Multi–phenotype epigenome-wide association analysis of fasting glucose and insulin in 981 Finns

Hermanus H.M. Draisma[1]; Matthias Wielscher[2]; Saqib Hassan[1]; Zhanna Balkhiyarova[1]; Marjo-Riitta Jarvelin[2, 4]; Marika Kaakinen[1, 3]; Inga Prokopenko[1]

1.) Section of Genomics of Common Disease, Department of Medicine; 2.) Department of Epidemiology and Biostatistics, School of Public Health; and 3.) Centre for Pharmacology and Therapeutics/Division of Experimental Medicine, Imperial College London, UK. 4.) Center for Life-Course Health Research and Northern Finland Cohort Center, Biocenter Oulu, University of Oulu, Finland. Contact: h.draisma@imperial.ac.uk

Background

- Multi-phenotype genome-wide association studies (MP-GWAS) of correlated traits are more powerful than single-trait GWAS for locus discovery
- We have previously developed a MP-GWAS method using 'reverse regression' approach in which genotype dosage is regressed on a linear combination of phenotypes, implemented into the software tool SCOPA and meta-analysis tool METASCOPA [1]
- There is no multi-phenotype method for epigenome-wide association analysis (MP-EWAS)

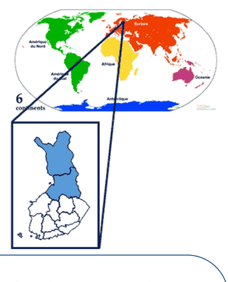
Aims

- To adapt the 'reverse regression' approach for epigenome-wide association analysis
- To test the method on metabolic traits

Material and Methods

- Cohorts: Northern Finland Birth Cohorts 1966 (NFBC1966, N=635) and 1986 (NFBC1986, N=346)
- Epigenetic data: Illumina Infinium HumanMethylation450K BeadChip array
- Phenotypic data: fasting glucose (FG) and fasting insulin (FI) ($r_{FG,FI}$=0.096)

Figure 1. Geographic location of the NFBC1966 and NFBC1986.[2]



Flowchart of the analysis process:

Development of methylSCOPA and META-methylSCOPA
- Extend the SCOPA software to allow for methylation data as input and to analyse methylation data analogous to genotype dosage data
- Extend METASCOPA software for meta-analysis of MP-EWAS results

Methylation data preprocessing & quality control (QC)
- Filter out unreliably detected methylation probes, and outlying individuals using principal component [PC] scores
- Subset quantile-normalize raw methylation signals
- Transform to "beta" values, inverse-normal transform
- Regress on PC scores (→ batch effect removal)[3] & 'Houseman estimates'[4] (→ leukocyte type composition correction)

Phenotype data preprocessing & QC
- Exclude non-fasting, pregnant, diabetic individuals
- Natural log-transform FI
- Regress FG & FI on sex, body mass index, fasting serum triglycerides & waist-hip ratio
- Use resulting residuals in the MP-EWAS

Cohort-specific analyses in methylSCOPA
- Model:

$\text{Methylation probe}_i = FG_{res} + \ln(FI)_{res} + e_i$,
$i=1,\ldots,459k$ in NFBC1966 and $i=1,\ldots,466k$ in NFBC1986

Meta-analysis
- Use METAmethylSCOPA for MP-EWAS results
- Use R packages metap and metafor for the meta-analysis of the phenotypic/methylation correlation
- Use $P<1\times10^{-7}$ to denote epigenome-wide significance

Annotation
- Map the results to CGCh37/hg19

Results

- The strongest association signal ($P=1.4\times10^{-7}$) is observed for a methylation marker at chromosome 22:49,088,813, mapping to *FAM19A5* gene
  - Single-trait EWAS results for the same marker: FI, P=0.13; FG, $P=4.0\times10^{-3}$
- We also observe an association ($P=9.84\times10^{-3}$) with a marker at chromosome 21:43,656,587, which maps to *ABCG1*, an established BMI, measures of glucose metabolism, and incident type 2 diabetes methylation locus
  - Single-trait EWAS results for the same marker : FI, P=0.44; FG, $P=8.25\times10^{-3}$

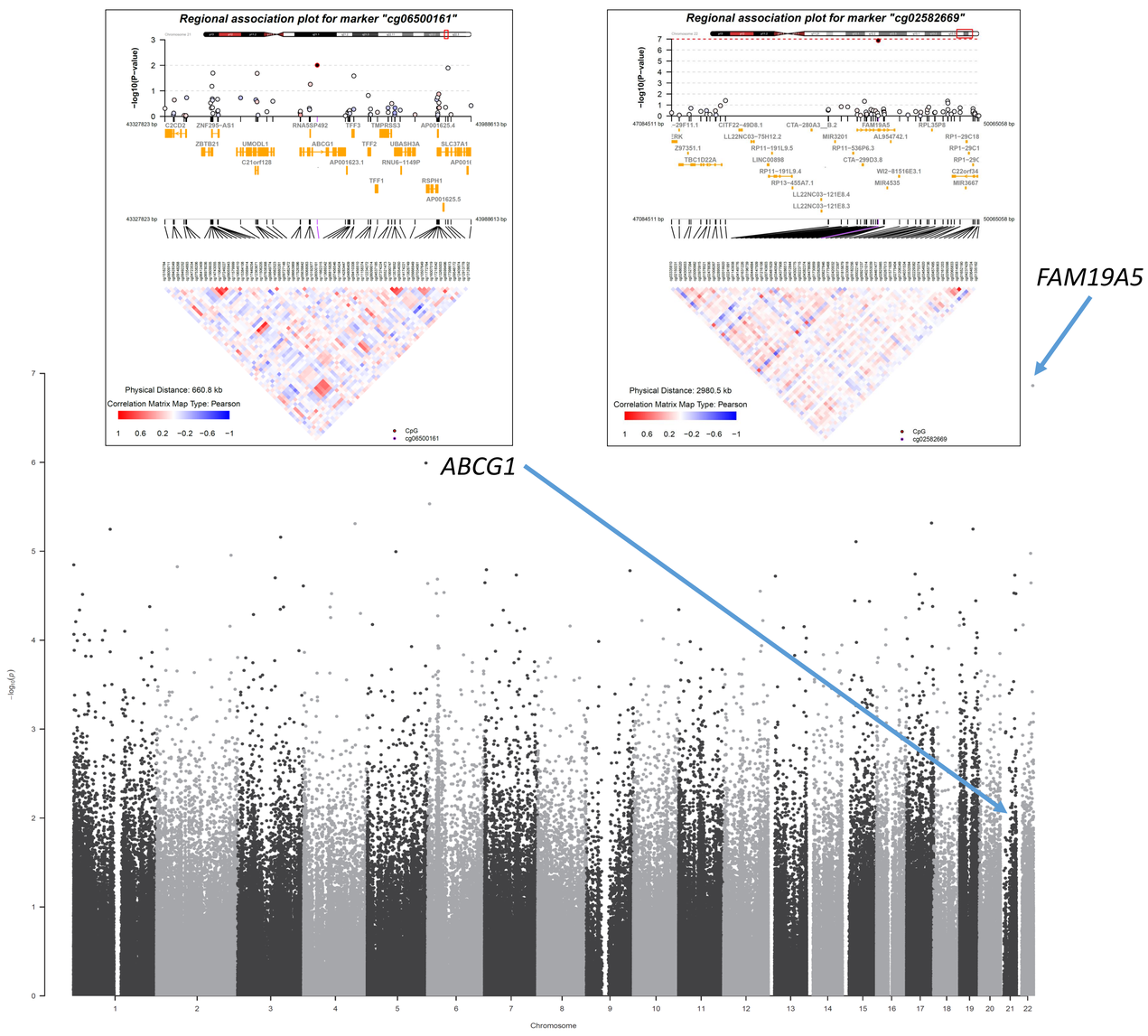


Figure 2. Manhattan plot displaying $-\log_{10}(P\text{-value})$ of association with FI and FG residuals in MP-EWAS meta-analysis for each methylation probe across the autosomal chromosomes. Insets show regional association plots for the 'top hit' at chr 22 and the established *ABCG1* locus at chr 21. The regional plots were made using the comet() function from the Bioconductor package 'coMET'.

Summary and conclusions

- We have successfully extended the multi-phenotype approach to methylation data
- We observed an association at an established locus for metabolic traits, as well as at a novel locus which needs further validation
- MP-EWAS with methylSCOPA is more powerful than single-trait EWAS for multi-phenotype epigenetic effect detection

References

1. Mägi *et al. BMC Bioinformatics* (2017) 18:25
2. http://www.oulu.fi/sites/default/files/86/continent.gif
3. Lehne *et al. Genome Biology* (2015) 16:37
4. Houseman *et al. BMC Bioinformatics* (2012) 13:86

Funding

This work is funded by Wellcome Trust Seed Award in Science (WT205915). The European Commission Marie Curie Intra-European Fellowship (project MARVEL, PIEF-GA-2013-626461, and the European Union's Horizon 2020 research and innovation programme (DYNAhealth, project number 633595). NFBC1966 received financial support from University of Oulu Grant no. 65354, Oulu University Hospital Grant no. 2/97, 8/97, Ministry of Health and Social Affairs Grant no. 23/251/97, 160/97, 190/97, National Institute for Health and Welfare, Helsinki Grant no. 54121, Regional Institute of Occupational Health, Oulu, Finland Grant no. 50621, 54231. NFBC1986 received financial support from EU QLG1-CT-2000-01643 (EUROBLCS) Grant no. E51560, NorFA Grant no. 731, 20056, 30167, USA / NIHH 2000 G DF682 Grant no. 50945.