Stem cell systems informatics for advanced clinical biodiagnostics: tracing molecular signatures from bench to bedside

Randolph S. Faustino[1], D. Kent Arrell[1], Clifford D.L. Folmes[1], Andre Terzic[1], Carmen Perez-Terzic[1,2]

[1]Division of Cardiovascular Diseases, Departments of Medicine, Molecular Pharmacology and Experimental Therapeutics, Mayo Clinic, Rochester, MN, USA

[2]Physical Medicine and Rehabilitation, Mayo Clinic College of Medicine, Rochester, MN, USA

Development of innovative high throughput technologies has enabled a variety of molecular landscapes to be interrogated with an unprecedented degree of detail. Emergence of next generation nucleotide sequencing methods, advanced proteomic techniques, and metabolic profiling approaches continue to produce a wealth of biological data that captures molecular frameworks underlying phenotype. The advent of these novel technologies has significant translational applications, as investigators can now explore molecular underpinnings of developmental states with a high degree of resolution. Application of these leading-edge techniques to patient samples has been successfully used to unmask nuanced molecular details of disease vs healthy tissue, which may provide novel targets for palliative intervention. To enhance such approaches, concomitant development of algorithms to reprogram differentiated cells in order to recapitulate pluripotent capacity offers a distinct advantage to advancing diagnostic methodology. Bioinformatic deconvolution of several "-omic" layers extracted from reprogrammed patient cells, could, in principle, provide a means by which the evolution of individual pathology can be developmentally monitored. Significant logistic challenges face current implementation of this novel paradigm of patient treatment and care, however, several of these limitations have been successfully addressed through continuous development of cutting edge *in silico* archiving and processing methods. Comprehensive elucidation of genomic, transcriptomic, proteomic, and metabolomic networks that define normal and pathological states, in combination with reprogrammed patient cells are thus poised to become high value resources in modern diagnosis and prognosis of patient disease.

Correspondence to:
C. Perez-Terzic
Mayo Clinic
200 First Street SW
Rochester, MN, USA 55905
terzic.carmen@mayo.edu

Disease anticipation prior to symptomatic presentation provides multiple pre-emptive opportunities for clinical management. Palliative or curative effectiveness is enhanced by early detection of disease-promoting factors, as early stage treatment can preclude catastrophic pathological progression. Translationally relevant risk factors can be identified through techniques that interrogate and assess biological marker status (1), where methods such as gene overexpression, peptide fragment enrichment, and metabolite concentration can be respectively used to diagnose diseases such as clinical carcinoma (2), cardiovascular pathologies (3), and diabetes (4).

Distinctions among high throughput methodologies confer specific advantages and disadvantages that determine efficiency and application of genomic, proteomic, or metabolomic approaches. For example, genes validated from tissue biopsies and measurement of glucose uptake with the glucose analogue tracer 18F-fluoro-2-deoxyglucose are among the tools that oncologists rely upon to diagnose and manage tumors (5). Experimental confirmation of gene expression differences in normal vs tumor samples can then be used to determine tumor aggressiveness and provide prognosis and susceptibility profiles (6).

Differential gene expression provides direct indicators of genetic background and transcriptome reprogramming as a consequence of evolving disease, but cannot provide information regarding dynamic functional changes. Peptide screening, a hallmark of proteomic analyses, can be employed to obtain this type of information (7,8). For example, palliation of cardiopathology is more effective when combined with a proteomic strategy to identify peptide fragments released into patient serum as a result of cardiac dysfunction (9,10).

Advantages of genomic/transcriptomic and proteomic screens provide relevant data for clinical management of disease, though the techniques themselves require several days to weeks for quality-controlled processing and data analyses. Metabolic screening offers a distinct advantage in that samples can be assayed to provide feedback on disease states with greater celerity than the "-omic" approaches described previously. Though the prototype for metabolite analysis is glucose monitoring used in diabetes management (11), other examples of clinical metabolomics exist (12,13). Significantly, stabilization and processing of evanescent metabolites for precise and accurate measurements of metabolic states in diseased vs non-diseased patients is critical to all metabolomic profiling approaches.

These techniques allow deep interrogation of the molecular complement that comprises the functional background of a cell or tissue sample. Clinical application of these approaches gains significant leverage when combined with a pluripotent cell platform that can provide readouts of the molecular origin of disease phenotype.

THE POST-GENOMIC ERA AND STEM CELLS

Current generation high throughput technology, originally developed to address rapid high volume sequencing needs (14), has galvanized genomic, transcriptomic, proteomic, and metabolomic platform development. Depth and resolution of biomarker composition analysis has increased significantly, which has necessitated development of commensurate *in silico* methods. Indeed, bioinformatic approaches that employ systems biology principles to parse multidimensional biodata extracts multiple levels of integrated information, and application of such techniques to targeted stem cell populations may offer novel modalities of advanced pathology diagnosis.

Stem cells possess a unique potential to anticipate disease phenotypes as they harbor the fundamental molecular baseline that gives rise to genomic content and its derived transcriptomic, proteomic, and metabolomic strata. Interrogation of these "-omic" layers provides rich data to deconvolute systems biology of developmental programs, elucidation of which is critical to understanding clinical pathology etiology. Molecular cartography of pluripotent disease-prone backgrounds would facilitate pre-emptive diagnosis through comparison of healthy wild type templates with diseased states (3,15,16), as well as offer a tool for dynamic prognosis to track changes concomitant with pathological progression or assessment of response to therapeutic intervention (3). While a variety of stem cells offer specific clinical advantages, pluripotent stem cells may be ideal candidates for patient applications (17,18).

Embryonic stem (ES) cells can be harvested and profiled to establish a molecular baseline for comparison to diseased genomes. Recent data indicate that in addition to patient-donated material, patient-specific ES cells can be successfully cloned for potential therapeutic applications (19). Partly in an attempt to circumvent moral, legal, and ethical disputes associated with human ES studies, discovery and development of a molecular algorithm to reprogram cells from a committed fate to a pluripotent state (20) may provide a viable alternative to, and significant advantages over, ES cells in development of systems biology strategies

for comprehensive disease resolution spanning embryonic to differentiated phenotypes (21). Generation and establishment of disease- and patient- specific embryonic stem cell lines is valuable for its potential to provide individualized experimental cell models, where tailored experimental design, with a focus on personalized therapeutic strategies, can be performed (22). Induced pluripotent stem (iPS) cells, as part of an advanced disease management strategy, have the potential to advance clinical diagnostics, though several obstacles face implementation of this strategy.

Full equivalency of induced compared to developmentally derived stem cells must be demonstrated to employ reprogrammed cells as proxies for a naturally occurring primordial baseline (23). Indeed, though genetic and epigenetic variation between ES and iPS cells confers a difference in tumorigenic potential (24), extraction of a conserved molecular signature may circumvent limitations associated with genetic and epigenetic heterogeneity (17,25). Furthermore, restriction of bioinformatic interrogation to iPS cell lines derived from the same patient can be used to limit influence of epigenetic variation among samples. In combination with quality control measures to ensure consistent cell culture technique and microenvironment exposure that may otherwise lead to respective chromosomal segregation and epigenetic changes (26,27), preparation of human iPS cell lines for clinical diagnostic applications may be feasible.

Systems-wide biodiagnostics, in which entire molecular complements can be assayed, provides an opportunity to comprehensively quantify molecular elements that underlie normal as well as clinically dysfunctional development. Current advances in cellular reprogramming methodology, along with the advent of high throughput technology and advanced bioinformatic computational approaches, presents a unique confluence of molecular tools and techniques by which diseased cells can be reprogrammed to a primordial state (28), then scrutinized for signature elements that segregate them from normal phenotypes. Thus, comprehensive bioinformatic dissection of phenotypic regressed ES cells offers opportunities to discover novel molecular targets for early intervention or palliation in advance of clinical manifestation, that may be found at genomic, transcriptomic, proteomic, and metabolomic levels, or any combination thereof.

GENOMICS

The genomic, transcriptomic, proteomic, and metabolomic elements of systems biology approaches provide bi-

ological information unique to each layer of the molecular network, and a variety of specialized techniques exist to interrogate each. From a molecular network ontogeny perspective, the genome is the origin of subsequent transcriptomes, proteomes, and metabolomes, and is a useful starting point for a discussion regarding roles the genome has played in translational diagnosis.

The genome is the molecular blueprint composed of genic and non-genic sequences that ultimately determine organismal phenotype (29), and the advent of whole genome sequencing has catapulted the field forward (30), as well as given rise to new specialties (31). This advance has spurred a renaissance of clinical genetics, as present genetic counseling techniques are based on reductionist "one gene, one disease" strategies, a paradigm which has successfully guided identification of the genetic basis for many clinically relevant diseases. Resolution of polygenic pathologies, as well as disease progression impacted by heritable epigenetic modifications, however, cannot be effectively addressed by this diagnostic modality. To resolve the problem of multigene and/or epigenetically triggered pathology, genome-wide association studies (GWAS) have been used to identify and define multiple susceptibility loci correlated with disease presentation (32). This approach is used to bioinformatically mine complex genomic data sampled from large populations to identify key genes that associate with disease. Paradoxically, the large sample sizes that provide robustness to a GWAS approach prevents direct application toward personalized therapeutic approaches, a limitation that must be considered when evaluating genomic diagnostic technologies for individualized patient application.

Implementation of genome-wide analyses to stem cell (re)programming in the context of therapeutic application has yielded a rich body of novel data that provides details on changes in genomic regulatory elements, epigenomic landscape transitions, and three-dimensional chromatin shifts critical to the process of phenotype (re)acquisition (28,33). These levels of genomic complexity imposed can now be resolved with high precision using modern methodology (34) to facilitate fine resolution of the full genomic blueprint, critical for translational application.

TRANSCRIPTOMICS

The transcriptome is the full complement of RNA produced in response to signaling cues processed by, and transcribed from, the underlying genome, and tech-

nology and methods employed for genomic deconstruction are applicable to transcriptome resolution. Comprehensive transcript analysis is an attractive option for biomarker identification, as panels of differentially expressed genes (DEGs) are used to establish indices of disease progression (35). Prioritized gene lists can be further analyzed for gene ontology enrichment and bionetwork analyses to respectively identify and quantitate the molecular gestalt underlying normal or diseased phenotype progression (36).

Transcriptome deconvolution has been used to identify contributions of specific genes during the process of somatic cell reprogramming (37), and a variety of criteria unique to high throughput RNA analysis that derive from diverse RNA heterogeneity (38) play critical roles in elucidating transcriptome dynamics of differentiation. An essential characteristic of the transcriptome is (auto)regulation facilitated by subtle intra-RNA dynamics (39,40), however interactions of regulatory non-coding RNA with molecular targets can be parsed using specific bioinformatic resources. Fate acquisition is driven by splice variance that occurs as a result of differential mRNA processing, and distinguishing these permutations from variations in background noise requires intensive computational resources (41). Development-dependent isoform switching presents another critical variable to temporal resolution of the transcriptome assembly during differentiation/reprogramming (42,43). In addition, high degrees of post-transcriptional regulation conferred by microRNA and long ncRNA mandates novel biostatistical models and *in silico* approaches to properly resolve transcriptome dynamics of fate commitment/reversion (44,45).

Ultimately, deeper data sets empower *in silico* tools, and is key to modern diagnostic techniques that employ next generation sequencing to leverage rich genomic and transcriptomic content against clinical diseases in order to facilitate high resolution etiology definition (46). Paired with quality biobank sample acquisition and leading-edge techniques for cell isolation and reprogramming, continuous development of high quality, publicly available computational tools will significantly advance transcriptome analysis, and will refine transcriptome interrogation as a modern tool for assaying gene expression dynamics associated with clinical pathologies.

PROTEOMICS

The proteome consists of all proteins expressed by a genome in a defined cell or tissue at a particular time,

whereas proteomics comprises an array of techniques for studying expression, abundance, structure (including post-translational modifications) and function, including their physical and functional interactions (3). Modern proteomic approaches involve high throughput protein separation and processing followed by mass spectrometry for peptide and protein identification, either as intact entities or as peptide fragments, defined respectively as top-down and bottom-up proteomics (3). As proteins form the molecular machinery of the cell, alterations in their abundance and activity translate into detectable changes in other "-omic" strata, such as epigenetic modifications, mRNA abundance or differential splicing, and altered metabolite levels. Thus, comprehension of stem cell proteomes and their dynamics may provide detailed systems understanding of pluripotency, how it differs from somatic cell states, and may yield important clues into mechanistic understanding of progenitor cells for therapeutic and diagnostic applications.

Proteomic studies have increased our understanding of protein complexes and dynamics contributing to cell fate determination, of cell state transitions during development and reprogramming cells to pluripotency, and of the extensive molecular impact mediated by disease-targeted stem cell-based therapy. An expanded pluripotency network described by proteomic assessment of protein-protein interactions between known transcription regulatory factors required for maintenance of ES cells identified combinatorial effects of transcriptional activator complexes required for pluripotency, together with repressor complexes necessary to prevent expression of proteins associated with differentiated cells (47-49).

Stoichiometric correlation between proteins and transcripts can vary (50), and proteomic deconvolution performed in conjunction with other "-omic" studies can enhance systems level analyses (51). For example, in a transcriptome/proteome study that employed chromatin occupancy interrogation with focused epigenomic tracking, discrete transcript and protein changes were revealed to be characteristic of cells released from pluripotency via NANOG depletion (52). This study also demonstrated that epigenetic and post-transcriptional effects targeted distinct subgroups of cellular processes and functions during differentiation, thus emphasizing fine dynamic regulation in and among transcriptome and proteome layers. Furthermore, proteomic characterization of reprogrammed somatic cells (53,54) has significantly identified a highly coordinated biphasic temporal dynamic driving induced pluripotency (55). In another methodological com-

bination study, (metabo)proteomic profiling was used to demonstrate enzymatic restructuring consistent with metabolomic transition from an oxidative to a glycolytic metabolomic phenotype that precedes and guides cell reprogramming (56).

Proteomic studies are also being used to define functional consequences of stem cell-based therapy. For example, extensive proteomic remodeling underlying structural and functional changes associated with onset of dilated cardiomyopathy was reversed by ES cell therapy, with derived protein networks exhibiting a pro-cardiogenic developmental response with concomitant demotion of dysfunctional disease-associated categories (57). Ultimately, the proteomic signature served as a diagnostic of stem cell repair in the setting of dilated cardiomyopathy. Collectively, these studies demonstrate the power of proteomics and of integrative systems biology strategies incorporating proteomics to elucidate molecular properties associated with maintenance or attainment of pluripotency, mechanistic underpinnings of the reprogramming process, and of stem cell therapeutic proteome remodeling in the setting of clinical disease.

METABOLOMICS

The metabolome consists of small molecular weight compounds that undergo chemical transformation within the cell. Metabolomics captures the complexity of global metabolism in the context of (patho)physiology, and a multitude of analytical tools have been developed to detect metabolite levels, such as enzymatic analysis, flame ionization, and Raman/Fourier transformed infrared and UV-VIS spectroscopy. Nuclear magnetic resonance (NMR) and mass spectrometry have become methods of choice due to their ability to resolve a wide range of chemical moieties in a high throughput manner (58). The majority of studies to date have utilized targeted approaches to interrogate a defined set of metabolites relevant to a specific biological question. However, with advances in instrumentation, data analysis and compound annotation, broadly inclusive shotgun approaches are routinely employed to reveal a more global profile (59). Resolution of intracellular (fingerprint) and extracellular (footprint) metabolomes has provided insight into metabolic restructuring that guides stem cell differentiation and dedifferentiation (56,60,61). Multiplexing of metabolomics technologies with stable-isotope assessment of metabolic fluxes will further enable dissection of intimate metabolite dynamics to establish metabolic maps defining cell fate.

To bridge the genotype-phenotype continuum, metabolite screening offers a minimally invasive diagnostic approach associated with high patient value and can provide a wealth of information as metabolite profiles serve as functional signatures of enzymatic activity (59). Metabolomic analysis of clinically relevant pathologies enables identification of key metabolites abnormal in identity or quantity. These disease state biomarkers offer fast, reproducible, and cost-effective identification of present or putative disease states (62). Presently, widespread and robust clinical applications of metabolite profiling include screening for inborn errors of metabolism, which is now routinely performed in most of the developed world (63). Targeted tandem mass spectrometry based screening of approximately 30-40 metabolites, with emphasis placed on amino acids and acylcarnitines, now enables diagnosis of over 30 different metabolic disorders, and with greater efficacy than clinical screening alone (64).

A growing number of studies have utilized metabolomic techniques to examine stem cell biology, critical for defining the baseline stem cell metabolic landscape and its perturbation in the diseased state, as well as identifying principal roles of energy metabolism in controlling stem cell fate (65-67). This fundamental work has laid the groundwork for application of stem cell metabolomics as a platform for pharmaceutical toxicity screening and identification of predictive biomarkers of toxicity. Indeed, ES cells treated with valproate were distinguished from vehicle-treated cells based upon a metabolic signature encompassing kynurenine and glutamate metabolism (68-70), demonstrating utility in applying metabolomics deconstruction against an ES cell derived investigative platform, that can ultimately be refined for clinical application.

TRANSLATIONAL APPLICATION OF NETWORK BIOLOGY

Genomic, transcriptomic, proteomic, and metabolomic networks possess unique traits that endow these strata with qualities suitable for use as clinical tools, yet an added level of diagnostic sophistication may be accomplished by leveraging the integrated molecular architecture of these networks against complex disease phenotypes. Network biology, as part of a clinical management strategy, can potentially be employed to identify molecular candidates for pharmacological intervention (71,72). Dissection of molecular interactions presents an opportunity to integrate these "-omic" layers to elucidate the systems etiology of disease occurrence and progression through study of the flow of biological information in and among these layers (73).

Biological networks possess a discriminant set of characteristics that can be quantified (74,75). This information can be used to identify critical genes, proteins, or metabolites in their respective networks that measures not only output of the system, ie, phenotype, but may provide information on biological robustness as well as informational flow (76). Thus, comprehensive molecular cartography through innovative and cutting edge high throughput methodologies can define the functional landscape, or interactome, underlying development and disease (77,78). Furthermore, biological network analysis provides dynamic metrics that permit navigation of this functional topography, serving as a molecular positioning system that identifies features of the interactome critical to maintenance of system function and crucial for targeted interventional translational strategies (9).

STEM CELL INFORMATICS: A NOVEL DIAGNOSTIC PARADIGM

Systems biology integrates multiple disciplines to create a novel area for translational application, yet faces a plethora of challenges that constrain full implementation. High throughput techniques are expensive and availability to clinical populations at present is financially prohibitive, but as technology improves and assay costs diminish, application of these techniques to the greater clinical population becomes feasible (79). For example, the speed and volume of whole genome sequencing has increased dramatically while concomitantly becoming less expensive in the past decade (80), and continuation of this trend with other "-omic" technologies (81) will ultimately realize translationally applicable personalized bioinformatics (82).
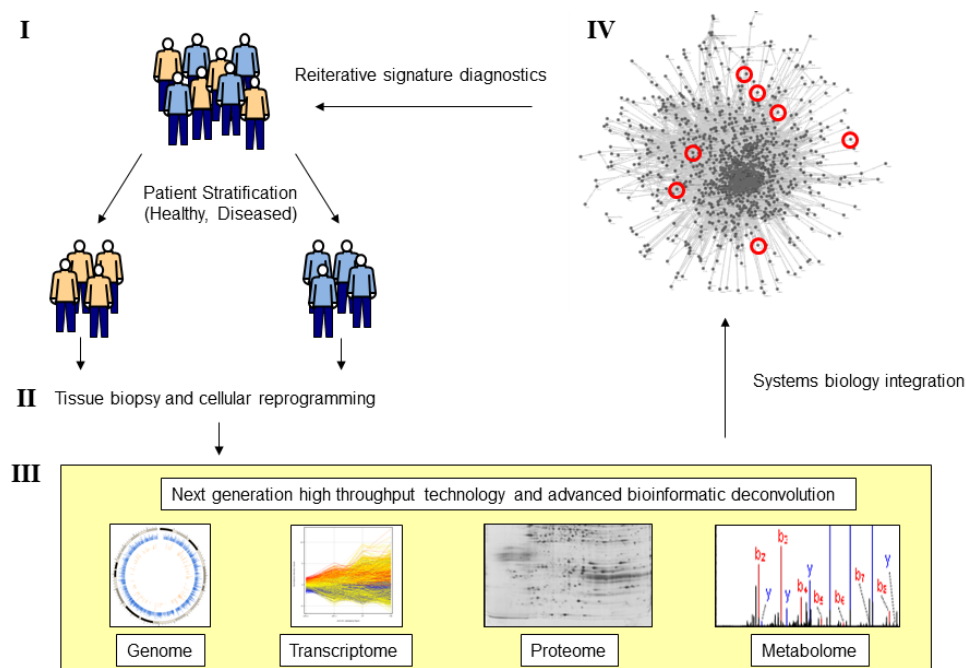
FIGURE 1. Implementation of advanced diagnostics facilitated by leading generation stem cell informatics. A comprehensive systems biology approach using integrated high throughput screening approaches in combination with the power of reprogrammed (induced pluripotent) stem cells can provide a depth of resolution that can be leveraged against poorly characterized disease etiology. In the illustrated scheme, patient stratification into healthy and diseased cohorts initiates this advanced diagnostic paradigm (**I**). Isolation of cells from patients with diseased tissue can be reprogrammed to an embryonic state (**II**), providing potential zero (embryonic) and end stage (differentiated) time points for longitudinal next generation assays. Generation of individualized and comprehensive multidimensional biological data sets at genome, transcriptome, proteome, and metabolome levels (**III**) can provide advanced clinical resources to track disease progression in real time (systems biology integration) that may be used in construction of an integrated and dynamic network signature (**IV**) to identify novel molecular targets for therapeutic intervention (red circles) in the original patient cohort (reiterative signature diagnostics).

Another parameter associated with high throughput approaches is generation of massive amounts of biodata (83). Logistics of data storage and recovery become critical factors for consideration as a clinically relevant resource, and to address this need, optimization of cloud server storage technology presents an attractive option as a readily accessible and dynamic electronic bioarchive (84). Indeed, continuous development of added-value electronic databases indicates that this growing need to store and parse intricate data sets is actively being addressed (85). It is anticipated that petabytes to exabytes of next generation biodata (86) will emerge in coming years, and for bioinformatic deconvolution to remain practical, cloud-based systems capable of handling these computationally intensive data sets are considered viable platforms to address this need (87). This is particularly critical for temporal profiling that requires integration of molecular data measured over a chronologically ordered series of developmental or disease stages (88,89).

TRACING MOLECULAR SIGNATURES FROM BENCH TO BEDSIDE ENVISIONED

Technology development in the post-genomic area has given rise to advanced high throughput methodologies and modern, integrative bioinformatic tools that permit an unprecedented level of molecular resolution. Dissection of these discrete, yet interrelated molecular strata, in combination with the ability to reprogram phenotypically committed cells to a pluripotent state that provides a unique and patient-specific embryonic pool of cells, enables feasible dynamic and individualized diagnostic strategies for translational application (90-93). For the first time, multiple molecular networks underlying clinical dysfunctions can be mapped and utilized to trace disease etiology (Figure 1). Furthermore, development of a prototype platform for stem cell diagnostics, which incorporates major "-omic" layers discussed here, provides the premise for incorporating other "-omes" for enhanced systems biology deconvolution of disease states. Implementation of this approach as part of a recursive diagnostic algorithm offers the potential for an enhanced modality of patient care made possible by a current medical zeitgeist constructed on an array of novel post-genomic knowledge (94). Ultimately, a collective and dynamic electronic repository for the diversity of biodata constantly generated by leading edge high throughput technology enables powerful meta-level analyses with unmatched precision applicable to multiple disease models (95-98).

Resetting diseased cells to a pluripotent state provides opportunities to track patient-specific changes at primary (genomic), secondary (transcriptomic), tertiary (proteomic), and quaternary (metabolomics) molecular network strata. Construction of an accessible electronic archive, to house large volumes of biodata produced from high grade bioinformatics analyses, is critical for establishing a dynamic clinical resource essential to fully realize comprehensive *ad hoc* diagnosis and real-time molecular tracking of patient pathology and disease progression.

References

1   Zaza G, Granata S, Rascio F, Pontrelli P, Dell'oglio MP, Cox SN, et al. A specific immune transcriptomic profile discriminates chronic kidney disease patients in predialysis from hemodialyzed patients. BMC Med Genomics. 2013;6:17. Medline:23663527 doi:10.1186/1755-8794-6-17

2   Verhaak RG, Tamayo P, Yang JY, Hubbard D, Zhang H, Creighton CJ, et al. Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. J Clin Invest. 2013;123:517-25. Medline:23257362

3   Arrell DK, Terzic A. Systems proteomics for translational network medicine. Circ Cardiovasc Genet. 2012;5:478. Medline:22896016 doi:10.1161/CIRCGENETICS.110.958991

4   Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics. 2013;9:280-99. Medline:23543913 doi:10.1007/s11306-012-0482-9

5   Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? Nat Rev Cancer. 2004;4:891-9. Medline:15516961 doi:10.1038/nrc1478

6   Savage SA, Mirabello L, Wang Z, Gastier-Foster JM, Gorlick R, Khanna C, et al. Genome-wide association study identifies two susceptibility loci for osteosarcoma. Nat Genet. 2013;45:799-803. Medline:23727862 doi:10.1038/ng.2645

7   Karisch R, Fernandez M, Taylor P, Virtanen C, St-Germain JR, Jin LL, et al. Global proteomic assessment of the classical protein-tyrosine phosphatome and "Redoxome". Cell. 2011;146:826-40. Medline:21884940 doi:10.1016/j.cell.2011.07.020

8   Tran JC, Zamdborg L, Ahlf DR, Lee JE, Catherman AD,

Durbin KR, et al. Mapping intact protein isoforms in discovery mode using top-down proteomics. Nature. 2011;480:254-8. Medline:22037311 doi:10.1038/nature10575

9   Arrell DK, Niederlander NJ, Faustino RS, Behfar A, Terzic A. Cardioinductive network guiding stem cell differentiation revealed by proteomic cartography of tumor necrosis factor alpha-primed endodermal secretome. Stem Cells. 2008;26:387-400. Medline:17991915 doi:10.1634/stemcells.2007-0599

10   Arrell DK, Neverova I, Van Eyk JE. Cardiovascular proteomics: evolution and potential. Circ Res. 2001;88:763-73. Medline:11325867 doi:10.1161/hh0801.090193

11   Vashist SK. Non-invasive glucose monitoring technology in diabetes management: a review. Anal Chim Acta. 2012;750:16-27. Medline:23062426 doi:10.1016/j.aca.2012.03.043

12   Trushina E, Nemutlu E, Zhang S, Christensen T, Camp J, Mesa J, et al. Defects in mitochondrial dynamics and metabolomic signatures of evolving energetic stress in mouse models of familial Alzheimer's disease. PLoS ONE. 2012;7:e32737. Medline:22393443 doi:10.1371/journal.pone.0032737

13   Nemutlu E, Zhang S, Juranic NO, Terzic A, Macura S, Dzeja P. 18O-assisted dynamic metabolomics for individualized diagnostics and treatment of human diseases. Croat Med J. 2012;53:529-34. Medline:23275318 doi:10.3325/cmj.2012.53.529

14   Bras J, Guerreiro R, Hardy J. Use of next-generation sequencing and other whole-genome strategies to dissect neurological disease. Nat Rev Neurosci. 2012;13:453-64. Medline:22714018 doi:10.1038/nrn3271

15   Folmes CD, Martinez-Fernandez A, Perales-Clemente E, Li X, McDonald A, Oglesbee D, et al. Disease-causing mitochondrial heteroplasmy segregated within induced pluripotent stem cell clones derived from a MELAS patient. Stem Cells. 2013;31:1298-308. Medline:23553816 doi:10.1002/stem.1389

16   Faustino RS, Chiriac A, Niederlander NJ, Nelson TJ, Behfar A, Mishra PK, et al. Decoded calreticulin-deficient embryonic stem cell transcriptome resolves latent cardiophenotype. Stem Cells. 2010;28:1281-91. Medline:20506533

17   Chiriac A, Nelson TJ, Faustino RS, Behfar A, Terzic A. Cardiogenic induction of pluripotent stem cells streamlined through a conserved SDF-1/VEGF/BMP2 integrated network. PLoS ONE. 2010;5:e9943. Medline:20376342 doi:10.1371/journal.pone.0009943

18   Nelson TJ, Martinez-Fernandez A, Yamada S, Perez-Terzic C, Ikeda Y, Terzic A. Repair of acute myocardial infarction by human stemness factors induced pluripotent stem cells. Circulation. 2009;120:408-16. Medline:19620500 doi:10.1161/CIRCULATIONAHA.109.865154

19   Tachibana M, Amato P, Sparman M, Gutierrez NM, Tippner-Hedges R, Ma H, et al. Human embryonic stem cells derived by somatic cell nuclear transfer. Cell. 2013;153:1228-38. Medline:23683578 doi:10.1016/j.cell.2013.05.006

20   Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126:663-76. Medline:16904174 doi:10.1016/j.cell.2006.07.024

21   Robinton DA, Daley GQ. The promise of induced pluripotent stem cells in research and therapy. Nature. 2012;481:295-305. Medline:22258608 doi:10.1038/nature10761

22   de Sa Silva F, Almeida PN, Rettore JV, Maranduba CP, de Souza CM, de Souza GT, et al. Toward personalized cell therapies by using stem cells: seven relevant topics for safety and success in stem cell therapy. J Biomed Biotechnol. 2012;2012:758102. Medline:23226945

23   Narsinh KH, Sun N, Sanchez-Freire V, Lee AS, Almeida P, Hu S, et al. Single cell transcriptional profiling reveals heterogeneity of human induced pluripotent stem cells. J Clin Invest. 2011;121:1217-21. Medline:21317531 doi:10.1172/JCI44635

24   Ben-David U, Benvenisty N. The tumorigenicity of human embryonic and induced pluripotent stem cells. Nat Rev Cancer. 2011;11:268-77. Medline:21390058 doi:10.1038/nrc3034

25   Nestor MW, Noggle SA. Standardization of human stem cell pluripotency using bioinformatics. Stem Cell Res Ther. 2013;4:37. Medline:23680084 doi:10.1186/scrt185

26   Lessing D, Anguera MC, Lee JT. X chromosome inactivation and epigenetic responses to cellular reprogramming. Annu Rev Genomics Hum Genet. Forthcoming 2013.

27   Hamada M, Malureanu LA, Wijshake T, Zhou W, van Deursen JM. Reprogramming to pluripotency can conceal somatic cell chromosomal instability. PLoS Genet. 2012;8:e1002913. Medline:22952451 doi:10.1371/journal.pgen.1002913

28   Polo JM, Anderssen E, Walsh RM, Schwarz BA, Nefzger CM, Lim SM, et al. A molecular roadmap of reprogramming somatic cells into iPS cells. Cell. 2012;151:1617-32. Medline:23260147 doi:10.1016/j.cell.2012.11.039

29   Gingeras TR. Origin of phenotypes: genes and transcripts. Genome Res. 2007;17:682-90. Medline:17567989 doi:10.1101/gr.6525007

30   Morozova O, Marra MA. Applications of next-generation sequencing technologies in functional genomics. Genomics. 2008;92:255-64. Medline:18703132 doi:10.1016/j.ygeno.2008.07.001

31   Morozova O, Hirst M, Marra MA. Applications of new sequencing technologies for transcriptome analysis. Annu Rev Genomics Hum Genet. 2009;10:135-51. Medline:19715439 doi:10.1146/annurev-genom-082908-145957

32   Ritchie MD, Denny JC, Zuvich RL, Crawford DC, Schildcrout JS, Bastarache L, et al. Genome- and phenome-wide analyses of cardiac conduction identifies markers of arrhythmia risk. Circulation. 2013;127:1377-85. Medline:23463857 doi:10.1161/CIRCULATIONAHA.112.000604

33   Zhu J, Adli M, Zou JY, Verstappen G, Coyne M, Zhang X, et al. Genome-wide chromatin state transitions associated with developmental and environmental cues. Cell. 2013;152:642-54.

Medline:23333102 doi:10.1016/j.cell.2012.12.033

34 Zhang JA, Mortazavi A, Williams BA, Wold BJ, Rothenberg EV. Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. Cell. 2012;149:467-82. Medline:22500808 doi:10.1016/j.cell.2012.01.056

35 Ruppert V, Maisch B. Molecular signatures and the study of gene expression profiles in inflammatory heart diseases. Herz. 2012;37:619-26. Medline:22918566 doi:10.1007/s00059-012-3662-5

36 Faustino RS, Behfar A, Perez-Terzic C, Terzic A. Genomic chart guiding embryonic stem cell cardiopoiesis. Genome Biol. 2008;9:R6. Medline:18184438 doi:10.1186/gb-2008-9-1-r6

37 Wang A, Huang K, Shen Y, Xue Z, Cai C, Horvath S, et al. Functional modules distinguish human induced pluripotent stem cells from embryonic stem cells. Stem Cells Dev. 2011;20:1937-50. Medline:21542696 doi:10.1089/scd.2010.0574

38 Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. Nature. 2012;489:101-8. Medline:22955620 doi:10.1038/nature11233

39 Yates LA, Norbury CJ, Gilbert RJ. The long and short of microRNA. Cell. 2013;153:516-9. Medline:23622238 doi:10.1016/j.cell.2013.04.003

40 Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. Cell. 2013;152:1298-307. Medline:23498938 doi:10.1016/j.cell.2013.02.012

41 Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics. 2009;25:1105-11. Medline:19289445 doi:10.1093/bioinformatics/btp120

42 Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28:511-5. Medline:20436464 doi:10.1038/nbt.1621

43 Han H, Irimia M, Ross PJ, Sung HK, Alipanahi B, David L, et al. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature. 2013;498:241-5. Medline:23739326 doi:10.1038/nature12270

44 Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 2011;25:1915-27. Medline:21890647 doi:10.1101/gad.17446611

45 van Rooij E, Olson EN. MicroRNA therapeutics for cardiovascular disease: opportunities and obstacles. Nat Rev Drug Discov. 2012;11:860-72. Medline:23080337 doi:10.1038/nrd3864

46 Mercer TR, Gerhardt DJ, Dinger ME, Crawford J, Trapnell C, Jeddeloh JA, et al. Targeted RNA sequencing reveals the deep complexity of the human transcriptome. Nat Biotechnol. 2012;30:99-104. Medline:22081020 doi:10.1038/nbt.2024

47 Wang J, Rao S, Chu J, Shen X, Levasseur DN, Theunissen TW, et al. A protein interaction network for pluripotency of embryonic stem cells. Nature. 2006;444:364-8. Medline:17093407 doi:10.1038/nature05284

48 Pardo M, Lang B, Yu L, Prosser H, Bradley A, Babu MM, et al. An expanded Oct4 interaction network: implications for stem cell biology, development, and disease. Cell Stem Cell. 2010;6:382-95. Medline:20362542 doi:10.1016/j.stem.2010.03.004

49 van den Berg DL, Snoek T, Mullin NP, Yates A, Bezstarosti K, Demmers J, et al. An Oct4-centered protein interaction network in embryonic stem cells. Cell Stem Cell. 2010;6:369-81. Medline:20362541 doi:10.1016/j.stem.2010.02.014

50 Hack CJ. Integrated transcriptome and proteome data: the challenges ahead. Brief Funct Genomic Proteomic. 2004;3:212-9. Medline:15642185 doi:10.1093/bfgp/3.3.212

51 O'Brien RN, Shen Z, Tachikawa K, Lee PA, Briggs SP. Quantitative proteome analysis of pluripotent cells by iTRAQ mass tagging reveals post-transcriptional regulation of proteins required for ES cell self-renewal. Mol Cell Proteomics. 2010;9:2238-51. Medline:20513800 doi:10.1074/mcp.M110.000281

52 Lu R, Markowetz F, Unwin RD, Leek JT, Airoldi EM, MacArthur BD, et al. Systems-level dynamic analyses of fate change in murine embryonic stem cells. Nature. 2009;462:358-62. Medline:19924215 doi:10.1038/nature08575

53 Phanstiel DH, Brumbaugh J, Wenger CD, Tian S, Probasco MD, Bailey DJ, et al. Proteomic and phosphoproteomic comparison of human ES and iPS cells. Nat Methods. 2011;8:821-7. Medline:21983960 doi:10.1038/nmeth.1699

54 Munoz J, Low TY, Kok YJ, Chin A, Frese CK, Ding V, et al. The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. Mol Syst Biol. 2011;7:550. Medline:22108792 doi:10.1038/msb.2011.84

55 Hansson J, Rafiee MR, Reiland S, Polo JM, Gehring J, Okawa S, et al. Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. Cell Rep. 2012;2:1579-92. Medline:23260666 doi:10.1016/j.celrep.2012.10.014

56 Folmes CD, Nelson TJ, Martinez-Fernandez A, Arrell DK, Zlatkovic Lindor J, Dzeja PP, et al. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. Cell Metab. 2011;14:264-71. Medline:21803296 doi:10.1016/j.cmet.2011.06.011

57 Zlatkovic-Lindor J, Arrell DK, Yamada S, Nelson TJ, Terzic A. ATP-sensitive K(+) channel-deficient dilated cardiomyopathy proteome remodeled by embryonic stem cell therapy. Stem Cells. 2010;28:1355-67. Medline:20572010 doi:10.1002/stem.465

58 Lindon JC, Nicholson JK. Spectroscopic and statistical techniques for information recovery in metabonomics and metabolomics. Annu Rev Anal Chem (Palo Alto Calif). 2008;1:45-69. Medline:20636074 doi:10.1146/annurev.anchem.1.031207.113026

59 Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: the apogee of the omics trilogy. Nat Rev Mol Cell Biol. 2012;13:263-9. Medline:22436749 doi:10.1038/nrm3314

60 Panopoulos AD, Yanes O, Ruiz S, Kida YS, Diep D, Tautenhahn R, et al. The metabolome of induced pluripotent stem cells reveals metabolic changes occurring in somatic cell reprogramming. Cell Res. 2012;22:168-77. Medline:22064701 doi:10.1038/cr.2011.177

61 Yanes O, Clark J, Wong DM, Patti GJ, Sanchez-Ruiz A, Benton HP, et al. Metabolic oxidation regulates embryonic stem cell differentiation. Nat Chem Biol. 2010;6:411-7. Medline:20436487 doi:10.1038/nchembio.364

62 Priori R, Scrivo R, Brandt J, Valerio M, Casadei L, Valesini G, et al. Metabolomics in rheumatic diseases: The potential of an emerging methodology for improved patient diagnosis, prognosis, and treatment efficacy. Autoimmun Rev. 2013;12:1022-30. Medline:23688955 doi:10.1016/j.autrev.2013.04.002

63 Xia J, Broadhurst DI, Wilson M, Wishart DS. Translational biomarker discovery in clinical metabolomics: an introductory tutorial. Metabolomics: Official journal of the Metabolomic Society. 2013;9:280-99. Medline:23543913 doi:10.1007/s11306-012-0482-9

64 Wilcken B, Wiley V, Hammond J, Carpenter K. Screening newborns for inborn errors of metabolism by tandem mass spectrometry. N Engl J Med. 2003;348:2304-12. Medline:12788994 doi:10.1056/NEJMoa025225

65 Folmes CD, Dzeja PP, Nelson TJ, Terzic A. Metabolic plasticity in stem cell homeostasis and differentiation. Cell Stem Cell. 2012;11:596-606. Medline:23122287 doi:10.1016/j.stem.2012.10.002

66 Zhang J, Nuebel E, Daley GQ, Koehler CM, Teitell MA. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. Cell Stem Cell. 2012;11:589-95. Medline:23122286 doi:10.1016/j.stem.2012.10.005

67 Folmes CD, Nelson TJ, Terzic A. Energy metabolism in nuclear reprogramming. Biomark Med. 2011;5:715-29. Medline:22103608 doi:10.2217/bmm.11.87

68 Cezar GG, Donley EL. Stemina biomarker discovery. Regen Med. 2008;3:665-9. Medline:18729791 doi:10.2217/17460751.3.5.665

69 Cezar GG, Quam JA, Smith AM, Rosa GJ, Piekarczyk MS, Brown JF, et al. Identification of small molecules from human embryonic stem cells using metabolomics. Stem Cells Dev. 2007;16:869-82. Medline:18042039 doi:10.1089/scd.2007.0022

70 West PR, Weir AM, Smith AM, Donley EL, Cezar GG. Predicting human developmental toxicity of pharmaceuticals using human embryonic stem cells and metabolomics. Toxicol Appl Pharmacol. 2010;247:18-27. Medline:20493898 doi:10.1016/j.taap.2010.05.007

71 Faustino RS, Terzic A. Bioinformatic networks: molecular reticles for pinpointing pharmacological target selection. Clin Pharmacol Ther. 2008;84:543-5. Medline:18941454 doi:10.1038/clpt.2008.181

72 Arrell DK, Terzic A. Network systems biology for drug discovery. Clin Pharmacol Ther. 2010;88:120-5. Medline:20520604 doi:10.1038/clpt.2010.91

73 Gulbahce N, Yan H, Dricot A, Padi M, Byrdsong D, Franchi R, et al. Viral perturbations of host networks reflect disease etiology. PLOS Comput Biol. 2012;8:e1002531. Medline:22761553 doi:10.1371/journal.pcbi.1002531

74 Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature. 1998;393:440-2. Medline:9623998 doi:10.1038/30918

75 Albert R. Scale-free networks in cell biology. J Cell Sci. 2005;118:4947-57. Medline:16254242 doi:10.1242/jcs.02714

76 Albert R, DasGupta B, Hegde R, Sivanathan GS, Gitter A, Gursoy G, et al. Computationally efficient measure of topological redundancy of biological and social networks. Phys Rev E Stat Nonlin Soft Matter Phys. 2011;84:036117. Medline:22060466 doi:10.1103/PhysRevE.84.036117

77 Vidal M, Cusick ME, Barabasi AL. Interactome networks and human disease. Cell. 2011;144:986-98. Medline:21414488 doi:10.1016/j.cell.2011.02.016

78 Faustino RS, Terzic A. Interactome of a cardiopoietic precursor. J Cardiovasc Transl Res. 2008;1:120-6. Medline:20559906 doi:10.1007/s12265-008-9019-z

79 Nocq J, Celton M, Gendron P, Lemieux S, Wilhelm BT. Harnessing virtual machines to simplify next generation DNA sequencing analysis. Bioinformatics. 2013;29:2075-83. Medline:23786767 doi:10.1093/bioinformatics/btt352

80 Kilpivaara O, Aaltonen LA. Diagnostic cancer genome sequencing and the contribution of germline variants. Science. 2013;339:1559-62. Medline:23539595 doi:10.1126/science.1233899

81 Altelaar AF, Munoz J, Heck AJ. Next-generation proteomics: towards an integrative view of proteome dynamics. Nat Rev Genet. 2013;14:35-48. Medline:23207911 doi:10.1038/nrg3356

82 Chen R, Snyder M. Promise of personalized omics to precision medicine. Wiley Interdiscip Rev Syst Biol Med. 2013;5:73-82. Medline:23184638 doi:10.1002/wsbm.1198

83 Baker M. Gene data to hit milestone. Nature. 2012;487:282-3. Medline:22810669 doi:10.1038/487282a

84 Chen J, Qian F, Yan W, Shen B. Translational biomedical informatics in the cloud: present and future. Biomed Res Int. 2013;2013:658925. Medline:23586054

85 Rung J, Brazma A. Reuse of public genome-wide gene expression data. Nat Rev Genet. 2013;14:89-99. Medline:23269463 doi:10.1038/nrg3394

86 Zerbino DR, Paten B, Haussler D. Integrating genomes. Science. 2012;336:179-82. Medline:22499938 doi:10.1126/science.1216830

87 Schadt EE, Linderman MD, Sorenson J, Lee L, Nolan GP. Computational solutions to large-scale data management and analysis. Nat Rev Genet. 2010;11:647-57. Medline:20717155 doi:10.1038/nrg2857

88 Kristensen LP, Chen L, Nielsen MO, Qanie DW, Kratchmarova I, Kassem M, et al. Temporal profiling and pulsed SILAC labeling identify novel secreted proteins during ex vivo osteoblast differentiation of human stromal stem cells. Mol Cell Proteomics. 2012;11:989-1007. Medline:22801418 doi:10.1074/mcp.

M111.012138

89 Mendoza-Parra MA, Walia M, Sankar M, Gronemeyer H. Dissecting the retinoid-induced differentiation of F9 embryonal stem cells by integrative genomics. Mol Syst Biol. 2011;7:538. Medline:21988834 doi:10.1038/msb.2011.73

90 Cashman CR, Lazzerini Ospri L. Induced pluripotent stem cells and motor neuron disease: toward an era of individualized medicine. J Neurosci. 2013;33:8587-9. Medline:23678103 doi:10.1523/JNEUROSCI.1062-13.2013

91 Sinnecker D, Goedel A, Laugwitz KL, Moretti A. Induced pluripotent stem cell-derived cardiomyocytes: a versatile tool for arrhythmia research. Circ Res. 2013;112:961-8. Medline:23569105 doi:10.1161/CIRCRESAHA.112.268623

92 Ferreira LM, Mostajo-Radji MA. How induced pluripotent stem cells are redefining personalized medicine. Gene. 2013;520:1-6. Medline:23470844 doi:10.1016/j.gene.2013.02.037

93 Henney AM. The promise and challenge of personalized medicine: aging populations, complex diseases, and unmet medical need. Croat Med J. 2012;53:207-10. Medline:22661132 doi:10.3325/cmj.2012.53.207

94 von Wulfingen BB. From re-pair and re-production to (re) generation: bio-objects as indicators of cultural change. Croat Med J. 2012;53:502-4. Medline:23100213 doi:10.3325/cmj.2012.53.502

95 Jeck WR, Siebold AP, Sharpless NE. Review: a meta-analysis of GWAS and age-associated diseases. Aging Cell. 2012;11:727-31. Medline:22888763 doi:10.1111/j.1474-9726.2012.00871.x

96 Van Hoof D, Heck AJ, Krijgsveld J, Mummery CL. Proteomics and human embryonic stem cells. Stem Cell Res. 2008;1:169-82. Medline:19383398 doi:10.1016/j.scr.2008.05.003

97 Morgan AA, Khatri P, Jones RH, Sarwal MM, Butte AJ. Comparison of multiplex meta analysis techniques for understanding the acute rejection of solid organ transplants. BMC Bioinformatics. 2010;11 Suppl 9:S6. Medline:21044364 doi:10.1186/1471-2105-11-S9-S6

98 Patti GJ, Tautenhahn R, Siuzdak G. Meta-analysis of untargeted metabolomic data from multiple profiling experiments. Nat Protoc. 2012;7:508-16. Medline:22343432 doi:10.1038/nprot.2011.454