Full text provided by www.sciencedirect.com

SCIENCE DIRECT

The role of analytical sciences in medical systems biology

Jan van der Greef[1,2,3,*], Paul Stroobant[2] and Rob van der Heijden[3]

Medical systems biology has generated widespread interest because of its bold conception and exciting potential, but the field is still in its infancy. Although there has been tremendous progress achieved recently in generating, integrating and analysing data in the medical and pharmaceutical field, many challenges remain, especially with respect to the crucial core technologies required for analytical characterization. This review briefly summarizes these aspects for metabolomics, proteomics, data handling and multivariate biostatistics.

Addresses
[1] TNO Systems Biology, PO Box 360, 3700 AJ Zeist, The Netherlands
[2] Beyond Genomics Inc., 40 Bear Hill road, Waltham, MA 02451, USA
[3] Center for Medical Systems Biology, LACDR, Gorleaus Laboratories, 3700 RA Leiden, The Netherlands
*e-mail: vandergreef@pharma.tno.nl

Current Opinion in Chemical Biology 2004, **8**:559–565

This review comes from a themed section on Analytical techniques
Edited by Renato Zenobi and Fred Regnier

Available online 25th August 2004

1367-5931/$ – see front matter
© 2004 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.cbpa.2004.08.013

Abbreviations

CE	capillary electrophoresis
ECD	electrochemical detector
HR-MAS	high resolution magic angle spinning
MALDI	matrix-assisted laser desorption ionization

Introduction

Interest in the application of systems biology to the life sciences has become widespread over the past five years, although systems perspectives have been applied to many sciences, ranging from biology to cosmology, since at least the beginning of the 20th century. There is a striking commonality between various scientific domains, which each describe the interconnectivity and interdependence of the components of the system and which each recognize that new properties of systems are revealed as increasing levels of complexity are studied. The traditional approach in biology of relating limited observational data to a holistic model is now being complemented with a new wave of life science technologies that are beginning to yield detailed molecular and mechanistic information on an unprecedented scale [1–3].

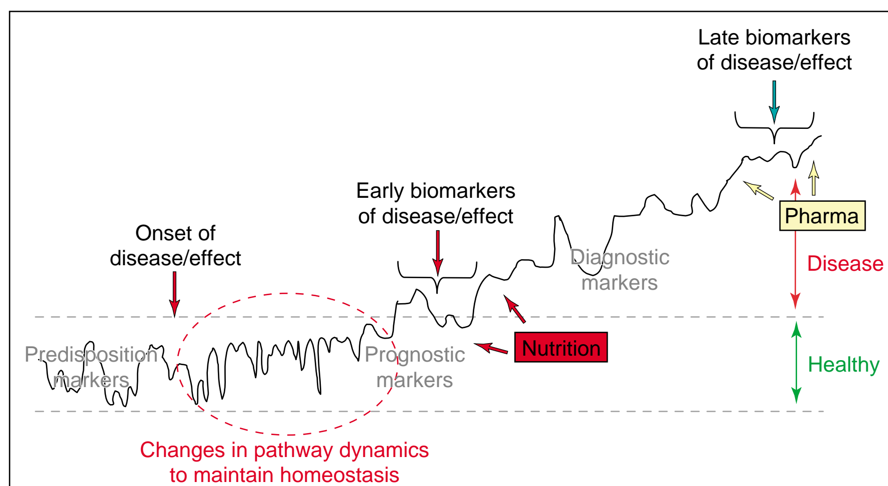
The definition of systems biology is highly variable in the literature. Our working definition for medical systems biology is 'studying biology as an integrated system of genetic, protein, metabolite, cellular and pathway events that are in flux and interdependent'. This does not imply that all elements need to be measured for systems studies but that systems thinking and systems-based analytical strategies are key. For medical systems biology, human samples are often restricted to body fluids and consequently pathway relationships between the components may not be revealed without further study, in contrast to the study of animal models or single cell types where informative pathway information can be gained at an early stage of analysis.

Systems biology requires the integration of biology, medicine, mathematics and chemistry with biostatistics and bioinformatics to transform complex and diverse datasets into useful knowledge [4–8,9••], and systems biology approaches are being increasingly applied in the fields of microbiology, and plant and medical sciences. We focus here on analytical sciences in medical systems biology; in particular, the status and challenges for metabolomics and proteomics. Data pre-processing before the important steps of biostatistics, bioinformatics and modeling is briefly considered and notable references for data evaluation and integration steps are given.

An important principal for the discovery of biomarkers for drug discovery, drug development and disease diagnosis [10••–12••,13,14] is that multi-factorial disease (Figure 1) involves studying complex and dynamic biomarker patterns rather than a single biomarker such as cholesterol, prostate-specific antigen or glucose. Although this evolution from single biomarker strategies towards those employing biomarker patterns involves managing new levels of complexity in data generation and analysis, such a change will be essential to adequately characterize and ultimately understand the etiology and progression of disease states. It is also important to recognize that different biomarker profiles are found at the onset of a disease versus the late stage where symptoms and indirect effects are prominent. Therefore understanding transitional biomarker profiles in terms of mechanism and validation is crucial.

Implicit in this assumption is that the analytical capabilities to reliably establish biomarker patterns are capable of identifying different levels of complexity that impact the practice of medicine. However, despite recently renewed interest in biomarkers, there are still critical gaps in the analytical capabilities at the levels of primary data acquisition and analysis, especially for proteomics [15••]. This review addresses critical issues

Figure 1



The development of disease from healthy (homeostasis within black dotted lines) to sub-optimal health and eventually an overt disease state. Biomarker patterns (for graphical reasons represented as a single line) are essential to describe the changes from normality to dysfunction. Adapted from [11••] with kind permission. Copyright 2003, Kluwer Academic Publishers.

confronting medical systems biology. In particular, we focus on analytical issues, describing the status of metabolomics, which has a longer history on profiling and pattern recognition [11••], followed by proteomics and data evaluation.

Analytical tools for metabolome analysis

Recent developments in both NMR and hyphenated MS technology have extended the crucial roles of these tools, although alternative approaches with great potential are also now emerging. Major improvements in NMR include the introduction of extremely high field magnets (currently up to 900 MHz), cryoprobe technology and the measurement of microsamples. When using the cryoprobe in a high field NMR (800 MHz) signal-to-noise ratios are improved by a factor of 15–20, as compared with a standard probe in a 600 MHz apparatus. Consequently, acquisition times can be reduced dramatically. This has opened the way to including ^{13}C NMR spectroscopy as an additional tool for metabolic profiling [16]. For this, spectra acquisition now requires 30 min per sample instead of overnight measurements.

Nanoprobe technology allows the analysis of samples in volumes as small as 25 μl, and has been applied in the metabolic profiling of cerebrospinal fluid microdialysates obtained from several brain regions of freely moving rats [17]. Additionally, high-resolution magic-angle spinning (HR-MAS) NMR spectroscopy provides information on metabolite concentration and compartmentalization in intact tissue samples. HR-MAS produces narrow-line-width spectra comparable to those from solid-state NMR. Some recent applications of NMR and HR-MAS NMR are found in [18–20].

The latest generation of Fourier transform ion cyclotron resonance mass spectrometers represents a quantum leap forward in the robust capability of mass spectrometers for high-resolution metabolite analysis. In standard analytical experiments, resolutions of > 100 000 (50% valley definition), yielding accurate masses within 0.5 ppm, are easily obtainable without compromising the chromatographic conditions (including nano-LC) and without the use of internal calibrants [21••,22••]. The accurate masses obtained give elemental compositions that provide a major step towards metabolite identification.

Although such high mass resolution might minimize the need for chromatographic separation before introduction to the mass spectrometer, this is probably not possible. Direct infusion of the sample into the mass spectrometer would be predicted to lead to suppression of ionization in the electrospray inlet and the formation of adducts. Interestingly, however, an optimized protocol including the use of the non-salt-based buffer tricine for metabolome analysis in yeast by direct infusion has been reported [23]. On the other hand, enhanced chromatographic resolution from the use of a long (90 cm) monolithic capillary column effectively reduced ionization suppression so that several hundred peaks could be detected in an *Arabidopsis thaliana* extract [24]. With limited sample availability, it is essential to reduce the dimensions of the LC columns to enable increased sensitivity. Capillary LC connected to a quadrupole-time-of-flight mass spectrometer generated about 2000 different mass signals in extracts of *Arabidopsis* [25]. Highly polar compounds, normally missed by standard reversed phase chromatography, are detected by hydrophilic interaction chromatography. In combination with electrospray

ionisation ion-trap MS (in positive and negative ion mode), oligosaccharides, aminosugars and sugar nucleotides were detected [26].

The highest separation efficiency before MS detection is obtained with capillary electromigration methods. From *Bacillus subtilis*, 1692 metabolites were analysed by using a combination of three capillary electrophoresis (CE)-MS methods. To maximize detection sensitivity, the mass spectrometer scan range was narrowed to 30 m/z; about 30 injections were necessary to cover the necessary mass range [27]. Pressure-assisted CE-MS allowed the analysis of multivalent ions as nucleotides and CoA-derivatives [28]. By applying air pressure to the inlet of a non-charged polymer-coated capillary a conductive liquid junction between the capillary and the MS was maintained. Selectivity originating from the mass spectrometer detector (i.e. by precursor ion scanning for specific fragment ions) was applied in the CE–MS analysis of sugar nucleotides [29]. Furthermore, CE is an excellent tool for the on-line pre-concentration of metabolites; dynamic pH junction, sweeping and dynamic pH junction sweeping are three recent, complementary methods for electrokinetic focusing (reviewed in [30••]). Further increase in separation efficiency is obtained by using CE in a two-dimensional approach [31] hyphenated micro-LC (with a monolithic column) and CE. Sweeping and a dynamic pH junction were employed to interface the two dimensions. The method was evaluated for 54 standard metabolites and applied to a *Bacillus subtilis* extract.

As an alternative to the commonly used MS and photodiode array detection systems, the use of an electrochemical detector (ECD) sheds new light on the metabolome by the very sensitive detection of redox active compounds. Recent developments in this area are the coulometric array technology (16 channels) including the option of combining it with gradient elution (so far, ECD is mainly used in isocratic systems) [32,33].

Another approach is the use of Fourier-transform infrared spectroscopy. This economical and rapid technique was applied in an animal model study for idiosyncratic toxicity [34,35]. For body fluid, tissue (and whole organism) profiling, Raman spectroscopy has several applications (reviewed in [36]). Stable isotope-based dynamic metabolic profiling (SIDMAP) proved to be a powerful approach, enabling a better understanding of changes in substrate flow as a basis for drug mechanisms and disease [37]. In conclusion, tremendous progress has been made in the dynamic range and coverage capabilities of metabolite profiling. However, despite the availability of high-resolution NMR and MS/MS, the comprehensive identification of metabolites remains a key challenge, which awaits the better integration of NMR and MS/MS data as well as the development of novel approaches. In medical systems biology, the metabolome coverage of body fluids is still limited and the integration of techniques in a platform is mandatory [11••].

Challenges in developing analytical tools for proteome analysis

Despite the crucial role that proteomics [38] will play in medical systems biology, the field currently faces huge technical challenges that are a consequence of the complex, dynamic, idiosyncratic and largely uncharacterized proteome present in most samples of human fluids, cells or tissues. Authoritative reviews [15••,39••,40] summarizing the situation for the analysis of human plasma and serum illustrate the magnitude of the problems. After extensive recent efforts to evolve better methodologies via such approaches as fractionation schemes or the isolation of targeted protein or peptide classes, the fact remains that even after over 30 years of productive work in isolating, characterizing and quantifying proteins, the large-scale proteomics vision of global and comprehensive protein analysis remains largely unfulfilled.

In particular, there are five critical, unsolved problems that need to be adequately addressed before proteomics can begin to realize its potential for medical systems biology, and before significant progress can be made with second-generation proteomics areas of study such as intermolecular interactions, biological function and structures. First, is the lack of accurate and reproducible quantification of all components detected, a requirement which is fundamental to systems biology approaches. Second, is the need for extensive coverage of the proteome, since most current global approaches probably detect less than 0.1% of the protein species present in a complex sample, largely due to the dynamic range of more than 10 orders of magnitude between the most abundant protein albumin, and the least abundant proteins measured clinically [39••]. Third, although protein catalogue information can be useful, the comparative nature of systems biology demands differential analyses, in which changes in multiple species across many samples are routinely measured. Fourth, is the important requirement for identifying differentially expressed protein species at the structural level, including precise amino acid sequence as well as the nature and structure of post-translational modifications. The widely adopted approach of attempting to map a limited complement of tryptic peptides to specific protein species using the currently substantially incomplete proteome databases will not meet the needs of medical systems biology for the future. Finally, the work flows for these primary capabilities need to be robust, cost- and time-effective, and capable of yielding data from hundreds of samples, even at a discovery stage, before focused and high-throughput assays can be developed for key biomarker subsets.

On a positive note, several potentially breakthrough approaches in proteomics are emerging related to

quantification improvements via labeling procedures [41••]. An example of these efforts is the recently presented [42••] new multiplexing tagging reagent termed iTRAQ (isobaric tags for relative and absolute quantification), reported to work via an efficient global peptide labeling strategy that can derivatize all peptides, including those which are post-translationally modified.

Other promising and innovative approaches include direct tissue profiling and imaging mass spectrometry, in which MALDI MS analysis of thin tissue sections allows a detailed assessment of the complex protein pattern within tissue samples [43••], and the development of natural protein microarrays for diagnosing cancer based on antibody responses to tumour antigens [44••].

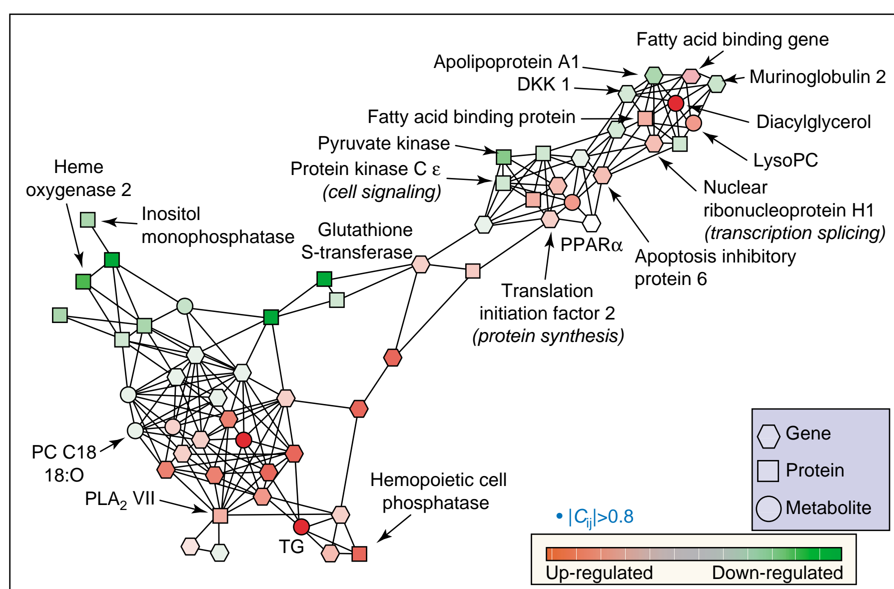
From data to information and knowledge

The avalanche of data resulting from applying '-omics' technologies to complex biological systems requires several crucial steps before integration to a systems level of understanding. Data pre-processing is critical to success in this area, the first issue being the peak (variable) alignment. A significant problem in MS and NMR measurements comes from the fact that the position of a peak has a measurement error and/or a shift in position as seen in the NMR dependency on the chemical composition of a sample. This is important when profiling samples of varying composition. With urine it has been necessary to align peaks using non-linear alignment procedures [45], peak picking [46] or by using binning techniques with variable bin sizes [47] to obtain useable data. In many cases, deconvolution or curve resolution plays an important role and can be effectively performed as in LC-MS and GC-MS [48–50]. Depending on the settings and the experiment, choosing a particular approach can be important to maintain optimum resolution [11••,47,51]. In addition, normalization and scaling of the data are crucial, and typical choices such as autoscaling, mean centering, or variable stability scaling [52] can only be made after examining the analytical techniques used in the experiment. Such choices are even more complex in combination with multi-way analysis [53].

It is important in most medical systems biology studies to apply unsupervised techniques such as pattern recognition (principal component analysis being the classical approach) for cluster analysis before using clinical diagnostic information and moving into non-supervised methods. New approaches for high dimensionality data have been described such as independent component analysis [54••]. Clustering based on a variety of distinct approaches is typically needed to obtain sufficient information to reveal biomarkers of disease.

When data are available at transcriptomics, proteomics and metabolomics levels, integration via correlation networks has been shown to yield information concerning the interconnectivity of biological elements. One such study focused on the detection of early markers of atherosclerosis in a transgenic model (Figure 2). Correlation networks not only demonstrate biomarker discovery but they also provide information on the biological pathways involved, even though they employ a *de novo* approach with broad profiling using '-omics' technologies. The

Figure 2



The correlation network of all biomarkers found in a study on the early onset of atherosclerosis in the ApoE3 transgenic mouse model. Adapted from [55••] with kind permission. Copyright 2004, Liebert publishers.

outcome of such an experiment demonstrates that developing a drug focused on a single target will only impact part of the disease and strongly supports studies centred on multi-dimensional pharmacology or multi-target pharmacology.

Although progress in the field of nutrition, which involves studying multi-component mixture responses in a complex biological system [56], has previously been limited by analytical challenges, newly developing systems biology platforms are now providing an excellent opportunity for further scientific progress. Systems biology in medical sciences has been mainly focused on statistical mining [57••] and developing or linking annotated databases of different origin (not discussed in this review). But it is important to note recent dynamic modeling innovations in cellular systems and the continuum between statistical mining and modeling. For transcriptomics and, in particular, in the analysis of metabolic networks, there are numerous excellent studies, which include the importance of flux considerations [58–60]. Studying dynamic systems also prompts the development of appropriate tools [61,62].

Conclusions and future perspectives

The rapid ongoing technological improvements in the fields of MS and NMR, in which instruments with higher resolution and sensitivity become available, will continue to improve the coverage of the metabolome and proteome from a chemical diversity, coverage and concentration perspective. The optimization of sample preparation and separation methods in hyphenated strategies as well as miniaturization to allow smaller sample volumes or to cover single cell (nano) systems biology studies [63], will allow a direct link to important modeling efforts at the cellular level.

Integrating new biological perspectives in the design of medical systems biology studies will be increasingly important. Because a disease state is the consequence of the deregulation of communication and subsequent changes in dynamics and loss of homeostasis, analytical methods capable of measuring dynamic systems and the appropriate statistical and modeling tools to analyse longitudinal studies must be further developed. Although human clinical study samples are mostly restricted to body fluids, limiting a full systems biology approach, the identification of candidate biomarkers to describe this loss of homeostasis will hopefully greatly facilitate the detection of early diagnostic and prognostic markers following additional focused studies.

A logical next step in the field will be to acknowledge the influence of the mind in human biology and to integrate quantitative techniques in neuroscience such as non-invasive brain imaging technologies and psychological profiling into medical systems biology [64]. Such approaches will include more detailed studies of extreme physiological stress, and will also involve understanding the mechanisms of effects of thought patterns and placebo effects on health status. Although the field is still in its infancy, analytical sciences integrated with many other biosciences will have a bright new future in life science research.

Acknowledgement

This study was supported by the Centre for Medical Systems Biology (CMSB), a centre of excellence approved by the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research (NWO).

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Auffray C, Imbeaud S, Roux-Rouquie M, Hood L: **From functional genomics to systems biology: concepts and practices**. *Compt Rend Biol* 2003, **326**:879-892.

2. Kell DB, Oliver SG: **Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era.** *Bioessays* 2004, **26**:99-105.

3. Stevens CF: **Systems biology versus molecular biology**. *Curr Biol* 2004, **14**:R51-R52.

4. Kitano H: **Looking beyond the details: a rise in system-oriented approaches in genetics and molecular biology**. *Curr Genet* 2002, **41**:1-10.

5. Kitano H: **Systems biology: a brief overview**. *Science* 2002, **295**:1662-1664.

6. Noble D: **The rise of computational biology**. *Nat Rev Mol Cell Biol* 2002, **3**:460-463.

7. Hood L: **Systems biology and cancer**. *Clin Canc Res* 2003, **9**:6267S-6267S.

8. Pennisi E: **Systems biology: tracing life's circuitry**. *Science* 2003, **302**:1646.

9. •• Ideker T: **Systems biology 101- what you need to know**. *Nat Biotechnol* 2004, **22**:473-475.
The research vision from statistical mining to modeling from a systems perspective is clearly illustrated and described.

10. •• Nicholson JK, Connelly J, Lindon JC, Holmes E: **Metabonomics: a platform for studying drug toxicity and gene function**. *Nat Rev Drug Disc* 2002, **1**:153-161.
A good conceptual overview on NMR-based metabolomics in the field of toxicology.

11. •• Van der Greef J, Davidov E, Verheij ER, Vogels J, van der Heijden R, Adourian AS, Oresic M, Marple EW, Naylor S: **The role of metabolomics in systems biology**. In *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Edited by Harrigan GG, Goodacre R. Boston, Dordrecht, London: Kluwer Academic Publishers; 2003:170-198.
This chapter describes the history of metabolomics and the relevant new developments for systems biology. The book as a whole gives an excellent perspective on width and innovation in metabolomics. See also [20,33,35].

12. •• Frank R, Hargreaves R: **Clinical biomarkers in drug discovery and development**. *Nat Rev Drug Disc* 2003, **2**:566-580.
Clinical biomarker research is described and illustrated with clear examples including imaging techniques and others.

13. Rosenkranz B: **Biomakers and surrogate endpoints in clinical drug development: as long as their limitations are kept in mind, biomarkers and surrogate endpoints have the potential to save sponsors time and money**. *Appl Clin Trials* 2003, **July**:30-40.

14. Lindon JC, Holmes E, Nicholson JK: **Metabonomics and its role in drug development and disease diagnosis**. *Expert Rev Mol Diagnostics* 2004, **4**:189-199.

15. Diamandis EP: **Analysis of serum proteomic patterns for early •• cancer diagnosis: Drawing attention to potential problems: drawing attention to potential problems**. *J Natl Cancer Inst* 2004, **96**:353-356.
Important aspects in biomarker research are well reviewed, highlighting potential shortcomings in cancer studies and focusing on the instrumental and biological underlying perspectives.

16. Keun HC, Beckonert O, Griffin JL, Richter C, Moskau D, Lindon JC, Nicholson JK: **Cryogenic probe C-13 NMR spectroscopy of urine for metabonomic studies**. *Anal Chem* 2002, **74**:4588-4593.

17. Khandelwal P, Beyer CE, Lin Q, McGonigle P, Schechter LE, Bach AC: **Nanoprobe NMR spectroscopy and *in vivo* microdialysis: new analytical methods to study brain neurochemistry**. *J Neurosci Methods* 2004, **133**:181-189.

18. 't Hart BA, Vogels JT, Spijksma G, Brok HP, Polman C, van der Greef J: **^{1}H-NMR spectroscopy combined with pattern recognition analysis reveals characteristic chemical patterns in urines of MS patients and non-human primates with MS-like disease**. *J Neurol Sci* 2003, **212**:21-30.

19. Bollard ME, Xu JS, Purcell W, Griffin JL, Quirk C, Holmes E, Nicholson JK: **Metabolic profiling of the effects of D-galactosamine in liver spheroids using H-1 NMR and MAS-NMR spectroscopy**. *Chem Res Toxicol* 2002, **15**:1351-1359.

20. Mortishire-Smith RJ, Skiles GL, Lawrence JW, Spence S, Nicholls AW, Johnson BA, Nicholson JK: **Use of metabonomics to identify impaired fatty acid metabolism as the mechanism of a drug-induced toxicity**. *Chem Res Toxicol* 2004, **17**:165-173.

21. Goodenowe D: **Metabolomic analysis with Fourier transform •• ion cyclotron resonance mass spectrometry**. In *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Edited by Harrigan GG, Goodacre R. Boston, Dordrecht, London: Kluwer Academic Publishers; 2003:125-140.
The power and potential of FTMS in profiling is well described.

22. Van der Greef J, van der Heijden R, Verheij ER: **The role of mass •• spectrometry in systems biology: data processing and identification strategies in metabolomics**. In *Advances in Mass Spectrometry, Volume 16.* Edited by Ashcroft AE, Brenton G, Monaghan JJ. Amsterdam: Elsevier Science; 2004:145-165.
Key factors in metabolomics data preprocessing issues are highlighted.

23. Castrillo JI, Hayes A, Mohammed S, Gaskell SJ, Oliver SG: **An optimized protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry**. *Phytochemistry* 2003, **62**:929-937.

24. Tolstikov VV, Lommen A, Nakanishi K, Tanaka N, Fiehn O: **Monolithic silica-based capillary reversed-phase liquid chromatography/electrospray mass spectrometry for plant metabolomics**. *Anal Chem* 2003, **75**:6737-6740.

25. von Roepenack-Lahaye E, Degenkolb T, Zerjeski M, Franz M, Roth U, Wessjohann L, Schmidt J, Scheel D, Clemens S: **Profiling of *Arabidopsis* secondary metabolites by capillary liquid chromatography coupled to electrospray ionization quadrupole time-of-flight mass spectrometry**. *Plant Physiol* 2004, **134**:548-559.

26. Tolstikov VV, Fiehn O: **Analysis of highly polar compounds of plant origin: Combination of hydrophilic interaction chromatography and electrospray ion trap mass spectrometry**. *Anal Biochem* 2002, **301**:298-307.

27. Soga T, Ohashi Y, Ueno Y, Naraoka H, Tomita M, Nishioka T: **Quantitative metabolome analysis using capillary electrophoresis mass spectrometry**. *J Proteome Res* 2003, **2**:488-494.

28. Soga T, Ueno Y, Naraoka H, Matsuda K, Tomita M, Nishioka T: **Pressure-assisted capillary electrophoresis electrospray ionization mass spectrometry for analysis of multivalent anions**. *Anal Chem* 2002, **74**:6224-6229.

29. Soo EC, Aubry AJ, Logan SM, Guerry P, Kelly JF, Young NM, Thibault P: **Selective detection and identification of sugar nucleotides by CE-electrospray-MS and its application to bacterial metabolomics**. *Anal Chem* 2004, **76**:619-626.

30. Britz-McKibbin P, Terabe S: **On-line preconcentration •• strategies for trace analysis of metabolites by capillary electrophoresis**. *J Chromatogr A* 2003, **1000**:917-934.
Differences in metabolite concentrations in biological samples range over at least nine orders of magnitude. For detection of the lower abundance metabolites, different pre-concentration is required. On-line strategies for electromigration techniques are reviewed.

31. Jia L, Liu BF, Terabe S, Nishioka T: **Two-dimensional separation method for analysis of Bacillus subtilis metabolites via hyphenation of micro-liquid chromatography and capillary electrophoresis**. *Anal Chem* 2004, **76**:1419-1428.

32. Shi H, Vigneau-Callahan KE, Matson WR, Kristal BS: **Attention to relative response across sequential electrodes improves quantitation of coulometric array**. *Anal Biochem* 2002, **302**:239-245.

33. Kaddurah-Daouk R, Beecher C, Kristal BS, Bogdanov M, Asa D: **Bioanalytical advances for metabolomics and metabolic profiling**. *Pharmagenomics* 2004, **Jan**:46-52.

34. Ellis DI, Harrigan GG, Goodacre R: **Metabolic fingerprinting with Fourier transform infrared spectroscopy**. In *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Edited by Harrigan GG, Goodacre R. Boston, Dordrecht, London: Kluwer Academic Publishers; 2003:111-124.

35. Harrigan GG, LaPlante RH, Cosma GN, Cockerell G, Goodacre R, Maddox JF, Luyendyk JP, Ganey PE, Roth RA: **Application of high-throughput Fourier-transform infrared spectroscopy in toxicology studies: contribution to a study on the development of an animal model for idiosyncratic toxicity**. *Toxicol Lett* 2004, **146**:197-205.

36. Clarke S, Goodacre R: **Raman spectroscopy for whole organism and tissue profiling**. In *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis.* Edited by Harrigan GG, Goodacre R. Boston, Dordrecht, London: Kluwer Academic Publishers; 2003:95-110.

37. Boros LG, Steinkamp MP, Fleming JC, Lee WNP, Cascante M, Neufeld EJ: **Defective RNA ribose synthesis in fibroblasts from patients with thiamine-responsive megaloblastic anemia (TRMA)**. *Blood* 2003, **102**:3556-3561.

38. Weston AD, Hood L: **Systems biology, proteomics, and the future of health care: Toward predictive, preventative, and personalized medicine**. *J Proteome Res* 2004, **3**:179-196.

39. Anderson NL, Anderson NG: **The human plasma proteome: •• history, character, and diagnostic prospects**. *Mol Cell Proteomics* 2002, **1**:845-867.
A landmark review integrating key information with insightful and important perspectives.

40. Anderson NL, Polanski M, Pieper R, Gatlin T, Tirumalai RS, Conrads TP, Veenstra TD, Adkins JN, Pounds JG, Fagan R, Lobley A: **The human plasma proteome: a non-redundant list developed by combination of four separate sources**. *Mol Cell Proteomics* 2004, **3**:311-326.

41. Julka S, Regnier F: **Quantification in proteomics through •• stable isotope coding: a review**. *J Proteome Res* 2004, **3**:350-363.
A thoughtful and critical review of progress and challenges in the important field of isotope labeling approaches in proteomics.

42. Ross P, Huang Y, Marchese J, Khainovski N, Williamson B, Hattan •• S, Juhasz P, Daniels S, Pillai S, Purkayastha S *et al.*: **Relative and absolute quantitation in yeast proteomics using multiplexed isobaric peptide tags**. *Proceedings 52nd ASMS, Nashville, 2004 May 23-27, 2004, Slot 392*. [http://www.inmerge.com/ASMS/DisplayAbstractList.aspx?Session=TPT].
iTRAQ design and methodology demonstrate a number of ingenious features that may provide breakthrough quantification capabilities.

43. Chaurand P, Schwartz SA, Caprioli RM: **Assessing protein •• patterns in disease using imaging mass spectrometry**. *J Proteome Res* 2004, **3**:245-252.
A promising approach that correlates protein expression with histological features observed by optical microscopy.

44. Qiu J, Madoz-Gurpide J, Misek DE, Kuick R, Brenner DE,
•• Michailidis G, Haab BB, Omenn GS, Hanash S: **Development of natural protein microarrays for diagnosing cancer based on an antibody response to tumor antigens**. *J Proteome Res* 2004, **3**:261-267.

A creative and potentially powerful approach to defining the repertoires of tumor antigens that elicit antibody responses in patients with specific cancers.

45. Lamers RJ, DeGroot J, Spies-Faber EJ, Jellema RH, Kraus VB, Verzijl N, TeKoppele JM, Spijksma GK, Vogels JT, van der Greef J, van Nesselrooij JH: **Identification of disease- and nutrient-related metabolic fingerprints in osteoarthritic Guinea pigs**. *J Nutr* 2003, **133**:1776-1780.

46. Andreev VP, Rejtar T, Chen HS, Moskovets EV, Ivanov AR, Karger BL: **A universal denoising and peak picking algorithm for LC-MS based on matched filtration in the chromatographic time domain**. *Anal Chem* 2003, **75**:6314-6326.

47. Brindle JT, Antti H, Holmes E, Tranter G, Nicholson JK, Bethell HWL, Clarke S, Schofield PM, McKilligin E, Mosedale DE *et al.*: **Rapid and noninvasive diagnosis of the presence and severity of coronary heart disease using ^{1}H-NMR-based metabonomics (vol 8, pg 1439, 2002)**. *Nat Med* 2003, **9**:477-477.

48. Idborg-Bjorkman H, Edlund PO, Kvalheim OM, Schuppe-Koistinen I, Jacobsson SP: **Screening of biomarkers in rat urine using LC/electrospray ionization-MS and two-way data analysis**. *Anal Chem* 2003, **75**:4784-4792.

49. Jonsson P, Gullberg J, Nordstrom A, Kusano M, Kowalczyk M, Sjostrom M, Moritz T: **A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS**. *Anal Chem* 2004, **76**:1738-1745.

50. Idborg H, Edlund PO, Jacobsson SP: **Multivariate approaches for efficient detection of potential metabolites from liquid chromatography/mass spectrometry data**. *Rapid Commun Mass Spectrom* 2004, **18**:944-954.

51. Duran AL, Yang J, Wang LJ, Sumner LW: **Metabolomics spectral formatting, alignment and conversion tools (MSFACTs)**. *Bioinformatics* 2003, **19**:2283-2293.

52. Keun HC, Ebbels TMD, Antti H, Bollard ME, Beckonert O, Holmes E, Lindon JC, Nicholson JK: **Improved analysis of multivariate data by variable stability scaling: application to NMR-based metabolic profiling**. *Anal Chim Acta* 2003, **490**:265-276.

53. Bro R, Smilde AK: **Centering and scaling in component analysis**. *J Chemometrics* 2003, **17**:16-33.

54. Scholtz M, Gatzek S, Sterling A, Fiehn O, Selbig J: **Metabolite**
•• **fingerprinting: detecting biological features by independent component analysis**. *Bioinformatics* 2004, in press.

Non-supervised methods are key to better understand biology. Here, a new strategy of independent component analysis is described to generate more meaningful components from metabolomics data than via traditional principle components analysis.

55. Clish CB, Davidov E, Oresic M, Plasterer TN, Lavine G, Londo T,
•• Meys M, Snell P, Stochaj W, Adourian A *et al.*: **Integrative biological analysis of the APOE*3-leiden transgenic mouse**. *OMICS* 2004, **1**:3-13.

The first example of a mammalian systems biology approach using transcriptomics, proteomics and metabolomics followed by an integration to a biological relevant correlation network, demonstrating the interconnectiveness and interdependency of a biological systems in disease development.

56. Jansen RC: **Studying complex biological systems using multifactorial perturbation**. *Nat Rev Genet* 2003, **4**:145-151.

57. Ideker T, Lauffenburger D: **Building with a scaffold: emerging**
•• **strategies for high-to low-level cellular modeling**. *Trends Biotechnol* 2003, **21**:255-262.

An important vision paper describing the range of approaches from statistical mining to cell modelling and the impact for computational strategies.

58. Fiehn O, Weckwerth W: **Deciphering metabolic networks**. *Eur J Biochem* 2003, **270**:579-588.

59. Kell DB: **Metabolomics and machine learning: explanatory analysis of complex metabolome data using genetic programming to produce simple, robust rules**. *Mol Biol Rep* 2002, **29**:237-241.

60. Hwang DH, Stephanopoulos G, Chan C: **Inverse modeling using multi-block PLS to determine the environmental conditions that provide optimal cellular function**. *Bioinformatics* 2004, **20**:487-499.

61. Antti H, Bollard ME, Ebbels T, Keun H, Lindon JC, Nicholson JK, Holmes E: **Batch statistical processing of H-1 NMR-derived urinary spectral data**. *J Chemometr* 2002, **16**:461-468.

62. Jansen JJ, Hoefsloot HC, Boelens HF, Van Der Greef J, Smilde AK: **Analysis of longitudinal metabolomics data**. *Bioinformatics* 2004, in press.

63. Heath JR, Phelps EP, Hood L: **Nano systems biology**. *Mol Imaging Biol* 2003, **5**:312-325.

64. Grant SGN: **Systems biology in neuroscience: bridging genes to cognition**. *Curr Opin Neurobiol* 2003, **13**:577-582.