COMMENTARY

COMPLEXITY

From genomics to epigenomics: a loftier view of life

Stephan Beck, Alexander Olek, and Jörn Walter

One of the goals of the human genome project (HGP) is to generate reference data on which to base further studies to explain life in all its complexity. With about 30% of the HGP completed, the time has come to initiate such studies. In the past few years, we have seen functional genomics and then pharmacogenomics each become established as a discipline in its own right. Earlier this year, the new single-nucleotide polymorphism consortium was born to take comparative genomic studies to a new level. Finally, a number of companies have begun to offer large-scale gene expression maps. Although each of these areas of inquiry is fruitful in itself, we believe none will answer the most important questions about the genetic underpinnings of life.

In a commentary in this journal, Strohmann[1] defined epigenesis as one system that provides explanations for the complex functional attributes of cells and organisms, a set of "informational systems and operating rules" that complement genes and genomes. In the same month, Jones and Laird[2] argued that it may now be time to include epigenetics in efforts toward global functional analysis.

At the risk of starting a nomenclature controversy, we prefer to term our field of interest "epigenomics." Although the terms "epigenetics" and "epigenesis" have scientific and historical significance, we believe "epigenomics" better captures the whole-genome scope of the phenomena we are observing.

Many previous and ongoing efforts have successfully characterized diverse epigenetic phenomena. Most prominent among these are recent studies on imprinting[3], on metabolic networks[4], on genetic hierarchies in embryonic development[5], and—perhaps most recently—on epigenetic mechanisms of gene activation in cancer[2,6,7].

Our belief is that an approach that views these and other such complex phenotypes from the genomic level down, rather than from the genetic level up, can provide powerful insights into the functional interrelationships of genes in health and disease. Consequently, together with the Centre National de Genotypage (CNG, Paris), the Technical University of Berlin (TU-Berlin), the German Cancer Research Centre (DKFZ, Heidelberg), The Society for Biotechnological Research (GBF, Braunschweig), Biopsytec GmbH (Berlin), and Medeea GmbH (Hamburg), we have formed a European alliance, the Human Epigenome Consortium (HEC).

The coordinating member of this network—a Berlin-based biotechnology company—chose the name "Epigenomics GmbH" to indicate its confidence in the promise of the new field. The company will provide enabling technologies for large-scale parallel epigenomic analysis by conventional high-throughput sequencing of bisulfite-treated DNA, high-density oligomer arrays for the detection of methylation signals, and DNA–mass spectrometry. This combination of technologies when fully developed will provide an organism's "epigenotype." The consortium plans to engage in several related inquiries:

Genome-scale mapping of the methylation status of CpG dinucleotides. This methylation map will be produced from selected sets of complex genome amplifications on DNA isolated from both healthy and selected diseased tissues. Mapping will begin with well-defined fragment populations of several thousand individual regions per experiment (e.g., from known genes) and include high-density snapshots of other regions spread evenly over the genome. Methylation sites were selected because they are the best-known epigenetic signals residing in genomic DNA, and we expect that methylation footprints on the chromatin will also yield valuable information on genome stability and overall chromatin packaging.

Identification and analysis of epigenomic loci in the MHC. The human major histocompatibility complex (MHC) presents an ideal model system for epigenomic pilot studies. The most polymorphic and gene-rich region of the human genome, the MHC is also associated with more diseases than any other genomic region. Yet the causes and mechanisms (some of which are likely to be of epigenomic origin) for most of these diseases are still unknown. Some MHC-linked genes, such as olfactory receptor genes, are also known to regulate expression through allelic inactivation. Differential allelic methylation has been suggested to be responsible for this phenomenon. Recently, the complete genomic sequence of the human MHC has become available, together with that of the chicken MHC and most of the mouse MHC. These sequences will be the basis for HEC participants and others to generate and compare, for example, high-density methylation patterns of orthologous and disease-associated regions.

Comparative analysis of epigenomic information from different organisms. This will parallel international efforts of comparative genomic sequencing as promoted by the US National Institutes of Health (Bethesda, MD) using the mouse as a model system. In the first studies, mouse and human methylation patterns will be compared in selected regions of interest and extended as the mouse sequencing project progresses. The mouse model serves as an excellent system to study regulation of developmental processes, induced pathological effects (like tumorigenesis), as well as environmental influences on epigenetic programs. Last, but not least, the whole spectrum of mutations produced in the mouse (e.g., knockout mice or mouse mutation models arising from the large-scale ENU mutagenesis programs launched in Europe and the United States) will provide an especially fertile area for differential epigenomic studies.

Initial studies will compare mouse and human methylation patterns for two selected regions, the MHC and an imprinted genomic chromatin domain, but once the techniques are established it will be possible—and desirable—to study the relationships of higher primates and other closely related species. Epigenetic information is amenable to much faster changes than the DNA sequence itself. Therefore, such studies might allow the isolation of the epigenomic drivers of recent evolution. For example, epigenesis could be used to explain the relatively rapid divergence of humans from other species.

The projects described here are just the beginning of what will undoubtedly be a long and fertile period of epigenomic research. It is hoped that such studies will ultimately reveal how epigenetic variation drives not only pathology, but ontogeny and phylogeny as well.

Stephan Beck is head of sequencing at The Sanger Centre, Cambridge, UK (beck@sanger.ac.uk). Alexander Olek is CEO of Epigenomics GmbH, Berlin, Germany (Olek@epigenomics.com). Jörn Walter is group leader at the Max Planck Institute for Molecular Genetics, Berlin, Germany (walter@mpimg-berlin-dahlem.mpg.de).

1. Strohmann, R.C. *Nat. Biotechnol.* **17**, 112 (1999).
2. Jones, P.A. & Laird, R.W. *Nat. Genet.* **21**, 163 (1999).
3. Reik, W. & Walter, J. *Curr. Opin. Dev. Biol.* **8**, 154 (1998).
4. Heinrich, R. & Schuster, S. *Biosystems* **47**, 61 (1998).
5. Gasser, S.M. et al. *Cell. Mol. Life Sci.* **54**, 1 (1998).
6. Rasnick, D. & Duesberg, P.H. *Biochem. J.* **340**, 621–630 (1999).
7. Reik, W. et al. *Nat. Genet.* **23**, 380–382 (1999).