

Integrative Biological Analysis For Neuropsychopharmacology

Mark R Emmett¹, Roger A Kroes², Joseph R Moskal², Charles A Conrad³, Waldemar Priebe⁴,
Fernanda Laezza⁵, Anke Meyer-Baese⁶ and Carol L Nilsson^{*5}

¹Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, Galveston, TX, USA; ²Department of Biomedical Engineering, Northwestern University, Falk Center for Molecular Therapeutics, Evanston, IL, USA; ³Department of Neuro-Oncology, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA; ⁴Department of Medicinal Chemistry, The University of Texas M.D. Anderson Cancer Center, Houston, TX, USA; ⁵Department of Pharmacology and Toxicology, University of Texas Medical Branch, Galveston, TX, USA; ⁶Department of Scientific Computing, Florida State University, Tallahassee, FL, USA

Although advances in psychotherapy have been made in recent years, drug discovery for brain diseases such as schizophrenia and mood disorders has stagnated. The need for new biomarkers and validated therapeutic targets in the field of neuropsychopharmacology is widely unmet. The brain is the most complex part of human anatomy from the standpoint of number and types of cells, their interconnections, and circuitry. To better meet patient needs, improved methods to approach brain studies by understanding functional networks that interact with the genome are being developed. The integrated biological approaches—proteomics, transcriptomics, metabolomics, and glycomics—have a strong record in several areas of biomedicine, including neurochemistry and neuro-oncology. Published applications of an integrated approach to projects of neurological, psychiatric, and pharmacological natures are still few but show promise to provide deep biological knowledge derived from cells, animal models, and clinical materials. Future studies that yield insights based on integrated analyses promise to deliver new therapeutic targets and biomarkers for personalized medicine.

Neuropsychopharmacology Reviews (2014) **39**, 5–23; doi:10.1038/npp.2013.156; published online 28 August 2013

Keywords: neurological disease; neuro-oncology; psychiatric disorder; systems biology

INTRODUCTION

Since the first draft of the human genome was published in 2001 (Venter *et al*, 2001), genome-wide studies of human variability have gone from dream to reality, covering the entire spectrum of human disease, including neurological and psychiatric diseases. Following the initial draft, scientific debates arose regarding the number of genes encoded by the human genome and the percentage of noncoding or ‘junk’ DNA. At first, the coding elements were deemed to be a low percent of the entire genome. The first general translation of human genome in 2012, delivered by the ‘Encyclopedia of DNA Elements’ or ENCODE project (Encode Project Consortium, 2012), showed that 80% or more of the human genome is biologically active and that the noncoding DNA regions in fact have the potential to modulate human health and disease. These new data will

provide a basis to increase the impact of genomics in human medicine and are highly synergistic with data from transcriptomic, proteomic, glycomic, and metabolomic initiatives. For instance, only one-third of the ENCODE-predicted protein-coding genes have been identified at the protein level. To meet the challenge of identifying all human proteins and assigning their functions, ENCODE has joined forces with the Chromosome-centric Human Proteome Project (C-HPP) (Paik and Hancock, 2012). Those combined efforts are expected to define the mechanisms by which polygenes, transcription factor networks, and single-nucleotide polymorphisms (SNPs) control the expression of groups of protein isoforms. The C-HPP will identify the expression of yet undetected proteins, including those encoded by newly identified coding regions. A third partner in the collaboration is the Genome-Tissue Expression Project (<http://www.genome.gov/gtex/>). The GTEx project employs optimized protocols for obtaining and storing a large range of organs and tissues, and for analyzing genetic variation and expression. These tissues and organs are collected and stored through the National Cancer Institute’s cancer Human Biobank initiative on behalf of GTEx. Microarray results for frontal and temporal cortex,

*Correspondence: Professor CL Nilsson, Department of Pharmacology and Toxicology, University of Texas Medical Branch, 301 University Boulevard, Galveston, TX 77555-0617, USA, Tel: +1 409 747 1840, FAX: +1 409 772 9648, E-mail: carol.nilsson@utmb.edu

Received 3 March 2013; revised 18 April 2013; accepted 19 April 2013; accepted article preview online 26 June 2013

cerebellum, and pons are already available. The main goal of the collaboration is to integrate genomic and proteomic data into functional modules that will describe systems-wide interactions in biological processes, in health, and in disease, including neuropsychiatric disorders (Califano *et al*, 2012).

Understanding the response of neurological cell populations within tissues (normal *vs* diseased, treated *vs* untreated, etc.) and evidence that these cell populations are ‘drivers’ of disease underscores the importance of understanding the detailed biology at the molecular level of cells and circuits. Utilizing an integrated biological approach to probe the response of cells and tissues will guide the rational design of therapeutic treatment. Because of the complexity of the approach, this research is most often accomplished through multidisciplinary and inter-institutional collaborations, as no single laboratory has expertise in all the technologies required. The C-HPP is an extreme example, because it brings together data sets acquired in hundreds of international laboratories.

Systems biology research is built on knowledge derived from global data sets measured in patterns of response of the transcriptome, proteome, glycome, lipidome, and metabolome. To obtain knowledge of the signaling pathways involved in the maintenance of the normal *vs* diseased state or treated *vs* untreated states, quantitative phosphoproteomic, glycomic, transcriptomic, lipidomic, and metabolomic data sets are collected on each sample. The data are compared with known, published results and pathways and then processed with innovative mathematical–computational algorithms tailored to define correlations between the data sets and graphical modeling to reconstruct pathways that are not yet defined in the scientific literature. The focus is to identify and characterize pathway reactions that will bridge the final gap and, for the first time, enable a mechanism to understand cellular responses, identify new targets, and design appropriate therapeutic interventions specifically targeting the disease.

We have applied this approach (He *et al*, 2010a,b; Kroes *et al*, 2010; Nilsson *et al*, 2010) to the study of GBM and glioma-derived cancer stem-like cells (GSCs). The research is based on several innovative biological systems and technologies, and has resulted in novel insights into GBM and GSC biology. Our approach permits a hypothesis-driven, carefully planned experimental design to generate the data that enable identification of novel therapeutic targets and biomarkers. We measured global changes in intact polar lipids (gangliosides, phospholipids, sulfatides, etc.) through an analytical methodology that employed innovative chromatographic separations, efficient ionization techniques, and high-resolution mass spectrometry (MS) analysis. Observed changes in lipids (lipidomics) and other metabolites in normal *vs* treated GBMs and/or GSCs were correlated to changes in glycogene transcripts. Those data were further correlated with changes in the phosphoproteome of treated GSCs. Computational–mathematical analysis (Goerke *et al*, 2010; Meyer-Baese *et al*, 2012a,b) further defined correlations in the large data sets (Figure 1). Although the subject of the study was within neuro-oncology, the same techniques are applicable to other diseases. However, experimental planning is more complicated in studies of neuropsychiatric illness, because those studies may include neurodevelopmental hypotheses or changes over a very long timespan. The appropriate samples, whether from human or from animal models, must be obtained in a standardized manner (Mattsson *et al*, 2012) and must be properly treated and stored, especially for biomolecular studies. Heat stabilization may be required at the time of sampling to preserve neuropeptide, metabolite, and protein, especially phosphoprotein, integrity (Ahmed and Gardiner, 2011)

This review focuses upon how the new knowledge base provided by advances in genomics can be leveraged in other genome-wide studies in neuroscience, especially when those data sets are combined. The four major tissues (human-derived or from animal models) used in molecular studies

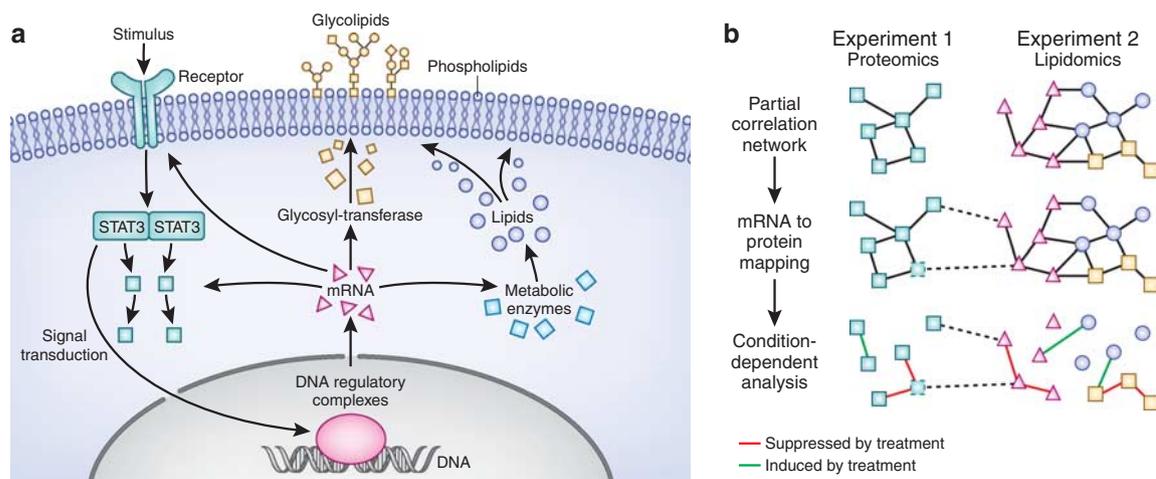


Figure 1. Computational analysis of comprehensive proteomic, transcriptomic, and lipidomic data sets can yield new insights derived from correlations hidden in the data.

of human CNS disease are cerebrospinal fluid, cell cultures, *post-mortem* brain, and spinal cord. Blood and urine samples are always part of the general medical analysis, but saliva and nasal secretions may prove to be valuable sample sources, especially for metabolite analysis. Biological tissues are a rich source of (poly)nucleotides, amino acid biopolymers (neuropeptides and proteins), lipids, and metabolites. Traditionally, bioanalytical methods in neuroscience have relied heavily on antibody-based detection of single or a few proteins by western blot or immunohistochemistry analysis. Although highly sensitive, antibody-based assays only recognize a linear or three-dimensional epitope of three to seven amino acids (Schechter, 1970) and often cross-react with molecules other than the one targeted. If the epitope is modified (such as by phosphorylation), the protein may go undetected. Furthermore, a single antibody likely will not deliver structural information outside the epitope and thus cannot resolve molecular isoforms that may have distinct biological functions. For these reasons, there has been a surge in interest in analytical techniques that can provide both high-confidence identification and structural assignments.

This review is organized into sections that describe the enabling technologies that are employed in integrated biological studies, including genome-wide determination of transcript, protein, lipid and metabolite expression, and computational strategies to find novel correlations between large biological data sets. Examples of synergistic results obtained in neuroscientific and neuro-oncological investigations are provided. We present methods relevant to neuroimaging connectomics and the cell biology of neuropsychiatric disorders. Although we do not include a separate discussion of the biomarker field, it is our experience that putative biomarker identification is a natural outcome of integrated biological studies. Future perspectives on the potential impact of integrated biological approaches in the development of neurotherapeutics are described in each section.

ESSENTIAL CONSIDERATIONS IN EXPERIMENTAL DESIGN

The transcription and translation of genetic information into proteins, lipids, and metabolites comprises highly nonlinear processes. Intra- and extracellular environments are dynamic systems that constantly adjust to signals in their surroundings. Thus, understanding the micro- and macroheterogeneity of functional molecules is essential to study disease processes. Recent advances in analytical and computational technologies show promise to characterize the full biochemical complement of neurological and other biological systems in health and disease. It is of importance to note that studies of biomolecules on a global scale often result in very large data sets (terabytes) and require a high degree of skilled experimental planning, execution, and data analysis in order to obtain biologically relevant results.

The pitfalls in large-scale studies can be eliminated by thoughtful study design (Oberge and Vitek, 2009). Systematic data biases result from different treatments between sample states (different sampling methods or storage conditions, for instance) and those that cannot be corrected by statistics. Data from adequate numbers of biological and technical replicates must be acquired. Every large-scale experiment does have sources of technical variation, which can be ameliorated by statistical analysis. Although large-scale investigations of lipids and metabolites are relatively rare, still, much discussion of statistical methods to correct for experimental biases in transcriptomic (Dupuy and Simon, 2007) and proteomic (Cappadona *et al*, 2012; Thompson *et al*, 2012) assays has been published. It is highly advisable to have a statistician participate in the design of large-scale studies, beginning at the planning stage.

The most important component of large-scale experimental design is a robust hypothesis, followed by the identification of sample populations, appropriate comparisons, and numbers of biological samples. There are special challenges associated with molecular research in the neuropsychiatric field. The privileged nature of human brain sampling, which most often occurs *post mortem*, the need to protect the samples from degradation to the greatest extent possible, and the challenges in making clinical diagnoses are especially prominent in this field. A strict control of factors that affect intersample variability is necessary. Once the biological samples are accumulated, experimental variations should be minimized by rigorous protocols to ensure that all samples are treated identically. Most samples are extracted and analyzed in a queue, injected sequentially into a mass spectrometer. When the samples are grouped and introduced sequentially, systematic quantitative differences may be observed that arise due to instrument sensitivity rather than real differences between samples. Block randomization of sample analysis is one way to overcome this issue.

Quantitative workflows in proteomics, lipidomics, and metabolomics record patterns of analyte abundances in samples derived from distinct conditions or states. Typically, very large data sets are generated, and are first extracted and processed by the use of commercial or in-house-developed software tools. A list of quantified analytes may then be further examined with knowledge-based tools, such as Ingenuity Pathways Analysis, to find biological circuits that may be relatively up- or downregulated between sample states.

TRANSCRIPTOMICS

One important principle underpinning the systems biological approach to understand complex neurological disorders is not merely the integration of each of the physical components and interactions of a system, but also the mapping of how the information propagates through this system in response to perturbations. If we hope to understand the dynamic biology of complex neurological

disorders, instead of looking at one transcript or protein at a time, which is historically how this has been approached, we need to understand the integration of thousands of proteins in a dynamically changing environment (Geschwind and Konopka, 2009; Villoslada *et al*, 2009; Kotaleski and Blackwell, 2010). How we understand core transcriptomic changes can occur at several different levels, depending upon the overarching nature and purpose of the question at hand: (a) the hypothesis-generating phase. When transcriptomics is employed as a single technique, it is an unbiased approach that requires no understanding of the system and occurs primarily at the level of gene expression patterns, and (b) the applications or target generation phase. The latter requires a systems approach and applies it in a way that can be used to identify better therapeutic targets or combinations of targets in the proper context for the treatment of disease. As many genes contribute to normal functions, research efforts are moving from the search for a ‘disease-specific gene’ towards the understanding of the biochemical and molecular functioning of gene families whose disrupted interaction in complicated networks can lead to the disease state.

Transcriptomics can address all (global approaches) or a segment (focused approaches) of the transcriptome, from normal or diseased single cells or tissues. Global approaches represent an important complement to the traditional reductionist single-gene studies. Whereas traditional single gene studies are hypothesis-driven, both global and focused approaches can be hypothesis-generating, discovery-based, or hypothesis-driven, and therefore have the potential to yield novel insight into brain function and dysfunction. As detailed below, depending on the techniques employed, it can be used as a snapshot to annotate a cell’s expression profile or fingerprint, including coding as well as noncoding transcripts; query the structure of the genes that gave rise to them, including intricacies of exon/intron boundaries, transcription start sites, splicing patterns, and even gene fusion events; and aid in mapping interactive networks. Networks are systems of interconnected entities. When networks are studied comprehensively, novel properties emerge that can’t be derived from the individual analysis of any of their components. And perhaps most familiarly, transcriptomics are used to determine how transcript expression patterns are altered under differing conditions (such as disease or drug treatment), and to identify and validate potential biomarkers or therapeutic targets.

MICROARRAY-BASED TRANSCRIPTOMICS

Focused Microarray Approaches

Focused microarrays are fabricated using a restricted number of carefully selected transcripts, and may include those targeting a particular biological pathway, a specific tissue, or even a scaled-down version of the transcriptome. Irrespective of their overall design (detailed below), this approach optimizes sensitivity and accuracy of the data,

while controlling the cost of experiments, thus facilitating the repeated measurement and replicate experiments required in analyses of complex tissues. The reduced number of probes comprising these arrays can be stringently designed, thus controlling their biophysical characteristics critical to hybridization efficiency (melting temperature, secondary structure, and stability). As such, these arrays are able to accurately identify differentially expressed transcripts with high sensitivity to low levels of differential expression. This approach reliably identifies transcripts having as low as 10% difference between RNA samples, reinforcing the applicability of this approach to analyze highly heterogeneous tissue sources. As with all transcriptomic platforms, the identified transcripts must be validated by an independent methodology (eg, higher throughput quantitative real-time PCR assays) (Canales *et al*, 2006).

We have utilized this focused microarray approach (FMA) platform to identify aberrant cell-surface glycosylation patterns in malignant human gliomas (Kroes *et al*, 2007a,b; Moskal *et al*, 2009; Kroes *et al*, 2010). We developed a custom-fabricated, focused oligonucleotide microarray representing all of the known human glyco-genes: transcripts encoding proteins responsible for glycan synthesis and glycan recognition, including glycosyltransferases, glycosylhydrolases, and other glycan processing enzymes, proteoglycans, and glycan-binding proteins. This platform was designed to provide comprehensive coverage of the glyco-associated gene family, the flexibility to add new sequences to the arrays as they become available, and to assure the most up-to-date coverage of these gene families. This platform has been rigorously evaluated in terms of dynamic range, discrimination power, accuracy, reproducibility, and specificity, which provides the ability to reliably measure even low levels of statistically significant differential gene expression. In particular, when combined with robust ontological analyses, qRT-PCR corroboration, and *in vivo* functional studies, these high-quality, comprehensively annotated, application-specific low-density microarray platforms provide an efficient strategy that can pinpoint gene expression changes and provide a robust database for pursuing the creation of novel therapeutics.

Global Microarray Approaches

The evolution of microarray platforms over the past two decades has resulted in multiple commercially available platforms for measuring genome-wide gene expression levels, including short oligonucleotide (25–30 base), long oligonucleotide (40–80 base), and cDNA (highly variable in length) arrays, in both single- and dual-channel formats. The continual increases in both the efficiency and economy of the oligonucleotide arrays seem to have made them the platform of choice for many studies. Performance characteristics between platforms, as well as inter- and intraplatform reproducibility of gene expression measurements, have been extensively evaluated over the years

(Petersen *et al*, 2005; Microarray Consortium, 2006; Patterson *et al*, 2006). Overall, such detailed platform comparisons reveal mixed results. Several studies demonstrate concordance between platforms (Kane *et al*, 2000; Hughes *et al*, 2001; Taniguchi *et al*, 2001; Guckenberger *et al*, 2002; Yuen *et al*, 2002; Barczak *et al*, 2003; Dabrowski *et al*, 2003; Lee *et al*, 2003; Wang *et al*, 2003; Bloom *et al*, 2004; Järvinen *et al*, 2004; Lee *et al*, 2004; Parmigiani *et al*, 2004; Thompson *et al*, 2004; Ulrich *et al*, 2004), whereas others suggest discordance (Kothapalli *et al*, 2002; Kuo *et al*, 2002; Li *et al*, 2002; Carter *et al*, 2003; Lenburg *et al*, 2003; Rogojina *et al*, 2003; Tan *et al*, 2003; Mah *et al*, 2004). Although the explanation for these discrepancies is not altogether consistent, platform-specific differences in the statistical algorithms used to discriminate signal from noise and calculate transcript abundance were most relevant. In fact, the majority of discrepancy appears to be because of the (in)ability to detect lower abundance transcripts. This, however, is an arena where FMA platforms appear to have significant benefit.

Although FMA, by definition, provides less than the total transcriptome coverage than a global array, this limitation can be partially overcome by the careful selection of the transcripts that are represented. The breadth of the transcriptome coverage can be achieved, eg, using approaches that select representative genes within the defined ontology of the organism under study. Thus, this approach is clearly limited to those with adequate genome annotation.

As part of a larger program aimed at utilizing such an approach in appropriate animal models to give some insight into the molecular mechanisms underlying the emergence of neurological disorders, we have created and utilized such annotated arrays (Kroes *et al*, 2007a,b, 2010). The genes comprising this oligonucleotide microarray were compiled from current databases and were strategically chosen to provide representation of greater than 90% of the major gene ontological categories. Downstream of the stringent data analysis, the incorporation of ontological algorithms, such as GOMiner (Zeeberg *et al*, 2003), Gene Set Enrichment Analysis (Subramanian *et al*, 2005, 2007), DAVID (Huang *et al*, 2008), or Expression Analysis Systematic Explorer (Hosack *et al*, 2003) provided statistical rigor to analyses of coregulation of multiple genes (gene sets) that are functionally related or related by involvement in given biological pathways. For a full description of the dynamic range, discrimination power, accuracy, reproducibility, and specificity of this platform, please see these complete references (Kroes *et al*, 2006, 2007a,b). We have identified expression patterns and functional gene families involved in models of epileptogenesis (Corcoran *et al*, 2011), depression (Kroes *et al*, 2006, 2007a,b), and autism (Moskal *et al*, 2011); all are underexamined areas for development of novel therapeutics. Such preclinical studies uncovering the role of NMDA receptor modulation in positive emotional learning and resilience to depression has reinforced the clinical development of GLYX-13, a functional partial agonist of the NMDA receptor currently in phase II clinical

development (clinicaltrials.gov identifier NCT01684163) as an adjunctive therapy for major depressive disorder (Burgdorf *et al*, 2011, 2013).

Direct Sequencing-Based Approaches

Although microarray technology has clearly dominated transcriptomics research, this situation appears to be changing rapidly with the recent technological advances in high-throughput RNA sequencing (RNAseq) (Janitz, 2008; Wang *et al*, 2009; Hawkins *et al*, 2010), albeit a technology still under active refinement; especially bioinformatic challenges related to storing, retrieving, and processing large amounts of data (Pepke *et al*, 2009; Martin and Wang, 2011). In general, a population of RNA is converted to a library of cDNA fragments with defined adaptors attached to one or both ends. Each molecule, with or without amplification, is then sequenced in a high-throughput manner to obtain short sequences; typically 30–400 bp, depending on the DNA-sequencing technology used. Following sequencing, the resulting reads are either aligned to a reference genome or reference transcripts, or assembled *de novo* without the genomic sequence to produce a genome-scale transcription map that consists of both the transcriptional structure and/or level of expression for each gene. It records the numerical frequency of a sequence in the library population, and thus provides an absolute measurement of transcript abundance rather than a relative measure.

RNAseq has some distinct advantages over microarrays (Courtney *et al*, 2010; Malone and Oliver, 2011). There is no prerequisite for genome sequence information in order to detect and ultimately evaluate transcripts; thus, RNAseq is most useful in organisms and situations where limited genome information exists. Unbiased sequence methods can detect novel transcripts, alternatively spliced isoforms, SNPs, and fusion transcript, again in the absence of genome annotation. It provides deep coverage and base-scale resolution (Cloonan and Grimmond, 2008). However, this technology also comes with some distinct trade-offs as well. The amount of time to retrieve results for a given sample is far greater. It is ~10–100 times more expensive when compared at the same resolution of accuracy (coverage *vs* cost). It still may suffer from unknown biases, such as those imposed by the required ligation steps. The ability to detect and quantify rare transcripts is obscured. By design, high-abundance transcripts (such as from housekeeping genes) are responsible for the majority of the sequencing data (5% of the genes give rise to 75% of the reads sequenced). As a consequence, it is difficult to measure the abundances of the remaining genes reliably, and the majority of transcript measurements that are made at this level are very noisy. In fact, some recent reports suggest that in a mammalian genome, ~700 million reads would be required to obtain an accurate quantification of >95% of the expressed transcripts (Blencowe *et al*, 2009) but, as yet, there has not been a systematic analysis on how sequencing coverage affects differential expression calls (Oshlack *et al*, 2010). Although

deep sequencing effectively enhances our view on the diversity of the transcriptome, the identification of true differential expression at a low-count range might not be so easy to achieve. More reads imply the detection of more genes, but also result in noisier data, which makes the assessment of differential expression increasingly difficult.

In sum, the diversity of transcriptomic platforms clearly provides investigators with many options. When sensitivity is not limiting, DNA microarrays seem to offer significant advantages. On the other hand, when sensitivity is everything, look toward the short-read sequencing technologies. These technologies are totally complementary: each strategy has its appropriate time and place for use. In the end, the use of multiple platforms can only lead to the generation of higher quality data.

GLOBAL ANALYSIS OF PROTEINS AND OTHER BIOPOLYMERS

Unlike transcriptomic experiments, which are readily performed in a chip format, sequence/structural analysis of proteins, peptides, lipids, and glycoconjugates relies heavily on MS. The application of large-scale MS studies has greatly increased in the past two decades because of the high sensitivity and structural specificity that MS and tandem MS (MS/MS) can provide. A mass spectrometer always contains two parts, a source that generates gas-phase ions and a mass analyzer. The increase in biological applications of MS is attributed to the discovery of new, 'soft' methods to ionize large biopolymers, the development of low-flow (nl/min) introduction of solubilized analytes into mass spectrometers, and improved figures of merit for speed of data acquisition, sensitivity, and resolution of the instruments themselves. For small molecules, such as metabolites, gas chromatography (GC) with electron ionization is one reliable ionization technique. For large analytes, such as peptides, proteins, and glycolipids, matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp, 1988) and electrospray ionization (ESI) (Fenn *et al*, 1989) provide intact molecular ions for measurements of intact mass by MS or fragmentation by MS/MS. MALDI-MS is a technique that is particularly tolerant of molar excesses of salts and buffers typically found in biological samples. MALDI-MS has been applied to analyze not only extracted analytes but has been used successfully in molecular imaging studies of *in situ* peptides, proteins, and lipids (Seeley *et al*, 2011) (see Andr n *et al*, this issue).

When an ion source is coupled to an ultra-high-resolution mass analyzer, such as a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR MS) (Marshall *et al*, 1998; Marshall and Hendrickson, 2008), high mass resolution (400 000 mass spectral peak width at half-maximal peak height) and high mass accuracy (ppb mass deviations) can be routinely attained. For molecules <400 Da, exact molecular formulae and isotopic fine structures can be determined. Thus, FT-ICR MS is highly suitable for

metabolomic studies. This type of mass spectrometer is often a shared resource in academic institutions because of its cost. For proteomic studies, lower-end instrumentation is often sufficient.

Proteomics

The human proteome is defined as the protein complement of the human genome. Proteomic studies benefit greatly from accurate and complete genome characterization. In comparative proteomic studies, global characterization without the use of antibodies of proteins expressed by cells or tissues in healthy or otherwise normal states, compared with proteins expressed in diseased states, can provide clues to underlying mechanisms of the underlying disease process. It is possible to identify putative biomarkers as a result of those studies, although extensive validation studies are required as a consequence. The proteomic samples are complex mixtures that can be simplified at the level of sampling by laser capture microdissection (P rez-Manso *et al*, 2006) to enrich specific cell types from tissues. Subcellular fraction or enrichment for proteins bearing a specific posttranslational modification (PTM) may be useful approaches, depending on the experimental hypothesis. An excellent overview with helpful facts on the application of proteomics to clinical materials was published recently by Tabb (2013).

Channelomics

The quest for new therapeutics that could ultimately lead to personalized treatments for psychiatric disorders remains a significant challenge. Large-scale genomic studies are likely to provide the initial data to identify potential disease mechanisms and disease-predictive biomarkers (Wong *et al*, 2011). The subsequent step is to capitalize on that new knowledge and evaluate systematically the role of relevant gene products in specific disease states through integrated biological technologies and informatics tools. This proposed pharmacogenetic pipeline has been proven successful in other disease contexts, such as cancer (Pitteri and Hanash, 2010) and cardiovascular diseases (Dubois *et al*, 2011) but it is at its nascent stage in psychiatric disorders. Over the past few years, a phenomenal number of genes have emerged as relevant in psychiatric disorders but very little is known about their exact role in psychiatric diseases. Some exemplary results will be discussed in the following paragraphs.

In neurons, voltage-gated sodium (Nav) and voltage-gated potassium (Kv) channels expressed at the axonal initial segment (AIS) drive the initiation and propagation of the action potential, the electrical signal underlying brain function and behaviors (Lai and Jan, 2006). Scaffolding and regulatory proteins at the AIS cluster with Nav and Kv channels ensure the integrity of electrical signaling (Ogawa and Rasband, 2008; Grubb and Burrone, 2010a,b) and interference with this protein network can lead to deficits in excitability and eventually neuronal death (Buffington and Rasband, 2012). The emerging association of mutations,

copy variants, or SNPs in genes coding for these ion-channel molecular complexes with psychiatric disorders is, in our opinion, one of the most exciting recent discoveries.

A large-scale genomic study in autistic patients identified two independent nonsense variants that interfere with the coding of Nav1.2 (Sanders *et al*, 2012), one of the Nav channel isoforms expressed at the AIS. Those results further support earlier studies that identified SNPs in the genes coding for Nav1.1 and Nav1.2 from another cohort of autistic patients (Weiss *et al*, 2003). Given the comorbidity of autism and epilepsy (Robinson, 2012), it is not surprising that the mutation in the Nav1.1-coding gene had been previously identified in a patient with juvenile myoclonic epilepsy (Weiss *et al*, 2003). Similar to Nav1.1, mutations in KCNQ2 (coding for Kv7.2, another channel expressed at the AIS) are recognized causes of human epilepsy and other excitability-driven inherited disorders (Singh *et al*, 1998; Mulley *et al*, 2003; Cooper, 2011; Mulley and Mefford, 2011). A new splice variant of the *KCNQ2* gene has been identified in bipolar disorder and its aberrant phenotype described in cellular studies (Borsotto *et al*, 2007). Whether all these genetic modifications of Nav and Kv channels converge on common cellular phenotypes with heterogeneous clinical manifestations remains to be determined, but it is a plausible hypothesis.

Other results from the collection of studies implicate ankyrin-G in bipolar disorder (Leussis *et al*, 2012). Ankyrin-G is the main scaffolding protein of the AIS required for clustering both Nav and Kv channels (Zhang and Bennett, 1998; Bennett and Lambert, 1999; Jenkins and Bennett, 2001; Brachet *et al*, 2010; Gasser *et al*, 2012) and necessary for establishing and maintaining neuronal polarity (Grubb and Burrone, 2010a,b; Galiano *et al*, 2012). Studies from different genetic cohorts have consistently ascribed the gene coding for ankyrin-G as a potential risk gene for bipolar disorder (Smith *et al*, 2009; Gella *et al*, 2011; Lett *et al*, 2011; Roussos *et al*, 2011; Tesli *et al*, 2011; Lee *et al*, 2012). Likewise, the gene coding for β -IV spectrin, a molecule associated with the ankyrin-G complex (Bergghs *et al*, 2000) and intimately linked to both the Nav and Kv channel complexes (Komada and Soriano, 2002; Kosaka *et al*, 2008; Hund *et al*, 2010), was found to contain microdeletions in an autistic patient (Griswold *et al*, 2011). Furthermore, a missense mutation in the *FGF14* gene, coding for a Nav channel regulatory protein expressed at the AIS (Lou *et al*, 2005; Laezza *et al*, 2007, 2009; Goetz *et al*, 2009; Shavkunov *et al*, 2012), is associated with cognitive impairment, neurodegeneration, and motor dysfunction (Van Swieten *et al*, 2003; Brusse *et al*, 2006), and SNPs in the same gene have been identified in schizophrenic patients (Need *et al*, 2009).

The biology of the mutations described above and their pathophysiological roles in psychiatric disorders is poorly understood. However, one could speculate that these ion channel macromolecular complexes are part of a nonredundant molecular network vital for neuronal communication, which, if aberrant, could lead to a wide range of

neurological or psychiatric syndromes. A systematic quantitative proteomic analysis of expression levels and regulatory PTMs (eg, phosphorylation, glycosylation, and ubiquitination) of these ion channel complexes in health and disease could provide guidance to formulate new hypotheses on disease mechanisms and identify reliable biomarkers for clinical applications (Filiou *et al*, 2012; Filiou and Turck, 2012). Those results will require further validations and proof of scalability, especially for membrane-associated proteins such as ion channels. A first step toward this goal would be to utilize animal models to evaluate the role of specific gene products into disease contexts and to optimize proteomic techniques towards clinical applications (Holsboer, 2008). The biological experimental data sets would then be properly integrated into predictive informatics networks and contribute to elucidate the complex mechanisms of psychiatric disorders. This process will require many further steps performed in translational collaborations between basic sciences and the clinical settings in order to create personalized treatments.

Protein PTMs

More than 100 PTMs of proteins have been described, including disulfide bonds, phosphorylation, glycosylation, and myristoylation. A polypeptide transcript may be processed differently in different tissues and depend on developmental stage and health of the organism. Epigenomics, the study of histone PTMs on a global scale, is an expanding area of research in neuroscience, recently reviewed in this journal (Volume 38, Number 1, 2013). Conversely, the quantitative study of a specific PTM on a global scale can elucidate the role of that PTM in a disease state. For instance, large-scale studies of protein phosphorylation are frequently employed to measure changes in intracellular signaling pathways (Nilsson, 2012).

Protein Phosphorylation

Reversible protein phosphorylation is a well-studied mechanism in the regulation of cell signaling. This PTM can increase or decrease protein activity, change protein subcellular localization or protein-protein binding characteristics. Important regulatory proteins are typically detected at substoichiometric concentrations in cells, which presents special challenges to their analysis.

Protein phosphorylation is by far the most-studied PTM in the area of the control of cell signaling. It is a reversible modification of serine, threonine, and tyrosine, enzymatically catalyzed by kinases (phosphorylation, 'writer') and phosphatases (dephosphorylation 'eraser'). SH2-domain binding proteins recognize and bind to sites of tyrosine phosphorylation, providing a specific 'reader' protein that can modulate tyrosine phosphorylation-mediated signaling pathways (Table 1). Recently, quantitative phosphoproteomic analysis of hippocampus samples from Alzheimer's disease (AD) patients compared with control samples found

changes in several proteins involved in metabolism and signal transduction (Di Domenico *et al*, 2011). An integrated transcriptomic–phosphoproteomic study of neuronal retrograde injury responses in a rat model identified several hundred redundant signaling networks; those may provide a relatively robust response to neuronal damage (Michaevlevski *et al*, 2010).

Protein Glycosylation

Unlike protein phosphorylation, protein glycosylation is not a single chemical entity. Glycan structures are among the most complex structures in the biochemical space. The coding density of sugars is significantly higher than that of amino acids or nucleotides (Laine, 1997). For a given saccharide of six sugar units, the number of possible structural isomers is 1.44×10^{15} ; for a six-amino-acid-long peptide, it is 6.4×10^7 ; and for a nucleotide of six, it is 4096. The branching nature of oligosaccharides, the substitution of sugars with phosphate or sulfate, and differing anomeric linkages all contribute to the structural diversity of carbohydrates, yielding rich complexity dubbed ‘the sugar code’ (Gabius *et al*, 2002, 2004) that is a key determinant of cell adhesion, receptor activation, and intracellular transduction. The complexity of carbohydrate epitopes on proteins (and lipids) modulates glycan–glycan interactions at the cell surface and intra- or extracellular glycan–protein interactions.

Proteins that recognize an extended glycoepitope but lack enzymatic or transporter activities are called lectins. The name lectin is derived from the past participle of *legere* (from Latin, to read) and are indeed readers of the sugar code (Table 1). Dysfunctional lectins and lectin polymorphisms have been implicated in neurological and psychiatric disorders. Galectins are defined by their affinity of β -galactosides and modulate cell cycle, apoptosis, angiogenesis, and cell motility. Galectin-3 has been associated in inflammatory diseases. Neuroinflammation has an important role in age-associated cognitive disorders. A recent study of genetic variation in the gene that encodes galectin-3 (*LGALS3*) showed an association of certain *LGALS3* SNPs to higher baseline C-reactive protein concentrations and subtle attention defects (Trompet *et al*, 2012). Galectin-1 is another lectin whose glycan interactions have been shown to be essential in tempering microglial activation, brain inflammation, and neurodegeneration, with therapeutic implications for treatment of multiple sclerosis (Starossom *et al*, 2012). This lectin was recently implicated in the disease mechanism of the hereditary disorder giant axonal neuropathy (Mussche *et al*, 2012). In addition to binding glycoproteins and glycolipids, and acting as a noncovalent crosslinker, lectins can activate intracellular signaling cascades. For instance, the attachment of galectin-3 to branched N-glycans modulates focal adhesion remodeling through activation of focal adhesion kinase and phosphoinositide 3-kinase (Ono *et al*, 2001).

TABLE 1 Examples of Writers, Readers and Erasers as Key Control Proteins in Biological Systems

Writers	Readers	Erasers
Histone acetylases	BRD proteins	Histone deacetylases
Kinases	SH2-domain-containing proteins	Phosphatases
Glycosyltransferases	Lectins	Glycosidases
Ubiquitin ligases	UB-binding proteins	Deubiquitinases

The interplay of their activities contributes to the complexity of biochemical signaling pathways in health and disease states. Thus, integrated biological studies benefit greatly from defining protein phosphorylation, glycosylation, histone modifications, and ubiquitin-mediated signaling.

There is biochemical cross-talk between aberrant cell surface glycosylation patterns and intracellular phosphorylation cascades; recent investigations have demonstrated that modulation of glycogene function by gene therapy can suppress tumor growth *in vivo* and induce measurable changes in kinase-mediated signaling (Moskal *et al*, 2009; Kroes *et al*, 2010). Reversible serine and threonine O-glycosylation by N-acetyl glucosamine has been demonstrated to act as an off-switch for protein phosphorylation (Hart *et al*, 2011).

Proteoglycans and glycoproteins are enriched in the extracellular matrix (ECM). In addition to disrupted extracellular lectin–matrix interactions, dysregulation of ECM components are associated with neuropsychiatric disorders. In a genome-wide association study of 682 patients and 1300 control subjects, genetic variation in the neurocan gene was demonstrated to be linked to bipolar disorder (Cichon *et al*, 2011). A follow-up study to derive main clinical factors associated with the neurocan risk allele was most strongly linked to mania, especially overactivity (Miro *et al*, 2012). Furthermore, neurocan knockout mice showed hyperactivity and other mania-related symptoms that were decreased by lithium administration (Miro *et al*, 2012). The aberrant organization of the ECM in AD was studied by Morawski *et al* (2012). ECM components, such as brevican and aggrecan, form perisynaptic axonal coats and perineural nets, respectively. Loss of brevican was correlated to synaptic loss, but other ECM components resisted degradation.

GLYCOMICS

Glycans decorate cell surface glycoproteins and glycolipids; as previously mentioned, these oligosaccharide structures are some of the most complex and diverse structures in vertebrate cells. The advancement of glycomics has been challenged by the need to develop analytical tools to characterize glycans on proteins and lipids in large-scale studies. Previously, definitive structural assignment, including anomeric sugar linkages, required nuclear magnetic resonance (NMR) methods, which are highly specific in their output but, in general, lack sensitivity for global

studies of cellular glycans. Advances in MS have recently enabled high-sensitivity analysis of both protein and lipid glycans. Low-flow liquid chromatography (LC) separations (nano-LC) coupled to low-flow ESI (micro and nano-electrospray) (Emmett and Caprioli, 1994; Wilm and Mann, 1994) with high-resolution MS, FT-ICR MS, and MS/MS analysis greatly facilitates identification of glycans from extremely complex biological mixtures. Lower-resolution MS and other ionization methods can also be used in glycoconjugate analysis, albeit with lower confidence in identification. MALDI-MS analysis is fast, tolerant of salts, and uses little sample, but often the MALDI process causes the loss of labile saccharides (especially sialylation), resulting in the loss of structural information. FT-ICR MS analysis provides the highest mass accuracy and resolution, which greatly aid in the identification of sugar residues, but the analysis time (scan speed) is slow compared with other MS analyses. The slower scan speed does not take full advantage of the chromatographic separation, but the high-resolution FT-ICR analysis permits identification of most overlapping peaks. Some direct glycomic applications with high-resolution FT-ICR MS will be discussed below.

To determine the importance of glycan changes in disease, it is necessary to coordinate many data sets. As proteins mediate the biosynthesis and degradation of complex glycoconjugates, it is necessary to merge transcriptomic, proteomic, glycomic, lipidomic, and even metabolomic data sets to chart pathways involved in altered glycoconjugate phenotypes in disease. The integrated biological approach has been successfully developed through the correlation of phenomic, transcriptomic, phosphoproteomic, and lipidomic data sets (He *et al*, 2007, 2010a,b; Kroes *et al*, 2010; Nilsson *et al*, 2010; Nilsson, 2012). Computational mathematical algorithms described in the next section were developed (Goerke *et al*, 2011; Meyer-Baese *et al*, 2012a,b) to extract useful correlations from these massive data sets to generate biochemical pathways to drive intelligent experimental design and to elucidate therapeutic targets. The goal of these studies is to further understand glycan changes to generate new therapeutic targets and identify novel biomarkers for disease treatment and early disease diagnosis.

Analysis of glycoproteins involves not only the structural identity of the attached glycan, but it is also important to know the site of the glycan attachment to the protein. Collisionally induced fragmentation in a mass spectrometer will often result in the removal of the glycan, resulting in the loss of the attachment site. Coupling a softer fragmentation technique first (such as electron capture dissociation (ECD) (Zubarev *et al*, 1998), with a heating fragmentation such as infrared multiphoton dissociation (IRMPD) (Little *et al* 1994) as demonstrated by Hakansson *et al* (2001), provides complementary data that identifies the site of glycosylation attachment with ECD and complementary carbohydrate structural information with IRMPD. Altered protein glycosylation has been detected in AD (Sihlbom *et al*, 2004, 2005). CSF from individual patients with AD was separated by narrow range two-dimensional (2D) gel electrophoresis,

in-gel-digested with trypsin, and analyzed by electrospray FT-ICR MS. Glycopeptides were fragmented with IRMPD, which produces abundant fragment ions through dissociation of glycosidic linkages. MS/MS with high mass accuracy facilitates the structural determination of site-specific N-linked glycosylation and possible identification of AD biomarkers. These analyses can be performed on enzymatically released glycans from the glycoproteins, or intact glycopeptides. The released glycans are collected from the peptides/proteins and further chromatographically purified on-line before high-resolution MS and MS/MS analysis (Wang *et al*, 2010). These analyses are information-rich, typically producing hundreds of glycan structures from a single sample.

LIPIDOMICS

About 50% of the brain and spinal cord is composed of lipid material. Bioidentical lipids have even been examined as therapeutics; for instance, intracerebroventricular administration of GM1 ganglioside to early onset Alzheimer patients showed positive effects, such as improved motor performance and neuropsychological assessments (Augustinsson *et al*, 1997; Svennerholm *et al*, 2002). Lipid profiling or lipidomics is often bundled into metabolomics, but lipids represent a distinct chemical and functional group of biomolecules. Lipids are enriched in the plasma membrane and lipid rafts. Lipids also are important in the compartmentalization of the cell by formation of organellar membranes. Polar lipids, such as gangliosides are cell-surface glycoconjugates that modulate cell adhesion, migration, and signal transduction through interactions, with growth factor tyrosine kinases and other kinase pathway members. Lipids are also responsible for altering membrane properties (eg, membrane permeability) and directly interact with effector proteins (eg, G-proteins) (Chun, 2005; Piomelli, 2005). Lipids are implemented as key factors in many neuroactive processes, such as second messenger signaling, potassium channel flux, cannabinoid receptor activity, and NMDA receptor function (Piomelli *et al*, 2007). Lipidomic data from neurological and psychiatric model systems can elucidate activation of enzymatic, receptor, and signaling pathways.

Polar lipid analysis has been documented in GBM and GSC cultures (He *et al*, 2007, 2010a,b) and correlated with phenotypic, transcriptomic, and functional proteomic changes (He *et al*, 2010a,b). Lipid profiling identifies not only the abundant phospholipids, but also the glycolipids in mammalian cells. Phospholipids are major components of the cell membrane, and their compositions and modifications affect the membrane fluidity and phospholipid-mediated signaling pathways. The methodology described by He *et al* (2007) was applied to identify phosphatidylinositol (PI), phosphatidylethanolamine, phosphatidylcholine (PC), sphingomyelin (SM), and many other polar lipids with a wide range of acyl chain lengths and varying

numbers of hydroxylations and unsaturations. Changes in phospholipid composition not only modulate the membrane fluidity, but also affect phospholipid-mediated signal transduction. For example, phosphorylated PI, including inositol 3,4-bisphosphate [PI(3,4)P2] and inositol 3,4,5-triphosphate [PI(3,4,5)P3], has been shown to affect cell proliferation, survival, and movement (Di Paolo and De Camilli, 2006). The presence of a double bond results in a 'kink' in the phospholipid fatty acyl chain. Thus, cells with a high concentration of unsaturated phospholipids may possess a more fluid cell membrane relative to cells with more saturated ones, because unsaturated phospholipids pack much more loosely. Moreover, differentiated tumor cells show an increase in PC and SM, which are predominantly localized on the outer plasma membrane and carry positive charges (Yamaji-Hasegawa and Tsujimoto, 2006). Together, these changes will alter membrane stability, modulate cell surface charge, and alter signal transduction via intracellular pathways.

MS analysis is the analytical detector of choice in lipid analysis. The primary methodology employed for lipid analysis involves a simple solvent/solvent extraction of the lipids and partitioning of the polar and nonpolar lipids into an aqueous and organic phase, respectively. Because of the complexity of the lipid mixture, a chromatographic separation is necessary before MS analysis. Reverse-phase chromatography is often used for the separation of the lipids, but is not the optimal choice because of the risk of irreversible binding of lipids to that resin. He *et al* (2007) optimized separation of the lipids over a chromatographic resin that focusses on π - π stacking (such as phenyl-hexyl), which greatly improves the recovery of polar lipids from the chromatographic column. Complex polar lipid mixtures are reasonably resolved chromatographically with phenyl-hexyl resins based on variations of oligosaccharide and aglycone moieties. Nano-LC was used to further increase sensitivity and was introduced by high-sensitivity microelectrospray (Emmett and Caprioli, 1994; Emmett *et al*, 1998) to a high-resolution FT-ICR MS to further aid in the detection of the complex lipid mixture (Schaub *et al*, 2007, 2008). Negative-ion micro-electrospray provided better sensitivity compared with positive-ion mode for analyzed polar lipids (Leverly, 2005). The lipid analysis methodology described above has been applied to analyses of cell culture, tissue, blood, and CSF samples. MALDI-MS imaging is showing great potential to analyze lipids in their histological context, which provides complementary data to information gained through analysis of total extracts (Sparvero *et al*, 2010; Hanada *et al*, 2012).

Accurate mass (typically better than 1 ppm), LC retention time trends, along with available tandem mass spectra generated from collisional-induced dissociation of precursor enable the determination of chemical composition and proposed structural assignment of 600–800 glycolipids and phospholipids per sample. A Kendrick mass analysis algorithm aids in assignment of structural families of polar lipids (Hughey *et al*, 2001). High mass accuracy and high

resolving power of high-field FT-ICR MS (Senko *et al*, 1997; Marshall *et al*, 1998; Leverly, 2005; Schaub *et al*, 2008) greatly improves the resolution and assignment of glycolipids and phospholipids in the complex mixture of polar lipid extracts from biological samples.

METABOLOMICS IN NEUROPHARMACOLOGY

To elucidate the etiology of neuropsychiatric disorders, metabolomic data can be quite complementary to proteomics. For example, Zhang *et al*, 2011 studied high and low anxiety-related behavior in mice by this combined approach. One conclusion of the integrated study suggested that the inositol pathway was implicated in mice with a genetic predisposition to anxiety. Taken together, their findings described proteins, metabolites, and pathways that could be related to the anxiety phenotype. Metabolites vary in size, polarity, charge, structure, and elemental composition, all of which reside as a highly complex mixture in a biological matrix; thus, comprehensive metabolite profiling is a daunting task that cannot be performed by a single analytical technique. Metabolomic profiling in neuroscience is discussed at length by Wood (2013) elsewhere in this issue. This section is meant to encourage metabolomic analysis, to inform the reader of the methodologies available and challenges of each.

NMR spectroscopy analysis provides the most detailed structural information for metabolites, but its application has limits of sensitivity and complexity (Steuer, 2006). One-dimensional (1D) NMR spectroscopy is often used for metabolite profiling, because the analysis is fast (min), but the fast scan time results in low resolution and low signal-to-noise ratios on metabolites. 2D NMR greatly increases resolution of metabolites, but the tradeoff is long acquisition times (several hours) per sample (Motta *et al*, 2010). NMR sensitivity is improving with the development of microcoil probes and the ability of coupling NMR detection with LC concentration/separation of biological metabolites, but for analysis of metabolites at biological concentrations, MS is currently the method of choice (Rubakhin *et al*, 2011).

For MS analysis of metabolites, GC-MS is the gold standard (Griffin and Kauppinen, 2007; Wibom *et al*, 2010; Garcia and Barbas, 2011). Even though GC-MS is characterized by high sensitivity, the separation of metabolites is greatly improved by 2D GC (GC \times GC) analysis, because often more than one compound occupies a peak as it elutes from the GC. The second dimension separation in GC \times GC often reveals lower-level components that were masked in a 1D separation. Recently, GC \times GC separations have been applied increasingly to the analysis of metabolites because of the complexity of the biological samples (Mohler *et al*, 2006; Ralston-Hooper *et al*, 2011). The added separation afforded by GC \times GC decreases the complexity of the MS analysis of small metabolites. Both GC-MS and GC \times GC-MS techniques are limited to the analysis of volatile compounds

or compounds derivatized chemically to make them volatile. On the other hand, LC separations before MS analysis allows analysis of compounds without derivatization (Urban *et al*, 2010; Vuckovic and Pawliszyn, 2011) and, similar to GC, have the ability to separate chiral molecules under optimal conditions (Oh *et al*, 2011). The nonpolar component of metabolomics has been largely overlooked. Many nonpolar analytes once derivatized can be analyzed by GC or GC \times GC MS, but not all nonpolar metabolites can be detected. Supercritical fluid chromatography (SFC) coupled to MS analysis adds another dimension to the analysis of these complex nonpolar metabolite mixtures (Farrell *et al*, 2009). In a very simplistic description, SFC separations combine attributes of both GC and LC separations, because the separations are performed in a gas that is in a supercritical or enhanced fluidity state (Taylor, 2008). Molecules that are not volatile and thus intractable by GC can often be dissolved in an enhanced fluidity gas and be separated by SFC. The same can be said of nonpolar compounds that are insoluble in aqueous solvents and can't be separated by conventional reverse-phase LC. In addition, SFC is readily interfaced with MS through atmospheric ionization sources and, with short separation times, is considered a high-throughput analysis (Farrell *et al*, 2009).

The application of SFC-MS for metabolomics has been demonstrated, but currently SFC-MS is underutilized in metabolomic analysis. Most of the methodologies described above involve different separation techniques coupled with MS, but all of the technologies are complementary and, when used together, will provide the most complete metabolomic data sets to date. Another factor involved in the MS analysis of metabolites is the efficient ionization of the compounds before MS detection. Multiple ionization modes are needed to detect comprehensively the diverse chemical compounds in the complex biological metabolome, including positive and negative ionization electrospray, chemical ionization in both positive and negative modes, and photoionization. Once separated and ionized, elemental composition assignment of metabolites is the first step of analysis. NMR analysis is superior to MS, because it provides exact structural information for compounds; however, the sensitivity of the technique can be insufficient for the analysis of complex biological mixtures. Low-resolution MS analysis necessitates MS/MS in most cases to provide elemental composition. Ultrahigh-resolution high mass accuracy MS (FT-ICR MS) often can assign elemental composition based on accurate mass, but sub-ppm mass accuracy alone maybe insufficient for elemental assignment because of the large number of possible atomic configurations for a given nominal mass. Recently, accurate mass coupled with ultrahigh-resolution isotopic fine structure information (as shown in Figure 2) has demonstrated that compounds can be correctly identified (Marshall *et al*, 1998; Shi *et al*, 1998; He *et al*, 2004; Marshall and Hendrickson, 2008; Miura *et al*, 2010). By using the hetero-atom content of the metabolite ion, nominal spacing

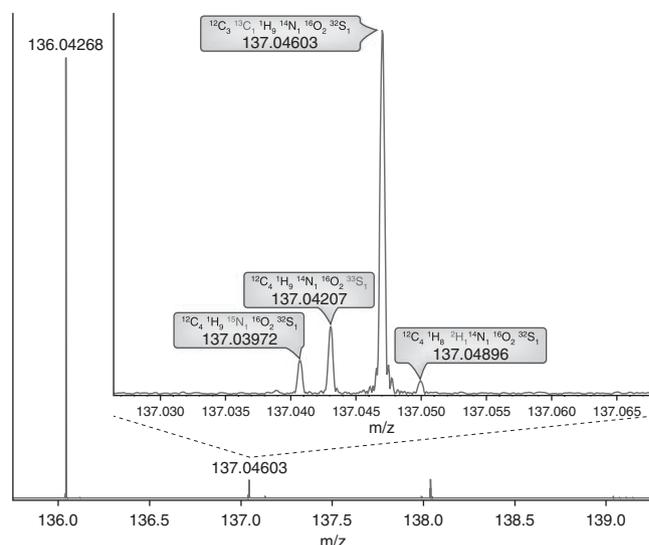


Figure 2. Ultrahigh-resolution isotopic fine structure of a metabolite, S-methyl-L-cysteine. Spectra were collected with a 12 T FT-ICR MS with a mass resolving power of ~ 340 K. The exact mass of S-methyl-L-cysteine $[M + H]^+$ is 136.04267. The zoomed inset shows the first ^{13}C isotopic fine structure. The natural isotope abundances are: ^{15}N , 0.36%; ^{33}S , 0.76%; ^{13}C , 1.1%; and ^2H , 0.015%. The exact mass, isotopic fine structure, and isotope abundances are used to assign unambiguously elemental compositions.

of isotopic intensities, and ultrahigh mass accuracy, metabolites' elemental composition can be determined unambiguously (Daltonics B, 2012).

In summary, the acquisition of comprehensive metabolite data sets requires an enhanced metabolomics toolbox of complementary analytical technologies and methods. There is no single assay that permits the detection and identification of all metabolites in a biological sample; thus, several analytical techniques must be employed to perform global metabolomic analysis. Because of the complexity of these analyses, a targeted metabolite approach is often adopted for specific experiments.

COMPUTATIONAL METHODS FOR INTEGRATIVE NEUROBIOLOGY

Computational approaches have facilitated quantitative analysis and data interpretation of large data sets stemming from EEG, MEG, PET, CT, MRI, and fMRI, showing the response of the brain to placebo and drug effects for substance use and cognitive disorders. The computational techniques range from (a) neural dynamical models describing synchronization to (b) multivariate pattern recognition techniques for emerging spatial and temporal maps as a response to a certain task or analyzing impulse control in connection to drug abuse, and finally to (c) graph theoretic approaches to elucidate the brain network topology and contrast structural and functional connectivity deficiencies associated with disorders (Stam *et al*, 2010; Fornito *et al*, 2012; Giessing and Thiel, 2012).

The latter, also known as ‘imaging connectomics’ (Fornito *et al*, 2012), represents a novel trend in neuroimaging techniques. It is based on analysis of graph networks, whose topologies aid in identifying areas of malfunctioning brain network architectures as targets for cognitive enhancing drugs. Existing studies, which are few, have shown that graph networks represent a powerful tool in the development of data-driven biomarkers of cognitive enhancing drug effects, on both local and global scales. Also, they provide a framework to understand the dynamics of neurodegenerative diseases. A new approach in drug development, network pharmacology, aims at addressing the whole brain instead of targeting specific areas, and represents a new route in therapeutics. Although neuroimaging scientists have embraced recently graph theoretic approaches, integrated biological studies in neuropharmacology based on large graphs are sparse (limited to only two studies of visualization in AD and Parkinsons diseases) (Paananen and Wong, 2009; You *et al*, 2010). Large-scale biological data analysis and their integration in neuropharmacology based on graph networks must overcome two major challenges: the high dimensionality of the data compared with neuroimaging and the disparate data sources to be integrated within the same network. Advances in graph theoretical analysis in the context of statistics and the development of visualization and navigation tools in large graphs have been combined to overcome these major challenges.

The first steps in this direction can be found in several approaches to mathematical modeling that have been developed to infer relationships between large data sets derived from complex biological systems. Biological systems are highly nonlinear and, thus, mathematical approaches are either qualitative, such as the graph network analysis based on correlations, or quantitative, such as the nonlinear multi-mode autonomous deterministic or stochastic system. However, one major issue needs to be addressed in order to facilitate the application of these techniques: the feature size of the measured data is far larger than the sample size, a phenomena known as the ‘curse of dimensionality’ in data mining. Effective feature reduction in case of qualitative approaches or model order reduction for quantitative techniques becomes imperative in order to reduce the complexity in computation and enhance the generalization capabilities of the proposed model. The field of transcriptomics has benefited greatly from the development of computational models to determine gene regulatory networks (GRNs) (Hecker *et al*, 2009), but the techniques are applicable to proteomic and glycomic data sets as well. In the following section, we will present some modern techniques that pertain to nonlinear quantitative and qualitative model reduction.

Approaches to Model GRNs

GRNs represent gene circuits that interact and regulate their own and the activity of other genes by protein expression. Many GRNs are described by complex models, which are

difficult to analyze and control (Hecker *et al*, 2009). The large-scale nature of these systems and the highly complex underlying models require reduced-order models to facilitate their analysis. Balanced truncation is known as a popular method for model reduction, as it is relatively simple and it yields a qualitatively accurate system. The interpretation of most balancing techniques (Scherpen, 1993) is based on the concept of past input energy (controllability) and future input energy (observability). For a GRN system to be controllable means to be able to reach a given state under the influence of an input signal, whereas to be observable means that there is enough information to observe all possible states of the system. The idea behind transforming a system into balanced form is to easily detect and remove a state component of the initial system to obtain a reduced-order model. Although for linear systems finding a balancing coordinate transformation via solutions of the controllability and observability, Lyapunov equations (Hamilton, 1994) are quite easy, but for nonlinear systems these equations are almost impossible to solve and, thus, balancing becomes in general a challenging task. Carleman bilinearization (Mohler, 1973; Brockett, 1976; Meyer-Baese *et al*, 2012a,b) facilitates the representation of a nonlinear system by a bilinear form. The idea behind the Carleman bilinearization procedure is to approximate the nonlinear function in the GRN by a Taylor expansion. This procedure preserves a certain contribution of the nonlinearity of the system in the resulting simplified form. The achieved reduction represents a simple estimate for the additional parameters employed and is, at the same time, computationally nonintensive for deterministic GRN under assumption of a bounded input. The determined model reduction has potential application for reverse engineering and robust biosynthetic GRN design.

Approaches to Model Phosphoprotein-Mediated Networks

Advances in MS techniques have facilitated the identification and quantification of very large numbers of phosphoproteins and corresponding phosphorylation sites. These identified and quantified phosphorylation sites provide a wealth of information about the biochemical status of cells and tissue, and have an increasing role to elucidate the mechanism and importance of novel therapeutics in cancer research and other diseases. However, the extensive data sets do not automatically provide deep insight into deciphering the phosphoproteome and, most importantly, how to use the acquired knowledge in reverse engineering and drug design. Quantitative data analysis methods, such as statistical data analysis and predictive analytical techniques, are essential to derive knowledge from the data.

For large data sets at the protein level, network-based perspectives offer a solution to this problem and many node-edge graphs are stored in public databases. Most protein networks are assembled from literature or inferred merely from data and cannot predict the response of cells to

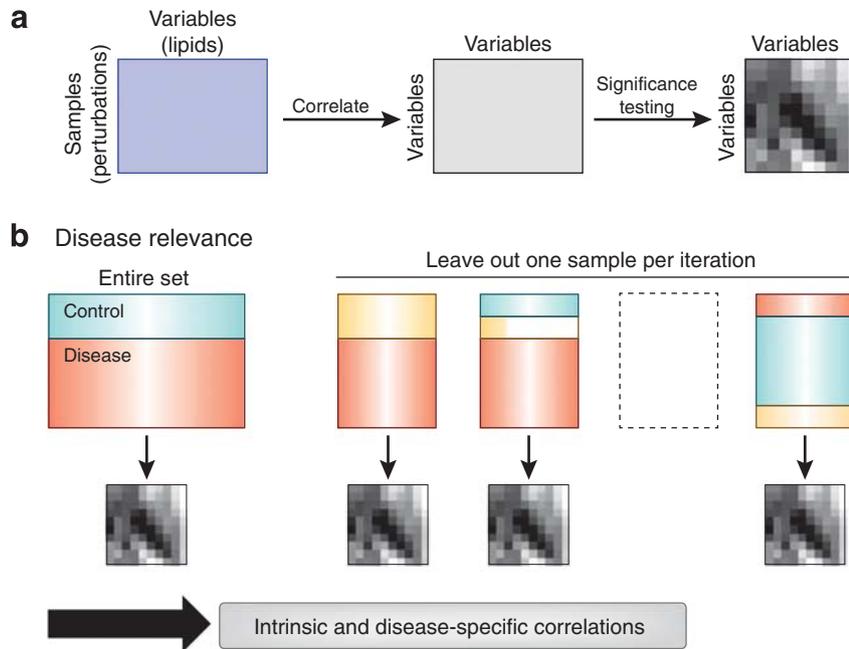


Figure 3. The raw lipidomic data is a matrix of samples over variables. (a) The samples are the individual perturbations, which are grouped into control samples and the sample(s) of interest to the study, here simply called 'disease'. Partial correlations of all variables are obtained and later evaluated with respect to statistical significance. (b) To investigate whether a significant partial correlation is specific for the disease sample, partial correlations (as in a) were calculated for the entire data set, as well as for data sets where each one sample was left out. Unless a correlation is significant in all Gaussian Graph Models (GGMs), it is considered disease-specific. This figure is reproduced with the permission of the publisher (Mueller *et al*, 2011).

aberrant signals. Few attempts to specify a pathway as a logical model emulating the signal input/output flow through gates in a parallel optimization against experimental data have been described. However, for site-specific phosphorylation data sets, both the theoretical and computational integration tools are very limited. The analysis of phosphoproteomic data highly pertinent to systems biology and disease informatics is either data-driven or topology-driven. Data-driven analysis requires *a priori* information of the signaling topology and includes standard techniques from chemometrics, such as principal component analysis and partial least square regression. Topology-driven techniques are based on ordinary differential equations, fuzzy logic, and Boolean logic, and employ canonical pathways as a scaffold. The only theoretical study for phosphoproteomic data constructs large signaling pathways using an adaptive perturbation approach (Melas *et al*, 2012). Pathway optimization is accomplished via integer linear programming. Two computational tools are available in the form of advanced bioinformatics platforms: NetworKIN (Linding *et al*, 2008) and PhosphoSiteAnalyzer (Bennetzen *et al*, 2012). NetworKIN predicts kinase families based on detailed linear motif analyses by employing an artificial neural network and scoring matrices. The prediction of kinase families responsible for the phosphorylation of a given substrate are determined with high accuracy and sensitivity. PhosphoSiteAnalyzer is currently the most advanced tool and enhances the capabilities of NetworKIN by using it to obtain kinase predictions from several phosphorylation

sites. A large-scale phosphoproteomic data analysis is accomplished based on integrated bioinformatics modules. These modules perform a statistical data analysis and extract relevant biological information.

Future research initiatives in theoretical phosphoproteomic modeling must focus on mathematical formulations that do allow a prediction of therapeutic targets or a desired response to a certain stimulus. Computational bioinformatic tools should provide improved solutions for facilitating complex kinase-substrate network analysis in a user-friendly integrative manner.

Identification of Novel Interactions Based on Compressed Graph Networks

The analysis of transcriptomic, proteomic, and metabolomic data sets with existing knowledge-based tools, such as IPA, can only reveal interconnected pathways that are already represented in the scientific literature. New models are needed to 'connect the dots,' making occult connections between data sets obvious. As physicist Richard Feynman expressed it, 'There is a rhythm and a pattern between the phenomena of nature which is not apparent to the eye, but only to the eye of [mathematical] analysis.' The Gaussian Graph Model (GGM) technique is one approach to create interrelated 'omics' graph network and determine novel correlations (Mueller *et al*, 2011). Reproducible quantitative changes in (phospho)proteins, transcripts, and metabolites in response to disease or other perturbations can contain correlations that cannot be derived based on intuitive

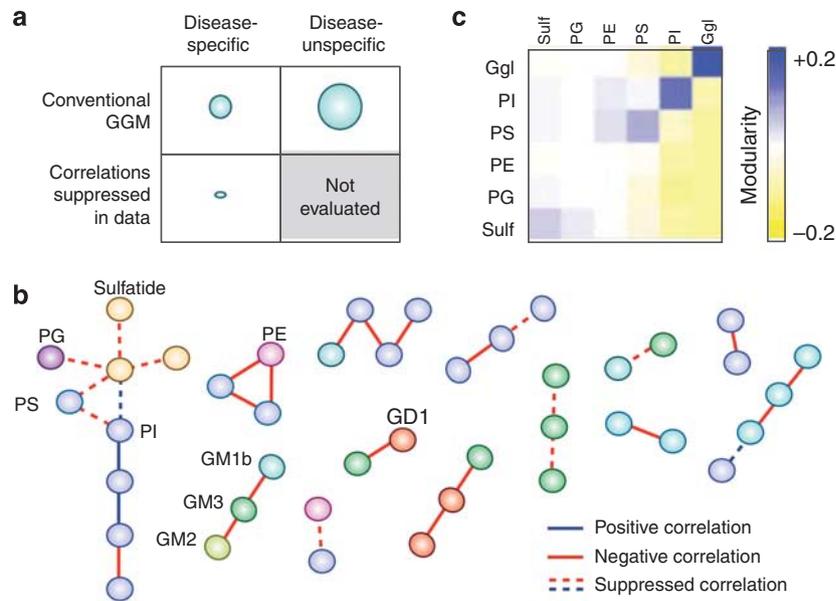


Figure 4. Lipids specifically regulated when a glioma cell was effectively perturbed. (a) Relative numbers of disease-specific and -unspecific lipid-lipid partial correlations in the Gaussian Graph Models (GGMs). Analysis of the entire data set is named 'conventional' GGM with respect to disease specificity. (b) Disease-relevant GGM, which is associated with the combined perturbation by WT-p53 adenoviral transfection before SN38 chemotherapy in U87 glioma cell lines. (c) Modularity matrix was calculated by using the lipid species as class label for the GGM in b. PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PI, phosphoinositol; PS, phosphoserine. This figure is reproduced with the permission of the publisher (Mueller *et al*, 2011).

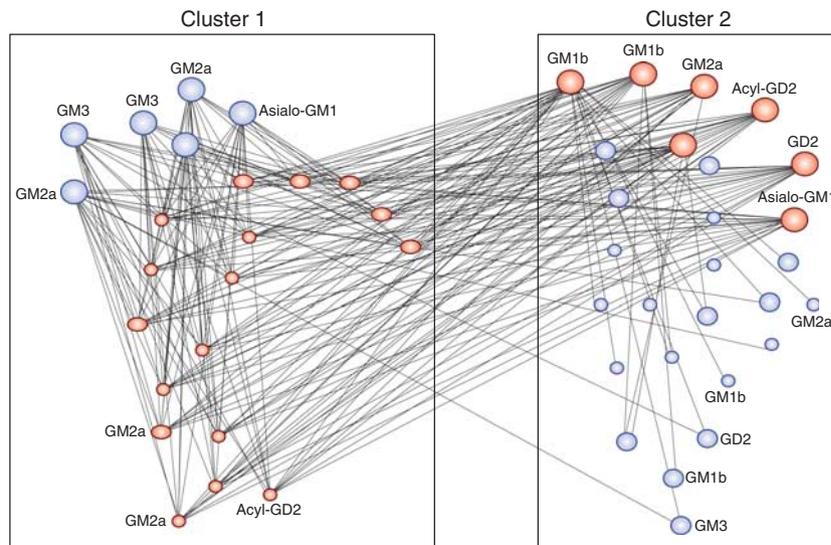


Figure 5. The clustered glycolipids in treatment with WP1193 (threshold of 0:9). The edges are colored on a linear HSV-scale (from blue = cold to red = hot) by the absolute value of the edge weight, and the size of the nodes is proportional to sum of their incident absolute edge weights. In this difference graph, a node is large if it differs strongly in its correlations to other nodes in the WP1193 treatment. Small nodes roughly preserve their correlation behavior in this treatment. This figure is reproduced with the permission of the publisher (Goerke *et al*, 2010).

reasoning. Visualization by generation of Gaussian Graph Maps is a powerful technique to find novel correlations and can be applied to derive new knowledge about the function of activators or inhibitors in brain-derived tissues or cells.

To identify the effects on lipid levels in glioma cells as a result of drug treatment, lipidomic data sets from control and treated cells were acquired (He *et al*, 2007). In contrast to simple Pearson's correlation networks the partial correlations of GGMs only identify direct correlations,

whereas indirect associations are eliminated as shown in Figure 3. By implementing disease-driven GGM, we revealed GBM-specific lipid correlations to advance biomedical research on novel gene therapies. A summary of our results is found in Figure 4.

We have previously applied graph-clustering techniques to determine correlations in the glycomic and glycotranscriptomic responses of glioma stem cells to pharmacological treatment with a STAT3 phosphorylation inhibitor or

fetal bovine serum (Goerke *et al*, 2011), yielding characteristic fingerprints of glioma stem cell responses that could not have been determined by searches of the scientific literature. We developed a novel technique to replace standard correlation networks to better reveal the topological properties reflecting the response to therapeutic interventions (Steuer *et al*, 2003; Steuer, 2006). Traditional correlation networks are based on computing the Pearson's correlation coefficients between the nodes of the graph corresponding to the distinct measurements and establishing an edge if the correlation value exceeds a specified threshold.

Our paradigm is based on graph-clustering techniques as an equivalent to standard correlation networks for the analysis of therapeutic outcomes for GBM cells. Global multivariate approaches become imperative to comprehend the potential dysregulation of normal cellular responses in disease and their response to various therapeutic interventions, and to identify causal influence relationships involving multiple biological analytes. Figure 5 shows an example of such an emerging graph network structure.

Algorithms that are robust and efficient for large-scale experiments and drug design are an active field of development in mathematics and will become standard tools in the future for biomedical scientists to identify novel correlations between dysregulated pathways in neurological and psychiatric disease.

FUTURE DIRECTIONS AND CLINICAL IMPLICATIONS

The recent advancements in human genomic research are essential to other global biological studies, but not completely sufficient to describe complex underlying phenomena associated with neurological and psychiatric disorders. The application of integrative biological techniques in the realm of neuropsychiatry and neuropharmacology is increasing in number and frequency in the scientific literature. The success of these endeavors will require new collaborations between basic science and clinical teams. Through the acquisition and integration of quantitative data sets from two or more biomolecular subsets, synergies arise and promise to provide the field with new therapeutic targets and biomarkers of disease and drug response.

FUNDING AND DISCLOSURE

Dr Priebe is the founder and a shareholder of Moleculin, IntertechBio, and Houston Pharmaceuticals. He is also the cofounding scientist, shareholder, and Scientific Advisory Board member of the Reata Pharmaceuticals, and has received financial compensation. Dr Moskal is the founder of Naurex. He has founders' shares of stock in the company. He receives financial compensation as a consultant. Dr Kroes has been a consultant for Naurex for the past 3 years, and has received financial compensation and stock. Dr Nilsson is a former employee of Pfizer Global Research

and Development (2007–11). Drs Conrad, Emmett, and Meyer-Baese have no conflicts of interest to disclose.

ACKNOWLEDGEMENTS

Dr Laezza is supported by NIMH R01 MH095995 (FL).

DISCLAIMER

All authors contributed to the composition of this review. The views expressed in this review are those of the authors and may not necessarily reflect the position of the University of Texas (CLN, MRE, FEL, CAC, and WP), Florida State University (AMB), or NIH (FEL, CAC).

REFERENCES

- Ahmed MM, Gardiner KJ (2011). Preserving protein profiles in tissue samples: Differing outcomes with and without heat stabilization. *J Neurosci Methods* **196**: 99–106.
- Augustinsson LE, Blennow K, Blomstrand C, Brane G, Ekman R, Fredman P *et al* (1997). Intracerebroventricular administration of GM1 ganglioside to presenile Alzheimer patients. *Dement Geriatr Cogn Disord* **8**: 128–136.
- Barczak A, Rodriguez MW, Hanspers K, Koth LL, Tai YC, Bolstad BM *et al* (2003). Spotted long oligonucleotide arrays for human gene expression analysis. *Genome Res* **13**: 1775–1785.
- Bennett V, Lambert S (1999). Physiological roles of axonal ankyrins in survival of premyelinated axons and localization of voltage-gated sodium channels. *J Neurocytol* **28**: 303–318.
- Bennetzen MV, Cox J, Mann M, Andersen JS (2012). PhosphoSiteAnalyzer: a bioinformatic platform for deciphering phosphoproteomes using kinase predictions retrieved from NetworkKIN. *J Proteome Res* **11**: 3480–3486.
- Berghs S, Aggujaro D, Dirx R Jr., Maksimova E, Stabach P, Hermel JM *et al* (2000). betaIV spectrin, a new spectrin localized at axon initial segments and nodes of ranvier in the central and peripheral nervous system. *J Cell Biol* **151**: 985–1002.
- Blencowe BJ, Ahmad S, Lee LJ (2009). Current-generation high-throughput sequencing: deepening insights into mammalian transcriptomes. *Genes Dev* **23**: 1379–1386.
- Bloom G, Yang IV, Boulware D, Kwong KY, Coppola D, Eschrich S *et al* (2004). Multi-platform, multi-site, microarray-based human tumor classification. *Am J Pathol* **164**: 9–16.
- Borsotto M, Cavarec L, Bouillot M, Romey G, Macciardi F, Delaye A *et al* (2007). PP2A-Bgamma subunit and KCNQ2 K⁺ channels in bipolar disorder. *Pharmacogenomics J* **7**: 123–132.
- Brachet A, Leterrier C, Irodelle M, Fache MP, Racine V, Sibarita JB *et al* (2010). Ankyrin G restricts ion channel diffusion at the axonal initial segment before the establishment of the diffusion barrier. *J Cell Biol* **191**: 383–395.
- Brockett R (1976). Volterra series and geometric control theory. *Automatica* **12**: 167–176.
- Brusse E, De Koning I, Maat-Kievit A, Oostra BA, Heutink Pandvan Swieten JC (2006). Spinocerebellar ataxia associated with a mutation in the fibroblast growth factor 14 gene (SCA27): a new phenotype. *Mov Disord* **21**: 396–401.
- Buffington SA, Rasband MN (2012). The axon initial segment in nervous system disease and injury. *Eur J Neurosci* **34**: 1609–1619.
- Burgdorf J, Kroes RA, Weiss C, Oh MM, Disterhoff JF, Brudzynski SM *et al* (2011). Positive emotional learning is regulated in the medial prefrontal cortex by GluN2B-containing NMDA receptors. *Neuroscience* **192**: 515–523.
- Burgdorf J, Zhang X-L, Nicholson KL, Balster RL, Leander DJ, Stanton PK *et al* (2013). GLYX-13, a NMDA receptor glycine-site functional partial agonist, induces antidepressant-like effects without ketamine-like side effects. *Neuropsychopharmacology* **38**: 729–742.
- Califano A, Butte AJ, Friend S, Ideker TandSchadt E (2012). Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nat Genet* **44**: 841–847.
- Canales RD, Luo Y, Willey JC, Austermiller B, Barbacioru CC, Boysen C *et al* (2006). Evaluation of DNA microarray results with quantitative gene expression platforms. *Nat Biotech* **24**: 1115–1122.

- Cappadona S, Baker PR, Cutillas PR, Heck AJ, RandBreukelen B (2012). Current challenges in software solutions for mass spectrometry-based quantitative proteomics. *Amino Acids* **43**: 1087–1108.
- Carter MG, Hamatani T, Sharov AA, Carmack CE, Qian Y, Aiba K et al (2003). In situ-synthesized novel microarray optimized for mouse stem cell and early developmental expression profiling. *Genome Res* **13**: 1011–1021.
- Chun J (2005). Lysophospholipids in the nervous system. *Prostaglandins Other Lipid Mediat* **77**: 46–51.
- Cichon S, Mühleisen TW, Degenhardt FA, Mattheisen M, Miró X, Strohmaier J et al (2011). Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet* **88**: 372–381.
- Cloonan N, Grimmond S (2008). Transcriptome content and dynamics at single-nucleotide resolution. *Genome Biol* **9**: 234.
- Cooper EC (2011). Made for 'anchoring': Kv7.2/7.3 (KCNQ2/KCNQ3) channels and the modulation of neuronal excitability in vertebrate axons. *Semin Cell Dev Biol* **22**: 185–192.
- Corcoran ME, Kroes RA, Burgdorf JS, SandMoskal JR (2011). Regional changes in gene expression after limbic kindling. *Cell Mol Neurobiol* **31**: 819–834.
- Courtney E, Kornfeld S, Janitz K, Janitz M (2010). Transcriptome profiling in neurodegenerative disease. *J Neurosci Methods* **193**: 189–202.
- Dabrowski M, Aerts S, Van Hummelen P, Craessaerts K, De Moor B, Annaert W et al (2003). Gene profiling of hippocampal neuronal culture. *J Neurochem* **85**: 1279–1288.
- Daltonics B (2012). *Definitive elemental formula determinatin debuts with Smart Formula-HR*, Bruker Daltonics Application Note. **FTMS-43**.
- Di Domenico F, Sultana R, Barone E, Perluigi M, Cini C, Mancuso C et al (2011). Quantitative proteomics analysis of phosphorylated proteins in the hippocampus of Alzheimer's disease subjects. *J Proteomics* **74**: 1091–1103.
- Di Paolo G, De Camilli P (2006). Phosphoinositides in cell regulation and membrane dynamics. *Nature* **443**: 651–657.
- Dubois E, Fertin M, Burdese J, Amouyel P, Bauters C, Pinet F (2011). Cardiovascular proteomics: translational studies to develop novel biomarkers in heart failure and left ventricular remodeling. *Proteomics Clin Appl* **5**: 57–66.
- Dupuy A, Simon RM (2007). Critical review of published microarray studies for cancer outcome and guidelines on statistical analysis and reporting. *J Natl Cancer Inst* **99**: 147–157.
- Emmett MR, Caprioli RM (1994). Microelectrospray mass spectrometry: ultra-high-sensitivity analysis of peptides and proteins. *J Am Soc Mass Spectrom* **5**: 605–613.
- Emmett MR, White FM, Hendrickson CL, Shi SD, Marshall AG (1998). Application of micro-electrospray liquid chromatography techniques to FT-ICR mass spectrometry to enable high-sensitivity biological analysis. *J Am Soc Mass Spectrom* **9**: 333–340. **This low-flow electrospray technique revolutionized biological mass spectrometry analysis.**
- Encode Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57–74. **The first "parts list" described for the human proteome.**
- Farrell WP, Aurigemma CM, Masters-Moore DF (2009). Advances in high throughput supercritical fluid chromatography. *J Liquid Chromatogr Related Technol* **32**: 1689–1710.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**: 64–71. **Electrospray ionization, a Nobel-prize awarded technology, was described for introducing samples to a mass spectrometer.**
- Filiou MD, Martins-de-Souza D, Guest PC, Bahn S, Turck CW (2012). To label or not to label: applications of quantitative proteomics in neuroscience research. *Proteomics* **12**: 736–747. **This review provides a comprehensive guide of quantitative proteomic methodologies, including labeling-based and label-free methods for basic and clinical neuroscience research.**
- Filiou MD, Turck CW (2012). Psychiatric disorder biomarker discovery using quantitative proteomics. *Methods Mol Biol* **829**: 531–539.
- Fornito A, Zalesky A, Pantelis C, Bullmore ET (2012). Schizophrenia, neuroimaging and connectomics. *Neuroimage* **62**: 2296–2314.
- Gabius HJ, Andre S, Kaitner H, Siebert H-C (2002). The sugar code: functional lectinomics. *Biochim Biophys Acta* **1572**: 165–177.
- Gabius HJ, Siebert H-C, Andre S, Jimenez-Barbero J, Rudiger H (2004). Chemical biology of the sugar code. *Chem Bio Chem* **5**: 740–764.
- Galiano MR, Jha S, Ho TS, Zhang C, Ogawa Y, Chang KJ et al (2012). A distal axonal cytoskeleton forms an intra-axonal boundary that controls axon initial segment assembly. *Cell* **149**: 1125–1139.
- Garcia A, Barbas C (2011). Gas chromatography-mass spectrometry (GC-MS)-based metabolomics. *Methods Mol Biol* **708**: 191–204.
- Gasser A, Ho TS, Cheng X, Chang KJ, Waxman SG, Rasband MN et al (2012). An ankyrinG-binding motif is necessary and sufficient for targeting Nav1.6 sodium channels to axon initial segments and nodes of ranvier. *J Neurosci* **32**: 7232–7243.
- Gella A, Segura M, Durany N, Pfuhlmann B, Stober G, Gawlik M (2011). Is Ankyrin a genetic risk factor for psychiatric phenotypes? *BMC Psychiatry* **11**: 103.
- Geschwind DH, Konopka G (2009). Neuroscience in the era of functional genomics and systems biology. *Nature* **461**: 908–915.
- Giessing C, Thiel CM (2012). Pro-cognitive drug effects modulate functional brain network organization. *Front Behav Neurosci* **6**.
- Goerke R, Meyer-Baese A, Plant C, He H, Emmett MR, Nilsson C et al (2011). Graph clustering techniques applied to the glycomic response in glioblastoma cells to treatments with STAT3 phosphorylation inhibition and fetal bovine serum. *Evol Bio-inspired Comput Theory Appl V* **8059**.
- Goerke R, Meyer-Baese A, Wagner D, He H, Emmett M, Conrad C (2010). Determining and interpreting correlations in lipidomic networks found in glioblastoma cells. *BMC Systems Biol* **4**: 126. **This paper sets the stage for all subsequent mathematical computational advances for the analysis of large systems biology data sets.**
- Goetz R, Dover K, Laezza F, Shtraizent N, Huang X, Tchetchik D et al (2009). Crystal structure of a fibroblast growth factor homologous factor (FHF) defines a conserved surface on FHF's for binding and modulation of voltage-gated sodium channels. *J Biol Chem* **284**: 17883–17896.
- Griffin JL, Kauppinen RA (2007). A metabolomics perspective of human brain tumors. *FEBS J* **274**: 1132–1139.
- Griswold AJ, Ma D, Sacharow SJ, Robinson JL, Jaworski JM, Wright HH et al (2011). A de novo 1.5 Mb microdeletion on chromosome 14q23.2–23.3 in a patient with autism and spherocytosis. *Autism Res* **4**: 221–227.
- Grubb MS, Burrone J (2010a). Activity-dependent relocation of the axon initial segment fine-tunes neuronal excitability. *Nature* **465**: 1070–1074.
- Grubb MS, Burrone J (2010b). Building and maintaining the axon initial segment. *Curr Opin Neurobiol* **20**: 481–488.
- Guckenberger M, Kurz S, Aepinus C, Theiss S, Haller S, Leimbach T et al (2002). Analysis of the heat shock response of *Neisseria meningitidis* with cDNA- and oligonucleotide-based DNA microarrays. *J Bacteriol* **184**: 2546–2551.
- Hakansson K, Cooper HJ, Emmett MR, Costello CE, Marshall AG, Nilsson CL (2001). Electron capture dissociation and infrared multiphoton dissociation MS/MS of an N-glycosylated tryptic peptide to yield complementary sequence information. *Anal Chem* **73**: 4530–4536. **Manuscript illustrates the need for multiple fragmentation techniques to identify the site of glycosylation on a biological protein and subsequent structural information on the carbohydrate.**
- Hamilton J (1994). *Time Series Analysis*. Princeton University Press: Princeton, NJ.
- Hanada M, Sugiura Y, Shinjo R, Masaki N, Imagama S, Ishiguro N et al (2012). Spatiotemporal alteration of phospholipids and prostaglandins in a rat model of spinal cord injury. *Anal Bioanal Chem* **403**: 1873–1884.
- Hart GW, Slawson C, Ramirez-Correa G, Lagerlof O (2011). Cross talk between O-GlcNAcylation and phosphorylation: Roles in signaling, transcription, and chronic disease. *Ann Rev Biochem* **80**: 825–858.
- Hawkins RD, Hon GC, Ren B (2010). Next-generation genomics: an integrative approach. *Nat Rev Genet* **11**: 476–486.
- He F, Emmett MR, Hakansson K, Hendrickson CL, Marshall AG (2004). Theoretical and experimental prospects for protein identification based solely on accurate mass measurement. *J Proteome Res* **3**: 61–67.
- He H, Conrad CA, Nilsson CL, Ji Y, Schaub TM, Marshall AG et al (2007). Method for lipidomic analysis: p53 expression modulates sulfatide, ganglioside and phospholipid composition of U87 MG glioblastoma cells. *Anal Chem* **79**: 8423–8430. **This paper describes methodology for extraction of, on-line chromatographic separation of, efficient microelectrospray ionization of and high-resolution FT-ICR MS analysis of full length polar lipids at biological concentrations.**
- He H, Nilsson CL, Emmett MR, Ji Y, Marshall AG, Kroes RA et al (2010a). Polar lipid remodeling and increased sulfatide expression are associated with the glioma therapeutic candidates, wild type p53 elevation and the topoisomerase-1 inhibitor, Irinotecan. *Glycoconj J* **27**: 27–38. **First description and application of an integrated glyco-transcriptomic and lipidomic approach to the study of glioblastoma.**
- He H, Nilsson CL, Emmett MR, Marshall AG, Kroes RA, Moskal JR et al (2010b). Glycomic and transcriptomic response of GSC11 glioblastoma stem cells to STAT3 phosphorylation inhibition and serum-induced differentiation. *J Proteome Res* **9**: 2098–2108.
- Hecker M, Lambeck S, Toepfer S, van Someren E, Guthke R (2009). Gene regulatory network inference: data integration in dynamic models—a review. *Biosystems* **96**: 86–103.
- Holsboer F (2008). How can we realize the promise of personalized antidepressant medicines? *Nat Rev Neurosci* **9**: 638–646.

- Hosack D, Dennis G, Sherman BT, Lane HC, Lempicki RA (2003). Identifying biological themes within lists of genes with EASE. *Genome Biol* **4**: R70.
- Huang DW, Sherman BT, Lempicki RA (2008). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57.
- Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW *et al* (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* **19**: 342–347.
- Hughey CA, Hendrickson CL, Rodgers RP, Marshall AG, Qian K (2001). Kendrick mass defect spectroscopy: a compact visual analysis for ultrahigh-resolution broadband mass spectra. *Anal Chem* **73**: 4676–4681.
- Hund TJ, Koval OM, Li J, Wright PJ, Qian L, Snyder JS *et al* (2010). A beta(IV)-spectrin/CaMKII signaling complex is essential for membrane excitability in mice. *J Clin Invest* **120**: 3508–3519.
- Janitz M (2008). *Next-Generation Genome Sequencing – Towards Personalized Medicine*. Wiley-VCH: Weinheim.
- Järvinen A-K, Hautaniemi S, Edgren H, Auvinen P, Saarela J, Kallioniemi O-P *et al* (2004). Are data from different gene expression microarray platforms comparable? *Genomics* **83**: 1164–1168.
- Jenkins SM, Bennett V (2001). Ankyrin-G coordinates assembly of the spectrin-based membrane skeleton, voltage-gated sodium channels, and L1 CAMs at Purkinje neuron initial segments. *J Cell Biol* **155**: 739–746.
- Kane MD, Jatke TA, Stumpf CR, Lu J, Thomas JD, Madore SJ (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res* **28**: 4552–4557.
- Karas M, Hillenkamp F (1988). Laser desorption of proteins with molecular masses exceeding 10 000 Daltons. *Anal Chem* **60**: 2299–2301.
- Komada M, Soriano P (2002). [Beta]IV-spectrin regulates sodium channel clustering through ankyrin-G at axon initial segments and nodes of Ranvier. *J Cell Biol* **156**: 337–348.
- Kosaka T, Komada M, Kosaka K (2008). Sodium channel cluster, betaIV-spectrin and ankyrin G positive 'hot spots' on dendritic segments of parvalbumin-containing neurons and some other neurons in the mouse and rat main olfactory bulbs. *Neurosci Res* **62**: 176–186.
- Kotaleski JH, Blackwell KT (2010). Modelling the molecular mechanisms of synaptic plasticity using systems biology approaches. *Nat Rev Neurosci* **11**: 239–251.
- Kothapalli R, Yoder S, Mane SM, Loughran T (2002). Microarray results: how accurate are they? *BMC Bioinformatics* **3**: 22.
- Kroes RA, Burgdorf J, Otto NJ, Panksepp J, Moskal JR (2007a). Social defeat, a paradigm of depression in rats that elicits 22-kHz vocalizations, preferentially activates the cholinergic signaling pathway in the periaqueductal gray. *Behav Brain Res* **182**: 290–300.
- Kroes RA, Dawson G, Moskal JR (2007b). Focused microarray analysis of glyco-gene expression in human glioblastomas. *J Neurochem* **103**: 14–24.
- Kroes RA, He H, Emmett MR, Nilsson CL, Leach FE, Amster IJ *et al* (2010). Overexpression of ST6GalNAcV, a ganglioside-specific α 2,6-sialyltransferase, inhibits glioma growth in vivo. *Proc Natl Acad Sci USA* **107**: 12646–12651.
- Kroes RA, Panksepp J, Burgdorf J, Otto NJ, Moskal JR (2006). Modeling depression: social dominance-submission gene expression patterns in rat neocortex. *Neuroscience* **137**: 37–49.
- Kuo WP, Jenssen T-K, Butte AJ, Ohno-Machado L, Kohane IS (2002). Analysis of matched mRNA measurements from two different microarray technologies. *Bioinformatics* **18**: 405–412.
- Laezza F, Gerber BR, Lou JY, Kozel MA, Hartman H, Craig AM *et al* (2007). The FGF14(F145S) mutation disrupts the interaction of FGF14 with voltage-gated Na⁺ channels and impairs neuronal excitability. *J Neurosci* **27**: 12033–12044.
- Through the molecular analysis of a naturally occurring rare mutation in the human FGF14 gene, this pioneering study provides evidence for the intracellular FGF14 as a multifaceted regulatory protein of Nav channel kinetics and gating, and expression at the AIS.**
- Laezza F, Lampert A, Kozel MA, Gerber BR, Rush AM, Nerbonne JM *et al* (2009). FGF14 N-terminal splice variants differentially modulate Nav1.2 and Nav1.6-encoded sodium channels. *Mol Cell Neurosci* **42**: 90–101.
- Lai HC, Jan LY (2006). The distribution and targeting of neuronal voltage-gated ion channels. *Nat Rev Neurosci* **7**: 548–562.
- Laine RA (1997). *Glycosciences, Status and Perspectives and Bilinear Control Processes* Gabius HJ, Gabius S (eds). Chapman & Hall: London, pp 1–14.
- Lee HS, Wang J, Tian L, Jiang H, Black MA, Madlung A *et al* (2004). Sensitivity of 70-mer oligonucleotides and cDNAs for microarray analysis of gene expression in Arabidopsis and its related species. *Plant Biotech J* **2**: 45–57.
- Lee J, Bussey K, Gwady F, Reinhold W, Riddick G, Pelletier S *et al* (2003). Comparing cDNA and oligonucleotide array data: concordance of gene expression across platforms for the NCI-60 cancer cells. *Genome Biol* **4**: R82.
- Lee KW, Woon PS, Teo YY, Sim K (2012). Genome wide association studies (GWAS) and copy number variation (CNV) studies of the major psychoses: what have we learnt? *Neurosci Biobehav Rev* **36**: 556–571.
- Lenburg M, Liou L, Gerry N, Frampton G, Cohen H, Christman M (2003). Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray data. *BMC Cancer* **3**: 31.
- Lett TA, Zai CC, Tiwari AK, Shaikh SA, Likhodi O, Kennedy J, LandMuller DJ (2011). ANK3, CACNA1C and ZNF804A gene variants in bipolar disorders and psychosis subphenotype. *World J Biol Psychiatry* **12**: 392–397.
- Leussis MP, Madison JM, Petryshen TL (2012). Ankyrin 3: genetic association with bipolar disorder and relevance to disease pathophysiology. *Biol Mood Anxiety Disord* **2**: 18. **This article revises the current evidence for genetic association of ankyrin-G with bipolar disorder, highlighting the new emerging role of the molecular components of the AIS in psychiatric disorders.**
- Leverly SB (2005). Glycosphingolipid Structural Analysis and Glycosphingolipidomics. *Methods Enzymol* **450**: 300–369.
- Li J, Pankratz M, Johnson JA (2002). Differential gene expression patterns revealed by oligonucleotide versus long cDNA arrays. *Toxicol Sci* **69**: 383–390.
- Linding R, Jensen LJ, Pasculescu A, Olhovskiy M, Colwill K, Bork P *et al* (2008). NetworkKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res* **36**(suppl 1): D695–D699.
- Little DP, Speir JP, Senko MW, O'Connor PB, McLafferty FW (1994). Infrared multiphoton dissociation of large multiply-charged ions for biomolecule sequencing. *Anal Chem* **66**: 2809–2815.
- Lou JY, Laezza F, Gerber BR, Xiao M, Yamada KA, Hartmann H *et al* (2005). Fibroblast growth factor 14 is an intracellular modulator of voltage-gated sodium channels. *J Physiol* **569**: 179–193.
- Mah N, Thelin A, Nikolaus S, Kuhbacher T, Gurbuz Y, Eickhoff H *et al* (2004). A comparison of oligonucleotide and cDNA-based microarray systems. *Physiol Genomics* **16**: 361–370.
- Malone J, Oliver B (2011). Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* **9**: 34.
- Marshall AG, Hendrickson CL (2008). High-resolution mass spectrometers. *Annu Rev Anal Chem (Palo Alto Calif)* **1**: 579–599.
- Marshall AG, Hendrickson CL, Jackson GS (1998). Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrom Rev* **17**: 1–35. **An excellent review of the theory, operation, and applications of high-resolution, high mass accuracy, Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS).**
- Martin JA, Wang Z (2011). Next-generation transcriptome assembly. *Nat Rev Genet* **12**: 671–682.
- Mattsson N, Zegers I, Andreasson U, Bjerke M, Blankenstein MA, Bowser R *et al* (2012). Reference measurement procedures for Alzheimer's disease cerebrospinal fluid biomarkers: definitions and approaches with focus on amyloid β 42. *Biomark Med* **6**: 409–417.
- Melas IN, Mitsos A, Messinis DE, Weiss TS, Rodriguez J-S, Alexopoulos LG (2012). Construction of large signaling pathways using an adaptive perturbation approach with phosphoproteomic data. *Mol Biosystems* **8**: 1571–1584.
- Meyer-Baeza A, Guthke R, Altwasser R, Vlaic S, Emmett MR, Conrad CA (2012a). Model order reduction of deterministic and stochastic gene regulatory networks. *BIOCOMP'12-Proc. 13th Int. Conf. Bioinformatics and Computational Biology*, Las Vegas, USA, 16–19 July 2012.
- Meyer-Baeza A, Guthke R, Altwasser R, Vlaic S, Emmett MR, Conrad CA (2012b). Almost sure stability of stochastic gene regulatory networks with mode-dependent interval delays. *BIOCOMP'12-Proc. 13th Int. Conf. Bioinformatics and Computational Biology*, Las Vegas, USA, 16–19 July 2012.
- Michalewski I, Segal-Ruder Y, Rozenbaum M, Medzihradzsky KF, Shalem O, Coppola G *et al* (2010). Signaling to transcription networks in the neuronal retrograde injury response. *Sci Signal* **3**: ra53.
- Microarray Consortium (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotech* **24**: 1151–1161.
- Miro X, Meier S, Dreisow ML, Frank J, Strohmaier J, Breuer R *et al* (2012). Studies in humans and mice implicate neurocan in the etiology of mania. *Am. J. Psychiatry* **169**: 982–990.
- Miura D, Tsuji Y, Takahashi K, Wariishi H, Saito K (2010). A strategy for the determination of the elemental composition by Fourier transform ion cyclotron resonance mass spectrometry based on isotopic peak ratios. *Anal. Chem* **82**: 5887–5891.
- Mohler RR (1973). *Bilinear Control Processes*. Academic Press: New York.
- Mohler RE, Dombek KM, Hoggard JC, Young ET, Synovec RE (2006). Comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry analysis of metabolites in fermenting and respiring yeast cells. *Anal Chem* **78**: 2700–2709.
- Morawski M, Brückner G, Jäger C, Seeger G, Matthews RT, Arendt T (2012). Involvement of perineuronal and perisynaptic extracellular matrix in Alzheimer's disease neuropathology. *Brain Pathol* **22**: 547–561.
- Moskal JR, Burgdorf J, Kroes RA, Brudzynski SM, Panksepp J (2011). A novel NMDA receptor glycine-site partial agonist, GLYX-13, has

- therapeutic potential for the treatment of autism. *Neurosci Biobehav Rev* **35**: 1982–1988.
- Moskal JR, Kroes RA, Dawson G (2009). The glycochemistry of brain tumors: disease relevance and therapeutic potential. *Expert Rev Neurother* **9**: 1529–1545.
- Motta A, Paris D, Melch D (2010). Monitoring real-time metabolism of living cells by fast two-dimensional NMR spectroscopy. *Anal Chem* **82**: 2405–2411.
- Mueller N, Krumsiek J, Boehm C, Theis F, Meyer-Baese A (2011). Gaussian graphical modeling reveals specific lipid correlations in glioblastoma cells. *SPIE Symp Comput Intell*.
- Mulley JC, Mefford HC (2011). Epilepsy and the new cytogenetics. *Epilepsia* **52**: 423–432.
- Mulley JC, Scheffer IE, Petrou S, Berkovic SF (2003). Channelopathies as a genetic cause of epilepsy. *Curr Opin Neurol* **16**: 171–176.
- Mussche S, De Paepe B, Smet J, Devreese K, Lissens W, Rasic VM *et al* (2012). Proteomic analysis in giant axonal neuropathy: New insights into disease mechanisms. *Muscle Nerve* **46**: 246–256.
- Need AC, Ge D, Weale ME, Maia J, Feng S, Heinzen EL *et al* (2009). A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* **5**: e1000373.
- Nilsson CL (2012). Advances in quantitative phosphoproteomics. *Anal Chem* **9**: 735–746.
- Nilsson CL, Dillon R, Devakumar A, Rogers JC, Krastins B, Rosenblatt MM *et al* (2010). Quantitative phosphoproteomic analysis of STAT3/IL-6/HIF1 α signaling network: an initial study in GSC11 glioblastoma stem cells. *J Proteome Res* **9**: 430–443.
- Oberg AL, Vitek O (2009). Statistical design of quantitative mass spectrometry-based proteomic experiments. *J Proteome Res* **8**: 2144–2156.
- Ogawa Y, Rasband MN (2008). The functional organization and assembly of the axon initial segment. *Curr Opin Neurobiol* **18**: 307–313.
- Oh SF, Vickery TW, Serhan CN (2011). Chiral lipidomics of E-series resolvins: Aspirin and the biosynthesis of novel mediators. *Biochim Biophys Acta* **11**: 737–747.
- Ono M, Handa K, Sonnino S, Withers DA, Nagai H, Hakomori S (2001). GM3 ganglioside inhibits CD9-facilitated haptotactic cell motility: coexpression of GM3 and CD9 is essential in the downregulation of tumor cell motility and malignancy. *Biochemistry* **40**: 6414–6421.
- Oshlack A, Robinson M, Young M (2010). From RNA-seq reads to differential expression results. *Genome Biol* **11**: 220.
- Paananen J, Wong G (2009). FOR3D: Force-directed 3D graph editor for visualization of integrated genome scale data. *BMC Syst Biol* **3**: 26.
- Paik Y-K, Hancock WS (2012). Uniting ENCODE with genome-wide proteomics. *Nat Biotechnol* **30**: 1065–1067.
- Parmigiani G, Garrett-Mayer ES, Anbazhagan R, Gabrielson E (2004). A cross-study comparison of gene expression studies for the molecular classification of lung cancer. *Clin Cancer Res* **10**: 2922–2927.
- Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, Collins PJ, Chu T-M, Bao W *et al* (2006). Performance comparison of one-color and two-color platforms within the Microarray Quality Control (MAQC) project. *Nat Biotech* **24**: 1140–1150.
- Pepke S, Wold B, Mortazavi A (2009). Computation for ChIP-seq and RNA-seq studies. *Nat Meth* **6**: S22–S32.
- Pérez-Manso M, Barroso-Chinea P, Aymerich MS, Lanciego JL (2006). 'Functional' neuroanatomical tract tracing: analysis of changes in gene expression of brain circuits of interest. *Brain Res* **1072**: 91–98.
- Petersen D, Chandramouli GVR, Geoghegan J, Hilburn J, Paarlberg J, Kim C *et al* (2005). Three microarray platforms: an analysis of their concordance in profiling gene expression. *BMC Genomics* **6**: 63.
- Piomelli D (2005). The challenge of brain lipidomics. *Prostaglandins Other Lipid Mediat* **77**: 23–34.
- Piomelli D, Astarita G, Rapaka R (2007). A neuroscientist's guide to lipidomics. *Nat Rev Neurosci* **8**: 743–754.
- Pitteri S, Hanash S (2010). A systems approach to the proteomic identification of novel cancer biomarkers. *Dis Markers* **28**: 233–239.
- Ralston-Hooper K, Jannasch A, Ademec J, Sepulveda M (2011). The use of two-dimensional gas chromatography-time-of-flight mass spectrometry (GC x GC-TOF-MS) for metabolomic analysis of polar metabolites. *Methods Mol Biol* **708**: 205–211.
- Robinson SJ (2012). Childhood epilepsy and autism spectrum disorders: psychiatric problems, phenotypic expression, and anticonvulsants. *Neuropsychol Rev* **22**: 271–279.
- Rogojina AT, Orr WE, Song BK, Geisert EEJ (2003). Comparing the use of Affymetrix to spotted oligonucleotide microarrays using two retinal pigment epithelium cell lines. *Mol Vis* **9**: 482–496.
- Roussos P, Giakoumakis SG, Georgakopoulos A, Robakis NK, Bitsios P (2011). The CACNA1C and ANK3 risk alleles impact on affective personality traits and startle reactivity but not on cognition or gating in healthy males. *Bipolar Disord* **13**: 250–259. **A combined transcriptomic-proteomic study reveals new molecular abnormalities associated with schizophrenia.**
- Rubakhin SS, Romanova EV, Nemes P, Sweedler JV (2011). Profiling metabolites and peptides in single cells. *Nat Meth* **8**: S20–S29.
- Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ *et al* (2012). De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**: 237–241. **A phenominal whole-exome sequencing analysis of 928 individuals identifies two independent nonsense variants disrupting the gene SCN2A, coding for the Nav1.2 channel, as a potential mechanistic link to autism.**
- Schaub TM, Blakney GT, Hendrickson CL, Quinn JP, Senko MW, Marshall AG (2007). *LC/MS, Proteins, and Petroleum: Performance Characteristics of a 14.5 Tesla LTQ FT-ICR Mass Spectrometer*. ASMS: Indianapolis, IN.
- Schaub TM, Hendrickson CL, Horning S, Quinn JP, Senko MW, Marshall AG (2008). High-performance mass spectrometry: fourier transform ion cyclotron resonance at 14.5 Tesla. *Anal Chem* **80**: 3985–3990.
- Schechter I (1970). Mapping of the combining sites of antibodies specific for poly-L-alanine determinants. *Nature* **228**: 639–641.
- Scherpen JMA (1993). Balancing for nonlinear systems. *Syst Control Lett* **21**: 143–153.
- Seeley EH, Schwamborn K, Caprioli RM (2011). Imaging of intact tissue sections: moving beyond the microscope. *J Biol Chem* **286**: 25459–25466.
- Senko MW, Hendrickson CL, Emmett MR, SD-H Shi, Marshall AG (1997). External accumulation of ions for enhanced electro-spray ionization Fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* **8**: 970–976.
- Shavkunov A, Panova N, Prasai A, Veselenak R, Bourne N, Stoilova-McPhie SandLaezza F (2012). Bioluminescence methodology for the detection of protein-protein interactions within the voltage-gated sodium channel macromolecular complex. *Assay Drug Dev Technol* **10**: 148–160. **The paper introduces a new methodology relying on the recombination of split-luciferase fragments as an in-cell read-out of protein-protein interactions in channelomics complexes; such method could be employed as a scalable large-scale assay for validation of genomic screens.**
- Shi SD-H, Hendrickson CL, Marshall AG (1998). Counting individual sulfur atoms in a protein by ultrahigh-resolution Fourier transform ion cyclotron resonance mass spectrometry: experimental resolution of isotopic fine structure in proteins. *Proc Natl Acad Sci USA* **95**: 11532–11537.
- Sihlbom C, Davidsson P, Emmett MR, Marshall AG, Nilsson CL (2004). Glycoproteomics of cerebrospinal fluid in neurodegenerative disease. *Int J Mass Spectrom* **234**: 145–152. **The paper describes the application of FT-ICR MS for glycoprotein analysis of CSF from Alzheimer's patients. This is article demonstrates the feasibility of glycoprotein analysis at true biological concentrations from patient samples.**
- Sihlbom C, Davidsson P, Nilsson CL (2005). Prefractionation of cerebrospinal fluid to enhance glycoprotein concentration prior to structural determination with FT-ICR mass spectrometry. *J Proteome Res* **4**: 2294–2301.
- Singh NA, Charlier C, Stauffer D, DuPont BR, Leach RJ, Melis R *et al* (1998). A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. *Nat Genet* **18**: 25–29.
- Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W *et al* (2009). Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* **14**: 755–763.
- Sparvero LJ, Amoscato AA, Kochanek PM, Pitt BR, Kagan VE, Bayir H (2010). Mass-spectrometry based oxidative lipidomics and lipid imaging: applications in traumatic brain injury. *J Neurochem* **115**: 1322–1336.
- Stam CJ, Hillebrand A, Wang H, Van Mieghem P (2010). Emergence of modular structure in a large-scale brain network with interactions between dynamics and connectivity. *Front Comput Neurosci* **4**.
- Starosom SC, Mascanfroni ID, Imitola J, Cao L, Raddassi K, Hernandez SF *et al* (2012). Galectin-1 deactivates classically activated microglia and protects from inflammation-induced neurodegeneration. *Immunity* **37**: 249–263.
- Steuer R (2006). On the analysis and interpretation of correlations in metabolomic data. *Briefings Bioinform* **7**: 151–158.
- Steuer R, Kurths J, Fiehn O, Weckwerth W (2003). Observing and interpreting correlations in metabolomic networks. *Bioinformatics* **19**: 1019–1026.
- Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP (2007). GSEA-P: a desktop application for gene set enrichment analysis. *Bioinformatics* **23**: 3251–3253.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert L, Gillette MA *et al* (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* **102**: 15545–15550.
- Svennerholm L, Brane G, Karlsson I, Lekman A, Ramstrom I, Wikkelso C (2002). Alzheimer disease-effect of continuous intracerebroventricular treatment with

- GM1 ganglioside and a systematic activation programme. *Dement Geriatr Cogn Disord* **14**: 128–136.
- Tabb DL (2013). Quality assessment for clinical proteomics. *Clinical Biochem* **46**: 411–420.
- Tan PK, Downey TJ, Spitznagel EL Jr, Xu P, Fu D, Dimitrov DS *et al* (2003). Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res* **31**: 5676–5684.
- Taniguchi M, Miura K, Iwao H, Yamanaka S (2001). Quantitative assessment of DNA microarrays—comparison with northern blot analyses. *Genomics* **71**: 34–39.
- Taylor LT (2008). Supercritical fluid chromatography. *Anal. Chem.* **80**: 4285–4294.
- Tesli M, Koefoed P, Athanasiu L, Mattingsdal M, Gustafsson O, Agartz I *et al* (2011). Association analysis of ANK3 gene variants in nordic bipolar disorder and schizophrenia case-control samples. *Am J Med Genet B Neuropsychiatr Genet* **156B**: 969–974.
- Thompson AJ, Abu M, Hanger DP (2012). Key issues in the acquisition and analysis of qualitative and quantitative mass spectrometry data for peptide-centric proteomic experiments. *Amino Acids* **43**: 1075–1085.
- Thompson KL, Afshari CA, Amin RP, Bertram TA, Car B, Cunningham M *et al* (2004). Identification of platform-independent gene expression markers of cisplatin nephrotoxicity. *Environ Health Perspect* **112**: 488–494.
- Trompet S, Jukema W, Mooijaart SP, Ford I, Stott DJ, Westendorp RGJ *et al* (2012). Genetic variation in galectin-3 gene associates with cognitive function at old age. *Neurobiol Aging* **33**: 2232.e1–2239.e2.
- Ulrich RG, Rockett JC, Gibson GG, Pettitt SD (2004). Overview of an interlaboratory collaboration on evaluating the effects of model hepatotoxicants on hepatic gene expression. *Environ Health Perspect* **112**: 423–427.
- Van Swieten JC, Brusse E, De Graaf BM, Krieger E, Van De Graaf R, De Koning I *et al* (2003). A mutation in the fibroblast growth factor 14 gene is associated with autosomal dominant cerebellar ataxia. *Am J Hum Genet* **72**: 191–199.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG *et al* (2001). The sequence of the human genome. *Science* **291**: 1304–1351.
- Villoslada P, Steinman L, Baranzini SE (2009). Systems biology and its application to the understanding of neurological diseases. *Ann. Neurol* **65**: 124–139.
- Vuckovic D, Pawliszyn J (2011). Systematic evaluation of solid-phase microextraction coatings for untargeted metabolomic profiling of biological fluids by liquid chromatography-mass spectrometry. *Anal Chem* **83**: 1944–1954.
- Wang H-Y, Malek R, Kwitek A, Greene A, Luu T, Behbahani B *et al* (2003). Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biol* **4**: R5.
- Wang X, Emmett MR, Marshall AG (2010). Liquid chromatography electrospray ionization fourier transform ion cyclotron resonance mass spectrometric characterization of N-linked glycans and glycopeptides. *Anal Chem* **82**: 6542–6548.
- Wang Z, Gerstein M, Snyder M (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet* **10**: 57–63.
- Weiss LA, Escayg A, Kearney JA, Trudeau M, MacDonald BT, Mori M *et al* (2003). Sodium channels SCN1A, SCN2A and SCN3A in familial autism. *Mol Psychiatry* **8**: 186–194.
- Wibom C, Surowiec I, Mören L, Bergström P, Johansson M, Antti H *et al* (2010). Metabolomic patterns in glioblastoma and changes during radiotherapy: A clinical microdialysis study. *J Proteome Res* **9**: 2909–2919.
- Wilm MS, Mann M (1994). Electrospray and Taylor-cone theory: Dole's beam of macromolecules at last? *Int J Mass Spectrom Ion Proc* **136**: 167–180. **This low-flow electrospray technique revolutionized biological mass spectrometry analysis.**
- Wong EH, Fox JC, Ng MY, Lee CM (2011). Toward personalized medicine in the neuropsychiatric field. *Int Rev Neurobiol* **101**: 329–349.
- Wood PL (2013). Mass spectrometry strategies for clinical metabolomics and lipidomics in psychiatry, neurology, and neuro-oncology. *Neuropsychol* 11 July 2013. doi:10.1038/npp.2013.167 [E-pub ahead of print].
- Yamaji-Hasegawa A, Tsujimoto M (2006). Asymmetric distribution of phospholipids in biomembranes. *Biol Pharm Bull* **29**: 1547–1553.
- You Q, Fang S, Chen JY (2010). GeneTerrain: Visual exploration of differential gene expression profiles organized in native biomolecular interaction networks. *Inform Visualization* **9**: 1–12.
- Yuen T, Wurmbach E, Pfeffer RL, Ebersole BJ, Sealfon SC (2002). Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res* **30**: e48.
- Zeeberg B, Feng W, Wang G, Wang M, Fojo A, Sunshine M *et al* (2003). GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* **4**: R28.
- Zhang X, Bennett V (1998). Restriction of 480/270-kD ankyrin G to axon proximal segments requires multiple ankyrin G-specific domains. *J Cell Biol* **142**: 1571–1581.
- Zhang Y, Fillou MD, Reckow S, Gormanns P, Maccarrone G, Kessler MS *et al* (2011). Proteomic and metabolomic profiling of a trait anxiety mouse model implicate affected pathways. *Mol Cell Proteomics* **10**: 12.
- Zubarev RA, Kelleher NL, McLafferty FW (1998). Electron capture dissociation of multiply charged protein cations. A nonergodic process. *J Am Chem Soc* **120**: 3265–3266.