FOR MORE INFORMATION:

ROSETTA BIOSOFTWARE
401 TERRY AVE N
SEATTLE, WA 98109 USA

TEL +1.206.926.1200
WWW.ROSETTABIO.COM
INFO@ROSETTABIO.COM

ROSETTA BIOSOFTWARE

Copyright © 2006, Rosetta Inpharmatics LLC. All rights reserved. The Rosetta Biosoftware icon and Syllego are trademarks of Rosetta Inpharmatics LLC.

Haploview was developed in and is maintained by Mark Daly's lab http://www.broad.mit.edu/personal/mjdaly/ at the Whitehead Institute http://www.wi.mit.edu by Jeffrey Barrett and Julian Maller. The current R is the result of a collaborative effort with contributions from all over the world. R was initially written by Robert Gentleman and Ross Ihaka—also known as "R & R" of the Statistics Department of the University of Auckland. R graphics are courtesy of the R statistical software, copyright R Foundation, from http://www.r-project.org.

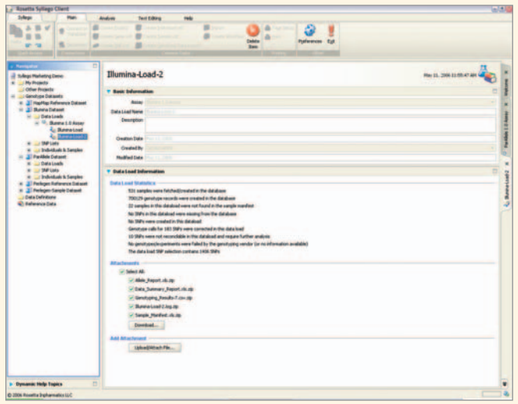
ROSETTA SYLLEGO™

Genetic Data Management and Analysis System

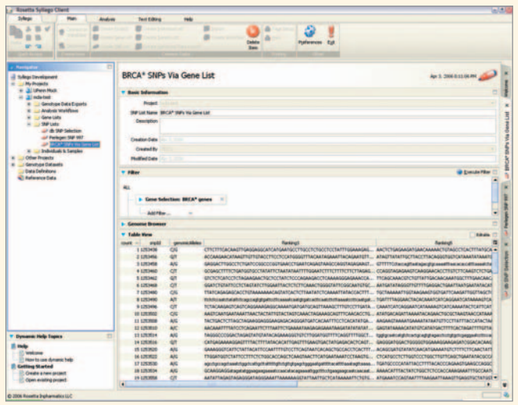
To meet the needs of scientists engaged in genetic studies, Rosetta Biosoftware built the Rosetta Syllego™ system. The system is designed to effectively manage genetic data from disparate sources and combine relevant genotypic and phenotypic study data so that scientists can spend time analyzing data and drawing conclusions instead of contending with data from multiple sources.

The number of published genome-wide association studies grew approximately 40% annually over the past five years. This increase is driven in part by growing acceptance of high throughput genotyping assay technologies such as the Illumina® beadchip and the Affymetrix® 500K GeneChip® mapping array and increased access to public, genetic information resources such as the International HapMap. Decreasing genotyping assay technology costs and a demand to unite genotypic and phenotypic assay data generated on a genome-wide scale in large study populations have also contributed to this increase. Consequently, genetics researchers across academia and industry now more than ever have the opportunity to conduct meaningful genetic studies that identify genetic variants and determine their impact on disease predisposition or drug response. Though researchers now have more sophisticated genetic analysis methods at their disposal, these researchers need a solution for genetic data management and analysis.

To address this need, Rosetta Biosoftware developed the Rosetta Syllego™ system.



Use the Syllego system with collaborators. Merge and combine data and publish analysis-ready data sets as soon as genotyping is complete or when all data are collected.



Identify SNP markers of interest.

The system is designed to overcome challenges presented by genetic study planning and design, quality control, data integration and analysis, and data accumulation and meta-analysis. Designed for scientists engaged in genetic studies, such as eQTL or genome-wide association or linkage, the system is a resource for managing genetic variation information. The system aggregates and manages genetic variation information from various sources so that researchers can better plan and execute genetic studies across a variety of commercial platforms to optimize results. QC tools help researchers verify that data are high-quality to prevent loss of time in downstream analyses. Application tools enable researchers to combine relevant genotypic and phenotypic study data to interface with the statistical environment of their choice. And automation tools enable researchers to analyze and catalog study data and analysis results, for incorporation into future studies and for collaboration purposes.

CHALLENGES INHERENT TO GENETIC STUDIES

Researchers engaged in genetic studies face a series of challenges. These challenges include how to:

- Plan and design a study that draws upon data from multiple technologies
- Ensure quality control to prevent loss of time in downstream analyses
- Integrate relevant study data and analyze them using the most suitable statistical tools
- Amass these data and catalog results for incorporating into future studies, for collaborating, or for publishing

To address these challenges, the Rosetta Syllego system includes the following capabilities:

STUDY PLANNING AND MANAGEMENT

- Support for multiple genotyping assay technologies available from Illumina and Affymetrix
- Project organization of study data, assay data and analysis results
- Selection of genes, markers (SNPs), and individuals/samples
- Ability to interface with genotyping service providers such as Illumina, Perlegen, and Parallele (Affymetrix)

QUALITY CONTROL

- Semi-automated execution of data quality control analyses
- Data quality control reporting and activity logging

DATA ANALYSIS AND INTEGRATION

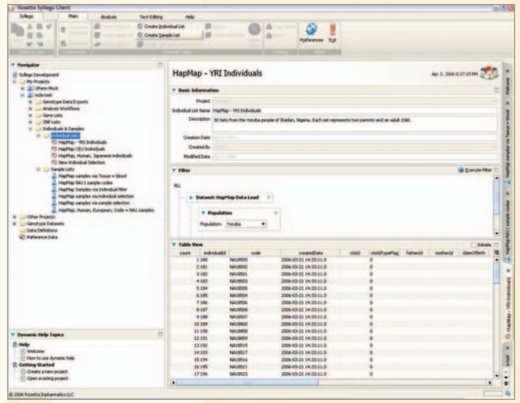
- Project and analysis templates for shared analysis methodology
- Common data repository with data access controls
- Integration with common statistics tools and genetics packages, such as R.
- Tools to link, script, and schedule analysis jobs
- Integration with the Rosetta Resolver® and Rosetta Elucidator® products

DATA ACCUMULATION AND META-ANALYSIS

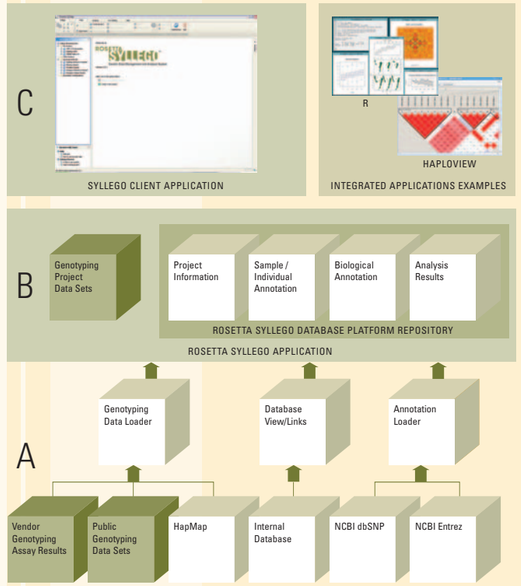
- Automated data loading and merging of public data sets, vendor data sets, and in-house assay results
- Support for multiple genome assemblies
- Reporting and publishing tools for study outcome
- Easy access to complete sets of interest within a common data repository as well as by project data

ROSETTA SYLLEGO SYSTEM ARCHITECTURE

The Syllego system facilitates the loading and acquisition of genotyping data sets from vendors as well as from public sources, such as the International HapMap Project. These data are annotated using information from public sources, such as NCBI EntrezGene or dbSNP (A). Data sets that have undergone quality control are published into the Syllego system's high-performance computing database (B). Users can access data and results by using the streamlined user interface or by using command-line tools. Finally, the system allows users to visualize and analyze data in tools of their choice, such as Haploview available from the Broad Institute, or R for statistical routines (C).



Manage multiple samples generated from study population, assist with study planning, and select specific samples for targeted downstream analysis.



Rosetta Syllego system architecture.