

manage & analyze your genetic data

Disease Miner™ Professional

With advances in high-throughput SNP genotyping technologies, researchers now need to manage and analyze vast amounts of data. The Disease Miner Professional solution shortens the time to discovery and enables clients to leverage their investments in genotyping and phenotype collection.

Added value for genotyping customers

Disease Miner is a highly advanced and user-friendly software tool designed for large scale genetic research. It has an intuitive graphical interface for data visualization, which makes it easy for researchers to perform statistical analysis and interpret results, even for those who are not experts in statistical genetics.

The system has a spread-sheet like report tool for data visualization, analysis, imports and exports, and a customizable folder structure for storing objects such as patient reports, pedigrees, analysis jobs, etc.

The Disease Miner client is Java based with a window manager for organizing multiple analysis panels and includes query tools and a framework for server side parallelization, queuing, and storage of computational intensive analysis jobs.

Disease Miner Professional allows researchers to:

- Easily manage and query large phenotype, genealogy and genotype datasets
- Quickly perform statistical analysis using proven statistical algorithms
- Perform genome-wide association of phenotypes with genotypes and copy number variations
- Analyze LD-blocks and perform multi-point haplotype association
- Explore statistical results, subject genotypes and CNVs together with genome annotations
- Calculate genotype statistics and easily filter subjects and markers based on quality metrics



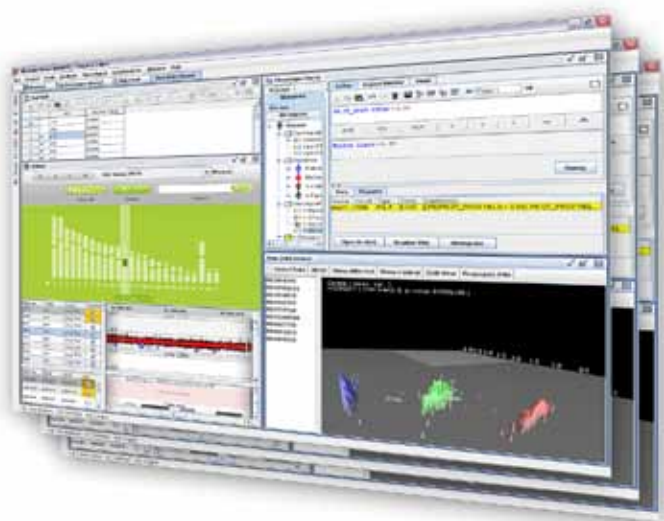
Use the Disease Miner genome browser for CNV analysis.

This example shows tracks from three subjects in a region on chromosome 15. The B allele frequency shows a clear indication of duplication for the subject in the upper track.

The Disease Miner allows you to easily search for individuals based on their CNV predictions and perform CNV case-control association analysis.



The Disease Miner Professional solution is offered to deCODE's large-volume genotyping clients. It is provided with the client's genotyping results on a powerful, preconfigured computer that runs an Oracle database and Disease Miner services on the Linux operating system. Additional data files from deCODE and other client data can be imported seamlessly.



Proof-of-principle

The Disease Miner has been developed and used by researchers at deCODE for over a decade and deployed to over two dozens of deCODE's biggest genotype clients. The statistical algorithms and the data structures used in the program have been optimized for a large number of clinical samples, as well as extremely high volumes of genotypes, to meet the output demands of modern high-throughput SNP genotyping technologies.

deCODE used the Disease Miner to quickly replicate the widely associated TCF7L2 gene for Type 2 diabetes, based on its genotypes using the Illumina HumanHap300 BeadChip on 1491 patients vs. 4712 controls. The results confirmed deCODE's previously published SNP (Nature Genetics Jan 2006) that was also based on analysis of microsatellites and selection of fine mapping SNPs using the Disease Miner.



Key features of the Disease Miner

Clinical data - phenotypes: Import and query large datasets using state-of-the art query and reporting tool based on a relational Set Definition Language (SDL) that can also include advanced SQL statements.

Family data: Import relationship data, cluster subfamilies and draw pedigrees for large extended families.

Genetic analysis: Genome-wide association (GWA) based on Fishers-Exact, TDT, S-TDT, CNVs, and multipoint EM based ML haplotype analysis and QTL models.

Genome browser: High speed visualization and browsing with multiple statistical analysis results, patient CNVs and genotypes and genomic features in one view.

Marker and genes: A database with close to 30,000 known genes and 9 million SNPs and public annotations that are easily searched with the SDL query tool. Proprietary annotations can be added easily.

Raw data viewer: SNP clusters are easily inspected for sets of cases and controls and outliers are easily identified in relation to CNV analysis.

