nature publishing group

Pediatric systems medicine: evaluating needs and opportunities using congenital heart block as a case study

Jesper Tegnér[1] and Imad Abugessaisa[1]

Medicine and pediatrics are changing and health care is moving from being reactive to becoming preventive. Despite rapid developments of new technologies for molecular profiling and systems analysis of diseases, significant hurdles remain. Here, we use the clinical setting of congenital heart block (CHB) to uncover and illustrate key informatics challenges impeding the development of a systems medicine approach emphasizing the prevention and prediction of disease. We find that there is a paucity of useful bioinformatics tools enabling the integrative analysis of different databases of molecular information and clinical sources in a disease context such as CHB, contrasting with the current emphasis on developing bioinformatics tools for the analysis of individual data types. Moreover, informatics solutions for managing data, such as the Integrating Biology and the Bedside (i2b2) or Stanford Translational Research Integrated Database Environment, require serious software engineering support for the maintenance and import of data beyond the capabilities of clinicians working with CHB. Hence, there is an urgent unmet need for user-friendly tools facilitating the integrative analysis and management of omics data and clinical information. Pediatrics represents an untapped potential to execute such a systems medicine program in close collaboration with clinicians and families who are keen to do what is needed for their children to prevent and predict diseases and nurture wellness.

THE LANDSCAPE OF MEDICINE AND PEDIATRICS IS UNDERGOING REVOLUTIONARY TRANSFORMATIONS: OPPORTUNITIES AND CHALLENGES

Medicine is changing and health care is on the verge of shifting from being reactive to becoming preventive. Drivers in this development include new technologies for molecular profiling and computational tools facilitating a systems analysis of disease. Systems biology approaches to uncover regulatory networks in biological model systems such as *Escherichia coli* and yeast have made significant contributions to our understanding and appreciation of biological complexity during the past decade. This body of work has produced new exciting network-based methods for how to connect, model, and analyze large-scale molecular data (1). It has been increasingly realized that these techniques for network analysis and modeling approaches can also be used to understand human diseases (2,3). Hence, the application of a systems biology approach in medical research and clinical practice has defined the rise of systems medicine over the past years. A large European FP7 consortium has recently (2012) joined forces in a Coordinating Action Systems Medicine (http://www.casym.eu/) to develop a roadmap (2016) for strategy and implementation of systems medicine. The core concept of a systems medicine approach is to intervene at an early stage to prevent the occurrence and reduce the suffering of the effects of disease, in contrast to chiefly targeting reactive measures only following the occurrence of disease. Such a vision has also been eloquently articulated by Hood and Flores (4) using the concept of "P4" medicine—a personalized, predictive, preventive, and participatory medicine. Personalized medicine has been in focus in the medical sciences since the completion of the draft human genome with the expectation that genomics could provide a basis for individualized treatment. The notion of participation is well in line with the growth and impact of social media in society and health. Hence, a systems medicine approach embraces and includes programs such as P4 medicine and personalized medicine. Potential benefits include not only early detection and prediction of disease but also stratification of patients into subgroups that enable the selection of optimal therapy, early assessment of individual drug responses thereby reducing adverse drug reactions, improvement of clinical trials by reduction of exposure time and failure rate, and development of tools enabling the clinician to shift the emphasis from reaction to prevention and from disease to wellness.

Of note, such a systems approach has recently been utilized in the molecular profiling of Michael Snyder (i.e., the Snyderome). In this approach, referred to as integrative personal omics profiling, transcripts, proteins, and antibodies were monitored over time in conjunction with the analysis of genetic variants and the state of health (5). This pioneering study revealed an unexpected dynamics of the molecular parts over time, such as RNA editing, as well as proving useful for the early detection of diabetes. Yet, despite these and other glimpses of success, there are significant hurdles to realizing systems medicine. In this review, we will first discuss the challenge(s) posed by the need to integrate different kinds of molecular data. However, the difficulty of identifying relevant data for integration in pediatric research projects and clinical practice is augmented by the

[1]Department of Medicine, Unit of Computational Medicine, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Solna, Stockholm, Sweden. Correspondence: Jesper Tegnér (jesper.tegner@ki.se)

Received 8 November 2012 accepted 22 January 2013; advance online publication 27 February 2013. doi:10.1038/pr.2013.19

Copyright © 2013 International Pediatric Research Foundation, Inc.

fact that the major data production that is publicly available is performed on cellular and animal model systems, or complex diseases targeting middle-aged or elderly people. The second, closely related translational challenge is one major unrecognized area of research: The translational informatics problem of how to access and perform complex search queries using data originating not only from molecular research but in particular clinical information currently residing within health care. We believe this is an untapped opportunity where pediatrics could take a lead.

We will illustrate these challenges and opportunities using congenital heart block (CHB) as a prototypical pediatric example. This is an area where we collaborate with Marie Wahren-Herlenius, a world authority on CHB, in addressing the challenges she and her colleagues are facing in her clinical research and practice.

CONGENITAL HEART BLOCK

CHB is a rare disease (6) occurring once in every 15,000–20,000 births. The conduction block occurs after damage to the fetal cardiac conduction system in gestational weeks 18–24 of pregnancy. CHB is a life-threatening condition; the mortality is 10–30% and survivors need pacemakers, and longitudinal studies are required (7). Half of the cases of CHB occur in the context of congenital heart disease, and in this review, we target isolated CHB. The risk for having a child with CHB is increased in women with Ro (Ro52 and/or Ro60) autoantibodies and La autoantibodies (La), where it occurs in 2–5% of all pregnancies (8). Of note, both Ro and La autoantibodies are common in women with rheumatic diseases such as systemic lupus erythematosus and Sjögren's syndrome. Because the Ro and La autoantibodies are major known risk factors for CHB, these are tested for in the clinic. In summary, there has been good progress, yet several challenges remain to predict and prevent CHB on a personalized level. Clearly, we need further advancements in technologies enabling sensitive measurements of biomarkers and improved methods for delivery of therapy in case of early detection of disease. Because both children and their mothers are involved in the disease, the family has to participate in research and care. To advance the current state of affairs, it is also evident that we need to develop informatics and bioinformatics tools meeting the challenges of molecular data integration and closing the gap between molecular research and healthcare. CHB constitutes an illustrative pediatric example for which we have some useful but fragmented knowledge and several practical issues to resolve, such as information about sample amounts and access in conjunction with research logistics, which are real concerns at the pediatric research floor.

THE CHALLENGE OF MOLECULAR DATA INTEGRATION

The sequencing of the human genome and its subsequent postgenomic acceleration of technological developments have resulted in immense large-scale data production. The technological advances have opened new windows into genomics beyond the DNA sequence (9). The vast majority of such data originate from the platforms producing a large number of different molecular data types, which is here referred to as "omics data." Moreover, a given omics data type, such as transcripts, can be measured on several platforms capturing different aspects of the transcriptome. Next-generation sequencing technologies produce omics data (DNA variations, RNA, epigenetic modifications, proteins, and metabolites) at a decreasing cost and increasing resolution as exemplified by the ENCODE project (10). Such data production at the peta/exabyte level generates enormous challenges with respect to data management, computing, security, and data analysis (11). The pace of the technology development and data production has relentlessly accelerated during the past decade and it appears that it will continue to do so over the next decade as genomics is poised to enter the clinic and pediatric practice.

However, it is not only the large volumes of DNA sequence data, storage, or computing problems generated from large-scale projects that pose new challenges. Since the surprising discovery at the turn of the century of the relatively small number of genes in the human genome, we have witnessed an explosion in the number of different types of data that can be generated today. These include DNA, single-nucleotide polymorphisms, copy-number variations, DNA methylations, protein-coding RNA, noncoding RNA, splice variants, histone modifications, nucleosome positions, transcription factors and their DNA-binding sites, transcription start sites, promoters, protein–protein interactions, protein localization, protein modifications (these are numerous), DNA-binding proteins, and metabolites. There are now varyingly mature "bioinformatics pipelines" at different stages of development available for the analysis of different types of omics: transcriptomics, proteomics, metabolomics, and the novel "seq" approaches: RNA-seq, ChIP-seq, and Methyl-seq (12–14). Furthermore, there are several different technological platforms that can measure the same data type. For example, to profile the pattern of DNA methylation, there exist more than five different technologies that capture slightly different aspects of this important "epigenetic" modification or reprogramming of DNA (15). To summarize, it is still a true bioinformatics challenge to analyze a "single" data type because there are different platforms and bioinformatics pipelines. Second, there is not yet a methodology for how to integrate different data types such as epigenetics, proteomics, and transcriptomics. Third, given that there are more than 1,400 well-curated public databases (16), we are still unsure how to efficiently and accurately integrate public information with data generated from clinical samples. There are, however, several success stories on linking genetic variants, transcripts, and disease phenotypes. Yet there are several challenges remaining on how to perform such a basic operation as pathway analysis (17).

Furthermore, at the core of a systems medicine analysis is the data integration from omics to the clinic, which requires standardization of data. This includes making data sets sharable in a standard and interoperable format. Standards are part of the infrastructure that is required for collecting, storing, retrieving, integrating, mining, and querying data generated

Copyright © 2013 International Pediatric Research Foundation, Inc.

by different researchers, clinicians, laboratories, and platforms. Moreover, the nature and amount of data generated by different omics platforms increases the need for the community to agree on guidelines, rules, and definitions for various terminologies and formats (e.g., genes and their products resulting from genome sequencing platforms). The practical implication of omics standards in translational and clinical setting is that they have the potential to facilitate the utilization of the omics approach in health care. Three types of data-sharing standards are essential:

1. Experiment description standards (how the data were generated, e.g., methods, protocols, and samples);
2. Data exchange standards (file formats, representation, e.g., structured vs. nonstructured); and
3. Terminology standards (genes and gene product names).

In functional genomics, there has been a continual effort to create standards that are agreed upon by the community. We recommend visiting BioStandard (http://biostandards.info/wiki/Main_Page), which provides consolidated information about the standards in omics and clinical databases.

These are fundamental issues of strong relevance for medicine. Yet, when trying to apply these technologies and bioinformatics pipelines in a real-world pediatric case such as CHB, the task becomes even more challenging. The bulk of public data originate from model systems or stages of disease, which are far from the situation with infants or children during development. Hence, which parts of the public data are relevant to CHB? How can we use such carefully selected public data to elucidate mechanisms of CHB or find predictive biomarkers? Similar to other pediatric diseases, genetic variants (single-nucleotide polymorphisms) have been associated with CHB, but it is yet unclear how to identify CHB-relevant genes from such disease loci because tag single-nucleotide polymorphisms only provides markers of loci and cannot identify individual genes (18). Because CHB is a rare disease, the study cohorts are as a rule small, and the availability of samples is limited, rendering omics approaches challenging. Furthermore, how can such molecular information be integrated with the finding of increased risk for CHB if the mother carries the Ro and La antibodies? How can we use the information on single-nucleotide polymorphisms and antibodies to explain or prevent heart block? From our viewpoint, there is an imbalance between current bioinformatics research on individual data types vs. required efforts on how to integrate different sources of molecular information, thereby potentially empowering translational studies targeting mechanisms or predictive biomarkers. Finally, stem cell therapy is a putative research area of high relevance for pediatric diseases including CHB (19). Instead of heart transplantation or insertion of a pacemaker, stem cell therapy may be used to rescue the tissue. However, to manipulate stem cells efficiently, we need deep molecular knowledge of their inner cellular workings, emphasizing the need for integrative bioinformatics to interpolate between distinct omics data types. Yet, analyzing the case of CHB, there are several additional challenges that are not solved by developing an integrative bioinformatics methodology:

- How to find other risk factors in addition to the Ro and La autoantibodies? Lifestyle, clinical history, and molecular parts and their combination may provide important clues.
- Why there are comorbidities with lupus erythematosus?
- Do other comorbidities exist?
- If the mother carries the Ro and La antibodies, she will be offered extra fetal surveillance during pregnancy at a pediatric cardiology unit, which can be stressful. Therefore, interviews and questionnaires to the mothers are used to assess their experiences during pregnancy (20,21). How can the healthcare be improved? Are there any correlations between different molecular markers and the experience of the mother?
- How do the CHB children thrive and develop during life?
- How is the pacemaker working?
- Do the children develop and/or have an increased risk of other diseases?

Hence, longitudinal clinical studies are required, and integration ranging from genetic variants to behavior is urgently needed. These clinical data clearly need to be collected in a systematic manner and integrated with available molecular information. This pediatric scenario emphasizes one of the most glaring gaps in translational research in general, and this is one area where pediatrics can possibly take a lead.

THE GAP BETWEEN MOLECULAR DATA AND INFORMATION IN HEALTH CARE

In parallel with the molecular data explosion, as described above, large amounts of data describing diseases, medications, environmental factors, and lifestyle-related information are generated from clinical practice and health care. Yet clinical data stored in electronic medical records are severely restricted and managed differently than research data, which is as a rule shared and available through public repositories or scientific journals. The data residing within health care, including electronic health records, images, lab data, and primary-care data, represent an underused data source that has much larger potential in translational research than is currently realized. Moreover, here we can identify important informatics research topics such as how to semiautomatically rank the different degrees of quality in such information sources; this is of importance given that it is a well-appreciated problem that the quality of data can be highly variable and that available data within health care are as a rule incomplete. In part, this is due to the lack of appropriate computing infrastructure, which could facilitate better research and clinical utilization of data sources. In particular, there is an urgent need for computing resources to connect molecular and health-care data. Yet there is a cultural dimension in this problem in that clinicians need to see the value of integration of clinical and molecular data. However, reuse of clinical data in the research setting brings data management challenges such as storage, computing, and, specifically, access restriction and control. This is complicated

Copyright © 2013 International Pediatric Research Foundation, Inc.

within a hospital system due to the numerous stakeholders and heterogeneous information technology environment, which appears to be a global rule. Yet clinical and pediatric researchers need to perform queries across different data sources (patient biosamples, genetics, and serology) and clinical data (diagnosis, medications, diseases activities, and lifestyle) from health-care facilities.

Turning to the example of CHB, we can identify the following practical needs originating from existing clinical workflows (**Figure 1**). First, families with CHB are identified and contacted for participation and a number of tasks as follows: informed consent, interviews, collection of medical records, collection of information from clinical registers, child health, school health, blood samples, DNA analysis, and serology. All this information needs to be continuously tracked over the years, which also includes revisits. Moreover, because CHB is a rare pediatric disease, clinical practice and research benefit from coordination across local counties and countries, and between countries. Clearly, to provide a secure infrastructure, serving as an enabler of these tasks is an urgent unmet need. The information needs to be deidentified, and various principal investigators and clinicians should have a secure role-based access to data. We can also expect an increased interest from children and their families in having access to limited amounts of data via smart phones and the Internet, as well as in the not too distant future using these devices for recording potentially valuable lifestyle data of benefit for translational research. Such an infrastructure in conjunction with integration with molecular data promises to facilitate research, solving the CHB problems described above, and unlocking mechanisms of CHB, thus propelling us toward a P4 medicine for CHB.

At this juncture, it is relevant to ask what tools are currently available that could support parts of this development and, perhaps as important, to ask what is currently lacking. Moreover, what implications do such a program have for assessing which bioinformatics problems are most pressing to address in relation to pediatric diseases?

Several technology platform solutions are available to manage biomedical data in translational research. However, there is no single solution that could manage the requirements from CHB as described above. Hence, in practice, the choice is between numerous different disconnected tools, which have to be managed more or less manually, and using some of the available solutions (22,23) described below, as backbones and then performing in-house professional software engineering to glue platforms together.

One of the commonly used open-source platforms is the Informatics for Integrating Biology and the Bedside (i2b2) (22). The i2b2 platform makes use of the International Statistical Classification of Diseases, 10th revision, as a taxonomic standard to classify diseases. The design principle fueling i2b2 is to provide a scalable software platform facilitating repurposing of clinical data in the research setting and to secure the access and management of patient information for research purposes. The predefined use cases that are supported by i2b2 (22), are, first, to explore patient data in order to find sets of patients that would be of interest for further research and second to make use of the detailed data provided by the electronic

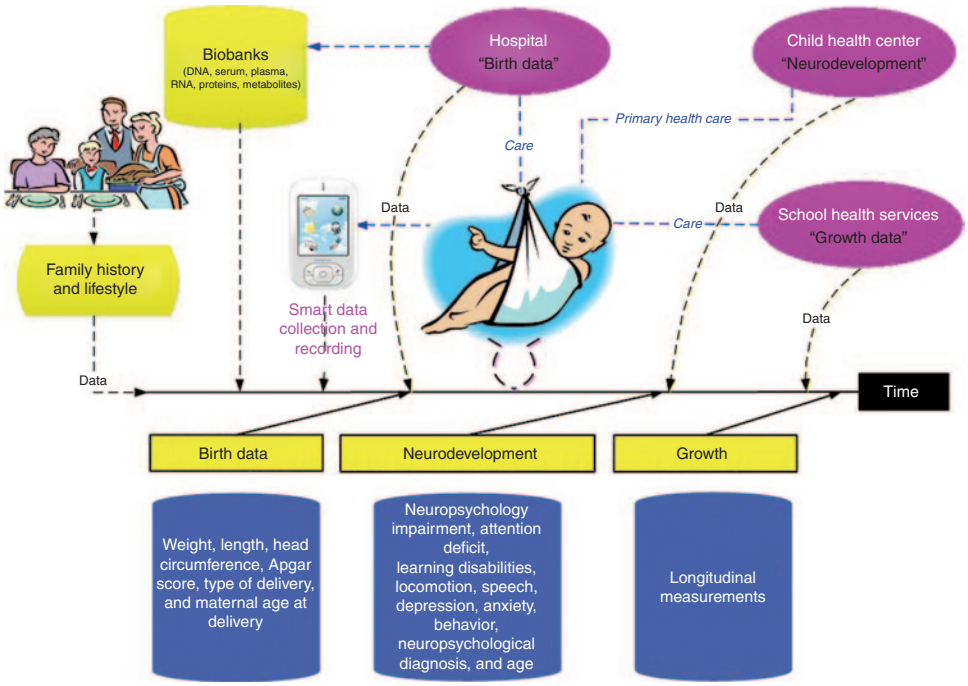

Figure 1. Flowchart diagram illustrating information flow for a CHB patient. At the center is the CHB patient receiving health care from multiple service providers. Data sources for clinical phenotypic data include hospital (birth data and biological samples are deposited in biobank for research), child health center (neurodevelopment), and school health service (growth parameters). Family history for the patient is collected via interviews and questionnaires. In the future, we may include the use of advanced tools, smart phones, and noninvasive technologies for collecting personalized data. CHB, congenital heart block.

Copyright © 2013 International Pediatric Research Foundation, Inc.

medical record to discover different phenotypes of the set of patients identified (first use case) in support of genomic, outcome, and environmental research. As a second example, we have the Stanford Translational Research Integrated Database Environment (STRIDE) (23), based on the Health Level 7 data model, representing an integrated standards–based translational research informatics platform. STRIDE provides a number of functionalities required in translational research (23), including the management of biobank data.

Using either i2b2 or STRIDE as a backbone simplifies clinical development and pediatric research but requires serious software engineering support for maintenance and communication with other bioinformatics and statistical platforms. However, neither i2b2 nor STRIDE has been designed for integrating with multiple types of next-generation sequence data. Once the clinical data, as in the case of CHB, have been captured in translational informatics platform, the clinical investigator can begin to interrogate the data using complex queries such as how many children with CHB have a mother with a La but not Ro antibodies, have a specific genetic variant, and develop neurological disease or cognitive deficits? Moreover, are there any blood samples left from these specific children over time sufficing to perform global transcriptomics? Such queries across several sources become more feasible provided there exists a proper computational infrastructure. Yet to have such an infrastructure to communicate with electronic medical record within the hospital domain is not trivial for practical, cultural, and legal reasons. Another clinical challenge is the problem of comorbidities. For example, in pediatrics, we would like to explore which comorbidities exist between CHB and other diseases. Here, we face at least three challenges. First, the task is simplified significantly if clinical data from other diseases are organized and accessible from an informatics platform. This is, however, rarely the case. Second, CHB has a single International Statistical Classification of Diseases, 9th revision code, whereas as in the case of CHB and other pediatric diseases, the emerging picture is that of several disease subtypes within a "single disease." Hence, the resolution provided by existing International Statistical Classification of Diseases versions is in many cases insufficient from a clinical standpoint, and the resulting analysis may therefore include an unknown number of false negatives. Third, we would like to assess molecular correlates underlying such comorbidities, thus access and integration with molecular data across diseases, in controlled or matched cohorts, become a critical, yet unresolved, issue.

Both i2b2 and STRIDE are examples of a data warehouse approach to consolidating and integrating multiple databases into a single database using schema translation. Data warehouses provide high access and response time performance while requiring regular updates of the source databases to keep the global data current. An alternative to having a central repository is to use a federated approach interconnecting different autonomous databases into a virtual composite database. This is a slower but highly popular mode of operation. A link-driven federation is based on a hyperlink protocol, in which the user begins by querying and retrieving a particular object and thereby obtains multiple related and/or relevant sources through uniform resource locators. An example of a link-driven system is LinkDB, a platform collecting link information from molecular biology databases (24). The relationships between biological entities are implemented as links (uniform resource locators) between objects in different data sources such as the Kyoto Encyclopedia of Genes and Genomes to signify equivalent contents (25). The federation integration approach acts as a mediator system in which the mediator translates queries from the user to match the query format in the target databases. The mediator system uses query optimization techniques to reduce the complexity of the query and to retrieve relevant and accurate results.

In conclusion, to systematically investigate a pediatric disease such as CHB, we need an infrastructure (data warehouse solution or a federated system) managing clinical data enabling complex queries. To this picture, we need to add connections to omics types of databases, either in-house or public databases, provided there are relevant bioinformatics techniques solving the problems of integrative analysis. This is not only needed in clinical care and research. The pharmaceutical company Johnson & Johnson has, for example, fused i2b2 and GenePattern, a suite of bioinformatics tools from the Broad Institute, in order to monitor clinical trials. However, the resulting system, denoted tranSMART (26), requires professional software engineering support and is not currently suited for distributed data management. Furthermore, electronic medical records or clinical or societal register information are central, as illustrated in the case of CHB. Current practice is largely limited by manual labor in transferring such information to clinical studies, provided all ethical and legal issues are solved. Yet, as demonstrated with the example of comorbidities, there are potentially huge benefits of mining electronic health records in a systematic manner and integrating such information with other clinical and molecular data. Machine-learning techniques can be used to develop decision-support tools for the clinicians, techniques for cohort querying, and clustering/stratification of patients (27). Recent work, using Alzheimer's disease as a case study, has demonstrated the feasibility of developing a decision-support system that integrates and computes using data from distinct sources (28). Key to their success was the close interaction among clinicians, software engineers, and bioinformaticians.

Finally, a major concern with using open-source software solutions is the security and authorization challenge. This requires a professional setup of the infrastructure bridging between the hospital and research environments. Authorization is essential to ensure that researchers/clinicians can only retrieve information they have been authorized to access. Study-level permission is among the best solutions. For example, researcher "X" is authorized to access the database created for a particular study. Identified patient data should not be used in a research setting, so the personal identification number must first be deidentified. The personal identification number should be removed, and therefore a new identification

Copyright © 2013 International Pediatric Research Foundation, Inc.

key has to be created to facilitate the link between patients' different records. This can be achieved using different degrees of deidentification in which direct and indirect identifiers (age) can be treated differently or in a statistical manner.

CONCLUSIONS

Clinical practice and research in pediatrics can potentially take the lead in performing developmental research in which the translational informatics challenges are addressed in conjunction with an emphasis on integration with and of molecular data. Of note, despite the fact that integrative bioinformatics targeting different omics data types and translational informatics challenges of bridging between different data sources face different subproblems, we believe that their respective designs must be considered and executed in close collaboration. Traditionally, e-health, bioinformatics, and medical informatics have been more isolated than necessary. Working closely with clinicians as in the case of CHB is key to developing practical operational solutions crossing such boundaries. This contention is strongly supported by the success story of Alzheimer's disease. Pediatrics represents an untapped potential in developing a methodology for collecting, managing, and analyzing information before and after birth together with families who are keen to do what is needed for their children. Hence, pediatrics may become one of the prime examples of P4 medicine preventing disease and nurturing wellness.

ACKNOWLEDGMENTS

We thank Marie Wahren-Herlenius and Amanda Skog for access to data and insights into the world of congenital heart block, enabling this work.

STATEMENT OF FINANCIAL SUPPORT

This research was supported by the Swedish Research Council (J.T.), Swedish Research Council, CERIC (J.T., I.A.), Swedish Research Council, SerC (J.T., I.A.), FP7 SYNERGY-COPD (J.T., I.A.), and Stockholm County Council (J.T.).

Disclosure: The authors declared no conflict of interest.

REFERENCES

1. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet 2004;5:101–13.
2. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. Cell 2011;144:986–98.
3. Hunter P, Coveney PV, de Bono B, et al. A vision and strategy for the virtual physiological human in 2010 and beyond. Philos Transact A Math Phys Eng Sci 2010;368:2595–614.
4. Hood L, Flores M. A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. N Biotechnol 2012;29:613–24.
5. Chen R, Mias GI, Li-Pook-Than J, et al. Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell 2012;148:1293–307.
6. Buyon JP, Clancy RM, Friedman DM. Cardiac manifestations of neonatal lupus erythematosus: guidelines to management, integrating clues from the bench and bedside. Nat Clin Pract Rheumatol 2009;5:139–48.
7. Bergman G, Eliasson H, Mohlkert LA, Wahren-Herlenius M, Sonesson SE. Progression to first-degree heart block in preschool children exposed in utero to maternal anti-SSA/Ro52 autoantibodies. Acta Paediatr 2012;101:488–93.
8. Skog A, Tingström J, Salomonsson S, Sonesson SE, Wahren-Herlenius M. Neurodevelopment in children with and without congenital heart block born to anti-Ro/SSA-positive mothers. Acta Paediatr 2013;102:40–6.
9. Leinonen R, Sugawara H, Shumway M; International Nucleotide Sequence Database Collaboration. The sequence read archive. Nucleic Acids Res 2011;39(Database issue):D19–21.
10. Dunham I, Kundaje A, Aldred SF, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.
11. 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.
12. Quackenbush J. From 'omes to biology. Anim Genet 2006;37:Suppl 1:48–56.
13. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 2008;5:621–8.
14. Nekrutenko A, Taylor J. Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. Nat Rev Genet 2012;13:667–72.
15. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet 2011;12:529–41.
16. Galperin MY, Fernández-Suárez XM. The 2012 Nucleic Acids Research Database Issue and the online Molecular Biology Database Collection. Nucleic Acids Res 2012;40(Database issue):D1–8.
17. Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol 2012;8:e1002375.
18. Ramos PS, Marion MC, Langefeld CD, Buyon JP, Clancy RM; International Consortium on Systemic Lupus Erythematosus Genetics. Brief report: enrichment of associations in genes with fibrosis, apoptosis, and innate immunity functions with cardiac manifestations of neonatal lupus. Arthritis Rheum 2012;64:4060–5.
19. Bernstein HS, Srivastava D. Stem cell therapy for cardiac disease. Pediatr Res 2012;71(4 Pt 2):491–9.
20. Tingström J, Barimani M, Sonesson SE, Wahren-Herlenius M, Henriksson EW. The experiences of pregnancy in women with SSA/Ro52 autoantibodies. Musculoskeletal Care 2010;8:215–23.
21. Tingström J, Welin Henriksson E, Sonesson SE, Wahren-Herlenius M. Ro52 autoantibody-positive women's experience of being pregnant and giving birth to a child with congenital heart block. Midwifery 2013;29:18–23.
22. Murphy SN, Weber G, Mendis M, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc 2010;17:124–30.
23. Lowe HJ, Ferris TA, Hernandez PM, Weber SC. STRIDE – an integrated standards-based translational research informatics platform. AMIA Annu Symp Proc 2009;2009:391–5.
24. Fujibuchi W, Goto S, Migimatsu H, et al. DBGET/LinkDB: an integrated database retrieval system. Pac Symp Biocomput 1998;683–94.
25. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28:27–30.
26. Szalma S, Koka V, Khasanova T, Perakslis ED. Effective knowledge management in translational medicine. J Transl Med 2010;8:68.
27. Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care. Nat Rev Genet 2012;13:395–405.
28. Mattila J, Koikkalainen J, Virkki A, van Gils M, Lötjönen J; Alzheimer's Disease Neuroimaging Initiative. Design and application of a generic clinical decision support system for multiscale data. IEEE Trans Biomed Eng 2012;59:234–40.

Copyright © 2013 International Pediatric Research Foundation, Inc.