

MEETING REPORT

A year of great leaps in genome research

Jingde Zhu*

Abstract

A report on the 6th International Conference on Genomics (ICG-VI), Shenzhen, China, 12-15 November 2011.

The International Conference of Genomics VI (<http://www.genomeconference.org/>) entertained more than 600 participants with around 90 quality oral presentations. The conference included five morning preliminary sessions and ten concurrent afternoon sessions comprehensively covering the latest breakthroughs in genomic research. For the first time, the conference introduced three panel discussions dedicated to non-academic topics, which included 'Improving biotech information flow from China', 'Ethical, legal and social issues' of genome research, and "Re-imagining" research: building new ties between industry, academia and research.

From 1,000 to 1 million genomes

The gigantic expansion over the past year in both sequencing and data handling capabilities, such as data retrieval, analysis, archiving and public accessibility, was well recognized, including its impact on today's life science research. On behalf of the Beijing Genomics Institute (BGI, Shenzhen, China), Jun Wang outlined the advances in sequencing-based research across the globe, highlighting those in which the BGI had a major role. Having become the largest sequencing center in the world over the past year, the BGI has sequenced 540 plant and animal reference genomes and 25,239 variant genomes, including 38,123 human samples (multiple samples of one variant were sequenced). In addition to the genomes of more than 600 microbial species, the BGI's achievements also included over 3,500 strains of microbes and 1,800 metagenomes. Continuing on these successes, three 'M projects' have been initiated. The 'Million Human Genomes Project' (the M1 project) was

announced as a natural extension of the 1000 Genomes Project initiated 3 years ago. The completion of the reference genomes of all living organisms is the goal of the 'Million Plant and Animal Genomes Project' (the M2 project). The third M project of the BGI is the 'The Million Micro-Ecosystem Genomes Project', which aims to sequence the metagenomes and cultured microbiomes of a spectrum of environments, such as the gut microbiomes in human and animal models.

Handling large volumes of data in genome research

One of the greatest challenges in today's research is to effectively handle the huge amounts of data being accumulated exponentially over time. Efforts to achieve early sharing of raw data have been frustrated for years, mainly by limited storage and archiving capabilities, despite researchers' willingness to release data early to the research community. Scott Edmunds (*GigaScience*, www.gigasciencejournal.com/) outlined the joint efforts of the BGI and BioMed Central to overcome this challenge with a new journal, *GigaScience*, which takes advantage of the BGI's extensive data hosting and cloud computing infrastructure to embrace a new publication format that integrates manuscript publication and data hosting. Supporting data associated with the articles will be permanently hosted in the *GigaDB* database (<http://GigaDB.org>). Moreover, these datasets are given digital object identifiers (DOIs) to enhance searching, tracking and citation.

Myles Axton (*Nature Genetics*) described his journal's carrot and stick approaches to encourage data sharing and community annotation in genetics. Microattribution is one potential way to encourage authors to share their data, and a good example was provided by Giardine *et al.* (*Nat Genet* 2011, 43:295-301): recording microcitations for data contributors to the HbVar database of hemoglobin variants and thalassemia mutations has led to a huge increase in the number of submissions since publication of the paper.

Advances in genomic research

Empowered by the greatly increased technological capabilities now available, the speakers shared their research results addressing cellular heterogeneity at genomic and gene expression levels. Jun Wang announced that

*Correspondence: zhujingde@gmail.com
Cancer Epigenetic Program, Shanghai Cancer Institute/Renji Hospital, Shanghai Jiaotong University, 200032 and The Center For Basic and Translational Epigenetic Research of Diseases, School of Life Science, Anhui Medical University, China

sequencing of single-cell genomes of several hundred cells from various regions of kidneys obtained from cancer patients has been completed, and that quantifying differences in the genome sequences among individual cells can help reveal the underlying mechanisms of kidney carcinogenesis. Richard Sandberg (Karolinska Institute, Stockholm, Sweden) demonstrated the global mRNA expression patterns of several dozen stem cells from early mouse embryos at different stages of development. Despite the limited coverage of low copy mRNA, Sandberg's research forecasts a broader use of such technologies to address more critical biological questions, particularly in systems in which cellular heterogeneity is notoriously prevalent, such as cancers and early embryos.

Genome-wide expression profiling of long non-coding RNAs was presented by Sumio Sugano (University of Tokyo, Japan). Sugano was able to comprehensively quantify the transcriptional start sites of all of the RNAs, coding and non-coding alike, utilizing oligo-capping - mapping of the transcription site. Taking advantage of second-generation sequencing technology, Sugano's team created a database for dozens of cell lines and tissues, healthy or diseased, and made them available to researchers with common interests.

Epigenetics is another rapidly advancing discipline in biomedical research today. Stephan Beck (University College London, UK) outlined a European Union initiative of a database for epigenome-wide association studies (EWAS), and presented some preliminary supportive data. Nevertheless, both conceptual and technological challenges need to be addressed in light of the huge inherent difference of EWAS from its precedent, genome-wide association studies (GWASs).

Ethical, legal and social issues

Along with the rapid technological advances, the scope and complexity of ethical, legal and social issues have expanded and now present greater challenges to genome research communities and beyond. There was an intensive discussion on subjects such as biobanking, genetic

predisposition, gene patterning, direct-to-consumer genetic testing and other related subjects by experts from relevant various disciplines.

Personal perspective

Needless to say, genome research today is facing even greater conceptual and technological challenges. There is a need for us to combine the conventional hypothesis-driven approach with the discovery-driven to the biological implication of genomic makeup in various biological systems.

I have attended this annual event consecutively for the past 5 years and witnessed the adjustments in the scientific programs in synchronization with the advances in genomics research. The willingness of the conference organizer to encompass both basic and translational genomic research and to address the important issues from all scientific, economic and social angles is strong, with an emphasis on collaboration and teamwork. In all aspects, the IGC conferences stand out among hundreds of genomics meetings each year not only by providing participants with cutting edge knowledge, but also by inspiring them to join in with current projects. If you have not attended this conference yet, the next meeting, IGC VII, will be held in the autumn of 2012 in Shenzhen, China and is well worth considering.

Abbreviations

BGI, Beijing Genomics Institute; EWAS, epigenome-wide association studies; ICG, International Conference on Genomics.

Competing interests

The author declares that they have no competing interests.

Published: 30 January 2012

doi:10.1186/gm303

Cite this article as: Zhu J: A year of great leaps in genome research. *Genome Medicine* 2012, **4**:4.