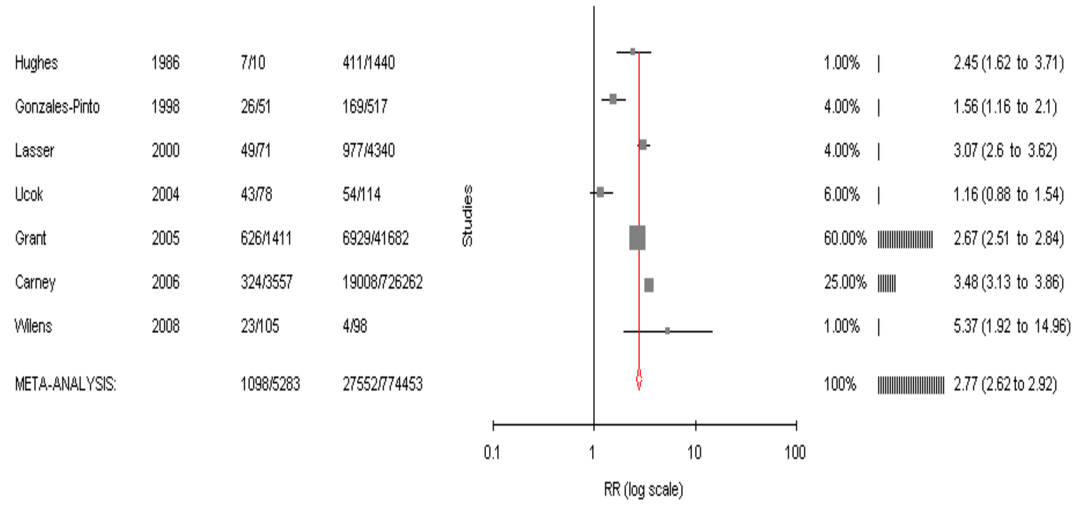
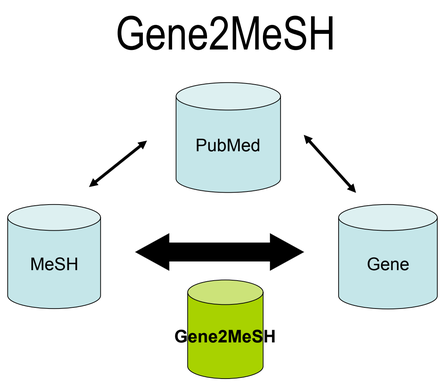
# Modeling Complex Genetic and Environmental Influences on Comorbid Bipolar Disorder with Tobacco Use Disorder

*McEachin RC[1], Saccone NL[2], Saccone, SF[2], Bierut LJ[2], Cavalcoli JC[1], McInnis MG[1]*
*[1]University of Michigan, National Center for Integrative Biomedical Informatics, [2] Washington University Saint Louis*

**Hypothesis:** Tobacco Use Disorder (TUD) disproportionately affects psychiatric patients in general and, for patients with Bipolar Disorder (BD), the risk for TUD is almost 3 times that for the general population based on our meta-analysis consistent with some common underlying etiology. Both BD and TUD show strong evidence of genetic influences on susceptibility. Given evidence of common etiology and evidence of genetic influences, we hypothesized common *genetic* influences on the etiology of BD and TUD.



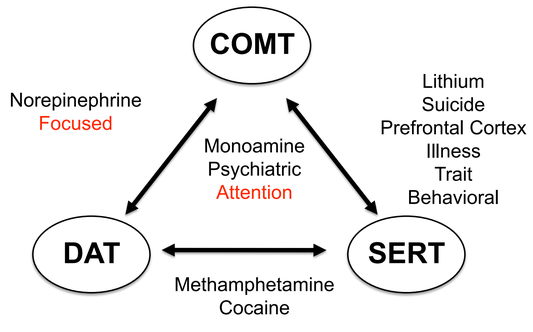

**Risk for Tobacco Use Disorder (TUD) among Bipolar Disorder (BD) patients is almost three times that of the general population.** Using the MIX program, we conducted a meta-analysis of the seven studies that have been published on risk for comorbid BD with TUD. Relative risk is estimated at 2.77, significant at p-value < 0.01, and 95% confidence interval 2.62 to 2.92, with smoking modeled as a fixed effect and Mantel-Haenszel weighting.




**Gene2MeSH provides a resource for candidate gene selection based on MeSH annotation of publications.** We queried Gene2MeSH to identify candidate genes for BD and TUD, evaluated the published evidence for the set of genes that Gene2MeSH nominated for each disorder, then selected the subset of three genes that show the strongest evidence of involvement in both BD and TUD: Catechol-O-Methyltransferase (COMT), Dopamine Transporter (SLC6A3), and Serotonin Transporter (SLC6A4)

**PDG-ACE identifies common elements of Entrez Gene text describing genes at pairs of genetic loci to help us understand how candidate genes work in complex disease.**

We used PDG-ACE to identify statistically significant commonality among our candidate genes. Results included the expected psychiatric and substance abuse related keywords, along with attention deficit. PDG-ACE also allowed us to identify significant gender specific effects of all three of these genes in both psychiatric and substance use disorders.

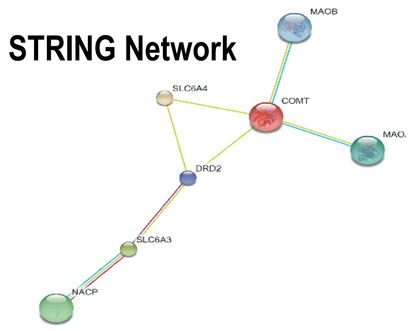



A conceptually related program, GRAIL, available at the Broad Institute, mines full text publications, rather than Entrez Gene records. In this analysis, GRAIL yields results consistent with PDG-ACE.

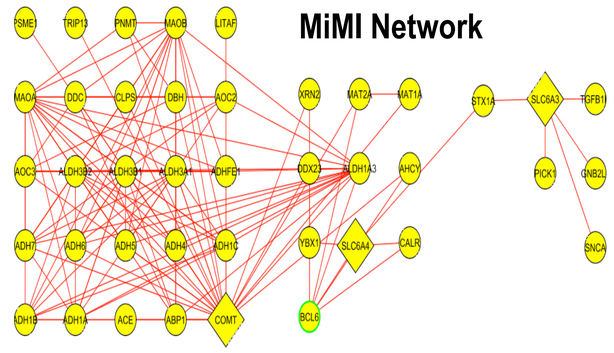
**Network Generation:** Given evidence that COMT, DAT and SERT play roles in comorbid BD with TUD, we next sought to understand how these genes work together in predisposing the comorbidity. We used MiMI (available at NCIBI) STRING (available at EMBL) and MetaCore (available at GeneGo Inc.) to create gene networks anchored by our validated candidate genes. MiMI focuses on direct protein-protein interactions, while STRING and MetaCore incorporate functional and text mining interactions into the networks

**Networks:** In each case, we set parameters to build the smallest network possible, using only the strongest evidence available. The STRING network (left) only includes 7 genes, while the MiMI network (right) includes 38 genes, and the GeneGo network (below) includes 66 genes.

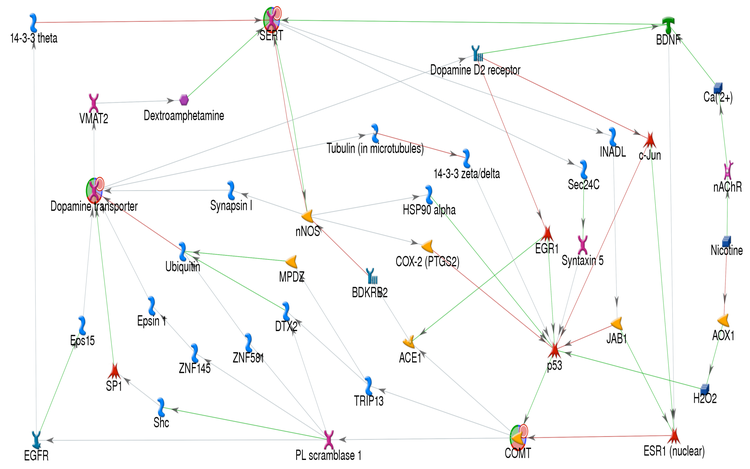
**STRING Network**




**MiMI Network**




**GeneGo Network**




Each network poses additional candidate genes, based on interactions with our validated candidate genes, as well as novel hypotheses on network disease association. For genes that influence disease susceptibility, any interacting genes may also influence disease susceptibility.

**Network hypothesis testing:** We tested each of these networks for association with BD and TUD using the Genetic Association Database (GAD), available through the DAVID interface at NIAID. We submitted the list of genes from each network and DAVID returned a table of GAD terms with the genes tagged for each annotation, p-value and FDR for over-representation of genes tagged with each term, and fold enrichment. For each network, we did the hypothesis testing both including the validated candidates and excluding them. Only the GeneGo network was significantly enriched (FDR < 5%) for both BD and TUD when we excluded the validated candidates.

| Term | Count | % | PValue | Genes | Fold Enrichment | FDR % |
|---|---|---|---|---|---|---|
| cognitive function | 14 | 20.59% | 2.77E-18 | 1137, 6531, 6532, 1138, 1312, 1813, 57053, 1135, 2099, 627, 55584, 1139, 1141, 1142, | 36.0 | 0.000 |
| bipolar disorder | 19 | 27.94% | 7.18E-15 | 1137, 6531, 1140, 1143, 8973, 6532, 1138, 1312, 1813, 1136, 57053, 1135, 1636, 627, 55584, 1134, 1139, 1141, 1142, | 10.2 | 0.000 |
| smoking behavior | 10 | 14.71% | 1.92E-10 | 6531, 627, 55584, 6532, 4842, 1139, 1312, 1141, 1813, 1142, | 21.9 | 0.000 |

**Candidate SNPs:** Based on the statistically significant association of the GeneGo network with both BD and TUD, we hypothesize that any of the genes in this network could impact susceptibility for the comorbidity. To date, no Genome Wide Association (GWA) studies have been conducted for comorbid BD with TUD. However, we used three lines of evidence to prioritize SNPs for follow on study. We combined evidence from the NicSNP GWA study for nicotine addiction, the GAIN GWA study for BD, and functional data based on the GIN algorithm. We weighted each SNP by the sum of $-\log_{10}$(p-value) in each of the GWA studies, plus weights for % evolutionary conservation with mouse and SNP function (i.e. non-synonymous substitution, synonymous substitution, intron, and gene locus). These prioritized SNPs are available for follow-on testing.

**Summary and Conclusions:** Based on our meta-analysis, we find strong evidence of bi-directional relative risk for BD and TUD, suggesting some common etiology for these disorders. Both disorders show evidence of genetic influences on susceptibility, consistent with a common *genetic* etiology. Gene2MeSH nominates COMT, DAT, and SERT as candidate genes for the comorbidity and we find strong evidence to support their roles. PDG-ACE identifies significant commonality among these genes consistent with psychiatric disorders, substance use disorders, attention deficit disorder, and gender specific effects in these disorders. No gene functions alone, so we used three algorithms to hypothesize gene networks consistent with roles that our candidate genes might play in the comorbidity. Based on hypothesis testing via the GAD, we find that the GeneGo network is significantly associated with both disorders, even when the established candidate genes are not included. All of the genes in this network are candidates for influencing the comorbidity and we have produced a set of prioritized SNPs for follow-on work.

**Acknowledgements:** This work was supported by NIH Grant #U54 DA021519