

New Era for Health Care and Genomics

Momiao Xiong*

Human Genetics Center, Division of Biostatistics, University of Texas School of Public Health, Houston, TX 77030, USA

Editorial

We witness the tremendous advances in sequencing technologies and dramatic reduction of the sequencing costs. Fast and cheaper Next Generation Sequencing (NGS) technologies will generate unprecedentedly massive (millions of individuals) and highly-dimensional (dozens or even hundreds of millions) genomic and epigenomic variation data that allow nearly complete evaluation of genomic and epigenomic variation including common and rare variants, insertion/deletion, CNVs, mRNA by sequencing (RNA-seq), microRNA by sequencing (miRNA-seq), methylation by sequencing (methylation-seq) and Chip-seq [1]. This will provide not only invaluable information on fully understanding the role of human genomic and epigenomic variation and its role in complex clinical phenotypes and evolution, but also powerful tools for clinical genomics: diagnosis of disease, classification of disease subtypes, prediction of clinical outcomes, characterization of disease progression, management of health care and development of treatments, and morphological evolution.

Recent rapidly developed technologies in sensing, communications and computers are revolutionizing clinical phenotype measurements and producing a deluge of physiological, environmental and image data. These clinical phenotype data offer solutions to non-invasive disease screening and diagnosis, and assessment for drug response. We are entering a new era for clinical genomics.

To meet great conceptual, analytical and computational challenges raised by NGS and new sensing technologies, several essential issues to the success of clinical genomics should be addressed.

First issue is how to combine genomic data with electronic medical records. Clinical genomics begins with carefully specified and accurately measured clinical phenotypes. Clinical phenotypes include measurements of physiological states like Electro Cardiogram (ECG) for diagnosing heart abnormalities, Electroencephalography (EEG), Magnetoencephalography (MEG), blood pressure, weight, results of laboratory testing, medical image such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Computed Tomography (CT) and ultrasound screening, and temporal changes in clinical phenotypes [2]. All these phenotypes are digitally recorded and can be accessed. These new sensing technologies have two remarkable features. First, they can accurately uncover the structures and function of the human body which allow detecting abnormal tissues and distinguishing normal individuals from patients. For example, MRI makes use of magnetism, radio waves, and a computer to produce images of body structures, reveal the bodies deepest secrets, and might offer a faster and more accurate way to diagnose diseases [3]. MRI scans have a wide variety of applications in disease diagnosis ranging from neuroscans, cardiac to liver scans. The development and progression of disease are dynamic processes. The ECG, MEG and EEG measure the temporal phenotypes and hence capture the dynamic features of physiological processes. They provide valuable dynamic information of biological processes in human body. The space-temporal phenotypes have not been systematically investigated in assessing the relationships between the genotype and phenotype. Much information on the disease development in the current genomic and epigenomic studies of complex phenotypes has not been explored. It is indispensable to

combine genomic and epigenomic studies with electronic medical record for holistically assessing the genomic and epigenomic structure of complex diseases.

Second issue is how to analyze extremely big genomic, epigenomic and physiological data. Due to advances in sequencing technologies, sensing and communications, two major types of biological information: the digital information of the genomes and environmental signals are generated. The current magnetic resonance imaging industry is producing over 2,000 units per year. Hundreds of thousands or even millions of medical images will be produced. Scientists are creating a global alliance for sharing of genomic and clinical Data. The genomes of millions of individuals will be sequenced in the near future. The emerging genomic and healthcare data are too large and too difficult for existing analytic tools to process. Analysis of these extremely big and diverse types of data sets provide invaluable information for holistic discovery of the genetic and epigenetic structure of disease, and for prediction, prevention, diagnosis and treatment of disease, but also pose great conceptual, analytical and computational challenges. To meet these challenges, innovative approaches and parallel computational platforms should be developed. A deluge of genomic and epigenomic data generated by NGS and enormous amounts of personal clinical phenotype data demand the paradigm shift in genomic and epigenomic data analysis from standard multivariate data analysis to functional data analysis, from low dimensional data analysis to high dimensional data analysis, from single type data analysis to integrated multiple types of data analysis, and from individual PC to parallel and cloud computing. The volume and complexity of sequences data in genomics and epigenomics, and real time measured health care data and three dimensional medical image data have begun to outpace the computing infrastructures used to calculate and store genomic, epigenomic and health monitor information [4,5]. Cloud is a metaphor for the Internet. Cloud computing is a type of Internet-based computing. It is a distributing computing platform. The data processing is moved from private PC to the remote computer clusters. Users access computational resources from a vendor over the internet [4]. The cloud is virtualization technology. It divides a server's hardware resources into multiple "computer devices", each running its own operating system in isolation from the other device which presents to the user as an entirely separate computer. A typical cloud computing begins by uploading data into the cloud storage, conducts computations on a cluster of virtual machines, output the results to the cloud storage and finally download the results back to the user's local

*Corresponding author: Momiao Xiong, Human Genetics Center, Division of Biostatistics, University of Texas School of Public Health, Houston, TX 77030, USA, E-mail: Momiao.Xiong@uth.tmc.edu

Received June 23, 2013; Accepted June 26, 2013; Published June 28, 2013

Citation: Xiong M (2013) New Era for Health Care and Genomics. J Phylogen Evolution Biol 1: e104. doi:10.4172/2329-9002.1000e104

Copyright: © 2013 Xiong M. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

computer. Since the pool of computational resources available 'in the cloud' is huge, we have enough computational power to analyze large amount of data. The cloud computing has been applied to manage the physiological data such as ECG [6], the deluge of 'big sequence data' in 1000 Genomes Project [7], comparative genomics [8], Chip-seq data analysis [9], translational medicine [10], transcriptome analysis [11], and image analysis [12].

The convergence of these two fields-sequenced genomics and epigenomics, and rapidly developed sensing, communication and computer technologies, with its flurry of innovative healthcare apps is causing a revolution in health care and medicine. Although there is heated debate about whether DNA variation has value to predict diseases, millions of sequenced individuals will provide invaluable information for not only disease prediction, prevention, diagnosis and treatment, but also for population evolution studies. We can expect that the emergence of NGS technologies and new development in sensing, communication and parallel computing, and publication in the *Journal of Phylogenetics & Evolutionary Biology* will stimulate the development of innovative algorithms and novel paradigm for big genomic, epigenomic and clinical data analysis as well as population evolutionary studies.

Acknowledgement

The project described was supported by Grant 1R01AR057120-01, 1R01HL106034-01 from National Institutes of Health and NHLBI.

References

1. Newsmakers. (2013) *Science* 340: 1270-1271.
2. Belmont J, McGuire AL (2009) The futility of genomic counseling: essential role of electronic health records. *Genome Med* 1: 48.
3. Francis S Collins (2010) New imaging technology is transforming medicine: revealing the body's deepest secrets. *Parade*: 6.
4. Schatz MC, Langmead B, Salzberg SL (2010) Cloud computing and the DNA data race. *Nat Biotechnol* 28: 691-693.
5. Pechette JM (2012) Transforming health care through cloud computing. *Health Care Law Mon* 5: 2-12.
6. Hsieh JC, Hsu MW (2012) A cloud computing based 12-lead ECG telemedicine service. *BMC Med Inform Decis Mak* 12: 77.
7. Waltz E (2012) 1000 genomes on Amazon's cloud. *Nat Biotechnol* 30: 376-376.
8. Wall DP, Kudtarkar P, Fusaro VA, Pivovarov R, Patil P, et al. (2011) Cloud computing for comparative genomics. *BMC Bioinformatics* 11: 259.
9. Feng X, Grossman R, Stein L (2011) PeakRanger: a cloud-enabled peak caller for ChIP-seq data. *BMC Bioinformatics* 12: 139.
10. Dudley JT, Pouliot Y, Chen R, Morgan AA, Butte AJ (2010) Translational bioinformatics in the cloud: an affordable alternative. *Genome Med* 2: 51.
11. Chouvarine P, Cooksey AM, McCarthy FM, Ray DA, Baldwin BS, et al. (2012) Transcriptome-based differentiation of closely-related *Miscanthus* lines. *PLoS One* 7: e29850.
12. Langer SG, Persons K, Erickson BJ, Blezek D (2013) Towards a more cloud-friendly medical imaging applications architecture: a modest proposal. *J Digit Imaging* 26: 58-64.