RFA-HL-07-010: Methods of Analysis of Gene-Environment Interactions in Complex Diseases: The Genes and Environment Initiative (R01)

Resources

Last updated: May 11, 2010

Objectives of the GxE Interaction Analysis Program

- Develop and test designs and analytical strategies for identifying GxE interactions in GWA, sequencing, linkage, or candidate gene studies in complex diseases
- Develop and validate algorithms and new computational approaches to identify individuals at high risk for developing disease based on both exposure patterns and genetic risk profiles

The following is a list of statistical software packages employed by the investigators to perform these analyses.

Investigator: Chris Amos

Project: Statistical Methods for Gene Environment Interactions in Lung Cancer

SimuPOP [simupop.sourceforge.net]

simuPOP is a general-purpose individual-based forward-time population genetics simulation environment. simuPOP provides an increasing number of built-in scripts that perform simulations ranging from implementation of basic population genetics models to generating datasets under complex evolutionary scenarios.

Investigator: Nilanjan Chatterjee

Project: Integrative Analysis of Genetic and Environmental Data in Epidemiologic Studies

CGEN-R Software [dceg.cancer.gov]

CGEN (Case-control.Genetics) is an R package for analyzing genetic data on case-control samples, with particular emphasis on novel methods for detecting Gene-Gene and Gene-Environment interactions.

Investigator: David Herrington

Project: Machine Learning to Identify Complex Interactions in Genome-wide Association Data

MECPM-SNP Software [cbil.ece.vt.edu]
Supplementary information for maximum entropy probability models of characterizing multilocus genomic interactions.

SNP Simulation [cbil.ece.vt.edu]
Supplementary information on SNP Simulation.

[Interacting SNP Detection](cbil.ece.vt.edu) [cbil.ece.vt.edu]

Reimplemented software of existing methods.

Investigator: Ed Iversen

Project: Bayesian Modeling and Optimal Design for Studies of Gene-Environment Association

[Multilevel Inference for SNP Association Studies Package (MISA)](stat.duke.edu) [stat.duke.edu]

The functions in this package focus on intermediate throughput case-control association studies, where the outcome of interest is often a binary disease state and where the genetic markers have been chosen to capture variation in a set of related genes, such as those involved in a specific biochemical pathway. The package incorporates functions that summarize the output of the model search algorithm, including image plots of the models with the highest posterior probability, marginal summaries of SNP and gene inclusion probabilities and Bayes Factors, and global summaries of the posterior probability and Bayes Factor giving evidence of an association in the set of SNPs of interest.

[Genetic Simulations for Case-Control or Survival Analyses (SimGbyE)](stat.duke.edu) [stat.duke.edu]

The functions in this package create simulated case/control or survival data sets with one or more of the following assumed effects: genetic main effects (G), environmental main effects (E), Gene by Gene interactions (GbyG), Gene by environment interactions (GbyE).

Investigator: Kung-Yee Liang

Project: Novel Statistical Methods for Gene-Environment Interactions in Complex Diseases

[Trio: Processing and Simulating Genotype Data for Trio Logic Regression](cran.r-project.org) [cran.r-project.org]

This package can be used to set up matched case-pseudo controls genotypes data for trios in order to run trio logic regression, to impute missing genotypes in trios, or to simulate case-parent trios with disease risk dependent on SNP-SNP interaction. This package furthermore contains functions for computing the values of pairwise LD measures and for identifying LD blocks.