

A genome-wide association study identifies novel risk loci for type 2 diabetes

Robert Sladek^{1,2,4}, Ghislain Rocheleau^{1*}, Johan Rung^{4*}, Christian Dina^{5*}, Lishuang Shen¹, David Serre¹, Philippe Boutin⁵, Daniel Vincent⁴, Alexandre Belisle⁴, Samy Hadjadj⁶, Beverley Balkau⁷, Barbara Heude⁷, Guillaume Charpentier⁸, Thomas J. Hudson^{4,9}, Alexandre Montpetit⁴, Alexey V. Pshzhetsky¹⁰, Marc Prentki^{10,11}, Barry I. Posner^{2,12}, David J. Balding¹³, David Meyre⁵, Constantin Polychronakos^{1,3} & Philippe Froguel^{5,14}

Type 2 diabetes mellitus results from the interaction of environmental factors with a combination of genetic variants, most of which were hitherto unknown. A systematic search for these variants was recently made possible by the development of high-density arrays that permit the genotyping of hundreds of thousands of polymorphisms. We tested 392,935 single-nucleotide polymorphisms in a French case-control cohort. Markers with the most significant difference in genotype frequencies between cases of type 2 diabetes and controls were fast-tracked for testing in a second cohort. This identified four loci containing variants that confer type 2 diabetes risk, in addition to confirming the known association with the *TCF7L2* gene. These loci include a non-synonymous polymorphism in the zinc transporter *SLC30A8*, which is expressed exclusively in insulin-producing β -cells, and two linkage disequilibrium blocks that contain genes potentially involved in β -cell development or function (*IDE-KIF11-HHEX* and *EXT2-ALX4*). These associations explain a substantial portion of disease risk and constitute proof of principle for the genome-wide approach to the elucidation of complex genetic traits.

The rapidly increasing prevalence of type 2 diabetes mellitus (T2DM) is thought to be due to environmental factors, such as increased availability of food and decreased opportunity and motivation for physical activity, acting on genetically susceptible individuals. The heritability of T2DM is one of the best established among common diseases and, consequently, genetic risk factors for T2DM have been the subject of intense research¹. Although the genetic causes of many monogenic forms of diabetes (maturity onset diabetes in the young, neonatal mitochondrial and other syndromic types of diabetes mellitus) have been elucidated, few variants leading to common T2DM have been clearly identified and individually confer only a small risk (odds ratio \approx 1.1–1.25) of developing T2DM¹. Linkage studies have reported many T2DM-linked chromosomal regions and have identified putative, causative genetic variants in *CAPN10* (ref. 2), *ENPP1* (ref. 3), *HNF4A* (refs 4, 5) and *ACDC* (also called *ADIPOQ*)⁶. In parallel, candidate-gene studies have reported many T2DM-associated loci, with coding variants in the nuclear receptor *PPARG* (P12A)⁷ and the potassium channel *KCNJ11* (E23K)⁸ being among the very few that have been convincingly replicated. The strongest known (odds ratio \approx 1.7) T2DM association⁹ was recently mapped to the transcription factor *TCF7L2* and has been consistently replicated in multiple populations^{10–20}.

Subjects and study design

The recent availability of high-density genotyping arrays, which combine the power of association studies with the systematic nature of a genome-wide search, led us to undertake a two-stage, genome-wide association study to identify additional T2DM susceptibility loci (Supplementary Fig. 1). In the first stage of this study, we obtained

genotypes for 392,935 single-nucleotide polymorphisms (SNPs) in 1,363 T2DM cases and controls (Supplementary Table 1). In order to enrich for risk alleles²¹, the diabetic subjects studied in stage 1 were selected to have at least one affected first degree relative and age at onset under 45 yr (excluding patients with maturity onset diabetes in the young). Furthermore, in order to decrease phenotypic heterogeneity and to enrich for variants determining insulin resistance and β -cell dysfunction through mechanisms other than severe obesity, we initially studied diabetic patients with a body mass index (BMI) $<30 \text{ kg m}^{-2}$. Control subjects were selected to have fasting blood glucose $<5.7 \text{ mmol l}^{-1}$ in DESIR, a large prospective cohort for the study of insulin resistance in French subjects²².

Genotypes for each study subject were obtained using two platforms: Illumina Infinium Human1 BeadArrays, which assay 109,365 SNPs chosen using a gene-centred design; and Human Hap300 BeadArrays, which assay 317,503 SNPs chosen to tag haplotype blocks identified by the Phase I HapMap²³. Of the 409,927 markers that passed quality control (Supplementary Tables 2 and 3), genotypes were obtained for an average of 99.2% (Human1) and 99.4% (Hap300) of markers for each subject with a reproducibility of $>99.9\%$ (both platforms). Forty-three subjects were removed from analysis because of evidence of intercontinental admixture (Supplementary Fig. 3) and an additional four because their genotype-determined gender disagreed with clinical records. In total, T2DM association was tested for 100,764 (Human1) and 309,163 (Hap300) SNPs representing 392,935 unique loci (Fig. 1). Because of unequal male/female ratios in our cases and controls, we analysed the 12,666 sex-chromosome SNPs separately for each gender.

¹Departments of Human Genetics, ²Medicine and ³Pediatrics, Faculty of Medicine, McGill University, Montreal H3H 1P3, Canada. ⁴McGill University and Genome Quebec Innovation Centre, Montreal H3A 1A4, Canada. ⁵CNRS 8090-Institute of Biology, Pasteur Institute, Lille 59019 Cedex, France. ⁶Endocrinology and Diabetology, University Hospital, Poitiers 86021 Cedex, France. ⁷INSERM U780-IFR69, Villejuif 94807, France. ⁸Endocrinology-Diabetology Unit, Corbeil-Essonnes Hospital, Corbeil-Essonnes 91100, France. ⁹Ontario Institute for Cancer Research, Toronto M5G 1L7, Canada. ¹⁰Montreal Diabetes Research Center, Montreal H2L 4M1, Canada. ¹¹Molecular Nutrition Unit and the Department of Nutrition, University of Montreal and the Centre Hospitalier de l'Université de Montréal, Montreal H3C 3J7, Canada. ¹²Polypeptide Hormone Laboratory and Department of Anatomy and Cell Biology, Montreal H3A 2B2, Canada. ¹³Department of Epidemiology & Public Health, Imperial College, Imperial College, St Mary's Campus, Norfolk Place, London W2 1PG, UK. ¹⁴Section of Genomic Medicine, Imperial College London W12 0NN, and Hammersmith Hospital, Du Cane Road, London W12 0HS, UK.

*These authors contributed equally to this work.