Integration of transcriptomics and GWAS data to study the biological processes involved in Type 2 Diabetes Mellitus

Manuel González

Supervisors
Martina Summer Kutmon, PhD
Lars Eijssen, PhD
Elisa Cirillo, MSc
Introduction

• Classical symptoms of Type 2 Diabetes Mellitus

• Hyperglycemia

• Insulin resistance

Figure 1 Risk factors for Type 2 diabetes. International Diabetes Federation
Introduction

Mitochondrial dysfunction → FFA → Hyperglycemia

Inflammatory mediators → Beta cell death

Adipose Tissue

Department of Bioinformatics
Methods - Workflow

Transcriptomics dataset

Pre-processing and statistical analysis

Gene Ontology

Pathway analysis

GWAS dataset

Variation prioritization

Merged pathways network

Integrative network analysis
Transcriptomics dataset

• Misu et al (2010)

• Aim: identify hepatic secretory proteins correlated with insulin resistance
  – GEO: GSE23343
  – Affymetrix Human Genome U133
  – Experimental groups:
    • 10 diabetic
    • 7 healthy controls
Pre-processing & statistical analysis

• Pre-processing:
  – ArrayAnalysis.org
  – GCRMA algorithm
  – 2 healthy samples were removed

• Statistical analysis
  – 311 up-regulated
    • logFC>0.58, pval<0.05
  – 555 down-regulated
    • logFC<-0.58, pval<0.05
Gene ontology analysis

- GO: biological processes, molecular functions and cellular components
- Gene annotations
- GO terms associated to gene input list
Gene ontology analysis

• Up&Down genes
  • Apoptotic Cell clearance (GO:0043277)

• Up-regulated genes
  • cholesterol biosynthetic process (GO:0006695)

• Down-regulated genes
  • organic acid metabolic process (GO:0006082)
  • base-excision repair, AP site formation (GO:00062285)
Pathway analysis

- Human Collection of WikiPathways
- Statistical Criteria:
  - $Z$-score $> 1.96$
  - P-value $< 0.05$
  - Minimum of 3 altered genes

- $Z$-score indicates less or more significantly altered genes than expected
## Pathway analysis

<table>
<thead>
<tr>
<th>Pathway</th>
<th>Z-score</th>
<th>P-value</th>
<th>#</th>
<th>Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cholesterol Biosynthesis</td>
<td>6.36</td>
<td>0.001</td>
<td>6/31</td>
<td>( \uparrow \text{HMGCR, HMGCS1, SC5DL, IDI1, SQLE, FDFT1} )</td>
</tr>
<tr>
<td>SREBF and miR33 in Cholesterol and Lipid homeostasis</td>
<td>4.7</td>
<td>0.001</td>
<td>5/19</td>
<td>( \uparrow \text{HMGCS1, HMGCR} ) ( \downarrow \text{NR1H3, SREBF1, MED15} )</td>
</tr>
<tr>
<td>NAD Biosynthesis</td>
<td>4.12</td>
<td>0.001</td>
<td>3/34</td>
<td>( \downarrow \text{KMO, HAAO, NADSYN1} )</td>
</tr>
<tr>
<td>Statin pathway</td>
<td>3.88</td>
<td>0.001</td>
<td>6/46</td>
<td>( \uparrow \text{HMGCR, SQLE, FDFT1} ) ( \downarrow \text{DGAT1, LRP1, ABCG8} )</td>
</tr>
<tr>
<td>Alanine and aspartate metabolism</td>
<td>3.3</td>
<td>0.005</td>
<td>3/62</td>
<td>( \downarrow \text{GPT, PC, ASPA} )</td>
</tr>
<tr>
<td>Osteopontin Signaling</td>
<td>3.3</td>
<td>0.012</td>
<td>3/14</td>
<td>( \uparrow \text{SPP1, IKK-alpha, ITGAV} )</td>
</tr>
<tr>
<td>Osteoblast signaling</td>
<td>3.1</td>
<td>0.005</td>
<td>3/19</td>
<td>( \uparrow \text{PDGFRA, ITGAV} ) ( \downarrow \text{PTH receptor} )</td>
</tr>
<tr>
<td>SREBP signaling</td>
<td>2.86</td>
<td>0.01</td>
<td>8/71</td>
<td>( \uparrow \text{HMGCR, HMGCS, FDFT, SQLE} ) ( \downarrow \text{SCAP, SREBP1a-c, nSREB, ARC105} )</td>
</tr>
</tbody>
</table>
Cholesterol Biosynthesis

Title: Cholesterol Biosynthesis
Last modified: 2/22/2013
Organism: Homo sapiens
Network building

1. Merged pathways network

2. Transcription factors (Cytargetlinker)
Merged pathways network

- Alanine and Aspartate metabolism
- Cholesterol Biosynthesis
- NAD Biosynthesis II
- Osteoblast Signaling
- Osteopontin Signaling
- SREBF and miR33 in cholesterol and lipid homeostasis
- SREPF Signaling
- Statin pathway
- Involved in at least 2 pathways
Transcription Factors
GWAS dataset

- *Johnson et al* (2009)
- Metanalysis (118 articles)
- Gene annotated database (56,411 SNPs)
- 2022 SNPs associated with T2DM
Altered pathways and SNP data
GWAS3D analysis

- Rs-number with P-value, cutoff P-value (10E-5)
- SNP data set (HapMap I+II+III)
- population of interest (CEU)
- SNPs-864 SNPs

Relations to
- BMI & Metabolic Syndrome
- Insulin exocytosis
- Hepatic steatosis
- Cholesterol score: rs693 in ApoB, rs328 in LPL, rs1800775 in CETP and rs4420638 in APOE cluster
Pathway-based analysis and prioritized variants

- Statin pathway
- Lipid metabolism
- Wnt signaling
- T2DM pathway
- SREBP signaling
- Il-1 signaling
- Apoptosis
- TLR signaling pathway
Conclusions

• Pathway analysis

• GO analysis
  – Apoptotic cell clearance
  – Cholesterol biosynthetic process
  – Organic Acid Metabolic process
  – Base Excision Repair
Conclusions

- PPARG, KCNJ11 and TCF7L2
- FTO, IGF2BP2 not linked to pathways
- CETP, APOB, APOC1, LPL, PPARG and CHUK
Future perspectives

• Consequence of variants and haplotypes
• Classification of subtypes of T2DM
• GWAS studies in non-European populations
• Environmental factors (epigenetics)
Thank you

- BiGCaT group
- Martina Summer Kutmon, PhD
- Lars Eijssen, PhD
- Elisa Cirillo, MSc