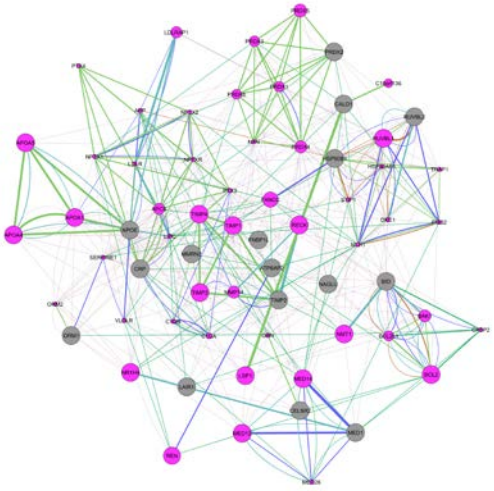
Protein Interactions and Network/Pathway Analysis

Network and pathways analysis was performed using the gene network tool GeneMANIA 2.4 Cytoscape 2.8.2 plugin, which contains association networks from multiple sources. We used 144 human interaction networks based on data from BIND, IntAct, INTERPRO, Pfam and other interaction and pathway databases, using association data from protein and genetic interactions, known and predicted pathways, coexpression, colocalization and protein domain. For our two interaction networks, our analysis was performed by, for the 16 putative biomarker network, query gene based weighting, in which the query genes interact as much as possible with each other; and, for the extended network, by Gene Ontology biological process and cellular component co-annotation patterns, where network weights were assigned based on the reproducibility of GO co-annotation patterns either for molecular function or cellular component [16].

Network and pathway analysis

To determine whether there were any functional associations between our putative biomarkers, a network and pathway analysis was performed using GeneMANIA plugin on Cytoscape 2.8.2. Protein-protein interactions with in 16 putative biomarkers were predicted by GeneMANIA from sixty one physical interaction data, 3 genetic interaction data, 6 pathway data, and 50 predicted interaction data. From our analysis, of our 16 putative biomarkers, only 4 (mediator of RNA polymerase II transcription subunit 14, orosomucoid, multimerin-2, caldesmon) did not participate in the interaction network (Figure 2).In order to find other proteins that were functionally associated with our 16 putative biomarkers, we constructed a network of these 16 proteins with the top 50 functionally associated proteins as determined by GeneMANIA's adaptive weighting with GO molecular function and cellular component. This analysis generated a network in which all proteins of our 16 putative biomarkers participated. The network analysis thus produces a single network involving all of our putative biomarkers, suggesting that our quantitative protein data may implicate a related set of pathways perturbed by HIV infection.

A.



B.

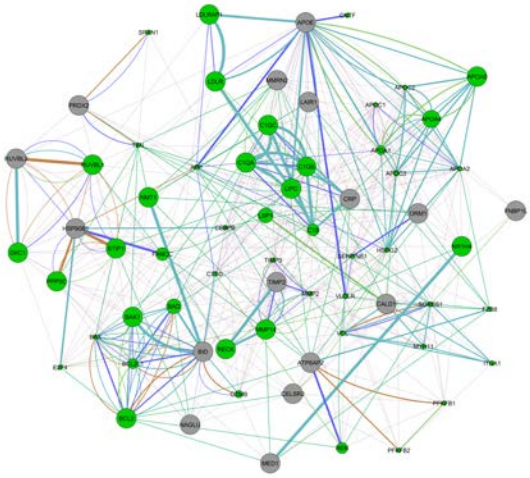


Figure 1. Functional association of 16 putative biomarker proteins with closely associated 50 genes. By using a set of input genes (gray), protein association network is created with a very large set of Gene Ontology A) molecular function and B) cellular components association data. In A) the association data include protein and genetic interactions (green), physical interaction (blue), pathways (red), co-expression (brown), co-localization (purple) and predicted protein domain (orange). Putative biomarker proteins are gray nodes and magenta nodes are top 50 related interacting genes.

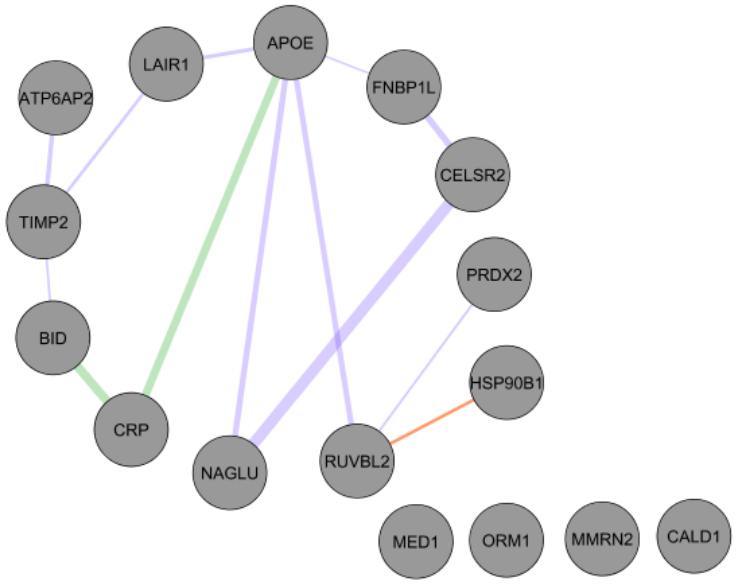


Figure 2. Functional association of 16 putative biomarker proteins. Illustrating the interactive functional relationship among genes based on co-expression (purple), co-localization (green) and predicted interactions (orange) association data. Gene APOE has the highest number of interactions with other genes based on co-expression and co-localization. The thickness of edge line indicates weight of the predictive value. Four genes have no interaction within the group.