a range of atopic conditions, including wheezing [26], eczema and allergies [27], and even hospitalization [28], all of which lead to a worse quality of life for many children. In developing countries millions of people are exposed to high levels of air pollution, due to smoke from inefficient and poorly ventilated solid fuel fires (biomass or coal) [21,22,29]. The smoke from these fires has been shown to have a similar composition to tobacco smoke [30] and cause health problems including COPD [22,31], strokes [32], ophthalmic disorders [33], TB [34] and cancer [31].

FUTURE SCIENCE part of fsg

Key Terms

Asthma: Disease characterized by reversible airway narrowing in response to nonspecific stimuli, such as allergens, irritants and environmental pollutants.

Chronic obstructive pulmonary disorder: Heterogeneous disease characterized by airflow obstruction associated with chronic bronchitis, bronchiolitis, emphysema or fibrosis.

Cystic fibrosis: Autosomal recessive genetic disorder that causes the body to produce abnormally thick and sticky mucus. The disease particularly affects the lung and digestive system and is characterized by abnormal transport of ions (sodium and chloride) across epithelia, resulting in viscous secretions.

Metabolomics: Analysis of the whole of the metabolome under a given set of conditions.

Metabolome: Complete compliment of all of the low-molecular weight molecules (<1500 amu) present in the biological compartment in a particular physiological state under a given set of environmental conditions.

Metabolite profiling: Targeted quantification of a predefined subset of metabolite components of the metabolome that usually are of related chemical structure and/or biological activity.

Current clinical diagnosis of respiratory diseases relies on a trained clinician making a decision based upon patient medical history and presentation of symptoms. A range of quantitative and semiquantitative tests, including radiological examinations, spirometry [35], sputum analysis [36] and, more recently, exhaled nitric oxide [36,37], have been used to improve clinical diagnosis of respiratory diseases. Carbon monoxide testing has been shown to be of utility in assessing the smoking status of patients with COPD [38]. There are also numerous chemical tests that can be used to improve clinical diagnosis of respiratory diseases, which quantify the concentrations of a range of markers [39]. Whilst tests that utilize a single biomarker can provide valuable information for clinical diagnosis, they tend to have relatively low specificity. In addition, they are often incapable of identifying and diagnosing specific disease subphenotypes [39,40]. This means that it is often challenging to accurately phenotype patients in terms of the diagnostic subphenotype [39] (e.g., aspirin-intolerant asthma [41]). However, MS-based techniques have been used to measure leukotriene B4 (LTB_4), a marker of inflammation [42], and 8-isoprostane, a marker of oxidative stress [43], in exhaled-breath condensate [44,45]. Nonetheless, disease subphenotypes are poorly described by existing diagnostic criteria and it is expected that approaches employing multiple chemical biomarkers will improve the accuracy and specificity of clinical diagnoses [39,46]. For example, different biomarkers of the same clinical condition are often only weakly, if at all, correlated [47,48], indicating that these markers do not provide a comprehensive picture of disease status.

There are a number of challenges that need to be addressed to improve diagnosis and treatment of respiratory diseases: disease subphenotypes are poorly described using existing diagnostic criteria [39], which means that there is a need for rigorous phenotyping to characterize specific disease states [48]; many current markers are not specific and only enable the generic diagnosis of the disease (e.g., asthma), which makes it necessary to identify sets of new biomarkers capable of diagnosing specific disease subphenotypes; and understanding of the disease processes is still poor, particularly at the metabolic level, and especially in terms of the onset of disease and the response to therapeutic intervention. One major unmet need in respiratory medicine is, therefore, to identify biomarkers that reflect specific pathobiologies. It is expected that the application of **metabolomics** approaches will provide partial solutions to these challenges [46,49–51]. Whereas the use of other 'omics' technologies is well established in studying respiratory disease [52], the application of metabolomics techniques is still in its infancy and is trailing behind their use in other diseases such as cancer [53–56], cardiovascular disease [57–59] and diabetes [60–63]. Much of the metabolomics work carried out to date examining respiratory diseases has focused on the development and validation of analytical approaches that can provide robust and reproducible data from a range of different biofluids [64–67]. The next phase will be to apply the wealth of metabolomics experience acquired in other fields to the study of respiratory diseases.

Introduction to metabolomics

Metabolomics is *"the analysis of the whole of the* ***metabolome*** *under a given set of conditions"* [68]. The exact definition of the metabolome is subject to some debate, but generally can be thought of as the complete compliment of all of the low-molecular weight molecules (<1500 amu) present in the biological compartment in a particular physiological state under a given set of environmental conditions (adapted from [69]). This definition of metabolomics requires that the entire chemical diversity of the metabolome is captured simultaneously; however, this is not possible with currently available technologies [70]. For example, detected metabolite composition in a given matrix is dependent upon multiple parameters, including extraction-solvent polarity [71], choice of chromatographic column (e.g., C_{18}, ion exchange, hydrophilic-interaction LC) [72] and detector (e.g., UV, Raman, mass spectrometer). Accordingly, a single analytical approach will only provide a snapshot of the system. The obstacles faced in analyzing the metabolome were well demonstrated in a recent study by Psychogios *et al.*, who used a multiplatform approach to examine the human serum metabolome [73]. In order to achieve a coverage of 4229 metabolites, six distinct analytical platforms were required, including high-resolution NMR, GC–MS, LC–MS and direct flow injection–MS. While representing one of the most comprehensive studies to date on the human serum metabolome, the identified metabolites most likely do not cover the entire potential metabolome, with the Human Metabolome Database containing over 7900 distinct metabolites to date [74,301]. In addition, it is not feasible to routinely apply six different

analytical platforms to every sample, which would be both expensive and time consuming.

Partially as a result of the difficulty associated with capturing the composition of the metabolome and the challenges posed by working with highly dimensional datasets, different levels of metabolic analysis have been used. These approaches can largely be split into two categories: metabolomics [66,75] and **metabolite profiling** [64,65,76], although the use of these terms differs greatly in the literature. For the purposes of this review, we will employ the following definitions: metabolomics experiments are defined as the *"global, unbiased analysis of the metabolite composition of the biological compartment in a specific physiological state under given environmental conditions"* [77]; metabolite profiling is defined as the *"targeted quantification of a predefined subset of metabolite components of the metabolome that usually are of related chemical structure and/or biological activity"* [77]; and metabolite fingerprinting approaches are *"rapid high-throughput techniques that group data according to shared biochemical characteristics, distinguishing these features from background variation without identifying individual metabolite annotations"* [78]. In a metabolite-profiling experiment, authentic analytical standards for each metabolite being analyzed are employed, enabling exact quantitation [79]. This is not feasible in global metabolomics due to the large number of variables, the exact metabolite identity of which is often unknown [80–82]. Of course, in a metabolite profiling experiment only changes in the focused set of metabolites will be observed, whereas the global nature of metabolomics experiments enables novel areas of metabolism to be identified [83,84].

Numerous analytical techniques can be used in metabolomics analysis; however, they can largely be split into two categories: NMR [54,85] and MS [86,87] (although applications with capillary electrophoresis [88] and other spectral approaches [UV, IR and Raman] are employed for selected applications). While it is generally less sensitive than MS [89], NMR requires minimal sample preparation prior to analysis and offers relatively short analytical run times [89], making it a robust high-throughput technology capable of rapidly analyzing large numbers of samples. NMR techniques are also nondestructive, allowing intact metabolites to be analyzed [90], which can simplify metabolite identification and enables the retention of samples for repeat or further in-depth follow-up analyses. These advantages, as well as relative ease of data interpretation, may explain why NMR techniques have been more widely utilized in the study of respiratory diseases [3,64,65,67,91–93] relative to MS [66,94].

There are multiple approaches to applying MS; however, they can be approximately categorized into two principle types: chromatography-coupled MS and direct- or flow-infusion MS. Chromatography-coupled techniques attempt to simultaneously detect and quantify metabolite peaks, following separation of the sample on a chromatographic column. Chromatographic separation can be achieved in either the gas or liquid phase, depending upon the target analytes. GC–MS techniques are routinely capable of resolving hundreds of metabolite peaks, with metabolite identifications commonly performed by matching electron impact fragmentation spectra and retention indices to established libraries [95]. 2D GC (GC × GC–MS) approaches have also been used in metabolomics [96–99], and because these techniques employ two orthogonal GC columns, an increased number of metabolites can be separated in a single run [100]. GC–MS techniques are limited to the analysis of volatile, thermally stable and relatively nonpolar compounds. Compound volatility can be increased using derivatization; however, this step is laborious and can potentially increase annotation complexity [82]. LC–MS systems are capable of analyzing a wider range of chemical species, including polar and nonvolatile compounds, over a greater mass range than GC–MS approaches and do not require sample derivatization [101]. A significant obstacle in LC–MS approaches is the lack of established spectral libraries, but efforts such as the METLIN [102] and FiehnLib [103] databases represent significant advances. Interested readers are directed to a number recent reviews for more information on the theory and application of MS [104–106].

Rather than performing time-consuming sample separation prior to analysis, it is also possible to utilize direct-infusion MS techniques, in which metabolites are represented simply as a mass variable (m/z) [75,107]. Direct-infusion metabolite fingerprints can be generated in two ways: first, by dissolving the sample in an appropriate solvent and injecting it directly into the ionization chamber (direct-injection MS); or second, the sample can be infused into the ionization chamber in a plug of solvent (flow infusion electrospray ionization-MS [FIE-MS]). Those methods that do not separate sample components prior to analysis are particularly susceptible to

ion suppression, which is caused by changes in the ionization-spray droplet due to the presence of high abundances of less- or non-volatile compounds [107]. The detailed comparison of different metabolomics technologies falls outside the scope of this review, although numerous review articles have addressed these topics specifically [85,107–110].

Choice of clinical material to be analyzed: the importance of matrix selection

One of the primary considerations in the design of a clinical metabolomics experiment is the choice of patient material. There are several biofluids that can be collected, including urine, plasma, serum, saliva, exhaled-breath condensate (EBC), induced sputum supernatants, bronchial wash and bronchoalveolar lavage fluid (BALF). The decision regarding which biofluid to sample is affected by a number of considerations [111], with each matrix possessing a unique combination of strengths and weaknesses (Table 1). Urine, plasma and serum are useful due to ease of collection and minimally invasive nature [39,109]. The standardized nature of collection protocols for these biofluids renders them particularly suitable for use in large multicenter studies [109]. Due to extensive analyses in multiple studies, the protein, lipid and metabolite composition of these matrices is relatively well documented. Urine is also attractive due to its low levels of proteins and cellular material [109], and normalization of total metabolite content using creatinine is well described [110,112–114]. Urine, plasma and serum are integrated biofluids. This offers the simultaneous advantage of reflecting both localized and systemic changes; however, it can be difficult to identify the origins of the observed metabolic changes. Plasma and serum also possess a number of analytical challenges: high protein concentrations means samples require deproteinization prior to analysis [109]; it is preferable to avoid the addition of anticoagulating agents such as heparin, citrate and EDTA because they can affect the metabolite composition of samples [115]; and there can be significant matrix effects associated with analysis [109]. It is also important to note

Table 1. Advantages and disadvantages of the use of different biofluids to characterize respiratory diseases.

Biofluid	Advantages	Disadvantages
Urine	▪ Collection easy and noninvasive ▪ Low protein and cellular material [52,64] ▪ Integrative biofluid ▪ Normalization for dilution (creatinine) ▪ Monitors whole body metabolism	▪ Not directly linked to lung tissue ▪ Presence of high salt levels
Plasma/serum	▪ Integrative biofluid [109] ▪ Easy to collect ▪ Well-described composition	▪ Requires deproteinization prior to analysis [109] ▪ Distant from tissue, biases towards systemic changes ▪ Can contain collection artefacts
Induced sputum	▪ Present in lower respiratory tract [225] ▪ Standardized collection and standard operating procedures	▪ Difficult to normalize for dilution ▪ Potential for plasma infiltration (especially in asthma) ▪ Saliva contamination during collection
Saliva	▪ Easy and noninvasive collection ▪ Low protein and cellular material ▪ Normalization for dilution (amylase)	▪ Susceptible to confounding factors [123] ▪ Limited reports in literature to guide method development
Exhaled-breath condensate	▪ Noninvasive collection ▪ Suitable for analyzing volatile and nonvolatile compounds [89] ▪ Contains mainly polar metabolites	▪ Very dilute, requires concentration (increases variability, insoluble precipitate can form) ▪ Whole of airway sampled, hard to localize changes
Bronchoalveolar lavage fluid	▪ Localized to specific region of lung ▪ Collected under reproducible clinical conditions	▪ Collection is highly invasive ▪ Very dilute, requires concentration (increases variability) ▪ No standardized protocols for normalizing for total metabolite content

that while working with plasma and serum is similar, the metabolite composition of these two biofluids is distinct [116]. In addition, the generation of plasma can result in the formation of artefacts and commensurate shifts in metabolite levels that are solely a result of the sample collection process [117].

EBC and saliva are utilized in metabolomics studies due to the ease and noninvasive nature of collection [2,42,44,89,92,93,118,119]. However, because they have yet to be widely employed, protocols for their collection are not widely standardized, leading to potentially high variability between laboratories. EBC is dilute and often requires concentration prior to analysis, which can lead to increased inter-sample variability. It has been suggested that EBC can be normalized for solute concentration using the sum of the sodium and potassium ions [120]. Saliva has been normalized using amylase concentrations [121]; however, saliva is susceptible to a range of confounding factors that affect the composition of the metabolome (a confounding factor is an internal or external source of metabolic variability that can obfuscate the detection of metabolic changes resulting from the factor being studied). Saliva is especially vulnerable to the effects of diet [122] and smoking [123], but changes resulting from factors including gender [123] and sampling time [67] have also been reported. A small number of metabolomics studies have analyzed BALF [124,125]; however, the highly invasive and expensive nature of sample collection renders it a less-attractive biofluid with which to work. Bronchoscopy can only be performed by well-trained specialists on a limited number of patients [126] and is, therefore, generally not suitable for the collection of temporally associated longitudinal studies. Variability in the volume of saline recovered during collection is a concern; however, normalizing the data using the ratio of instilled versus recovered volume has been shown to slightly reduce variability [127]. BALF samples are also dilute and require concentration prior to analysis, which can introduce further variability. Whilst some metabolite profiling work has been performed on sputum [128], global metabolomics techniques have yet to be applied, which may be due to a number of challenges faced in working with this matrix (Table 1).

Following the choice of clinical material, it is important to design a sampling strategy that will minimize variability and remove sampling bias. A number of factors present in a daily routine significantly affect the metabolome composition (Table 2), including gender [123], smoking [123], sampling time [67], diet [122], environment [15,16], age [129,130], exercise [131,132] and associated comorbidities. Some reported studies chose not to control these confounding factors, under the rationale that *"the metabolites of interest would be altered sufficiently between disease and nondisease groups that such intrapersonal variability would be superseded"* [64,65]. Saude *et al.* were able to classify stable asthmatic patients with an accuracy of 94%, yet only matched the age and gender composition. Numerous strategies have been used to correct for intra- and inter-personal variability:

- Fasting prior to blood or urine collection can reduce the impact of diet-related effects [104,133,134];
- Food frequency questionnaires have been widely used in nutrition studies, and allow compositional changes to be linked to changes in food intake [135]; however, food frequency questionnaire compliance has been shown to be a problem [136];
- Standardization of urine collection protocols is important with different strategies required depending upon the study aim (e.g., first evacuation of the day, collection in the clinic, collection in the home, collection and combination of multiple samples).

Table 2. Metabolites reported as changing in abundance in response to environmental factors in healthy patients.

Metabolite	Comment	Confounding factor
Citrate	Increased in smokers [123]	Smoking
Formate	Decreased in smokers [123]	Smoking
Lactate	Increased in smokers [123]	Smoking
Pyruvate	Increased in smokers[123]	Smoking
Sucrose	Increased in smokers[123]	Smoking
Acetate	Higher in men [123]	Gender
Formate	Higher in men [123]	Gender
Glycine	Higher in men [123]	Gender
Lactate	Higher in men [123]	Gender
Methanol	Higher in men [123]	Gender
Propionate	Higher in men [123]	Gender
Propylene glycol	Higher in men [123]	Gender
Pyruvate	Higher in men [123]	Gender
Succinate	Higher in men [123]	Gender
Taurine	Higher in men [123]	Gender
Alanine	Lower at night [67]	Sampling time
Choline	Lower at night [67]	Sampling time
Methanol	Higher at night [67]	Sampling time
N-acetyl groups	Higher at night [67]	Sampling time
Propionate	Lower at night [67]	Sampling time
Trimethyamine oxide	Lower at night [67]	Sampling time

Achievements of metabolomics in the study of respiratory disease: progress in the field to date

Initial applications of metabolomics have been promising and have successfully classified several respiratory diseases, including asthma [2,64,65,94,99,137], COPD [3,91,93,138] and cystic fibrosis [66,92,125,137] with a high degree of accuracy. No known studies have, to date, investigated sarcoidosis or other respiratory diseases. The high classification accuracy of these models, generated from sample material that was collected noninvasively (e.g., urine and EBC) suggests that metabolomics approaches can play a central role in diagnosing/characterizing respiratory diseases. Metabolomics approaches have also identified the individual candidate metabolites that are responsible for discriminating respiratory disease patients and healthy controls within these experiments (Tables 3–6). These lists of metabolites render it possible to tentatively identify distinct areas of metabolism and the pathways that characterize the individual disease metabolic phenotypes. We present these studies on a disease-specific basis below.

■ Asthma

NMR metabolite profiling techniques were used to classify both stable and exacerbated asthma compared with healthy controls, with an accuracy of 94% [65]. Partial least squares–discriminant analysis (PLS-DA) modeling of these data identified 23 metabolites present at different abundances in asthma patients compared with

Table 3. Metabolites reported as characterizing stable asthma patients compared with healthy controls.

Metabolite	Biofluid	Platform	Urine†	EBC†
1-methylhistamine	Urine [64,65]	NMR [64,65]	Increased	NP
1-methylnicotinamide	Urine [65]	NMR [65]	Increased	NP
2-hydroxyisobutyrate	Urine [64,65]	NMR [64,65]	Increased	NP
2-oxoglutarate	Urine [64,65]	NMR [64,65]	Decreased	NP
3-OH-3-methylglutarate	Urine [64,65]	NMR [64,65]	Decreased	NP
Acetone	Urine [65]	NMR [65]	Increased	NP
Alanine	Urine [65] and EBC [93]	NMR [65,93]	Increased	Decreased
Carnitine	Urine [64,65]	NMR [64,65]	NP	NP
Creatine	Urine [64,65]	NMR [64,65]	Increased	NP
Dimethylamine	Urine [64,65]	NMR [64,65]	Increased	NP
Formate	Urine [65] and EBC [93]	NMR [65,93]	Increased	Increased
Glucose	Urine [64,65]	NMR [64,65]	Increased	NP
Glycolate	Urine [64,65]	NMR [64,65]	Decreased	NP
Hippurate	Urine [65]	NMR [65]	Decreased	NP
Lactate	Urine [65] and EBC [93]	NMR [65,93]	Increased	Decreased
1-methylimidazolacetic acid	Urine [94,150]	UPLC–Q-tof [94]	Decreased	NP
Methylamine	Urine [64,65]	NMR [64,65]	Decreased	NP
O-acetylcarnitine	Urine [65]	NMR [65]	Decreased	NP
Phenylacetylglycine	Urine [64,65]	NMR [64,65]	Increased	NP
Phenylalanine	Urine [64,65] and EBC [93]	NMR [64,65,93]	Increased	Increased
Propionate	EBC [93]	NMR [93]	NP	Decreased
Pyruvate	EBC [93]	NMR [93]	NP	Increased
Succinate	Urine [64,65] and EBC [93]	NMR [64,65,93]	Increased	Increased
Taurine	Urine [65]	NMR [65]	Increased	NP
Threonine	Urine [65] and EBC [93]	NMR [65,93]	Decreased	Decreased
Trans-aconitate	Urine [65]	NMR [65]	Increased	NP
Trigonelline	Urine [65]	NMR [65]	Increased	NP
Trimethylamine	Urine [64] and EBC [93]	NMR [65,93]	Increased	Increased
Trimethylamine *N*-oxide	Urine [65]	NMR [65]	Decreased	NP
Tryptophan	Urine [65]	NMR [65]	Increased	NP
Urocanic acid	Urine [94,155]	UPLC–Q-tof [94]	Decreased	NP

† *'Increased' indicates higher levels in asthmatics and 'decreased' indicates lower levels in asthmatics.*
EBC: Exhaled-breath condensate; NP: The metabolite was not found in this matrix.

Table 4. Metabolites characterizing the differences in the metabolic phenotype of patients with stable and exacerbated asthma versus healthy controls.

Metabolite	Healthy versus stable†	Healthy versus exacerbated†	Stable versus exacerbated†
1-methylhistamine	Increased	Decreased	Increased
1-methylnicotinamide	Increased	ND	Decreased
2-hydroxybutyrate	ND	Increased	Increased
2-hydroxyisobutyrate	Increased	Increased	Increased
2-methylglutarate	ND	Increased	Increased
2-oxoglutarate	Decreased	Increased	Increased
3-methyladipate	ND	Increased	Increased
3-OH-3-methylglutarate	Decreased	Increased	Increased
4-aminohippurate	ND	Decreased	Decreased
4-pyridoxate	ND	Increased	Increased
Acetone	Increased	Increased	Decreased
Adenosine	ND	Increased	Increased
Alanine	Increased	Increased	Increased
Carnitine	Decreased	Increased	Increased
Cis-aconitate	ND	Increased	Increased
Creatine	Increased	Increased	Increased
Dimethylamine	Increased	Increased	Increased
Formate	Increased	Increased	ND
Fumarate	ND	Increased	Increased
Glucose	ND	Increased	Increased
Glycolate	Decreased	Increased	Increased
Hippurate	Decreased	Increased	Increased
Homovanilate	ND	Increased	Increased
Lactate	ND	Increased	Increased
Methylamine	Decreased	Decreased	ND
Myo-inositol	ND	Increased	Increased
O-acetylcarnitine	Decreased	Increased	Increased
Oxaloacetate	ND	Increased	Increased
Phenylacetylglycine	Increased	Decreased	Decreased
Phenylalanine	Increased	ND	Decreased
Succinate	Increased	Increased	Increased
Taurine	Increased	Increased	Decreased
Threonine	Decreased	Decreased	ND
Trans-aconitate	Increased	ND	Decreased
Trigonelline	Increased	Increased	ND
Trimethylamine	Decreased	Decreased	ND
Trimethylamine *N*-oxide	Decreased	Decreased	Decreased
Tryptophan	Increased	ND	Decreased
Tyrosine	ND	Increased	Increased

†*'Increased' indicates metabolite present at higher levels in patient (e.g., stable or exacerbated asthmatic) and 'decreased' indicates metabolite present at lower levels in patient (e.g., stable or exacerbated asthmatic).*
ND: No difference.

healthy controls (TABLES 3 & 4). Mattarucchi *et al.* classified a range of atopic asthma phenotypes using orthogonal projections to latent structures discriminant analysis (OPLS-DA) models generated from nontargeted LC–MS metabolomics data generated from urine samples [94]. Three models were generated – the first model classified asthmatics against healthy controls with an accuracy of 98%; the second model classified medicated versus non-medicated asthmatics with an accuracy of 96%; and the third model classified patients with well-controlled asthma from poorly controlled asthma with an accuracy of 100%. Deeper investigation of these data revealed three metabolites that were excreted at lower levels in asthma patients: urocanic acid,

Table 5. Metabolites characterizing patients with chronic obstructive pulmonary disorder compared with healthy controls.

Metabolite	Biofluid	Platform	Comment†
3-hydroxybutyrate	Serum [91]	NMR	Increased
3-hydroxyisobutyrate	Serum [91]	NMR	Decreased
3-methylhistidine	Serum [91]	NMR	Increased
Acetate	EBC [89]	NMR	Increased
Acetoacetate	Serum [91]	NMR	Increased
Alanine	EBC [93]	NMR	Increased
Ascorbate	Serum [91]	NMR	Increased
Butyrate	EBC [93]	NMR	Decreased
Creatine	Serum [91]	NMR	Decreased
Formate	Urine [3] and EBC [93]	NMR	Positive correlation with lung function
Glutamine	Serum [91] and EBC [89]	NMR	Increased
Glycerol	Serum [91]	NMR	Decreased
Glycine	Serum [91]	NMR	Decreased
Hippurate	Urine [3]	NMR	Positive correlation with lung function
Isobutyrate	Serum [91]	NMR	Decreased
Isoleucine	Serum [91]	NMR	Decreased
Methionine	Serum [91]	NMR	Decreased
Dimethylglycine	Serum [91]	NMR	Decreased
Phenylalanine	Serum [91]	NMR	Increased
Pyruvate	EBC [93]	NMR	Increased
Threonine	EBC [93]	NMR	Increased
Trigonelline	Urine [3]	NMR	Positive correlation with lung function
Trimethylamine	Serum [91]	NMR	Decreased
Valine	Serum [91]	NMR	Decreased

†'Increased' indicates higher levels in chronic obstructive pulmonary disorder patients and 'decreased' indicates lower levels in chronic obstructive pulmonary disorder patients.
EBC: Exhaled-breath condensate.

methyl-imidazolacetic acid and a chemical species resembling an isoleucine-proline fragment. Metabolomics techniques are capable of improving asthma classification compared with traditional diagnostic techniques. Carraro *et al.* used NMR analysis of EBC to classify stable asthma in pediatric patients with an accuracy of 86%, which was an improvement over the 81% accuracy achieved using exhaled nitric oxide and forced expiration volume in 1 s (FEV_1) [2].

In individuals with asthma who have recently suffered an exacerbation, five metabolites acting in the TCA cycle, succinate, fumarate, oxaloacetate, *cis*-aconitate and 2-oxoglutarate have all been reported as being present at higher abundances in urine compared with controls (Figure 1) [65]. This increase in the abundance of TCA cycle intermediates suggests an upregulation in TCA-cycle metabolism as the result of a greater effort to breathe during exacerbation and/or hypoxic stress due to poor oxygenation as a result of exacerbation. Similar shifts in TCA-cycle metabolism have also been reported during exercise [139,140], supporting the hypothesis that increased abundances of these metabolites during exacerbation is a result of the effort to breath and hypoxic stress. Higher levels of lactate were also reported during exacerbation (Table 4), further supporting the idea that patients are undergoing hypoxic stress, because increased levels of this metabolite have been reported during periods of anaerobic exercise [141]. However, during these periods of exercise, increases in the abundance of lactate have been reported as occurring in conjunction with a decrease in the abundance of citrate [140], which has not been observed in individuals with asthma. Nicholson *et al.* hypothesized that this increase in the abundance of lactate would lead to lactic acidosis, which would in turn cause renal tubular acidosis leading to lower levels of glycine and hippurate in urine [142], as both of these compounds have been described as biomarkers of reversible renal malfunction [143]. Neither of these metabolites has been reported as being reduced in abundance in urine after

Table 6. Metabolites characterizing patients with cystic fibrosis compared with healthy controls.

Metabolite	Biofluid	Platform	Comment†
1-methylnicotinamide	PHAECC [66]	UHPLC–MS [66]	Increased
2-propanol	EBC [92]	NMR [92]	Discriminate CF and HC
Acetate	EBC [92]	NMR [92]	Discriminate CF and HC
Acetone	EBC [92]	NMR [92]	Discriminate CF and HC
Adenosine	PHAECC [66]	UHPLC–MS [66]	Decreased
Anthranilate	PHAECC [66]	UHPLC–MS [66]	Increased
Cytidine	PHAECC [66]	UHPLC–MS [66]	Decreased
Formate	EBC [92]	NMR [92]	Discriminate CF and HC
Fructose	PHAECC [66]	UHPLC–MS [66]	Decreased
Fructose-6-phosphate	PHAECC [66]	UHPLC–MS [66]	Decreased
Glucose	PHAECC [66]	UHPLC–MS [66]	Decreased
Glucose-6-phosphate	PHAECC [66]	UHPLC–MS [66]	Decreased
Glutamate	EBC [92]	NMR [92]	Discriminate CF and HC
Glutamine	EBC [92]	NMR [92]	Discriminate CF and HC
Glutathione (oxidized GSSG)	PHAECC [66]	UHPLC–MS [66]	Decreased
Glutathione (reduced GSH)	PHAECC [66]	UHPLC–MS [66]	Decreased
Glycerophosphorylcholine	PHAECC [66]	UHPLC–MS [66]	Decreased
Guanosine	PHAECC [66]	UHPLC–MS [66]	Decreased
Hypoxanthine	PHAECC [66]	UHPLC–MS [66]	Decreased
Inosine	PHAECC [66]	UHPLC–MS [66]	Decreased
Kynurenine	PHAECC [66]	UHPLC–MS [66]	Increased
Lactate	PHAECC [66] and EBC [92]	UHPLC–MS [66] and NMR [92]	Decreased
Malate	PHAECC [66]	UHPLC–MS [66]	Decreased
Mannose-6-phosphate	PHAECC [66]	UHPLC–MS [66]	Decreased
Nicotinamide	PHAECC [66]	UHPLC–MS [66]	Decreased
Opthalmate	PHAECC [66]	UHPLC–MS [66]	Decreased
Propionate	EBC [92]	NMR [92]	Discriminate CF and HC
Ribulose-5-phosphate	PHAECC [66]	UHPLC–MS [66]	Decreased
S-lactoylglutathione	PHAECC [66]	UHPLC–MS [66]	Decreased
Sorbitol	PHAECC [66]	UHPLC–MS [66]	Decreased

†'Increased' indicates higher levels in cystic fibrosis patients and 'decreased' indicates lower levels in cystic fibrosis patients.
CF: Cystic fibrosis, EBC: Exhaled-breath condensate; HC: Healthy controls; PHAECC: Primary human airway epithelial cell cultures.

exacerbation, conversely, hippurate is reported as increasing in abundance (TABLE 4). This suggests that if hypoxic stress is occurring, then it is not at sufficient levels to cause renal malfunction. Hippurate is reported at lower concentrations in the urine of stable asthmatics, in conjunction with increased levels of lactate and some intermediates of the TCA cycle (e.g., succinate). These data might potentially suggest that individuals with asthma may permanently suffer from some degree of hypoxic stress leading to low-level renal malfunction; however, this is highly speculative and there are no reports in the literature to support this theory.

1-methylhistamine is a major downstream metabolite of histamine, a significant pro-inflammatory agent [144,145] and mediator of inflammation [146,147], as well as an important component of allergic metabolism. 1-methylhistamine is produced via the methylation of histamine by histamine methyltransferase [148,149]. 1-methyl-imidazolacetic acid is a downstream product of the histamine methylation pathway (FIGURE 2), with both 1-methylhistamine and 1-mehtyl-imidazolacetic acid occurring at modified abundances in asthma patients (TABLE 3). 1-methylhistamine is reported at higher abundances in asthma patients [65], with lower levels of 1-methyl-imidazolacetic observed in the urine of individuals with asthma [94,150]. In experiments using ^{14}C-labeled histamine, methylation was the primary route of histamine metabolism in humans [123,151], suggesting that 1-methylhistamine and 1-methyl-imidazolacetic acid are appropriate urinary markers of histamine metabolism [152] and potential biomarkers for the diagnosis of pathological inflammation. Histamine release has been reported

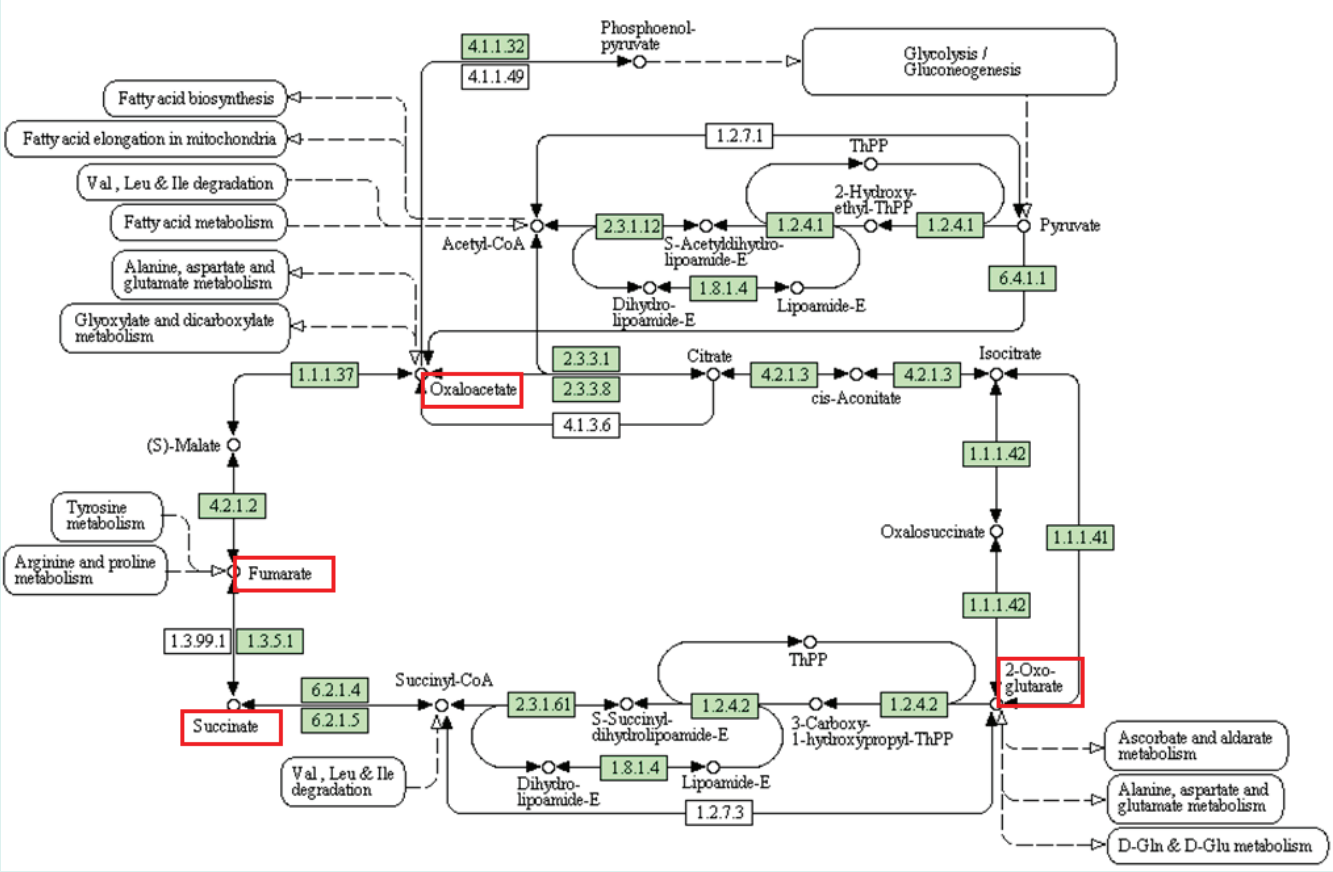


Figure 1. Section of the TCA cycle. Highlighted metabolites have been reported to increase in abundance shortly after asthma exacerbation.
Adapted with permission from [309].

during mast-cell activation in individuals with allergic asthma [153,154]. This might suggest that the changes in the abundance of 1-methylhistamine, 1-methyl-imidazolacetic and urocanic acid observed are also linked to mast-cell activation. The change in abundance of these two metabolites along with decreased levels of urocanic acid in individuals with asthma [94,155] is consistent with a reduction in the metabolism of free histamine [150]. Kerr [150] suggested that the methylation step is the most likely location for the blockage in histamine metabolism; however, the reported increase in the levels of 1-methylhistamine (Table 3) in asthmatics makes this unlikely because if the blockage was upstream of this metabolite, it would not accumulate.

■ COPD

Analysis of serum samples using NMR has been successfully used to classify severe COPD (Global Initiative for Chronic Obstructive Lung Disease [GOLD] stage III [156]) and very severe COPD (GOLD stage IV) patients with an accuracy of 82% [91], where decreased levels of branched-chain amino acids (BCAA) and their associated metabolites are responsible for the discrimination of COPD patients and healthy controls [91]. LC–MS analysis of plasma samples successfully classified emphysematous COPD and non-emphysematous COPD patients with an accuracy of 64% using hierarchical clustering [138]. This classification accuracy is relatively low; however, when using linear discriminant analysis on a subset of biomarkers classification accuracy was improved. When using the seven optimal biomarkers, classification accuracy was improved to 97% [138], although the biomarkers were not identified.

The most significant area of metabolism identified as discriminating COPD patients from healthy controls is the metabolism of BCAA (Table 5) [91,157], with valine, isoleucine and their degradation product 3-hydroxyisobutyrate reported at lower abundances in

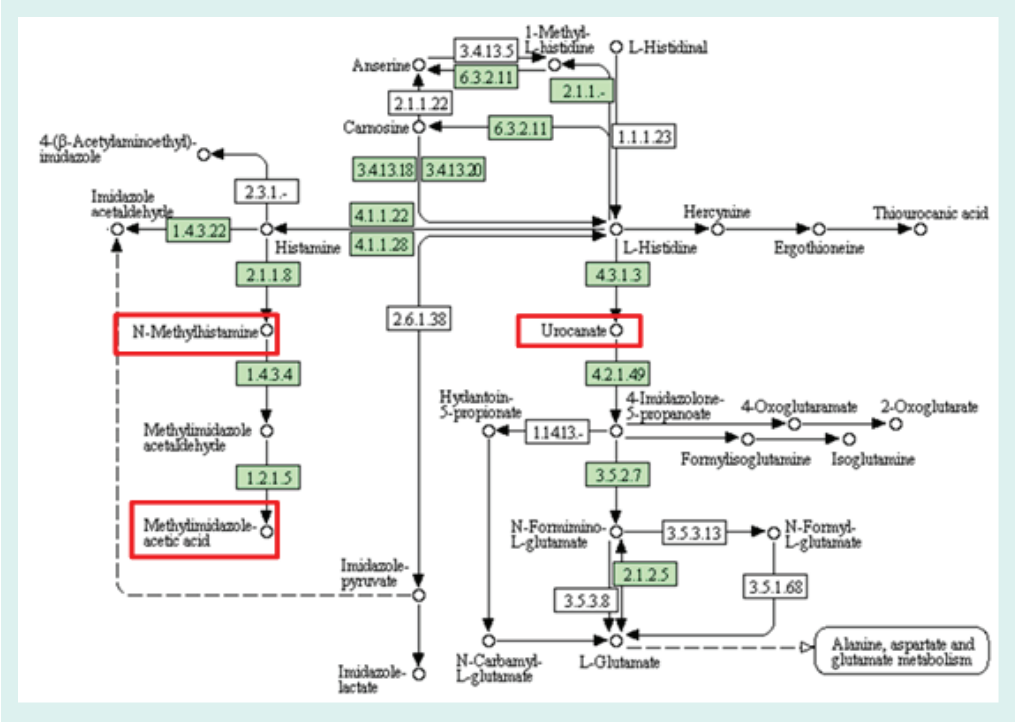


Figure 2. Section of the histidine metabolism pathway. Highlighted metabolites have been reported to characterize patients with stable asthma.
Adapted with permission from [310].

patients suffering from COPD (Figure 3). This modification of BCAA metabolism might be the result of cachexia (a wasting syndrome characterized by muscle atrophy, fatigue, weakness and significant weight loss) in COPD patients. During extended periods of fasting the proteolysis of skeletal muscle and the transamination of BCAA by branched-chain aminotransferase (BCKHD) is an important metabolic response providing a resource for gluconeogenesis [91]. Increased levels of gluconeogenesis that are not suppressed by glucose have been reported in patients suffering from cachexic weight loss [158]. This hypothesis is consistent with the results reported in Ubhi *et al.*, which showed that the reduction in BCAA's correlated with the patient body mass index [91]. Increased levels of BCAA catabolism have also been reported in urine samples collected after exercise [140], where the increase in BCAA catabolism is thought to be the result of proteolysis in skeletal muscle and the visceral region [159]. This might be the result of increased gluconeogenesis due to increased demand for energy during exercise. If accurate, these findings would support the hypothesis that the lower levels of BCAAs in COPD patients are the result of proteolysis, potentially due to wasting.

Cystic fibrosis

The analysis of NMR spectra generated from EBC has been used to classify patients with cystic fibrosis relative to healthy controls with a classification accuracy of 96%. Validation of this model showed that it had a sensitivity of 91% and a specificity of 96% [92]. A second model was generated from these data, classifying stable versus unstable cystic fibrosis patients with an accuracy of 95%. Validation showed that this model had a sensitivity of 86% and a specificity of 94% (the metabolites responsible for the discrimination in both of these models are provided in Table 6) [92]. Preliminary metabolomics analysis has been performed looking at BALF collected from pediatric cystic fibrosis patients, which could be classified into high and low inflammation groups [125]. An OPLS-DA model of NMR spectra was generated that was able to discriminate the two groups of patients. This model had an R^2Y of 0.96 (i.e., the model explained 96% of the variation observed within the data) and a Q^2 of 0.80, indicating that the model had good predictive power.

As with asthma and COPD, the metabolomics studies looking at cystic fibrosis performed to date have successfully identified numerous metabolites and areas of metabolism that characterize the disease state (Table 6). Sorbitol, fructose, glucose,

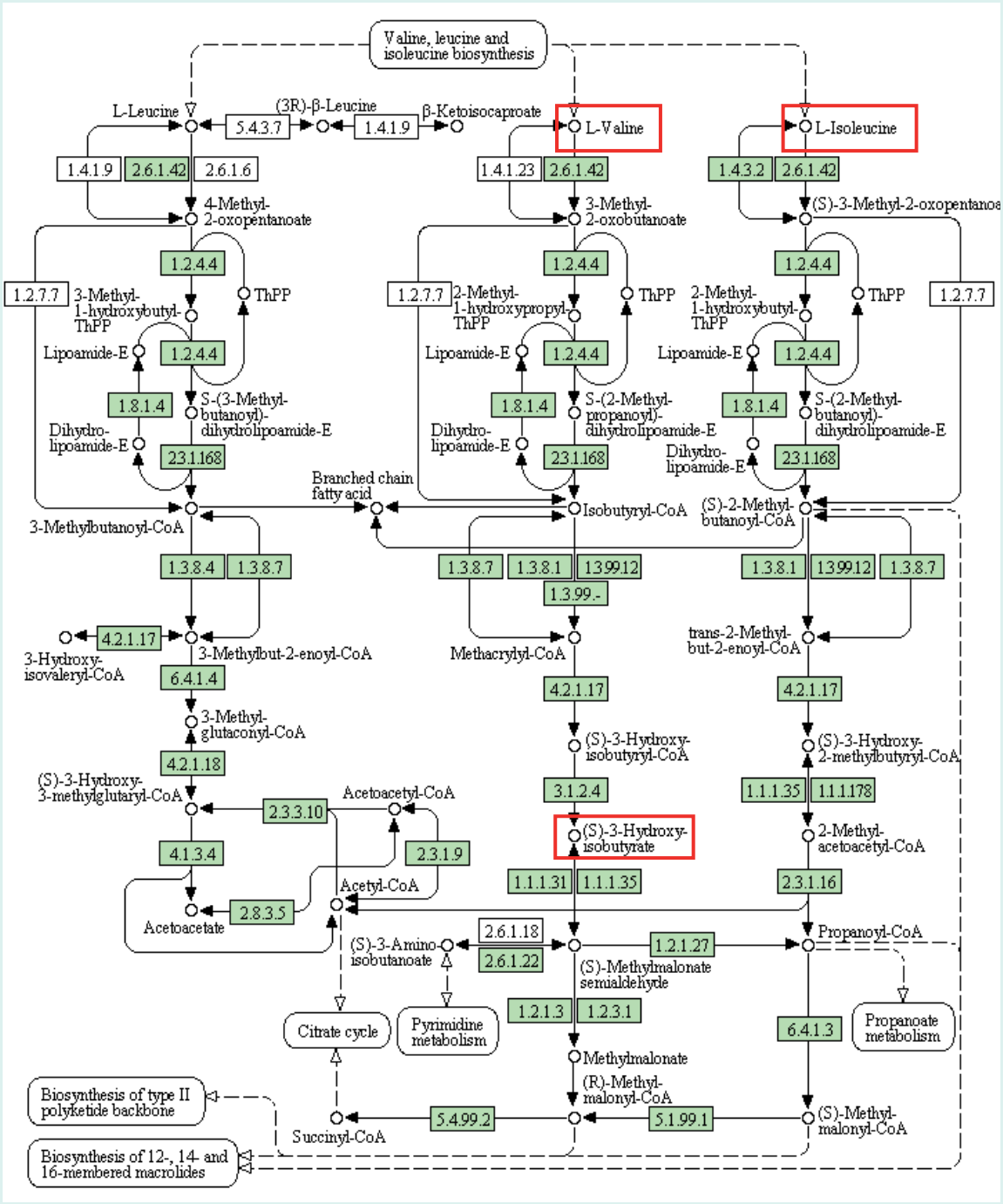


Figure 3. Section of the branched-chain amino acid degradation pathway. Highlighted metabolites have been reported to characterize patients with chronic obstructive pulmonary disorder.
Adapted with permission from [311].

mannose-6-phosphate, fructose-6-phosphate, ribulose-5-phosphate and glucose-6-phosphate have all been reported as occurring at lower abundances in primary human airway epithelial cell cultures [66]. These metabolites act in at least one of three linked metabolic pathways (glycolysis/gluconeogenesis, pentose phosphate pathway and fructose/mannose metabolism) that are involved in the metabolism of glucose (Figure 4). Glucose metabolism plays a central role in normal cellular

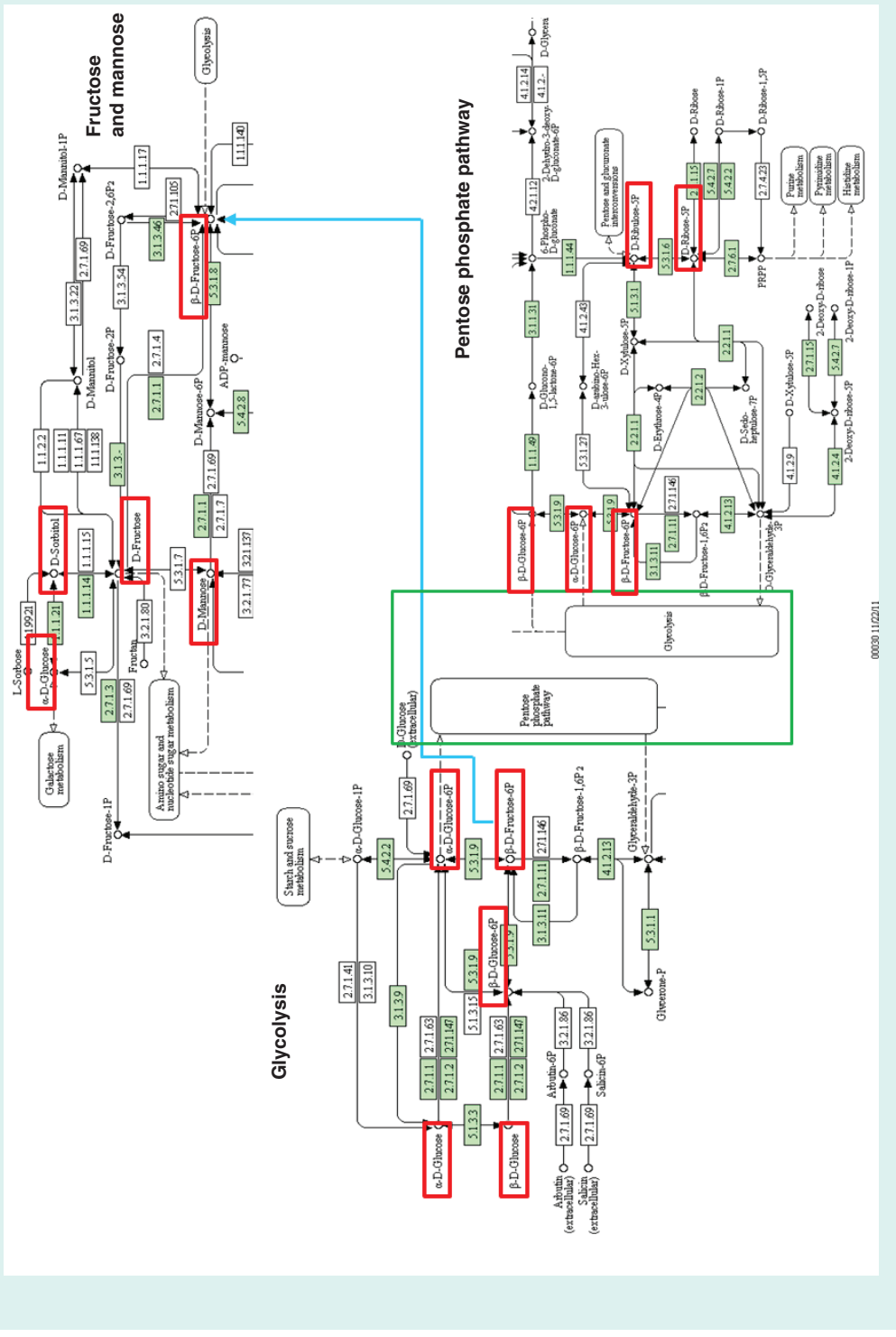


Figure 4. Sections of the pentose phosphate, fructose mannose metabolism and glycolysis pathways. Highlighted metabolites have been reported to characterize patients with cystic fibrosis.
Adapted with permission from [312–314].

function, providing both energy [66] and biosynthetic precursors for the synthesis of a range of biomolecules including nucleotides and fatty acids. It is also interesting to note that suppression of glucose metabolism, as indicated by the reduced abundances of these metabolites, could be responsible for reduced cellular abundances of sorbitol (FIGURE 4). Sorbitol is an important cellular osmolyte and plays a crucial role in maintaining cellular volume and fluid balance [160]. Sorbitol has also been described as promoting the proper maturation of CFTR protein [161], which

is an ATP-gated anion-selective channel that transports chloride ions across the epithelial cell membrane [162]. The transport of chloride ions plays a crucial role in maintaining airway surface liquid volume. Accelerated Na^+ absorption and defective Cl^- secretion in airway epithelial cells is the underlying cause of the pathogenic thickening of airway mucus observed in cystic fibrosis patients [163].

The abundances of five purines and purine metabolites (adenosine, guanosine, inosine, hypoxanthine and adenine) have also been reported to be lower in cystic fibrosis patients (Table 6), suggesting that these patients exhibit lower levels of purine synthesis relative to healthy controls (Figure 5). This is biologically important because purines have been shown to act as signaling molecules [164,165]. Adenosine is especially relevant to cystic fibrosis, because it has been reported to play a role in controlling the viscosity of airway mucus by regulating airway surface liquid volume through activation of CFTR receptors [166,167]. In studies of rat lung tissue, the addition of adenosine directly increased chloride ion flux through CFTR ion channels [168]. Three metabolites relating to glutathione – reduced glutathione, oxidized glutathione and *S*-lactoylglutathione – have been reported at lower concentrations in primary human airway epithelial cell cultures (Table 6). A number of metabolites related to glutathione, including *S*-nitrosoglutathione and *S*-nitrosylating compounds act to promote CFTR cell-surface expression and channel activation in cystic fibrosis epithelial cells [169,170]. As well as these global metabolomics efforts, metabolite profiling of oxylipins has also been used to study cystic fibrosis [171,172].

Challenges in applying metabolomics to the study of respiratory diseases

The application of metabolomics techniques to the study of respiratory diseases faces a number of significant challenges, some of which are common to all metabolomics experiments and some of which are more specific to clinical studies. The most significant of these are discussed below.

■ Biobanks

The generation of biobanks is part of the growing trend towards the use of large-scale biology (such as 'omics' and systems biology) to address questions in clinical science. The term biobank generally refers to repositories of biological material that are relevant to the study of disease [173]. Biobanks are likely to play an integral part in the future of biomedicine because they can supply sufficient numbers of samples to power large-scale systems-based studies. Biobanks will also provide the opportunity to study the etiology of diseases from large populations in a longitudinal fashion, greatly increasing the sensitivity for detecting shifts relating to the onset of disease. For example, the 'LifeGene' biobank aims to collect samples from 500,000 patients, which represents approximately 5% of the Swedish population, thereby enabling diseases to be studied at the population level [174]. However, a number of important ethical issues with using biobanks have been raised [175–177]. There are also a range of analytical challenges faced in using metabolomics techniques to analyze biobank samples. One of the foremost issues is the reproducibility of sample collection strategy employed in the studies, making the development of a standardized collection protocol important [178]. A second major obstacle is the effect of long-term storage on the sample metabolite composition. Even samples stored at -80°C have been demonstrated to exhibit metabolite loss during storage [71,179], with the rate of loss dependent upon the matrix and analyte. For these vast libraries of samples that are being collected in biobanks to be useful in the metabolomics study of clinical diseases, it is necessary to develop and implement a range of standard operating procedures for sample collection [178] and prestorage processing [180]. It is also important to assess the effects of different storage strategies on the metabolite compositions of samples to identify the strategy that best preserves the true metabolic phenotype of the sample, and to identify the limits of viable storage so that time and resources are not wasted on analyzing poor-quality samples. Towards this end, reliable degradation markers should be identified to monitor sample integrity. Another issue involves sample ownership, with potential conflicts between researchers who deposit samples in biobanks and the clinical entities that maintain the biobanks. Interested readers are directed to the Biobanking and Biomolecular Resources Research Infrastructure Initiative [302]. It has also been suggested that the interpretation of results generated from biobank samples may be problematic. This is due to the varied nature of respiratory diseases, which exhibit large deviations in pathological phenotype. These differences will make it potentially difficult to link experimental results to clinical aspects of the disease [173] – further illustrating the need for rigorous and standardized clinical phenotyping of patients.

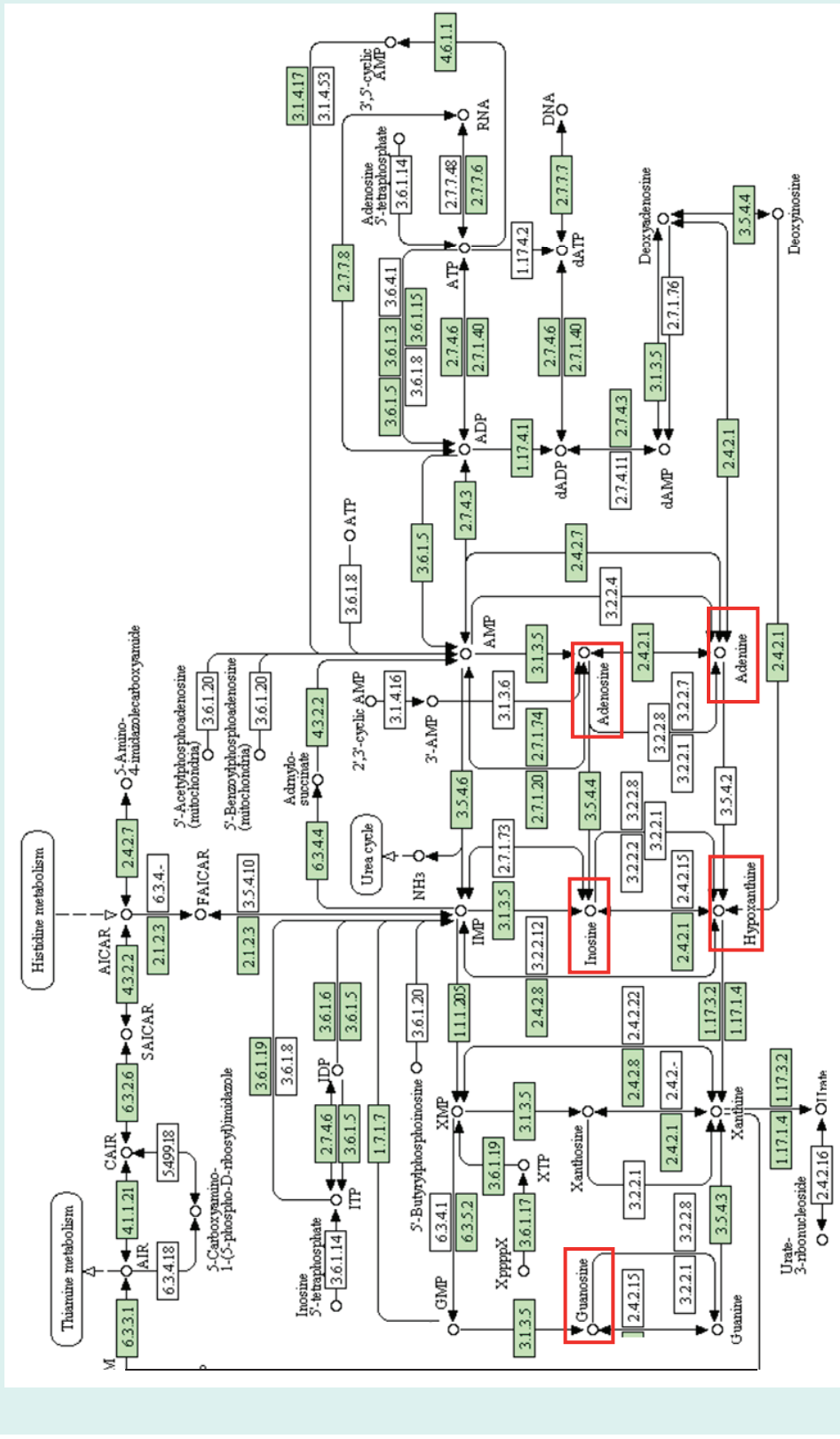


Figure 5. Purine metabolism pathway highlighting metabolites modified in patients with cystic fibrosis.
Adapted with permission from [315].

Experimental design

The high intra- and inter-individual variability need to be addressed in clinical studies [181–183]. Strategies for controlling confounding factors (e.g., diet [122] and smoking [123]) have been described, including matching the age

and gender composition of sample groups [65], smoking history of patients [91], and patient fasting prior to sample collection [104,133,134,184]. However, these approaches only reduce the variability introduced by these confounding factors and do not completely remove it. These strategies also introduce problems related to patient conformity (i.e., patient compliance with study protocols regarding dietary and behavioral restrictions). It would be advantageous to develop analytical strategies to account for this variability; for example, metabolites that are modified in abundance as the result of disease should correlate across all patients with the diseased state, and metabolite abundances that are modified as the result of confounding factors should be excluded.

Metabolomics experiments also suffer from problems of poor laboratory-to-laboratory reproducibility, this is especially true for nontargeted MS-based techniques. To facilitate collaborative research between multiple laboratories, strategies to reduce this variability should be implemented. Initially, mRNA microarray experiments suffered from similar limitations, although studies have now shown that inter-laboratory variability can be reduced via use of common platforms and procedures [185,186]. It has also been shown (for microarray analysis) that data normalization is a vital component in reducing variability [185]. The analysis of standard reference materials is crucial in reducing inter-laboratory variability [187]. The inclusion of reference materials also enables comparison of instrumental accuracy/precision between sample batches, experiments and analytical platforms [187]. One strategy to address this variability is to normalize metabolite abundance by total metabolite composition, which can remove effects of sample dilution in biofluids. Methods for normalizing for sample dilution are well described for saliva using amylase abundance [121], and using creatinine in urine [110,112–114]. It is especially important for urine because its concentration can vary widely [114]; however, it is less of an issue in plasma and serum as solute concentration is tightly controlled [114]. For NMR spectroscopy it is possible to normalize data for total metabolite content by reference to the total spectral area of each sample [119]. More detailed descriptions of strategies for the normalization of metabolite concentration for total metabolite content [114,188], and general discussions of metabolomics and metabolomics workflows [77,109,189], are widely available. While a separate subject in and of itself, the field of design of experiments (DOE) is extremely important, but is unfortunately rarely discussed in either clinical or analytical papers. DOE selects a diverse and representative set of experiments in which all factors are independent of each other despite being varied simultaneously. The result is a causal predictive model showing the importance of all factors and their interactions [303]. Increased effort placed on appropriate DOE will enable the use of decreased sample numbers while maintaining experimental power and will be a useful tool in clinical metabolomics [190].

Data analysis

Data analysis is a critical part of all metabolomics experiments, regardless of the biological system or analytical instrument in question, and presents a major challenge to the development and application of metabolomics techniques. The statistical analysis of metabolomics data is challenging for multiple reasons, with detailed discussion of these problems being outside the scope of this review. A number of articles have dealt with these problems in detail [191–194] and only a few of these issues will be highlighted here. Metabolomics experiments, along with most large-scale biology studies in general, have few degrees of freedom, consisting of low sample number combined with high-dimensionality datasets. It is accordingly challenging to derive meaningful biological knowledge via visual analysis of the dataset [195]. Identifying the appropriate statistical approaches is therefore one of the foremost challenges in data analysis. Univariate statistics, such as analysis of variance (ANOVA) and Student's t-test, identify variables as being significant at a given probability (i.e., $\alpha = 0.05$, which means that there is a 5% chance that any given variable is identified as being significant by chance, a so-called false-positive or type I error.) Within large metabolomics datasets containing thousands of variables [189,196,197], there are potentially hundreds of false-positives, which can be referred to as the false-discovery rate (FDR) [198,199]. Many studies within the 'omics' paradigm employ various approaches to data analysis in addressing the FDR issue, including Tukey's honestly significant difference [200], the Bonferroni correction [201], Holm-Bonferroni [202], the Šidák correction [203,204] and Benjamini and Hochberg [198]. The drawback to many of these approaches for correcting for multiple hypothesis testing relates to the issue of false negatives (type II errors). There is a strong

likelihood that application of an FDR-based correction will result in the generation of false negatives, which is particularly problematic in a hypothesis-generating experiment. Accordingly, any statistical treatment of the data will have to balance type I and II errors appropriately to address the aim of the study.

Another approach that has been particularly employed in the metabolomics community is the use of multivariate statistical analysis. There are numerous multivariate statistical techniques that can be applied to metabolomics data, including unsupervised techniques (e.g., principal component analysis [51,205]), supervised techniques (e.g., PLS-DA [143,205], OPLS-DA [206,207], bidirectional OPLS [O2PLS] [208,209]) and more specialized approaches such as ensemble classifiers (e.g., random forest [210]). For any multivariate analysis, it is important to report the appropriate model statistics. At the minimum, the R^2 and Q^2 values should be provided for principal component analysis, PLS, OPLS-DA and O2PLS models, and eigen values for linear discriminant analysis. In addition, the number of components, and CV-ANOVA p-values, should be provided where appropriate, such as in discriminant analysis. The values of R^2 and Q^2 are commonly utilized model statistics to evaluate a multivariate model. The R^2 represents the percentage of variation within a dataset that can be explained by the model, which is often referred to as a measure of fit. The Q^2 is the percentage variation of the response predicted by the model according to the cross validation, in other words, how accurately the model can predict new data. Unfortunately, many papers omit these values, making it impossible for the reader to assess the quality of the model. Visual inspection of the model does not give evidence of significant separation, especially for supervised methods.

To improve the reporting of results in metabolomics experiments it would be advantageous to develop a set of 'minimum reporting standards' for metabolomics data. Ideally, these standards would apply to both the analytical methodology as well as the statistical analyses. Goodacre *et al.* proposed a set of reporting standards aimed at addressing two main issues: formalize a 'reporting' scheme to prevent confusion over the use of terminology; and define a set of 'minimum reporting standards' for all stages of analysis from data preprocessing to validation of initial hypothesis [195]. These reporting standards represent a starting point; however, they are unfortunately not yet widely adhered to in the scientific literature. The development and implementation of data-reporting standards to enable the exchange of information is an important goal in metabolomics [184]. The Metabolomics Society has initiated the metabolomics standards initiative [304]. These reporting standards outline the minimum information content that should be reported for a metabolomics study, including clear and accurate description of the design and implementation of the study, the subjects involved, biological material sampled and the data collected. This standardization will increase the accessibility of the information and enable the maximum amount of knowledge to be extracted from a dataset [211].

Data interpretation

A significant bottleneck in an 'omics' experiment is interpretation of the acquired data. It is useful to map the metabolites that describe respiratory diseases on to metabolic pathways to identify both areas of metabolism and specific metabolic pathways to develop an improved understanding of the underlying biological perturbation [212]. One of the most powerful tools is the Kyoto Encyclopedia of Genes and Genomes (KEGG) [213,305]; however, there are others, including WikiPathways [306], Ingenuity Pathway Analysis, MetaCyc [307] and the Human Metabolome Database [214,301]. Metabolomics data can be mapped onto KEGG pathways either manually or using the KegArray tool [215], which is a Java-based application that enables metabolomics data to be considered in isolation or in conjunction with both transcriptomics and proteomics data. A relatively new function in the KEGG suite is the DISEASE utility, which highlights known metabolic pathways in a range of diseases including asthma (H00079 [308]). This functionality could play a useful role in interpretation of metabolomics data generated for asthma; however, a closer look at the asthma pathway reveals that there is a significant need for expansion of the mechanistic information to be of utility in data interpretation [127].

Future perspective

The use of metabolomics approaches in the study of respiratory diseases is still in its infancy and lags behind applications in other diseases. To date both NMR- [3,64,65,67,91–93] and MS-based [66,94] metabolomics techniques have been used to analyze respiratory diseases including asthma [2,64,65,94], COPD [3,91,93,138] and cystic fibrosis [66,92,125] in biofluids including urine [3,64,65,94], EBC [2,92,93], serum [91], plasma [138], BALF [125]

and sputum supernatants [128]. These studies have demonstrated the power of metabolomics techniques to classify and potentially diagnose patients suffering from multiple respiratory diseases. In particular, metabolomics approaches have the potential to diagnose respiratory diseases with a higher degree of accuracy than traditional diagnostic methods [2]. Metabolomics approaches have identified several areas of metabolism and metabolic pathways that describe the metabolic phenotype of respiratory diseases. The TCA cycle and histamine metabolism can characterize asthma, with COPD characterized by BCAA catabolism and cystic fibrosis by glucose, purine and glutathione metabolism. Examining the metabolites that are remodeled in each of these respiratory diseases, it can be seen that whilst some metabolites are remodeled in multiple respiratory diseases compared with controls, the metabolic phenotype of each of the three diseases is unique. It is also interesting to note that the metabolic phenotype of an asthma exacerbation is unique to that of stable asthma [65], rather than it being a simple worsening of the underlying asthma phenotype. This is also true for the comparison of the metabolic phenotype of stable and unstable cystic fibrosis, with unstable cystic fibrosis possessing a unique phenotype rather than being a worsening of the stable phenotype. Finally, in the context of studies in asthmatics, the influence of treatments, and in particular the widely used glucocorticosteriods, must be determined in future studies.

The use of metabolomics to study respiratory diseases has produced some promising results and its future applications can largely be split into two areas. The first is to identify and describe the shifts in metabolite composition and metabolism associated with different respiratory diseases. Initial efforts in applying metabolomics approaches should focus on gathering as much information as possible to provide an overview of disease-specific biochemistry. This should be done using both global metabolomics approaches and targeted metabolic-profiling techniques, because global metabolomics methods often lack the sensitivity of metabolic-profiling methods. Accordingly, in order to examine the complete underlying metabolic picture, there is a need for targeted metabolite profiling methods to be performed in parallel to detect low-abundance compounds (e.g., eicosanoids and other oxylipins) [216]. Targeted-profiling methods have identified numerous lipid mediators that would not otherwise be detected using global, nontargeted-metabolomics methods [79]. These mediators play vital roles in the pathology of a range of respiratory diseases including asthma [25,79,217], COPD [218,219] and cystic fibrosis [219–221]. Once sufficient information has been gathered to accurately describe respiratory diseases, it will be necessary to develop targeted methods that are both quantitative and high-throughput to process large numbers of samples to further refine disease models. The second area to which metabolomics can contribute to improved disease management is improving our understanding of the metabolic mechanisms of disease pathology and how this underlying metabolic phenotype responds to therapeutic intervention. For example, it is unlikely that pharmacological treatment shifts the metabolic phenotype of diseased state patients back to that of healthy controls, instead, it is much more probable that the phenotype is shifted to a third state that could be considered as the pharmacological phenotype [222–224]. Accordingly, the true power of metabolomics in identifying and quantifying phenotypes will come into play for classifying subphenotypes of disease and their homeodynamic shifts between disease and pharmacological phenotypes relative to healthy individuals. The acquisition of flux data will enable the development of quantitative models for these dynamic interactions, greatly increasing our ability to monitor disease and predict patient response to interventions. Metabolites are the end products of cellular processes and can be thought of as the ultimate response of a system to genetics and environment [77]. Studying the metabolite composition of systems using metabolomics approaches will provide a powerful tool in understanding the pathological mechanisms of respiratory diseases and for developing new therapeutic strategies for their treatment.

The next 5–10 years should see some exciting developments in the field of metabolomics. It is expected that the use of metabolomics techniques will become more routine, both in general and in applications to respiratory disease. One of the greatest challenges will remain obtaining comprehensive coverage of the metabolome, especially in determining endogenous from exogenous metabolites, in particular those related to diet or microbiota production. While it is not expected that a single analytical platform will evolve in the near future to address this need, the ability of MS to acquire a greater portion of the metabolome will increase. These efforts will most likely involve multidimensional chromatography equipped

with the ability to analyze several stationary phases in a single analysis (e.g., hydrophilic interaction LC, C_{18} and ion exchange). These complex multistationary phase instruments in combination with extremely high pressures, long analytical columns and potentially microfluidics (i.e., nanoflow) will provide the resolution necessary to chromatographically separate a metabolome. These instrument configurations will require high-resolution mass spectrometers with very fast scanning rates and the ability to perform multiple MS/MS experiments without a significant loss in signal. There will also be a concomitant increased emphasis placed on metabolomics kits (e.g., Biocrates and their MetaDisIDQ® Kit) and chip-based technologies with, for example, a lipidomics-based chip designed to capture the lipidome (e.g., Agilent's HPLC-Chip/MS system). One can envision the production of a respiratory kit or an asthma kit capable of identifying or classifying subphenotypes, such as poor glucocorticosteriod responders. One of the major challenges will continue to be the quantitative acquisition of a global metabolome. Given that it is not feasible to have internal standards for all potential metabolites, these efforts will require other novel approaches besides the use of stable isotope dilution or external standards. The field also requires the development of a comprehensive set of analytical standards available as a commercial kit. Such a product could be used for interlaboratory comparisons to produce quantitative data and for routine quality assurance/quality control. An additional need is for standardized reference material, such as the NIST SRM 1950 plasma. This material consists of a plasma pool collected from an equal number of men and women and with a racial distribution that reflects the US population. The inclusion of this material in each published dataset would provide a standardized metric for data normalization. There is a distinct need for a public repository or metabolomics data along the lines of Gene Expression Omnibus; however, to make these data useful, standardization will be necessary. Accordingly, there are multiple challenges remaining in the widespread application of metabolomics methods, but the future is very bright and the field of metabolomics should continue to grow as well as its application to the study of respiratory diseases.

Financial & competing interests disclosure

This work was funded by the Swedish Medical Research Council, the Swedish Heart–Lung Foundation, Vinnova Chronic Inflammation – Diagnostic and Therapy, the Swedish Foundation for Strategic Research, the Stockholm County Council Research Funds, IMI (U-BIOPRED) and Karolinska Institutet. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Executive summary

Respiratory diseases: the unmet need for new biomarkers & increased understanding of pathobiology

- Respiratory diseases are a major cause of global morbidity and mortality.
- Metabolomics has yet to be extensively applied to the field of respiratory disease.
- Applications of metabolomics methodologies represent an opportunity to gain insight into the pathobiology of respiratory diseases, characterize disease subphenotypes and develop new diagnostic tools.

Choice of clinical material to be analyzed: the importance of matrix selection

- Clinical samples have matrix-specific strengths and weaknesses. The appropriate choice depends on the biological question and analytical method being applied. The selection of matrix should be carefully considered during the experimental design phase of any clinical study.

Achievements of metabolomics in the study of respiratory disease: progress in the field to date

- Metabolomics approaches can differentiate healthy controls from individuals with a variety of respiratory diseases, including asthma, chronic obstructive pulmonary disorder and cystic fibrosis. These results suggest that there is potential for diagnostic applications.
- Metabolomics has implicated numerous metabolites that associate with a range of respiratory diseases, providing insight into potential areas of metabolism responsible for disease pathology.

Challenges in applying metabolomics to the study of respiratory diseases

- The most important challenges in a metabolomics experiment include metabolome coverage, data analysis, annotation and subsequent interpretation.
- The primary clinical challenges for respiratory diseases include rigorous clinical phenotyping of patients, identification of disease-specific subphenotypes and normalization of biofluids (e.g., saliva, sputum, bronchoalveolar lavage fluid and exhaled-breath condensate).

References

1 Mukherjee AB, Zhang Z. Allergic asthma: influence of genetic and environmental factors. *J. Biol. Chem.* 286(38), 32883–32889 (2011).

2 Carraro S, Rezzi S, Reniero F *et al.* Metabolomics applied to exhaled breath condensate in childhood asthma. *Am. J. Respir. Crit. Care Med.* 175(10), 986–990 (2007).

3 Mcclay JL, Adkins DE, Isern NG *et al.* ^{1}H nuclear magnetic resonance metabolomics analysis identifies novel urinary biomarkers for lung function. *J. Proteome Res.* 9(6), 3083–3090 (2010).

4 Mastrangelo G, Tartari M, Fedeli U, Fadda E, Saia B. Ascertaining the risk of chronic obstructive pulmonary disease in relation to occupation using a case–control design. *Occup. Med.* 53(3), 165–172 (2003).

5 Holgate ST, Polosa R. The mechanisms, diagnosis, and management of severe asthma in adults. *Lancet* 368(9537), 780–793 (2006).

6 Fireman P. Understanding asthma pathophysiology. *Allergy Asthma Proc.* 24(2), 79–83 (2003).

7 Gibson RL, Burns JL, Ramsey BW. Pathophysiology and management of pulmonary infections in cystic fibrosis. *Am. J. Respir. Crit. Care Med.* 168(8), 918–951 (2003).

8 To T, Wang C, Guan J, Mclimont S, Gershon AS. what is the lifetime risk of physician-diagnosed asthma in Ontario, Canada? *Am. J. Respir. Crit. Care Med.* 181(4), 337–343 (2010).

9 Lopez AD, Shibuya K, Rao C *et al.* Chronic obstructive pulmonary disease: current burden and future projections. *Eur. Resp. J.* 27(2), 397–412 (2006).

10 Rubinsztajn R, Chazan R. An analysis of the causes of mortality and co-morbidities in hospitalised patients with chronic obstructive pulmonary disease. *Pneumonol. Alergol. Pol.* 79(5), (2011).

11 Davies JM, Li H, Green M, Towers M, Upham JW. Subtropical grass pollen allergens are important for allergic respiratory diseases in subtropical regions. *Clin. Transl Allergy* 2(1), 4 (2012).

12 Dreborg S, Foucard T. Allergy to apple, carrot and potato in children with birch pollen allergy. *Allergy* 38, 167–172 (1983).

13 De Swert LF, Gadisseur R, Sjolander S, Raes M, Leus J, Van Hoeyveld E. Secondary soy allergy in children with birch pollen allergy may cause both chronic and acute symptoms. *Pediatr. Allergy Immunol.* 23(2), 118–124 (2012).

14 Lau S, Illi S, Sommerfeld C *et al.* Early exposure to house-dust mite and cat allergens and development of childhood asthma: a cohort study. *Lancet* 356(9239), 1392–1397 (2000).

15 Sunyer J, Spix C, Quénel P *et al.* Urban air pollution and emergency admissions for asthma in four European cities: the APHEA Project. *Thorax* 52(9), 760–765 (1997).

16 Halonen JI, Lanki T, Yli-Tuomi T, Kulmala M, Tiittanen P, Pekkanen J. Urban air pollution, and asthma and COPD hospital emergency room visits. *Thorax* 63(7), 635–641 (2008).

17 Larsson B-M, Grunewald J, Sköld CM *et al.* Limited airway effects in mild asthmatics after exposure to air pollution in a road tunnel. *Resp. Med.* 104(12), 1912–1918 (2010).

18 Stenfors N, Bosson J, Helleday R *et al.* Ozone exposure enhances mast-cell inflammation in asthmatic airways despite inhaled corticosteroid therapy. *Inhal. Toxicol.* 22(2), 133–139 (2010).

19 Sehlstedt M, Behndig AF, Boman C *et al.* Airway inflammatory response to diesel exhaust generated at urban cycle running conditions. *Inhal. Toxicol.* 22(14), 1144–1150 (2010).

20 Sehlstedt M, Dove R, Boman C *et al.* Antioxidant airway responses following experimental exposure to wood smoke in man. *Part. Fibre Toxicol.* 7, 21 (2010).

21 Agrawal S. Effect of indoor air pollution from biomass and solid fuel combustion on prevalence of self-reported asthma among adult men and women in India: findings from a nationwide large-scale cross-sectional survey. *J. Asthma* 49(4), 355–365 (2012).

22 Salvi S, Barnes P. Is exposure to biomass smoke the biggest risk factor for COPD globally? *Chest* 138(1), 3–6 (2010).

23 Tzivian L. Outdoor air pollution and asthma in children. *J. Asthma* 48(5), 470–481 (2011).

24 Trasande L, Thurston GD. The role of air pollution in asthma and other pediatric morbidities. *J. Allergy Clin. Immunol.* 115(4), 689–699 (2005).

25 Kuehl FA, Zanetti ME, Soderman DD, Miller DK, Ham EA. Cyclic AMP-dependent regulation of lipid mediators in white cells: a unifying concept for explaining the efficacy of theophylline in asthma. *Am. J. Respir. Crit. Care Med.* 136(1), 210–213 (1987).

26 Bernstein DI. Traffic-related pollutants and wheezing in children. *J. Asthma* 49(1), 5–7 (2012).

27 Pénard-Morand C, Raherison C, Charpin D *et al.* Long-term exposure to close-proximity air pollution and asthma and allergies in urban children. *Eur. Resp. J.* 36(1), 33–40 (2010).

28 Iskandar A, Andersen ZJ, Bønnelykke K, Ellermann T, Andersen KK, Bisgaard H. Coarse and fine particles but not ultrafine particles in urban air trigger hospital admission for asthma in children. *Thorax* 67(3), 252–257 (2012).

29 Harris SA, Weeks JB, Chen JP, Layde P. Health effects of an efficient vented stove in the highlands of Guatemala. *Glob. Public Health* 6(4), 421–432 (2010).

30 Perez-Padilla R, Schilmann A, Riojas-Rodriguez H. Respiratory health effects of indoor air pollution. *Int. J. Tuberc. Lung Dis.* 14(9), 1079–1086 (2010).

31 Kurmi OP, Lam KBH, Ayres JG. Indoor air pollution and the lung in low and medium income countries. *Eur. Resp. J.* 40(1), 239–254 (2012).

32 Hong YC, Lee J, Kim H, Ha EH, Schwartz J, Christiani DC. Effects of air pollutants on acute stroke mortality. *Environ. Health Persp.* 110(2), 187–191 (2002).

33 Pokhrel AK, Smith KR, Khalakdina A, Deuja A, Bates MN. Case–control study of indoor cooking smoke exposure and cataract in Nepal and India. *Int. J. Epidemiol.* 34(3), 702–708 (2005).

34 Lin H-H, Ezzati M, Murray M. Tobacco smoke, indoor air pollution and tuberculosis: a systematic review and meta-analysis. *PLoS Med.* 4(1), e20 (2007).

35 Celli BR. The importance of spirometry in COPD and asthma: effect on approach to management. *Chest* 117(90020), 15S–19S (2000).

36 Smith AD, Cowan JO, Filsell S *et al.* Diagnosing asthma: comparisons between exhaled nitric oxide measurements and conventional tests. *Am. J. Respir. Crit. Care Med.* 169(4), 473–478 (2004).

37 Dupont LJ. Prospective evaluation of the validity of exhaled nitric oxide for the diagnosis of asthma. *Chest* 123(3), 751–756 (2003).

38 Sandberg A, Sköld CM, Grunewald J, Eklund A, Wheelock ÅM. Assessing recent smoking status by measuring exhaled carbon monoxide levels. *PLoS ONE* 6(12), e28864 (2011).

39 Atzei A, Atzori L, Moretti C *et al.* Metabolomics in paediatric respiratory diseases and bronchiolitis. *J. Matern. Fetal Neonatal Med.* 24(Suppl. 2), 59–62 (2011).

40 Slupsky CM, Rankin KN, Fu H *et al.* Pneumococcal pneumonia: potential for

diagnosis through a urinary metabolic profile. *J. Proteome Res.* 8(12), 5550–5558 (2009).

41 Kim S-H, Hur G-Y, Choi J-H, Park H-S. Pharmacogenetics of aspirin-intolerant asthma. *Pharmacogenomics* 9(1), 85–91 (2007).

42 Montuschi P, Ragazzoni E, Valente S *et al.* Validation of leukotriene B4 measurements in exhaled breath condensate. *Inflamm. Res.* 52(2), 69–73 (2003).

43 Montuschi P, Currò D, Ragazzoni E, Preziosi P, Ciabattoni G. Anaphylaxis increases 8-iso-prostaglandin F2α release from guinea-pig lung *in vitro*. *Eur. J. Pharmacol.* 365(1), 59–64 (1999).

44 Montuschi P, Martello S, Felli M, Mondino C, Chiarotti M. Ion trap liquid chromatography/tandem mass spectrometry analysis of leukotriene B4 in exhaled breath condensate. *Rapid Commun. Mass Spectrom.* 18(22), 2723–2729 (2004).

45 Carpenter CT, Price PV, Christman BW. Exhaled breath condensate isoprostanes are elevated in patients with acute lung injury or ARDS. *Chest* 114(6), 1653–1659 (1998).

46 Madsen R, Lundstedt T, Trygg J. Chemometrics in metabolomics – a review in human disease diagnosis. *Anal. Chim. Acta* 659(1–2), 23–33 (2010).

47 Lemiere C, Ernst P, Olivenstein R *et al.* Airway inflammation assessed by invasive and noninvasive means in severe asthma: eosinophilic and noneosinophilic phenotypes. *J. Allergy Clin. Immunol.* 118(5), 1033–1039 (2006).

48 Bel EH, Sousa A, Fleming L *et al.* Diagnosis and definition of severe refractory asthma: an international consensus statement from the Innovative Medicine Initiative (IMI). *Thorax* 66(10), 910–917 (2011).

49 Ellis DI, Goodacre R. Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy. *Analyst* 131(8), 875–885 (2006).

50 Shyur LF, Yang NS. Metabolomics for phytomedicine research and drug development. *Curr. Opin. Chem. Biol.* 12(1), 66–71 (2008).

51 Griffin J. Metabonomics: NMR spectroscopy and pattern recognition analysis of body fluids and tissues for characterisation of xenobiotic toxicity and disease diagnosis. *Curr. Opin. Chem. Biol.* 7(5), 648–654 (2003).

52 Wheelock Ce, Goss V, Balgoma D *et al.* Application of 'omics technologies to biomarker discovery in inflammatory lung diseases. *Eur. Resp. J.* (2012) (In press).

53 Claudino WM, Quattrone A, Biganzoli L, Pestrin M, Bertini I, Di Leo A. Metabolomics: available results, current research projects in breast cancer, and future applications. *J. Clin. Oncol.* 25(19), 2840–2846 (2007).

54 Kim K, Aronov P, Zakharkin SO *et al.* Urine metabolomics analysis for kidney cancer detection and biomarker discovery. *Mol. Cell. Proteomics* 8(3), 558–570 (2009).

55 Fan TW, Lane AN, Higashi RM *et al.* Altered regulation of metabolic pathways in human lung cancer discerned by (13)C stable isotope-resolved metabolomics (SIRM). *Mol. Cancer* 8, 41 (2009).

56 Sugimoto M, Wong D, Hirayama A, Soga T, Tomita M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics* 6(1), 78–95 (2010).

57 Lewis GD, Asnani A, Gerszten RE. Application of metabolomics to cardiovascular biomarker and pathway discovery. *J. Am. Coll. Cardiol.* 52(2), 117–123 (2008).

58 Sabatine MS, Liu E, Morrow DA *et al.* Metabolomic identification of novel biomarkers of myocardial ischemia. *Circulation* 112(25), 3868–3875 (2005).

59 Brindle JT, Antti H, Holmes E *et al.* Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using 1H-NMR-based metabonomics. *Nat. Med.* 8(12), 1439–1444 (2002).

60 Suhre K, Meisinger C, Döring A *et al.* Metabolic footprint of diabetes: a multiplatform metabolomics study in an epidemiological setting. *PLoS ONE* 5(11), e13953 (2010).

61 Bain JR, Stevens RD, Wenner BR, Ilkayeva O, Muoio DM, Newgard CB. Metabolomics applied to diabetes research: moving from information to knowledge. *Diabetes* 58(11), 2429–2443 (2009).

62 Lanza IR, Zhang S, Ward LE, Karakelides H, Raftery D, Nair KS. Quantitative metabolomics by ^{1}H-NMR and LC–MS/MS confirms altered metabolic pathways in diabetes. *PLoS ONE* 5(5), e10538 (2010).

63 Patterson AD, Bonzo JA, Li F *et al.* Metabolomics reveals attenuation of the SLC6A20 kidney transporter in nonhuman primate and mouse models of Type 2 diabetes mellitus. *J. Biol. Chem.* 286(22), 19511–19522 (2011).

64 Saude EJ, Obiefuna IP, Somorjai RL *et al.* Metabolomic biomarkers in a model of asthma exacerbation: urine nuclear magnetic resonance. *Am. J. Respir. Crit. Care Med.* 179(1), 25–34 (2009).

65 Saude EJ, Skappak CD, Regush S *et al.* Metabolomic profiling of asthma: diagnostic utility of urine nuclear magnetic resonance spectroscopy. *J. Allergy Clin. Immunol.* 127(3), 757–764 (2011).

66 Wetmore DR, Joseloff E, Pilewski J *et al.* Metabolomic profiling reveals biochemical pathways and biomarkers associated with pathogenesis in cystic fibrosis cells. *J. Biol. Chem.* 285(40), 30516–30522 (2010).

67 Bertram HC, Eggers N, Eller N. Potential of human saliva for nuclear magnetic resonance-based metabolomics and for health-related biomarker identification. *Anal. Chem.* 81(21), 9188–9193 (2009).

68 Fiehn O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp. Funct. Genomics* 2(3), 155–168 (2001).

69 Oliver SG, Winson MK, Kell DB, Baganz F. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* 16(9), 373–378 (1998).

70 Dunn WB, Goodacre R, Neyses L, Mamas M. Integration of metabolomics in heart disease and diabetes research: current achievements and future outlook. *Bioanalysis* 3(19), 2205–2222 (2011).

71 Dunn WB, Ellis DI. Metabolomics: current analytical platforms and methodologies. *Trends Analyt. Chem.* 24(4), 285–294 (2005).

72 Gika HG, Theodoridis GA, Wilson ID. Hydrophilic interaction and reversed-phase ultra-performance liquid chromatography TOF-MS for metabonomic analysis of Zucker rat urine. *J. Sep. Sci.* 31(9), 1598–1608 (2008).

73 Psychogios N, Hau DD, Peng J *et al.* The human serum metabolome. *PLoS ONE* 6(2), e16957 (2011).

74 Wishart DS, Knox C, Guo AC *et al.* HMDB: a knowledgebase for the human metabolome. *Nucleic Acids Res.* 37(Suppl. 1), D603–D610 (2009).

75 Parker D, Beckmann M, Zubair H *et al.* Metabolomic analysis reveals a common pattern of metabolic re-programming during invasion of three host plant species by *Magnaporthe grisea*. *Plant J.* 59(5), 723–737 (2009).

76 Fiehn O, Kopka J, Dörmann P, Altmann T, Trethewey RN, Willmitzer L. Metabolite profiling forplant functional genomics. *Nat. Biotechnol.* 18, 1157–1161 (2000).

77 Fiehn O. Metabolomics – the link between genotypes and phenotypes. *Plant Mol. Biol.* 48(1), 155–171 (2002).

78 Scholz M, Gatzek S, Sterling A, Fiehn O, Selbig J. Metabolite fingerprinting: detecting biological features by independent component

analysis. *Bioinformatics* 20(15), 2447–2454 (2004).

79 Lundström SL, Levänen B, Nording M *et al.* asthmatics exhibit altered oxylipin profiles compared to healthy individuals after subway air exposure. *PLoS ONE* 6(8), e23864 (2011).

80 Sana TR, Roark J, Li X, Waddell K, Fischer SM. Molecular formula and metlin personal metabolite database matching applied to the identification of compounds generated by LC–TOF-MS. *J. Biomol. Tech.* 9, 258–256 (2008).

81 Kind T, Fiehn O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics* 7, 234 (2006).

82 Draper J, Enot DP, Parker D *et al.* Metabolite signal identification in accurate mass metabolomics data with MZedDB, an interactive *m/z* annotation tool utilising predicted ionisation behaviour 'rules'. *BMC Bioinformatics* 10, 227 (2009).

83 Chen C, Krausz KW, Idle JR, Gonzalez FJ. Identification of novel toxicity-associated metabolites by metabolomics and mass isotopomer analysis of acetaminophen metabolism in wild-type and Cyp2e1-null mice. *J. Biol. Chem.* 283(8), 4543–4559 (2008).

84 Viant MR, Rosenblum ES, Tjeerdema RS. NMR-Based Metabolomics: a powerful approach for characterizing the effects of environmental stressors on organism health. *Environ. Sci. Technol.* 37(21), 4982–4989 (2003).

85 Reo NV. NMR-based metabolomics. *Drug Chem. Toxicol.* 25(4), 375–382 (2002).

86 Dettmer K, Aronov PA, Hammock BD. Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* 26(1), 51–78 (2007).

87 Katajamaa M, Oresic M. Data processing for mass spectrometry-based metabolomics. *J. Chromatogr. A* 1158(1–2), 318–328 (2007).

88 Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J. Proteome Res.* 2(5), 488–494 (2003).

89 De Laurentiis G, Paris D, Melck D *et al.* Metabonomic analysis of exhaled breath condensate in adults by nuclear magnetic resonance spectroscopy. *Eur. Resp. J.* 32(5), 1175–1183 (2008).

90 Wishart DS. Quantitative metabolomics using NMR. *Trends Analyt. Chem.* 27(3), 228–237 (2008).

91 Ubhi BK, Riley JH, Shaw PA *et al.* Metabolic profiling detects biomarkers of protein degradation in COPD patients. *Eur. Resp. J.* 40(2), 345–355 (2011).

92 Montuschi P, Paris D, Melck D *et al.* NMR spectroscopy metabolomic profiling of exhaled breath condensate in patients with stable and unstable cystic fibrosis. *Thorax* 67(3), 222–228 (2012).

93 Sofia M, Maniscalco M, De Laurentiis G, Paris D, Melck D, Motta A. Exploring airway diseases by NMR-based metabonomics: a review of application to exhaled breath condensate. *J. Biomed. Biotechnol.* 2011, 403260 (2011).

94 Mattarucchi E, Baraldi E, Guillou C. Metabolomics applied to urine samples in childhood asthma; differentiation between asthma phenotypes and identification of relevant metabolites. *Biomed. Chromatogr.* 26(1), 89–94 (2012).

95 Lisec J, Schauer N, Kopka J, Willmitzer L, Fernie AR. Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protocols* 1(1), 387–396 (2006).

96 Almstetter M, Oefner P, Dettmer K. Comprehensive two-dimensional gas chromatography in metabolomics. *Anal. Bioanal. Chem.* 402(6), 1993–2013 (2012).

97 Lankinen M, Schwab U, Seppänen-Laakso T *et al.* Metabolomic analysis of plasma metabolites that may mediate effects of rye bread on satiety and weight maintenance in postmenopausal women. *J. Nutr.* 141(1), 31–36 (2011).

98 Ralston-Hooper K, Hopf A, Oh C, Zhang X, Adamec J, Sepúlveda MS. Development of GC × GC/TOF–MS metabolomics for use in ecotoxicological studies with invertebrates. *Aquat. Toxicol.* 88(1), 48–52 (2008).

99 Caldeira M, Perestrelo R, Barros AS *et al.* Allergic asthma exhaled breath metabolome: a challenge for comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* 1254, 87–97 (2012).

100 Van Mispelaar VG, Tas AC, Smilde AK, Schoenmakers PJ, Van Asten AC. Quantitative analysis of target components by comprehensive two-dimensional gas chromatography. *J. Chromatogr. A* 1019(1–2), 15–29 (2003).

101 De Vos RC, Moco S, Lommen A, Keurentjes JJ, Bino RJ, Hall RD. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 2(4), 778–791 (2007).

102 Smith CA, Maille GO, Want EJ *et al.* METLIN: a metabolite mass spectral database. *Ther. Drug Monit.* 27(6), 747–751 (2005).

103 Kind T, Wohlgemuth G, Lee DY *et al.* FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* 81(24), 10038–10048 (2009).

104 Scalbert A, Brennan L, Fiehn O *et al.* Mass-spectrometry-based metabolomics: limitations and recommendations for future progress with particular focus on nutrition research. *Metabolomics* 5(4), 435–458 (2009).

105 Koulman A, Cao M, Faville M, Lane G, Mace W, Rasmussen S. Semi-quantitative and structural metabolic phenotyping by direct infusion ion trap mass spectrometry and its application in genetical metabolomics. *Rapid Commun. Mass Spectrom.* 23(15), 2253–2263 (2009).

106 Lee DY, Bowen BP, Northen TR. Mass spectrometry-based metabolomics, analysis of metabolite-protein interactions, and imaging. *Biotechniques* 49(2), 557–565 (2010).

107 Beckmann M, Parker D, Enot DP, Duval E, Draper J. High-throughput, nontargeted metabolite fingerprinting using nominal mass flow injection electrospray mass spectrometry. *Nat. Protoc.* 3(3), 486–504 (2008).

108 Oldiges M, Lutz S, Pflug S, Schroer K, Stein N, Wiendahl C. Metabolomics: current state and evolving methodologies and tools. *Appl. Microbiol. Biotechnol.* 76(3), 495–511 (2007).

109 Dunn WB, Broadhurst D, Begley P *et al.* Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* 6(7), 1060–1083 (2011).

110 Viau C, Lafontaine M, Payan JP. Creatinine normalization in biological monitoring revisited: the case of 1-hydroxypyrene. *Int. Arch. Occup. Environ. Health* 77(3), 177–185 (2004).

111 Bos LDJ, Sterk PJ, Schultz MJ. Measuring metabolomics in acute lung injury: choosing the correct compartment? *Am. J. Respir. Crit. Care Med.* 185(7), 789 (2012).

112 Wagner S, Scholz K, Sieber M, Kellert M, Voelkel W. Tools in metabonomics: an integrated validation approach for LC–MS metabolic profiling of mercapturic acids in human urine. *Anal. Chem.* 79(7), 2918–2926 (2007).

113 Jentzmik F, Stephan C, Miller K *et al.* Sarcosine in urine after digital rectal examination fails as a marker in prostate cancer detection and identification of aggressive tumours. *Eur. Urol.* 58(1), 12–18; discussion 20–11 (2010).

114 Warrack BM, Hnatyshyn S, Ott K-H *et al.* Normalization strategies for metabonomic

analysis of urine samples. *J. Chromatogr. B* 877(5–6), 547–552 (2009).

115 Boernsen KO, Gatzek S, Imbert G. Controlled protein precipitation in combination with chip-based nanospray infusion mass spectrometry. an approach for metabolomics profiling of plasma. *Anal. Chem.* 77(22), 7255–7264 (2005).

116 Yu Z, Kastenmüller G, He Y *et al.* Differences between human plasma and serum metabolite profiles. *PLoS ONE* 6(7), e21230 (2011).

117 Goodfriend TL, Pedersen TL, Grekin RJ, Hammock BD, Ball DL, Vollmer A. Heparin, lipoproteins, and oxygenated fatty acids in blood: a cautionary note. *Prostaglandins Leukot. Essent. Fatty Acids* 77(5–6), 363–366 (2007).

118 Battaglia S, Den Hertog H, Timmers MC *et al.* Small airways function and molecular markers in exhaled air in mild asthma. *Thorax* 60(8), 639–644 (2005).

119 Motta A, Paris D, Melck D *et al.* Nuclear magnetic resonance-based metabolomics of exhaled breath condensate: methodological aspects. *Eur. Resp. J.* 39(2), 498–500 (2012).

120 Effros RM. Saving the breath condensate approach. *Am. J. Respir. Crit. Care Med.* 168(9), 1129–1132 (2003).

121 Voho A, Chen J, Kumar M, Rao M, Wetmur J. lipase expression and activity in saliva in a healthy population. *Epidemiology* 17(6), S336 (2006).

122 Walsh MC, Brennan L, Malthouse JPG, Roche HM, Gibney MJ. Effect of acute dietary standardization on the urinary, plasma, and salivary metabolomic profiles of healthy humans. *Am. J. Clin. Nutr.* 84(3), 531–539 (2006).

123 Takeda I, Stretch C, Barnaby P *et al.* Understanding the human salivary metabolome. *NMR Biomed.* 22(6), 577–584 (2009).

124 Fabiano A, Gazzolo D, Zimmermann LJ *et al.* Metabolomic analysis of bronchoalveolar lavage fluid in preterm infants complicated by respiratory distress syndrome: preliminary results. *J. Matern. Fetal Neonatal Med.* 24(Suppl. 2), 55–58 (2011).

125 Wolak JE, Esther CR, Jr., O'Connell TM. Metabolomic analysis of bronchoalveolar lavage fluid from cystic fibrosis patients. *Biomarkers* 14(1), 55–60 (2009).

126 Montravers P, Gauzit R, Dombret MC, Blanchet F, Desmonts JM. Cardiopulmonary effects of bronchoalveolar lavage in critically ill patients. *Chest* 104(5), 1541–1547 (1993).

127 Lundstrom S, Balgoma D, Wheelock AM, Haeggstrom JX, Dahlen SE, Wheelock CE. Lipid mediator profiling in pulmonary disease. *Curr. Pharm. Biotechnol.* 12(7), 1026–1052 (2011).

128 Saude EJ, Lacy P, Musat-Marcu S *et al.* NMR analysis of neutrophil activation in sputum samples from patients with cystic fibrosis. *Magn. Resonan. Med.* 52(4), 807–814 (2004).

129 Holm G, Jacobsson B, Björntorp P, Smith U. Effects of age and cell size on rat adipose tissue metabolism. *J. Lipid Res.* 16(6), 461–464 (1975).

130 Salmon E, Collette F, Degueldre C, Lemaire C, Franck G. Voxel-based analysis of confounding effects of age and dementia severity on cerebral metabolism in Alzheimer's disease. *Hum. Brain Mapp.* 10(1), 39–48 (2000).

131 Krug S, Kastenmuller G, Stuckler F *et al.* The dynamic range of the human metabolome revealed by challenges. *FASEB J.* 26(6), 2607–2619 (2012).

132 Chorell E, Svensson MB, Moritz T, Antti H. Physical fitness level is reflected by alterations in the human plasma metabolome. *Mol. Biosyst.* 8(4), 1187–1196 (2012).

133 Lloyd AJ, Beckmann M, Fave G, Mathers JC, Draper J. Proline betaine and its biotransformation products in fasting urine samples are potential biomarkers of habitual citrus fruit consumption. *Br. J. Nutr.* 106(6), 812–824 (2011).

134 Lloyd AJ, Favé G, Beckmann M *et al.* Use of mass spectrometry fingerprinting to identify urinary metabolites after consumption of specific foods. *Am. J. Clin. Nutr.* 94(4), 981–991 (2011).

135 Wishart DS. Metabolomics: applications to food science and nutrition research. *Trends Food Sci. Technol.* 19(9), 482–493 (2008).

136 Brantsaeter AL, Haugen M, Rasmussen SE, Alexander J, Samuelsen SO, Meltzer HM. Urine flavonoids and plasma carotenoids in the validation of fruit, vegetable and tea intake during pregnancy in the Norwegian Mother and Child Cohort Study (MoBa). *Public Health Nutr* 10(8), 838–847 (2007).

137 Adamko DJ, Sykes BD, Rowe BH. The metabolomics of asthma. *Chest* 141(5), 1295–1302 (2012).

138 Paige M, Burdick MD, Kim S, Xu J, Lee JK, Michael Shim Y. Pilot analysis of the plasma metabolite profiles associated with emphysematous chronic obstructive pulmonary disease phenotype. *Biochem. Biophys. Res. Commun.* 413(4), 588–593 (2011).

139 Enea C, Seguin F, Petitpas-Mulliez J *et al.* (1)H NMR-based metabolomics approach for exploring urinary metabolome modifications after acute and chronic physical exercise. *Anal. Bioanal. Chem.* 396(3), 1167–1176 (2010).

140 Pechlivanis A, Kostidis S, Saraslanidis P *et al.* ^{1}H NMR-based metabonomic investigation of the effect of two different exercise sessions on the metabolic fingerprint of human urine. *J. Proteome Res.* 9(12), 6405–6416 (2010).

141 Nelson DL, Cox MM. *Lehninger Principles of Biochemistry (5th Edition)*. WH Freeman, NY, USA (2008).

142 Nicholson JK, Timbrell JA, Sadler PJ. Proton NMR spectra of urine as indicators of renal damage. Mercury-induced nephrotoxicity in rats. *Mol. Pharmacol.* 27(6), 644–651 (1985).

143 Bairaktari E, Seferiadis K, Liamis G, Psihogios N, Tsolas O, Elisaf M. Rhabdomyolysis-related renal tubular damage studied by proton nuclear magnetic resonance spectroscopy of urine. *Clin. Chem.* 48(7), 1106–1109 (2002).

144 Bachert C. The role of histamine in allergic disease: re-appraisal of its inflammatory potential. *Allergy* 57(4), 287–296 (2002).

145 Sharif NA, Xu SX, Magnino PE, Pang IH. Human conjunctival epithelial cells express histamine-1 receptors coupled to phosphoinositide turnover and intracellular calcium mobilization: role in ocular allergic and inflammatory diseases. *Exp. Eye Res.* 63(2), 169–178 (1996).

146 Singh TS, Lee S, Kim HH, Choi JK, Kim SH. Perfluorooctanoic acid induces mast cell-mediated allergic inflammation by the release of histamine and inflammatory mediators. *Toxicol. Lett.* 210(1), 64–70 (2012).

147 Macglashan D. Histamine: a mediator of inflammation. *J. Allergy Clin. Immunol.* 112(Suppl. 4), S53–S59 (2003).

148 Furukawa F, Yoshimasu T, Yamamoto Y, Kanazawa N, Tachibana T. Mast cells and histamine metabolism in skin lesions from MRL/MP-lpr/lpr mice. *Autoimmun. Rev.* 8(6), 495–499 (2009).

149 Ogasawara M, Yamauchi K, Satoh Y-I *et al.* Recent advances in molecular pharmacology of the histamine systems: organic cation transporters as a histamine transporter and histamine metabolism. *J. Pharmacol. Sci.* 101(1), 24–30 (2006).

150 Kerr JW. Identification of histamine metabolite 1-methyl,4-imidazole acetic acid in human urine and its absence in status asthmaticus. *Br. Med. J.* 2, 3 (1964).

151 Schayer RW, Cooper JAD. Metabolism of C14 histamine in man. *J. Appl. Physiol.* 9(3), 481–483 (1956).

152 Winterkamp WM, Otte P, Stolper J, Schwab D, Hahn EG, Raithel M. Urinary excretion of

N-methylhistamine as a marker of disease activity in inflammatory bowel disease. *Am. J. Gastroenterol.* 97(12), 3071–3077 (2002).

153 Wenzel SE, Fowler AA, Schwartz LB. Activation of pulmonary mast cells by bronchoalveolar allergen challenge: *in vivo* release of histamine and tryptase in atopic subjects with and without asthma. *Am. J. Respir. Crit. Care Med.* 137(5), 1002–1008 (1988).

154 Brown MJ, Ind PW, Causon R, Lee TH. A novel double-isotope technique for the enzymatic assay of plasma histamine: application to estimation of mast cell activation assessed by antigen challenge in asthmatics. *J. Allergy Clin. Immunol.* 69(1, Pt 1), 20–24 (1982).

155 Kerr J. Identification of urocanic acid in the urine in status asthmaticus. *Lancet* 282(7310), 709–710 (1963).

156 Pauwels RA, Buist AS, Calverley PMA, Jenkins CR, Hurd SS. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 163(5), 1256–1276 (2001).

157 Hofford JM, Milakofsky L, Vogel WH, Sacher RS, Savage GJ, Pell S. The nutritional status in advanced emphysema associated with chronic bronchitis: a study of amino acid and catecholamine levels. *Am. J. Respir. Crit. Care Med.* 141(4, Pt 1), 902–908 (1990).

158 Baracos VE, Mackenzie ML. Investigations of branched-chain amino acids and their metabolites in animal models of cancer. *J. Nutr.* 136(1), 237S-242S (2006).

159 Poortman JR. Chapter 8. In: *Principles of Exercise Biochemistry.* Karger Publishers, Basel, Switzerland, 227–278 (2004).

160 Zhang XM, Wang XT, Yue H *et al.* Organic solutes rescue the functional defect in delta F508 cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* 278(51), 51232–51242 (2003).

161 Lang F. Mechanisms and significance of cell volume regulation. *J. Am. Coll. Nutr.* 26(Suppl. 5), S613–S623 (2007).

162 Miller C. CFTR: break a pump, make a channel. *Proc. Natl Acad. Sci. USA* 107(3), 959–960 (2010).

163 Yancey PH. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* 208(Pt 15), 2819–2830 (2005).

164 Haskó G. Adenosine: an endogenous regulator of innate immunity. *Trends Immunol.* 25(1), 33–39 (2004).

165 Blackburn MR. Too much of a good thing: adenosine overload in adenosine-deaminase-deficient mice. *Trends Pharmacol. Sci.* 24(2), 66–70 (2003).

166 Com G, Clancy JP. *Adenosine Receptors, Cystic Fibrosis, and Airway Hydration Adenosine Receptors in Health and Disease*. Wilson CN, Mustafa SJ (Eds). Springer, Berlin, Germany, 363–381 (2009).

167 Kreda SM, Okada SF, Van Heusden CA *et al.* Coordinated release of nucleotides and mucin from human airway epithelial Calu-3 cells. *J. Physiol.* 584(Pt 1), 245–259 (2007).

168 Factor P, Mutlu GM, Chen L *et al.* Adenosine regulation of alveolar fluid clearance. *Proc. Natl Acad. Sci. USA* 104(10), 4083–4088 (2007).

169 Zaman K, Carraro S, Doherty J *et al.* S-nitrosylating agents: a novel class of compounds that increase cystic fibrosis transmembrane conductance regulator expression and maturation in epithelial cells. *Mol. Pharmacol.* 70(4), 1435–1442 (2006).

170 Servetnyk Z, Krjukova J, Gaston B *et al.* Activation of chloride transport in CF airway epithelial cell lines and primary CF nasal epithelial cells by S-nitrosoglutathione. *Respir. Res.* 7, 124 (2006).

171 Yang J, Eiserich JP, Cross CE, Morrissey BM, Hammock BD. Metabolomic profiling of regulatory lipid mediators in sputum from adult cystic fibrosis patients. *Free Radic. Biol. Med.* 53(1), 160–171 (2012).

172 Eiserich JP, Yang J, Morrissey BM, Hammock BD, Cross CE. Omics approaches in cystic fibrosis research: a focus on oxylipin profiling in airway secretions. *Ann. NY Acad. Sci.* 1259(1), 1–9 (2012).

173 Founti P, Topouzis F, Van Koolwijk L, Traverso CE, Pfeiffer N, Viswanathan AC. Biobanks and the importance of detailed phenotyping: a case study – the European Glaucoma Society GlaucoGENE project. *Br. J. Ophthalmol.* 93(5), 577–581 (2009).

174 Almqvist C, Adami H-O, Franks P *et al.* LifeGene – a large prospective population-based study of global relevance. *Eur. J. Epidemiol.* 26(1), 67–77 (2011).

175 Budimir D, Polašek O, Marušić A *et al.* Ethical aspects of human biobanks: a systematic review. *Croat. Med. J.* 52(3), 262–279 (2011).

176 Hansson MG. Ethics and biobanks. *Br. J. Cancer* 100(1), 8–12 (2009).

177 Petersen. The ethics of expectations: biobanks and the promise of personalised medicine. *Monash Bioeth. Rev.* 28(1), 1–12 (2009).

178 Peakman TC, Elliott P. The UK Biobank sample handling and storage validation studies. *Int. J. Epidemiol.* 37(Suppl. 1), i2–i6 (2008).

179 Wittmann C, Krömer JO, Kiefer P, Binz T, Heinzle E. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Anal. Biochem.* 327(1), 135–139 (2004).

180 Jackson DH, Banks RE. Banking of clinical samples for proteomic biomarker studies: a consideration of logistical issues with a focus on pre-analytical variation. *Proteomics Clin. Appl.* 4(3), 250–270 (2010).

181 Rasmussen LG, Savorani F, Larsen TM, Dragsted LO, Astrup A, Engelsen SB. Standardization of factors that influence human urine metabolomics. *Metabolomics* 7(1), 71–83 (2010).

182 Shurubor YI, Matson WR, Willett WC, Hankinson SE, Kristal BS. Biological variability dominates and influences analytical variance in HPLC-ECD studies of the human plasma metabolome. *BMC Clin. Pathol.* 7, 9 (2007).

183 Lundstedt T, Hedenström M, Soeria-Atmadja D *et al.* Dynamic modelling of time series data in nutritional metabonomics – a powerful complement to randomized clinical trials in functional food studies. *Chemometr. Intell. Lab. Syst.* 104(1), 112–120 (2010).

184 Fiehn O, Robertson D, Griffin J *et al.* The metabolomics standards initiative (MSI). *Metabolomics* 3(3), 175–178 (2007).

185 Bammler T, Beyer RP, Bhattacharya S *et al.* Standardizing global gene expression analysis between laboratories and across platforms. *Nat. Methods* 2(5), 351–356 (2005).

186 Kohlmann A, Kipps TJ, Rassenti LZ *et al.* An international standardization programme towards the application of gene expression profiling in routine leukaemia diagnostics: the microarray innovations in leukemia study prephase. *Br. J. Haematol.* 142(5), 802–807 (2008).

187 Bino RJ, Hall RD, Fiehn O *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* 9(9), 418–425 (2004).

188 Sysi-Aho M, Katajamaa M, Yetukuri L, Oresic M. Normalization method for metabolomics data using optimal selection of multiple internal standards. *BMC Bioinformatics* 8, 93 (2007).

189 Patti GJ, Yanes O, Siuzdak G. Innovation: metabolomics: the apogee of the omics trilogy. *Nat. Rev. Mol. Cell. Biol.* 13(4), 263–269 (2012).

190 Eliasson M, Rännar S, Madsen R *et al.* Strategy for optimizing LC–MS data processing in metabolomics: a design of experiments Approach. *Anal. Chem.* 84(15), 6869–6876 (2012).

191 Tikunov Y, Lommen A, De Vos CH *et al.* A novel approach for nontargeted data analysis for metabolomics. Large-scale profiling of tomato fruit volatiles. *Plant Physiol.* 139(3), 1125–1137 (2005).

192 Mendes P, Camacho D, De La Fuente A. Modeling and simulation for metabolomics data analysis. *Biochem. Soc. Trans.* 33, 1427–1429 (2005).

193 Jansen JJ, Hoefsloot HC, Boelens HF, Van Der Greef J, Smilde AK. Analysis of longitudinal metabolomics data. *Bioinformatics* 20(15), 2438–2446 (2004).

194 Verhoeckx KC, Bijlsma S, Jespersen S *et al.* Characterization of anti-inflammatory compounds using transcriptomics, proteomics, and metabolomics in combination with multivariate data analysis. *Int. Immunopharmacol.* 4(12), 1499–1514 (2004).

195 Goodacre R, Broadhurst D, Smilde AK *et al.* Proposed minimum reporting standards for data analysis in metabolomics. *Metabolomics* 3(3), 231–241 (2007).

196 Want EJ, O'Maille G, Smith CA *et al.* Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry. *Anal. Chem.* 78(3), 743–752 (2005).

197 Wikoff WR, Anfora AT, Liu J *et al.* Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc. Natl Acad. Sci. USA* 106(10), 3698–3703 (2009).

198 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57(1), 289–300 (1995).

199 Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* 29(4), 1165–1188 (2001).

200 Miller RG. *Simultaneous Statistical Inference (2nd Edition)*. Springer Verlag, NY, USA (1981).

201 Broadhurst D, Kell D. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* 2(4), 171–196 (2006).

202 Wright Sp. Adjusted P-values for simultaneous inference. *Biometrics* 48, 1005–1013 (1992).

203 Sidak Z. On multivariate normal probabilities of rectangles: their dependence on correlations. *Ann. Math. Statist.* 39, 1425–1434 (1968).

204 Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other. *Am. J. Hum. Genet.* 74(4), 765–769 (2004).

205 Ciosek P, Brzozka Z, Wroblewski W, Martinelli E, Di Natale C, D'Amico A. Direct and two-stage data analysis procedures based on PCA, PLS-DA and ANN for ISE-based electronic tongue-effect of supervised feature extraction. *Talanta* 67(3), 590–596 (2005).

206 Bylesjö M, Rantalainen M, Cloarec O, Nicholson JK, Holmes E, Trygg J. OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification. *J. Chemometr.* 20(8–10), 341–351 (2006).

207 Wiklund S, Johansson E, Sjostrom L *et al.* Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Anal. Chem.* 80(1), 115–122 (2007).

208 Bylesjö M, Eriksson D, Kusano M, Moritz T, Trygg J. Data integration in plant biology: the O_2PLS method for combined modeling of transcript and metabolite data. *Plant J.* 52(6), 1181–1191 (2007).

209 Kirwan GM, Johansson E, Kleemann R *et al.* Building multivariate systems biology models. *Anal. Chem.* 84(16), 7064–7071 (2012).

210 Enot DP, Lin W, Beckmann M, Parker D, Overy DP, Draper J. Preprocessing, classification modeling and feature selection using flow injection electrospray mass spectrometry metabolite fingerprint data. *Nat. Protoc.* 3(3), 446–470 (2008).

211 Quackenbush J. Data standards for 'omic' science. *Nat. Biotech.* 22(5), 613–614 (2004).

212 Kelder T, Conklin BR, Evelo CT, Pico AR. Finding the right questions: exploratory pathway analysis to enhance biological discovery in large datasets. *PLoS Biol.* 8(8), e1000472 (2010).

213 Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 27(1), 29–34 (1999).

214 Wishart DS, Tzur D, Knox C *et al.* HMDB: the Human Metabolome Database. *Nucleic Acids Res.* 35(Suppl. 1), D521–D526 (2007).

215 Wheelock CE, Wheelock AM, Kawashima S *et al.* Systems biology approaches and pathway tools for investigating cardiovascular disease. *Mol. Biosyst.* 5(6), 588–602 (2009).

216 Kupczyk M, Lundström S, Dahlén B, Balgoma D, Wheelock CE, Dahlén S-E. Lipid mediators in severe asthma. In: *Difficult-to-Treat Severe Asthma*. European Respiratory Society Journals Ltd, Lausanne, Switzerland, 218–235 (2011).

217 Luster AD, Tager AM. T-cell trafficking in asthma: lipid mediators grease the way. *Nat. Rev. Immunol.* 4(9), 711–724 (2004).

218 Montuschi P, Kharitonov SA, Ciabattoni G, Barnes PJ. Exhaled leukotrienes and prostaglandins in COPD. *Thorax* 58(7), 585–588 (2003).

219 Zakrzewski JT, Barnes NC, Costello JF, Piper PJ. Lipid mediators in cystic fibrosis and chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 136(3), 779–782 (1987).

220 Konstan MW, Walenga RW, Hilliard KA, Hilliard JB. Leukotriene B4 markedly elevated in the epithelial lining fluid of patients with cystic fibrosis. *Am. J. Respir. Crit. Care Med.* 148(4 Pt 1), 896–901 (1993).

221 Reid DW, Misso N, Aggarwal S, Thompson PJ, Walters EH. Oxidative stress and lipid-derived inflammatory mediators during acute exacerbations of cystic fibrosis. *Respirology* 12(1), 63–69 (2007).

222 Johnson DR, Finch RA, Lin ZP, Zeiss CJ, Sartorelli AC. The pharmacological phenotype of combined multidrug-resistance mdr1a/1b- and mrp1-deficient Mice. *Cancer Res.* 61(4), 1469–1476 (2001).

223 Van Der Greef J, Martin S, Juhasz P *et al.* The art and practice of systems biology in medicine: mapping patterns of relationships. *J. Proteome Res.* 6(4), 1540–1559 (2007).

224 Van Der Greef J, Hankemeier T, Mcburney RN. Metabolomics-based systems biology and personalized medicine: moving towards n = 1 clinical trials? *Pharmacogenomics* 7(7), 1087–1094 (2006).

225 Samuel CM, Whitelaw A, Corcoran C *et al.* Improved detection of *Pneumocystis jirovecii* in upper and lower respiratory tract specimens from children with suspected pneumocystis pneumonia using real-time PCR: a prospective study. *BMC Infect. Dis.* 11, 329 (2011).

■ Websites

301 Human Metabolome Database. www.hmdb.ca

302 Biobanking and Biomolecular Resources Research Infrastructure Initiative. www.bbmri.eu

303 Umetrics. www.umetrics.com

304 Metabolomics Standards Initiative. www.msi-workgroups.sourceforge.net

305 Kyoto Encyclopedia of Genes and Genomes. www.genome.jp/kegg

306 WikiPathways. www.wikipathways.org/index.php/WikiPathways

307 MetaCyc.
www.metacyc.org

308 Kyoto Encyclopedia of Genes and Genomes. Diseases.
www.kegg.jp/dbget-bin/www_bget?ds:H00079

309 Kyoto Encyclopedia of Genes and Genomes. Citrate cycle (TCA cycle).
www.kegg.jp/kegg-bin/show_pathway?ko00020

310 Kyoto Encyclopedia of Genes and Genomes. Histidine metabolism.
www.kegg.jp/kegg-bin/show_pathway?ko00340

311 Kyoto Encyclopedia of Genes and Genomes. Valine, leucine and isoleucine degradation.
www.kegg.jp/kegg-bin/show_pathway?ko00280

312 Kyoto Encyclopedia of Genes and Genomes. Pentose phosphate pathway.
www.kegg.jp/kegg-bin/show_pathway?ko00030

313 Kyoto Encyclopedia of Genes and Genomes. Fructose and mannose metabolism.
www.kegg.jp/kegg-bin/show_pathway?ko00051

314 Kyoto Encyclopedia of Genes and Genomes. Glycolysis and gluconeogenesis.
www.kegg.jp/kegg-bin/show_pathway?ko00010

315 Kyoto Encyclopedia of Genes and Genomes. Purine metabolism.
www.kegg.jp/kegg-bin/show_pathway?ko00230