Published in final edited form as:

Curr Opin Biotechnol. 2016 June ; 39: 192–197. doi:10.1016/j.copbio.2016.04.013.

Targeting the untargeted in molecular phenomics with structurally-selective ion mobility-mass spectrometry

Jody C May, Randi L Gant-Branum, and John A McLean

Department of Chemistry, Center for Innovative Technology, Vanderbilt Institute of Chemical Biology, Vanderbilt Institute for Integrative Biosystems Research and Education, Vanderbilt University, Nashville, TN 37235, USA

Abstract

Systems-wide molecular phenomics is rapidly expanding through technological advances in instrumentation and bioinformatics. Strategies such as structural mass spectrometry, which utilizes size and shape measurements with molecular weight, serve to characterize the sum of molecular expression in biological contexts, where broad-scale measurements are made that are interpreted through big data statistical techniques to reveal underlying patterns corresponding to phenotype. The data density, data dimensionality, data projection, and data interrogation are all critical aspects of these approaches to turn data into salient information. Untargeted molecular phenomics is already having a dramatic impact in discovery science from drug discovery to synthetic biology. It is evident that these emerging techniques will integrate closely in broad efforts aimed at precision medicine.

Introduction

One of the emerging paradigms in systems, synthetic, and chemical biology is the ability to discern the entire molecular expression of a phenotypic sample, termed molecular phenomics, whereby the sum of factors influencing the biology (e.g. exposures, diseases, nutrition, social influences, etc.) are manifest in the consequential molecular landscape corresponding to that phenotype. Massive efforts in molecular phenomics are targeted in precision medicine to tailor prediction of medical outcome and treatment at an individual level [1,2]. To discern the systems-wide response requires massive numbers of measurements, and as such, approaches were largely developed around genome science, where parallel measurements are facilitated using gene-array technologies [3]. In medicine, recent advances include the development of coordinated data workflows correlating genomics data linked to the electronic medical records (EMR) of large cohorts of patients [4,5] and corresponding phenotype [6].

In large part, these efforts have centered on genomics data because of practical technological challenges associated with rapidly and broadly characterizing other molecular types (e.g. proteins, lipids, and metabolites) as a part of the overall diagnostic of biological status. This challenge is considerably exacerbated by the number of samples and/or sampling frequency

necessary to characterize the systems-wide molecular complement, because it can vary dramatically depending on location (e.g. cell, tissue, or regional pathology) and time (e.g. health versus disease, disease stage of progression, acute versus chronic exposure). These broad scale measurement challenges are being addressed with advances in several key technologies, chief among them mass spectrometry (MS, [7–9]). These instrumental advances when tied to advances in bioinformatics have resulted in new strategies aimed at molecular phenomics using a breadth of molecular type [10–12], or the sum of molecular expression resulting in specific phenotype. Note that EMR in concert with molecular phenomics measurement is critical in the translation of molecular characterization with predication and interpretation of health status, outcomes, and the development of treatment regimens for improved health outcomes. Massive population scale efforts for integrating MS with individual patient phenotyping, in concert with genomics, have recently been described [13]. Molecular phenomics data are then interpreted through identifying the molecular changes that are important to phenotype and/or by mapping the molecular complement onto biological pathways, or networks, to understand alterations in phenotype. Importantly, specific phenotypic studies demonstrating the potential for integrating other data modalities with EMR motivate incorporation of MS data into precision medicine workflows [13]. This report highlights some recent advances in structural MS aimed at obtaining rapid and untargeted molecular information in both spatially-resolved and temporally-resolved analyses and emerging bioinformatics approaches for phenotypic interpretation.

Data density and data dimensionality: opportunities in space and time

To assess molecular phenotype using systems-wide approaches requires cataloging the signals and abundances for as many molecular species as possible, which is the purview of integrated omics studies (i.e. the near simultaneous measurement of proteomics, lipidomics, metabolomics, etc.) using a combination of liquid-chromatography or gas-chromatography coupled with MS-based detection (LC–MS and GC–MS, respectively). Contemporary approaches in omics measurements typically reduce the complexity of the sample to one biomolecular class, where steps for protein or lipid purification are often performed so that the resulting MS signals will correspond primarily to those species, which facilitates identification. However, when multiple classes of molecules are of interest this conventional approach requires the analysis of multiple samples in series and adds significant time to both the sample preparation steps and also to the sample analysis workflows. Structural separations on the basis of ion mobility-MS (IM-MS, [14]) can mitigate this limitation through performing the molecular class separations in the instrument itself with minimal sample pretreatment. This is accomplished in IM-MS through separating these species on the basis on structure where mobility-mass correlations emerge from the data as a result of the structural preferences for different classes of molecules (Figure 1a) [15,16,17•]. As illustrated in Figure 1a, lipids on average adopt larger structures than peptides and likewise peptides adopt larger structures than carbohydrates. Through inspection of collision cross section versus mass-to-charge plots, rapid identification of the different species can be performed in the instrument in parallel without purification of an individual class of molecules before analysis. Recent reports have demonstrated high interlaboratory reproducibility for IM-MS collision cross-section measurements (<5% RSD for metabolites [18•] and <3% RSD for lipids [19•]) and the utility of cataloging correlations for different

studies, because of differences in physiology. Likewise, a significant fraction of drug candidates that potentially could be efficacious in humans are also deemed inappropriate to move to humans, because they fail in the animal model, again because of differences in physiology. Thus, human organoids to replace the animal models in this workflow are of critical importance. In these strategies, the synthetic human organ is exposed to drug or toxin and the molecular response to this perturbation is then read out directly from the organ as the secretome of the organ into the waste stream. Fast time resolution and integrated omics are critically important in these studies as the precise diagnostic molecular targets are oftentimes not known and acute versus chronic response is also desirable. Thus, IM-MS is rapidly finding wide spread adoption for the analytics to probe organoid response [45].

Data projection and data interrogation

In both spatially-resolved and time-resolved integrated omics measurements, it is typical to obtain 10s to 100s of thousands of molecular signals with IM-MS representing a myriad of molecular types from proteins to lipids to metabolites. Owing to this high data density, visualizing the data through low density projections is critical to understand prevailing, and sometimes subtle, molecular patterns [46]. Depending on the aims of a particular study, different strategies project high dimensional data in two dimensions to show different aspects of the information and/or statistical significance. We briefly highlight here three approaches of notable interest to distinguish patterns and guide the analysis of broad scale untargeted data. Although principle component analysis (PCA) techniques are often used for gross discrimination of samples on the basis of the most highly contributing factors to variance, PCA can mask subtle patterns of the individual molecular species in large datasets, that is, small molecular variations that are very impactful. A recently described method for projecting high dimensional temporally-resolved metabolomics data are cloud plots as illustrated in Figure 1c [47]. In this representation, LC retention time is projected on the *x*-axis where positive and negative abundance changes are plotted on the *y*-axis, statistical significance projected as a bubble, where diameter and shading of the of the bubble corresponds to log-fold change and statistical significance by a Welch *t*-test, respectively. Alternatively, patterns in complex data can be observed using neural network approaches to self-organize the data to reveal patterns from multidimensional descriptors [48^{••},49,50^{••},51]. One such approach termed molecular expression dynamics inspector (MEDI) is illustrated in Figure 1d highlighting molecular patterns observed in an animal model of drug addiction [48^{••}]. These techniques prioritize which molecules exhibit similar responses in their patterns across many experiments. Molecules which exhibit similar response patterns are clustered together in the two-dimensional heatmap and clusters of molecules that exhibit different response patterns are projected further apart. In this manner, variance is dispersed throughout the self-organized map so that small variations are equally as visually pronounced as large variations.

Following the characterization of molecular response to perturbation, it is oftentimes beneficial to understand how specific molecular species fit within canonical pathway and network models of biology. A noteworthy recent approach termed mummichog [52^{••}], fits the broadscale data into biological networks and then uses statistical means to guide the validation of molecular identity and affected pathways (Figure 1e). In subsequent network

analysis, this approach is significant in that the identification/validation step is partially coupled, rather than using strictly conservative *bona fide* identifications, which is a time intensive step. In particular, for metabolic analyses, this approach can rapidly facilitate prioritizing which of a multitude of molecules corresponding to the same or similar atomic compositions that specific signals correspond, which at present is challenging to perform without additional complementary analytical data.

Concluding remarks and outlook

MS approaches such as IM-MS facilitate integrated omics and in particular, molecular phenomics is redefining the approaches used to understand health, disease, and discovery. The main challenge is the massive data density that these broad scale and rapid measurement techniques provide. Thus, integrating these measurements with bioinformatic and biostatistical methods to recognize the patterns quickly is a difficult task. Contemporary innovations in both measurement instrumentation and data interrogation drive the expansion of what is possible in molecular phenomics and ultimately should have far reaching implications in areas such as personalized medicine.

Acknowledgements

We thank Stacy D. Sherrod, Cody R. Goodwin, collaborators, and former colleagues for stimulating discussions that have shaped our thinking in these areas. Support for aspects of this work was provided by the National Institutes of Health/National Institute for General Medical Sciences (5R01GM092218), the Vanderbilt University Center for Innovative Technology, Vanderbilt Institute of Chemical Biology; the Vanderbilt Institute for Integrative Biosystems Research and Education; and Vanderbilt University.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. *Science*. 2004; 306:640–643. [PubMed: 15499008]
2. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. *Per Med*. 2013; 10:565–576. [PubMed: 25342952]
3. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L. Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science*. 2001; 292:929–934. [PubMed: 11340206]
4. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, Masys DR. Development of a large-scale deidentified DNA biobank to enable personalized medicine. *Clin Pharmacol Ther*. 2008; 84:362–369. [PubMed: 18500243]
5. Ritchie MD, Denny JC, Crawford DC, Ramirez AH, Weiner JB, Pulley JM, Basford MA, Brown-Gentry K, Balser JR, Masys DR, et al. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. *Am J Hum Genet*. 2010; 86:560–572. [PubMed: 20362271]
6. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, Wang D, Masys DR, Roden DM, Crawford DC. PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*. 2010; 26:1205–1210. [PubMed: 20335276]
7. Patti GJ, Yanes O, Siuzdak G. Innovation: metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol*. 2012; 13:263–269. [PubMed: 22436749]

8. Fuhrer T, Zamboni N. High-throughput discovery metabolomics. *Curr Opin Biotechnol.* 2015; 31:73–78. [PubMed: 25197792]
9. May JC, Goodwin CR, McLean JA. Ion mobility-mass spectrometry strategies for untargeted systems, synthetic, and chemical biology. *Curr Opin Biotechnol.* 2015; 31:117–121. [PubMed: 25462629]
10. Holmes E, Wijeyesekera A, Taylor-Robinson SD, Nicholson JK. The promise of metabolic phenotyping in gastroenterology and hepatology. *Nat Rev Gastroenterol Hepatol.* 2015; 12:458–471. [PubMed: 26194948]
11. Nicholson JK, Holmes E, Kinross JM, Darzi AW, Takats Z, Lindon JC. Metabolic phenotyping in clinical and surgical environments. *Nature.* 2012; 491:384–392. [PubMed: 23151581]
12. Holmes E, Loo RL, Stamler J, Bictash M, Yap IK, Chan Q, Ebbels T, De Iorio M, Brown IJ, Veselkov KA, Daviglus ML, Kesteloot H, Ueshima H, Zhao L, Nicholson JK, Elliott P. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature.* 2008; 453:396–400. [PubMed: 18425110]
13. Bowton EA, Collier SP, Wang X, Sutcliffe CB, Van Driest SL, Couch LJ, Herrera M, Jerome RN, Slebos RJ, Alborn WE, Liebler DC, McNaughton CD, Mernaugh RL, Wells QS, Brown NJ, Roden DM, Pulley JM. Phenotype-driven plasma biobanking strategies and methods. *J Pers Med.* 2015; 5:140–152. [PubMed: 26110578]
14. May JC, McLean JA. Ion mobility-mass spectrometry: time-dispersive instrumentation. *Anal Chem.* 2015; 87:1422–1436. [PubMed: 25526595]
15. Woods AS, Ugarov M, Egan T, Koomen J, Gillig KJ, Fuhrer K, Gonin M, Schultz JA. Lipid/peptide/nucleotide separation with MALDI-ion mobility-TOFMS. *Anal Chem.* 2004; 76:2187–2195. [PubMed: 15080727]
16. Fenn LS, McLean JA. Biomolecular structural separations by ion mobility-mass spectrometry. *Anal Bioanal Chem.* 2008; 391:905–909. [PubMed: 18320175]
- 17•. May JC, Goodwin CR, Lareau NM, Leapfrog KL, Morris CB, Kurulugama RT, Mordehai A, Klein C, Barry W, Darland E, Overney G, Imatani K, Stafford GC, Fjeldsted JC, McLean JA. Conformational ordering of biomolecules in the gas-phase: nitrogen collision cross-sections measured on a prototype high resolution drift tube ion mobility-mass spectrometer. *Anal Chem.* 2014; 86:2107–2116. [PubMed: 24446877] [An ion mobility atlas of collision cross sections for many molecular classes measured in nitrogen using a prototype commercial IM-MS platform.]
- 18•. Paglia G, Williams JP, Menikarachchi L, Thompson JW, Tyldesley-Worster R, Halldórsson S, Rolfsson O, Moseley A, Grant D, Langridge J, Palsson BO, Astarita G. Ion mobility derived collision cross sections to support metabolomics applications. *Anal Chem.* 2014; 86:3985–3993. [PubMed: 24640936] [Interlaboratory comparison of the reproducibility of ion mobility measurements in metabolomics.]
- 19•. Paglia G, Angel P, Williams JP, Richardson K, Olivos HJ, Thompson JW, Menikarachchi L, Lai S, Walsh C, Moseley A, Plumb RS, Grant DF, Palsson BO, Langridge J, Geromanos S, Astarita G. Ion mobility-derived collision cross section as an additional measure for lipid fingerprinting and identification. *Anal Chem.* Interlaboratory comparison of the reproducibility of ion mobility measurements in lipidomics. 2015; 87:1137–1144. [PubMed: 25495617]
- 20•. Hofmann J, Hahm HS, Seeberger PH, Pagel K. Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature.* 2015; 526:241–244. [PubMed: 26416727] [Presentation of a unified strategy for rapid characterization of carbohydrates and synthetic carbohydrates.]
- 21•. Both P, Green AP, Gray CJ, Sardžik R, Voglmeir J, Fontana C, Austeri M, Rejzek M, Richardson D, Field RA, Widmalm G, Flitsch SL, Eyers CE. Discrimination of epimeric glycans and glycopeptides using IM-MS and its potential for carbohydrate sequencing. *Nat Chem.* 2014; 6:65–74. [PubMed: 24345949] [The utility of structural separations to distinguish carbohydrate epimers in fine structural analyses.]
22. Gaye MM, Kurulugama R, Clemmer DE. Investigating carbohydrate isomers by IMS-CID-IMS-MS: precursor and fragment ion cross-sections. *Analyst.* 2015; 140:6922–6932. [PubMed: 26306702]
23. Li H, Bendiak B, Siems WF, Gang DR, Hill HH Jr. Determining the isomeric heterogeneity of neutral oligosaccharide-alditols of bovine submaxillary mucin using negative ion traveling wave ion mobility mass spectrometry. *Anal Chem.* 2015; 87:2228–2235. [PubMed: 25594283]

24. Struwe WB, Benesch JL, Harvey DJ, Pagel K. Collision cross sections of high-mannose N-glycans in commonly observed adduct states — identification of gas-phase conformers unique to [M-H](S) ions. *Analyst*. 2015; 140:6799–6803. [PubMed: 26159123]
25. Harvey DJ, Sobott F, Crispin M, Wrobel A, Bonomelli C, Vasiljevic S, Scanlan CN, Scarff CA, Thalassinou K, Scrivens JH. Ion mobility mass spectrometry for extracting spectra of N-glycans directly from incubation mixtures following glycan release: application to glycans from engineered glycoforms of intact, folded HIV gp120. *J Am Soc Mass Spectrom*. 2010; 22:568–581. [PubMed: 21472575]
- 26•• Kyle, JE.; Zhang, X.; Weitz, KK.; Monroe, ME.; Ibrahim, YM.; Moore, RJ.; Cha, J.; Sun, X.; Lovelace, ES.; Wagoner, J.; Polyak, SJ.; Metz, TO.; Dey, SK.; Smith, RD.; Burnum-Johnson, KE.; Baker, ES. Uncovering biologically significant lipid isomers with liquid chromatography, ion mobility spectrometry and mass spectrometry.. *Analyst*. 2016. <http://dx.doi.org/10.1039/C5AN02062J> [A compendium of the separation characteristics of lipid isomers using ion mobility-mass spectrometry.]
27. Milic I, Kipping M, Hoffmann R, Fedorova M. Separation and characterization of oxidized isomeric lipid-peptide adducts by ion mobility mass spectrometry. *J Mass Spectrom*. 2015; 50:1386–1392. [PubMed: 26634972]
28. Zhang F, Guo S, Zhang M, Zhang Z, Guo Y. Characterizing ion mobility and collision cross section of fatty acids using electrospray ion mobility mass spectrometry. *J Mass Spectrom*. 2015; 50:906–913. [PubMed: 26349645]
29. Groessl M, Graf S, Knochenmuss R. High resolution ion mobility-mass spectrometry for separation and identification of isomeric lipids. *Analyst*. 2015; 140:6904–6911. [PubMed: 26312258]
30. Caprioli RM. Imaging mass spectrometry: enabling a new age of discovery in biology and medicine through molecular microscopy. *J Am Soc Mass Spectrom*. 2015; 26:850–852. [PubMed: 25801587]
31. Heeren RM, Smith DF, Stauber J, Kükrer-Kaletas B, MacAleese L. Imaging mass spectrometry: hype or hope? *J Am Soc Mass Spectrom*. 2009; 20:1006–1014. [PubMed: 19318278]
32. Trim PJ, Henson CM, Avery JL, McEwen A, Snel MF, Claude E, Marshall PS, West A, Princivale AP, Clench MR. Matrix-assisted laser desorption/ionization-ion mobility separation-mass spectrometry imaging of vinblastine in whole body tissue sections. *Anal Chem*. 2008; 80:8628–8634. [PubMed: 18847214]
33. Cole LM, Clench MR. Mass spectrometry imaging tools in oncology. *Biomark Med*. 2015; 9:863–868. [PubMed: 26330284]
34. Spraggins JM, Rizzo DG, Moore JL, Rose KL, Hammer ND, Skaar EP, Caprioli RM. MALDI FTICR IMS of intact proteins: using mass accuracy to link protein images with proteomics data. *J Am Soc Mass Spectrom*. 2015; 26:974–985. [PubMed: 25904064]
35. McLean JA, Ridenour WB, Caprioli RM. Profiling and imaging of tissues by imaging ion mobility-mass spectrometry. *J Mass Spectrom*. 2007; 42:1099–1105. [PubMed: 17621390]
36. Jackson SN, Ugarov M, Egan T, Post JD, Langlais D, Albert Schultz J, Woods AS. MALDI-ion mobility-TOFMS imaging of lipids in rat brain tissue. *J Mass Spectrom*. 2007; 42:1093–1098. [PubMed: 17621389]
37. Bennett RV, Gamage CM, Galhena AS, Fernández FM. Contrast-enhanced differential mobility-desorption electrospray ionization-mass spectrometry imaging of biological tissues. *Anal Chem*. 2014; 86:3756–3763. [PubMed: 24650201]
38. Jackson SN, Barbacci D, Egan T, Lewis EK, Schultz JA, Woods AS. MALDI-ion mobility mass spectrometry of lipids in negative ion mode. *Anal Methods*. 2014; 6:5001–5007. [PubMed: 24999374]
39. Matusch A, Fenn LS, Depboylu C, Kliezt M, Strohmer S, McLean JA, Becker JS. Combined elemental and biomolecular mass spectrometry imaging for probing the inventory of tissue at a micrometer scale. *Anal Chem*. 2012; 84:3170–3178. [PubMed: 22413784]
40. Djidja MC, Claude E, Snel MF, Scriven P, Francese S, Carolan V, Clench MR. MALDI-ion mobility separation-mass spectrometry imaging of glucose-regulated protein 78 kDa (Grp78) in

- human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections. *J Proteome Res.* 2009; 8:4876–4884. [PubMed: 19673544]
41. Stauber J, MacAleese L, Franck J, Claude E, Snel M, Kaletas BK, Wiel IM, Wisztorski M, Fournier I, Heeren RM. On-tissue protein identification and imaging by MALDI-ion mobility mass spectrometry. *J Am Soc Mass Spectrom.* 2010; 21:338–347. [PubMed: 19926301]
42. Baker ES, Burnum-Johnson KE, Jacobs JM, Diamond DL, Brown RN, Ibrahim YM, Orton DJ, Piehowski PD, Purdy DE, Moore RJ, Danielson WF 3rd, Monroe ME, Crowell KL, Slys GW, Gritsenko MA, Sandoval JD, Lamarche BL, Matzke MM, Webb-Robertson BJ, Simons BC, McMahon BJ, Bhattacharya R, Perkins JD, Carithers RL Jr, Strom S, Self SG, Katze MG, Anderson GA, Smith RD. Advancing the high throughput identification of liver fibrosis protein signatures using multiplexed ion mobility spectrometry. *Mol Cell Proteomics.* 2014; 13:1119–1127. [PubMed: 24403597]
43. Marasco CC, Enders JR, Seale KT, McLean JA, Wikswo JP. Real-time cellular exometabolome analysis with a microfluidic-mass spectrometry platform. *PLOS ONE.* 2015; 10:e0117685. [PubMed: 25723555]
44. Shi M, Majumdar D, Gao Y, Brewer BM, Goodwin CR, McLean JA, Li D, Webb DJ. Glia co-culture with neurons in microfluidic platforms promotes the formation and stabilization of synaptic contacts. *Lab Chip.* 2013; 13:3008–3021. [PubMed: 23736663]
45. Sherrod SD, McLean JA. Systems-wide high-dimensional data acquisition and informatics using structural mass spectrometry strategies. *Clin Chem.* 2016; 62:77–83. [PubMed: 26453699]
46. May, JC.; McLean, JA. Advanced multidimensional separations in mass spectrometry: navigating the big data deluge.. *Annu Rev Anal Chem.* 2016. <http://dx.doi.org/10.1146/annrev-anchem-071015-041734>
47. Ivanisevic J, Elias D, Deguchi H, Averell PM, Kurczy M, Johnson CH, Tautenhahn R, Zhu Z, Watrous J, Jain M, Griffin J, Patti GJ, Siuzdak G. Arteriovenous blood metabolomics: a readout of intra-tissue metabostasis. *Sci Rep.* 2015; 12757
- 48•. Goodwin CR, Sherrod SD, Marasco CC, Bachmann BO, Schramm-Saptya N, Wikswo JP, McLean JA. Phenotypic mapping of metabolic profiles using self-organizing maps of high-dimensional mass spectrometry data. *Anal Chem.* 2014; 86:6563–6571. [PubMed: 24856386] [A description of the MEDI workflow for constructing self-organizing maps of high-dimensional ion mobility-mass spectrometry data.]
49. Derewacz DK, Covington BC, McLean JA, Bachmann BO. Mapping microbial response metabolomes for induced natural product discovery. *ACS Chem Biol.* 2015; 10:1998–2006. [PubMed: 26039241]
- 50•. Goodwin CR, Covington BC, Derewacz DK, McNeese CR, Wikswo JP, McLean JA, Bachmann BO. Structuring microbial metabolic responses to multiplexed stimuli via self-organizing metabolomics maps. *Chem Biol.* 2015; 22:661–670. [PubMed: 25937311] [A critical comparison of conventional PCA analyses in the context of MEDI self-organizing map protocols. This work introduces new concepts for natural product discovery from microbial systems using MEDI to prioritize features of interest.]
51. Kumar KK, Goodwin CR, Uhouse MA, Bornhorst J, Schwerdtle T, Aschner M, McLean JA, Bowman AB. Untargeted metabolic profiling identifies interactions between Huntington's disease and neuronal manganese status. *Metallomics.* 2015; 7:363–370. [PubMed: 25599126]
- 52•. 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. [PubMed: 23861661] [A new approach for identifying and validating metabolite data using pathway and network mapping. This work potentially provides a means by which false discovery rates may be estimated in metabolomics.]

Figure 1.

The development of comprehensive and untargeted strategies for molecular characterization are driven by advances in both the analytical technologies used to obtain multidimensional measurements and the informatics techniques used to visualize and distill the data into phenotypical information, with some relevant examples shown here. **(a)** The integration of IM-MS with imaging studies results in a comprehensive size-mass dispersion at each spatial coordinate. **(b)** This capability is particularly useful for mapping the location of analytes which cannot be differentiated on the basis of MS alone, as in the case with two nominally-same mass peptides present in different spatial locations of brain tissue. **(c)** Novel data

visualization strategies such as this cloud plot highlight metabolite features from LC-MS data which vary significantly between arterial and venous blood ($p < 0.01$). (d) MEDI plots utilize self-organizing maps to organize data into patterns of molecules that behave similarly. In this example, metabolite differences between behaviorally cocaine-addicted and cocaine-naïve rats self-organize into different areas within the map. (e) Translating untargeted metabolomics data into useful systems information traditionally requires that all metabolites be identified before pathway analysis. The *mummichog* approach bypasses the time-intensive identification step by mapping all possible metabolite matches, using clustering of enrichment to validate true versus false positives. [Panel (b) adapted from Ref. [41]. With permission of Springer. Panel (c) adapted and reproduced with permission from Ref. [47] ©Nature Publishing group. Panel (d) adapted from Ref. [48**] ©2014 American Chemical Society. Panel (e) adapted and reproduced with permission from Ref. [52**]. ©2013 Li *et al.*].