AMERICAN THORACIC SOCIETY DOCUMENTS

New Strategies and Challenges in Lung Proteomics and Metabolomics

An Official American Thoracic Society Workshop Report

Russell P. Bowler, Chris H. Wendt, Michael B. Fessler, Matthew W. Foster, Rachel S. Kelly, Jessica Lasky-Su, Angela J. Rogers, Kathleen A. Stringer, and Brent W. Winston; on behalf of the American Thoracic Society Workgroup on Metabolomics and Proteomics

This official workshop Report of the American Thoracic Society was approved October 2017

Abstract

This document presents the proceedings from the workshop entitled, "New Strategies and Challenges in Lung Proteomics and Metabolomics" held February 4th–5th, 2016, in Denver, Colorado. It was sponsored by the National Heart Lung Blood Institute, the American Thoracic Society, the Colorado Biological Mass Spectrometry Society, and National Jewish Health. The goal of this workshop was to convene, for the first time, relevant experts in lung proteomics and metabolomics to discuss and overcome specific challenges in these fields that are unique to the lung. The main objectives of this workshop were to identify, review, and/or understand: (*1*) emerging technologies in metabolomics and proteomics as applied to the study of the lung; (*2*) the unique composition and challenges of lung-specific biological specimens for metabolomic and proteomic analysis; (*3*) the diverse informatics approaches and databases unique to metabolomics and proteomics, with special emphasis on the lung; (*4*) integrative platforms across genetic and genomic databases that can be applied to lung-related metabolomic and proteomic studies; and (*5*) the clinical applications of proteomics and metabolomics. The major findings and conclusions of this workshop are summarized at the end of the report, and outline the progress and challenges that face these rapidly advancing fields.

Keywords: mass spectrometry; nuclear magnetic resonance; systems biology; biomarkers; lung diseases

ORCID IDs: 0000-0003-4651-363X (R.P.B.); 0000-0002-0924-8745 (C.H.W.); 0000-0003-0212-2346 (M.W.F.); 0000-0003-3023-1822 (R.S.K.); 0000-0003-0238-7774 (K.A.S.).

Correspondence and requests for reprints should be addressed to Chris H. Wendt, M.D., University of Minnesota, 1 Veterans Drive, Minneapolis, MN 55417. E-mail: wendt005@umn.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Ann Am Thorac Soc Vol 14, No 12, pp 1721–1743, Dec 2017

DOI: 10.1513/AnnalsATS.201710-770WS
Internet address: www.atsjournals.org

Contents

Overview: The Role of Proteomics and Metabolomics in Systems Biology

Personalized disease risk and drug response predictions based on genomic sequences now represent a cornerstone of precision medicine, and have also been successful at informing therapeutic decisions. However, genomics remains relatively limited in its ability to predict the onset of most complex diseases, largely because genomic information does not account for dynamic environmental influences (1). To better understand lung disease, one needs to examine the downstream changes occurring at the level of proteins and metabolites.

Proteins are the main effectors of cellular physiology. Therefore, the proteome, alone or through its integration with other systems sciences, is a particularly informative tool for understanding pulmonary diseases (Figure 1) (2). Until recently, technology was limited to studying the role of single proteins. Although mass spectrometry (MS) has been an important tool for decades, new technologies and strategies in peptide/protein separation, MS analysis, quantitative protein analysis, and databases now enable the simultaneous analysis of dozens to even thousands of proteins in a single biological sample. In parallel, advances in statistical and bioinformatics tools now allow insight into protein pathways and networks involved in lung disease. Consequently, there has been a surge in the number of proteomics publications related to lung disease (Figure 2) (3).

Metabolites, small biological compounds with a low molecular weight (typically ⩽1,500 Daltons), reflect the activity of proteins, and serve as signaling molecules for processes that include gene and protein regulation (4, 5) (Figure 1). As such, the added value of metabolomics (i.e., the simultaneous measurement of small molecules in a biological sample) is that it reflects, complements, and informs data acquired by other systems biology sciences (Figure 1) (6). Because metabolic profiles change rapidly with the biologic state, metabolomics permits unique insight into both the pathogenesis of disease and drug response (pharmacometabolomics), as well as often unapparent phenotypes and endotypes. The trans-omic approach provides a unique opportunity to gain insights in to how genetic

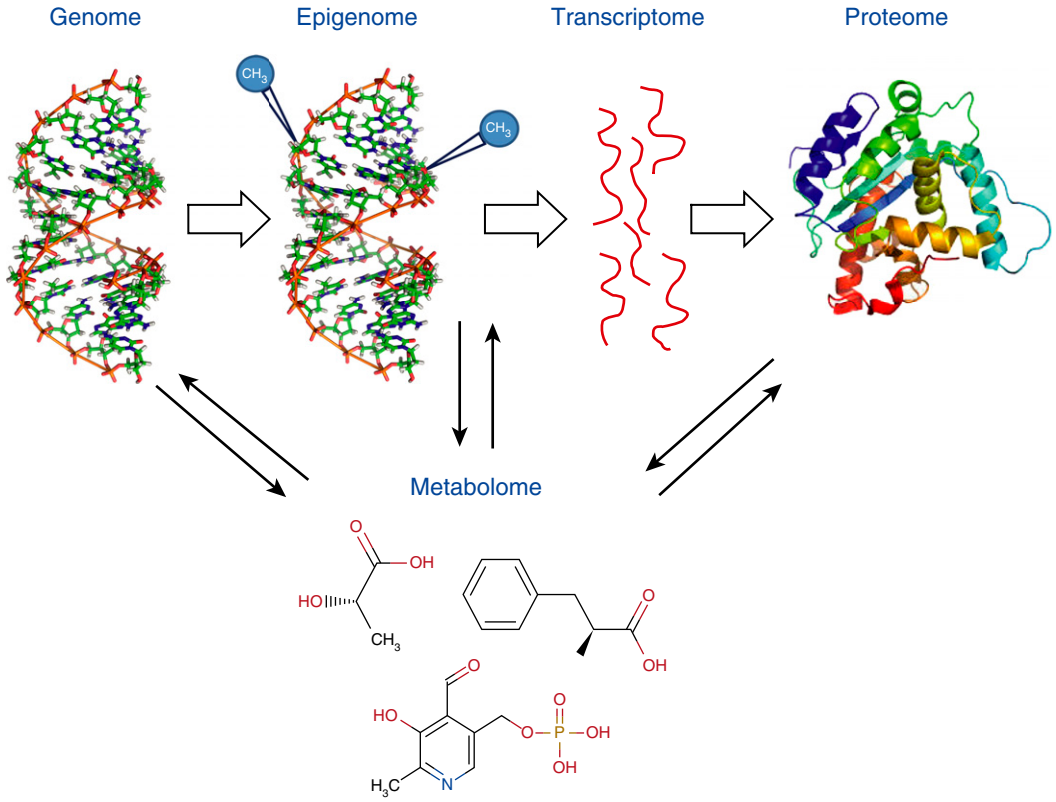

Figure 1. Proteomics and metabolomics are members of systems biology science that includes genomics, epigenomics, and transcriptomics. Metabolomics is particularly reflective of gene and protein activity. *DNA structure*: https://commons.wikimedia.org/wiki/File:A-DNA,_B-DNA_and_Z-DNA.png. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled *GNU Free Documentation License*. *Protein structure*: Structure of the C3 protein. Emw (https://commons.wikimedia.org/wiki/File:Protein_C3_PDB_1c3d.png), "Protein C3 PDB 1c3d," https://creativecommons.org/licenses/by-sa/3.0/legalcode. *Images of metabolites* are publically available from: http://www.hmdb.ca/ with citation of: Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, *et al.*, *HMDB 3.0 — The Human Metabolome Database in 2013.* Nucleic Acids Res. 2013. Jan 1;41(D1):D801-7. 23161693.

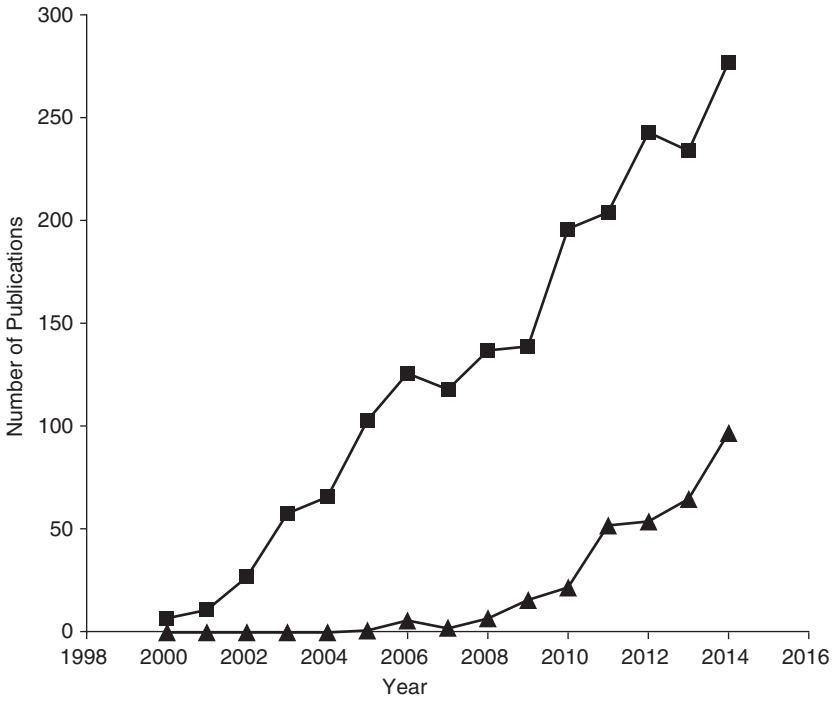

Figure 2. Temporal increase in the number of lung proteomics and metabolomics publications in the PubMed database. *Squares* represent proteomic publications; *triangles* represent metabolomic publications.

programs are translated into biological function and how alterations in the program associate with the onset of diseases. As such, metabolomics may reveal early biomarkers that could improve risk assessment and diagnosis of complex diseases. Furthermore, metabolic profiles are influenced by exogenous factors, including medications, lifestyle, and the environment, so metabolomic profiling has the potential to unravel the impact of both genetic and nongenetic factors on disease onset, progression, and severity.

This report highlights the significant progress and continued major challenges in the fields of proteomics and metabolomics, with particular focus on lung health and disease.

Methods

The conference was created by the American Thoracic Society Respiratory Cell Molecular Biology Assembly Working Group on Proteomics and Metabolomics. The goal of the workshop was to convene *for the first time* relevant experts in lung proteomics and metabolomics to discuss specific challenges that are unique to the lung. The main objectives were to:

- Identify existing and emerging technologies in metabolomics and proteomics as applied to the study of the lung.
- Understand the inherent challenges associated with metabolomics and proteomics, with a specific focus on challenges associated with lung biological specimens.
- Identify informatics approaches and online databases relating to metabolomics and proteomics.
- Discuss systems biology approaches, including integrative platforms across databases that can be applied to metabolomic and proteomic studies.
- Discuss the potential clinical applications of proteomics and metabolomics for lung-related disorders.

The conference included combined sessions relevant to proteomics and metabolomics, and breakout sessions highlighting challenges specific to each. The conference report is a summation of the presentations and discussions. Potential conflicts of interest were disclosed and managed in accordance with the policies and procedures of the American Thoracic Society.

Issues Common to Proteomics and Metabolomics

Unique Challenges of Lung Biospecimens

The lung has several features that make metabolomics and proteomics both a unique opportunity and challenge compared with other organs and blood. One unique quality is the lung's exposure to the environment (air pollution, pollen, etc.); however, as part of an environmental defense, the lung is covered by a complex epithelial defensive barrier and epithelial lining fluid (ELF) consisting of both solute and gel phases containing mucus and lipids (e.g., surfactants), along with inflammatory cells. These features make sample preparation challenging. The collection of lung-specific specimens includes tissue biopsy or pathology specimens, transbronchial biopsies, bronchoalveolar lavage fluid (BALF), and airway brushings. The most frequent lung sampling technique is bronchoscopy to obtain BALF and/or bronchial biopsy. Of note, as in all tissues, biopsies lead to blood content contamination that may influence sample analysis. This may be less of an issue for airway wall biopsies. BALF is typically performed using normal saline (0.9% NaCl), often requiring techniques to remove the high salt content. Furthermore, BALF dilutes the ELF up to 100-fold, and must be taken into account when performing quantitative analysis. Dilution can be minimized by using a single-cycle lavage and corrected by normalizing to urea (7, 8). ELF is rich in plasma-derived proteins (albumin, transferrin, etc.), along with proteins specifically expressed by airway cells, such as surfactant proteins and club cell secretory protein (9).

Other less frequently used biosamples for lung investigations include exhaled breath condensate and epithelial brushes or biopsies. The former is limited by difficulty in standardization and dilute samples, and the latter by bleeding and invasiveness of procedures. Sputum is another lung sample with its own unique challenges, such as viscosity and different layers (gel and sol). To minimize variance and maximize reproducibility, we have included in this document recommended sample protocols for lung and other biofluid samples (Table 1).

Table 1. Sample preparation for proteomics and metabolomics studies*

Sample Type	Preparation (Recommended)	Comments
BALF	Centrifuge to remove cells and debris (800 × g at 4°C for 10 min); remove supernatant. Aliquot and freeze (−80°C) until the time of assay.	Recommend first cycle lavage; correct for dilution (typically 100×) with urea ratio (e.g., BALF:plasma); consider buffer exchange to remove salt (0.9% or 154 mM NaCl) and methanol or acetone precipitation for protein separation.
Sputum (*see* https://www2.cscc.unc.edu/spiromics/system/files/documents/sputmcmtmop5sputumcollectionandprocessing09172014_0.pdf for full protocol)	Blow nose, gargle, and rinse mouth using room temperature water; separate sputum plugs; weigh and measure 0.3 g of whole sputum sample and consider solubilizing an aliquot with DTT (0.1%) or 6 M guanidine reduction buffers to solubilize.	Healthy subject will likely need nebulization with hypertonic (3%) saline
Exhaled breath condensate (60)	Consider commercial equipment, such as Ecoscreen or RTube; collect during tidal breathing using a nose clip and a saliva trap; define cooling temperature and collection time (10 min is generally sufficient to obtain 1–2 ml of sample and is well tolerated by patients); use inert material for condenser; do not use resistor and do not use filter between the subject and the condenser (61).	For proteomics, it is difficult to standardize; very dilute specimens (<1 μg/ml protein).
Lung biopsy or transbronchial biopsy	After obtaining samples they are snap frozen in liquid nitrogen; 1–5 g of tissue is processed using a bead-based tissue maceration method (such as a genogrinder). Once this is done, one can use the usual preparation metabolomics methods for NMR, GC-MS, or LC-MS as used for liquid samples.	
Bronchial wall brushings	Again, after obtaining samples they are snap frozen in liquid nitrogen. It is best if 1–5 g or the equivalent of 10,000 cells are processed using a bead-based tissue maceration method (such as a genogrinder). Once this is done, one can use the usual preparation metabolomics methods for NMR, GC-MS or LC-MS as used for liquid samples.	The sample can be placed in saline and then centrifuged, snap frozen and then processed.
Plasma	Collect blood by direct venipuncture, if possible, into a vacutainer tube containing either EDTA or sodium heparin. Immediately invert the tube several times to ensure mixture with anticoagulant. Within 30 min of blood collection, centrifuge balanced tubes (15 min at 1,300 × *g*) with no brake to ensure proper plasma separation. Refrigeration before or during centrifugation is recommended for metabolomic studies but not recommended for proteomic studies. After centrifugation, the blood should be separated into 3 visible layers, the upper layer is generally clear and pale yellow in color and is the plasma. The second, thin, whitish layer sits at the interface between the plasma and the red blood cells, and is called the buffy coat. The third or bottom layer is dark red and consists of red blood cells. Carefully collect the plasma layer, aliquot and freeze (−80°C) in cryovial.	For proteomics, can consider protease inhibitors in blood draw tube (e.g., BD P100; BD Biosciences, San Jose, CA); in general this is not recommended for all tubes. Recommend heparin or EDTA (not for NMR), not citrate-containing tubes, for metabolomics studies.
Serum	Collect blood by direct venipuncture, if possible, into a vacutainer tube (no additive). Allow the blood to clot at room temperature for at least 30 minutes then centrifuge the balanced vacutainer tube(s) (15 min at 1,300 × *g*) with low brake. Remove the serum, being careful not to disturb the clot at the bottom of the tube. Aliquot and freeze serum (−80°C) in cryovial.	For proteomics, consider the addition of protease inhibitors in blood draw tube (e.g., BD P100); in general, this is not recommended for all tubes and is not recommended for metabolomics.

(Continued)

Table 1. (*Continued*)

Sample Type	Preparation (Recommended)	Comments
Whole blood (62)	Collect blood by direct venipuncture, if possible, into a vacutainer tube containing sodium heparin. Immediately place tube into an ice water bath, aliquot into tubes and flash freeze in liquid nitrogen as soon as possible but preferably within 30 minutes of collection. Store (−80°C or in liquid nitrogen) until the time of assay.	
Urine (63)	Collect fresh urine via clean catch, centrifuge (1,000–3,000 × g for 5 min at 4°C) to precipitate the cells. Collect the supernatant, aliquot and freeze (−80°C) until the time of assay.	For metabolomics, consider the addition of sodium azide to the collection cup to retard bacterial growth.

Definition of abbreviations: BALF = bronchoalveolar lavage fluid; DTT = dithiothreitol; EDTA = ethylenediaminetetraacetic acid; GC = gas chromatography; LC = liquid chromatography; MS = mass spectroscopy; NMR = nuclear magnetic resonance.
*Sample handling and storage: transport samples on ice, avoid letting samples warm to room temperature, aliquot in useable sample volumes to avoid future freeze and thaw cycles, freeze immediately (−80°C) until the time of assay. For long-term storage (>6 mo), consider storing samples in liquid nitrogen.

Invasive lung samples are not often readily accessible, so blood (whole blood, serum, plasma), urine, and isolated cells (e.g., airway epithelial brushings or alveolar macrophages) are often used to indirectly study the lung proteome and metabolome. Although these are more readily accessible, their relevance to lung disease is often less clear.

Shared Resources

Dr. Shankar Subramaniam (University of California, San Diego, CA) discussed the challenges of shared databases, repositories, and software for metabolomics studies. The National Institutes of Health (NIH, Bethesda, MD) supports resources to organize and store national and international metabolomics data and analysis tools through the Metabolomics Workbench program (http://www.metabolomicsworkbench.org//nihmetabolomics/index.html), housed and managed by the University of California San Diego Supercomputer Center. In addition, the site is a resource for analytical standards important for confirming metabolite identities and lipid map classifications. On the horizon is the expected publication of the NIH-sponsored Ring Trial in Metabolomics, in which several metabolomics laboratories across the United States assayed technical replicates of samples to assess standardized processes, analyte detection, and data reproducibility across centers.

Data Integrity

Dr. Arthur Moseley (Duke University, Durham, NC) discussed the importance of project-to-project and laboratory-to-laboratory reproducibility in proteomics and metabolomics. Although not unique to the lung, the lack of standardization in proteomics and metabolomics analyses presents unique challenges to data integrity. There are strategies to improve standardization. As an example, the plate-based targeted metabolomic platform (Biocrates Absolute IDQ p180; Innsbruck, Tirol, Austria) uses internal standards and calibration curves for precise metabolite quantitation, and has been validated across all major MS vendors. Similarly, quality control pools containing a mixture of study-specific samples and reference standards (e.g., human plasma from Golden West Biologicals, Inc., Temecula, CA) can be used to measure intra- and interstudy reproducibility. These approaches can help to overcome batch effects, and should ensure that identical results can be achieved across laboratories and instrument platforms. However, ultimately, the harmonization of analytical and quality control methods will improve and ensure metabolomic and proteomic data integrity (10, 11).

Statistical Approaches

Careful analysis of complex data is essential to fully capture all potential opportunities to explore biological systems and disease. Although proteomics and metabolomics enable accurate detection and quantification in an unbiased manner, there are some unique statistical challenges in assessing these data. Dr. Katerina Kechris (University of Colorado, Denver, CO) discussed how untargeted MS, by definition, does not include chemical standards and, therefore, measurements reflect relative abundances. When studies are completed over multiple experiments, there is frequently drift in retention time (RT) and sensitivity requiring batch correction by using methods, such as Combat or Remove Unwanted Variation (12–14). Metabolites, in particular, tend to be highly correlated within a class. This requires class analysis or the use of dimension reduction through principal component analysis, partial least squares projection to latent structures, clustering, or other multivariate methods (15). Correction for multiple comparisons using methods, such as the false discovery rate, is essential, because many proteins and metabolites can be simultaneously tested (16). In addition, proteins and metabolites are most often not normally distributed and require data transformations (e.g., logarithmic) to normalize the data. It is also important to adjust for covariates, such as age, sex, and smoking history, that can influence metabolite and protein expression. Finally, unique challenges include handling missing values, mass spectra acquisition, identification, and multiple sources of variability (17).

Current State of Metabolomics of Lung Diseases

The fundamental premise in metabolomics is that changes, whether physiological or pathological, cause alterations of the metabolome that are detected as variations

Table 2. Metabolomics biomarker publications in acute respiratory distress syndrome, chronic obstructive pulmonary disease, and asthma

Year	Disease	Metabolic Profiling Approach	Candidate Biomarkers	Reference
1998	ARDS	Untargeted GC-MS exhaled breath	Isoprene	64
2011	ARDS	NMR of Plasma	Sphingomyelin	65
2013	ARDS	Untargeted LC-MS BAL	Higher levels of many amino acids and glycolysis products; lower lipid intermediates	66
2014	ARDS	GC-MS of exhaled breath	Octane, acetaldehyde, 3-methylheptane	67
2017	ARDS	GC-MS of undiluted pulmonary edema	Subphenotype with widespread metabolic differences	68
2010	COPD	NMR of urine	Trigonelline, hippurate and formate	69
2012	COPD	Targeted LC-MS/MS of plasma	Glutamine, aspartate and arginine	70
2012	COPD	NMR and LC-MS of serum	Branched-chain amino acids, glutamine and 3-methylhistidine	71
2013	COPD	NMR of serum and urine	Branched-chain amino acids, glycerol-phosphocholine, 1-methylnicotinamide, creatinine, lactate, acetate, ketone bodies, carnosine, m-hydroxyphenylacetate, phenylacetyglycine, pyruvate and α-ketoglutarate	72
2013	COPD	NMR of EBC	Acetate and 1-methylimidazole	73
2013	COPD	LC-MS of plasma	Ceramides	74
2015	COPD	LC-MS of plasma	Sphingolipids	75
2016	COPD	LC-MS of serum	Myoinositol, glycerophopshoinositol, fumarate, cysteinesulfonic acid, a modified version of fibrinogen peptide B	76
2009	Asthma	MS	Adenosine	77
2011	Asthma	NMR	Succinate	78
2011	Asthma	NMR	Threonine	78
2011	Asthma	NMR	Trans-aconitate	78
2012	Asthma	LC-MS	Urocanic acid	79
2012	Asthma	MS	Adenosine	80
2013	Asthma	NMR	Formate	81
2013	Asthma	NMR	Methanol	81
2013	Asthma	NMR	Arginine	81
2013	Asthma	LC-MS	Acetate	81
2013	Asthma	GC-MS	1,4-dichloro-benzene	82
2013	Asthma	GC-MS	1,4-dichloro-benzene	83
2013	Asthma	GC-MS	2,4-dimethyl-1-heptene	82
2014	Asthma	NMR	Formate	84
2014	Asthma	GC-MS	Hippurate	85
2014	Asthma	NMR	Hippurate	84
2014	Asthma	NMR	Methanol	84
2014	Asthma	GC-MS	Alanine	85
2014	Asthma	NMR	Alanine	84
2014	Asthma	NMR	Arginine	84
2014	Asthma	NMR	Phenylalanine	84
2014	Asthma	GC-MS	Threonine	85
2014	Asthma	NMR	Urocanic acid	84
2014	Asthma	NMR	Trans-aconitate	85
2014	Asthma	NMR	Adenosine	84
2014	Asthma	GC-MS	Acetate	85
2014	Asthma	GC-MS	Acetate	86
2014	Asthma	NMR	Acetate	84
2014	Asthma	GC-MS	2,4-dimethyl-1-heptene	86
2015	Asthma	GC-MS	Succinate	87
2015	Asthma	GC-MS	Phenylalanine	87
2015	Asthma	MS	Adenosine	88
2016	Asthma	LC-MS	Threonine	89

Definition of abbreviations: ARDS = acute respiratory distress syndrome; BAL = bronchoalveolar lavage; COPD = chronic obstructive pulmonary disease; EBC = exhaled breath condensate; GC = gas chromatography; LC = liquid chromatography; MS = mass spectroscopy; NMR = nuclear magnetic resonance.

in metabolite concentrations. Metabolomics studies are generally based on either of two approaches: a global profiling strategy (i.e., untargeted metabolomics) or a selective measurement strategy (targeted metabolomics). The untargeted method is useful as an initial evaluation to shortlist key metabolites with distinct alterations. This is especially helpful if there is no prior knowledge about the metabolic disturbances involved. On the other hand, targeted metabolomics is useful for measuring specific metabolites of particular relevance to the condition being researched.

Most published metabolomic studies have focused on three major lung diseases: acute respiratory distress syndrome (ARDS), asthma, and chronic obstructive

pulmonary disease (COPD), as reviewed by Dr. Kathleen Stringer (University of Michigan, Ann Arbor, MI) and summarized in Table 2. These illnesses are notoriously heterogeneous, and an absence of predictive and prognostic biomarkers has stalled the identification of new drug targets and hindered the implementation of precision medicine. Recent studies demonstrate that metabolomics enhances the diagnostic accuracy of ARDS, COPD, and asthma (18–21). Metabolomics is particularly useful in the challenging area of respiratory disease endotyping and phenotyping (20), and in the discrimination of individuals with differing levels of severity and exacerbations, allowing for more targeted treatment regimens (20, 21). These studies highlight the potential of metabolomics to "deep phenotype" pulmonary diseases. However, more work is needed to confirm and validate links between metabolic changes, clinical phenotypes, and biological processes to further understanding about disease pathogenesis, and ultimately to drive drug discovery and achieve precision pulmonary medicine.

Challenges in Metabolomics

Major challenges in metabolomics include data processing, identification and validation of metabolites, and data visualization. Importantly, because no one analytical

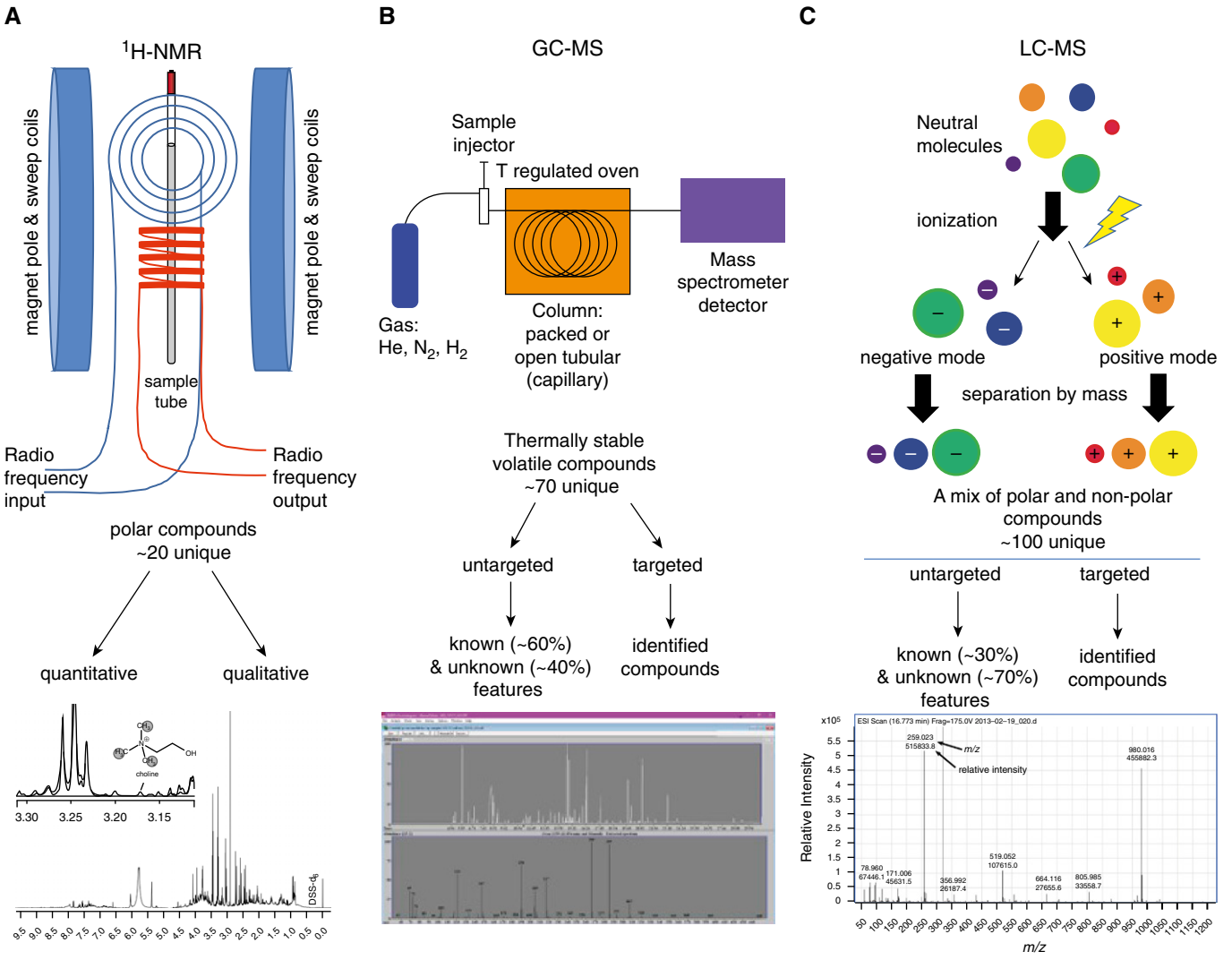

Figure 3. The most commonly used analytical platforms for metabolomics are: (*A*) proton (^{1}H) nuclear magnetic resonance (NMR); (*B*) gas chromatography (GC)–mass spectroscopy (MS); and (*C*) liquid chromatography (LC)-MS. (*A*) NMR is ideal for the detection of polar compounds like amino acids and for smaller molecular weight (⩽100 Da) metabolites that LC-MS can miss. NMR is routinely quantitative when an internal standard, such as 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), is added to the sample. (*B*) GC-MS most often requires the derivatization of volatile compounds that are separated by a gas carrier phase and elute based on retention time in the column. After ionization, compounds are detected by MS. The graphic printout shows a typical serum readout of abundance versus time (*top*) and abundance versus mass/charge ratio (*bottom*). (*C*) For LC-MS metabolomics, molecules are ionized, typically by electrospray ionization, and the resulting positive and negative ions are detected by MS. This results in a mass-to-charge ratio (*m/z*) versus relative peak intensity graphical representation of the data. More details about the advantages and disadvantage of each approach can be found in Table 2. By K. Murray (Kkmurray) (Own work) [GFDL (http://www.gnu.org/copyleft/fdl.html), CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/) or CC BY-SA 2.5-2.0-1.0 (http://creativecommons.org/licenses/by-sa/2.5-2.0-1.0)], via Wikimedia Commons.

Table 3. Examples of strengths and weaknesses of nuclear magnetic resonance and MS methods used in metabolomics studies (6, 32)

Technique	Strengths	Weaknesses
NMR		
	Nondestructive technique (several analyses can be conducted on the same samples)	Low sensitivity (only metabolites with μM concentrations can be reliably detected)
	Can be done with minimal sample preparation (addition of deuterated solvent and a buffer)	Overlap in peaks and high chemical degeneracy (different metabolites have resonances in the same spectral region)
	Versatility for analyzing metabolites in biofluids, tissues or *in vivo*	
	Reproducible and repeatable	
	Routinely quantitative	
	Detects small molecular weight (<100 Da) compounds that LC-MS may miss	
GC-MS		
	High-resolution capacity	Thermostable and volatile and nonpolar metabolites may require derivatization
	High spectral resolution	High molecular weight analytes
	Very sensitive	Low reproducibility (within and across labs)
	Low limit of detection	Fragmentation in MS
	Required technical skill	Extensive sample preparation steps
	High mass accuracy to detect compounds	Not routinely quantitative
	Reproducible retention time	Possible variation due to sample preparation
	Highly developed compound libraries	Matrix effect
	Linear dynamic range	Compound degradation (high temperature)
	High separation efficiency	Destructive to sample
LC-MS		
	Short separation time	Ideal for nonvolatile compounds
	High resolution	High solvent consumption and lower separation power
	Very sensitive	Low reproducibility (within and across labs)
	Reasonable robustness	Ionization of metabolites
	Selective	Not routinely quantitative
	High mass accuracy to detect compounds	Lower retention time reproducibility with different chromatography systems or columns versus GC-MS
	Simple sample preparation	Destructive to sample
	Detects a wider range of chemical classes of compounds than GC-MS	No universal compound library
	Can use a small sample volume (~100 μl)	

Definition of abbreviations: GC = gas chromatography; LC = liquid chromatography; MS = mass spectroscopy; NMR = nuclear magnetic resonance.
Adapted from Reference 32.

platform captures the entire metabolome, there are a number of analytical platforms that can be used (Figure 3), each of which has its own advantages and disadvantages (Table 3). An overall consensus was a call for greater synchrony of research methods to enable development of "big data" resources shared across institutions that can be applied to lung diseases.

Metabolite Identification

MS data are often reported as features, which represent analytes with discrete mass-to-charge ratios (m/z) and RTs. Dr. Richard Reisdorph (University of Colorado, Denver, CO) discussed the importance of metabolite identification, rather than a simple description of these features, as many represent breakdown products of known metabolites. Differential features are identified, and extensive follow-up work is needed to identify the specific compound. This is particularly challenging for lipids, as large numbers of lipids have very similar or identical m/z and RT. Gas chromatography annotation libraries are more developed than liquid chromatography (LC) libraries, but, to date, a standard metabolite library for untargeted LC– tandem MS (MS/MS) is not readily available for all researchers. This limits downstream pathway analysis, which can only be as good as the annotation of known metabolites. To move the field forward, a description of the annotation confidence is important, ranging from low confidence (e.g., mass of analyte matched to a database) to high (MS/MS spectrum matches to an MS/MS library) to highest "gold standard" (confirmed with purchased standards, RT, and MS/MS spectrum). To this end, there is a need for analytical reference standards for metabolomics (http://www.metabolomicsworkbench.org/standards/index.php). Data validation, as reviewed by Dr. Nichole Reisdorph (University of Colorado, Denver, CO), with a targeted assay in an independent population is the gold standard, but requires a substantial investment of time to confirm metabolite identifications.

Data Processing

Data processing needs, which differ significantly for nuclear magnetic resonance (NMR) spectroscopy and MS approaches, were discussed by Dr. Dean Jones (Emory University, Atlanta, GA). The analysis of NMR spectra can be challenging due to peak overlap, but analysis can be optimized with the use of software that permits the identification and quantification of metabolites (6). Spectral peak overlap is more readily deconvoluted for LC-MS than for NMR. For both sources of data, chemometric methods can be used in which peak or feature signals

Table 4. Publicly available tools for metabolite identification*

Program	Website	Description	Reference
MetLin	https://metlin.scripps.edu/index.php	METLIN includes >900,000 molecules ranging from lipids, steroids, plant and bacteria metabolites, small peptides, carbohydrates, exogenous drugs/metabolites, central carbon metabolites, and toxicants. Over 14,000 metabolites have been individually analyzed and another 200,000 have *in silico* tandem mass spectroscopy data.	90
HMDB	http://www.hmdb.ca/	The Human Metabolome Database (HMDB) contains detailed information about small molecule metabolites found in the human body. It is designed to contain or link three kinds of data: (*1*) chemical data; (*2*) clinical data; and (*3*) molecular biology/biochemistry data. The database contains >42,000 metabolite entries, including chemical/clinical/enzymatic data, and links to proteins and other databases (KEGG, PubChem, MetaCyc, ChEBI, PDB, UniProt, and GenBank) and a variety of structure and pathway viewing applets. The HMDB database supports extensive text, sequence, chemical structure, and relational query searches. Four additional databases, DrugBank, T3DB, SMPDB, and FooDB, are also part of the HMDB suite of databases.	91

*Information about instrument specific software packages can be found at each respective manufacturer's website.

represent signatures of biologic response without knowledge of or the quantification of specific compounds, as a means of an initial survey of the data (22). After extraction of signals for individual metabolites, many of the downstream methods for NMR spectroscopy data and MS metabolomics data are the same. In general, these involve common univariate and multivariate biostatistical methods, as well as a range of bioinformatics approaches with data clustering and correlation analyses, and also encompassing knowledgebase searching and integration of multi-omics data (23). Tools available for metabolite identification, spectral analysis, and the assessment of data integrity and data interpretation are summarized in Tables 4–6, respectively, including available reference standards, through the Metabolomics Workbench website (http://www.metabolomicsworkbench.org/).

Mapping, Modeling, and Visualization of Metabolomics Data

Dr. Alla Karnovsky (University of Michigan Ann Arbor, MI) highlighted the recent advances in analytical technologies, as metabolomic datasets are becoming increasingly large and complex (6). With respect to lung health and disease, it is important to put the observed metabolite changes into biological and disease context. A growing number of tools and resources for pathway mapping and enrichment analysis of metabolomics data, as well as data-driven network analysis methods and software, are being developed, including multiple open source options, such as ConceptMetab (24), Metscape (25), and metabolite set enrichment analysis (26) (Table 6).

Lipidomics

Dr. Oliver Fiehn by (University of California–Davis, Davis, CA) discussed the rapidly expanding field of lipidomics. Peak selection and identification are particularly challenging in lipidomics. The open-source software, MS-DIAL (http://prime.psc.riken.jp/Metabolomics_Software/MS-DIAL/index2.html; Table 5), de-convolutes overlapping spectra and aids in the identification and quantification of compounds (Table 4) (27). For reversed-phase LC-MS lipid identification, MS-DIAL uses data converted to an axon binary file format and the Lipidblast library (28), a computer-generated *in silico* MS/MS spectral library of 26 lipid compound classes, including phospholipids, glycerolipids, bacterial lipoglycans, and plant glycolipids.

Cutting-Edge Metabolomics

Advanced Imaging Using MS Technology

Dr. Richard Caprioli (Vanderbilt University, Nashville, TN) described how matrix-assisted laser desorption/ionization imaging MS produces molecular maps of peptides, proteins, lipids, and metabolites present in intact tissue sections (29, 30). This technique employs desorption of molecules by direct laser irradiation to map the location of specific molecules from fresh-frozen and formalin-fixed tissue sections without the need of target-specific reagents, such as antibodies. Molecular images of this nature are produced in specific m/z values, or ranges of values. Each specimen gives rise to many hundreds of specific molecular images from a single raster of the tissue. In a complementary approach, where only discrete areas within the tissue are of interest, a histology-directed approach that integrates MS and microscopy has been developed. Thus, mass spectra are collected from only selected areas of cells within the tissue after laser ablation and analysis. Clinically relevant studies include advanced diabetic nephropathy involving both proteins and lipids (29). In addition, imaging MS has been applied to drug targeting and metabolic studies, both in specific organs and also in intact whole animal sections after drug administration. These techniques, though promising, have yet to be applied to the lung outside of lung cancer.

The Future of Lung Metabolomics

Dr. Brent Winston (University of Calgary, Calgary, AB, Canada) discussed the key strategies for advancing lung metabolomics. Data sharing and adopting standards of practice (SOPs) are key for future studies. In addition, a focus on targeted mechanistic studies will enable the field to move beyond hypothesis–agnostic discovery science. The

Table 5. Publicly available and commercial software/tools for spectral processing and analysis for metabolomics*

Program	Website	Description	Reference
LC-MS			
XMSanalyzer	http://clinicalmetabolomics.org/init/default/index	xMSanalyzer is a package of utilities for data extraction, quality control assessment, detection of overlapping and unique metabolites in multiple datasets, and batch annotation of metabolites. The program was designed to integrate with existing packages, such as apLCMS and XCMS, but the framework can also be used to enhance data extraction for other LC-MS data software.	92
Hybrid apLCMS	http://web1.sph.emory.edu/aplcms/	The R package apLCMS is designed for the processing of LC-MS based metabolomics data. It starts with a group of LC-MS files in the same folder, and generates a table with features in the rows and intensities in the columns. Data can be analyzed using unsupervised methods (*de novo* peak detection from the data) or hybrid analysis, combines *de novo* peak detection with existing knowledge.	93
Camera	http://www.bioconductor.org/packages/release/bioc/html/camera.html	Collection of annotation related methods for MS data.	94
MS-DIAL	http://prime.psc.riken.jp/Metabolomics_Software/MS-DIAL/index2.html	Deconvolutes mass spectra. Identification and quantification of small molecules.	27
XCMS	https://metlin.scripps.edu/xcms/	The XCMS software reads and processes LC-MS data stored in netcdf, mzXML, mzData, and mzML files. It provides methods for feature detection, nonlinear retention time alignment, visualization, relative quantization and statistics. XCMS is capable of simultaneously preprocessing, analyzing, and visualizing the raw data from hundreds of samples. XCMS is freely available under an open-source license.	95
MetSign	http://metaopen.sourceforge.net/metsign.html	For LC-MS–based metabolomics data, MetSign provides a set of data preprocessing algorithms for peak detection and peak list alignment. For spectrum deconvolution, peak picking is achieved at the selected ion chromatogram (XIC) level.	96
MZmine 2	http://mzmine.github.io/	MZmine 2 is an open-source software for MS data processing, with the main focus on LC-MS data, with the goal to provide a user-friendly, flexible and easily extendable software with a complete set of modules covering the entire LC-MS data analysis workflow.	97
MAIT	https://www.bioconductor.org/packages/release/bioc/html/mait.html	The MAIT package contains functions to perform end-to-end statistical analysis of LC-MS Metabolomic Data. Special emphasis is put on peak annotation and in modular function design of the functions.	98

(Continued)

Table 5. (*Continued*)

Program	Website	Description	Reference
GC-MS			
Metabolite Detector	http://metabolitedetector.tu-bs.de/	Metabolite Detector is a QT4 based software package for the analysis of GC-MS based metabolomics data. The software is especially intended for the analysis of high resolution GC-MS chromatograms which accumulate during high throughput based metabolomics experiments. For this purpose, Metabolite Detector features a nearly fully automated data analysis pipeline starting from the raw GC-MS data and ending in a principal component analysis. Currently, Metabolite Detector is able to import GC-MS data in NetCDF and FastFlight format.	99
NMR			
Automics	http://www.softpedia.com/get/science-cad/automics.shtml	Automics is a highly integrated platform for NMR-based metabonomics or metabolomics spectral processing and data analysis. It is targeted to aid researchers for processing high dimensional NMR spectroscopic data.	100
Bayesil	http://bayesil.ca/	Bayesil is a web system that automatically identifies and quantifies metabolites using 1D ^{1}H NMR spectra of ultrafiltered plasma, serum or cerebrospinal fluid. Bayesil performs all spectral processing steps, then deconvolutes the resulting NMR spectrum using a reference spectral library, which here contains the signatures of more than 60 metabolites. This deconvolution process determines both the identity and quantity of the compounds in the biofluid mixture.	101
Chenomx	http://www.chenomx.com/software/	A commercial software, Chenomx NMR Suite is an integrated set of tools for identifying and quantifying metabolites in NMR spectra.	102
FOCUS	http://www.urr.cat/focus/	FOCUS is a complete workflow for processing NMR metabolomics data that provides efficient methodological advances both on peak alignment and metabolite identification.	103
Speaq	https://cran.r-project.org/web/packages/speaq/index.html	Suite of informatics tools for the quantitative analysis of NMR metabolomic profile data. The core of the processing cascade is a peak alignment algorithm, called hierarchical CluPA. The algorithm aligns a target spectrum to the reference spectrum.	104
MetaboLab	http://metabolab.uk/	MetaboLab is a software package for NMR processing written in MATLAB. It contains of a series of processing algorithms for 1D, 2D and 3D processing, including apodization functions, linear prediction, Fourier transformation, and baseline correction. It also includes routines for wavelet denoising of spectra. A simple interface is available to set up processing parameters. The current version supports Bruker and Varian file formats.	105

(Continued)

Table 5. (*Continued*)

Program	Website	Description	Reference
rNMR	http://rnmr.nmrfam.wisc.edu/	rNMR is an open source software package for visualizing and interpreting one and two-dimensional NMR data. rNMR is specifically designed for high-throughput assignment and quantification of small molecules. As a result, rNMR supports extensive batch manipulation of plotting parameters and has numerous tools for expediting repetitive resonance assignment and quantification tasks.	106
MetabMiner	http://wishart.biology.ualberta.ca/metabominer/	Easy-to-use software tool, MetaboMiner, aids in rapid and efficient metabolite identification from complex mixtures using 2D NMR spectroscopy.	107
BATMAN	http://batman.r-forge.r-project.org/	BATMAN is an R package for estimating metabolite concentrations from NMR spectral data using a specialized MCMC algorithm. It deconvolutes peaks from 1D NMR spectra, automatically assigns them to specific metabolites from a target list and obtains concentration estimates. Uses Bayesian analysis and a MCMC algorithm to obtain concentration estimates with reduced error.	108
MetaboHunter	http://www.nrcbioinformatics.ca/metabohunter/about.php	MetaboHunter is a web server application for semiautomatic assignment of 1D NMR spectra of metabolites. MetaboHunter provides tools for metabolite identification based on spectra or peak lists with three different search methods and with possibility for peak drift in a user defined spectral range. The assignment is performed by comparison with two major publically available databases (HMDB, MMCD) of NMR metabolite standard measurements.	109

Definition of abbreviations: 1D–3D = one to three dimensional; CluPA = cluster-based peak alignment; GC = gas chromatography; LC = liquid chromatography; MCMC = Markov chain Monte Carlo; MS = mass spectroscopy; NMR = nuclear magnetic resonance.
*Information about instrument specific software packages can be found at each respective manufacturer's website.

overall objectives of "omics" technology in human disease (31) are to:

- Facilitate subcategorizing into specific endotypes to target therapeutic interventions
- Improve triage decisions
- Provide a means to follow response to therapy
- Establish new therapeutic targets
- Provide ways to identify patients amenable to tailored therapies

The use of existing samples from major lung-related studies is highly desirable, but is only possible if samples are collected with proper SOPs and appropriate storage of samples in an internationally recognized fashion (32). This will require cooperation from major funding agencies for all phase 1, 2, and 3 studies involving the respiratory system. For example, the Precision Medicine Initiative (https://ghr.nlm.nih.gov/primer/precisionmedicine/initiative) and the Million Veteran Program (http://www.research.va.gov/mvp/) plan to collect samples from one million subjects, each using uniform methodology. To do this well would include the development and harmonization of SOPs, storage methods, sample annotation, data sharing, and exploration of age-related storage material degradation (33, 34).

Current State of Proteomics in Lung Diseases

Proteomic Approaches in Studying Lung Disease

MS has revolutionized the study of proteins, as it allows the measurement of hundreds to thousands of proteins in complex systems in a very precise and reproducible manner. In the first presentation of the proteomics breakout session, Dr. Chris Wendt (University of Minnesota, Minneapolis, MN) discussed the current state of proteomics in lung disease, specifically focusing on ARDS, COPD, and idiopathic pulmonary fibrosis (Table 7). Many of these studies used two-dimensional gel electrophoresis and/or a combination of LC and MS, common techniques for the identification of disease biomarkers and disease-related signaling pathways using gene ontology analysis (35–39). Specific advances and resources in databases, repositories, methods, and protocols have greatly advanced the field of proteomics (Table 8); however, challenges remain.

Table 6. Publicly available and commercial software/tools for analysis and interpretation of metabolomics data*

Program	Website	Description	Reference
Metaboanalyst	http://www.metaboanalyst.ca/	A comprehensive tool suite for metabolomics data analysis. There are additional tools embedded with MetaboAnalyst including MSEA.	110
MetScape for Cytoscape	http://metscape.ncibi.org/	MetScape is a plugin for Cytoscape. It provides a bioinformatics framework for the visualization and interpretation of metabolomic and expression profiling data in the context of human metabolism. It allows users to build and analyze networks of genes and compounds, identify enriched pathways from expression profiling data, and visualize changes in metabolite data. Integrates data from KEGG and EHMN.	25
ConceptMetab	http://conceptmetab.med.umich.edu/	ConceptMetab, is a tool for mapping and exploring the relationships among 16,069 biologically defined metabolite sets developed from Gene Ontology, KEGG and Medical Subject Headings. It uses both KEGG and PubChem compound identifiers, and based on statistical tests for association.	24
Metabolome Express	https://www.metabolome-express.org/	MetabolomeExpress houses both private and public uncurated repositories to process, interpret and share GC-MS metabolomics datasets, as well as a quality-controlled database of highly annotated metabolite response statistics submitted by MetabolomeExpress users. The quality-controlled database of metabolite response statistics can be queried to find relevant experiments using tools in the *Database Explorer* and examined in detail using the in-built *Experiment Explorer* which includes integrated tools for raw data visualization, processing and statistical analysis.	111
Astream	http://www.urr.cat/astream/astream.html	AStream is an R statistical software package for the curation and identification of feature peaks extracted from LC-MS metabolomics data. Compounds are identified and subsequently linked to metabolite databases.	112
MetExplore	http://metexplore.toulouse.inra.fr/joomla3/index.php	MetExplore is a free academic service that enables: Importing/storing/sharing genome scale metabolic networks Mapping polyomics data Pathway enrichment Visualizing networks Mining networks based on data and network structure Computing fluxes	113
IMPaLA	http://impala.molgen.mpg.de/	Pathway overrepresentation and enrichment analysis with expression and/or metabolite data.	114
PaintOmics	http://bioinfo.cipf.es/paintomics/	Paintomics is a web tool for the integration and visualization of transcriptomics and metabolomics data.	115
MAVEN	http://genomics-pubs.princeton.edu/mzroll/index.php	The aim of this software package is to reduce complexity of metabolomics analysis through an interface that enables exploring and validating metabolomics data. The program features multifile chromatographic aligner, peak-feature detector, isotope and adduct calculator, formula predictor, pathway visualizer, and isotopic flux animator. Data from both triple quadropole and full spectrum instruments is supported.	116
MAIT	https://www.bioconductor.org/packages/release/bioc/html/MAIT.html	The MAIT package contains functions to perform end-to-end statistical analysis of LC-MS metabolomics data. Special emphasis is put on peak annotation and in modular function design of the functions.	98

(Continued)

Table 6. (*Continued*)

Program	Website	Description	Reference
VANTED	https://immersive-analytics.infotech.monash.edu/vanted/	VANTED is a Java-based extendable network visualization and analysis tool with focus on applications in the life sciences. It allows users to create and edit networks, as well as mapping experimental data onto networks. Experimental datasets can be visualized on network elements as graphical charts to show time series data or data of different treatments, as well as environmental conditions in the context of the underlying biological processes.	117
METABOX	http://kwanjeeraw.github.io/metabox/	An R-based web applications for data processing, statistical analysis, integrative visual exploration and functional analysis with several approaches (such as functional class scoring, overrepresentation analysis and WordCloud generation).	118
Mummichog	http://clinicalmetabolomics.org/init/default/mummichog_manual	Mummichog is a Python program for analyzing data from high-throughput, untargeted metabolomics. It leverages metabolic networks to predict functional activity directly from feature tables, bypassing metabolite identification. The features include: (*1*) computing significantly enriched metabolic pathways; (*2*) identifying significant modules in the metabolic network; (*3*) visualization of top networks in web browser; (*4*) visualization that also plugs into Cytoscape; (*5*) tentative annotations; and (*6*) metabolic models for different species through plugins	119

Definition of abbreviations: LC = liquid chromatography; MS = mass spectroscopy; MSEA = metabolite set enrichment analysis
*Tools such as XCMS and MetSign in Table I include data analysis and visualization options.

Challenges in Quantitative Proteomics

Challenges remain in the ability to accurately quantify changes in protein abundance. Both label and label-free strategies exist for proteome quantitation (Table 9). Dr. Alexey Nesvizhskii (University of Michigan, Ann Arbor, MI) discussed the application of label-free quantitative methods, including software tools for data-dependent acquisition with quantitation by spectral counting or ion abundance (e.g., QSpec/QProt) (40, 41), as well as data-independent acquisition (DIA; e.g., DIA-Umpire) (42, 43) (Figure 4). Spectral counting is defined as the number of MS/MS sequencing attempts made on a precursor (i.e., intact peptide) during a single LC-MS/MS analysis, whereas intensity-based quantitation is the measurement of the area under the curve of each precursor in a sample. Intensity-based methods have greater accuracy, but both methods suffer from missing data across replicate analyses. The missing data problem can be mostly eliminated by aligning (or matching) of precursors across batched analyses (44).

Data-Independent Methods for Quantitative Proteomics

DIA methods use "MS2-based" quantitation of the ions that are produced by MS/MS fragmentation (45) and can have greater selectivity (and greater signal-to-noise) than precursor/MS1-based quantitation. DIA also allows for matching of each sample to an external library of MS/MS spectra, thus largely eliminating the missing data problem. DIA has shown promise for analysis of biofluids, including BALF (46). DIA-Umpire software can furthermore extract MS/MS spectra for peptide/protein identification using conventional database searching (42, 43).

Quantifying the Secretome Proteome

LC-MS/MS allows the identification and quantification of hundreds of proteins in cellular secretions (secretomes) of airway cells. This includes the analysis of airway cells *in vitro*, which allows for the identification of key mechanistic biochemical insights, and thus plays a pivotal role in translational lung research. However, there remain challenges in quantifying the secretome, as discussed by Dr. Kristy Brown (Children's National Health System, Washington, D.C.) and Dr. Mehmet Kesimer (University of North Carolina–Chapel Hill, Chapel Hill, NC).

Dr. Brown studies the altered secretome of human bronchial epithelial cells (HBECs) obtained from patients with cystic fibrosis (47, 48). Here, she introduced the concept of stable isotope labeling of amino acids in cell culture (SILAC), which uses cells grown in isotopically labeled amino acids (typical ^{13}C- and ^{15}N-labeled arginine, lysine, and/or leucine) to synthesize the "heavy" forms of proteins that can be mixed with their "light" counterparts before trypsinization and/or peptide/protein fractionation (Figure 5). Advantages of SILAC include accuracy in quantitation and the ability to simultaneously quantify and differentiate two proteomes. This method is also well suited to studying proteome-wide protein synthesis and decay, and a "SuperSILAC" mix, such as that generated from ARDS secretions, can be used as a common reference standard (49). This approach could have general utility for airway secretomics; however, SILAC techniques usually require numerous passages with

Table 7. Proteomic biomarker publications in acute respiratory distress syndrome, chronic obstructive pulmonary disease, and idiopathic pulmonary fibrosis

Year	Disease	Proteomics Approach	Biomarkers	Reference
2004	ARDS	2D Gel with MALDI-TOF MS	Surfactant protein A	9
2006	ARDS	SELDI-TOF, 2D Gel with MALDI-TOF MS	Apolipoprotein A1, S100 calcium-binding proteins A8 and A9	120
2006	ARDS	Label-free LC-MS/MS	IGFBP-3, IGF	121
2008	ARDS	DIGE with MALDI MS	22 proteins, including S100A8/9, HSPG2, FTL	122
2013	ARDS	iTRAQ label with MS/MS	5 apolipoproteins, complement factor H, haptoglobin, serotransferrin, A_1-antitrypsin, antichymotrypsin, CRP, amyloid A, leucine-rich 2 glycoprotein	123
2013	ARDS	2D Gel with MALDI-TOF MS	27 proteins, including S100A8/9, IL1RA, γ-actin	124
2014	ARDS	iTRAQ label with MS/MS	Plasminogen, factor 12, antithrombin III, ceruloplasmin, S10049, thioredoxin	36
2016	ARDS	iTRAQ label with MS/MS	DMBT1	125
2005	COPD	SELDI MS	Neutrophil defensins 1 and 2, calgranulin A and B, salivary proline-rich peptide, club cell secretory protein	126
2008	COPD	2D Gel with MALDI-TOF MS	RAGE	127
2008	COPD	SELDI-TOF	Serum amyloid	128
2009	COPD	2D Gel with MALDI-TOF MS	MMP-13, thioredoxin-like 2	129
2010	COPD	2D Gel with MALDI-TOF MS	Apolipoprotein A1, lipocalin-1	37
2011	COPD	2D Gel with MALDI-TOF MS	Hsp27, CyPA	130
2011	COPD	DIGE with MALDI MS	2-macroglobulin, haptoglobin, ceruloplasmin, hemopexin	131
2013	COPD	iTRAQ label with MS/MS	Lactotransferrin, HMGB1, α-1 antichymotrypsin, cofilin-1	132
2014	COPD	GeLC-MS/MS	GRP78, soluble CD163, IL1AP and MST1	133
2014	COPD	Label-free MS/MS	423 proteins, ADH1B, ALDH2, and ALDH3A1	38
2015	COPD	Label-free MS/MS	203 proteins, mucin 5AC	134
2015	COPD	Labeled MS/MS	TIMP1, APOA1, C6orf58, BP1FB1	135
2008	IPF	2D Gel with MALDI-TOF MS	Calgranulin B	136
2013	IPF	2D Gel with MALDI-TOF MS	Stress-induced, antiapoptotic, and antifibrotic proteins	137
2013	IPF	2D Gel with MALDI-TOF MS	Plastin 2, annexin A3, calcyclin	138
2014	IPF	Data-independent MS/MS	Osteopontin, MMP7, CXCL7, CCL18	55
2016	IPF	2D Gel with MALDI-TOF MS	22 unique proteins	139

Definition of abbreviations: 2D = two dimensional; ADH = antidiuretic hormone; ALDH = aldehyde dehydrogenase; APOA = apolipoprotein A; ARDS = acute respiratory distress syndrome; BP1FB11 = bacterial/permeability-increasing fold-containing B1; C6orf58 = chromosome 6 open reading frame 58; COPD = chronic obstructive pulmonary disease; CRP = C-reactive protein; CyPA = cyclophilin; DMBT1 = deleted in malignant brain tumors 1; FTL = ferritin light chain; GRP78 = 78 kDa glucose-regulated protein; HMGB1 = high mobility group box 1 protein; Hsp = heat shock protein; HSPG = heparan sulfate proteoglycan; IGF = insulin-like growth factor; IGFBP = insulin-like growth factor-binding protein; IPF = idiopathic pulmonary fibrosis; LC = liquid chromatography; MALDI-TOF = matrix-assisted laser desorption/ionization–time of flight; MMP = matrix metalloproteinase; MS = mass spectroscopy; MST1 = hepatocyte growth factor-like protein; RAGE = receptor for advanced glycation endproducts; TIMP1 = tissue inhibitor of metalloproteinases 1.

media containing the heavy amino acids, limiting its applicability to cells in culture.

Label-Free Quantitative Proteomics

Dr. Kesimer emphasized the utility of label-free quantitation of lung secretions. In a typical label-free proteomic experiment, samples are normalized to total protein before trypsinization, and are intensity normalized during data analysis. He discussed the potential pitfalls of this approach for comparing health to disease (e.g., how large increases in mucin-5B secretion by HBECs were compressed when data were intensity normalized across all samples). Urea has been long used as a normalizing factor in protein quantitation of BALF. It stands to reason that a similar approach (normalization to an "unperturbed" metabolite or protein) might improve quantitative accuracy of HBEC secretomics.

Proteomic Approaches to Cell Signaling

As is the case with epithelial cells, proteomic evaluation of isolated myeloid cells has the virtue of providing focused molecular insights into physiologic events that occur *in vivo* in the multicellular lung. In particular, *ex vivo* culture of sentinel immune cells, such as macrophages, allows for study of time-resolved cell signaling events induced by environmental stimuli, such as bacterial LPS. Dr. Michael Fessler (National Institute of Environmental Health Sciences, NIH) discussed proteomic approaches to cell signaling using primary murine macrophages (50), primary human neutrophils (51), and immortalized macrophage cell lines (52, 53). Whereas phosphoproteomic strategies partnered with chemical inhibitors or RNA interference can be used to map out kinase cascades, subcellular fractionation and immunoprecipitation can permit focused insight into compartmentalization of signaling events within the cell. SILAC has been applied successfully to both primary (e.g., bone marrow-derived) and immortalized macrophages, and can be used in multiple signaling applications, including detecting changes in post-translational modification, localization, and interaction of signaling proteins. SILAC has been particularly valuable in kinetic studies (e.g., measurement of protein turnover) and analysis of signaling events (e.g., phosphorylation). In addition, label-free approaches, such as spectral counting, can be used in a semiquantitative manner to monitor targeted signaling events within the cell, and can also be used to help validate specificity in pulldown assays.

Table 8. Resources for (lung-specific) proteomics database and repositories

Resource	URL	Description
BioGRID	https://thebiogrid.org	Repository of protein–protein interaction (interactome) datasets
Human Protein Atlas	http://www.proteinatlas.org	Database of human protein expression with immunohistochemistry and gene expression data from normal lung parenchyma (140, 141).
Human Proteome Project	https://hupo.org/human-proteome-project	Information resources. International project organized by the HUPO, recently published list of organ system-specific proteins for targeted proteomics (including lung) (142).
Lung Map	http://www.lung-map.org/	Data from NIH-funded Molecular Atlas of Lung Development program, including proteomes of developing mouse lung (143, 144).
MaxQB	http://maxqb.biochem.mpg.de/mxdb	Database of studies from Max Plank Institute, including an in-depth, time-resolved proteomic analysis of the bleomycin model of pulmonary fibrosis (145, 146).
Phosphomouse	https://phosphomouse.hms.harvard.edu	Protein and phosphopeptide data from mouse tissue, including lung (147).
Proteomics DB	https://www.proteomicsdb.org	Database of large-scale proteome studies. Analysis of human cells and tissues, including lung (148).
ProteomeXChange	http://www.proteomexchange.org	Catalog Searchable data from multiple proteomics repositories, including Peptide Atlas and MassIVE, with numerous airway cell datasets.
REPRINT	https://reprint-apms.org/	Resource for Evaluation of Protein Interaction Networks provides a database for the contaminant repository.
General methods and protocols		
Duke Proteomics and Metabolomics Shared Resource	https://genome.duke.edu/cores-and-services/proteomics-and-metabolomics/protocols-reagents	Methods for sample preparation, including tissue lysis, affinity purification and proteolytic digestion
Max Planck Institute (Mann Laboratory)	https://www.biochem.mpg.de/221814/Sample-preparation	Protocols for filter aided sample preparation (FASP) methods
UMN Proteomics Resource and shared protocols	http://cbs.umn.edu/cmsp/protocols	Protocols for sample preparation
UW Proteomics Resource	http://proteomicsresource.washington.edu/methods.php	Methods for sample preparation and data acquisition
UMN Proteomics Resource and shared protocols	http://cbs.umn.edu/cmsp/protocols	Protocols for sample preparation
Open-source software		
DIA-Umpire	http://diaumpire.sourceforge.net/	Software for analysis of swath-type DIA data
Galaxy-P project	https://usegalaxyp.org/	A multi-omics informatics platform for integrative analysis of mass spec–based proteomics, genomic and transcriptomic data.
MacCoss Laboratory	https://sites.google.com/a/uw.edu/maccoss/home/software	Numerous software tools, including PECAN (for searching DIA data) and Percolator (for determining peptide FDR from decoy database searches)
MaxQuant	http://www.biochem.mpg.de/5111795/maxquant	Quantitative proteomics software package supports all main labeling techniques as well as label free.
Motif-X	http://motif-x.med.harvard.edu/	Online tool for generating consensus motifs from proteomic datasets containing post-translationally modified peptides
NIST Mass Spectrometry Data Center	http://chemdata.nist.gov/	Repository of tandem mass spectrum libraries
Skyline	https://skyline.ms	Vendor-neutral software for targeted proteomics
PNNL Omics Software	https://omics.pnl.gov/software	Open-source proteomics software
Seattle Proteomics Tools	http://tools.proteomecenter.org/software.php	Software tools that form the basis for the TPP
STRING	http://string-db.org/	Online tool for visualizing protein interaction networks
Yates Laboratory	http://fields.scripps.edu/yates/wp/?page_id=17	Proteomics software tools, including the search engine Sequest

Definition of abbreviations: DIA = data-independent acquisition; FDR = false discovery rate; HUPO = Human Proteome Organization; NIH = National Institutes of Health; TPP = trans-proteomic pipeline; UMN = University of Minnesota; UW = University of Washington.

Table 9. Overview of quantitation strategies for shotgun proteomic analyses

Analysis Type	Quantification Type	Quantification Method	Description	MS1- or MS2-Based?	Reference
Discovery-based proteomics	Labeled	Isobaric (iTRAQ/TMT)	Tryptic peptides from ⩽10 samples are labeled with isobaric tags containing reporter ions of differing molecular weight. Labeled samples are combined before LC-MS/MS.	MS2 (reporter ion)	36
		Isotopic (SILAC)	Cells are grown in media containing light or heavy (^{13}C,^{15}N-labeled) amino acids. Two states are mixed before trypsin digestion and LC-MS/MS.	MS1 (heavy/light ratio)	49
	Label free	Spectral counting	Estimation of protein abundance based on number of MS/MS sequence attempts are performed on a peptide/protein.	MS1/MS2	43
		Ion abundance or Ion current	AUC analysis used to quantify the expression of peptides/protein identified by LC-MS/MS. MS/MS performed on topN most abundant ions (i.e., data-dependent acquisition). Can use alignment across samples to reduce missing data.	MS1 (precursor intensity)	149
		DIA (e.g., DIA, SWATH, MS^E)	AUC analysis of peptides/proteins identified by matching to spectral library. MS/MS performed on all ions within in a defined mass range.	MS1 or MS2	46
Targeted proteomics	Relative to internal standard	MRM/SRM/PRM	Ratio of AUC for native versus SIL internal standard. Can be applied to large panels or to a few select targets.	MS2	55

Definition of abbreviations: AUC = area under the curve; DIA = data-independent acquisition; iTRAQ = isobaric tags for relative and absolute quantification; LC = liquid chromatography; MRM = multiple reaction monitoring; MS = mass spectroscopy; PRM = parallel reaction monitoring; SIL = stable isotope-labeled; SILAC = stable isotope labeling of amino acids in cell culture; SRM = selected reaction monitoring; TMT = tandem mass tags.

The application of LC-MS/MS for the identification and quantitation of select peptides/proteins (targeted proteomics) is an alternative to immunoassay-based protein quantitation (10, 54). Although immunoassays may have greater sensitivity, targeted proteomic assays use internal standards to achieve high specificity, offer a high degree of multiplexing, and enable facile quantitation of post-translational modifications. These advantages were discussed in a presentation by Dr. Matt Foster (Duke University Durham, NC), which also served as an introduction to the design and application of targeted proteomic assays to airway cells and biofluids. Dr. Foster has employed targeted proteomic assays for the quantitation of cytokines and chemokines in BALF (55), and the quantitation of allelic variants and isoforms of surfactant protein A (56). Additional applications include quantitation of genetic lineages of human metapneumovirus from cell culture and nasal lavage specimens (57). Targeted proteomics is also a powerful tool for quantitation of post-translationally modified peptides beyond phosphorylation (e.g., small ubiquitin-like modifier (SUMO) modification, methylation, acetylation, ubiquitinylation, acylation, and oxidation) that can now be measured by the thousands in discovery-based proteomics studies. To this end, he presented data on a targeted proteomic assay for quantitation of newly discovered phosphorylation sites in basal cell cytokeratins (58). Targeted proteomic assay development has been a major focus of the NIH-funded Clinical Proteomic Tumor Analysis Consortium (CPTAC) (10), and this technique will likely have an important future role in clinical diagnostic and prognostic assays for lung diseases.

Summary Integration

This symposium focused on current challenges in applying emerging metabolomics and proteomics methodologies to lung disease. Unlike genetics and genomics, where advanced sequencing technology allows independent laboratories to achieve highly similar results, metabolomics and proteomic profiling remains challenging for several key reasons. First, there are the inherent challenges to metabolomics and proteomics, such as identification and quantification of both peptides and metabolites, along with "big data" analyses that aggregate samples and data across many laboratories and impair the feasibility of systems biology data integration from multiple sources. In addition, there are lung-specific issues, such as lack of uniform SOPs specifically for the lung, leading to operator and protocol variability with sample attainment. Working to advocate for an NIH/NHLBI investment in publicly available, well phenotyped biobanks that include diverse sample types (e.g., plasma, BALF, lung biopsies) would help standardize proteomic and metabolomic methods and further the field of biomarker development in lung diseases. This emerged as a major goal of this symposium.

Although there may be challenges to standardize large population metabolomics and proteomics analyses across different platforms, the potential benefit for

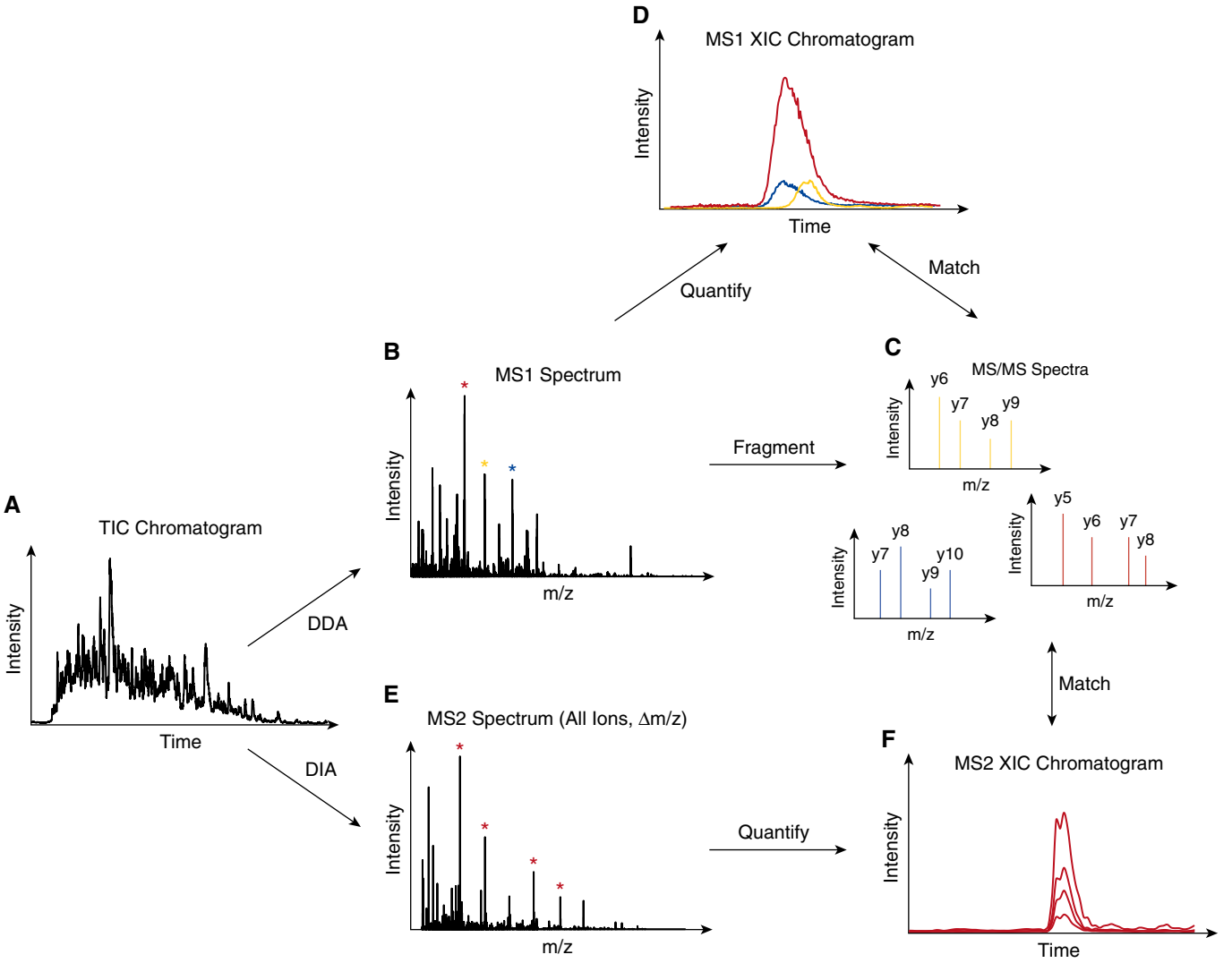

Figure 4. Summary of two common proteomic data acquisition methods. In a typical unbiased proteomic analysis, tryptic peptides are separated using liquid chromatography and introduced into the mass spectrometer using electrospray ionization. The sum intensity of detected peptides is often visualized as total ion current over time, as in *A*. (*B*) In a typical cycle of a data-dependent analysis (DDA), a "full scan" of all precursor (MS1) ions present is performed followed by (*C*) tandem mass spectrometry (MS/MS) analysis of the topN (e.g., top3; *starred peaks*) most abundant ions. The MS/MS spectra are used for database searching to identify the corresponding peptides. (*D*) Finally, identified peptides are quantified based on the area under the curve (AUC) of the MS1 intensity. (*E*) In a data-independent analysis (DIA), all ions within a selected mass range are subjected to MS/MS fragmentation. (*F*) Quantitation is performed by AUC of the fragment ions (MS2) that belong to a particular peptide. *m/z* = mass-to-charge ratio; MS1 = mass spectrometry analyzer 1; MS2 = mass spectrometry analyzer 2; TIC = total ion chromatogram; XIC = extracted ion chromatogram.

overcoming these challenges can be illustrated through more focused approaches on single platforms by individual laboratories. For example, Dr. Michael Snyder (Stanford University, Stanford, CA) highlighted his own experience using a longitudinal integrative personal omics profile (iPOP), examining genomic, transcriptomic, proteomic, metabolomic, and autoantibody profiles from a single individual over a 14-month period (59). This strategy revealed how iPOP can detect prediabetes as well as routine viral infections. This type of approach in which proteomics and metabolomics are integrated with other "omics" at the individual level is technically feasible and is an outstanding example of precision medicine; however, there remain cost and bioinformatic challenges that need to be overcome before the iPOP becomes routine in clinical practice.

In conclusion, the symposium demonstrated how proteomics and metabolomics can be used to better understand and track lung diseases, but there remain important necessary steps to bring these fields into routine clinical practice. ■

This official Workshop Report was prepared by an *ad hoc* subcommittee of the American Thoracic Society Workgroup on Metabolomics and Proteomics.

Members of the Subcommittee are as follows:
Russell P. Bowler, M.D., Ph.D. (*Co-Chair*)
Chris H. Wendt, M.D. (*Co-Chair*)
Michael B. Fessler, M.D.
Matthew W. Foster, Ph.D.
Rachel S. Kelly, M.P.H., Ph.D.
Jessica Lasky-Su, Sc.D.

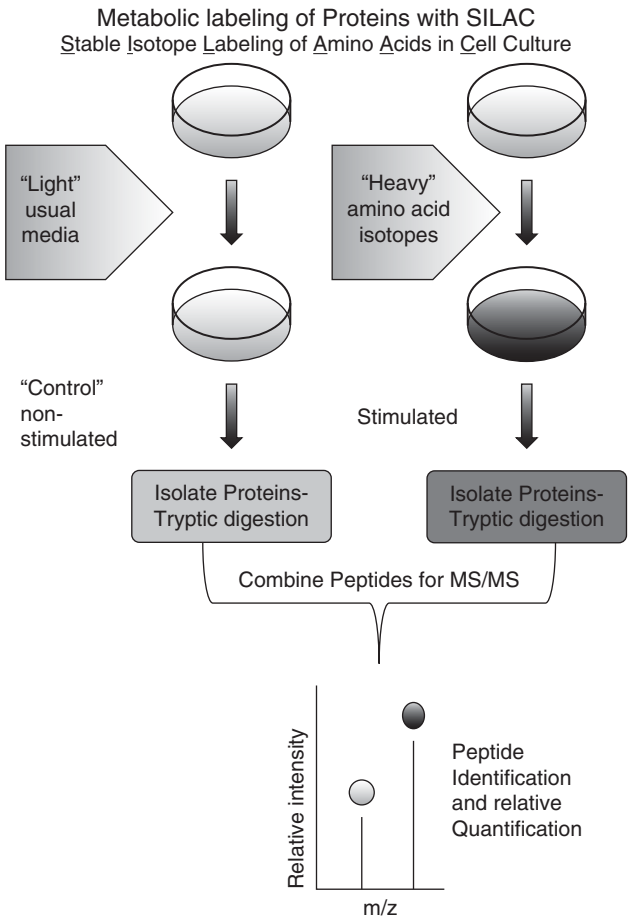

Figure 5. Stable isotope labeling with amino acids in cell culture (SILAC) technology uses cells grown in isotopically labeled amino acids to synthesize the "heavy" forms of proteins that can be mixed with their "light" counterparts before trypsinization and/or peptide/protein fractionation and identification by tandem mass spectroscopy. *m/z* = mass-to-charge ratio; MS/MS = tandem mass spectrometry.

Angela J. Rogers, M.D.
Kathleen A. Stringer, Pharm.D.
Brent W. Winston, M.D.

Author disclosures: R.P.B. received research support and served on an advisory committee for GlaxoSmithKline. J.L.-S. served as a consultant and received personal fees from Metabolon, Inc. M.B.F., M.W.F., R.S.K., A.J.R., K.A.S., C.H.W., B.W.W. reported no relationships with relevant commercial interests.

Members of the Workshop Organizing Committee are as follows:
Russell P. Bowler, M.D., Ph.D. (*Co-Chair*)
Chris H. Wendt, M.D. (*Co-Chair*)
Evgeny Berdyshev, Ph.D.
Maneesh Bhargava, M.D., Ph.D.
Michael B. Fessler, M.D.
Matthew W. Foster, Ph.D.
Dean Jones, Ph.D.
Mehmet Kesimer, Ph.D.
Dinesh Pillai, M.D.
Angela J. Rogers, M.D.
Mary Rose, Ph.D.
Kathleen A. Stringer, Pharm.D.
Brent W. Winston, M.D.

A list of participants, poster award winners, speaker summary statements, and poster abstracts is included in the online supplement.

Acknowledgment: The authors gratefully acknowledge support from the following organizations and individuals who made the workshop possible: the American Thoracic Society (ATS), the National Institutes of Health (R13HL127973), National Jewish Health, and Colorado Biological Mass Spectrometry Society, including unrestricted funding from: Agilent Technologies, Inc., Parion Sciences, Inc., Avanti Polar Lipids, Inc., Bruker Co., Chenomx Inc., SomaLogic Inc., and Waters Co. Special thanks to Aubrey Shoe, National Jewish Health, for administrative assistance, and ATS administrative staff Breana Portelli and Kimberly Lawrence.

References

1 Contrepois K, Liang L, Snyder M. Can metabolic profiles be used as a phenotypic readout of the genome to enhance precision medicine? *Clin Chem* 2016;62:676–678.
2 Auffray C, Adcock IM, Chung KF, Djukanovic R, Pison C, Sterk PJ. An integrative systems biology approach to understanding pulmonary diseases. *Chest* 2010;137:1410–1416.
3 Kelly RS, Dahlin A, McGeachie MJ, Qiu W, Sordillo J, Wan ES, Wu AC, Lasky-Su J. Asthma metabolomics and the potential for integrative omics in research and the clinic. *Chest* 2017;151:262–277.
4 Li X, Snyder M. Metabolites as global regulators: a new view of protein regulation: systematic investigation of metabolite-protein interactions may help bridge the gap between genome-wide association studies and small molecule screening studies. *BioEssays* 2011;33:485–489.

5 Towle HC. Metabolic regulation of gene transcription in mammals. *J Biol Chem* 1995;270:23235–23238.
6 Stringer KA, McKay RT, Karnovsky A, Quémerais B, Lacy P. Metabolomics and its application to acute lung diseases. *Front Immunol* 2016;7:44.
7 Rennard SI, Basset G, Lecossier D, O'Donnell KM, Pinkston P, Martin PG, Crystal RG. Estimation of volume of epithelial lining fluid recovered by lavage using urea as marker of dilution. *J Appl Physiol* 1985;1986: 532–538.
8 van der Vliet A, O'Neill CA, Cross CE, Koostra JM, Volz WG, Halliwell B, Louie S. Determination of low-molecular-mass antioxidant concentrations in human respiratory tract lining fluids. *Am J Physiol* 1999;276:L289–L296.
9 Bowler RP, Duda B, Chan ED, Enghild JJ, Ware LB, Matthay MA, Duncan MW. Proteomic analysis of pulmonary edema fluid and plasma in patients with acute lung injury. *Am J Physiol Lung Cell Mol Physiol* 2004;286:L1095–L1104.
10 Carr SA, Abbatiello SE, Ackermann BL, Borchers C, Domon B, Deutsch EW, Grant RP, Hoofnagle AN, Hüttenhain R, Koomen JM, *et al*. Targeted peptide measurements in biology and medicine: best practices for mass spectrometry-based assay development using a fit-for-purpose approach. *Mol Cell Proteomics* 2014;13: 907–917.
11 Grant RP, Hoofnagle AN. From lost in translation to paradise found: enabling protein biomarker method transfer by mass spectrometry. *Clin Chem* 2014;60:941–944.
12 Wehrens R, Hageman JA, van Eeuwijk F, Kooke R, Flood PJ, Wijnker E, Keurentjes JJ, Lommen A, van Eekelen HD, Hall RD, *et al*. Improved batch correction in untargeted MS-based metabolomics. *Metabolomics* 2016;12:88.
13 Gagnon-Bartsch JA, Speed TP. Using control genes to correct for unwanted variation in microarray data. *Biostatistics* 2012;13: 539–552.
14 Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 2007; 8:118–127.
15 Worley B, Powers R. Multivariate analysis in metabolomics. *Curr Metabolomics* 2013;1:92–107.
16 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995;57:289–300.
17 Johnson CH, Ivanisevic J, Benton HP, Siuzdak G. Bioinformatics: the next frontier of metabolomics. *Anal Chem* 2015;87: 147–156.
18 Rogers AJ, Matthay MA. Applying metabolomics to uncover novel biology in ARDS. *Am J Physiol Lung Cell Mol Physiol* 2014;306: L957–L961.
19 Thille AW, Esteban A, Fernández-Segoviano P, Rodriguez JM, Aramburu JA, Peñuelas O, Cortés-Puch I, Cardinal-Fernández P, Lorente JA, Frutos-Vivar F. Comparison of the Berlin definition for acute respiratory distress syndrome with autopsy. *Am J Respir Crit Care Med* 2013;187:761–767.
20 Comhair SA, McDunn J, Bennett C, Fettig J, Erzurum SC, Kalhan SC. Metabolomic endotype of asthma. *J Immunol* 2015;195:643–650.
21 Adamko DJ, Nair P, Mayers I, Tsuyuki RT, Regush S, Rowe BH. Metabolomic profiling of asthma and chronic obstructive pulmonary disease: a pilot study differentiating diseases. *J Allergy Clin Immunol* 2015;136:571–580.e573.
22 Trygg J, Holmes E, Lundstedt T. Chemometrics in metabonomics. *J Proteome Res* 2007;6:469–479.
23 Uppal K, Walker DI, Liu K, Li S, Go YM, Jones DP. Computational metabolomics: a framework for the million metabolome. *Chem Res Toxicol* 2016;29:1956–1975.
24 Cavalcante RG, Patil S, Weymouth TE, Bendinskas KG, Karnovsky A, Sartor MA. ConceptMetab: exploring relationships among metabolite sets to identify links among biomedical concepts. *Bioinformatics* 2016;32:1536–1543.
25 Karnovsky A, Weymouth T, Hull T, Tarcea VG, Scardoni G, Laudanna C, Sartor MA, Stringer KA, Jagadish HV, Burant C, *et al*. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* 2012;28: 373–380.
26 Xia J, Wishart DS. MSEA: a web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res* 2010;38:W71–W77.
27 Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, VanderGheynst J, Fiehn O, Arita M. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nat Methods* 2015;12:523–526.
28 Kind T, Liu KH, Lee DY, DeFelice B, Meissen JK, Fiehn O. LipidBlast in silico tandem mass spectrometry database for lipid identification. *Nat Methods* 2013;10:755–758.
29 Caprioli RM. Imaging mass spectrometry: molecular microscopy for the new age of biology and medicine. *Proteomics* 2016;16: 1607–1612.
30 Van de Plas R, Yang J, Spraggins J, Caprioli RM. Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping. *Nat Methods* 2015;12:366–372.
31 Snowden S, Dahlén SE, Wheelock CE. Application of metabolomics approaches to the study of respiratory diseases. *Bioanalysis* 2012; 4:2265–2290.
32 Banoei MM, Donnelly SJ, Mickiewicz B, Weljie A, Vogel HJ, Winston BW. Metabolomics in critical care medicine: a new approach to biomarker discovery. *Clin Invest Med* 2014;37: E363–E376.
33 Fang ZZ, Gonzalez FJ. LC-MS–based metabolomics: an update. *Arch Toxicol* 2014;88:1491–1502.
34 Weckwerth W, Fiehn O. Can we discover novel pathways using metabolomic analysis? *Curr Opin Biotechnol* 2002;13:156–160.
35 Bantscheff M, Lemeer S, Savitski MM, Kuster B. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem* 2012;404:939–965.
36 Bhargava M, Becker TL, Viken KJ, Jagtap PD, Dey S, Steinbach MS, Wu B, Kumar V, Bitterman PB, Ingbar DH, *et al*. Proteomic profiles in acute respiratory distress syndrome differentiates survivors from non-survivors. *PLoS One* 2014;9:e109713.
37 Nicholas BL, Skipp P, Barton S, Singh D, Bagmane D, Mould R, Angco G, Ward J, Guha-Niyogi B, Wilson S, *et al*. Identification of lipocalin and apolipoprotein A1 as biomarkers of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2010; 181:1049–1060.
38 Tu C, Mammen MJ, Li J, Shen X, Jiang X, Hu Q, Wang J, Sethi S, Qu J. Large-scale, ion-current–based proteomics investigation of bronchoalveolar lavage fluid in chronic obstructive pulmonary disease patients. *J Proteome Res* 2014;13:627–639.
39 Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR III. Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* 2013;113: 2343–2394.
40 Choi H, Fermin D, Nesvizhskii AI. Significance analysis of spectral count data in label-free shotgun proteomics. *Mol Cell Proteomics* 2008;7:2373–2385.
41 Choi H, Kim S, Fermin D, Tsou CC, Nesvizhskii AI. QPROT: statistical method for testing differential expression using protein-level intensity data in label-free quantitative proteomics. *J Proteomics* 2015;129:121–126.
42 Tsou CC, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras AC, Nesvizhskii AI. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* 2015;12:258–264. [257 p. following 264.]
43 Tsou CC, Tsai CF, Teo GC, Chen YJ, Nesvizhskii AI. Untargeted, spectral library-free analysis of data-independent acquisition proteomics data generated using Orbitrap mass spectrometers. *Proteomics* 2016;16:2257–2271.
44 Jaitly N, Monroe ME, Petyuk VA, Clauss TR, Adkins JN, Smith RD. Robust algorithm for alignment of liquid chromatography–mass spectrometry analyses in an accurate mass and time tag data analysis pipeline. *Anal Chem* 2006;78:7397–7409.
45 Gillet LC, Navarro P, Tate S, Rost H, Selevsek N, Reiter L, Bonner R, Aebersold R. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* 2012;11:O111.016717.

46 Ortea I, Rodríguez-Ariza A, Chicano-Gálvez E, Arenas Vacas MS, Jurado Gámez B. Discovery of potential protein biomarkers of lung adenocarcinoma in bronchoalveolar lavage fluid by SWATH MS data-independent acquisition and targeted data extraction. *J Proteomics* 2016;138:106–114.

47 Peters-Hall JR, Brown KJ, Pillai DK, Tomney A, Garvin LM, Wu X, Rose MC. Quantitative proteomics reveals an altered cystic fibrosis *in vitro* bronchial epithelial secretome. *Am J Respir Cell Mol Biol* 2015;53:22–32.

48 Pillai DK, Sankoorikal BJ, Johnson E, Seneviratne AN, Zurko J, Brown KJ, Hathout Y, Rose MC. Directional secretomes reflect polarity-specific functions in an *in vitro* model of human bronchial epithelium. *Am J Respir Cell Mol Biol* 2014;50:292–300.

49 Val S, Burgett K, Brown KJ, Preciado D. SuperSILAC quantitative proteome profiling of murine middle ear epithelial cell remodeling with NTHi. *PLoS One* 2016;11:e0148612.

50 Chowdhury SM, Zhu X, Aloor JJ, Azzam KM, Gabor KA, Ge W, Addo KA, Tomer KB, Parks JS, Fessler MB. Proteomic analysis of ABCA1-null macrophages reveals a role for stomatin-like protein-2 in raft composition and Toll-like receptor signaling. *Mol Cell Proteomics* 2015;14:1859–1870.

51 Fessler MB, Malcolm KC, Duncan MW, Worthen GS. A genomic and proteomic analysis of activation of the human neutrophil by lipopolysaccharide and its mediation by p38 mitogen-activated protein kinase. *J Biol Chem* 2002;277:31291–31302.

52 Dhungana S, Merrick BA, Tomer KB, Fessler MB. Quantitative proteomics analysis of macrophage rafts reveals compartmentalized activation of the proteasome and of proteasome-mediated ERK activation in response to lipopolysaccharide. *Mol Cell Proteomics* 2009;8:201–213.

53 Merrick BA, Dhungana S, Williams JG, Aloor JJ, Peddada S, Tomer KB, Fessler MB. Proteomic profiling of S-acylated macrophage proteins identifies a role for palmitoylation in mitochondrial targeting of phospholipid scramblase 3. *Mol Cell Proteomics* 2011;10: M110.006007.

54 Picotti P, Bodenmiller B, Aebersold R. Proteomics meets the scientific method. *Nat Methods* 2013;10:24–27.

55 Foster MW, Morrison LD, Todd JL, Snyder LD, Thompson JW, Soderblom EJ, Plonk K, Weinhold KJ, Townsend R, Minnich A, *et al*. Quantitative proteomics of bronchoalveolar lavage fluid in idiopathic pulmonary fibrosis. *J Proteome Res* 2015;14:1238–1249.

56 Foster MW, Thompson JW, Ledford JG, Dubois LG, Hollingsworth JW, Francisco D, Tanyaratsrisakul S, Voelker DR, Kraft M, Moseley MA, *et al*. Identification and quantitation of coding variants and isoforms of pulmonary surfactant protein A. *J Proteome Res* 2014; 13:3722–3732.

57 Foster MW, Gerhardt G, Robitaille L, Plante PL, Boivin G, Corbeil J, Moseley MA. Targeted proteomics of human metapneumovirus in clinical samples and viral cultures. *Anal Chem* 2015;87: 10247–10254.

58 Foster MW, Gwinn WM, Kelly FL, Brass DM, Valente AM, Moseley MA, Thompson JW, Morgan DL, Palmer SM. Proteomic analysis of primary human airway epithelial cells exposed to the respiratory toxicant diacetyl. *J Proteome Res* 2017;16:538–549.

59 Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, Miriami E, Karczewski KJ, Hariharan M, Dewey FE, *et al*. Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell* 2012;148:1293–1307.

60 Hayes SA, Haefliger S, Harris B, Pavlakis N, Clarke SJ, Molloy MP, Howell VM. Exhaled breath condensate for lung cancer protein analysis: a review of methods and biomarkers. *J Breath Res* 2016; 10:034001.

61 Horváth I, Hunt J, Barnes PJ, Alving K, Antczak A, Baraldi E, Becher G, van Beurden WJ, Corradi M, Dekhuijzen R, *et al*.; ATS/ERS Task Force on Exhaled Breath Condensate. Exhaled breath condensate: methodological recommendations and unresolved questions. *Eur Respir J* 2005;26:523–548.

62 Stringer KA, Younger JG, McHugh C, Yeomans L, Finkel MA, Puskarich MA, Jones AE, Trexel J, Karnovsky A. Whole blood reveals more metabolic detail of the human metabolome than serum as measured by 1H-NMR spectroscopy: implications for sepsis metabolomics. *Shock* 2015;44:200–208.

63 Emwas AH, Luchinat C, Turano P, Tenori L, Roy R, Salek RM, Ryan D, Merzaban JS, Kaddurah-Daouk R, Zeri AC, *et al*. Standardizing the experimental conditions for using urine in NMR-based metabolomic studies with a particular focus on diagnostic studies: a review. *Metabolomics* 2015;11:872–894.

64 Schubert JK, Müller WP, Benzing A, Geiger K. Application of a new method for analysis of exhaled gas in critically ill patients. *Intensive Care Med* 1998;24:415–421.

65 Stringer KA, Serkova NJ, Karnovsky A, Guire K, Paine R III, Standiford TJ. Metabolic consequences of sepsis-induced acute lung injury revealed by plasma ^{1}H-nuclear magnetic resonance quantitative metabolomics and computational analysis. *Am J Physiol Lung Cell Mol Physiol* 2011;300:L4–L11.

66 Evans CR, Karnovsky A, Kovach MA, Standiford TJ, Burant CF, Stringer KA. Untargeted LC-MS metabolomics of bronchoalveolar lavage fluid differentiates acute respiratory distress syndrome from health. *J Proteome Res* 2014;13:640–649.

67 Bos LD, Weda H, Wang Y, Knobel HH, Nijsen TM, Vink TJ, Zwinderman AH, Sterk PJ, Schultz MJ. Exhaled breath metabolomics as a noninvasive diagnostic tool for acute respiratory distress syndrome. *Eur Respir J* 2014;44:188–197.

68 Rogers AJ, Contrepois K, Wu M, Zheng M, Peltz G, Ware LB, Matthay MA. Profiling of ARDS pulmonary edema fluid identifies a metabolically distinct subset. *Am J Physiol Lung Cell Mol Physiol* 2017;312:L703–L709.

69 McClay JL, Adkins DE, Isern NG, O'Connell TM, Wooten JB, Zedler BK, Dasika MS, Webb BT, Webb-Robertson BJ, Pounds JG, *et al*. (1)H nuclear magnetic resonance metabolomics analysis identifies novel urinary biomarkers for lung function. *J Proteome Res* 2010;9: 3083–3090.

70 Ubhi BK, Cheng KK, Dong J, Janowitz T, Jodrell D, Tal-Singer R, MacNee W, Lomas DA, Riley JH, Griffin JL, *et al*. Targeted metabolomics identifies perturbations in amino acid metabolism that sub-classify patients with COPD. *Mol Biosyst* 2012;8: 3125–3133.

71 Ubhi BK, Riley JH, Shaw PA, Lomas DA, Tal-Singer R, MacNee W, Griffin JL, Connor SC. Metabolic profiling detects biomarkers of protein degradation in COPD patients. *Eur Respir J* 2012;40:345–355.

72 Wang L, Tang Y, Liu S, Mao S, Ling Y, Liu D, He X, Wang X. Metabonomic profiling of serum and urine by (1)H NMR-based spectroscopy discriminates patients with chronic obstructive pulmonary disease and healthy individuals. *PLoS One* 2013;8:e65675.

73 de Laurentiis G, Paris D, Melck D, Montuschi P, Maniscalco M, Bianco A, Sofia M, Motta A. Separating smoking-related diseases using NMR-based metabolomics of exhaled breath condensate. *J Proteome Res* 2013;12:1502–1511.

74 Bahr TM, Hughes GJ, Armstrong M, Reisdorph R, Coldren CD, Edwards MG, Schnell C, Kedl R, LaFlamme DJ, Reisdorph N, *et al*. Peripheral blood mononuclear cell gene expression in chronic obstructive pulmonary disease. *Am J Respir Cell Mol Biol* 2013;49: 316–323.

75 Bowler RP, Jacobson S, Cruickshank C, Hughes GJ, Siska C, Ory DS, Petrache I, Schaffer JE, Reisdorph N, Kechris K. Plasma sphingolipids associated with chronic obstructive pulmonary disease phenotypes. *Am J Respir Crit Care Med* 2015;191: 275–284.

76 Chen Q, Deeb RS, Ma Y, Staudt MR, Crystal RG, Gross SS. Serum metabolite biomarkers discriminate healthy smokers from COPD smokers. *PLoS One* 2015;10:e0143937.

77 Esther CR Jr, Boysen G, Olsen BM, Collins LB, Ghio AJ, Swenberg JW, Boucher RC. Mass spectrometric analysis of biomarkers and dilution markers in exhaled breath condensate reveals elevated purines in asthma and cystic fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2009; 296:L987–L993.

78 Saude EJ, Skappak CD, Regush S, Cook K, Ben-Zvi A, Becker A, Moqbel R, Sykes BD, Rowe BH, Adamko DJ. Metabolomic profiling of asthma: diagnostic utility of urine nuclear magnetic resonance spectroscopy. *J Allergy Clin Immunol* 2011;127:757–764.e751–756.

79 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* 2012;26:89–94.
80 Carraro S, Giordano G, Reniero F, Carpi D, Stocchero M, Sterk PJ, Baraldi E. Asthma severity in childhood and metabolomic profiling of breath condensate. *Allergy* 2013;68:110–117.
81 Jung J, Kim SH, Lee HS, Choi GS, Jung YS, Ryu DH, Park HS, Hwang GS. Serum metabolomics reveals pathways and biomarkers associated with asthma pathogenesis. *Clin Exp Allergy* 2013;43: 425–433.
82 van de Kant KDG, van Berkel JJBN, Jöbsis Q, Lima Passos V, Klaassen EMM, van der Sande L, van Schayck OCP, de Jongste JC, van Schooten FJ, Derks E, *et al*. Exhaled breath profiling in diagnosing wheezy preschool children. *Eur Respir J* 2013;41:183–188.
83 Gahleitner F, Guallar-Hoyas C, Beardsmore CS, Pandya HC, Thomas CP. Metabolomics pilot study to identify volatile organic compound markers of childhood asthma in exhaled breath. *Bioanalysis* 2013;5: 2239–2247.
84 Motta A, Paris D, D'Amato M, Melck D, Calabrese C, Vitale C, Stanziola AA, Corso G, Sofia M, Maniscalco M. NMR metabolomic analysis of exhaled breath condensate of asthmatic patients at two different temperatures. *J Proteome Res* 2014;13:6107–6120.
85 Loureiro CC, Duarte IF, Gomes J, Carrola J, Barros AS, Gil AM, Bousquet J, Bom AT, Rocha SM. Urinary metabolomic changes as a predictive biomarker of asthma exacerbation. *J Allergy Clin Immunol* 2014;133:261–263.e1–5.
86 Smolinska A, Klaassen EMM, Dallinga JW, van de Kant KDG, Jobsis Q, Moonen EJC, van Schayck OCP, Dompeling E, van Schooten FJ. Profiling of volatile organic compounds in exhaled breath as a strategy to find early predictive signatures of asthma in children. *PLoS One* 2014;9:e95668.
87 Chang C, Guo ZG, He B, Yao WZ. Metabolic alterations in the sera of Chinese patients with mild persistent asthma: a GC-MS–based metabolomics analysis. *Acta Pharmacol Sin* 2015;36:1356–1366.
88 Comhair SAA, McDunn J, Bennett C, Fettig J, Erzurum SC, Kalhan SC. Metabolomic endotype of asthma. *J Immunol* 2015;195: 643–650.
89 Fitzpatrick AM, Park Y, Brown LA, Jones DP. Children with severe asthma have unique oxidative stress-associated metabolomic profiles. *J Allergy Clin Immunol* 2014;133:258–261.e1–8.
90 Tautenhahn R, Cho K, Uritboonthai W, Zhu Z, Patti GJ, Siuzdak G. An accelerated workflow for untargeted metabolomics using the METLIN database. *Nat Biotechnol* 2012;30:826–828.
91 Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, Liu Y, Djoumbou Y, Mandal R, Aziat F, Dong E, *et al*. HMDB 3.0—the Human Metabolome Database in 2013. *Nucleic Acids Res* 2013;41: D801–D807.
92 Uppal K, Soltow QA, Strobel FH, Pittard WS, Gernert KM, Yu T, Jones DP. xMSanalyzer: automated pipeline for improved feature detection and downstream analysis of large-scale, non-targeted metabolomics data. *BMC Bioinformatics* 2013;14:15.
93 Yu T, Park Y, Johnson JM, Jones DP. apLCMS—adaptive processing of high-resolution LC/MS data. *Bioinformatics* 2009; 25:1930–1936.
94 Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal Chem* 2012;84:283–289.
95 Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 2006;78:779–787.
96 Wei X, Sun W, Shi X, Koo I, Wang B, Zhang J, Yin X, Tang Y, Bogdanov B, Kim S, *et al*. MetSign: a computational platform for high-resolution mass spectrometry-based metabolomics. *Anal Chem* 2011;83:7668–7675.
97 Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* 2010;11:395.
98 Fernández-Albert F, Llorach R, Andrés-Lacueva C, Perera A. An R package to analyse LC/MS metabolomic data: MAIT (Metabolite Automatic Identification Toolkit). *Bioinformatics* 2014;30: 1937–1939.
99 Hiller K, Hangebrauk J, Jäger C, Spura J, Schreiber K, Schomburg D. MetaboliteDetector: comprehensive analysis tool for targeted and nontargeted GC/MS based metabolome analysis. *Anal Chem* 2009; 81:3429–3439.
100 Wang T, Shao K, Chu Q, Ren Y, Mu Y, Qu L, He J, Jin C, Xia B. Automics: an integrated platform for NMR-based metabonomics spectral processing and data analysis. *BMC Bioinformatics* 2009;10:83.
101 Ravanbakhsh S, Liu P, Bjorndahl TC, Mandal R, Grant JR, Wilson M, Eisner R, Sinelnikov I, Hu X, Luchinat C, *et al*. Accurate, fully-automated NMR spectral profiling for metabolomics. *PLoS One* 2015;10:e0124219. [Published erratum appears in *PLoS One* 10: e0132873.]
102 Lacy P, McKay RT, Finkel M, Karnovsky A, Woehler S, Lewis MJ, Chang D, Stringer KA. Signal intensities derived from different NMR probes and parameters contribute to variations in quantification of metabolites. *PLoS One* 2014;9:e85732.
103 Alonso A, Rodríguez MA, Vinaixa M, Tortosa R, Correig X, Julià A, Marsal S. Focus: a robust workflow for one-dimensional NMR spectral analysis. *Anal Chem* 2014;86:1160–1169.
104 Vu TN, Valkenborg D, Smets K, Verwaest KA, Dommisse R, Lemière F, Verschoren A, Goethals B, Laukens K. An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. *BMC Bioinformatics* 2011;12:405.
105 Ludwig C, Günther UL. MetaboLab—advanced NMR data processing and analysis for metabolomics. *BMC Bioinformatics* 2011;12:366.
106 Lewis IA, Schommer SC, Markley JL. rNMR: open source software for identifying and quantifying metabolites in NMR spectra. *Magn Reson Chem* 2009;47:S123–S126.
107 Xia J, Bjorndahl TC, Tang P, Wishart DS. MetaboMiner—semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. *BMC Bioinformatics* 2008;9:507.
108 Hao J, Liebeke M, Astle W, De Iorio M, Bundy JG, Ebbels TM. Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nat Protoc* 2014;9: 1416–1427.
109 Tulpan D, Léger S, Belliveau L, Culf A, Cuperlović-Culf M. MetaboHunter: an automatic approach for identification of metabolites from 1H-NMR spectra of complex mixtures. *BMC Bioinformatics* 2011;12:400.
110 Xia J, Wishart DS. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr Protoc Bioinformatics* 2016;55: 14.10.1–14.10.91.
111 Carroll AJ, Badger MR, Harvey Millar A. The MetabolomeExpress Project: enabling web-based processing, analysis and transparent dissemination of GC/MS metabolomics datasets. *BMC Bioinformatics* 2010;11:376.
112 Alonso A, Julià A, Beltran A, Vinaixa M, Díaz M, Ibañez L, Correig X, Marsal S. AStream: an R package for annotating LC/MS metabolomic data. *Bioinformatics* 2011;27:1339–1340.
113 Cottret L, Wildridge D, Vinson F, Barrett MP, Charles H, Sagot MF, Jourdan F. MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res* 2010;38:W132–W137.
114 Cavill R, Kamburov A, Ellis JK, Athersuch TJ, Blagrove MS, Herwig R, Ebbels TM, Keun HC. Consensus-phenotype integration of transcriptomic and metabolomic data implies a role for metabolism in the chemosensitivity of tumour cells. *PLoS Comput Biol* 2011;7: e1001113.
115 García-Alcalde F, García-López F, Dopazo J, Conesa A. Paintomics: a web based tool for the joint visualization of transcriptomics and metabolomics data. *Bioinformatics* 2011;27:137–139.
116 Melamud E, Vastag L, Rabinowitz JD. Metabolomic analysis and visualization engine for LC-MS data. *Anal Chem* 2010;82: 9818–9826.
117 Rohn H, Junker A, Hartmann A, Grafahrend-Belau E, Treutler H, Klapperstück M, Czauderna T, Klukas C, Schreiber F. VANTED v2: a framework for systems biology applications. *BMC Syst Biol* 2012;6: 139.

118 Wanichthanarak K, Fan S, Grapov D, Barupal DK, Fiehn O. Metabox: a toolbox for metabolomic data analysis, interpretation and integrative exploration. *PLoS One* 2017;12:e0171046.

119 Li S, Park Y, Duraisingham S, Strobel FH, Khan N, Soltow QA, Jones DP, Pulendran B. Predicting network activity from high throughput metabolomics. *PLoS Comput Biol* 2013;9: e1003123.

120 de Torre C, Ying SX, Munson PJ, Meduri GU, Suffredini AF. Proteomic analysis of inflammatory biomarkers in bronchoalveolar lavage. *Proteomics* 2006;6:3949–3957.

121 Schnapp LM, Donohoe S, Chen J, Sunde DA, Kelly PM, Ruzinski J, Martin T, Goodlett DR. Mining the acute respiratory distress syndrome proteome: identification of the insulin-like growth factor (IGF)/IGF-binding protein-3 pathway in acute lung injury. *Am J Pathol* 2006;169:86–95.

122 Chang DW, Hayashi S, Gharib SA, Vaisar T, King ST, Tsuchiya M, Ruzinski JT, Park DR, Matute-Bello G, Wurfel MM, *et al*. Proteomic and computational analysis of bronchoalveolar proteins during the course of the acute respiratory distress syndrome. *Am J Respir Crit Care Med* 2008;178:701–709.

123 Chen X, Shan Q, Jiang L, Zhu B, Xi X. Quantitative proteomic analysis by iTRAQ for identification of candidate biomarkers in plasma from acute respiratory distress syndrome patients. *Biochem Biophys Res Commun* 2013;441:1–6.

124 Dong H, Li J, Lv Y, Zhou Y, Wang G, Hu S, He X, Yang P, Zhou Z, Xiang X, *et al*. Comparative analysis of the alveolar macrophage proteome in ALI/ARDS patients between the exudative phase and recovery phase. *BMC Immunol* 2013;14:25.

125 Ren S, Chen X, Jiang L, Zhu B, Jiang Q, Xi X. Deleted in malignant brain tumors 1 protein is a potential biomarker of acute respiratory distress syndrome induced by pneumonia. *Biochem Biophys Res Commun* 2016;478:1344–1349.

126 Merkel D, Rist W, Seither P, Weith A, Lenter MC. Proteomic study of human bronchoalveolar lavage fluids from smokers with chronic obstructive pulmonary disease by combining surface-enhanced laser desorption/ionization–mass spectrometry profiling with mass spectrometric protein identification. *Proteomics* 2005;5:2972–2980.

127 Ohlmeier S, Mazur W, Salmenkivi K, Myllärniemi M, Bergmann U, Kinnula VL. Proteomic studies on receptor for advanced glycation end product variants in idiopathic pulmonary fibrosis and chronic obstructive pulmonary disease. *Proteomics Clin Appl* 2010;4: 97–105.

128 Bozinovski S, Hutchinson A, Thompson M, Macgregor L, Black J, Giannakis E, Karlsson AS, Silvestrini R, Smallwood D, Vlahos R, *et al*. Serum amyloid a is a biomarker of acute exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2008;177:269–278.

129 Lee EJ, In KH, Kim JH, Lee SY, Shin C, Shim JJ, Kang KH, Yoo SH, Kim CH, Kim HK, *et al*. Proteomic analysis in lung tissue of smokers and COPD patients. *Chest* 2009;135:344–352.

130 Hu R, Ouyang Q, Dai A, Tan S, Xiao Z, Tang C. Heat shock protein 27 and cyclophilin A associate with the pathogenesis of COPD. *Respirology* 2011;16:983–993.

131 Verrills NM, Irwin JA, He XY, Wood LG, Powell H, Simpson JL, McDonald VM, Sim A, Gibson PG. Identification of novel diagnostic biomarkers for asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2011;183:1633–1643.

132 Franciosi L, Govorukhina N, Fusetti F, Poolman B, Lodewijk ME, Timens W, Postma D, ten Hacken N, Bischoff R. Proteomic analysis of human epithelial lining fluid by microfluidics-based nanoLC-MS/MS: a feasibility study. *Electrophoresis* 2013;34:2683–2694.

133 Merali S, Barrero CA, Bowler RP, Chen DE, Criner G, Braverman A, Litwin S, Yeung A, Kelsen SG. Analysis of the plasma proteome in COPD: novel low abundance proteins reflect the severity of lung remodeling. *COPD* 2014;11:177–189.

134 Baraniuk JN, Casado B, Pannell LK, McGarvey PB, Boschetto P, Luisetti M, Iadarola P. Protein networks in induced sputum from smokers and COPD patients. *Int J Chron Obstruct Pulmon Dis* 2015;10:1957–1975.

135 Titz B, Sewer A, Schneider T, Elamin A, Martin F, Dijon S, Luettich K, Guedj E, Vuillaume G, Ivanov NV, *et al*. Alterations in the sputum proteome and transcriptome in smokers and early-stage COPD subjects. *J Proteomics* 2015;128:306–320.

136 Bargagli E, Olivieri C, Prasse A, Bianchi N, Magi B, Cianti R, Bini L, Rottoli P. Calgranulin B (S100A9) levels in bronchoalveolar lavage fluid of patients with interstitial lung diseases. *Inflammation* 2008; 31:351–354.

137 Korfei M, von der Beck D, Henneke I, Markart P, Ruppert C, Mahavadi P, Ghanim B, Klepetko W, Fink L, Meiners S, *et al*. Comparative proteome analysis of lung tissue from patients with idiopathic pulmonary fibrosis (IPF), non-specific interstitial pneumonia (NSIP) and organ donors. *J Proteomics* 2013;85:109–128.

138 Landi C, Bargagli E, Bianchi L, Gagliardi A, Carleo A, Bennett D, Perari MG, Armini A, Prasse A, Rottoli P, *et al*. Towards a functional proteomics approach to the comprehension of idiopathic pulmonary fibrosis, sarcoidosis, systemic sclerosis and pulmonary Langerhans cell histiocytosis. *J Proteomics* 2013;83:60–75.

139 Carleo A, Bargagli E, Landi C, Bennett D, Bianchi L, Gagliardi A, Carnemolla C, Perari MG, Cillis G, Armini A, *et al*. Comparative proteomic analysis of bronchoalveolar lavage of familial and sporadic cases of idiopathic pulmonary fibrosis. *J Breath Res* 2016; 10:026007.

140 Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, *et al*. Proteomics: tissue-based map of the human proteome. *Science* 2015;347:1260419.

141 Lindskog C, Fagerberg L, Hallström B, Edlund K, Hellwig B, Rahnenführer J, Kampf C, Uhlén M, Pontén F, Micke P. The lung-specific proteome defined by integration of transcriptomics and antibody-based profiling. *FASEB J* 2014;28:5184–5196.

142 Lam MP, Venkatraman V, Xing Y, Lau E, Cao Q, Ng DC, Su AI, Ge J, Van Eyk JE, Ping P. Data-driven approach to determine popular proteins for targeted proteomics translation of six organ systems. *J Proteome Res* 2016;15:4126–4134.

143 Clair G, Piehowski PD, Nicola T, Kitzmiller JA, Huang EL, Zink EM, Sontag RL, Orton DJ, Moore RJ, Carson JP, *et al*. Spatially-resolved proteomics: rapid quantitative analysis of laser capture microdissected alveolar tissue samples. *Sci Rep* 2016;6:39223.

144 Dautel SE, Kyle JE, Clair G, Sontag RL, Weitz KK, Shukla AK, Nguyen SN, Kim YM, Zink EM, Luders T, *et al*. Lipidomics reveals dramatic lipid compositional changes in the maturing postnatal lung. *Sci Rep* 2017;7: 40555.

145 Schiller HB, Fernandez IE, Burgstaller G, Schaab C, Scheltema RA, Schwarzmayr T, Strom TM, Eickelberg O, Mann M. Time- and compartment-resolved proteome profiling of the extracellular niche in lung injury and repair. *Mol Syst Biol* 2015;11:819.

146 Schamberger AC, Schiller HB, Fernandez IE, Sterclova M, Heinzelmann K, Hennen E, Hatz R, Behr J, Vašáková M, Mann M, *et al*. Glutathione peroxidase 3 localizes to the epithelial lining fluid and the extracellular matrix in interstitial lung disease. *Sci Rep* 2016;6:29952.

147 Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, Villén J, Haas W, Sowa ME, Gygi SP. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* 2010;143:1174–1189.

148 Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, *et al*. Mass-spectrometry–based draft of the human proteome. *Nature* 2014; 509:582–587.

149 Liu H, Sadygov RG, Yates JR III. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 2004;76:4193–4201.