Combined methods to explore genetic etiology of related complex diseases

Shefali S Verma[1], Anurag Verma[1], Anastasia Lucas[1], Jim Linneman[2], Peggy Peissig[2], Murray Brilliant[2], Catherine A. McCarty[3], Jonathan L. Haines[4], Tamara R. Vrabec[5], Gerard Tromp[5], Jyoti Pathak[6], Christopher Chute[6], David R. Crosslin[7,8], Gail P. Jarvik[7,8], Eric B. Larson[9],M. Geoffrey Hayes[9], Marylin D. Ritchie[1]

1. Center for Systems Genomics, Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA
2. Marshfield Clinic, Marshfield, WI
3. Essentia Rural Health, Duluth, MN
4. Case Western University, Cleveland, OH
5. Geisinger Health System, Danville, PA
6. Mayo Clinic, Rochester, MN
7. Department of Medicine, Division of Medical Genetics, University of Washington, Seattle, WA
8. Group Health Research Institute, Seattle, WA
9. Northwestern University, Chicago, IL

Abstract

Genome-wide association studies (GWAS) have identified several SNPs associated with either glaucoma or ocular hypertension (OHT). However, these susceptibility loci explain a small fraction of the genetic risk. Gene-gene interaction (GxG) studies are considered a potential avenue to identify this missing heritability. Using a dataset from the eMERGE (electronic Medical Records and Genomics) Network, which included GWAS data imputed using the 1000 Genomes, we were able to explore the genetic etiology of two very related common eye- diseases: glaucoma and OHT. OHT is one of the leading risk factors for glaucoma, thus we explored the relationships between these two traits at the molecular level. A total of 5,032 (glaucoma) and 3,154 (OHT) unrelated samples of ages 40-90 were extracted from the eMERGE study bio repositories. First, we performed GWAS and GxG studies for each trait using the imputed dataset and identified several main effects and GxG models that meet bonferroni significance. Secondly, from the obtained GWAS with main effect p< 0.01, we also performed a pathway-enrichment analysis using KEGG database on both of these traits combined.

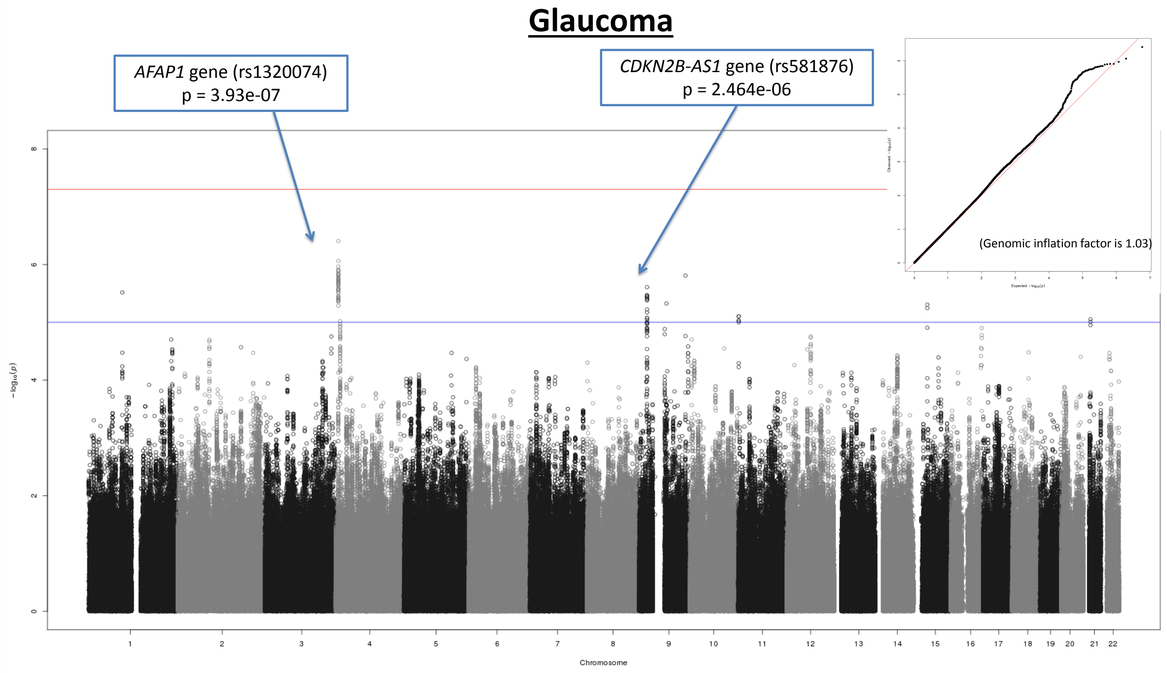
Dataset and Quality Control

Glaucoma and ocular hypertension samples were extracted from eMERGE imputed dataset that includes 55,292 samples in total. After extracting the samples, we performed following quality control on the data:

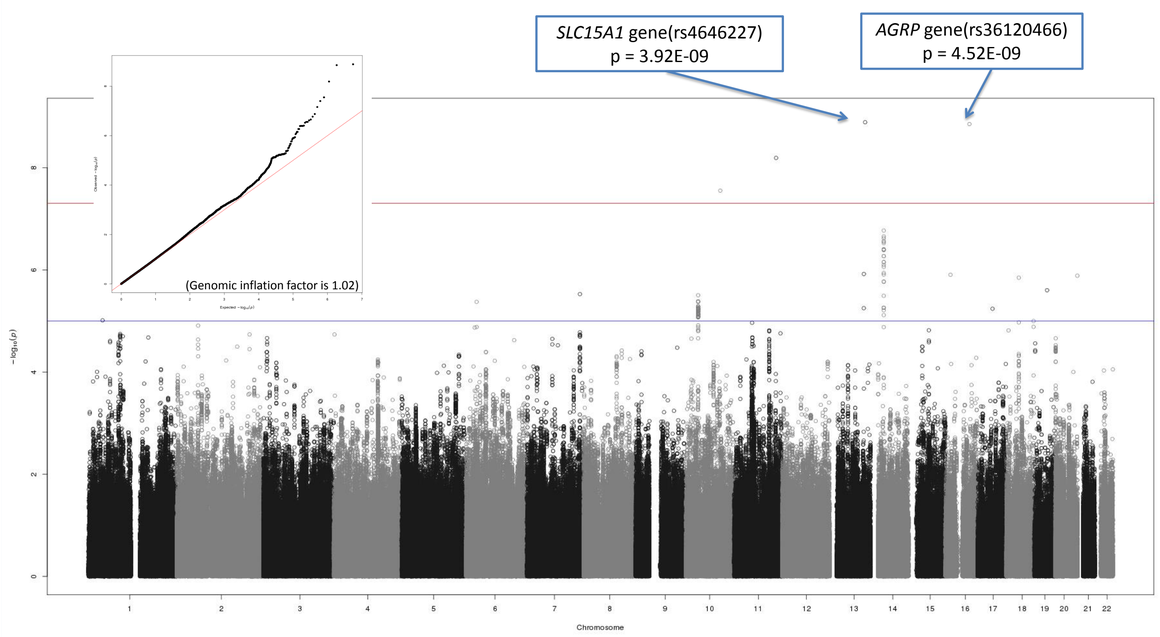
- Info score ≥ 0.7
- Marker call rate ≥ 99%
- Sample call rate ≥ 99%
- MAF ≥ 0.05
- Relateds removed (kinship ≥ 0.125)

	#Cases	#Controls	#Males	#Females	Total Samples	Total Markers
Glaucoma	962	4068	2229	2803	5032	2,951,107
Ocular Hypertension	868	2286	1369	1785	3154	2,785,382

Genome Wide Association Study



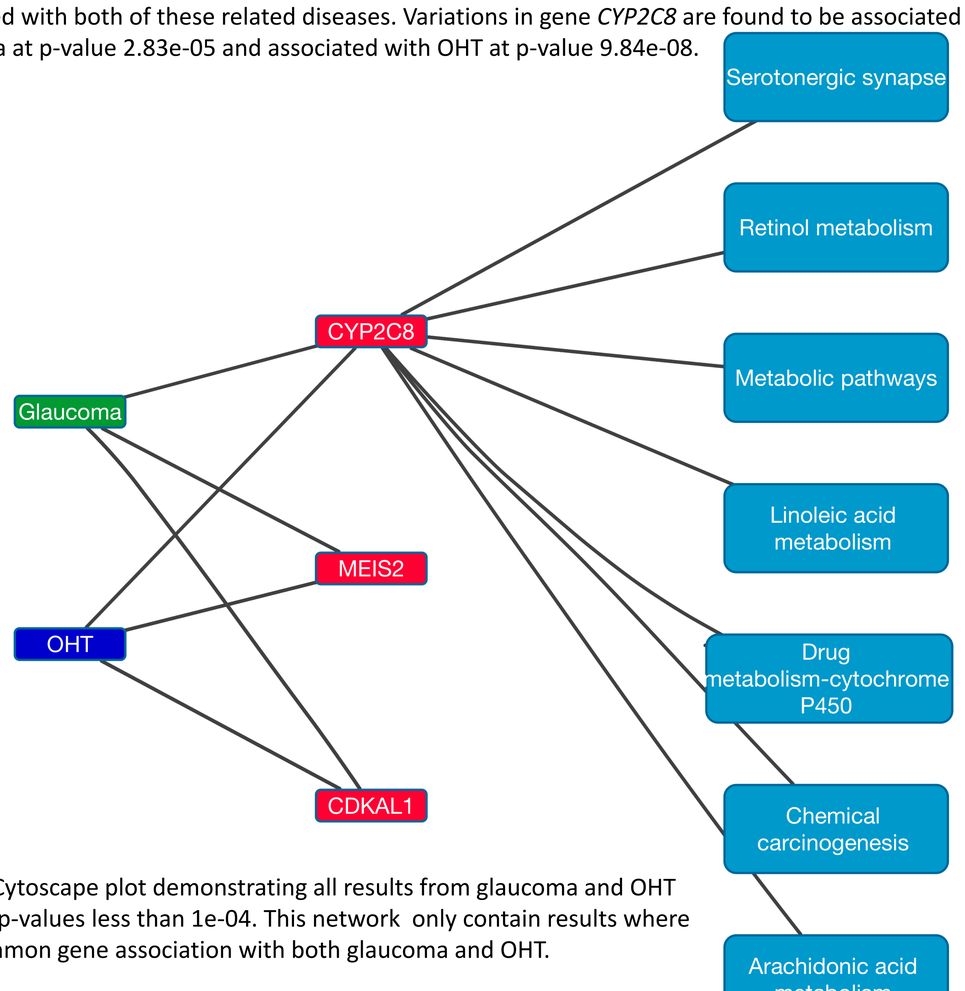
^Figure 1: Manhattan and QQ plots of genome wide association study after adjusting the additive model by age, sex, eMERGE site, genotyping platform and first 3 principal components.
Manhattan Plot: Each SNP is represented as a point on the plot and chromosomes are colored as black and grey alternatively as labeled on the X-axis. The –log10(p-value) is shown on Y-axis. Two loci are denoted as our most significant associations with glaucoma.



^Figure 2: Manhattan and QQ plots of genome wide association study after adjusting the additive model by age, sex, eMERGE site, genotyping platform, and first 3 principal component.
Manhattan Plot: Each SNP is represented as a point on the plot and chromosomes are colored as black and grey alternatively as labeled on the X-axis. The –log10(P-value) is shown on Y-axis. Two loci are denoted as our most significant associations with intra-ocular pressure.

Pathway Analysis

From both glaucoma and OHT main effect GWAS results, all SNPs at p-value<1e-04 were selected and mapped to genes and then the genes were mapped to pathways. Pathway enrichment analysis done by merging results for both glaucoma and OHT resulted in common genes and pathways that are found to be associated with both of these related diseases. Variations in gene *CYP2C8* are found to be associated with glaucoma at p-value 2.83e-05 and associated with OHT at p-value 9.84e-08.



^Figure 3: Cytoscape plot demonstrating all results from glaucoma and OHT GWAS with p-values less than 1e-04. This network only contain results where there is common gene association with both glaucoma and OHT.

Gene-Gene Interaction Analysis

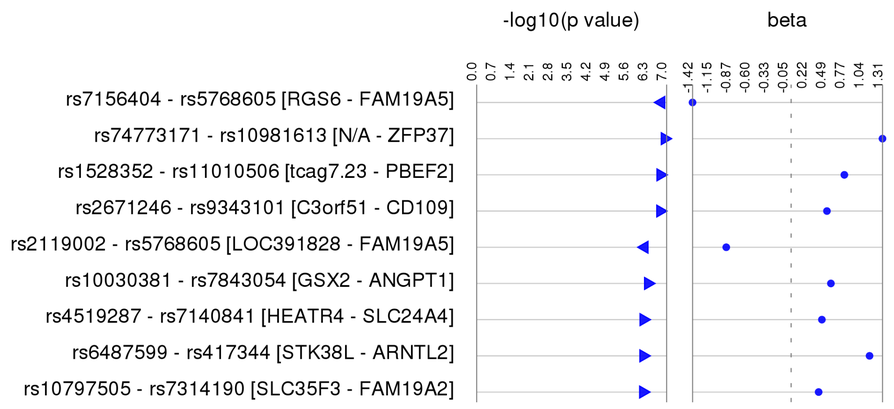
SNPs from main effect analysis with p-values <0.001 were selected and then LD pruned to r^2 0.6 to perform regression using PLATO software to look for interactions exhaustively. Likelihood ratio test p-value represent interaction effect above and beyond the main effect for each SNP.

#SNPs tested in glaucoma: 4405
#SNPs tested in OHT: 4441

Figures 4 and 5 represent gene-gene interaction results for glaucoma and ocular hypertension respectively.

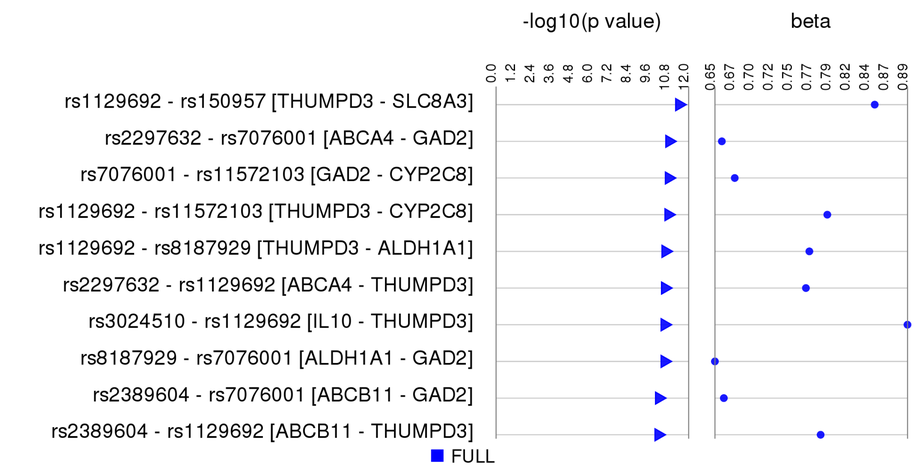
>Figure 4: Synthesis view plot representing top 10 GXG results for glaucoma analysis.

Models are listed on the left and two tracks shown are for –log10(p-value) and direction of effects



>Figure 5: Synthesis view plot representing top 10 GXG results for Ocular hypertension analysis.

Models are listed on the left and two tracks shown are for –log10(p-value) and direction of effects



ACKNOWLEDGEMENTS

The eMERGE Network was initiated and funded by NHGRI through the following grants: U01HG006828 (Cincinnati Children's Hospital Medical Center/Harvard); U01HG006830 (Children's Hospital of Philadelphia); U01HG006389 (Essentia Institute of Rural Health); U01HG006382 (Geisinger Clinic); U01HG006375 (Group Health Cooperative); U01HG06379 (Mayo Clinic).); U01HG006380 (Mount Sinai School of Medicine); U01HG006388 (Northwestern University); U01HG006378 (Vanderbilt University); and U01HG006385 (Vanderbilt University serving as the Coordinating Center).