

# “A New Method for Genetic Network Reconstruction in Expression QTL Data Sets”

**Dr. Christine Duarte**  
Research Scientist  
Nature Source Genetics, Ithaca, NY

Expression QTL (eQTL) studies involve the collection of microarray gene expression data and genetic marker data from segregating individuals in a population in order to search for genetic determinants of differential gene expression. Previous studies have found large numbers of trans-regulated genes that link to a single locus or eQTL “hotspot”. It would be of great interest to discover the mechanism of co-regulation for these groups of genes. However, many difficulties exist with current network reconstruction algorithms such as low power and high computational cost. It is commonly observed that biological networks have a scale-free or power-law architecture in which there exist a few highly influential nodes that have many connections to other nodes. If we assume that this type of architecture applies to genetic networks, then we can simplify the problem of genetic network reconstruction by focusing on the discovery of the key regulatory genes at the top of the network.

In our new method we introduce the concept of “shielding” in which a gene is conditionally independent of the QTL given the shielder gene, and we iteratively build networks from the QTL down using tests of conditional independence. We evaluate the confidence level of shielders using a two-part strategy of requiring a threshold number of genes to be shielded and requiring a high level of bootstrap support for shielders. We have performed a set of simulations to test the sensitivity and specificity of our method as a function of method parameters. We have found that even with a small sample size (100) and a large number of network genes (as many as 600), our algorithm succeeds in finding a large number of key network regulators (47% on average) with high confidence (95% specificity on average). We have applied our network reconstruction algorithm to a yeast expression QTL data set in which microarray and marker data were collected from the progeny of a backcross of two species of *Saccharomyces cerevisiae*. Networks have been reconstructed for 11 of the largest eQTL hotspots in this data set. The regulation of shielder gene expression has been found to be primarily in trans, although about 10% of shielder genes are found to be regulated in cis. Bioinformatic analysis of three networks generated different hypotheses for mechanisms of regulation of the shielded genes by the primary shielders. One common theme was that the shielders modulated the effect of transcription factors of which they were themselves targets. Overall our method has created a large list of potentially important regulatory genes in various yeast biological processes, and further bioinformatic analysis or laboratory experiments could lead to the generation and testing of many important hypotheses.