DiaGWAS: Systems biology genome-wide association studies to identify novel type 2 diabetes molecular mechanisms

Description

Despite tremendous investments to identify susceptibility genes for type 2 diabetes (T2D) through genome-wide association studies (GWAS), only around 10% of the variance attributable to genetic factors can be explained with the currently identified variants. In addition, each of the identified locus frequently maps to several genes, remaining a challenge to assign a causal gene to each locus. On the other hand, while a huge amount of biological knowledge is available from several datasets derived from both omics experiments and data mining of the scientific literature, these data have not been fully exploited to better guide and interpret the analysis of GWAS.

The main goal of this project was to overcome these limitations taking advantage of available GWAS data and biological databases and using advanced statistical and systems biology methodologies to identify causal genes and pathways involved in the pathophysiology of T2D.

Using enhanced genotype imputation techniques applied to available GWAS datasets comprising up to 11.549 T2D cases and 12.276 controls, we aimed to identify new loci, as well as fine mapping the existing ones. We then used manually curated T2D specific pathways and networks reconstructed by our group through integration and data mining of several omics and literature sources. Using novel in-house developed pathway and network analysis methods for GWAS, we aimed to identify new key biological processes and genes perturbed in T2D subjects.

Finally, we experimentally validated the findings by analysing the effect of the discovered variants on gene expression in independent cohorts for which DNA and tissue banks are available. This project aimed to allow the discovery of novel molecular mechanisms involved in T2D, opening new lines of research, to provide a novel framework to better exploit the existing available GWAS data to identify the molecular basis of complex diseases.

Barcelona Supercomputing Center - Centro Nacional de Supercomputación