Application of the genome topographer system to the study of complex human disease.

Thomas G. Marr
marr@cshl.org

Cold Spring Harbor Laboratory
1 Bungtown Road, Cold Spring Harbor, NY 11724, U.S.A.

The next phase of the Human Genome Project (HGP) will be characterized by the availability of (1) significant amounts of DNA sequence from human and model organisms, (2) physical maps of increasing resolution of the human and mouse genomes using standardized, verifiable mapping technologies, (3) high density maps of polymorphic genetic markers for the human and mouse genomes, and (4) major advances in methods for high throughput genotyping, positional cloning and/or exon trapping, and sequencing.

The methods and materials resulting from the HGP are well suited to meet the challenges in the study of complex human genetic diseases, which require the ability to detect and sort out heterogeneous, polygenic, and epistatic genetic effects, modulated by environmental factors. With these issues in mind, we have developed the Genome Topographer (GT) informatics system. The main focus of system development currently is:

1. To optimize the ability of users of GT to gather and merge highly detailed genetic, molecular and biochemical data from the major genome centers, various public databases (e.g. GenBank, GDB, CHLC, Mouse Genome Database, various ACeDB-based files), and the scientific literature.

2. To integrate a state-of-the-art sequence analysis and interpretation tool within the GT framework, called Sequence Analyst (SA). SA is designed to help the user perform and interpret sensitive sequence analyses using methods that are known to be mathematically optimal, while remaining compatible with existing approaches, such as BLAST and FASTA.

3. To integrate a new tool called Genetic Analyst (GA). GA will provide a way to store, display, analyze, and manipulate data required for genetic segregation and linkage analyses.

4. To maximize the effectiveness of annotation of data as an aid in the design of key experiments needed to understand gene function and dysfunction.

In order to test the effectiveness of the design of GT, we have begun application of GT to study two complex genetic diseases, Manic Depressive disorder and cancer.

Note: While I will be presenting this work, it is the result of work by many individuals in my laboratory, including Lisa Catapano, William Chang, Steven Cozza, Martha Hiller, Rebecca Koskela, Moira Mallison, Corp Reed, Jacqueline Salit, Adam Tracy.