Systems Epigenomics for Cancer Detection and Diagnosis

Tim H.-M. Huang, Ph.D.

Division of Human Cancer Genetics
Department of Molecular Virology, Immunology, and Medical Genetics
The Ohio State University
420 W. 12th Street
Columbus, Ohio 43210

Abstract:

The development of high-throughput technologies has provided powerful tools for comprehensive analysis of gene expression, chromosomal alterations, and nucleotide mutations in cancer genomes. These technologies generate a vast amount of molecular information, which can be used for data mining and computational modeling. The systems-scale approach provides a new paradigm for today's research of complex cancer genomes. The outcomes are potentially enormous for the improvement of cancer diagnosis and prognosis. Until recently, such a high-throughput technological development had been lacking for the emerging phenomenon of epigenetic alteration in cancer. Unlike genetic mutations, this molecular change does not affect DNA sequences and yet can stably influence gene expression during cancer development.

Our laboratory has developed a novel microarray platform for comprehensive interrogation of altered DNA methylation in solid tumors and leukemia. This chemical change is an epigenetic hallmark associated with transcriptional silencing of multiple genes. In addition, the microarray platform can be used to analyze chromatin remodeling in the regulatory regions of multiple genes. Altered chromatin structure usually occurs at the protruding histone tails that are shown to have a profound influence on genes' transcriptional activities.

This epigenetic microarray has been used to define molecular subtypes of primary breast and ovarian tumors. Computational modeling has further uncovered molecular sequences of chromatin repressors being recruited to establish epigenetic silencing of tumor suppressor genes. Phylogenetic clustering algorithms have been used to recapitulate complex DNA methylation patterns related to tumor progression. This systems approach has led to the identification of novel genes, the hypermethylation of which is associated with chemo-resistance in ovarian cancer.

To facilitate data sharing of our epigenetic microarray system, we have established two databases, ArrayAnnotator and ChipArray, through the support of NCI's Integrative Cancer Biology Program. These databases are developed via the NCI caBIG (cancer Biomedical Informatics Grid) infrastructure. Specifically, ArrayAnnotator has online access to gene annotation and methylation probe sequences while ChipArray is an integrated system for epigenetic microarray data storage, retrieval, visualization and analysis. Genome Data Visualization Toolkit (GDVTK) is the key framework of this data structure. Within GDVTK, we have used Java™ tools for back-end data-management and front-end visual presentation on the web. These databases have a web interface for researchers to search and download methylation microarray data. Furthermore, the goal of this integrated and collaborative research is to establish a common set of cancer models and vocabularies and standards for epigenomic and genomic data in multi-institutional environment.