JOURNAL OF NEUROTRAUMA 30:1101–1116 (July 1, 2013)

DOI: 10.1089/neu.2012.2631

Review

Systems Biology Approaches for Discovering Biomarkers for Traumatic Brain Injury

Jacob D. Feala,[1] Mohamed Diwan M. AbdulHameed,[1] Chenggang Yu,[1] Bhaskar Dutta,[1] Xueping Yu,[1] Kara Schmid,[2] Jitendra Dave,[2] Frank Tortella,[2] and Jaques Reifman[1]

Abstract

The rate of traumatic brain injury (TBI) in service members with wartime injuries has risen rapidly in recent years, and complex, variable links have emerged between TBI and long-term neurological disorders. The multifactorial nature of TBI secondary cellular response has confounded attempts to find cellular biomarkers for its diagnosis and prognosis or for guiding therapy for brain injury. One possibility is to apply emerging systems biology strategies to holistically probe and analyze the complex interweaving molecular pathways and networks that mediate the secondary cellular response through computational models that integrate these diverse data sets. Here, we review available systems biology strategies, databases, and tools. In addition, we describe opportunities for applying this methodology to existing TBI data sets to identify new biomarker candidates and gain insights about the underlying molecular mechanisms of TBI response. As an exemplar, we apply network and pathway analysis to a manually compiled list of 32 protein biomarker candidates from the literature, recover known TBI-related mechanisms, and generate hypothetical new biomarker candidates.

Key words: biomarker; pathway analysis; protein-protein interaction; systems biology; traumatic brain injury

Introduction

The clinical significance and long-term effects of traumatic brain injury (TBI) have garnered great attention in recent years,[1,2] partly as a result of a rapidly increasing population of U.S. warfighters suffering injuries to the head. The overall rate of TBI in service members nearly tripled from 2000 to 2010, driven by a 400% increase in cases of mild TBI (mTBI).[3] In fact, some degree of TBI has been diagnosed in 16% of wounded warfighters returning from Iraq.[4] This widespread, increasing prevalence of brain injuries is of great concern, especially in light of recent evidence that TBI may lead to serious long-term neurological deficits and disease.[1]

Traumatic brain injuries can be classified by severity as mild, moderate, or severe, each of which poses unique medical challenges. mTBI is the most prevalent, representing 77% of military TBI cases in 2011.[3] However, mild cases are frequently undiagnosed because they escape detection by brain imaging, can be overlooked because of more-immediate medical concerns, and can have delayed presentation of symptoms.[5] Moderate and severe cases of TBI are less common and relatively easier to detect, but prognosis of short-term secondary complications or long-term disease progression remains a challenge. Early detection and treatment of TBI may improve outcome[6,7] and help reduce long-term cognitive deficits and occurrence of related neurological diseases.[1]

However, to date, there are no U.S. Food and Drug Administration (FDA)-approved biomarkers for the diagnosis or prognosis of TBI, and the molecular mechanisms of TBI response remain poorly understood. This lack of understanding reflects the complex, multifactorial nature of secondary cellular responses to TBI, which are believed to involve a network of interweaving molecular pathways that mediate cellular response. The emerging field of systems biology attempts to harness complex, multi-gene systems by computationally integrating gene-level data with molecular pathways and networks to extract new biological insight. Systems biology may combine and augment current strategies to biomarker discovery, generating novel, experimentally testable candidates.

Challenges in TBI Biomarker Discovery

Existing TBI biomarker candidates

Molecular biomarkers generally consist of biomolecules measured from biofluids or from the affected tissue that provide diagnostic, prognostic, or therapeutic information.[8] There are several

[1]Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command (MRMC), Fort Detrick, Maryland.
[2]Department of Brain Trauma Neuroprotection and Neurorestoration, Center for Military Psychiatry and Neuroscience, Walter Reed Army Institute of Research, Silver Spring, Maryland.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 JUL 2013		**00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Systems Biology Approaches for Discovering Biomarkers for Traumatic Brain Injury	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
US Army Medical Research and Materiel Command,DoD Biotechnology High Performance Computing Software Applications Institute,Telemedicine and Advanced Technology Research Center,Fort Detrick,MD,21702	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The rate of traumatic brain injury (TBI) in service members with wartime injuries has risen rapidly in recent years, and complex, variable links have emerged between TBI and long-term neurological disorders. The multifactorial nature of TBI secondary cellular response has confounded attempts to find cellular biomarkers for its diagnosis and prognosis or for guiding therapy for brain injury. One possibility is to apply emerging systems biology strategies to holistically probe and analyze the complex interweaving molecular pathways and networks that mediate the secondary cellular response through computational models that integrate these diverse data sets. Here, we review available systems biology strategies, databases and tools. In addition, we describe opportunities for applying this methodology to existing TBI data sets to identify new biomarker candidates and gain insights about the underlying molecular mechanisms of TBI response. As an exemplar we apply network and pathway analysis to a manually compiled list of 32 protein biomarker candidates from the literature recover known TBI-related mechanisms, and generate hypothetical new biomarker candidates.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**16**	

Table 1. Network Properties and Pathway Associations of 32 TBI Biomarker Candidates[a]

Gene symbol(s)	*Gene name*	*Interactions in the PPI network*	*Associated KEGG pathways*
GFAP	Glial fibrillary acidic protein	27	NA
S100B	S100 calcium-binding protein B	20	NA
UCHL1	Ubiquitin carboxyl-terminal esterase L1	27	Parkinson's disease
ENO2, NSE	Enolase 2 (gamma, neuronal)	17	Glycolysis/gluconeogenesis, metabolic pathways, RNA degradation
SPTAN1 (SBDP)[b]	Spectrin, alpha, non-erythrocytic 1 (alpha-fodrin)	59	NA
MBP	Myelin basic protein	48	NA
MAPT, TAU	Microtubule-associated protein tau	54	MAPK-signaling pathway, Alzheimer's disease
FABP7, B-FABP	Fatty-acid–binding protein 7, brain	0	PPAR-signaling pathway
HSPD1, HSP60	Heat shock 60kDa protein 1	43	RNA degradation, type I diabetes mellitus
HSPA4, HSP70	Heat shock 70kDa protein 4	64	Antigen processing and presentation
HMOX1, HO-1	Heme oxygenase (decycling) 1	10	Porphyrin and chlorophyll metabolism, mineral absorption
CYCS, CYC	Cytochrome c, somatic	33	Viral myocarditis, small-cell lung cancer, colorectal cancer, pathways in cancer, toxoplasmosis, Huntington's disease, amyotrophic lateral sclerosis, Parkinson's disease, Alzheimer's disease, apoptosis, p53-signaling pathway
BCL2	B-cell CLL/lymphoma 2	90	Protein processing in endoplasmic reticulum, apoptosis, focal adhesion, neurotrophin signaling pathway, amyotrophic lateral sclerosis, toxoplasmosis, pathways in cancer, colorectal cancer, prostate cancer, small-cell lung cancer
IL6	Interleukin-6 (interferon, beta 2)	5	Cytokine-cytokine receptor interaction, Toll-like receptor-signaling pathway, nucleotide oligomerization domain (NOD)-like receptor signaling pathway, cytosolic DNA-sensing pathway, Jak-STAT-signaling pathway, hematopoietic cell lineage, intestinal immune network for IgA production, prion diseases, Chagas disease (American trypanosomiasis), African trypanosomiasis, malaria, amoebiasis, measles, pathways in cancer, rheumatoid arthritis, graft-versus-host disease, hypertrophic cardiomyopathy
APOE	Apolipoprotein E	16	Alzheimer's disease
APP, ABPP	Amyloid beta (A4) precursor protein	120	Alzheimer's disease
NGF	Nerve growth factor (beta polypeptide)	7	MAPK-signaling pathway, apoptosis, Neurotrophin-signaling pathway
CRP	C-reactive protein, pentraxin-related	17	NA
ADM	Adrenomedullin	4	NA
CP	Ceruloplasmin (ferroxidase)	8	Porphyrin and chlorophyll metabolism
CHI3L1, YKL40	Chitinase 3-like 1 (cartilage glycoprotein-39)	0	Amino sugar and nucleotide sugar metabolism
CASP9	Caspase-9, apoptosis-related cysteine peptidase	40	p53-signaling pathway, apoptosis, vascular endothelial growth factor-signaling pathway, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, toxoplasmosis, pathways in cancer, colorectal cancer, pancreatic cancer, endometrial cancer, prostate cancer, small-cell lung cancer, non-small-cell lung cancer, viral myocarditis
BDKRB1	Bradykinin receptor B1	2	Calcium-signaling pathway, neuroactive ligand-receptor interaction, complement and coagulation cascades, regulation of actin cytoskeleton

(*continued*)

Table 1. (Continued)

Gene symbol(s)	*Gene name*	*Interactions in the PPI network*	*Associated KEGG pathways*
BDKRB2	Bradykinin receptor B2	12	Calcium-signaling pathway, neuroactive ligand-receptor interaction, complement and coagulation cascades, regulation of actin cytoskeleton, endocrine and other factor-regulated calcium reabsorption, Chagas disease (American trypanosomiasis)
BECN1	Beclin-1, autophagy related	7	Regulation of autophagy
BMP6	Bone morphogenetic protein 6	10	Hedgehog-signaling pathway, transforming growth factor-beta-signaling pathway
BDNF	Brain-derived neurotrophic factor	10	MAPK-signaling pathway, neurotrophin-signaling pathway, Huntington's disease
CASP7	Caspase-7, apoptosis-related cysteine peptidase	51	Apoptosis, Alzheimer's disease
AVEN	Apoptosis, caspase activation inhibitor	4	NA
CNTFR	Ciliary neurotrophic factor receptor	13	Cytokine-cytokine receptor interaction, Jak-STAT-signaling pathway
AIMP1, EMAPII	Aminoacyl tRNA synthetase complex-interacting multifunctional protein 1	11	NA
NEFH, NFH	Neurofilament, heavy polypeptide	5	Amyotrophic lateral sclerosis

[a]Ordered by the number of citations that we have collected; see Supplementary Table 1.
[b]SPTAN1 encodes αII-spectrin and αII-spectrin breakdown products (SBDPs), which are considered as TBI biomarkers.
TBI, traumatic brain injury; PPI, protein-protein interaction; KEGG, Kyoto Encyclopedia of Genes and Genomes; NA, not available; MAPK, mitogen-activated protein kinase; PPAR, peroxisome proliferator-activated receptor; Jak-STAT, Janus kinase/signal transducer and activator of transcription; IgA, immunoglobulin A.

successful examples of molecular biomarkers that are currently the clinical standard for diagnostic screening in several diseases, for example, in myocardial infarction[9] and certain cancers,[10] and the search for novel molecular biomarkers continues to be a major research thrust in many biomedical fields. Most new biomarkers proposed in the literature never reach the clinic, however, often because of a lack of reproducibility. In a meta-analysis of highly cited articles announcing new biomarker candidates for a variety of diseases, it was shown that follow-up experiments with greater statistical power generally fail to reproduce the same effect size as the original studies.[11]

TBI has not been entirely immune from such criticism. To date, many candidate molecular biomarkers of TBI have been identified and some are being further investigated in ongoing clinical studies, but none are in clinical use in the United States.[2] An ideal biomarker would always be present in biofluids in cases of TBI (sensitivity), would never be present in its absence (specificity), and would provide prognostic information on secondary complications that are important factors of clinical outcome. This would include severity level, ischemic versus traumatic nature of injury, intracranial pressure levels, and status of the blood–brain barrier. Though some candidate biomarkers can predict clinical outcome with either high sensitivity or high specificity in severe TBI (sTBI), the challenge is to be able to display both in a clinical evaluation. S100B is a case in point. S100B has been one of the most extensively studied biomarkers,[12] which, though not approved in the United States, is currently being used in Europe as a screening tool because of its high sensitivity.[13] However, S100B is not unique to the nervous system because it can rise in response to other traumas in the absence of brain injury.[14–16] Because of its low specificity for brain injury, its diagnostic value for military-relevant TBI (where polytrauma is likely) is constrained, and in civilian TBI its value as a clinical diagnostic tool is limited to its high sensitivity for computed tomography (CT)-positive injuries.[13,17,18] As another example, postinjury cerebral spinal fluid levels of the protein Tau (official gene symbol, MAPT) have been shown to predict clinical outcome and intracranial pressure for sTBI with high sensitivity and specificity,[19,20] but have large standard deviations[19] and show no significant changes during mTBI.[21]

However, significant progress has been made toward identifying TBI biomarkers and developing antibodies (Abs) and assays with the required sensitivity to yield clinically meaningful, FDA-acceptable guidelines. More recently, the results of several clinical studies in mild-to-severe TBI patients have emerged in support of previous preclinical research efforts,[22,23] including the glial marker, GFAP (glial fibrillary acidic protein), and the neuronal marker, UCHL1 (ubiquitin carboxy-terminal hydrolase L1). GFAP is a monomeric intermediate filament protein that is mainly expressed by astrocytes in the central nervous system (CNS). Though an early study showed high sensitivity (85%), but only moderate specificity (<60%), for serum GFAP in predicting the outcome of sTBI patients,[24] more-recent studies observed significantly higher specificity (93%) and sensitivity (71%).[12,25] In addition, another recent study showed strong association between levels of serum GFAP breakdown products and CT-detectable lesions for mild and moderate TBI,[26] suggesting that GFAP could also serve as a potential marker for less-severe brain injury. Unlike GFAP, which is highly abundant in glial cells, UCHL1 is highly abundant in neuronal cells and is involved in enzymatic ubiquitination and deubiquitination processes of metabolic pathways. Recent clinical studies have shown that, for sTBI, the concentration of UCHL1 is significantly elevated in both cerebrospinal fluid and serum[27–29] and that the use of UCHL1 serum level as a predictor of in-hospital mortality of patients with sTBI yields a 96% specificity and a 52% sensitivity.[25]

We have compiled these and other TBI biomarker candidates from the literature into a list of 32 proteins (Table 1), to which we will refer throughout this article. Molecular information and clinical findings for this list are summarized in Supplementary Table 1 (see online supplementary material at http://www.liebertpub.com).

Although these biomarker candidates have been heavily studied, much remains unknown about how changes in their expression levels relate to mechanisms of injury and clinical outcome. Molecular-level responses to injury are linked to clinical outcomes through poorly understood cascades of interacting pathways, and thus one-to-one relationships between genes and TBI phenotypes are unlikely. Therefore, the current thinking is that there may not be an ideal single biomarker, but rather that a panel or signature of markers may provide more-accurate information about injury status and clinical outcome.[2,8,30] Along these lines, Mondello and colleagues recently investigated the use of the ratio between GFAP and UCHL1 as a differential indicator of TBI.[31] However, given the high dimensionality of the search space for biomarker discovery, the identification of ideal combinations of multiple biomarkers requires a systematic, systems-level approach that is inherently capable of discovering multidimensional signatures from complex molecular interactions.

The complex, system-wide consequences of TBI hinders biomarker discovery

TBI is composed of ''primary'' and ''secondary'' injury components, but it is the multi-cellular, heterogeneous nature of the secondary injury that makes predicting outcomes and designing therapies for TBI exceedingly difficult.[32] The primary insult can be focal damage, resulting from contact injury, or diffuse axonal damage.[1,33] The tissue then undergoes secondary injury, a complex series of biochemical events to mediate cell damage evolving over hours to weeks after the initial trauma. These secondary events are often more damaging and can lead to tissue-level pathologies, such as ischemia, apoptosis cascades, increased intracranial pressure, and inflammation.[34]

Tissue-level secondary injuries emerge from imbalances at the neuron level. Early stages of injury lead to altered cellular metabolism and ''ischemia-like'' activity of the anaerobic glycolysis pathway.[33] The resulting adenosine triphospahte imbalance causes energy-dependent ion pumps to fail, depolarizing the neural membrane and causing an influx of calcium and sodium, release of neurotransmitters (i.e., excitotoxicity), and initiation of catabolic processes. This early disruption of metabolic pathways triggers the release of reactive oxygen species, activating apoptotic death pathways.[33] Inflammation is also a prominent feature of TBI,[33] adding a multi-cellular layer of complexity to the mechanisms of secondary injury.

The phenotypic effects of secondary brain injury emerge through a currently intractable, not well-understood multi-cellular system involving hundreds of interacting molecular components. Conversely, traditional research approaches require some tractable conceptual model of the system of interest to transform observations into hypotheses. As a result, it is difficult to generate hypotheses for TBI biomarker candidates from these large, complex systems. Systems biology helps distill unmanageably complex biological phenomena into experimentally testable hypotheses using computational methods[35] and may overcome limitations in current approaches for biomarker discovery.

Current methods for discovering TBI biomarkers

Noorbakhsh and colleagues categorize current methods for biomarker discovery into two main approaches: ''top-down'' and ''bottom-up'' methods.[36] The most commonly used method for discovering new molecular biomarkers has been by the top-down method, in which conceptual models of disease mechanisms and observed biological interactions are mentally combined to construct new hypotheses. Hypothetical markers are then tested by applying molecular biology methods to model organisms or clinical samples. This approach can lead to experimental bias, favoring the further study of already well-known systems, and can overlook the involvement of important biological mechanisms outside the realm of current knowledge. The method is also ''low throughput,'' in that only a few hypotheses can be tested at a time, by time-consuming methods. Most of the biomarker candidates listed in Table 1 were discovered using such a top-down method.

In contrast, the bottom-up method[36] is unbiased, using high-throughput omics technologies to attempt to quantify all biomolecules of a given type within a cell or tissue. Generally, the top differentially expressed biomolecules discovered in a high-throughput data set are proposed as biomarker candidates. This approach, however, usually results in overwhelmingly large lists of candidate genes or proteins, which makes interpretation and hypothesis generation difficult. The maturation and widespread use of these technologies, which can include complementary DNA (cDNA) or oligonucleotide microarrays, proteomics, and metabolomics, has resulted in many such bottom-up studies. The sole example from our biomarker candidate list in Table 1 identified by such a bottom-up method, EMAPII, emerged from proteomics in injured rat brain tissue[37] and was later validated in cerebral spinal fluid and plasma.[38]

Both top-down and bottom-up methods have inherent limitations. Top-down methods are inefficient for exploring the thousands of biomolecules potentially available as biomarkers. Additionally, these methods rely heavily on sparse existing knowledge and the limited ability of researchers to form accurate mental models of large biological networks. Bottom-up methods are noisy and result in an intractably large list of molecular candidates for follow-up. Further, such a method provides few explicit links to the underlying mechanism of action, whereas an ideal biomarker should directly relate to injury or disease progression. However, both methods provide essential biological information that should be combined in a more global, systems-level approach to biomarker discovery.

Opportunities for Systems Biology in TBI Biomarker Discovery

Systems biology is a natural approach to investigate such complex molecular and cellular interactions. It allows for a holistic, systematic, and unbiased analysis of integrated experiment-specific, high-throughput genomics and proteomics data with canonical biological networks.[35] It integrates top-down knowledge of molecular mechanisms and processes embedded in the biological networks with bottom-up data generated by high-throughput techniques, facilitating the generation of novel hypotheses. Ultimately, systems biology should be used to generate a testable hypothesis that can be experimentally validated.[39]

In a systems biology approach, hypotheses are generated by the construction and analysis of genome-scale, data-driven models of biomolecules and their interactions. To this end, biological systems are abstracted as networks represented by ''nodes'' (biomolecules) and ''links'' (biochemical interactions). Nodes in a network model generally represent genes or gene products, although they can also represent metabolites,[40] drugs,[41] and diseases.[42] Nodes can be

assigned values specific to a biological condition, using bottom-up concentration measurements or top-down knowledge about a gene. For example, node values can represent the concentration of a gene's products, phenotypes induced by its perturbation, or mutation of its sequence.[43] Links can represent measurements of physical interaction, computationally predicted binding, phenotypic relationships, or other connections between nodes.[43–45] Thus, using molecular networks as a scaffold and overlaying data on the nodes, top-down and bottom-up data can be integrated into a unified structure,[46] bridging existing knowledge and discovery-based assays. Once data are converted to a network, algorithms from mathematics and physics, such as graph theory, systems science, and statistical mechanics, can be applied to extract network-level insights.

One possibility to exploit the promises of a systems approach is to integrate TBI high-throughput molecular data with two types of complementary biological networks, canonical pathways and protein-protein interaction (PPI) maps, with the goal of identifying TBI-specific pathways and protein interaction modules, respectively, that emerge within the context of the specific omics data. For example, a TBI gene expression data set can be integrated with pathways and PPI networks to add biological context, suggest new interrelationships, and hypothesize novel biomarkers (Fig. 1). Importantly, many genes may be unmeasured or nonsignificant in the original gene expression data set, but their significance may emerge within the context of the network connectivity information.

Available high-throughput data sets for TBI

Several high-throughput data sets are publicly available for constructing data-driven systems biology models of TBI. The most applicable and widely available high-throughput data for this purpose are gene microarrays and proteomics. Microarrays measure expression levels of messenger RNA (mRNA) for thousands of predefined genes within a target genome, whereas proteomics attempts to identify and quantify all of the proteins expressed within a cell. Because protein abundance does not always correlate well with mRNA levels,[47] protein expression profiles cannot simply be inferred from microarray data and must be measured independently.

In a microarray experiment, RNA from a biological sample is labeled with fluorescent tags and then hybridized to a microscale grid of nucleotide (nt) sequences corresponding to target genes. This grid is then imaged to quantify mRNA levels for all genes simultaneously. The ubiquitous use of this technology over the last decade led to the establishment of public repositories for microarray data, including the widely used Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) and ArrayExpress (http://www.ebi.ac.uk/arrayexpress). Table 2 compiles large-scale microarray studies from animal models of TBI gathered from these repositories, with their respective accession numbers.[48–56] Most of these TBI microarray studies used oligonucleotide platforms, such as Affymetrix (six studies) or Agilent (two studies), in which multiple short nt sequences matching a portion of each target gene are chemically bound to a surface, whereas two studies use cDNA platforms, in which a single cDNA sequence for the entire gene is spotted to a glass slide. Oligonucleotide platforms are more common, have standardized data-processing pipelines,[57] and are more reproducible than cDNA microarrays.[58]

The majority of the studies in Table 2 consist of microarray data of different rodent models of TBI, which measure mRNA expression levels in control and injury conditions for thousands of genes. Five *in vitro* studies measured gene expression from primary rodent cortical or hippocampal neurons, after either stretching or transecting the axons. *In vivo* microarray studies generally used either fluid percussion injury (FPI), in which injury is produced by the impact of a pendulum onto a fluid reservoir, or controlled cortical impact (CCI), in which a rigid, computer-controlled, pneumatically driven impactor strikes the dural surface.[59,60] The studies of Natale and colleagues[53] and Babikian and colleagues[54] have provided rich microarray data sets covering different animals (mouse and rat), models of TBI (FPI and CCI), severity levels (moderate to severe), and brain tissues (cortex and hippocampus) collected at distinct time points. Natale and colleagues, using an FPI rat model and a CCI mouse model, identified 82 genes differentially expressed in

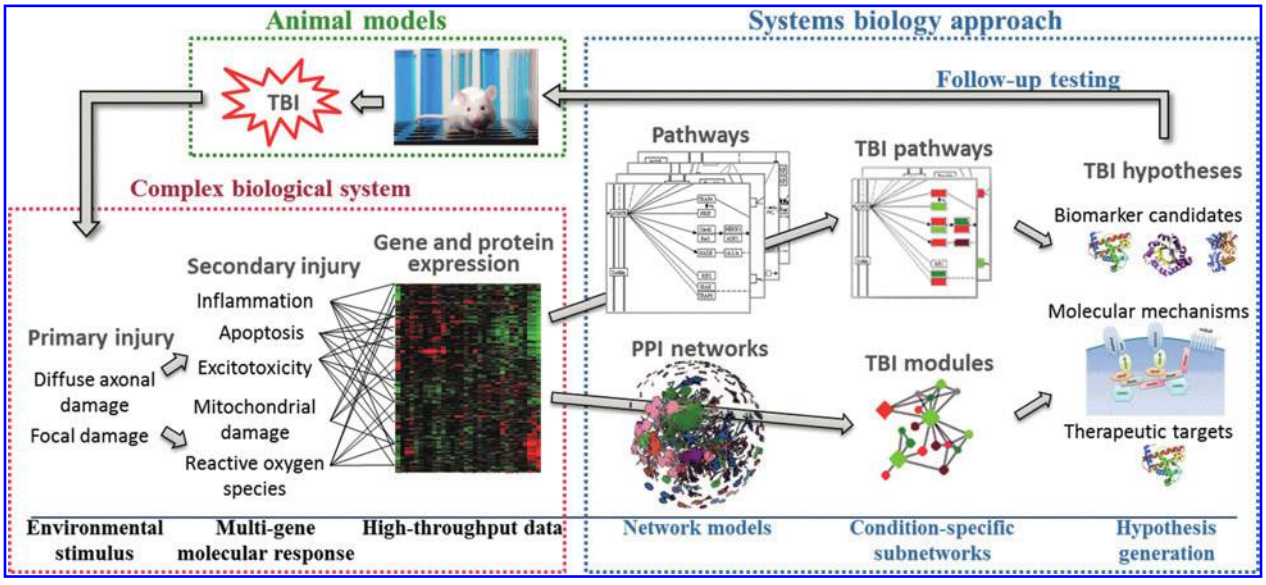

FIG. 1. Schematic representation of a systems biology approach to TBI. Pathways and protein interaction networks act as a scaffold to integrate heterogeneous information from high-throughput molecular data sets, distilling the complex molecular TBI response into testable hypotheses. PPI, protein-protein interaction; TBI, traumatic brain injury.

Table 2. Microarray Data Sets for TBI

		Conditions					
Accession	*Reference no.*	*Species*	*Tissues (measured)*	*Injury type*	*Severity*	*Time point(s)*	*Assay platform*
GSE21854	48	*Homo sapiens*	Plasma	Clinical	Severe (GCS, < 8)	68 ± 8 h	uParaflo (microRNA)
GSE16735	*	*Rattus norvegicus*	CA3 hippocampal neurons	*In vitro*	Death/survival	24 h	Agilent
GSE24047	49	*Rattus norvegicus*	Cerebral cortex	FPI—lateral moderate	Moderate (3.3 atm)	3, 6, 12, 48 h	Affy 230.2
E-MEXP-2235	50	*Rattus norvegicus*	Hippocampal neurons	*In vitro* diffuse axonal injury	Mild (10%), severe (50%) stretch	24 h	Agilent
GSE11730	51	*Rattus norvegicus*	Matured cortical axons	*In vitro* microfluidic axotomy	Axotomy	13 days	Affy 230.2
GSE5034	*	*Rattus norvegicus*	CA1, CA3 hippocampal 2D cultures	*In vitro* axonal transection	Axotomy	7 days	Affy 230.2
GSE1911	52	*Rattus norvegicus*	Ipsilateral hippocampus	CCI—lateral	2.3 mm depth, 6 m/s	3, 24 h	Affy U34A
GSE2392	53	*Mus musculus, Rattus norvegicus*	Parietal cortex below injury	FPI—lateral (rat), CCI—lateral (mouse)	Rat: 2.5 atm, Mouse: 1.5 mm depth, 6 m/s	4, 8, 24, 72 h	Affy U34A 74Av2
GSE2871	54	*Rattus norvegicus*	Cortex, hippocampus	FPI—lateral	Mild: toe pinch response < 45 sec, severe: > 120 sec	0.5, 4, 24 h	Affy U34A
E-CBIL-4	55	*Mus musculus*	CA3, dentate hippocampal neurons	CCI	Injured/uninjured	24 h	cDNA dye-swap
GSE3396	56	*Rattus norvegicus*		CCC	Injured/sham, maximum 3 mm compression	1, 4 days	cDNA

Databases: GSE, National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo); E, ArrayExpress (http://www.ebi.ac.uk/arrayexpress). Clinical samples in GSE21854 were collected at varying times postinjury, within the specified range. Asterisk under Reference no. column indicates that published reference was not available beyond the NCBI GEO entry.

TBI, traumatic brain injury; 2D, two-dimensional; FPI, fluid percussion injury; CCI, controlled cortical impact; CCC, controlled cortical contusion; GCS, Glasgow Coma Scale; atm, atmospheres; cDNA, complementary DNA.

both rat and mouse in at least one time point, whereas Babikian and colleagues, using an FPI rat model, discovered 269 unique genes up- or down-regulated in at least one of the experimental conditions (brain tissues, time after injury, and severity). Each of these studies provides lists of statistically significant genes and results from functional annotations of these lists of genes [i.e., enrichment analysis of Gene Ontology (GO) terms]; however, both remain otherwise unexplored by more-sophisticated, emerging systems biology techniques.

Proteomics data sets exist for TBI, but are much less common. Because proteins have more structural and chemical heterogeneity than mRNA, proteomics technologies have been slower to develop and require more specialized expertise. However, many labs in academia and industry have acquired these capabilities in two main areas: protein mixture separation and protein identification and quantification. These two areas are usually applied in tandem in proteomics studies. The techniques for protein mixture separation include gel electrophoresis and liquid chromatography (LC). Gel electrophoresis, such as sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS-PAGE) and two-dimensional gel electrophoresis (2DGE), separates proteins by mass and charge using electrical and pH gradients in a gel. The LC technique separates proteins according to their differential moving speeds in a flowing liquid (mobile phase) while passing through solid materials (stationary phase). The techniques for protein identification/quantification include immunoblotting and tandem mass spectrometry (MS/MS). The immunoblotting technique identifies proteins through the binding of protein-specific Abs and the subsequent radioactive, or fluorescent, detection of these Abs by linked reporter enzymes. The MS/MS technique identifies proteins by determining the mass-to-charge ratios of proteins (or fragmented peptides) and the subsequent matching of these ratios to a mass spectra database of known proteins (or peptides).[61] MS/MS can determine protein abundance in one sample or abundance changes between two samples, such as TBI and control samples. This can be achieved by various labeling techniques, such as isotope-coded affinity tag (ICAT) and isobaric tagging for relative and absolute quantification (iTRAQ).

A variety of combinations of the above-mentioned techniques have been used in the discovery of TBI biomarker candidates. For example, Jenkins and colleagues used 2DGE of young mice after CCI, staining with an Ab for protein kinase B (PKB) substrates, to identify 120 PKB substrate proteins that changed more than 5-fold after TBI.[62] Yao and colleagues used SDS-PAGE with a panel of 998 Abs, followed by Western blot analysis, to discover 18 proteins differentially expressed in a rat model of penetrating TBI.[37] Kobeissy and colleagues used a workflow combining cation/anion chromatography, SDS-PAGE, and LC-MS/MS to identify 59 proteins with changes in abundance in a mouse model of TBI.[63] In addition, Haqqani and associates used ICAT-MS/MS to identify 95 proteins differentially expressed in serum of patients with sTBI,[64] and Crawford and associates identified 35 proteins that are significantly related to TBI, using combinations of iTRAQ and LC-MS/MS in transgenic mice.[65] Although early proteomic studies were limited to the identification of a small number of differentially expressed proteins, technological advances have significantly increased this number. For example, recently, Cortes and colleagues identified 484 differentially expressed proteins in rat brain tissue using a CCI model.[66]

As evident in these studies involving bottom-up methods for biomarker discovery, microarray and proteomics experiments often identify hundreds of genes and proteins, which would be impossible to study one by one, especially when considering multiple time points or conditions. With the rapid improvement and increased availability of these and other genome-scale technologies, the major bottleneck is therefore in the analysis, rather than collection, of molecular data. The integration of such high-throughput data with biological pathways and networks provides a mechanism to further interpret and screen these large gene lists through contextual ''biological filters.''

Table 3. Publicly Available Systems Biology Databases and Web Tools

Interaction databases		*Systems biology tools*	
Database	*URL*	*Tool*	*URL*
Pathways		*Web services*	
Database of Cell Signaling	http://stke.sciencemag.org/cm/	DAVID	http://david.abcc.ncifcrf.gov
KEGG	http://www.genome.jp/kegg/	GENECODIS	http://genecodis.dacya.ucm.es
MSigDB	http://broadinstitute.org/gsea/msigdb/	Genetic Association Database	http://geneticassociationdb.nih.gov
WikiPathways	http://wikipathways.org	MIMI	http://mimi.ncibi.org/MimiWeb/
Networks		*Downloadable software*	
BIND	http://bond.unleashedinformatics.com	Cytoscape	http://cytoscape.org
BioGRID	http://thebiogrid.org	DisGeNet	http://ibi.imim.es/DisGeNET/
DIP	http://dip.doe-mbi.ucla.edu/dip/	Expander	http://acgt.cs.tau.ac.il/expander/
HPRD	http://hprd.org	GenePattern	http://genepattern.org
IntAct	http://www.ebi.ac.uk/intact		
MINT	http://mint.bio.uniroma2.it		
MIPS	http://mips.helmholtz-muenchen.de/proj/ppi/		
PDZBase	http://icb.med.cornell.edu/services/pdz/		
Reactome	http://reactome.org		

BIND, Biomolecular Interaction Network Database; BioGRID, Biological General Repository for Interaction Datasets; DAVID, Database for Annotation, Visualization and Integrated Discovery; DIP, Database of Interacting Proteins; DisGeNet, Disease Gene Networks; GENECODIS, GENE Annotations CO-occurrence DIScovery; HPRD, Human Protein Reference Database; KEGG, Kyoto Encyclopedia of Genes and Genomes; MIMI, MIchigan Molecular Interactions; MINT, Molecular INTeraction database; MIPS, Munich Information center for Protein Sequences; MSigDB, Molecular Signatures DataBase.

Pathways

Well-studied canonical pathways provide ''wiring diagrams'' describing how gene products and other biomolecules (e.g., lipids or metabolites) interact, relate, and regulate each other to perform biological functions. Canonical pathway diagrams are often used as a knowledge base to help design experiments and derive conclusions.

Pathways are often manually curated from the literature into large online compendia (see Table 3 for a list), which can be exploited to link disease- or injury-specific differentially expressed genes to biological processes and identify pathways associated with the studied disease or injury condition. The most commonly used pathway database is the Kyoto Encyclopedia of Genes and Genomes (KEGG),[67] which provides dynamic, hyperlinked maps connecting genes, biochemical reactions, and small molecules in 414 pathways from four categories: metabolism; cell signaling; disease mechanisms; and chemical compound synthesis. Reactome[68] is a cross-referenced pathway database similar in scope to KEGG, but with fewer organisms and a larger number of pathways. Unlike KEGG, each reaction in Reactome is annotated with GO terms, text descriptions, PubMed cross-references, and author information. The Molecular Signatures Database (MSigDB)[69] is a curated database of annotated gene sets, but does not provide wiring diagrams for each set. MSigDB is divided into five collections: (1) 326 gene sets from the same chromosome or cytogenetic band; (2) 3272 pathways compiled by experts from publications; (3) 836 gene sets thought to be targeted by a shared transcription factor or microRNA; (4) 881 gene sets gathered by mining cancer-related expression data; and (5) 1454 genes with shared functional annotations. WikiPathways[70] is an effort to extend the crowdsourcing approach of Wikipedia to construct consensus biological pathways, thus far resulting in 1668 pathways containing over 9500 edits submitted by users. Additionally, some companies have compiled large, proprietary pathway databases for which licenses are available for purchase, including Ingenuity Pathway Analysis (IPA), Ariadne Pathway Studio, and GeneGo Metacore.

One approach to integrate gene expression data with canonical pathways and identify significant pathways associated with the condition represented in the expression data is to perform statistical tests.[71] Such tests assess whether the number of differentially expressed genes in a pathway is significantly higher than what would be expected by chance. The development of statistical methods for automated pathway analysis is a rich area of research and there are several competing algorithms[71,72] and publicly available tools (Table 3). In the simplest form of pathway analysis, pathways from a selected database are tested for associations with a list of differentially expressed genes to identify pathways whose genes are represented in the list at a higher rate than expected by chance. Such statistical analysis invariably involves some variant of the Fisher's exact test (also called the ''hypergeometric test'' because of the use of the hypergeometric distribution). An example of such an application from Table 3 is the commonly used DAVID Web tool,[73] which calculates adjusted hypergeometric p values for both KEGG and Reactome pathways, given a gene list of interest. However, results from the hypergeometric test depend considerably on the subset of genes selected as significant (i.e., differentially expressed). Gene Set Enrichment Analysis (GSEA) addresses this problem by using expression values from an entire high-throughput experiment, without the need to select a subset of differentially expressed genes.[69,74] The MSigDB collection of gene sets was originally constructed for use with the GSEA algorithm, and the MSigDB Web site in Table 3 allows users to run GSEA on uploaded data. One drawback to GSEA and the hypergeometric test is that these methods treat pathways as unordered collections of genes and neither capitalizes on the topology, or connectivity patterns, among genes or proteins in a pathway. To address this limitation, algorithms such as signaling pathway impact analysis[75] and our group's PathNet[76] use the connectivity information of a pathway to determine its significance within the context of microarray data. In validation experiments using Alzheimer's disease (AD) microarray data sets, PathNet achieved better performance than non-topology-based algorithms.[76]

A few examples of pathway analysis have been performed for high-throughput data sets of TBI. Shojo and colleagues applied GSEA to microarray data from several time points after FPI in rats.[49] Their pathway analysis revealed time-dependent patterns in expression response of five pathways from the apoptosis and inflammatory systems, suggesting a causal temporal relationship between the two systems during the acute phase of TBI (<6 h), which faded after 48 h. They also integrated these pathways to propose the following systems-level hypothesis: an immediate inflammatory response by macrophages, triggered by the cytokines, interleukin (IL)-1α, IL-1β, and tumor necrosis factor, and mediated by inflammatory nuclear factor kappa B and mitogen-activated protein kinase signaling, induces an apoptosis program in neurons. Independently, Kobeissy and colleagues applied Pathway Studio to their TBI proteomics data set described above, reaffirming the involvement of inflammatory and survival signaling pathways.[77] In addition, their analysis identified novel pathways, especially synaptic plasticity, for further study for their association with TBI. Recently, Mondello and colleagues analyzed the function of proteins in their corresponding pathways to down-select TBI biomarkers from a list of potential candidates.[23]

Pathway analysis has also been applied to high-throughput studies of AD and its potential links to TBI. Chen and associates[78] used pathway analysis to reduce false positives in selecting biomarker candidates from a genomic data set of peripheral blood leukocytes from Alzheimer's patients. They used reverse-transcription polymerase chain reaction to validate expression of genes appearing in enriched pathways, resulting in 13 of 18 genes successfully validated *in vivo*. Crawford and colleagues[79] used IPA to examine networks involved in genomic response to TBI in rats with and without overexpression of the AD-related β-amyloid peptide. They concluded (similarly to Shojo and colleagues[49] above) that the AD rat model showed exacerbated immune response and cell death pathways after TBI.

Although pathway analysis is widely used in systems biology research, it has some limitations. One limitation of pathway diagrams is that they are constructed manually by experts to reflect consensus opinions. Accordingly, they are biased toward well-studied genes and interactions and are therefore inherently unable to discover novel biological mechanisms. Further, because pathway databases can only contain existing knowledge, they necessarily exclude any genes with unknown function, limiting their range of applicability. For example, although KEGG and Reactome are two of the largest, most widely used and freely available pathway databases, they contain only 5633 and 4437, respectively, of the nearly 20,000 human genes. Thus, in a whole-genome microarray experiment, only a fraction of genes can be investigated in pathway analysis. Another limitation of pathways is that they share a considerable number of genes. For example, of 130 nonmetabolic pathways from KEGG, 88 have only 20% or fewer genes unique to a pathway, and all pathways share at least one gene with another

pathway.[76] This ''promiscuity'' of genes across pathways may lead to false-positive pathway inferences when, by chance, a pathway happens to share many of its genes with the pathways that are truly active.

PPI networks

Recently developed high-throughput methods that capture protein-binding events have enabled researchers to systematically establish PPI maps for a large number of species. In contrast to the manually curated pathway databases, PPIs are now being detected through whole-genome, high-throughput experimental assays. Therefore, they cover a much broader range of proteins and can reveal novel biological mechanisms of action characterized by the underlying PPI network, where network nodes represent proteins and a link between two nodes indicates a PPI. The two most commonly used experimental assays to identify PPIs are (1) yeast two-hybrid (Y2H), which measures binary pairwise interactions in a yeast model, and (2) affinity purification followed by mass spectroscopy (AP/MS), which identifies protein complexes that associate with a bait protein in the biological system of interest.[44]

In Y2H interactome mapping, two candidate proteins (''bait'' and ''prey'') are fused to separate domains of a yeast transcription factor and expressed in yeast cells.[80] When the bait and prey interact, the transcription factor becomes functional and a reporter gene is expressed. This process has been automated for genome-scale throughput, resulting in large-scale interactome maps for yeast.[81,82] Importantly, proteins from other organisms can also be cloned into Y2H constructs, and they have been used to construct large-scale PPI maps for humans.[83,84] However, only a fraction of the estimated 100,000–130,000 human PPIs are thought to have been mapped by Y2H thus far.[85] In contrast, in AP/MS, the bait protein is tagged with a sequence recognizable by an Ab, expressed in the cell of interest, and isolated by a set of affinity purification steps.[86] Isolated complexes are then passed to a proteomics analysis pipeline (e.g., the LC-MS/MS technique described above) to identify interacting proteins.[87]

Both methods can produce high-quality interactions, but each provides fundamentally different information with unique limitations.[88,89] Protein complexes measured by AP/MS have ambiguous network interpretations because they can be represented either by the spoke model, in which interactions are inferred only between the bait and each prey protein in the purified complex, or by the fully connected model, in which each protein in the complex is assumed to interact with all other proteins. In contrast, interactions measured by Y2H are more naturally interpreted as binary, pairwise interactions. Though AP/MS identifies interactions in the endogenous system at the approximate physiological protein levels, protein concentrations in Y2H screens are not necessarily comparable to those found in their native environment, and, for the interactions to be detected, the interacting proteins must be localized to the nucleus. In addition, Y2H is more sensitive to low-affinity interactions that would not survive the purification process of AP/MS.[86] The reliability of each technique has been extensively reviewed in the literature, and comprehensive analyses have often resulted in contrasting conclusions.[88,90–94] For example, the overlap of Y2H screens by different laboratories is often small,[94] suggesting high false-negative rates, whereas AP/MS screens can infer a substantial number of indirect interactions, depending on the interaction model,[88] suggesting high false-positive rates. Further, the distribution of connectivity (i.e., links per node or degree distribution) in these networks reflects a probabilistic nature, perhaps because of abundance bias from intrinsic randomness in the interaction detection methods,[95] or the entropic effects of shuffling during their evolutionary construction.[96]

Currently available PPI data sets are of three types: (1) genome-scale screens aimed at probing all possible PPIs[83,84,87]; (2) semi-large-scale screens investigating interactions within a specific pathway or biological system[97,98]; and (3) small-scale, traditional studies aimed at detecting specific interactions among proteins of interest. Many databases compile PPIs from all three types of studies, which, together, form networks of thousands of proteins and tens of thousands of interactions. In these databases, interactions from the third type of study (small-scale) comprise 80% of interactions, although genome- and semi-large-scale interactome mapping are becoming increasingly common. In Table 3, we have compiled nine databases that include primary protein interactions (i.e., not a collection of aggregated data sets), collected solely from experimental measurements (i.e., not predicted computationally or mined from the literature). These data sets are known to be noisy, but many groups, including our own, have devised methods to distill them into high-confidence subsets. For example, Yu and colleagues consolidated three Y2H datasets into a single high-confidence network and showed that this set is more enriched with interactions found in a manually curated gold-standard set than a combined set from two AP/MS studies.[88] Our group has developed a statistical method, called Interaction Detection Based on Shuffling,[93,99] that generates high-confidence subsets by correcting for biases toward frequently studied proteins, effectively allowing the construction of protein interaction networks with a given false-positive rate (e.g., 5%).

Gene expression data have been integrated with PPI networks to identify regions of the original network associated with the condition represented in the microarray study.[100–105] Such analysis recovers coregulated, highly connected subnetworks (or functional protein interaction modules) that have been found to characterize biological processes[89] or to work together to produce a cellular phenotype.[80]

Several algorithms exist for decomposing PPI networks into functional modules. Seminal work by Ideker and colleagues devised a method to score the aggregate expression of a given subnetwork of genes and applied the stochastic optimization-simulated annealing method to the global network to identify the highest-scoring subnetworks.[101] Since then, other groups have devised competing methods that incorporate graph theory, engineering optimization, and heuristics.[104–106] Some of these algorithms have been implemented in downloadable software tools, such as Cytoscape, Expander, and Matisse,[104,107,108] with graphical user interfaces for use by biologists (Table 3). These techniques have been applied to biological systems, such as the DNA-damage response in yeast,[109] prediction of metastatic potential in cancer patients,[102,110,111] and genes altered in type 2 diabetes.[112] In an application of the approach to neurological disease, Ma and colleagues used protein interaction networks and well-known AD disease genes to prioritize genes that were differentially expressed in AD microarray studies.[113] However, this approach has not yet been applied to discover new protein interaction modules, and thus new molecular mechanisms of action, in TBI.

Application of Systems Biology to Identify TBI Biomarker Candidates

In this section, we provide an example to illustrate and provide a specific context for the systems biology concepts discussed above.

Using some of the systems biology resources in Table 3, we integrated a list of 32 previously reported protein TBI biomarker candidates (Table 1) with publicly available canonical pathways and human PPI networks to illustrate how to systematically generate new, testable hypotheses and identify candidate biomarkers for TBI.

In an actual analysis, one should start from a list of condition-specific, high-throughput genomics or proteomics data, instead of a small list of predetermined biomarkers as in this illustrative example, and project them onto injury-independent pathways and PPI scaffolds to delineate the subset of protein interactions associated with the specific condition. Thus, by repeating such an analysis for distinct conditions (e.g., injury severity level and time postinjury), one could potentially identify patterns that stratify secondary injury response for each of the conditions represented in the high-throughput data.

A literature-derived list of TBI proteins

Table 1 lists the 32 TBI biomarker candidate proteins that we compiled from the literature, ordered by the number of identified citations, with the top eight proteins (GFAP, S100B, UCHL1, ENO2, SPTAN1, MBP, MAPT, and FABP7) garnering multiple citations (see Supplementary Table 1). The proteins in this list have diverse roles across cellular metabolism, cytoskeleton, calcium binding, and other functions. Although many are specific to the CNS, these proteins share little else in common and show no directly obvious relationship to TBI injury mechanisms.

Enrichment analysis discovers unifying biological themes from a list of genes or proteins of interest, based on commonly occurring gene annotations. Using the GENECODIS[114,115] tool in Table 3, which performs enrichment analysis for diverse types of annotations simultaneously, the biomarker candidate list was found to have statistically significant enrichment with GO biological process terms related to apoptosis and neurogenesis. We also used the Genetic Association Database[116] to find disease terms associated with proteins in the list that were observed to a higher degree than what would be expected by chance. This analysis uncovered associations with several neurological and CNS diseases, including AD and schizophrenia (Supplementary Table 2; see online supplementary material at http://www.liebertpub.com). Associations with AD reflect multiple emerging lines of evidence for long-term neurological disease after TBI. For example, brain injury induces altered subanatomical features resembling AD, such as amyloid-β deposits, neurofibrillary tangles, and acetylcholine deficiency.[117,118] Retired football players with a history of chronic mTBI (i.e., multiple concussions) have increased cognitive impairment and earlier onset of AD.[119]

It must be noted that this analysis is only for the purpose of demonstration, because disease annotations of genes are themselves ultimately derived from experimental results reported in the literature. Therefore, it may be somewhat circular to apply enrichment analysis to a literature-derived set of genes. However, when analyzing unbiased lists of differentially expressed genes from proteomics or microarray data, statistical enrichment of biological annotations can be used to formulate new hypotheses about molecular mechanisms.

Pathway analysis of candidate genes

As Table 1 shows, many TBI biomarker candidates appear in multiple KEGG pathways, making it difficult to identify significant trends. For very large pathways, it might be expected that any list of randomly selected genes would contain multiple genes associated with that pathway. Therefore, statistical methods must be applied to discover the most relevant pathways significantly associated with a gene list.

We explored our 32 biomarker candidates for pathway enrichment, applying the hypergeometric test to 130 nonmetabolic pathways from the KEGG database. Only four KEGG pathways were significantly enriched ($p < 0.05$): legionellosis; AD; amyotrophic lateral sclerosis (ALS); and apoptosis (Supplementary Table 3; see online supplementary material at http://www.liebertpub.com). Legionellosis is an infection caused by *Legionella* bacteria and not likely to be relevant to TBI, whereas the other three results are more closely related to neural function and will be the focus of this analysis. A closer look at the wiring diagram of the enriched pathways can help clarify the function of TBI biomarker candidates within each well-understood biological context, and can drive hypothesis generation, both for targets of companion therapeutics and for novel biomarker candidates with similar biological roles. As an illustration, Supplementary Figures 1, 2, and 3 (see online supplementary material at http://www.liebertpub.com) depict the three significant, neural-related pathways, annotated with symbols designating known TBI biomarker candidates, known drug targets, and proteins that interact with multiple TBI biomarker candidates. Notably, there is considerable overlap of apoptosis-related TBI biomarker candidates in the two neurological disease pathways. These proteins, including BCL-2 in the ALS pathway, CASP7 in the AD pathway, and CASP9 and CYCS (CytC in Supplementary Figs. 2 and 3; see online supplementary material at http://www.liebertpub.com) in both, are well-known downstream effectors of apoptosis that, taken individually, were each found to have only one citation as a biomarker candidate in the TBI literature. However, their relevance becomes clearer in the aggregate context of pathways. Apoptosis proteins comprise all but one (NEFH) of the TBI biomarker candidates found within the ALS pathway, confirming the importance of apoptosis as a postinjury mechanism. At the same time, however, this illustrates the possible danger that promiscuous genes may cause certain pathways, which may be, on the whole, unrelated to the condition of interest, to emerge as statistically significant. By contrast, half of the biomarker candidates associated with AD were unique to that pathway, supporting the association of these proteins with postinjury mechanisms of progression to neurological disease.

PPI network analysis of candidate genes

Although our pathway analysis recapitulated the known biology of the cellular response to TBI and provided a mechanistic context for known biomarker candidates, this approach is inherently unable to reveal new interactions among these and other proteins. To this end, we overlaid the biomarker candidate list onto a high-confidence PPI network to reveal previously unknown interactions among TBI biomarker candidates, discover novel protein candidates, and generate biological hypotheses from patterns of connectivity.

We created a comprehensive PPI network of 11,789 proteins and 74,376 interactions by combining all nine PPI databases in Table 3. Among the 32 TBI biomarker candidates, 30 had nodes represented in the network (Table 1) and there were 15 interactions among them. In sharp contrast, had we randomly selected 30 proteins from the set of 11,789, on average, we would have observed 0.39 interactions among them ($N = 1000$ random samples), indicating that the biomarker candidates are highly interconnected within the PPI network. We also identified a number of other proteins that interact with these TBI biomarker candidates, including 35 proteins known to have three or more interactions with them (Supplementary Table 4; see online supplementary material at http://www.liebertpub.com). Among these 35 proteins, seven (ABL1, IKBKE, UBC, PSEN1, CASP3, CASP8, and BCL2L1) were found to be highly connected (having five or more interactions) with this set of biomarker candidates and may be

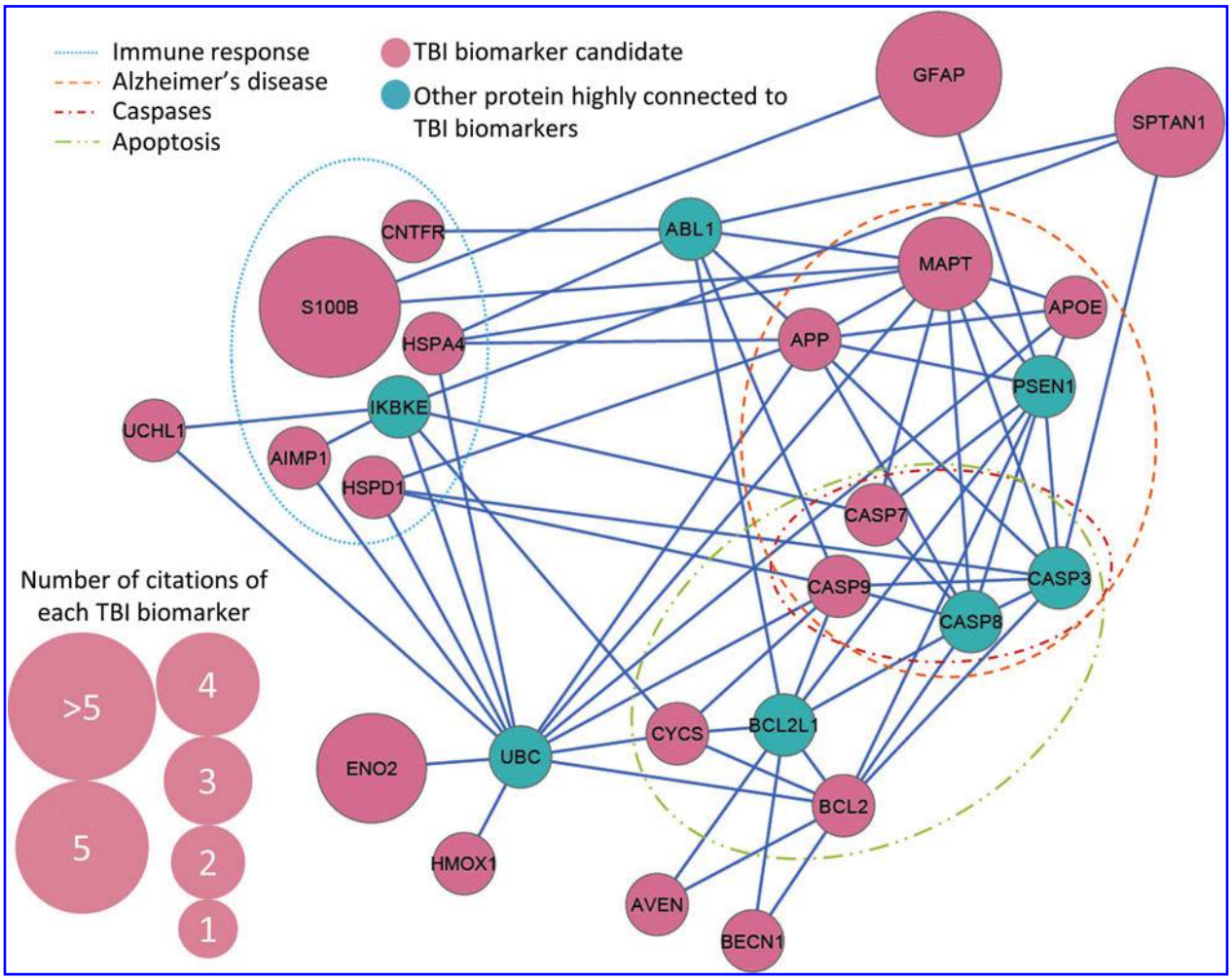

FIG. 2. Projection onto the PPI network of all interconnected TBI biomarker candidates (pink) plus seven novel proteins found to interact with at least five biomarker candidates in the network (blue). The number of citations for each biomarker candidate is denoted by the size of the node. Proteins were categorized by biological function by manual inspection of KEGG, Reactome, and Gene Ontology annotations for each protein. Immune system proteins have many direct interactions with proteins associated with Alzheimer's disease and apoptosis, as well as indirect interactions with these proteins through well-studied biomarkers SPTAN1 and GFAP. PPI, protein-protein interaction; TBI, traumatic brain injury; KEGG, Kyoto Encyclopedia of Genes and Genomes; SPTAN1, spectrin, alpha, non-erythrocytic 1 (alpha-fodrin); GFAP, glial fibrillary acidic protein.

potential candidates themselves. Two of these seven proteins (UBC and ABL1) are hubs, or proteins that are several orders of magnitude more highly connected than the average protein. Exploration of immediate neighbors of interacting proteins is biased toward the discovery of hubs because of their large connectivity. However, further statistical tests showed that both UBC and ABL1 have significantly more interactions with TBI biomarkers than would be expected by random chance (hypergeometic test $p < 10^{-6}$ for both) and were therefore included in our analysis.

Figure 2 depicts the wiring diagram of a core network containing 19 TBI biomarker candidates and the other seven proteins highly connected to these biomarker candidates. TBI biomarker candidates with high numbers of citations are emphasized in Figure 2 by node size. The proteins in this network can be roughly divided into four groups using GO biological process, KEGG, and Reactome pathway annotations: immune response; caspases; apoptosis; and AD. Caspases are responsible for effecting protein cleavage during the final steps of apoptosis, and network analysis identified caspase -3 and -8 as having similar connectivity to TBI biomarkers as the previously studied candidate proteins, caspase -7 and -9. Although our pathway analysis implicated AD as a shared pathway for TBI biomarker candidates, PPI network analysis further revealed new interactions with known AD proteins, such as presenilin (PSEN1). Two of the best-studied biomarker candidates (S100B and GFAP) have less well-known associations with AD in the literature; however, they directly interacted with AD proteins Tau (MAPT) and PSEN1, respectively, in the PPI network.

Importantly, an immune-related cluster of TBI biomarker candidates was directly connected to several AD-related proteins, as well as indirectly connected through the well-studied biomarkers, GFAP and SPTAN1 (αII-Spectrin Breakdown Products; SBDPs), and through hub proteins UBC and ABL1. As mentioned above, recent evidence has emerged suggesting that TBI-induced early inflammation cascades may trigger neuronal apoptosis events,[49,79] and our network analysis supports the possibility of mechanistic interactions between these pathways. Additionally, well-studied biomarker candidates GFAP and SPTAN1 (i.e., SBDP) may be involved in mediating this response.

Of the seven novel proteins emerging from this integrative network analysis, the protein kinase, ABL1, may be the most interesting. A DNA translocation event common to chronic myeloid leukemia connects the ABL1 and BCR genes, producing an oncogenic fusion protein (BCR/ABL) that is selectively targeted by the existing, FDA-approved drug, imatinib.[120] ABL1 is also known to be associated with AD,[121] but not with TBI. Thus, ABL1 is a tractable drug target that represents a possible therapeutic opportunity for intervening in the progression to neurodegenerative disease after TBI.

Conclusion

TBI is a complex, multicellular neurological condition that has confounded previous attempts to discover molecular biomarkers. Systems biology may help distill high-throughput data from the complex TBI response into novel hypotheses, and existing high-throughput data sets and publicly available tools provide new opportunities for applying such systems approaches.

A well-known challenge of biomarker discovery in TBI is the difficulty of acquiring clinical samples of injured tissue. One reason for the successful clinical application of high-throughput techniques to cancer, for example, in the development of prognostic gene signatures for breast cancer,[122,123] has been the wide availability of tumor samples from routine biopsies. TBI researchers, by contrast, are forced to rely instead on animal models of brain injury. Systems biology may help address this challenge. Rather than viewing model organisms as a limitation, systems biology relies on them by definition,[35,43] leveraging the reproducibility and controllability of animal experiments for iterative cycles of hypothesis generation, experimental testing, and model refinement.

Animal experiments do not always reproduce the same results across studies. This is primarily because of variations in animal species, injury type and severity, time course of collection, and sampled tissue. Nevertheless, biomarker candidates are more likely to have clinical applicability if they are insensitive to these experimental variations. Systems biology can be valuable for this purpose as well, in that expression patterns of network modules and pathways have been shown to be more reproducible across data sets than individual genes. For example, two microarray studies of breast cancer reported distinct sets of genes predictive of clinical outcome, but with little overlap between them. Systems biology analysis of these same data sets, however, showed considerable overlap in the expression of pathways and network modules associated with these gene lists.[102,110] Network-based modules of interacting genes have also been shown to be more conserved across species than the individual member genes in the modules.[124–126]

We illustrated the application of a typical systems biology approach using a manually compiled list of candidate TBI biomarkers, rather than a high-throughput data set. We integrated this top-down knowledge of disease-related markers and pathways with a bottom-up, unbiased network approach to hypothesize potential new biomarkers for further research. Our analysis identified several potential candidate biomarkers for further study, including ABL1, which also has potential as a tractable therapeutic target.

Acknowledgments

The authors were supported, in part, by the U.S. Department of Defense, Defense Medical Research and Development Program managed by the U.S. Army Medical Research and Materiel Command, Ft. Detrick, Maryland (project no.: D61_I_10_J6_126).

The opinions and assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Army or the U.S. Department of Defense. This article has been approved for public release with unlimited distribution.

Author Disclosure Statement

No competing financial interests exist.

References

1. DeKosky, S.T., Ikonomovic, M.D., and Gandy S. (2010). Traumatic brain injury—football, warfare, and long-term effects. N. Engl. J. Med. 363, 1293–1296.
2. Hergenroeder, G.W., Redell, J.B., Moore, A.N., and Dash, P.K. (2008). Biomarkers in the clinical diagnosis and management of traumatic brain injury. Mol. Diagn. Ther. 12, 345–358.
3. Defense and Veterans Brain Injury Center. (2012). DOD Worldwide numbers for traumatic brain injury. Available at: http://www.dvbic.org/dod-worldwide-numbers-tbi. (Last accessed February 23, 2012.)
4. MacGregor, A.J., Shaffer, R.A., Dougherty, A.L., Galarneau, M.R., Raman, R., Baker, D.G., Lindsay, S.P., Golomb, B.A., and Corson, K.S. (2010). Prevalence and psychological correlates of traumatic brain injury in Operation Iraqi Freedom. J. Head Trauma Rehabil. 25, 1–8.
5. Powell, J.M., Ferraro, J.V., Dikmen, S.S., Temkin, N.R., and Bell, K.R. (2008). Accuracy of mild traumatic brain injury diagnosis. Arch. Phys. Med. Rehabil. 89, 1550–1555.
6. Ponsford, J., Willmott, C., Rothwell, A., Cameron, P., Ayton, G., Nelms, R., Curran, C., and Ng, K. (2001). Impact of early intervention on outcome after mild traumatic brain injury in children. Pediatrics 108, 1297–1303.
7. Ponsford, J., Willmott, C., Rothwell, A., Cameron, P., Kelly, A.M., Nelms, R., and Curran, C. (2002). Impact of early intervention on outcome following mild head injury in adults. J. Neurol. Neurosurg. Psychiatry 73, 330–332.
8. Dash, P.K., Zhao, J., Hergenroeder, G., and Moore, A.N. (2010). Biomarkers for the diagnosis, prognosis, and evaluation of treatment efficacy for traumatic brain injury. Neurotherapeutics 7, 100–114.
9. Morrow, D.A., Cannon, C.P., Rifai, N., Frey, M.J., Vicari, R., Lakkis, N., Robertson, D.H., Hille, D.A., DeLucca, P.T., DiBattiste, P.M., Demopoulos, L.A., Weintraub, W.S., and Braunwald, E.; TACTICS-TIMI 18 Investigators. (2001). Ability of minor elevations of troponins I and T to predict benefit from an early invasive strategy in patients with unstable angina and non-ST elevation myocardial infarction: results from a randomized trial. JAMA 286, 2405–2412.
10. Kulasingam, V., and Diamandis, E.P. (2008). Strategies for discovering novel cancer biomarkers through utilization of emerging technologies. Nat. Clin. Pract. Oncol. 5, 588–599.
11. Ioannidis, J.P., and Panagiotou, O.A. (2011). Comparison of effect sizes associated with biomarkers reported in highly cited individual articles and in subsequent meta-analyses. JAMA 305, 2200–2210.
12. Vos, P.E., Jacobs, B., Andriessen, T.M., Lamers, K.J., Borm, G.F., Beems, T., Edwards, M., Rosmalen, C.F., and Vissers, J.L. (2010). GFAP and S100B are biomarkers of traumatic brain injury: an observational cohort study. Neurology 75, 1786–1793.
13. Undén, J., and Romner, B. (2010). Can low serum levels of S100B predict normal CT findings after minor head injury in adults?: an evidence-based review and meta-analysis. J. Head Trauma Rehabil. 25, 228–240.
14. Anderson, R.E., Hansson, L.O., Nilsson, O., Dijlai-Merzoug, R., and Settergren, G. (2001). High serum S100B levels for trauma patients without head injuries. Neurosurgery 48, 1255–1258; discussion, 1258–1260.
15. Bloomfield, S.M., McKinney, J., Smith, L., and Brisman, J. (2007). Reliability of S100B in predicting severity of central nervous system injury. Neurocrit. Care 6, 121–138.
16. Savola, O., Pyhtinen, J., Leino, T.K., Siitonen, S., Niemela, O., and Hillbom, M. (2004). Effects of head and extracranial injuries on

serum protein S100B levels in trauma patients. J. Trauma 56, 1229–1234; discussion, 1234.
17. Honda, M., Tsuruta, R., Kaneko, T., Kasaoka, S., Yagi, T., Todani, M., Fujita, M., Izumi, T., and Maekawa, T. (2010). Serum glial fibrillary acidic protein is a highly specific biomarker for traumatic brain injury in humans compared with S-100B and neuron-specific enolase. J. Trauma 69, 104–109.
18. Ruan, S., Noyes, K., and Bazarian, J.J. (2009). The economic impact of S-100B as a pre-head CT screening test on emergency department management of adult patients with mild traumatic brain injury. J. Neurotrauma 26, 1655–1664.
19. Liliang, P.C., Liang, C.L., Weng, H.C., Lu, K., Wang, K.W., Chen, H.J., and Chuang, J.H. (2010). Tau proteins in serum predict outcome after severe traumatic brain injury. J. Surg. Res. 160, 302–307.
20. Zemlan, F.P., Jauch, E.C., Mulchahey, J.J., Gabbita, S.P., Rosenberg, W.S., Speciale, S.G., and Zuccarello, M. (2002). C-tau biomarker of neuronal damage in severe brain injured patients: association with elevated intracranial pressure and clinical outcome. Brain Res. 947, 131–139.
21. Kavalci, C., Pekdemir, M., Durukan, P., Ilhan, N., Yildiz, M., Serhatlioglu, S., and Seckin, D. (2007). The value of serum tau protein for the diagnosis of intracranial injury in minor head trauma. Am. J. Emerg. Med. 25, 391–395.
22. Mondello, S., Robicsek, S.A., Gabrielli, A., Brophy, G.M., Papa, L., Tepas, J., Robertson, C., Buki, A., Scharf, D., Jixiang, M., Akinyi, L., Muller, U., Wang, K.K., and Hayes, R.L. (2010). alphaII-spectrin breakdown products (SBDPs): diagnosis and outcome in severe traumatic brain injury patients. J. Neurotrauma 27, 1203–1213.
23. Mondello, S., Muller, U., Jeromin, A., Streeter, J., Hayes, R.L., and Wang, K.K. (2011). Blood-based diagnostics of traumatic brain injuries. Expert Rev. Mol. Diagn. 11, 65–78.
24. Vos, P.E., Lamers, K.J.B., Hendriks, J.C.M., van Haaren, M., Beems, T., Zimmerman, C., van Geel, W., de Reus, H., Biert, J., and Verbeek, M.M. (2004). Glial and neuronal proteins in serum predict outcome after severe traumatic brain injury. Neurology 62, 1303–1310.
25. Mondello, S., Papa, L., Buki, A., Bullock, M.R., Czeiter, E., Tortella, F.C., Wang, K.K., and Hayes, R.L. (2011). Neuronal and glial markers are differently associated with computed tomography findings and outcome in patients with severe traumatic brain injury: a case control study. Crit. Care 15, R156.
26. Papa, L., Lewis, L.M., Falk, J.L., Zhang, Z., Silvestri, S., Giordano, P., Brophy, G.M., Demery, J.A., Dixit, N.K., Ferguson, I., Liu, M.C., Mo, J., Akinyi, L., Schmid, K., Mondello, S., Robertson, C.S., Tortella, F.C., Hayes, R.L., and Wang, K.K. (2012). Elevated levels of serum glial fibrillary acidic protein breakdown products in mild and moderate traumatic brain injury are associated with intracranial lesions and neurosurgical intervention. Ann. Emerg. Med. 59, 471–483.
27. Brophy, G.M., Mondello, S., Papa, L., Robicsek, S.A., Gabrielli, A., Tepas, J., 3rd, Buki, A., Robertson, C., Tortella, F.C., Hayes, R.L., and Wang, K.K. (2011). Biokinetic analysis of ubiquitin C-terminal hydrolase-L1 (UCH-L1) in severe traumatic brain injury patient biofluids. J. Neurotrauma 28, 861–870.
28. Mondello, S., Linnet, A., Buki, A., Robicsek, S., Gabrielli, A., Tepas, J., Papa, L., Brophy, G.M., Tortella, F., Hayes, R.L., and Wang, K.K. (2012). Clinical utility of serum levels of ubiquitin C-terminal hydrolase as a biomarker for severe traumatic brain injury. Neurosurgery 70, 666–675.
29. Papa, L., Akinyi, L., Liu, M.C., Pineda, J.A., Tepas, J.J., 3rd, Oli, M.W., Zheng, W., Robinson, G., Robicsek, S.A., Gabrielli, A., Heaton, S.C., Hannay, H.J., Demery, J.A., Brophy, G.M., Layon, J., Robertson, C.S., Hayes, R.L., and Wang, K.K. (2010). Ubiquitin C-terminal hydrolase is a novel biomarker in humans for severe traumatic brain injury. Crit. Care Med. 38, 138–144.
30. Ho, L., Sharma, N., Blackman, L., Festa, E., Reddy, G., and Pasinetti, G.M. (2005). From proteomics to biomarker discovery in Alzheimer's disease. Brain Res. Brain Res. Rev. 48, 360–369.
31. Mondello, S., Jeromin, A., Buki, A., Bullock, R., Czeiter, E., Kovacs, N., Barzo, P., Schmid, K., Tortella, F., Wang, K.K., and Hayes, R.L. (2012). Glial neuronal ratio: a novel index for differentiating injury type in patients with severe traumatic brain injury. J. Neurotrauma 29, 1096–1104.
32. Loane, D.J., and Faden, A.I. (2010). Neuroprotection for traumatic brain injury: translational challenges and emerging therapeutic strategies. Trends Pharmacol. Sci. 31, 596–604.
33. Werner, C., and Engelhard, K. (2007). Pathophysiology of traumatic brain injury. Br. J. Anaesth. 99, 4–9.
34. Bayir, H., Clark, R.S., and Kochanek, P.M. (2003). Promising strategies to minimize secondary brain injury after head trauma. Crit. Care Med. 31, S112–S117.
35. Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. Annu. Rev. Genomics Hum. Genet. 2, 343–372.
36. Noorbakhsh, F., Overall, C.M., and Power, C. (2009). Deciphering complex mechanisms in neurodegenerative diseases: the advent of systems biology. Trends Neurosci. 32, 88–100.
37. Yao, C., Williams, A.J., Ottens, A.K., May Lu, X.C., Chen, R., Wang, K.K., Hayes, R.L., Tortella, F.C., and Dave, J.R. (2008). Detection of protein biomarkers using high-throughput immunoblotting following focal ischemic or penetrating ballistic-like brain injuries in rats. Brain Inj. 22, 723–732.
38. Yao, C., Williams, A.J., Ottens, A.K., Lu, X.C., Liu, M.C., Hayes, R.L., Wang, K.K., Tortella, F.C., and Dave, J.R. (2009). P43/pro-EMAPII: a potential biomarker for discriminating traumatic versus ischemic brain injury. J. Neurotrauma 26, 1295–1305.
39. Kell, D.B., and Oliver, S.G. (2004). Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. Bioessays 26, 99–105.
40. Papin, J.A., Price, N.D., Wiback, S.J., Fell, D.A., and Palsson, B.O. (2003). Metabolic pathways in the post-genome era. Trends Biochem. Sci. 28, 250–258.
41. Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabasi, A.L., and Vidal, M. (2007). Drug-target network. Nat. Biotechnol. 25, 1119–1126.
42. Goh, K.I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., and Barabasi, A.L. (2007). The human disease network. Proc. Natl. Acad. Sci. U. S. A. 104, 8685–8690.
43. Joyce, A.R., and Palsson, B.O. (2006). The model organism as a system: integrating 'omics' data sets. Nat. Rev. Mol. Cell. Biol. 7, 198–210.
44. Bork, P., Jensen, L.J., von Mering, C., Ramani, A.K., Lee, I., and Marcotte, E.M. (2004). Protein interaction networks from yeast to human. Curr. Opin. Struct. Biol. 14, 292–299.
45. Barabasi, A.L., and Oltvai, Z.N. (2004). Network biology: understanding the cell's functional organization. Nat. Rev. Genet. 5, 101–113.
46. Ideker, T., and Lauffenburger, D. (2003). Building with a scaffold: emerging strategies for high- to low-level cellular modeling. Trends Biotechnol. 21, 255–262.
47. Rabani, M., Levin, J.Z., Fan, L., Adiconis, X., Raychowdhury, R., Garber, M., Gnirke, A., Nusbaum, C., Hacohen, N., Friedman, N., Amit, I., and Regev, A. (2011). Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. Nat. Biotechnol. 29, 436–442.
48. Redell, J.B., Moore, A.N., Ward, N.H., 3rd, Hergenroeder, G.W., and Dash, P.K. (2010). Human traumatic brain injury alters plasma microRNA levels. J. Neurotrauma 27, 2147–2156.
49. Shojo, H., Kaneko, Y., Mabuchi, T., Kibayashi, K., Adachi, N., and Borlongan, C.V. (2010). Genetic and histologic evidence implicates role of inflammation in traumatic brain injury-induced apoptosis in the rat cerebral cortex following moderate fluid percussion injury. Neuroscience 171, 1273–1282.
50. Di Pietro, V., Amin, D., Pernagallo, S., Lazzarino, G., Tavazzi, B., Vagnozzi, R., Pringle, A., and Belli, A. (2010). Transcriptomics of traumatic brain injury: gene expression and molecular pathways of different grades of insult in a rat organotypic hippocampal culture model. J. Neurotrauma 27, 349–359.
51. Taylor, A.M., Berchtold, N.C., Perreau, V.M., Tu, C.H., Li Jeon, N., and Cotman, C.W. (2009). Axonal mRNA in uninjured and regenerating cortical mammalian axons. J. Neurosci. 29, 4697–4707.
52. Matzilevich, D.A., Rall, J.M., Moore, A.N., Grill, R.J., and Dash, P.K. (2002). High-density microarray analysis of hippocampal gene expression following experimental brain injury. J. Neurosci. Res. 67, 646–663.
53. Natale, J.E., Ahmed, F., Cernak, I., Stoica, B., and Faden, A.I. (2003). Gene expression profile changes are commonly modulated across models and species after traumatic brain injury. J. Neurotrauma 20, 907–927.
54. Babikian, T., Prins, M.L., Cai, Y., Barkhoudarian, G., Hartonian, I., Hovda, D.A., and Giza, C.C. (2010). Molecular and physiological

responses to juvenile traumatic brain injury: focus on growth and metabolism. Dev. Neurosci. 32, 431–441.
55. Marciano, P.G., Brettschneider, J., Manduchi, E., Davis, J.E., Eastman, S., Raghupathi, R., Saatman, K.E., Speed, T.P., Stoeckert, C.J., Jr., Eberwine, J.H., and McIntosh, T.K. (2004). Neuron-specific mRNA complexity responses during hippocampal apoptosis after traumatic brain injury. J. Neurosci. 24, 2866–2876.
56. von Gertten, C., Flores Morales, A., Holmin, S., Mathiesen, T., and Nordqvist, A.C. (2005). Genomic responses in rat cerebral cortex after traumatic brain injury. BMC Neurosci. 6, 69.
57. Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics 21, 3439–3440.
58. Consortium, M., Shi, L., Reid, L.H., Jones, W.D., Shippy, R., Warrington, J.A., Baker, S.C., Collins, P.J., de Longueville, F., Kawasaki, E.S., Lee, K.Y., Luo, Y., Sun, Y.A., Willey, J.C., Setterquist, R.A., Fischer, G.M., Tong, W., Dragan, Y.P., Dix, D.J., Frueh, F.W., Goodsaid, F.M., Herman, D., Jensen, R.V., Johnson, C.D., Lobenhofer, E.K., Puri, R.K., Schrf, U., Thierry-Mieg, J., Wang, C., Wilson, M., Wolber, P.K., Zhang, L., Amur, S., Bao, W., Barbacioru, C.C., Lucas, A.B., Bertholet, V., Boysen, C., Bromley, B., Brown, D., Brunner, A., Canales, R., Cao, X.M., Cebula, T.A., Chen, J.J., Cheng, J., Chu, T.M., Chudin, E., Corson, J., Corton, J.C., Croner, L.J., Davies, C., Davison, T.S., Delenstarr, G., Deng, X., Dorris, D., Eklund, A.C., Fan, X.H., Fang, H., Fulmer-Smentek, S., Fuscoe, J.C., Gallagher, K., Ge, W., Guo, L., Guo, X., Hager, J., Haje, P.K., Han, J., Han, T., Harbottle, H.C., Harris, S.C., Hatchwell, E., Hauser, C.A., Hester, S., Hong, H., Hurban, P., Jackson, S.A., Ji, H., Knight, C.R., Kuo, W.P., LeClerc, J.E., Levy, S., Li, Q.Z., Liu, C., Liu, Y., Lombardi, M.J., Ma, Y., Magnuson, S.R., Maqsodi, B., McDaniel, T., Mei, N., Myklebost, O., Ning, B., Novoradovskaya, N., Orr, M.S., Osborn, T.W., Papallo, A., Patterson, T.A., Perkins, R.G., Peters, E.H., Peterson, R., Philips, K.L., Pine, P.S., Pusztai, L., Qian, F., Ren, H., Rosen, M., Rosenzweig, B.A., Samaha, R.R., Schena, M., Schroth, G.P., Shchegrova, S., Smith, D.D., Staedtler, F., Su, Z., Sun, H., Szallasi, Z., Tezak, Z., Thierry-Mieg, D., Thompson, K.L., Tikhonova, I., Turpaz, Y., Vallanat, B., Van, C., Walker, S.J., Wang, S.J., Wang, Y., Wolfinger, R., Wong, A., Wu, J., Xiao, C., Xie, Q., Xu, J., Yang, W., Zhang, L., Zhong, S., Zong, Y., and Slikker, W., Jr. (2006). The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. Nat. Biotechnol. 24, 1151–1161.
59. Morales, D.M., Marklund, N., Lebold, D., Thompson, H.J., Pitkanen, A., Maxwell, W.L., Longhi, L., Laurer, H., Maegele, M., Neugebauer, E., Graham, D.I., Stocchetti, N., and McIntosh, T.K. (2005). Experimental models of traumatic brain injury: do we really need to build a better mousetrap? Neuroscience 136, 971–989.
60. Williams, A.J., Hartings, J.A., Lu, X.C., Rolli, M.L., Dave, J.R., and Tortella, F.C. (2005). Characterization of a new rat model of penetrating ballistic brain injury. J. Neurotrauma 22, 313–331.
61. Gstaiger, M., and Aebersold, R. (2009). Applying mass spectrometry-based proteomics to genetics, genomics and network biology. Nat. Rev. Genet. 10, 617–627.
62. Jenkins, L.W., Peters, G.W., Dixon, C.E., Zhang, X., Clark, R.S., Skinner, J.C., Marion, D.W., Adelson, P.D., and Kochanek, P.M. (2002). Conventional and functional proteomics using large format two-dimensional gel electrophoresis 24 hours after controlled cortical impact in postnatal day 17 rats. J. Neurotrauma 19, 715–740.
63. Kobeissy, F.H., Ottens, A.K., Zhang, Z., Liu, M.C., Denslow, N.D., Dave, J.R., Tortella, F.C., Hayes, R.L., and Wang, K.K. (2006). Novel differential neuroproteomics analysis of traumatic brain injury in rats. Mol. Cell. Proteomics 5, 1887–1898.
64. Haqqani, A.S., Hutchison, J.S., Ward, R., and Stanimirovic, D.B. (2007). Biomarkers and diagnosis; protein biomarkers in serum of pediatric patients with severe traumatic brain injury identified by ICAT-LC-MS/MS. J. Neurotrauma 24, 54–74.
65. Crawford, F., Crynen, G., Reed, J., Mouzon, B., Bishop, A., Katz, B., Ferguson, S., Phillips, J., Ganapathi, V., Mathura, V., Roses, A., and Mullan, M. (2012). Identification of plasma biomarkers of TBI outcome using proteomic approaches in an APOE mouse model. J. Neurotrauma 29, 246–260.
66. Cortes, D.F., Landis, M.K., and Ottens, A.K. (2012). High-capacity peptide-centric platform to decode the proteomic response to brain injury. Electrophoresis 33, 3712–3719.
67. Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., and Yamanishi, Y. (2008). KEGG for linking genomes to life and the environment. Nucleic Acids Res. 36, D480–D484.
68. Croft, D., O'Kelly, G., Wu, G., Haw, R., Gillespie, M., Matthews, L., Caudy, M., Garapati, P., Gopinath, G., Jassal, B., Jupe, S., Kalatskaya, I., Mahajan, S., May, B., Ndegwa, N., Schmidt, E., Shamovsky, V., Yung, C., Birney, E., Hermjakob, H., D'Eustachio, P., and Stein, L. (2011). Reactome: a database of reactions, pathways, and biological processes. Nucleic Acids Res. 39, D691–D697.
69. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A. 102, 15545–15550.
70. Pico, A.R., Kelder, T., van Iersel, M.P., Hanspers, K., Conklin, B.R., and Evelo, C. (2008). WikiPathways: pathway editing for the people. PLoS Biol. 6, e184.
71. Werner, T. (2008). Bioinformatics applications for pathway analysis of microarray data. Curr. Opin. Biotechnol. 19, 50–54.
72. Ackermann, M., and Strimmer, K. (2009). A general modular framework for gene set enrichment analysis. BMC Bioinformatics 10, 47.
73. Huang, da W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4, 44–57.
74. Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstrale, M., Laurila, E., Houstis, N., Daly, M.J., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B., Lander, E.S., Hirschhorn, J.N., Altshuler, D., and Groop, L.C. (2003). PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. 34, 267–273.
75. Tarca, A.L., Draghici, S., Khatri, P., Hassan, S.S., Mittal, P., Kim, J.S., Kim, C.J., Kusanovic, J.P., and Romero, R. (2009). A novel signaling pathway impact analysis. Bioinformatics 25, 75–82.
76. Dutta, B., Wallqvist, A., and Reifman, J. (2012). PathNet: A tool for pathway analysis using topological information. Source Code Biol. Med. 7, 10.
77. Kobeissy, F.H., Sadasivan, S., Oli, M.W., Robinson, G., Larner, S.F., Zhang, Z., Hayes, R.L., and Wang, K.K. (2008). Neuroproteomics and systems biology-based discovery of protein biomarkers for traumatic brain injury and clinical validation. Proteomics Clin. Appl. 2, 1467–1483.
78. Chen, K.D., Chang, P.T., Ping, Y.H., Lee, H.C., Yeh, C.W., and Wang, P.N. (2011). Gene expression profiling of peripheral blood leukocytes identifies and validates ABCB1 as a novel biomarker for Alzheimer's disease. Neurobiol. Dis. 43, 698–705.
79. Crawford, F.C., Wood, M., Ferguson, S., Mathura, V.S., Faza, B., Wilson, S., Fan, T., O'Steen, B., Ait-Ghezala, G., Hayes, R., and Mullan, M.J. (2007). Genomic analysis of response to traumatic brain injury in a mouse model of Alzheimer's disease (APPsw). Brain Res. 1185, 45–58.
80. Ideker, T., and Sharan, R. (2008). Protein networks in disease. Genome Res. 18, 644–652.
81. Uetz, P., Giot, L., Cagney, G., Mansfield, T.A., Judson, R.S., Knight, J.R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J.M. (2000). A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. Nature 403, 623–627.
82. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. Proc. Natl. Acad. Sci. U. S. A. 98, 4569–4574.
83. Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D.S., Zhang, L.V., Wong, S.L., Franklin, G., Li, S., Albala, J.S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R.S., Vandenhaute, J., Zoghbi, H.Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M.E., Hill, D.E., Roth, F.P., and Vidal, M. (2005). Towards a proteome-scale map of the human protein-protein interaction network. Nature 437, 1173–1178.

84. Stelzl, U., Worm, U., Lalowski, M., Haenig, C., Brembeck, F.H., Goehler, H., Stroedicke, M., Zenkner, M., Schoenherr, A., Koeppen, S., Timm, J., Mintzlaff, S., Abraham, C., Bock, N., Kietzmann, S., Goedde, A., Toksoz, E., Droege, A., Krobitsch, S., Korn, B., Birchmeier, W., Lehrach, H., and Wanker, E.E. (2005). A human protein-protein interaction network: a resource for annotating the proteome. Cell 122, 957–968.
85. Venkatesan, K., Rual, J.F., Vazquez, A., Stelzl, U., Lemmens, I., Hirozane-Kishikawa, T., Hao, T., Zenkner, M., Xin, X., Goh, K.I., Yildirim, M.A., Simonis, N., Heinzmann, K., Gebreab, F., Sahalie, J.M., Cevik, S., Simon, C., de Smet, A.S., Dann, E., Smolyar, A., Vinayagam, A., Yu, H., Szeto, D., Borick, H., Dricot, A., Klitgord, N., Murray, R.R., Lin, C., Lalowski, M., Timm, J., Rau, K., Boone, C., Braun, P., Cusick, M.E., Roth, F.P., Hill, D.E., Tavernier, J., Wanker, E.E., Barabasi, A.L., and Vidal, M. (2009). An empirical framework for binary interactome mapping. Nat. Methods 6, 83–90.
86. Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature 422, 198–207.
87. Ewing, R.M., Chu, P., Elisma, F., Li H., Taylor, P., Climie, S., McBroom-Cerajewski, L., Robinson, M.D., O'Connor, L., Li, M., Taylor, R., Dharsee, M., Ho, Y., Heilbut, A., Moore, L., Zhang, S., Ornatsky, O., Bukhman, Y.V., Ethier, M., Sheng, Y., Vasilescu, J., Abu-Farha, M., Lambert, J.P., Duewel, H.S., Stewart, I.I., Kuehl, B., Hogue, K., Colwill, K., Gladwish, K., Muskat, B., Kinach, R., Adams, S.L., Moran, M.F., Morin, G.B., Topaloglou, T., and Figeys, D. (2007). Large-scale mapping of human protein-protein interactions by mass spectrometry. Mol. Syst. Biol. 3, 89.
88. Yu, H., Braun, P., Yildirim, M.A., Lemmens, I., Venkatesan, K., Sahalie, J., Hirozane-Kishikawa, T., Gebreab, F., Li, N., Simonis, N., Hao, T., Rual, J.F., Dricot, A., Vazquez, A., Murray, R.R., Simon, C., Tardivo, L., Tam, S., Svrzikapa, N., Fan, C., de Smet, A.S., Motyl, A., Hudson, M.E., Park, J., Xin, X., Cusick, M.E., Moore, T., Boone, C., Snyder, M., Roth, F.P., Barabasi, A.L., Tavernier, J., Hill, D.E., and Vidal, M. (2008). High-quality binary protein interaction map of the yeast interactome network. Science 322, 104–110.
89. Yu, X., Ivanic, J., Memisevic, V., Wallqvist, A., and Reifman, J. (2011). Categorizing biases in high-confidence high-throughput protein-protein interaction data sets. Mol. Cell. Proteomics 10, M111.012500.
90. Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H., and Michnick, S.W. (2008). An *in vivo* map of the yeast protein interactome. Science 320, 1465–1470.
91. von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S.G., Fields, S., and Bork, P. (2002). Comparative assessment of large-scale data sets of protein-protein interactions. Nature 417, 399–403.
92. Ivanic, J., Wallqvist, A., and Reifman, J. (2008). Probing the extent of randomness in protein interaction networks. PLoS Comput. Biol. 4, e1000114.
93. Yu, X., Ivanic, J., Wallqvist, A., and Reifman, J. (2009). A novel scoring approach for protein co-purification data reveals high interaction specificity. PLoS Comput. Biol. 5, e1000515.
94. Rajagopala, S.V., Hughes, K.T., and Uetz, P. (2009). Benchmarking yeast two-hybrid systems using the interactions of bacterial motility proteins. Proteomics 9, 5296–5302.
95. Ivanic, J., Wallqvist, A., and Reifman, J. (2008). Evidence of probabilistic behaviour in protein interaction networks. BMC Syst. Biol. 2, 11.
96. Minnhagen, P., and Bernhardsson, S. (2008). The blind watchmaker network: scale-freeness and evolution. PLoS One 3, e1690.
97. Jeronimo, C., Forget, D., Bouchard, A., Li, Q., Chua, G., Poitras, C., Therien, C., Bergeron, D., Bourassa, S., Greenblatt, J., Chabot, B., Poirier, G.G., Hughes, T.R., Blanchette, M., Price, D.H., and Coulombe, B. (2007). Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. Mol. Cell 27, 262–274.
98. Wang, J., Huo, K., Ma, L., Tang, L., Li, D., Huang, X., Yuan, Y., Li, C., Wang, W., Guan, W., Chen, H., Jin, C., Wei, J., Zhang, W., Yang, Y., Liu, Q., Zhou, Y., Zhang, C., Wu, Z., Xu, W., Zhang, Y., Liu, T., Yu, D., Zhang, Y., Chen, L., Zhu, D., Zhong, X., Kang, L., Gan, X., Yu, X., Ma, Q., Yan, J., Zhou, L., Liu, Z., Zhu, Y., Zhou, T., He, F., and Yang, X. (2011). Toward an understanding of the protein interaction network of the human liver. Mol. Syst. Biol. 7, 536.
99. Yu, X., Wallqvist, A., and Reifman, J. (2012). Inferring high-confidence human protein-protein interactions. BMC Bioinformatics 13, 79.
100. Ulitsky, I., Krishnamurthy, A., Karp, R.M., and Shamir, R. (2010). DEGAS: de novo discovery of dysregulated pathways in human diseases. PLoS One 5, e13367.
101. Ideker, T., Ozier, O., Schwikowski, B., and Siegel, A.F. (2002). Discovering regulatory and signalling circuits in molecular interaction networks. Bioinformatics 18, Suppl. 1, S233–S240.
102. Chuang, H.Y., Lee, E., Liu, Y.T., Lee, D., and Ideker, T. (2007). Network-based classification of breast cancer metastasis. Mol. Syst. Biol. 3, 140.
103. Begley, T.J., Rosenbach, A.S., Ideker, T., and Samson, L.D. (2004). Hot spots for modulating toxicity identified by genomic phenotyping and localization mapping. Mol. Cell 16, 117–125.
104. Ulitsky, I., and Shamir, R. (2007). Identification of functional modules using network topology and high-throughput data. BMC Syst. Biol. 1, 8.
105. Breitling, R., Amtmann, A., and Herzyk, P. (2004). Graph-based iterative Group Analysis enhances microarray interpretation. BMC Bioinformatics 5, 100.
106. Dittrich, M.T., Klau, G.W., Rosenwald, A., Dandekar, T., and Muller, T. (2008). Identifying functional modules in protein-protein interaction networks: an integrated exact approach. Bioinformatics 24, i223–i231.
107. Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., Hanspers, K., Isserlin, R., Kelley, R., Killcoyne, S., Lotia, S., Maere, S., Morris, J., Ono, K., Pavlovic, V., Pico, A.R., Vailaya, A., Wang, P.L., Adler, A., Conklin, B.R., Hood, L., Kuiper, M., Sander, C., Schmulevich, I., Schwikowski, B., Warner, G.J., Ideker, T., and Bader, G.D. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. 2, 2366–2382.
108. Ulitsky, I., Maron-Katz, A., Shavit, S., Sagir, D., Linhart, C., Elkon, R., Tanay, A., Sharan, R., Shiloh, Y., and Shamir, R. (2010). Expander: from expression microarrays to networks and functions. Nat. Protoc. 5, 303–322.
109. Workman, C.T., Mak, H.C., McCuine, S., Tagne, J.B., Agarwal, M., Ozier, O., Begley, T.J., Samson, L.D., and Ideker, T. (2006). A systems approach to mapping DNA damage response pathways. Science 312, 1054–1059.
110. Lee, E., Chuang, H.Y., Kim, J.W., Ideker, T., and Lee, D. (2008). Inferring pathway activity toward precise disease classification. PLoS Comput. Biol. 4, e1000217.
111. Valsesia, A., Rimoldi, D., Martinet, D., Ibberson, M., Benaglio, P., Quadroni, M., Waridel, P., Gaillard, M., Pidoux, M., Rapin, B., Rivolta, C., Xenarios, I., Simpson, A.J., Antonarakis, S.E., Beckmann, J.S., Jongeneel, C.V., Iseli, C., and Stevenson, B.J. (2011). Network-guided analysis of genes with altered somatic copy number and gene expression reveals pathways commonly perturbed in metastatic melanoma. PLoS One 6, e18369.
112. Liu, M., Liberzon, A., Kong, S.W., Lai, W.R., Park, P.J., Kohane, I.S., and Kasif, S. (2007). Network-based analysis of affected biological processes in type 2 diabetes models. PLoS Genet. 3, e96.
113. Ma, X., Lee, H., Wang, L., and Sun, F. (2007). CGI: a new approach for prioritizing genes by combining gene expression and protein-protein interaction data. Bioinformatics 23, 215–221.
114. Carmona-Saez, P., Chagoyen, M., Tirado, F., Carazo, J.M., and Pascual-Montano, A. (2007). GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. Genome Biol. 8, R3.
115. Nogales-Cadenas, R., Carmona-Saez, P., Vazquez, M., Vicente, C., Yang, X., Tirado, F., Carazo, J.M., and Pascual-Montano, A. (2009). GeneCodis: interpreting gene lists through enrichment analysis and integration of diverse biological information. Nucleic Acids Res. 37, W317–W322.
116. Becker, K.G., Barnes, K.C., Bright, T.J., and Wang, S.A. (2004). The genetic association database. Nat. Genet. 36, 431–432.
117. Ikonomovic, M.D., Uryu, K., Abrahamson, E.E., Ciallella, J.R., Trojanowski, J.Q., Lee, V.M., Clark, R.S., Marion, D.W., Wisniewski, S.R., and DeKosky, S.T. (2004). Alzheimer's pathology in human temporal cortex surgically excised after severe brain injury. Exp. Neurol. 190, 192–203.

118. Jordan, B.D. (2000). Chronic traumatic brain injury associated with boxing. Semin. Neurol. 20, 179–185.
119. Guskiewicz, K.M., Marshall, S.W., Bailes, J., McCrea, M., Cantu, R.C., Randolph, C., and Jordan, B.D. (2005). Association between recurrent concussion and late-life cognitive impairment in retired professional football players. Neurosurgery 57, 719–726.
120. Capdeville, R., Buchdunger, E., Zimmermann, J., and Matter, A. (2002). Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. Nat. Rev. Drug Discov. 1, 493–502.
121. Schlatterer, S.D., Acker, C.M., and Davies, P. (2011). c-Abl in neurodegenerative disease. J. Mol. Neurosci. 45, 445–452.
122. van de Vijver, M.J., He, Y.D., van't Veer, L.J., Dai, H., Hart, A.A., Voskuil, D.W., Schreiber, G.J., Peterse, J.L., Roberts, C., Marton, M.J., Parrish, M., Atsma, D., Witteveen, A., Glas, A., Delahaye, L., van der Velde, T., Bartelink, H., Rodenhuis, S., Rutgers, E.T., Friend, S.H., and Bernards, R. (2002). A gene-expression signature as a predictor of survival in breast cancer. N. Engl. J. Med. 347, 1999–2009.
123. Wang, Y., Klijn, J.G., Zhang, Y., Sieuwerts, A.M., Look, M.P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M.E., Yu, J., Jatkoe, T., Berns, E.M., Atkins, D., and Foekens, J.A. (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365, 671–679.
124. Sharan, R., and Ideker, T. (2006). Modeling cellular machinery through biological network comparison. Nat. Biotechnol. 24, 427–433.
125. Stuart, J.M., Segal, E., Koller, D., and Kim, S.K. (2003). A gene-coexpression network for global discovery of conserved genetic modules. Science 302, 249–255.
126. Zinman, G.E., Zhong, S., and Bar-Joseph, Z. (2011). Biological interaction networks are conserved at the module level. BMC Syst. Biol. 5, 134.

Address correspondence to:
Jaques Reifman, PhD
U.S. Army Medical Research and Materiel Command
MCMR-TT
504 Scott Street
Fort Detrick, MD 21702

E-mail: Jaques.Reifman.civ@mail.mil