

# Editorial

To many researchers who have gone through rigorous training to conduct hypothesis-driven projects, genomics is a science that is characterised as fishing with some particularly expensive poles. This was probably accurate when genomics was in its infancy and when data generation was its primary focus. Since then, genomics has expanded and diversified, and now covers studies such as examining the architectural features of genes and chromosomes, the relationship between the genes of different species and the function of genes and their protein products. While data generation remains a vital component of genomics, we are now beginning an era in which genomics approaches are being integrated with classical genetics, clinical research and other areas.

Instead of focusing on single genes, it is possible to gather information from the entire genome in order to study basic biological processes and to describe the genetic and physiological bases of threats to human health. It is now routine practice to screen large segments of the genome for variants influencing disease predisposition or variable drug reaction, while expression profiling and proteomics hold out the promise of much finer molecular descriptions of biological processes, including pathological ones. This is the direction in which human genomics must go, but it requires further development of the techniques and analytical tools that are specifically designed for biomedical research.

The primary goal of this peer-reviewed journal, *Human Genomics*, is to promote and review the genomics approaches that are tailored to providing answers to the most important and interesting questions in biomedical research. More specifically, *Human Genomics* provides a forum for the publication of primary research and reviews on the application of genomics research in drug discovery and medicine. Recognising the critical role of genomics data analysis and modelling, the Journal will pay particular attention to the statistical and computational tools required for interpreting genomics data. The inclusion of a diverse range of computational approaches applied to genomics data in a single forum will facilitate the flow of information across different aspects of human genomics, currently a critical bottleneck in human genomics research. This philosophy is fully reflected by the expertise of the Editorial Board, which is extensive and covers bioinformatics, genomics, proteomics, neurogenetics, human genetics, clinical genetics, cancer genetics, epidemiology, genomic diversity, human evolution and primate genetics, population genetics, statistical genetics, microarrays, pharmacogenetics and evolutionary genomics.

The field of human genomics is rapidly evolving so, given this protean nature, we have opted to be inclusive in coverage and not to provide a strict description of what the Journal will cover. We certainly do not feel in a position to predict which directions are likely to be the most promising in the coming years. It is easier to recognise the kind of work that should be included than to define it. Obvious examples at the moment include the following:

- Linkage disequilibrium association mapping of complex and multifactorial diseases and traits, and relevant methods
- Comparisons of the human genome with chimpanzee, gorilla and mouse genomes
- Assessment of the functional consequences of gene variations associated with disease and drug response
- Biology of repetitive elements
- Origin and divergence of human genes
- Protein structure and human variation
- Quantitative genetics and development
- Human variation and disease
- Genomic redundancy
- Human gene families
- Statistical analyses of comparative genomics data
- Regular updates on human genome completion and annotations
- Human proteomics, as applied to specific questions as a clinical diagnostic

In this inaugural issue of the Journal, we are publishing three research papers and two reviews, these are accompanied by two book reviews, a software review and a gene annotation update.

One of the most important challenges in human genomics is to localise and identify the genetic variation underlying the diversity of phenotypes such as disease susceptibilities and response to drugs using linkage disequilibrium. The inference of gametic phase in population samples, and the assessment of association between phenotypes and multiple polymorphic sites, are pivotal steps addressed by two of the papers in this issue. Excoffier *et al.* introduce their computationally fast algorithm, ELB, for inferring gametic phase in population samples of multilocus genotypes. This algorithm is well suited to problems involving many loci in large genomic regions. The authors show, through computer simulation, that their method provides a better local estimation of gametic phase than the existing popular methods such as PHASE and HTYPER, while its global accuracy is close to the best.

In the second of these papers, Zhang *et al.* tackle a challenging problem in association studies: developing association statistics for multiple single nucleotide polymorphisms (SNPs) in genomic regions. It is striking that, despite the huge number of association studies being published every month, the vast majority of papers simply test each SNP individually for association with the phenotype, often employing a correction for multiple comparisons as if they were independent. Here, Zhang *et al.* adopt a recently developed non-parametric regression approach known as Bayesian adaptive regression splines (BARS). It integrates the linkage disequilibrium of single-locus statistics into a single test by examining the class of smooth curves consistent with the data. This robust method, which is applicable to a wide range of data types, including genotype frequencies estimated from pooled samples, constitutes a potentially powerful alternative to classical tests of association, and also circumvents the multiple testing problems inherent in those tests.

The third paper in this issue deals with a different aspect of genomics: comparative genomics. Murphy *et al.* started their quest for reconstructing homologous syntenies of mammalian ancestors of two three-species groups (human-mouse-cat and human-cat-cattle) based on the distribution of genes and markers in these species. This is the first attempt to reconstruct the evolutionary exchanges of the genome that predates modern placental mammals. The 70–100 conserved segments shared across the genomes were primarily generated by inversions and translocations.

Halder and Shriver take on the formidable challenge of reviewing the state of science of measuring admixture linkage disequilibrium and admixture mapping in populations such as

African American and Hispanic American. The stratification introduced by population admixture can create false positives (or indeed false negatives) in case-control association studies, but it can also generate useful associations in certain study designs. The authors review these aspects of population structure and include a careful discussion of methods to detect for and control stratification in case-control studies.

Makowski and Rodi present an impressive review of the current state of phage display technology that is currently used for characterising the binding repertoire of small molecule drugs. A successful application of this technology with the drug paclitaxel is carefully discussed as an example of the technical and theoretical benefits and pitfalls of the method.

Also in this issue, Dudbridge provides a concise and yet informative survey of current software for linkage analysis, while Nebert and Wain present an insightful update on human genome completion and annotations, along with a witty discussion on gene nomenclature. Two books, *DNA Microarrays and Gene Expression — From Experiments to Data Analysis and Modelling* and *Nature via Nurture*, are also reviewed.

After a glorious decade of technological progress in genomics, the quest to dissect complex human diseases and traits seems more realistic, although still daunting. The Publisher, the Editorial Board and the Managing Editor of the Journal are committed to serve and interact with the research community in building a bridge between genomics and biological questions of human genetics.

Welcome to *Human Genomics*!

Li Jin, PhD and David B. Goldstein, PhD