Uncovering Genetic Factors Contributing to Type 2 Diabetes and Diabetic Nephropathy

Y. Bai,1 V. Nair,2 S. Martini,2 J. Cavalcoti,1 G. Omenn,1 M. Kretzler1,2
1 National Center for Integrative Biomedical Informatics, and Center for Computational Medicine and Biology; 2 Department of Internal Medicine
University of Michigan, Ann Arbor, MI 48109

Introduction

Recent genome-wide association studies [1-4] have made major discoveries in identifying Type 2 diabetes (T2D) associated regions and loci, but the specific sequence variances responsible for the associations remain elusive. To define putative causative gene sets from GWAS we employed a promoter modeling approach based on the hypothesis that promoter regions integrate upstream signaling cascades towards coordinated transcription of functionally interdependent mRNAs. Defining T2D dependent promoter models in GWAS candidate promoters might thereby facilitate identification of putative causative transcription alterations.

Here we studied the proximal promoter regions of 13 genes selected from T2D associated regions in the 3-way FUSION-DGI-WTCCC meta-analysis [1, 2, 4]. In particular, we used computational methods to identify shared putative regulatory promoter modules in the proximal promoter regions that we investigated. Specific potential regulatory promoter modules containing three transcription factor (TF) binding motifs in a defined order and spacing were identified in a subset of genes chosen from GWAS associated regions. These promoter modules helped elucidate other module sharing genes in the GWAS, which are possibly regulated in a similar fashion.

Our study provides TF binding module data that can putatively activate a subset of T2D GWAS genes.

Methods

1. Selecting significantly regulated transcripts from Diabetic Endorgan Damage (Diabetic Nephropathy) and module genes from T2D GWAS
2. Feeding selected genes into TFBS/promoter module analysis pipeline [6]
3. Training test T2D GWAS genes for identified promoter modules

We chose a region within 200kb of each SNP with p<.001 in the 3-way FUSION_DGI_WTCCC meta-analysis [1, 2, 4]. Two additional genes from large scale studies were added: Wolfram Syndrome (WFS1) and HNF1B.

The filtered PIMA gene expression dataset resulted in ~11000 genes to be expressed above background. 21 out of the 29 T2D associated genes were mapped to this dataset. SAM analysis on these 21 genes in the three datasets (glomeruli, tubulointerstitium and cortex) resulted in 12 significantly regulated genes. We additionally incorporated "CDKN2B" into the input gene list.

TF binding modules enrichment in 48 selected genes

Results

1. Initial modules identified

2. TF binding modules enrichment in 48 selected genes

Future work

1. Develop methods for incorporating the interactions for TF factors linked to putative module genes into cellular functional context via NCBI-MIMI database
2. Run NCBI-ConceptGen software on 61 genes
3. Apply NCBI tools (MIMI, Gene2MeSH, and SAGA) to look for detailed protein interactions, possible MeSH term enrichment, and pathway involvement for a subset of genes which share four TF binding modules to define functional context of co-regulation

References


Acknowledgement

This work was supported by National Institutes of Health. Grant #U54 DA021519