Boja et al. Clinical Proteomics 2014, 11:45
http://www.clinicalproteomicsjournal.com/content/11/1/45

CLINICAL PROTEOMICS

MEETING REPORT Open Access

Integration of omics sciences to advance biology and medicine

Emily S Boja[1*], Christopher R Kinsinger[1*], Henry Rodriguez[1] and Pothur Srinivas[2*] on behalf of Omics Integration Workshop Participants

Abstract

In the past two decades, our ability to study cellular and molecular systems has been transformed through the development of omics sciences. While unlimited potential lies within massive omics datasets, the success of omics sciences to further our understanding of human disease and/or translating these findings to clinical utility remains elusive due to a number of factors. A significant limiting factor is the integration of different omics datasets (i.e., integromics) for extraction of biological and clinical insights. To this end, the National Cancer Institute (NCI) and the National Heart, Lung and Blood Institute (NHLBI) organized a joint workshop in June 2012 with the focus on integration issues related to multi-omics technologies that needed to be resolved in order to realize the full utility of integrating omics datasets by providing a glimpse into the disease as an integrated "system". The overarching goals were to (1) identify challenges and roadblocks in omics integration, and (2) facilitate the full maturation of 'integromics' in biology and medicine. Participants reached a consensus on the most significant barriers for integrating omics sciences and provided recommendations on viable approaches to overcome each of these barriers within the areas of technology, bioinformatics and clinical medicine.

Keywords: Omics integration, Omics science, Clinical application, Risk prediction, Proteomics, Metabolomics, Genomics

Introduction

The past two decades have been witness to an explosion of data stemming from the development and gradual maturation of 'omics' technologies and bioinformatics. Today, whole-genome sequencing has become a routine research tool, and state-of-the-art proteomic technologies have caught up to genomics in the past few years in terms of coverage as evidenced by their ability to identify a large percentage of all observed human gene products, including functionally significant alternative splice variants [1-4]. Nevertheless, the omics mindset has not yet permeated the broad biological and clinical community. Of the ~20,000 genes in the human genome, only 10% have 5 or more publications [5], while one gene, p53 that regulates the cell cycle and functions as a tumor suppressor, is the subject of over 56,000 articles in scientific literature. Clearly, our technological abilities to generate large amounts of data from molecular systems have advanced enormously, but the ability to translate this information for use in the clinic remains elusive due to a number of factors. One key reason postulated is that while individual omics domains yield distinct and important information, no single omics science is sufficient to facilitate a comprehensive understanding of the complex human biology and physiology. Additionally, there are logical scientific steps missing in leaping from a lack of information on 90% of the proteins to clinical use. The integration of omics sciences bioinformatically remains a challenge and thus a limiting factor in fully extracting biological meaning from the mounds of data being generated. For instance, the NCI's The Cancer Genome Atlas (TCGA) integrated multiple data types to identify three mutually exclusive pathways that affect the development of glioblastoma multiforme (e.g., RTK, TP53, RB) [6], suggesting that the presence of one aberration removes the selective pressure for a second aberration. This example demonstrates the immediate value of data integration since these pathways were not observed from data in isolation (either from mutations, copy number

* Correspondence: bojae@mail.nih.gov; kinsingc@mail.nih.gov; srinivap@nhlbi.nih.gov
[1]Office of Cancer Clinical Proteomics Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
[2]Division of Cardiovascular Sciences, National Heart, Lung and Blood Institute, Bethesda, MD, USA

changes, or other measurements). Omics integration is the next logical and necessary step in propelling systems biology and medicine forward and potentially allowing for its use in the clinic. NCI's Clinical Proteomic Tumor Analysis Consortium (CPTAC) is one such multi-institutional initiative that employs proteogenomic integration to better enhance our understanding of cancer biology using genomically characterized tumors [7], and there are similar international efforts such as uniting the chromosome-centric human proteome project with the Encyclopedia of DNA Elements (ENCODE) [8].

Executive summary

In light of previous workshops addressing the challenges and opportunities of clinical proteomics in biology and medicine [9,10] and the advancement of proteogenomic science, the NCI and NHLBI organized a workshop focusing the topic of integrating omics datasets obtained from multi-omics technologies to provide broader insights into disease pathophysiology. The workshop was held on the National Institute of Health (NIH) campus in Bethesda, MD on June 19 and 20, 2012 with participants from a diverse variety of scientific expertise. Herein, this report summarizes the major challenges and proposes solutions for omics integration in an effort to raise support and awareness of omics integration within the scientific community. It is hoped that this report will initiate new collaborative efforts that harness the vast amount of knowledge embedded in disparate data sets and promote training of more multidisciplinary scientists better positioned in the science of omics integration (integromics).

Workshop overview

To identify key limiting factors and challenges in integromics and provide actionable solutions to overcome such roadblocks in the context of biology and diseases, the workshop was structured to ground discussions upon three case studies - personal omics profiling [11], multi-omics pathway analysis of cardiovascular-specific circadian clock [12], and glycoproteomics [13]. In addition, experts from the Framingham Heart Study presented a "lessons learned" talk on identifying risk factors for heart disease and its associated studies using omics-based technologies on a much larger patient population [14,15]. Next, workshop participants broke off into multi-disciplinary groups for further discussion in order to develop integrative solutions to address three major areas of challenges (clinical, informatics, and technology) identified. For example, questions were raised by the participants during rounds of discussions, including: (1) Can omics improve the odds ratio for diabetes or heart disease prediction in cardiovascular research? (2) Can omics science provide the context for cancers that begins as genetic aberrations? Collectively, six major recommendations for facilitating omics integration were put forth and summarized below.

Case studies

Personal omics profiling (case study 1)

The case study described by Dr. Michael Snyder from Stanford University illustrated how integration of different omics data can facilitate a shift from disease treatment to prevention based on his own experience. Discussed was how longitudinal personalized omics profiling (POP) from analysis of the genome, epigenome, transcriptome, proteome and metabolome ("Snyderome") can collectively provide useful information that otherwise could not be gleaned from any single individual omics domain (data sets) alone. The "Snyderome" included routine measurements interspersed with dense sampling during states of infection. Integrative analyses of the data revealed an increased insulin biosynthetic pathway that spiked during states of viral infections [11]. The data further indicated Dr. Michael Snyder to be at an increased risk of type 2 diabetes, despite having no known family history of the disease, which subsequently proved true. This highlights the fact that following multiple omics components longitudinally may provide valuable information about disease risk, drug sensitivity, and other components of personalized medicine.

This POP study simultaneously illustrated the potential of omics integration. Clearly, methods exist to shift less studied areas of medicine from hearsay and conjecture to data-established-truth. Yet, POP studies are hardly scalable across a population due to an analysis cost of $10,000 per sample. Furthermore, progress in POP research requires people to allow the collection of their omics profiles. This is a delicate subject as the collection of so much data will increase the likelihood of false positives and induce undue or premature emotional strain. The so-called, "democratization of data", namely the shift from expert protectionism to people governing their own data, has led to the possibility of better decision-making which might significantly impact the choices they make day-to-day. Although this can be done in medicine, the challenge remains to protect human subjects without hindering research, while restraining clinical adoption until clear data-driven-truths have been clinically validated.

Pathways and targets to modulate clocks (case study 2)

Dr. John Hogenesch from University of Pennsylvania discussed the utility of omics integration to identify clock-modifying genes and pathways. The circadian clock regulates many aspects of biology, including core body temperature, organ function, heart rate, and blood pressure, among others. Clocks are present in most of the

body's cells and interestingly most cancers appear to have lost their circadian clocks.

Omics approaches that include whole-genome siRNA circadian genomic screens, gene expression data, and protein-protein interaction data are used to identify clock-modifying genes and define their mechanistic and functional attributes [16]. The insulin signaling pathway is one of the most significant clock-modifying pathways identified by such an approach. Dr. John Hogenesch discussed the use of Bayesian integration strategies to help assess whether the evidence provided by a given result indicates that the gene is a core clock component. Additional discussion on major challenges for integrating omics results include the use of different synonyms by the scientific community (e.g., multiple names for a given gene and/or its variants, and access to high-quality standardized data sets for "trustworthy" analyses).

Glycoproteomics (case study 3)

Drs. Gerald Hart and Jennifer Van Eyk from Johns Hopkins University discussed the fields of glycobiology, highlighting the critical nature of integrative approaches since one omics domain cannot adequately explain the underlying biology. Dr. Gerald Hart estimated that 90% of proteins are glycosylated, and glycosylation is involved in nearly all cellular activities and metabolic processes. He also noted that post-translational modifications (PTMs), such as glycosylation, greatly expand the genetic code's chemical diversity, and hence, function cannot be inferred through genomics approaches alone. "Glycomics" is defined as the study to characterize or quantify the glycome of a cell, tissue, or organ. Glycome complexity is a reflection of cellular complexity and the collective tools of genomics, proteomics, lipidomics and metabolomics are required for functional characterization. Challenges to the integration of glycomics include a lack of integration of glycan data into mainstream databases, a lack of standardization across existing glycomic databases, and a lack of clarity regarding different levels of glycan "structure" in published literature. A further challenge is the paucity of measurement tools for site-specific identification and quantitation of glycoproteomics.

The Framingham heart study (lessons learned)

The Framingham Heart Study was initiated in Framingham, Massachusetts in 1948 to understand the underlying causes of cardiovascular disease (CVD). The study aimed to investigate the expression of coronary disease in a normal population, determine factors that predispose individuals to develop CVD, and evaluate new screening tests (e.g., electrocardiography, blood metabolites). Currently, the Framingham Study incorporates a systems biology approach to biomarker research [i.e., CVD Systems Approach to Biomarker Research (SABRe) initiative], aiming to identify biomarker signatures of CVD and its major risk factors using omics technologies. Dr. Andrew Johnson from NHLBI summarized omics data collected to date, in which studies have profiled three generations of families across thousands of phenotypes with many of them being longitudinal. Specific data collected include 8,500 genome-wide association studies, 7,000 cell line analyses, 300 whole exome sequences, 1,000 whole-genome sequences, 5,000 DNA microarrays, 2,000 metabolomics analyses, and ongoing data collection with induced pluripotent stem cells, DNA methylation, computed tomography scans, and magnetic resonance imaging. Challenges identified in the Framingham Heart Study include data acquisition (e.g., throughput, cost, and sample tracking/batch effects), storage (e.g., results, storage demands, raw data in one place for cross-comparison, etc.), and limitations with data processors, competing needs on servers, costly renewal of outdated resources, and security issues.

Roadblocks in integrating omics knowledge in biology and medicine

Discussions regarding roadblocks and challenges in omics science that took place following the presented case studies are outlined below with a focus on three main areas - clinical utility, informatics, and technology.

Clinical utility challenges

Two fundamental challenges that were identified for the integration of omics into medicine included (1) disseminating, managing, and interpreting omics data in a clinical context, and (2) ensuring that omics results have added value to existing paradigms of patient care. Providing a solution to these problems should allow for enhanced preventative, diagnostic, and prognostic procedures [17]. The democratization of multi-omics data is a key aspect of the integration of omics data in medicine. While the physical barriers to access, management, and transfer of data have been removed through the digitalization of data files, clinical utility of research data is limited by privacy and other barriers, justly placed to prohibit the abuse of protected health information. However, the ease of disseminating, managing, and interpreting massive amounts of omics data would allow for quicker application of integrative omics knowledge to clinical practice.

Transforming and incorporating data derived from different omics approaches into a defined clinical context is essential, but remains complex and problematic [18,19]. Genomic scans, for example, have started to identify more and rarer variants in addition to common SNP variants [20], and when different commercial platforms are used to molecularly analyze a common sample, variability is often found in their risk prediction capacities [21]. This variability most likely lies in data interpretation models that incorporate different assumptions during data processing and widespread problems of overfitting high dimensional data with an extremely large number of

molecular measurements relative to limited sample size [19]. This begs the question of how well a genetic variant correlates to a specific disease condition and whether predicted disease risks have any clinical validity. In the age of declining genotyping costs and retail genome sequencing kit, consumers can now obtain data on their own personal DNA, and patient expectations of clinicians providing useful genetic information are soaring. Therefore, a disconnect is growing between the realistic, operable utilities of omics sciences and the expectations of patients with little clarity on how to bridge the gap. Finally, legitimate concerns about how to keep data and results private and secure are becoming more prominent.

The second major clinical challenge lies in determining, through appropriate studies, whether the new omics findings add incremental value to current clinical practices or clinical decision making. While multiple omics technologies can potentially discover a host of biological candidates from samples, their clinical utility requires rigorous validation. Hence, discovery-based omics research should seek to maximize the signal-to-noise ratio of a biomarker candidate(s) in order to produce fewer false leads [19]. Furthermore, it is important to distinguish the causes of pathogenesis versus markers that indicate disease phenotypes, since causes are often treatable and have robust associations (e.g., LDL and atherosclerosis [22]), whereas markers of disease are the often most powerful predictors. Although the markers of diseases can guide diagnosis and treatment, their effects are not a direct target for treatment (e.g., you can treat LDL, but you do not treat Troponin). Cholesterol was studied for over 100 years prior to becoming a clinically useful biomarker. However, it is uncertain that any new biomarker candidates from omics studies alone or in combination to cholesterol perform better than cholesterol alone. Such complex barriers need to be adequately addressed to be of help in actionable clinical decision-making.

Informatics challenges

Three major challenges identified in informatics that limit the integration of omics data in the clinic were (1) the development of more mature models of cellular processes that incorporate non-commensurate omics data types [23,24], (2) data storage limitations and organization of fragmented data sets, and (3) a shortage of multidisciplinary scientists with training in biology, computer science, informatics and statistics.

Omics integration includes the incorporation of multiple omics data types into a comprehensive model that accurately describes biological processes. The simplest model assumes the "central dogma" and maps transcripts and proteins to gene sequences. Slightly more sophisticated models entail quantitative information and use correlations across molecular entities. As each "ome" reflects a distinct biological domain (e.g., transcripts, proteins, metabolites), the resulting datasets represent the measurements of various underlying variables on different scales. For example, transcriptional and translational profiles for mRNA transcripts and corresponding proteins are often but not always the same [25-27]. To capture both the temporal and spatial dynamics of biomolecules embedded within complex biological relationships, the most complex models must appropriately integrate all pertinent, distinct measurements of the various Omes. However, the modeling of non-commensurate data types comprised of nonlinear relationships and multivariate signals is extremely complex, and current computational algorithms and statistical procedures are limited in this capacity. Additionally, the non-synonymous naming systems for the myriad of biological molecules in the various Omes further complicate algorithm development and inhibit omics integration. As discussed previously, modeling would be greatly aided by the standardization of gene names (e.g., circadian clock genes). Once a model is established, faster and more efficient methods are required to validate computational results in cellular and animal model systems, representing a huge challenge in the field of integrative omics science [28].

This specific challenge is particularly difficult to address, involving many aspects of the scientific and clinical disciplines dependent on the diseases, including but not limited to:

a) relative risk of disease or adverse outcome is often arbitrarily assigned,
b) association does not necessarily equal prediction,
c) insufficient sample numbers in some studies,
d) difficult to extrapolate from n = 1 to a population and to model the environment, and
e) modeling needs to be performed by computers and not by physicians, with results translated to a scale that physicians can easily understand (e.g., 10-year coronary heart disease risk).

The second bioinformatics challenge for omics integration involves the storage of large, heterogeneous datasets generated from multiple high-throughput omics platforms. With the continued development of more sophisticated instrumentation for data acquisition, the amount of data generated is exponentially rising, along with the demand for data storage. As the usage of stored data occurs at distinct levels (e.g., raw data vs. mass spectrometry search results files in proteomics, or raw nucleotide sequence reads vs. variant calls in vcf format in genomics) specific to a particular expertise in the multidisciplinary end user pool (e.g., computer scientists vs. genome biologists), data storage infrastructure should be stratified and specifically tailored to meet the needs of end

users. If storing all data is cost-prohibitive, the difficulty lies in determining which data are the most valuable to keep. Furthermore, datasets are heterogeneous with respect to both intra-omics (e.g., proteomic datasets from different file formats) and inter-omics (e.g., genomic vs. proteomic datasets) acquisition protocols. This results in a storage infrastructure that is fragmented and disjointed, thereby hindering cross-comparison and retrograde use by the scientific community. Security and privacy of stored clinical data is an additional issue for avoiding ethical concerns.

The participants collectively put forth recommendations to overcome informatics barriers by:

a) establishing data standards for all types of omics data files (e.g., cite genomics and proteomics papers),
b) changing access to data [29] to protect research subjects without hindering valuable research opportunities,
c) completing the incomplete reference databases (~1/3 of SNPs in dbSNP), such as using proteomics data to confirm/verify gene annotation [30], and adding PTMs that are not routinely integrated in mainstream databases,
d) calculating some key parameters for data processing and storage, such as how many times will a raw file be processed? How long will it need to be stored? How frequently do data analysis methods change?
e) providing sufficient incentive to data generators for data deposition into publicly accessible repositories although great stride has been made in the past few years such as dbGAP and ProteomeXchange [31], and
f) overcoming data storage and computing power limitations.

The third major bioinformatics challenge is primarily driven by technology. Rapidly evolving analytical methods unleash new measurements which in turn give rise to new types of data and data analysis. Hence, there is a constant requirement for scientists including bioinformaticians to keep up with the developing technologies and methodologies. Most experts in the field have experience in a single omics technology, such as calling mutations in next-generation sequencing data or extracting peptides from mass spectra, and those who specialize in the next higher level of data integration are rare. A combination of reasons contribute to this dearth including: rapidly changing technologies that keep bioinformaticians from continually specializing in the analysis of one molecular moiety, insufficient biomedical informatics training opportunities, and the transient nature of the interface between technology development and disease-specific research. Major adjustments to the vision and expanding the training of medical bioinformatics research community are highly recommended and required to surpass these obstacles, even though informatics training opportunities related to NIH's BD2K initiative and others have been added more recently to address this challenge.

Technological challenges

Two major technological challenges that were recognized to limit omics integration into medicine were (1) a lack of reproducibility of data acquired through non-uniformly standardized sample preparation, including a lack of understanding of the impact of pre-analytical variables on samples [32], and inconsistent instrument performance [19], and (2) a lack of high-throughput and multiplexing methods that make parallel measurements of multiple types of analytes for handling large clinical studies. Addressing such obstacles, the scientific community has come a long way to demonstrate the analytical robustness of genomic, proteomic, and metabolomic workflows, including data analysis pipelines as witnessed by a flurry of standardization/harmonization activities during the last two decades in several omics areas including Genomic Standards Consortium, CPTAC, HUPO and ABRF [33-40]. Furthermore, there have been significant technological advances in measuring genomic variants, proteins and peptides, and small molecule metabolites that include next-generation genomic sequencing, immuno-multiple reaction monitoring mass spectrometry, flow cytometry, and protein microarrays [41-44]. There is no doubt that technologies will continue to be improved/developed to increase sensitivity, specificity and throughput, making it feasible to measure every molecule at the single cell level. To apply multiplexing and high throughput methods in clinical studies, researchers need to ensure that the appropriate technologies/platforms and bioinformatic analyses are analytically robust and standardized, and can be validated in an independent lab and/or in a separate set of clinical samples.

Recommendations for successful omics integration

Following rounds of discussions, six major recommendations for facilitating omics integration to address the identified roadblocks described above were put forth by workshop participants and summarized below.

1) Committed funding for the education of multi-disciplinary teams is needed. Clinicians, clinical scientists, basic scientists, and bioinformaticians need to be educated in these disciplines, and form collaborative, multi-disciplinary teams to carry out omics integration from discovery to the patient. Omics sciences are inherently integrative of multiple

specialties. Therefore, all phases of discovery efforts, including sample procurement, experimental design and bio-interpretation, and all phases of clinical translation including clinical trials and implementation into clinical procedures must be performed by a multi-disciplinary team of investigators. From this, appropriate epidemiological and statistical measures should be applied for determining whether a newly discovered marker or panel of markers adds value to pre-existing clinical regimes of risk prediction, diagnosis and prognosis. Furthermore, end users need to be educated on the realistic utilities of omics results at each stage of omics development. This can be accomplished via public seminars or via genetic counselors acting as a liaison between clinicians and patients. This will lessen unrealistic expectations of the public for physicians to infer patient risk from the results of omics studies. In the long term, committed funding to create a new discipline of omics sciences is needed, providing rigorous training in the omics sciences in order to create a group of specialized experts to propel the field forward. Fellowships are needed for young scientists in the field of omics sciences to train future experts. Specifically, there is a need for the development of informatics training centers that produce experts who derive meaning from large omics datasets, including data curators and wranglers.

2) Committed and sustained funding for technology development is needed. In particular, further developments are needed in mass spectrometry instruments and technologies (e.g., top-down MS) in order to sequence deeper proteomes and/or metabolomes, and to allow for high throughput multiplexed analysis.
3) Sample preparatory procedures and acquisition must be standardized to allow for reliable cross-comparison, sharing and integration of large omics datasets and for whole-omics profiling from the same sample.
4) The development of an unifying resource is needed to permanently store data in a coordinated and structured manner. This resource would provide security, privacy and consensus on how data are stored and accessed by the community. This is critical for the integration of omics sciences and one where the National Institutes of Health (NIH) can play a significant role.
5) Mature models for integrating non-commensurate data types are needed. Algorithms must be developed for data compression, integration, querying and display to handle the distinct data types of omics sciences. Quality control algorithms should be developed for data format and exchange, and natural language data mining.
6) A consensus needs to be developed in order to create validity and value for integrating omics findings into clinical guidelines. Useful, reliable and valid metrics for establishing association and prediction in diagnostic and prognostic studies need to be utilized. Moreover, calculations for diagnostic and prognostic purposes need to be locked down and automated within a laboratory in order to remove any inconsistencies stemming by physician bias or interpretation. Translating scores to a scale that physicians can understand and converting to a single scale that can be modified over time is very important in this process [19,45].

Conclusion

Omics science has transformed biology and has the potential of transforming medicine. This workshop was a first step on opening a dialogue amongst scientists and clinicians in relevant omics disciplines to (1) update recent progress and further emphasize the importance of omics science and its potential in transforming biology and future clinical practice, (2) discuss barriers in omics integration existent in a variety of forms, and (3) put forth recommendations to overcome such barriers to enable the science to move forward.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

EB carried out the majority of manuscript writing, editing and proofreading, and participated in the planning of the workshop with CK, PS and HR. CK and PS played a major role in the planning and execution of the workshop, in addition to the editing of the manuscript. HR participated in the planning and execution of the workshop and manuscript proofreading. All authors read and approved the final manuscript.

Acknowledgement

We thank all the workshop participants for their contribution to the discussions and recommendations with regards to the topic of omics integration.

Workshop participant list

Bishow B. Adhikari, Ph.D.
Program Director
Heart Failure and Arrhythmias Branch
Division of Cardiovascular Diseases
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, RKL2 BG RM 8186
Bethesda, MD 20892 USA
Phone: +1 (301) 435-0504
Fax: +1 (301) 480-7404
Email: bishow.adhikari@nih.gov
Ivor J. Benjamin, M.D.
Chair and Professor
Departments of Medicine and Biochemistry
School of Medicine

Health Sciences Center
University of Utah
30 North 1900 East, Room 4A100
Salt Lake City, UT 84132 USA
Phone: +1 (801) 587-9785
Fax: +1 (801) 585-1082
Email: ivor.benjamin@hsc.utah.edu
Aruni Bhatnagar, Ph.D.
Professor
Institute of Molecular Cardiology
Department of Medicine/Cardiology
University of Louisville
580 South Preston St.
Delia Baxter Building, Room 421 F
Louisville, KY 40202 USA
Phone: +1 (502) 852-5966
Fax: +1 (502) 852-3663
Email: aruni@louisville.edu
Emily Boja, Ph.D.
Program Manager
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 31, Suite 10A52
Bethesda, MD 20892 USA
Phone: +1 (301) 451-8883
Fax: +1 (301) 496-7808
Email: bojae@mail.nih.gov
Kimberly Bunje
Senior Administrative Analyst
NHLBI Proteomics Coordinating and Administration Center
Departments of Physiology and Medicine/Cardiology
University of California at Los Angeles
CHS 14-142
10833 LeConte Ave.
Los Angeles, CA 90095 USA
Phone: +1 (310) 825-5175
Fax: +1 (310) 267-5623
Email: kbunje@medmet.ucla.edu
Sonia L. Calcagno
Science Program Coordinator
Office of Cancer Nanotechnology Research
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
31 Center Drive, Suite 10A52, MSC 2580
Bethesda, MD 20892
Phone: +1 (301) 594-5612
Fax: +1 (301) 496-7807
Email: calcagnosl@mail.nih.gov
Josef Coresh, M.D., Ph.D., M.H.S.
Director and Professor
Cardiovascular Epidemiology & Comstock Center Departments of Epidemiology/Biostatistics
Bloomberg School of Public Health
Johns Hopkins University
2024 E. Monument Street, Suite 2-600
Baltimore, MD 21287 USA
Phone: +1 (410) 955-0495
Fax: +1 (410) 955-0476
Email: jcoresh@jhsph.edu
James M. Deleo
Section Chief
Department of Clinical Research Informatics
Scientific Computing Section
National Institutes of Health
9000 Rockville Pike, Building 10
Bethesda, MD 20892 USA
Phone: +1 (301) 496-3848
Fax: +1 (301) 496-3848
Email: jdeleo@nih.gov
Leslie K. Derr, Ph.D.
Program Director
Office of Strategic Coordination
Office of the Director
National Institutes of Health
9000 Rockville Pike, Building 1, Room 201B
Bethesda, MD 20892 USA
Phone: +1 (301) 594-8174
Fax: +1 (301) 480-6641
Email: leslie.derr@nih.gov
Valentina Di Francesco
Senior Program Officer
Bioinformatics, Structural Genomics and Systems Biology Division of Microbiology and Infectious Diseases
National Institute of Allergry and Infectious Diseases
National Institutes of Health
6610 Rockledge Dr., MSC 6603; Room 4802
Bethesda, MD 20892 USA
Phone: +1 (301) 496-1884
Fax: +1 (301) 480-4528
Email: vdifrancesco@niaid.nih.gov
Kay Fleming, Ph.D.
Writer Editor
Center for Biomedical Informatics and Information Technology
National Cancer Institute
National Institutes of Health
2115 E. Jefferson Street, Room 6047
Rockville, MD 20852
Phone: +1 (301) 594-3602
Email: flemingl@mail.nih.gov
Nancy Fournier, Ph.D., M.B.A.
Director
Génome Québec
630, boul. René-Lévesque Ouest
Bureau 2660
Montréal, Québec QC H3B 1S6 Canada
Phone: +1 (514) 398-0668 x224
Fax: +1 (514) 398-0883
Email: nfournier@genomequebec.com
Weiniu Gan, Ph.D.
Program Director
Division of Airway Biology and Disease
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, Room 10164
Bethesda, MD 20892 USA
Phone: +1 (301) 435-0202
Fax: +1 (301) 480-1336
Email: ganw2@mail.nih.gov
Scott Geromanos
Waters Corporation
5 Technology Drive
Milford, MA 01757 USA
Phone: +1 (508) 482-2904
Fax: +1 (508) 482-4524
Email: scott_geromanos@waters.com
Morgan Giddings, Ph.D.
Research Professor
Department of Biochemistry and Biophysics
College of Arts and Sciences
Boise State University
1910 University Dr., Boise, ID 83725 USA
Phone: +1 (919) 240-7007
Fax: +1 (919) 240-7356
Email: morgan@giddingslab.org
Charles A. Goldthwaite Jr, Ph.D.
Science Writer
Goldthwaite & Associates
254 Leo Avenue
Shreveport, LA 71105 USA

Phone: +1 (318) 865-5058
Fax: +1 (318) 865-5058
Email: charlesgoldthwaite@gmail.com
Gerald Hart, Ph.D.
Director and Professor
Department of Biological Chemistry
School of Medicine
Johns Hopkins University
725 N. Wolfe St. 515 WBSB
Baltimore, MD 21205 USA
Phone: +1 (410) 614-5993
Fax: +1 (410) 614-8804
Email: gwhart@jhmi.edu
Henning Hermjakob
Team Leader
Proteomics Services
EMBL - European Bioinformatics Institute
Wellcome Trust Genome Campus
Hinxton, Cambridge CB10 1SD UK
Phone: +44 (1223) 49 4671
Fax: +44 (1223) 49 4468
Email: hhe@ebi.ac.uk
Joseph A. Hill, M.D., Ph.D.
Chair and Professor
Departments of Medicine / Molecular Biology
The University of Texas Southwestern Medical Center
UT Southwestern Medical Center
5323 Harry Hines Boulevard
Dallas, TX 75390 USA
Phone: +1 (214) 645-8300
Email: joseph.hill@utsouthwestern.edu
Tara Hiltke, Ph.D.
Program Manager
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 31, Suite 10A52
Bethesda, MD 20892 USA
Phone: +1 (301) 451-8883
Fax: +1 (301) 496-7808
Email: hiltket@mail.nih.gov
John B. Hogenesch, Ph.D.
Associate Professor
Department of Pharmacology
Perelman School of Medicine
University of Pennsylvania
Translational Research Center 10-124
3400 Civic Center Blvd., Bldg. 421
Philadelphia, PA 19104-5158 USA
Phone: +1 (484) 842-4232
Email: hogenesc@mail.med.upenn.edu
Andrew D. Johnson, Ph.D.
Tenure Track Investigator
Framingham Heart Study
National Heart, Lung, and Blood Institute
73 Mt. Wayte Avenue
Framingham, MA 01702 USA
Phone: +1 (508) 663-4082
Fax: +1 (508) 626-1262
Email: andrew.johnson@nih.gov
Youngsoo Kim, Ph.D.
Professor
Departments of Biomedical Sciences and Biomedical Engineering
Seoul National University College of Medicine/Hospital
103 Daehangno Chongno-gu
Seoul 110-799 South Korea
Phone: +82 (2) 740-8073
Fax: +82 (2) 741-0253
Email: biolab@snu.ac.kr
Christopher Kinsinger, Ph.D.
Program Manager
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 31, Suite 10A52
Bethesda, MD 20892 USA
Phone: +1 (301) 451-8883
Fax: +1 (301) 496-7808
Email: kinsingc@mail.nih.gov
John R. Knowlton, Ph.D.
Program Director
Structural Biology and Molecular Applications Branch
Division of Cancer Biology
National Cancer Institute
National Institutes of Health
6130 Executive Blvd
Rockville, MD 20852 USA
Phone: +1 (301) 435-5226
Fax: +1 (301) 480-2854
Email: knowltoj@mail.nih.gov
Cheolju Lee, Ph.D.
Principal Researcher
Life Sciences Division
Korea Institute of Science and Technology
39-1 Hawolgok-dong
Seongbuk-gu, Seoul 136-791 Republic of Korea
Phone: +82 (2) 958-6788
Fax: +82 (2) 958-6919
Email: clee270@kist.re.kr
Daniel Levy, M.D.
Director and Professor
Framingham Heart Study
Center of Population Studies
National Heart, Lung, and Blood Institute
Boston University School of Medicine
73 Mt. Wayte Avenue
Framingham, MA 01702 USA
Phone: +1 (508) 935-3458
Fax: +1 (508) 626-1262
Email: levyd@nih.gov
Aldons J. Lusis, Ph.D.
Director and Professor
Departments of Medicine/Cardiology, Human Genetics, Microbiology, Immunology & Molecular Genetics
University of California at Los Angeles
675 CE Young Dr. MRL Bldg. RM 3730
Los Angeles, CA 90095 USA
Phone: +1 (310) 825-1359
Fax: +1 (310) 825-1595
Email: jlusis@mednet.ucla.edu
Pamela Marino, Ph.D.
Program Director
Pharmacology, Physiology and Biological Chemistry
National Institute of General Medical Sciences
National Institutes of Health
9000 Rockville Pike, Building 45
Bethesda, MD 20892 USA
Phone: +1 (301) 594-3827
Fax: +1 (301) 402-0224
Email: marinop@nigms.nih.gov
Mehdi Mesri, Ph.D.
Program Manager
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 31, Suite 10A52
Bethesda, MD 20892 USA
Phone: +1 (301) 451-8883
Fax: +1 (301) 496-7808

Email: mesrim@mail.nih.gov
Ken Miller, Ph.D.
Vice President, Marketing
Life Sciences Mass Spectrometry
Thermo Fisher Scientific
355 River Oaks Parkway
San Jose, CA 95134
Phone: +1 (408) 965-6336
Fax: +1 (408) 965-6132
Email: ken.miller@thermo.com
Larry G. Moss, M.D.
Associate Professor
Division of Endocrinology, Metabolism & Nutrition
Sarah W. Stedman Nutrition & Metabolism Center
School of Medicine, Duke University
2100 Erwin Road, Durham
Durham, NC 27710 USA
Phone: +1 (919) 479-2310
Email: larry.moss@duke.edu
Peter J. Munson, Ph.D.
Chief, Center for Information Technology
Mathematical and Statistical Computing Laboratory
National Institutes of Health
Bldg 12A, Rm. 2039
Bethesda, MD 20892-5626 USA
Phone: +1 (301) 496-2972
Fax: +1 (301) 402-2172
Email: munson@helix.nih.gov
Larry A. Nagahara, Ph.D.
Director, Office of Physical Sciences Oncology
National Cancer Institute
National Institutes of Health
9000 Rockville Pike, Building 31, Suite 10A03
Bethesda, MD 20892 USA
Phone: +1 (301) 451-3388
Fax: +1 (301) 496-7807
Email: larry.nagahara@nih.gov
Susan E. Old, Ph.D.
Special Assistant (Detail) to Deputy Director
Office of Extramural Research
National Center for Advancing Translational Sciences
National Institutes of Health
6705 Rockledge Drive, Room 5162
Bethesda, MD 20892
Phone: +1 (301) 435-1961
Fax: +1 (301) 402-1798
Email: susan.old@nih.gov
Samuel Payne, Ph.D.
Scientist
Pacific Northwest National Laboratory
P.O.Box 999
MSIN: k8-98
Richland, WA 99352 USA
Phone: +1 (509) 371-6513
Fax: +1 (509) 371-6564
Email: samuel.payne@pnnl.gov
Peipei Ping, Ph.D.
Director and Professor
Departments of Physiology and Medicine/Cardiology
University of California at Los Angeles
MRL Building, Suite 1-609
675 Charles E. Young Dr.
Los Angeles, CA 90095-1760 USA
Phone: +1 (310) 206-0058
Fax: +1 (310) 267-5623
Email: pping@mednet.ucla.edu
Dennis A. Przywara, Ph.D.
Program Director
Heart Failure and Arrhythmias Branch
Division of Cardiovascular Diseases
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, RKL2 BG RM 8182
Bethesda, MD 20817 USA
Phone: +1 (301) 435-0506
Fax: +1 (301) 480-7404
Email: przywarad@nhlbi.nih.gov
Mona A. Puggal, M.P.H.
Epidemiology Branch
National Institute of Enviromental Sciences
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, RKL2 BG RM 10199
Bethesda, MD 20817 USA
Phone: +1 (301) 435-0704
Fax: +1 (301) 480-1455
Email: puggalma@mail.nih.gov
Robert Rivers, Ph.D.
Program Manager
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 31, Suite 10A52
Bethesda, MD 20892 USA
Phone: +1 (301) 451-8883
Fax: +1 (301) 496-7808
Email: riversrc@mail.nih.gov
Henry Rodriguez, Ph.D., M.B.A.
Director
Office of Cancer Clinical Proteomics Research
National Cancer Institute
National Institutes of Health
9000 Rockville Pike
Building 31, Suite 10A52
Bethesda, MD 20892 USA
Phone: +1 (301) 451-8883
Fax: +1 (301) 496-7807
Email: rodriguezh@mail.nih.gov
Lisa Schwartz, Ph.D.
Program Director
Heart Failure and Arrhythmias Branch
Division of Cardiovascular Diseases
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, RKL2 BG RM 8166
Bethesda, MD 20817 USA
Phone: +1 (301) 402-4826
Fax: +1 (301) 480-1336
Email: schwartzlongal@mail.nih.gov
Belinda L. Seto, Ph.D.
Deputy Director
Office of the Director
National Institute of Biomedical Imaging and Bioengineering
National Institutes of Health
9000 Rockville Pike, Building 31, Suite 1C18
Bethesda, MD 20817 USA
Phone: +1 (301) 496-8859
Fax: +1 (301) 480-4515
Email: setob@mail.nih.gov
Svati H. Shah, M.D., M.H.S.
Associate Professor
Department of Medicine/Cardiology
Duke Center for Human Genetics
Duke University Medical Center
DUMC Box 3445
Durham, NC 27710 USA
Phone: +1 (919) 684-2859
Fax: +1 (919) 684-0928
Email: svati.shah@duke.edu
Douglas M. Sheeley, SC.D.
Senior Scientific Officer

Center for Bioinformatics and Computational Biology
National Institute of General Medical Sciences
National Institutes of Health
9000 Rockville Pike, Building 45
Bethesda, MD 20892 USA
Phone: +1 (301) 435-0755
Fax: +1 (301) 402-0224
Email: douglas.sheeley@nih.gov
Phyliss Sholinsky, M.S.P.H.
Senior Advisor
Prevention and Population Sciences Program
Division of Cardiovascular Diseases
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, RKL2 BG RM 10120
Bethesda, MD 20892-7936 USA
Phone: +1 (301) 435-0703
Fax: +1 (301) 480-1864
Email: sholinsp@nhlbi.nih.gov
Gary Siuzdak, Ph.D.
Director and Professor
Scripps Center for Metabolomics and Mass Spectrometry
Departments of Chemistry/Molecular Biology
The Scripps Research Institute
Mailcode: SR-15
10550 North Torrey Pines Road
La Jolla, CA 92037 USA
Phone: +1 (858) 784-9415
Fax: +1 (858) 784-9496
Email: siuzdak@scripps.edu
Steven Skates, Ph.D.
Associate Professor
Department of Medicine
Harvard Medical School
Massachusetts General Hospital
50 Staniford Street, Suite 560
Boston, MA 02114 USA
Phone: +1 (617) 726-4309
Fax: +1 (617) 724-9878
Email: sskates@partners.org
Michael Snyder, Ph.D.
Chair and Professor
Department of Genetics
Stanford University
300 Pasteur Dr., M-344
Stanford, CA 94305-5120 USA
Phone: +1 (650) 736-8099
Fax: +1 (650)796-6378
Email: mpsnyder@stanford.edu
Heidi J. Sofia, Ph.D., M.P.H.
Program Director
Computational Biology
National Human Genome Research Institute
National Institutes of Health
5635 Fishers Lane, Suite 4076
Bethesda, MD 20892 USA
Phone: +1 (301) 496-7531
Fax: +1 (301) 480-2770
Email: heidi.sofia@nih.gov
Pothur Srinivas, Ph.D., M.P.H.
Lead Program Director
Division of Cardiovascular Sciences
National Heart, Lung, and Blood Institute
National Institutes of Health
6701 Rockledge Drive, Room 10184, MSC 7936
Bethesda, MD 20892 USA
Phone: +1 (301) 402-3712
Fax: +1 (310) 480-2858
Email: srinivap@nhlbi.nih.gov
Sudhir Srivastava, Ph.D., M.P.H.
Chief, Cancer Biomarkers Research Group
Division of Cancer Prevention
National Cancer Institute
National Institutes of Health
6130 Executive Boulevard, Suite 3142
Rockville, MD 20852 USA
Phone: +1 (301) 496-3983
Fax: +1 (301) 402-8990
Email: sudhir.srivastava@nih.gov
Michael B. Strader, Ph.D.
Researcher
Laboratory of Biochemistry and Vascular Biology
Center for Biologics Evaluation and Research
Food and Drug Administration
9000 Rockville Pike, Building 29, Room B26
Bethesda, MD 20892 USA
Phone: +1 (301) 827-0288
Fax: +1 (301) 451-5780
Email: michael.strader@fda.hhs.gov
Danilo A. Tagle, Ph.D.
Program Director, Neurogenetics
National Institute of Neurological Disorders and Stroke National Institutes of Health
6001 Executive Blvd, Room 2114
Bethesda, MD 20892 USA
Phone: +1 (301) 496-5745
Fax: +1 (301) 402-1501
Email: danilo.tagle@nih.gov
Magdalena Thurin, Ph.D.
Program Director
Cancer Diagnosis Program
National Cancer Institute
National Institutes of Health
6130 Executive Blvd, Room 6044
Rockville, MD 20852 USA
Phone: +1 (301) 496-1591
Fax: +1 (301) 402-7819
Email: magdalena.thurin@nih.gov
Jennifer E. Van Eyk, Ph.D.
Director and Professor
JHU Bayview Proteomics Center
Departments of Medicine/Biological Chemistry and Biomedical Engineering
School of Medicine
Johns Hopkins University
5200 Eastern Avenue
Mason F. Lord Building, Center Tower, Room 602
Baltimore, MD 21224 USA
Phone: +1 (410) 550-8511
Fax: +1 (410) 550-8512
Email: jvaneyk1@jhmi.edu
John N. Weinstein, M.D., Ph.D.
Chair and Professor
Division of Quantitative Sciences
Department of Bioinformatics and Computational Biology
The University of Texas M.D. Anderson Cancer Center
Unit 1410 P.O. Box 301402
Houston, TX 77230 USA
Phone: +1 (713) 563-9296
Fax: +1 (713) 563-4242
Email: jweinste@mdanderson.org
John Yates III, Ph.D.
Director and Professor
Department of Chemical Physiology
The Scripps Research Institute
10550 North Torrey Pines Rd.
Department of Chemical Physiology, SR11
La Jolla, CA 92037 USA
Phone: +1 (858) 784-8862
Fax: +1 (858) 784-8883
Email: jyates@scripps.edu
Jun Zhang, Ph.D.
Director

NHLBI Proteomics Coordinating and Administration Center
Departments of Physiology and Medicine/Cardiology
University of California at Los Angeles
MRL Building, Suite 1-619, 675 Charles E. Young Dr.
Los Angeles, CA 90095-1760 USA
Phone: +1 (310) 794-1348
Fax: +1 (310) 267-5623
Email: jzhang@mednet.ucla.edu

Received: 3 September 2014 Accepted: 2 December 2014
Published: 15 December 2014

References

1. Lane L, Bairoch A, Beavis RC, Deutsch EW, Gaudet P, Lundberg E, Omenn GS: **Metrics for the Human Proteome Project 2013-2014 and strategies for finding missing proteins.** *J Proteome Res* 2014, **13**(1):15–20.
2. Kim MS, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabuddhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LD, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, *et al*: **A draft map of the human proteome.** *Nature* 2014, **509**(7502):575–581.
3. Wilhelm M, Schlegl J, Hahne H, Moghaddas Gholami A, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber F, Kuster B: **Mass-spectrometry-based draft of the human proteome.** *Nature* 2014, **509**(7502):582–587.
4. Low TY, van Heesch S, van den Toorn H, Giansanti P, Cristobal A, Toonen P, Schafer S, Hübner N, van Breukelen B, Mohammed S, Cuppen E, Heck AJ, Guryev V: **Quantitative and qualitative proteome characteristics extracted from in-depth integrated genomics and proteomics analysis.** *Cell Rep* 2013, **5**(5):1469–1478.
5. Aebersold R, Bader GD, Edwards AM, van Eyk JE, Kussmann M, Qin J, Omenn GS: **The biology/disease-driven human proteome project (B/D-HPP): enabling protein research for the life sciences community.** *J Proteome Res* 2013, **12**(1):23–27.
6. Ciriello G, Cerami E, Sander C, Schultz N: **Mutual exclusivity analysis identifies oncogenic network modules.** *Genome Res* 2012, **22**(2):398–406.
7. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S, Davies SR, Wang S, Wang P, Kinsinger CR, Rivers RC, Rodriguez H, Townsend RR, Ellis MJ, Carr SA, Tabb DL, Coffey RJ, Slebos RJ, Liebler DC, NCI CPTAC: **Proteogenomic characterization of human colon and rectal cancer.** *Nature* 2014, **513**(7518):382–387.
8. Paik YK, Hancock WS: **Uniting ENCODE with genome-wide proteomics.** *Nat Biotechnol* 2012, **30**(11):1065–1067.
9. Vidal M, Chan DW, Gerstein M, Mann M, Omenn GS, Tagle D, Sechi S, Workshop Participants: **The human proteome – A scientific opportunity for transforming diagnostics, therapeutics, and healthcare.** *Clin Proteomics* 2012, **9**(1):6.
10. Hood LE, Omenn GS, Moritz RL, Aebersold R, Yamamoto KR, Amos M, Hunter-Cevera J, Locascio L, Workshop Participants: **New and improved proteomics technologies for understanding complex biological systems: addressing a grand challenge in the life sciences.** *Proteomics* 2012, **12**(18):2773–2783.
11. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, Miriami E, Karczewski KJ, Hariharan M, Dewey FE, Cheng Y, Clark MJ, Im H, Habegger L, Balasubramanian S, O'Huallachain M, Dudley JT, Hillenmeyer S, Haraksingh R, Sharon D, Euskirchen G, Lacroute P, Bettinger K, Boyle AP, Kasowski M, Grubert F, Seki S, Garcia M, Whirl-Carrillo M, Gallardo M, *et al*: **Personal omics profiling reveals dynamic molecular and medical phenotypes.** *Cell* 2012, **148**(6):1293–1307.
12. Hogenesch JB, Ueda HR: **Understanding systems-level properties: timely stories from the study of clocks.** *Nat Rev Genet* 2011, **12**(6):407–416.
13. Ball LE, Hart GW: **Post-translational modifications: a major focus for the future of proteomics.** *Mol Cell Proteomics* 2013, **12**(12):3443.
14. Joehanes R, Johnson AD, Barb JJ, Raghavachari N, Liu P, Woodhouse KA, O'Donnell CJ, Munson PJ, Levy D: **Gene expression analysis of whole blood, peripheral blood mononuclear cells, and lymphoblastoid cell lines from the Framingham Heart Study.** *Physiol Genomics* 2012, **44**(1):59–75.
15. Sen SK, Boelte KC, Barb JJ, Joehanes R, Zhao X, Cheng Q, Adams L, Teer JK, Accame DS, Chowdhury S, Singh LN, NISC Comparative Sequencing Program; CHARGE Consortium, Kavousi M, Peyser PA, Quigley L, Priel DL, Lau K, Kuhns DB, Yoshimura T, Johnson AD, Hwang SJ, Chen MY, Arai AE, Green ED, Mullikin JC, Kolodgie FD, O'Donnell CJ, Virmani R, Munson PJ, *et al.*: **Integrative DNA, RNA, and Protein Evidence Connects TREML4 to Coronary Artery Calcification.** *Am J Hum Genet* 2014, **95**(1):66–76.
16. Zhang EE, Liu AC, Hirota T, Miraglia LJ, Welch G, Pongsawakul PY, Liu X, Atwood A, Huss JW 3rd, Janes J, Su AI, Hogenesch JB, Kay SA: **A genome-wide RNAi screen for modifiers of the circadian clock in human cells.** *Cell* 2009, **139:**199–210.
17. Shankar A, Sun L, Klein BE, Lee KE, Muntner P, Nieto FJ, Tsai MY, Cruickshanks KJ, Schubert CR, Brazy PC, Coresh J, Klein R: **Markers of inflammation predict the long-term risk of developing chronic kidney disease: a population-based cohort study.** *Kidney Int* 2011, **80**(11):1231–1238.
18. Ng PC, Murray SS, Levy S, Venter JC: **An agenda for personalized medicine.** *Nature* 2009, **461**(7265):724–726.
19. Committee on the Review of Omics-Based Tests for Predicting Patient Outcomes in Clinical Trials, Board on Health Care Services, Board on Health Sciences Policy, Institute of Medicine, Micheel CM, Nass SJ, Omenn GS: *Evolution of Translational Omics: Lessons Learned and the Path Forward.* Washington (DC): National Academies Press (US); 2012.
20. Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ: **Exome sequencing identifies the cause of a mendelian disorder.** *Nat Genet* 2010, **42**(1):30–35.
21. Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA: **Evaluation of next generation sequencing platforms for population targeted sequencing studies.** *Genome Biol* 2009, **10**(3):R32.
22. Imes CC, Austin MA: **Low-density lipoprotein cholesterol, apolipoprotein B, and risk of coronary heart disease: from familial hyperlipidemia to genomics.** *Biol Res Nurs* 2013, **15**(3):292–308.
23. Shankavaram UT, Varma S, Kane D, Sunshine M, Chary KK, Reinhold WC, Pommier Y, Weinstein JN: **Cell Miner: a relational database and query tool for the NCI-60 cancer cell lines.** *BMC Genomics* 2009, **10:**277.
24. Maier CW, Long JG, Hemminger BM, Giddings MC: **Ultra-Structure database design methodology for managing systems biology data and analyses.** *BMC Bioinformatics* 2009, **10:**254.
25. Taniguchi Y, Choi PJ, Li GW, Chen H, Babu M, Hearn J, Emili A, Xie XS: **Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells.** *Science* 2010, **329**(5991):533–538.
26. Vreede FT, Chan AY, Sharps J, Fodor E: **Mechanisms and functional implications of the degradation of host RNA polymerase II in influenza virus infected cells.** *Virology* 2010, **396**(1):125–134.
27. Brar GA, Yassour M, Friedman N, Regev A, Ingolia NT, Weissman JS: **High-resolution view of the yeast meiotic program revealed by ribosome profiling.** *Science* 2012, **335**(6068):552–557.
28. Vinayagam A, Stelzl U, Foulle R, Plassmann S, Zenkner M, Timm J, Assmus HE, Andrade-Navarro MA, Wanker EE: **A directed protein interaction network for investigating intracellular signal transduction.** *Sci Signal* 2011, **4**(189):rs8.
29. Homer N, Szelinger S, Redman M, Duggan D, Tembe W, Muehling J, Pearson JV, Stephan DA, Nelson SF, Craig DW: **Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays.** *PLoS Genet* 2008, **4**(8):e1000167.
30. Tanner S, Shen Z, Ng J, Florea L, Guigó R, Briggs SP, Bafna V: **Improving gene annotation using peptide mass spectrometry.** *Genome Res* 2007, **17**(2):231–239.
31. Perez-Riverol Y, Alpi E, Wang R, Hermjakob H, Vizcaíno JA: **Making proteomics data accessible and reusable: current state of proteomics databases and repositories.** *Proteomics* 2014. doi:10.1002/pmic.201400302 [Epub ahead of print].
32. Tarvin KA, Sandusky GE: **Using molecular profiled human tissue to accelerate drug discovery.** *Expert Opin Drug Discov* 2014, **12:**1–5 [Epub ahead of print].
33. Rogersa S, Cambrosiob A: **Making a new technology work: the standardization and regulation of microarrays.** *Yale J Biol Med* 2007, **80**(4):165–178.

34. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen MJ, Angiuoli SV, Ashburner M, Axelrod N, Baldauf S, Ballard S, Boore J, Cochrane G, Cole J, Dawyndt P, De Vos P, DePamphilis C, Edwards R, Faruque N, Feldman R, Gilbert J, Gilna P, Glöckner FO, Goldstein P, Guralnick R, Haft D, Hancock D, *et al*: **The minimum information about a genome sequence (MIGS) specification.** *Nature Biotech* 2008, **26:**541–547.
35. Martínez-Bartolomé S: **Guidelines for reporting quantitative mass spectrometry based experiments in proteomics.** *J Proteomics* 2013, **95:**84–88.
36. Ivanov AR, Colangelo CM, Dufresne CP, Friedman DB, Lilley KS, Mechtler K, Phinney BS, Rose KL, Rudnick PA, Searle BC, Shaffer SA, Weintraub ST: **Interlaboratory studies and initiatives developing standards for proteomics.** *Proteomics* 2013, **13**(6):904–909.
37. Rudnick PA, Clauser KR, Kilpatrick LE, Tchekhovskoi DV, Neta P, Blonder N, Billheimer DD, Blackman RK, Bunk DM, Cardasis HL, Ham AJ, Jaffe JD, Kinsinger CR, Mesri M, Neubert TA, Schilling B, Tabb DL, Tegeler TJ, Vega-Montoto L, Variyath AM, Wang M, Wang P, Whiteaker JR, Zimmerman LJ, Carr SA, Fisher SJ, Gibson BW, Paulovich AG, Regnier FE, Rodriguez H, *et al*: **Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses.** *Mol Cell Proteomics* 2010, **9**(2):225–241.
38. Addona TA, Abbatiello SE, Schilling B, Skates SJ, Mani DR, Bunk DM, Spiegelman CH, Zimmerman LJ, Ham AJ, Keshishian H, Hall SC, Allen S, Blackman RK, Borchers CH, Buck C, Cardasis HL, Cusack MP, Dodder NG, Gibson BW, Held JM, Hiltke T, Jackson A, Johansen EB, Kinsinger CR, Li J, Mesri M, Neubert TA, Niles RK, Pulsipher TC, Ransohoff D, *et al*: **Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma.** *Nat Biotechnol* 2009, **27**(7):633–641.
39. MSI Board Members, Sansone SA, Fan T, Goodacre R, Griffin JL, Hardy NW, Kaddurah-Daouk R, Kristal BS, Lindon J, Mendes P, Morrison N, Nikolau B, Robertson D, Sumner LW, Taylor C, van der Werf M, van Ommen B, Fiehn O: **The metabolomics standards initiative.** *Nat Biotechnol* 2007, **25**(8):846–848.
40. Mertens P, Yang F, Liu T, Mani DR, Petyuk VA, Gillette MA, Clauser KR, Qiao JW, Gritsenko MA, Moore RJ, Levine DA, Townsend R, Erdmann-Gilmore P, Snider JE, Davies SR, Ruggles KV, Fenyo D, Kitchens RT, Li S, Olvera N, Dao F, Rodriguez H, Chan DW, Liebler D, White F, Rodland KD, Mills GB, Smith RD, Paulovich AG, Ellis M, *et al*: **Ischemia in tumors induces early and sustained phosphorylation changes in stress kinase pathways but does not affect global protein levels.** *Mol Cell Proteomics* 2014, **13**(7):1690–1704.
41. Kuhn E, Whiteaker JR, Mani DR, Jackson AM, Zhao L, Pope ME, Smith D, Rivera KD, Anderson NL, Skates SJ, Pearson TW, Paulovich AG, Carr SA: **Interlaboratory evaluation of automated, multiplexed peptide immunoaffinity enrichment coupled to multiple reaction monitoring mass spectrometry for quantifying proteins in plasma.** *Mol Cell Proteomics* 2012, **11**(6):M111.013854.
42. Bjornson ZB, Nolan GP, Fantl WJ: **Single-cell mass cytometry for analysis of immune system functional states.** *Curr Opin Immunol* 2013, **25**(4):484–494.
43. Qiu J, LaBaer J: **Nucleic acid programmable protein array a just-in-time multiplexed protein expression and purification platform.** *Methods Enzymol* 2011, **500:**151–163.
44. Newman RH, Hu J, Rho HS, Xie Z, Woodard C, Neiswinger J, Cooper C, Shirley M, Clark HM, Hu S, Hwang W, Jeong JS, Wu G, Lin J, Gao X, Ni Q, Goel R, Xia S, Ji H, Dalby KN, Birnbaum MJ, Cole PA, Knapp S, Ryazanov AG, Zack DJ, Blackshaw S, Pawson T, Gingras AC, Desiderio S, Pandey A, *et al*: **Construction of human activity-based phosphorylation networks.** *Mol Syst Biol* 2013, **9:**655.
45. Boja E, Jortani SA, Ritchie J, Hoofnagle AN, Težak Ž, Mansfield E, Keller P, Rivers RC, Rahbar A, Anderson NL, Srinivas P, Rodriguez H: **The journey to regulation of protein-based multiplex quantitative assays.** *Clin Chem* 2011, **57**(4):560–567.

doi:10.1186/1559-0275-11-45
Cite this article as: Boja *et al.*: Integration of omics sciences to advance biology and medicine. *Clinical Proteomics* 2014 **11**:45.