A distinct metabolic signature predicts development of fasting plasma glucose

Manuela Hische
Charite Universitätsmedizin Berlin

Abdelhalim Larhlimi
University of Potsdam

Franziska Schwarz
Charite Universitätsmedizin Berlin

Antje Fischer-Rosinsky
Charite Universitätsmedizin Berlin

Thomas Bobbert
Charite Universitätsmedizin Berlin

Anke Assmann
Charite Universitätsmedizin Berlin

Gareth Catchpole
Max-Planck-Institute for Molecular Plant Physiology

Andreas Pfeiffer
Charite Universitätsmedizin Berlin

Lothar Willmitzer
Max-Planck-Institute for Molecular Plant Physiology

Joachim Selbig
University of Potdam

Joachim Spranger
Charite Universitätsmedizin Berlin

Abstract

Type 2 diabetes reduces life expectancy worldwide and numerous studies have been performed to identify risk factors of type 2 diabetes. Nevertheless, this is a topic that is subject to continuing discussion. Established classical markers include: family history of diabetes, markers of adiposity, age and glycemic control itself. In recent years, high-throughput methods have been increasingly applied in clinical research. We here applied multivariate statistical methods to account for dependencies within the metabolome.

Fasting plasma glucose levels were measured at baseline and at follow-up after an average of five years in subjects who participated in the prospective MetabolicSyndrome BerlinPotsdam (MESY-BEPO) study. We compared the quality of prediction between this metabolic pattern and established risk markers. We analysed metabolic profiles of baseline fasting plasma samples in a random sub-cohort (n=172) using Gas Chromatography coupled with time-of-flight Mass Spectrometry (GC-MS). We measured in total 286 metabolites, some of them are not yet

*hische@mpimp-golm.mpg.de

identified. The measurements cover various biochemical classes, such as amino acids, carbohydrates, organic acids, fatty acids and steroids. The chromatographic peaks were picked and identified using the Golm Metabolome Database (GMD) and the R package TargetSearch. Log-transformation and normalisation of the measured relative intensities were performed according to previously published methods. Since missing values only occurred if metabolite concentration went below detection limit, these values were replaced by a value 0.7 times the minimum measured value. We used Random Forests for a regression analysis to predict the development of fasting glucose levels. To evaluate the performance of this prediction we calculated the Pearson product-moment correlation coefficient between the real and calculated delta-glucose for all samples on the one hand and for the *out-of-bag* samples (the test samples of a form of cross-validation) on the other hand.

In doing so, we were able to define a complex pattern of metabolites that predicts future development of fasting plasma glucose levels with high accuracy. Notably, not single, but a complex pattern of metabolites propels the prediction and therefore reflects the complexity of the underlying molecular mechanisms.