

# **Automated Selection of Genes for Translational Research on Comorbidity of Bipolar Disorder with Substance Abuse**



Raphael D. Isokpehi<sup>1</sup>, Sharon A. Lewis<sup>2</sup>, Tolulola O. Oyeleye<sup>2</sup>, Wellington K. Ayensu<sup>2</sup> and Tonya M. Gerald<sup>3</sup>

2008-2009 Pilot Project Award

<sup>1</sup>Center for Bioinformatics & Computational Biology, Jackson State University, Jackson, MS 39217 <sup>2</sup>Department of Chemistry, Langston University, OK <sup>3</sup>Department of Chemistry, North Carolina Central University, NC Corresponding Author: raphael.isokpehi@jsums.edu

#### INTRODUCTION:

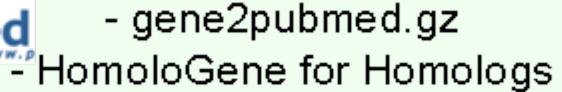
- Bipolar Disorder is a highly heritable mental illness. The global burden of bipolar disorder is complicated by its comorbidity with substance abuse.
- Several genome-wide linkage and association studies on Bipolar Disorder as well as Substance Abuse have focused on the identification and/or prioritization of candidate disease genes (18247375; 18694166; 18711365; 18628681).
- A useful step for translational research of these identified/prioritized genes is to identify sets of genes that have particular kinds of publicly available data.
- Therefore, we have leveraged the availability of links to related resources in the Entrez Gene database to develop a web-based resource for selecting genes based on presence or absence in particular biological data resources.

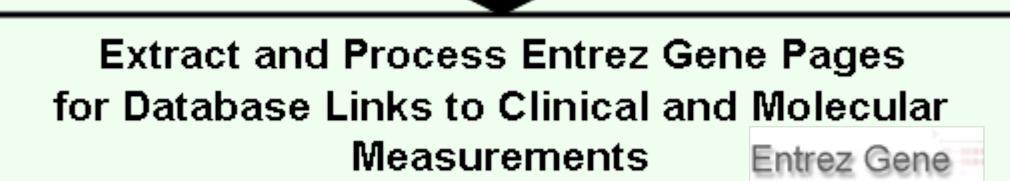
# **METHODS:**

#### Computational Pipeline

Compile Multi-Organism Gene Set on Bipolar Disorder and Substance Abuse - MeSH defined PubMed Citations

Publoged - gene2p





Develop a Binary Profile of Availability of Database Link for each Gene

Gene Selection Resource for Genes Linked to Bipolar Disorder and/or Substance Abuse (BPD)



Linking Query Outputs to the Michigan Molecular Interactions (MiMI) resource

#### **RESULTS:**

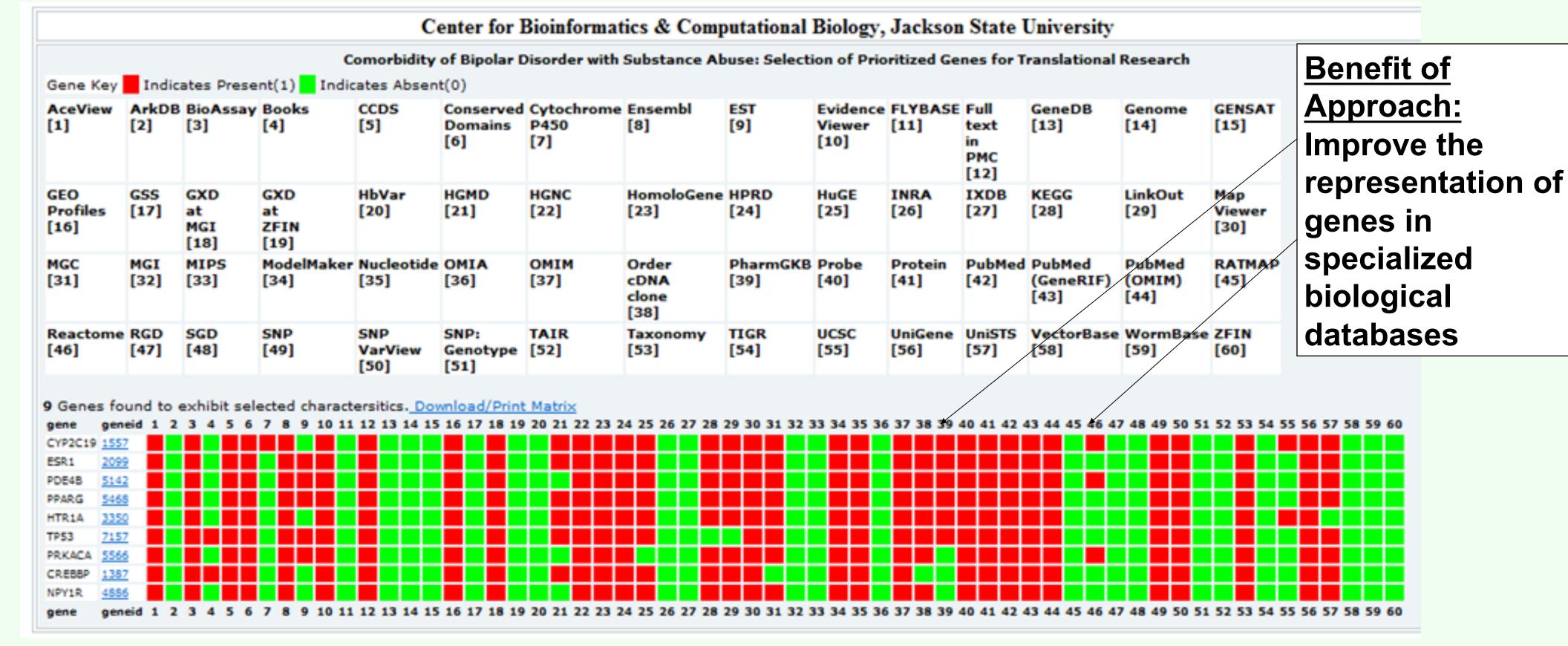
• The utility of our approach is demonstrated using a set of 3,399 genes from multiple eukaryotes that have been studied in the context of Bipolar Disorder and/or Substance Abuse. A web resource to automate the selection of genes that contain certain database links is available at http://compbio.jsums.edu/bpd

### Multi-organism Dataset

Organism	Gene
	Count
Anopheles gambiae str. PEST	124
Arabidopsis thaliana	46
Ashbya gossypii ATCC 10895	22
Bos taurus	326
Caenorhabditis elegans	101
Canis lupus familiaris	330
Danio rerio	321
Drosophila melanogaster	142
Gallus gallus	290
Homo sapiens	388
Kluyveromyces lactisNRRL Y-1140	30
Macaca mulatta	6
Magnaporthe grisea70-15	36
Mus musculus	433
Neurospora crassa	36
Oryza sativa Japonica Group	44
Pan troglodytes	294
Plasmodium falciparum3D7	10
Rattus norvegicus	352
Saccharomyces cerevisiae	31
Schizosaccharomyces pombe	37

## Database Selection (60)

Comorbidity of Bipolar Disorder with Substance Abuse: Selection of Prioritized Genes for Translational Research									
1. AceView	2. ArkDB	3. BioAssay	4. Books	5. CCDS	6. Conserved Domains	7. Cytochrome P450 Homepage	8. Ensembl	9. EST	10. Evidence Viewer
Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼
11. FLYBASE	12. Full text in PMC	13. GeneDB	14. Gene Expression Database at ZFIN	15. Gene Expression Database (GXD) at MGI	16. Genome	17. GENSAT	18. GEO Profiles	19. GSS	20. HbVar: A Databas of Human Hemoglobin Variants and Thalassemias
Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼
21. HGMD	22. HGNC	23. HomoloGene	24. HPRD	25. HuGE Navigator	26. INRA	27. Integrated X Chromosome Database (IXDB)	28. KEGG	29. LinkOut	30. Map Viewer
Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼
31. MGC	32. MGI	33. MIPS	34. ModelMaker	35. Nucleotide	36. OMIA	37. OMIM	38. Order cDNA clone	39. PharmGKB	40. Probe
Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼
41. Protein	42. PubMed	43. PubMed (GeneRIF)	44. PubMed (OMIM)	45. RATMAP	46. Reactome	47. RGD	48. SGD	49. SNP	50. SNP: Genotype
Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼
51. SNP /arView	52. TAIR	53. Taxonomy	, 54. TIGR	55. UCSC	56. UniGene	57. UniSTS	58. VectorBase	59. WormBase	60. ZFIN
Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼	Ignore ▼



# Top 20 databases with "Links" for Bipolar Disorder and Substance Abuse Gene Set

Database	Gene	Digit in
	count	Signature*
Map Viewer	3362	30
Taxonomy	3352	53
HomoloGene	3307	23
Nucleotide	3285	35
Protein	3279	41
Genome	3203	16
Conserved Domains	3063	6
LinkOut	2808	29
KEGG	2780	28
Evidence Viewer	2664	10
ModelMaker	2664	34
UniGene	2503	56
PubMed	2349	42
GEO Profiles	2064	18
SNP	1765	49
Full text in PMC	1729	12
Ensembl	1579	8
SNP: Genotype	1411	50
Probe	1400	40
UniSTS	1325	57

#### **FUTURE WORK:**

A future development goal of the resource is to facilitate systems biology analyses by linking query outputs to the Michigan Molecular Interactions (MiMI) resource.

