**Aqwa grants**

**06-HG-101 New computational and statistical methods for the analysis of large data sets from next-generation sequencing technologies.**

The introduction of new methods for DNA sequencing has opened new avenues, including large-scale sequencing studies, metagenomics, transcriptomics, genetic network analysis, and determination of the relationship of sequence variation and phenotypes to disease, to address heretofore unapproachable problems in biomedical research. However, since the large amounts (terabases) of data generated overwhelm existing computational resources and analytic methods, urgent action is needed to enable the translation of this rich new source of genomic information into medical benefit. Contact: Dr. Lisa Brooks, 301 496-7531, brooksl@mail.nih.gov

**06-GM-115 High-end computing software. Upgrading of biomedical computing software to high-end computing (HEC).**

This developmental effort will seek to expand the domain areas to the macromolecular, cell, tissue, organ, whole-organism, and population levels. The program would support grants to upgrade and port software to run and perform experiments on new generation HEC supercomputers. Contact: Dr. Peter Lyster, 301-594-3928, lysterp@mail.nih.gov

**Related grants**

**06-CA-106 Data integration and visualization methods and tools.**

Cancer research is increasingly complex and data-rich. In order for biologists to view their data in the context of other similar data and to view it against the complex background of other data types, new data integration and visualization methods are needed. These can be in the form of software modules that can be plugged into existing portals or viewers and can include the adaptation of existing data visualization and integration methods now tailored to cancer research. Contact: Dr. Jennifer Couch, 301-435-5226, couchj@mail.nih.gov

**06-CA-111 Integrative analysis of genomic data sets generated by TCGA and TARGET.**

Methods for the unsupervised analysis of large and varied data sets that are predictive of cancer formation and can determine regulatory points in pathways and circuits. Contact: Dr. Joseph Vockley, 301-435-3881, vockleyj@mail.nih.gov

**06-CA-112 Development of high throughput mechanisms for genomic analysis.**

This includes methods to improve the throughput of next gen methods for genomic analysis. Methods could be either laboratory based or bioinformatics based improvements with the goal of decreasing the amount of time it takes to analyze a sample. Contact: Dr. Joseph Vockley, 301-435-3881, vockleyj@mail.nih.gov

**06-AG-106 New computational and statistical methods for the analysis of large data sets from genome-wide association studies (GWAS) and the use of next-generation sequencing technologies.**

Develop new tools to enable the translation of vast amounts of genomic information into medical benefit to address large amounts of data generated (e.g., terabases of sequence) that overwhelm existing computational resources and analytic methods. These new approaches include very large-scale genotyping and sequencing studies, metagenomics, transcriptomics, and genetic network analysis.  
Contact: Dr. Marilyn Miller, 301-496-9350, MillerM@nia.nih.gov

**Exome grants**

**08-DA-102 Improved Bioinformatics Analysis for Deep Sequencing.**

The current estimate of sequencing an entire human genome is $5000 and can be accomplished in a few months. However, current bioinformatic and analytic capabilities are inadequate to analyze the volumes of data that would be generated by deep sequencing many individuals. Specifically, RC1 applications are sought to (1) optimize base calls from next-generation sequencing machines, (2) develop and improve optimal alignment/mapping methods that tackle uncertainty and multiple potential placements, (3) identify methods for SNP calling from multiple reads and multiple samples, (4) identify copy-number variation calling from next-generation sequencing data, and (5) develop automated methods for searching sequence databases that could be used to give probabilities that a variant is real. Contact: Dr. Jonathan D. Pollock, 301-435-1309, jpollock@mail.nih.gov