

# Identification of marker genes for type 2 diabetes

A. Rasche · R. Herwig

Published online: 19 September 2007  
© Springer-Verlag and NuGO 2007

**Keywords** Nutritional systems biology · Data integration

Diabetes is a complex trait, reflecting interactions of a network of genes with environmental factors, such as obesity, age and nutrition [1]. The polygenic nature of type 2 diabetes is now well established. The individual contributions of each individual gene to the phenotype is likely to be small. The primary objective of our work is to provide identification and data integration for obesity and type 2 diabetes mellitus.

Microarrays have become a standard technology platform for genome screening. This technique fails in the integration of other data sources relevant for diagnostic purposes. For these tasks, data integration methods are indispensable. In our poster we present strategies to incorporate these sources within the framework of R/Bio-Conductor [2, 3].

To elucidate the most important genes for diabetes we employ scoring to gain new knowledge on the pathways and marker genes relevant. Pathway information from databases like KEGG [4] is comprised to map the genes into their function/framework.

Ultimately, the proposed marker genes will enable us to develop models for the disease. Medical researcher can

combine molecular biological data with clinically relevant parameters to predict quantitatively individual predispositions to type 2 diabetes and include physiological data.

**Acknowledgments** Thanks goes to Christoph Wierling and Atanas Kamburov for providing the pathway data in a convenient way. We also thank Hadi Al-Hasani from the DIFE for the QTL's and his biological knowledge. A.R. is funded by a Max-Planck Stipend for foreign students, R.H. is employee of the Max-Planck-Society.

## References

1. Stumvoll M, Goldstein BJ, van Haeften TW (2005) Type 2 diabetes: principles of pathogenesis and therapy. *The Lancet* 365:1333–1346
2. Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28(1):27–30
3. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-07-0
4. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JYH, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5:R80

---

A. Rasche (✉) · R. Herwig  
Department Vertebrate Genomics, Bioinformatics Group,  
Max Planck Institute for Molecular Genetics,  
Inhnestr. 63–73, 14195 Berlin-Dahlem, Germany  
e-mail: rasche@molgen.mpg.de