Original Article

MOLECULAR METABOLISM

Expression and functional assessment of candidate type 2 diabetes susceptibility genes identify four new genes contributing to human insulin secretion

Fatou K. Ndiaye [1,6], Ana Ortalli [1,6], Mickaël Canouil [1,6], Marlène Huyvaert [1], Clara Salazar-Cardozo [1], Cécile Lecoeur [1], Marie Verbanck [1], Valérie Pawlowski [1], Raphaël Boutry [1], Emmanuelle Durand [1], Iandry Rabearivelo [1], Olivier Sand [1], Lorella Marselli [2], Julie Kerr-Conte [3], Vikash Chandra [4], Raphaël Scharfmann [4], Odile Poulain-Godefroy [1], Piero Marchetti [2], François Pattou [3], Amar Abderrahmani [1,5], Philippe Froguel [1,5,,7], Amélie Bonnefond [1,5,*,7]**

ABSTRACT

Objectives: Genome-wide association studies (GWAS) have identified >100 loci independently contributing to type 2 diabetes (T2D) risk. However, translational implications for precision medicine and for the development of novel treatments have been disappointing, due to poor knowledge of how these loci impact T2D pathophysiology. Here, we aimed to measure the expression of genes located nearby T2D associated signals and to assess their effect on insulin secretion from pancreatic beta cells.

Methods: The expression of 104 candidate T2D susceptibility genes was measured in a human multi-tissue panel, through PCR-free expression assay. The effects of the knockdown of beta-cell enriched genes were next investigated on insulin secretion from the human EndoC-βH1 beta-cell line. Finally, we performed RNA-sequencing (RNA-seq) so as to assess the pathways affected by the knockdown of the new genes impacting insulin secretion from EndoC-βH1, and we analyzed the expression of the new genes in mouse models with altered pancreatic beta-cell function.

Results: We found that the candidate T2D susceptibility genes' expression is significantly enriched in pancreatic beta cells obtained by laser capture microdissection or sorted by flow cytometry and in EndoC-βH1 cells, but not in insulin sensitive tissues. Furthermore, the knockdown of seven T2D-susceptibility genes (*CDKN2A*, *GCK*, *HNF4A*, *KCNK16*, *SLC30A8*, *TBC1D4*, and *TCF19*) with already known expression and/or function in beta cells changed insulin secretion, supporting our functional approach. We showed first evidence for a role in insulin secretion of four candidate T2D-susceptibility genes (*PRC1*, *SRR*, *ZFAND3*, and *ZFAND6*) with no previous knowledge of presence and function in beta cells. RNA-seq in EndoC-βH1 cells with decreased expression of *PRC1*, *SRR*, *ZFAND6*, or *ZFAND3* identified specific gene networks related to T2D pathophysiology. Finally, a positive correlation between the expression of *Ins2* and the expression of *Prc1*, *Srr*, *Zfand6*, and *Zfand3* was found in mouse pancreatic islets with altered beta-cell function.

Conclusions: This study showed the ability of post-GWAS functional studies to identify new genes and pathways involved in human pancreatic beta-cell function and in T2D pathophysiology.

Keywords EndoC-βH1; Expression analysis; Genome-wide association study; Insulin secretion; RNAi screening; Type 2 diabetes

[1]CNRS UMR 8199, European Genomic Institute for Diabetes (EGID), Institut Pasteur de Lille, University of Lille, 59000 Lille, France [2]Department of Clinical and Experimental Medicine, Islet Cell Laboratory, University of Pisa, 56100 Pisa, Italy [3]Inserm U1190, EGID, CHU Lille, University of Lille, 59000 Lille, France [4]Inserm U1016, Institut Cochin, Faculté de Médecine, Paris Descartes University, Sorbonne Paris Cité, 75014 Paris, France [5]Department of Genomics of Common Disease, Imperial College London, W12 0NN London, United Kingdom

[6] Co-first authors.
[7] Co-senior authors.

*Corresponding author. CNRS UMR8199, Institut Pasteur de Lille, 1 Rue du Professeur Calmette B.P. 245, F-59019 Lille, France. E-mail: amelie.bonnefond@inserm.fr (A. Bonnefond).

**Corresponding author. CNRS UMR8199, Institut Pasteur de Lille, 1 Rue du Professeur Calmette B.P. 245, F-59019 Lille, France. E-mail: p.froguel@imperial.ac.uk (P. Froguel).

Received February 18, 2017 • Revision received March 17, 2017 • Accepted March 24, 2017 • Available online xxx

http://dx.doi.org/10.1016/j.molmet.2017.03.011

Original Article

1. INTRODUCTION

Type 2 diabetes (T2D) is a complex, multifactorial disorder with an estimated heritability ranging between 40% and 70% [1]. Genetic studies have identified 28 genes causing monogenic diabetes due to impaired insulin secretion. These genes are critically important for pancreatic beta-cell lineage, phenotype, and function [1]. Genome-wide association studies (GWAS) have shown that common T2D is highly polygenic, with more than 100 loci contributing to T2D risk. GWAS of related quantitative traits have shown that a significant part of the T2D-associated single nucleotide polymorphisms (SNPs) also modulates fasting plasma glucose levels and the homeostatic model assessment (HOMA) of beta-cell function in non-diabetic individuals, suggesting that their main diabetogenic impact was on insulin secretion and not on insulin resistance [2]. Surprisingly indeed, only a few T2D-associated SNPs were found to have an effect on fasting serum insulin levels, the HOMA of insulin resistance, or the insulin sensitivity index in general populations. About half of the T2D susceptibility SNPs did not have any apparent effect on glucose homeostasis in non-diabetics [2].

Elucidating the function of T2D-associated SNPs (and target genes) and their involvement in T2D pathophysiology may have major translational implications for precision medicine and for the development of novel treatments. Post-GWAS studies are challenging due to the fact that the vast majority of these SNPs are non-coding and often intergenic, which has not facilitated functional investigations of genes located in these chromosome regions. In this context, a recent study has demonstrated that T2D-associated SNPs are significantly enriched in clusters of enhancers that are active in human pancreatic islets, and most of these enhancers map less than 500 kb from transcription start sites (TSS) of nearby genes [3]. Moreover, it has been recently shown that most genes located nearby loci associated with blood pressure (systolic, diastolic, pulse pressure) are highly expressed in vascular tissues [4]. These data suggested that focusing the expression studies and functional investigations on genes nearby GWAS-identified SNPs may be a good first strategy, as they may drive pathways involved in disease pathophysiology.

Here, we aimed to identify novel genes contributing to T2D risk through a dominant effect on insulin secretion in human. To do so, we first performed a comprehensive expression study of candidate T2D susceptibility genes closest to all GWAS-identified T2D SNPs in a large panel of human organs, tissues, and cells, followed by the functional analysis of the knockdown of these genes in human pancreatic beta-cell lines. Using this strategy, we first found that the expression of tested candidate T2D susceptibility genes was significantly and specifically enriched in pancreatic beta cells, and we reported functional evidence for a role in insulin secretion of four T2D susceptibility genes (*PRC1*, *SRR*, *ZFAND3*, and *ZFAND6*) with previously unknown presence and function in pancreatic beta cells.

2. MATERIAL AND METHODS

2.1. Samples included in the panel expression study

Total RNA from human colon, liver, kidney, adipose tissue, lung, skeletal muscle, heart, brain, small intestine, substantia nigra, hippocampus, dorsal root ganglion, and insula, and Poly A + RNA from human hypothalamus, pituitary gland, caudate nucleus, and frontal lobe were purchased from Clontech Laboratories (Palo Alto, CA, USA). Pancreatic islets ($n = 8$; average purity: 70.0%; average viability: 94.3%) and exocrine pancreas ($n = 2$) were isolated from adult brain-dead donors without diabetes, in accordance with French and Italian local ethics committee approval [5,6]. Pancreatic beta cells were obtained by laser capture microdissection (LCM beta cells; $n = 2$) [7] or were sorted by flow cytometry (FACS sorted beta cells; $n = 5$), as previously described [8]. Primary pre-adipocytes (Lonza, Basel, Switzerland) came from subcutaneous fat of a non-diabetic Caucasian female patient. These pre-adipocytes were differentiated into mature adipocytes in PBM-2 medium supplemented with insulin, dexamethasone, phosphodiesterase inhibitor 3-isobutyl-1-methylxanthine (IBMX), and indomethacin (all supplied by Lonza) for 10 days according to the manufacturer's recommendations. RNA samples from human EndoC-βH1 cells ($n = 3$), pre-adipocytes, and mature adipocytes were extracted as described below (see Section 2.9).

2.2. Probe design and selection for the panel expression study

Capture probes (containing a biotin affinity tag) and reporter probes (containing a color-coded molecular barcode) were designed to detect 148 targets, including five housekeeping genes for normalization, genes whose expression is known to be enriched ('markers') in pancreatic islet including beta cell ($n = 13$), gut ($n = 3$), kidney ($n = 4$), lung ($n = 3$), and adipose tissue ($n = 3$), 28 genes known to be involved in monogenic diabetes, and 104 genes located nearby GWAS-identified SNPs associated with T2D risk (Table A) [1]. The expression of the present housekeeping genes was previously shown to be uniform across a large panel of human tissues using RNA-sequencing (RNA-seq) [9]. Here, we found that the expression of the five housekeeping genes was highly correlated across the present panel of human tissues ($R^2 > 0.85$). The genes for which expression is known to be enriched in specific tissues were chosen according to the literature, the GTEx portal [10] and/or the BioGPS portal [11].

2.3. Panel gene expression study and statistical analysis

The detection of transcripts was carried out in multiplexed hybridization reactions using NanoString Technologies (Seattle, WA, USA), following manufacturer's protocol. Each hybridization reaction contained 25–100 ng total RNA or 5 ng Poly A + RNA at a final concentration of 0.8–3.3 ng/ul and 0.16 ng/ul respectively. Reagents were mixed and incubated at 65 °C in a thermocycler block with a heated lid for 16 h. All post-hybridization steps were handled robotically on a custom liquid-handling robot (Prep Station, NanoString Technologies). Finally, samples were loaded into a microfluidic chamber (nCounter Cartridge, NanoString Technologies) and imaged in a Digital Analyzer (NanoString Technologies). Barcodes were counted in 555 fields of view per sample. Experimental quality control was performed with nSolver (NanoString Technologies) to flag failed samples. Normalized expression values were first obtained by considering the logarithm of the ratio of the expression of a given gene over the average expression of the set of five housekeeping genes (Table A) in corresponding samples. The expression profiles of relevant gene sets were analyzed through heat map representations. A double hierarchical clustering of the tissues and gene expressions in each set was performed, using Ward's method. Log2 expression was centered and scaled. Each cell was colored to quantitatively reflect the relative expression: from green if the gene is over-expressed in the tissue compared to all other tissues, to red otherwise. For enrichment analyses, we established an expression threshold to classify 'under-expressed' and 'over-expressed' sets of genes (*i.e.* genes involved in monogenic diabetes, candidate susceptibility genes for T2D, and the different markers) in each panel tissue. For this purpose, the threshold was defined as the average gene expression across all tissues, plus 1.5 standard deviation (SD). We built a contingency table to count the number of genes from each set that were under- and over-expressed

