Integration of methylation QTL and enhancer-target gene maps with schizophrenia GWAS summary results identifies novel genes

Chong Wu
Florida State University

IGES 2018

Oct. 15, 2018

Introduction

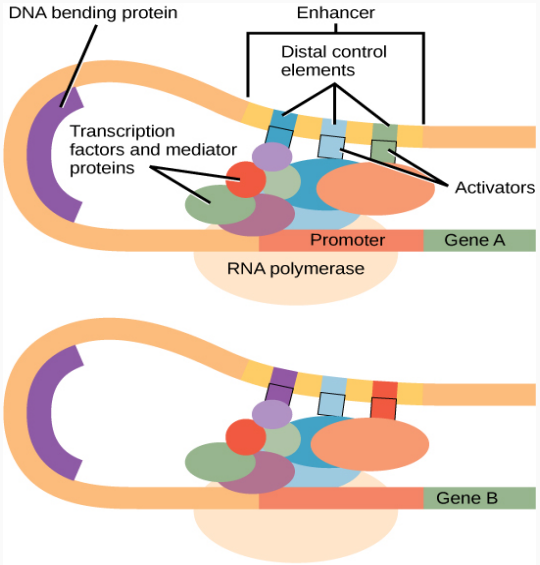
Two directions for gene-based analysis in GWAS

- Constructing a powerful test based on the GWAS data itself
  - Sum test; SSU test; adaptive test (aSPU)
- Integrating external information
  - PrediXcan/TWAS: integrating eQTL data sets with GWAS individual data or summary results
  - "E + G": integrating enhancer-promoter interactions

Goal

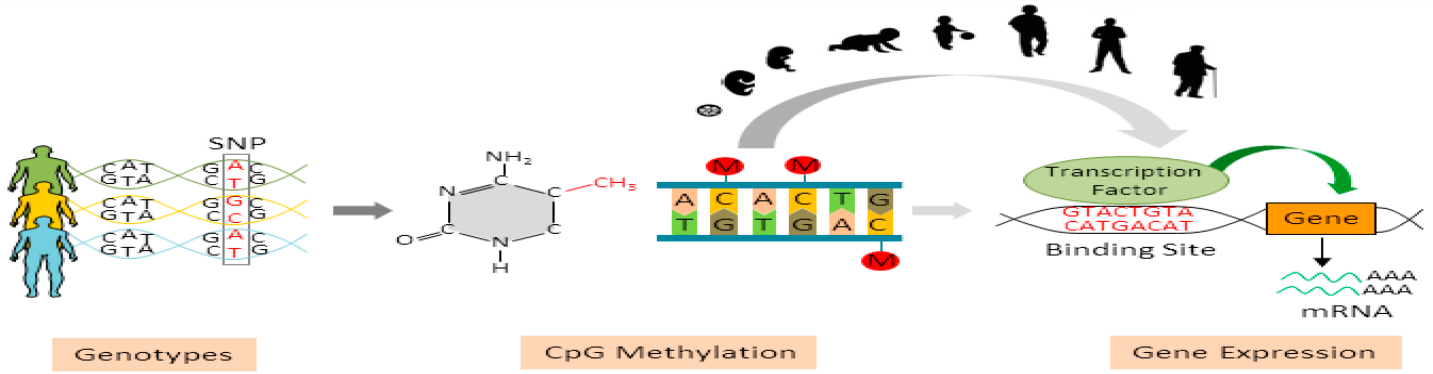
Develop a new gene based test by integrating external regulatory information to improve statistical power and enhance interpretability.

How does the enhancer-promoter interaction inform GWAS



- Enhancers: regions that help increase or enhance transcription
- May as far as 2 or 3 Mbp away from the gene
- GWAS risk loci are enriched in enhancers
- Recent biotechnological advances made enhancer-promoter interactions data available

How does the mQTL inform GWAS



- DNA methylation: epigentic; affect gene expression
- mQTL: locus associated with DNA methylation
- Genetic variation influences level of DNA methylation at regulatory regions and can module gene expression
- Some mQTL databases are publicly available

Method

New method: "E + G + Methyl"

- "E + G + Methyl": integrates enhancer-target gene maps, mQTL databases, and GWAS summary results to identify significant and novel genes
  - Use only mQTLs (and exclude other SNPs) located in enhancers, promoters, and coding (including introns) regions
  - Apply some well known gene-based tests, such as SPU(1) and SPU(2).

New method "E + G + Methyl"

- Suppose that $Z_j = \hat{\beta}_j / \mathrm{SE}_j$ is the Z-statistic for association between the GWAS trait and SNP j
- $\mathrm{SPU}(1) = \sum_{j=1}^{p} Z_j$;

  $\mathrm{SPU}(2) = \sum_{j=1}^{p} Z_j^2$
- We use a reference sample (e.g. the 1000 Genome Project samples) to estimate linkage disequilibrium (LD) among the SNPs and thus the correlation matrix for Z

Results

Schizophrenia GWAS summary data

copyright @ Havis 2017

- SCZ is a chronic and severe brain disease; affects about 1% of the worldwide population
- Highly heritable (70%–85%)
- Only a few hundred loci have been identified; enriched in non-coding regions
- SCZ1: 8,832 cases and 12,067 controls;
  SCZ2: 36,989 cases and 113,075 controls

"E + G + Methyl" identifies new significant and novel genes

Figure 1: Venn diagrams of the **significant genes** (left panel), and the **significant and novel** (right panel) genes identified by the different methods applied to the SCZ1 data

- novel gene: one that does not cover any GWAS risk variant within an 500 Kbp extension in the same dataset

Validation analysis of "E + G + Methyl" results

- Highly significant replication rates
  - 'E + G + Methyl" with SPU(1) identified 10 novel genes in the SCZ1 data, of which 6 (60%) contained genome-wide significant SNPs in the larger SCZ2 data ($p = 9.6 \times 10^{-6}$ by the hypergeometric test)
- Reported by other studies
  - Identified 22 significant and novel genes; 14 out of 22 ($p = 1.1 \times 10^{-14}$) have been reported by other studies

Adaptive test results

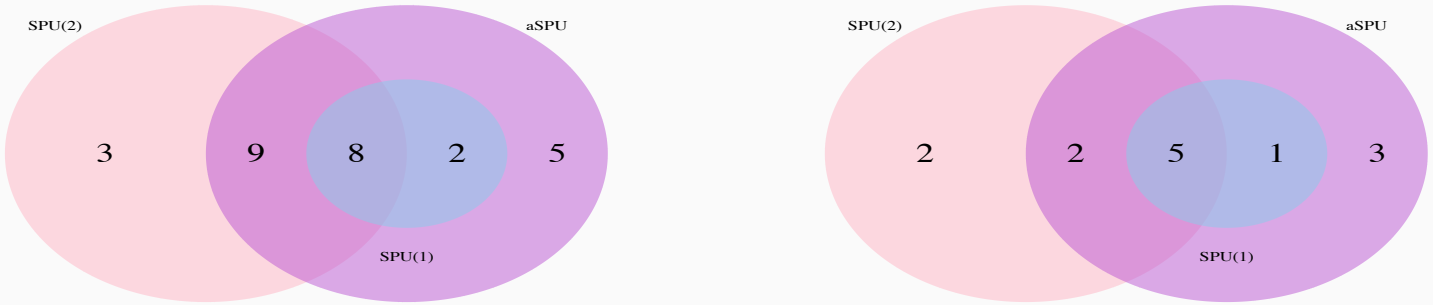


Figure 2: Venn diagrams of the **significant** genes(left panel), and the **significant and novel** (right panel) genes identified by the "E+G+Methyl" with different methods applied to the SCZ1 data

Conclusion

Conclusion

- Propose a simple but powerful gene-base test by integrating enhancer-promoter interactions and mQTL data with GWAS summary results
- Will be most useful when the enhancers, especially those far away from a gene, contain trait-associated mQTLs.
- It is complementary to the current methods

Thank you!

Reference

- Wu, Chong, and Wei Pan. "Integration of enhancer-promoter interactions with GWAS summary results identifies novel schizophrenia-associated genes and pathways." *Genetics* 209.3 (2018): 699-709.
- Wu, Chong, and Wei Pan. "Integration of methylation QTL and enhancer-target gene maps with schizophrenia GWAS summary results identifies novel genes." *Bioinformatics* (2019).
- Xu, Zhiyuan, et al. "A powerful framework for integrating eQTL and GWAS summary data." *Genetics* 207.3 (2017): 893-902.