Introduction:

GWAS are used to link a specific genomic locus to a measurable phenotype.

Other phenotypic associations from existing GWAS can inform a connection between two phenotypes.

CSF functions in protection, transport, and removal of neuronal metabolic waste products.

CSF also contains circulating metabolites passing through the blood brain barrier.

CSF metabolites may inform brain function.

Our study uses the largest set of CSF data to date.

Abstract:

Genome-wide association studies (GWAS) are used to identify genetic loci that are associated with a trait of interest in order to decipher underlying biological mechanisms. Metabolite measures in cerebrospinal fluid (CSF), the fluid that surrounds the brain, provides insight into brain function that may be relevant for neurobehavioral traits and neuropsychiatric disorders. We performed a GWAS of 1,000 metabolites in a sample of 600 human subjects, the largest set of CSF data used in a GWAS to date. We applied standard quality control of genetic data including missingness, minor allele frequency cut-off, and population stratification. Phenotype data were checked for outliers, and non-normal data were transformed using inverse rank normalization. A linear association, including age and sex as additional covariates, was performed using the PLINK toolset. Significant SNP associations were found in 40 metabolites. Analysis via FUMA found that many of these SNPs are associated with other traits. Further steps include analysis of other CSF metabolites and the use of dosage genotype data for potentially more detailed results.

Methods:

Sample Acquisition

• European ancestry
• Exclude neuropsychiatric disorders
• Blood samples

Quality Control

• Cerebrospinal Fluid (CSF)
• Spinal tap during unrelated surgery
• Hil-Pos, C8-Pos, C18-Neg platforms

GWAS and analysis

• Match genotype samples to phenotype samples: 480 matches
• Plot: linear association
• Variables: Sex, Age, 3 Covariates (PCs)
• > 4 million variants
• Manhattan and QQ plots
• FUMA: downstream analyses
• GWAS Catalog, mapped genes

Results:

Discussion:

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Linked Traits (GWAS Catalog)</th>
<th>Mapped Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydroxyproline/proline: amino acid</td>
<td>Blood Metabolites, Serum Metabolites, Urine Metabolites</td>
<td>SLC6A20: kidney cross-membrane transport proteins</td>
</tr>
<tr>
<td>Carnitine: fatty acid transport in mitochondria, waste removal</td>
<td>SDMA: renal disease marker</td>
<td></td>
</tr>
<tr>
<td>N-acetylcarnitine: amino acid derivative, common metabolite</td>
<td>Cognitive Performance, Educational Attainment, Math Ability, Intelligence, Smoking Cessation</td>
<td></td>
</tr>
<tr>
<td>Phosphocholine: choline metabolic pathway, phospholipid synthesis</td>
<td>Liver enzymes, Folate pathway, Non-small cell lung cancer</td>
<td></td>
</tr>
<tr>
<td>Betaine: osmoregulation, liver detoxification reactions, metabolic syndromes</td>
<td>Bilirubin levels, LDL Cholesterol</td>
<td></td>
</tr>
<tr>
<td>N-carbamoyl-beta-alanine: Uracil metabolism derivative</td>
<td>Irritable Mood, Blood Protein levels</td>
<td></td>
</tr>
</tbody>
</table>

Conclusion:

A GWAS analysis can be successfully performed with CSF metabolite data, resulting in multiple significant SNP associations. Significant SNPs can be mapped to genes and compared to other GWAS trait results for analysis.

Some significant SNPs were linked to blood metabolite levels in other GWAS. These metabolites may be part of general processes occurring in all cells, or crossing the blood-brain barrier.

Links to brain-related traits may warrant further study.

Future Direction:

Increase statistical power by completely matching all samples.

Analyze all 10,000 metabolites to find additional links.

Identify potential loci to study functionally in wet lab.

Acknowledgements:

A sincere thank you to: The B.I.G. program, espically Lana Martin, for providing us this opportunity for research. The Ophoff lab, for their invaluable advice and support. The study participants.