Integrative transcriptomic analysis of pancreatic islets from patients with prediabetes/type 2 diabetes

María V. Mencucci | Luis E. Flores | Juan J. Gagliardino | Martín C. Abba | Bárbara Maiztegui

1CENEXA. Centro de Endocrinología Experimental y Aplicada (UNLP-CONICET-CEAS CICPBA), Facultad de Ciencias Médicas UNLP, La Plata, Argentina
2CINIBA. Centro de Investigaciones Inmunológicas Básicas y Aplicadas (UNLP-CICPBA), Facultad de Ciencias Médicas UNLP, La Plata, Argentina

Correspondence
Bárbara Maiztegui, 60 y 120 (s/n) 4to piso Fac. Cs. Médicas (UNLP), 1900 La Plata, Argentina.
Email: barmaiztegui@hotmail.com

Funding information
Consejo Nacional de Investigaciones Científicas y Técnicas, Grant/Award Number: PUE 2017 Nº 22920170100082CO

Abstract
Aim: To identify new transcriptomic alterations in pancreatic islets associated with metabolic dysfunctions in people with prediabetes (PD)/type 2 diabetes (T2D).

Materials and methods: We collected information from public data repositories T2D related microarray datasets from pancreatic islets. We identified Differential Expressed Genes (DEGs) in non-diabetic (ND) vs people with T2D in each study. To identify relevant DEGs in T2D, we selected those that varied consistently in the different studies for further meta-analysis and functional enrichment analysis. DEGs were also evaluated at the PD stage.

Results: A total of seven microarray datasets were collected and analysed to find the DEGs in each study and meta-analysis was performed with 245 ND and 96 T2D cases. We identified 55 transcriptional alterations potentially associated with specific metabolic dysfunctions in T2D. Meta-analysis showed that 87% of transcripts identified as DEGs (48 out of 55) were confirmed as having statistically significant up- or down-modulation in T2D compared to ND. Notably, nine of these DEGs have not been previously reported as dysregulated in pancreatic islets from people with T2D. Consistently, the most significantly enriched pathways were related to the metabolism and/or development/maintenance of β-cells. Eighteen of the 48 selected DEGs (38%) showed an altered expression in islets from people with PD.

Conclusions: These results provide new evidence to interpret the pathogenesis of T2D and the transition from PD to T2D. Further studies are necessary to validate its potential use for the development/implementation of efficient new strategies for the prevention, diagnosis/prognosis and treatment of T2D.

KEYWORDS
gene expression microarray, pancreatic islets, Prediabetes, type 2 diabetes

1 INTRODUCTION

Type 2 diabetes (T2D) is a worldwide, continuously increasing disease characterized by metabolic dysfunction diagnosed from persistent hyperglycemia resulting from impaired insulin secretion or action. It represents a serious public health problem due to its frequent association with other cardiovascular risk factors and the development/progression of chronic complications that decrease the quality of life of the patients and significantly increase their cost of care. Although they can be prevented, these complications result mainly from delayed diagnosis and poor metabolic control of the disease.

Abbreviations: DEG, differential expressed genes; ND, non-diabetic; PD, prediabetes; T2D, type 2 diabetes.
The clinical manifestations of T2D are preceded by a period of metabolic dysfunction known as prediabetes (PD), characterized by impaired fasting glucose, impaired glucose tolerance, or the association of these alterations.\(^4\) The transition from PD to T2D can be significantly prevented (up to 58%) by adopting healthy lifestyles.\(^5,6\)

A key factor in the pathogenesis of T2D is the early and progressive decrease in pancreatic \(\beta\)-cell mass and function which results in deficient insulin secretion frequently associated with decreased response of target tissues to this hormone’s action.\(^7,8\) Although several molecular alterations have been identified in the pancreatic \(\beta\)-cells of people with T2D,\(^7\) its gene expression profiling and the underlying molecular alterations of its pathogenesis are still subjects of active research.

In recent years, comprehensive assessments of molecular alterations have been obtained by studying the transcriptomic features of isolated islets by microarray or RNA-sequencing. However, these studies require high technology and economic resources not always easily available. In this regard, some databases compile data from these types of assays, such as ArrayExpress (https://www.ebi.ac.uk/arrayexpress/) and GEO (https://www.ncbi.nlm.nih.gov/gds). These data are freely available for utilization by the research community, and their adequate integration/interpretation facilitates elucidation of significant biological alterations.

On account of this situation we have currently collected, integrated, and analysed datasets of microarray studies attempting to identify new relevant transcriptomic alterations in pancreatic islets/\(\beta\)-cells from people without diabetes (ND), with T2D and with PD. We assume that these data may provide new evidence to understand the pathogenesis of PD/T2D, facilitating the development of effective strategies for early diagnosis of both diseases, and also prevent the transition of PD to T2D.

2 | MATERIALS AND METHODS

2.1 | Selection of microarray datasets from GEO and ArrayExpress

T2D related microarray datasets from pancreatic islets were collected from public data repositories: Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/) and ArrayExpress (http://www.ebi.ac.uk/arrayexpress/). The keywords used to perform the search were “type 2 diabetes” and “pancreatic islets/pancreatic \(\beta\)-cells” and “\textit{Homo sapiens}”.

We selected only those studies from transcription profiling by array in which both ND and T2D people were included and each group contained more than five samples. The search was expanded using PubMed with the same selection criteria mentioned above. In the case of PD, we also searched related microarray datasets from pancreatic islets using the keywords: “prediabetes/impaired glucose tolerance/impaired fasting glucose”, “pancreatic islets/pancreatic \(\beta\)-cells”, and “\textit{Homo sapiens}”. These searches were updated in April 2019.

Using the GEO and ArrayExpress databases and the keywords mentioned above, we found 22 different studies containing samples of people without (ND) and with T2D. Although six of them were array-based transcriptomics studies, one was excluded because it included only one patient with diabetes. Therefore, we used five studies from this source: GSE50397 (Fadista et al.\(^{10}\) Taneera et al.\(^{11}\) and Taneera et al.\(^{12}\)), GSE38642 (Taneera et al.\(^{13}\) Taneera et al.\(^{14}\) and Kanatsuna et al.\(^{15}\)), GSE25724 (Dominguez et al.\(^{16}\)), GSE20966 (Marselli et al.\(^{17}\)), E-CBIL-20 (Gunton et al.\(^{18}\)), plus two others from PubMed search: GSE76894 and GSE76895 (Solimena et al.\(^ {19}\)). Their main data characteristics are summarized in Table 1.

2.2 | Identification of differential expressed genes across T2D studies

Raw data (CEL. format files) of the following studies were downloaded from GEO/ArrayExpress: GSE50397,\(^{10-12}\) GSE38642,\(^ {13-15}\) GSE25724,\(^ {16}\) GSE20966,\(^{17}\) E-CBIL-20,\(^ {18}\) GSE76894 and GSE76895.\(^ {19}\) These datasets were preprocessed using R/Bioconductor package Oligo\(^ {20}\) (GSE50397, GSE38642) or Affy\(^ {21}\)R packages (GSE25724, GSE20966, E-CBIL-20, GSE76894, GSE76895) according to the platform of each dataset.

Data were subjected to background correction, normalization and calculation of expression values using the robust multi-array average algorithm.\(^ {22}\)

Differential Expressed Genes (DEGs) between ND and T2D people were detected employing the LIMMA package (Linear Models for Microarray data).\(^ {23}\) \(P\) value \(<.05\) was considered statistically significant and fold change (FC) \(>1.5\) was considered biologically significant. Genes that met both criteria were considered DEGs in this study. Those that varied consistently in the same direction (down- or up-regulated) in at least three studies were considered relevant DEGs in T2D. This identification was effected using Venn diagrams (http://bioinformatics.psb.ugent.be/webtools/Venn/). To integrate the information of the different studies, a random effect size meta-analysis for each relevant DEG identified across the seven studies was performed using Comprehensive Meta-Analysis Software (https://www.meta-analysis.com/). Since more than one probe can map a gene, to perform this analysis we choose the probe that presented expression level values with the greatest variance (among the significant ones, if there were).

2.3 | Pathway and molecular interaction analysis

Pathway analysis of the relevant DEGs in T2D were performed using the resource InnateDB (https://www.innatedb.com),\(^ {24}\) one of the most comprehensive sources of pathways available. This type of analysis allows determination of those biological pathways that are significantly over-represented (represented more than expected by chance) in a list of certain genes/proteins.

Since molecular interactions are important for studies of regulation of biological systems, we have built, visualized, and analysed
molecular interactions among proteins encoded by the relevant DEGs using the platform NetworkAnalyst (http://www.networkanalyst.ca). Specifically, a protein-protein interaction network was done with IMEx Interactoma database (International Molecular Exchange Consortium), a non-redundant set of physical molecular interaction data from a broad taxonomic range of organisms. The proteins involved in the most interactions were key nodes in the network.

### 2.4 | Analysis of selected DEGs expression in ND vs PD cases

The expression level of the relevant DEGs was determined in ND, people with PD and people with T2D from GSE50397 and GSE76895 studies. Since more than one probe can map a gene, to perform this analysis we choose the probe that presented expression level values with the greatest variance. Scatter dot plots and statistical analyses were performed using GraphPad Prism version 5.01 (GraphPad Software). Data are presented as the mean ± SE of the mean (SEM). Analyses between two groups (ND vs PD) were done using one-tailed t test. P values <.05 were considered statistically significant for each comparison.

### 3 | RESULTS

#### 3.1 | Identification of relevant differentially expressed genes in T2D

Each selected microarray study was analysed using R/Bioconductor to find the DEGs comparing islets from people with/out diabetes

---

**Table 1** Characteristics of selected microarray for analysis

| Study information | Patient information | Characteristics | ND | PD | DT2 |
|-------------------|---------------------|----------------|-----|----|-----|
| GSE25724 (GEO)    | Gender              | 4/3            | 7   | —  | 3/3 |
| (Reference 16)    | Age                 | 58 ± 17        | 71 ± 9 |
|                   | BMI                 | 24.8 ± 2.5     | 26.0 ± 2.2 |
| E-CBIL-20 (ArrayExpress) | Gender | 2/5            | 0/5  |
| (Reference 18)    | Age                 | 48 ± 13 (n = 6)| 47 ± 8 (n = 4) |
| GSE38642 (GEO)    | Gender              | 25/29          | 4/5  |
| (Reference 13-15) | Age                 | 59 ± 9         | 57 ± 4 |
|                   | BMI                 | 25.9 ± 3.5     | 28.5 ± 4.5 |
| GSE20966 (GEO)    | Gender              | 6/4            | 7/3  |
| (Reference 17)    | Age                 | 60 ± 5         | 67 ± 7 |
|                   | BMI                 | 30.5 ± 6.5     | 30.9 ± 6.2 (n = 9) |
| GSE76894 (GEO)    | Gender              | 38/46          | 13/6 |
| (Reference 19)    | Age                 | 60 ± 16        | 72 ± 7 |
|                   | BMI                 | 25.8 ± 4.2 (n = 83) | 26.5 ± 3.6 |
| GSE50397 (GEO)    | Gender              | 33/18          | 9/6  |
| (Reference 10-12) | Age                 | 56 ± 12        | 61 ± 11 |
|                   | BMI                 | 25.6 ± 2.2     | 29.8 ± 3.1 |
| GSE76895 (GEO)    | Gender              | 16/16          | 23/13|
| (Reference 19)    | Age                 | 60 ± 14        | 66 ± 12 |
|                   | BMI                 | 24.9 ± 3.4     | 25.8 ± 5.0 |

Note: All the studies were performed on human pancreatic islets obtained from cadaverous donors except for the GSE76895 study that used material from pancreatectomized patients while the study GSE20966 was performed specifically on β-cell enriched tissue. Genders are expressed as males/females. Ages are expressed as mean ± SD (SD) in years. BMI are expressed as mean ± SD in Kg/m².
Following this procedure, we found a total of 756, 576, 492, 381, 246, 196 and 78 DEGs in studies GSE76894, GSE25724, GSE20966, E-CBIL-20, GSE76895, GSE38642 and GSE50397, respectively. Each study showed different proportions of down- or up-regulated genes (Figure 1A). Supplementary Table 1 shows all genes (and all probes corresponding to each gene) that have been identified as differently expressed in each study.

In order to identify transcriptional alterations associated with metabolic dysfunctions in people with T2D, we selected those DEGs that varied consistently in the same direction (down- or up-regulated) in at least three studies. Based on this criterion, we identified 55 relevant DEGs: 36 down-regulated and 19 up-regulated as shown in Figure 1B and Table 2. Whereas the 19 upregulated DEGs were simultaneously found in only three studies, 25 of those downregulated were commonly found in three studies, seven in four studies and four in five studies.

Hierarchical clustering analysis showed a separation among the studies and identified two main groups: E-CBIL-20 and GSE25724, GSE38642, GSE50397, GSE76894, GSE20966, GSE76895 (Figure 1B). As a result of meta-analysis, 87% of DEGs (48 out of 55) showed a significant alteration by integrating all studies. Notably, 9 of these DEGs have not been previously reported as dysregulated in T2D (Table 2, highlighted in grey); the corresponding *P*-values are also shown in that Table.

### 3.2 Functional enrichment analysis of relevant DEGs in T2D

Attempting to find a functional association between the 55 relevant DEGs in T2D, we performed an analysis of the pathway and the protein-protein interaction network. As shown in Figure 2A, we found that the most significantly enriched pathways were directly related to metabolism and/or development of β-cells.

The protein-protein interaction network represented in Figure 2B showed that among the upregulated DEGs, IL7R and IL6, were involved in the most interactions (12 and 6, respectively), thereby becoming key nodes in the network. Among the downregulated DEGs, NR0B1, SCD and PFKFB2 were involved in four interactions each. A principal node in the network was UBC which encodes Polyubiquitin-C, which in turn participates in protein recycling, interacting primarily with several downregulated DEGs.

**FIGURE 1** Identification of relevant DEGs in T2D/PD. A, Bar plot representing the number of DEGs (down- or up-regulated genes) in islets from ND vs T2D people of each selected microarray study. B, HeatMap plot of relevant DEGs in T2D and Hierarchical clustering analysis: each row represents DEGs that vary consistently in the same way (down- or up-regulated) in at least 3 studies, and each column represents a selected microarray study. Blue indicates down-regulated genes, red indicates up-regulated genes, and white indicates unaltered genes. The colour intensity is proportional to the fold-change. Grey indicates relevant DEGs that present *P*-value of meta-analysis < .05. Black indicates relevant DEGs in islets from ND vs T2D people that also present altered expression in people with PD.
| N° | Gene symbol | Entrez ID | Full name                                      | Hits | Average FC | P-value meta-analysis |
|----|-------------|-----------|------------------------------------------------|------|------------|-----------------------|
|    |             |           | **Downregulated DEGs**                          |      |            |                       |
| 1  | CHL1        | 10752     | Cell adhesion molecule L1 like                  | 5    | −2.78      | <.001                 |
| 2  | SLC2A2      | 6514      | Solute carrier family 2 member 2               | 5    | −2.74      | <.001                 |
| 3  | PPP1R1A     | 5502      | Protein phosphatase 1 regulatory inhibitor subunit 1A | 5    | −1.98      | <.001                 |
| 4  | ARG2        | 384       | Arginase 2                                     | 5    | −1.83      | <.001                 |
| 5  | PFKFB2      | 5208      | 6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 2 | 4    | −2.06      | <.001                 |
| 6  | GLRA1       | 2741      | Glycine receptor alpha 1                       | 4    | −2.01      | .01                   |
| 7  | RASGRP1     | 10125     | RAS guanyl releasing protein 1                  | 4    | −1.88      | .003                  |
| 8  | FFAR4       | 338557    | Free fatty acid receptor 4                     | 4    | −1.81      | <.001                 |
| 9  | PPM1E       | 22843     | Protein phosphatase, Mg2+/Mn2+ dependent 1E    | 4    | −1.79      | <.001                 |
| 10 | CAPN13      | 92291     | Calpain 13                                     | 4    | −1.64      | <.001                 |
| 11 | IAPP        | 3375      | Islet amyloid polypeptide                      | 4    | −1.59      | <.001                 |
| 12 | MYCN        | 4613      | MYCN proto-oncogene, bHLH transcription factor | 4    | −1.59      | <.001                 |
| 13 | LINC01933   | 101927115 | Long intergenic non-protein coding RNA 1933    | 3    | −2.15      | <.001                 |
| 14 | HHTL        | 57467     | Hedgehog acyltransferase like                  | 3    | −2.14      | <.001                 |
| 15 | PLA1A       | 51365     | Phospholipase A1 member A                      | 3    | −2.01      | <.001                 |
| 16 | EDN3        | 1908      | Endothelin 3                                   | 3    | −2.00      | .045                  |
| 17 | LOC101929550| 101929550 | ncRNA uncharacterized                          | 3    | −1.95      | .003                  |
| 18 | TMEM37      | 140738    | Transmembrane protein 37                       | 3    | −1.91      | <.001                 |
| 19 | HS6ST2      | 90161     | Heparan sulfate 6-O-sulfotransferase 2         | 3    | −1.90      | <.001                 |
| 20 | ABCC8       | 6833      | ATP binding cassette subfamily C member 8      | 3    | −1.89      | .003                  |
| 21 | RASGRF1     | 5923      | Ras protein specific guanine nucleotide releasing factor 1 | 3    | −1.83      | <.001                 |
| 22 | GLP1R       | 2740      | Glucagon like peptide 1 receptor               | 3    | −1.81      | .01                   |
| 23 | TMED6       | 146456    | Transmembrane p24 trafficking protein 6         | 3    | −1.80      | <.001                 |
| 24 | TAGLN3      | 29114     | Transgelin 3                                   | 3    | −1.79      | <.001                 |
| 25 | SCD         | 6319      | Stearoyl-CoA desaturase                        | 3    | −1.74      | .004                  |
| 26 | NROB1       | 190       | Nuclear receptor subfamily O group B member 1  | 3    | −1.68      | .008                  |
| 27 | HADH        | 3033      | Hydroxyacyl-CoA dehydrogenase                  | 3    | −1.65      | <.001                 |
| 28 | G6PC2       | 57818     | Glucose-6-phosphatase catalytic subunit 2      | 3    | −1.64      | <.001                 |
| 29 | PLCB4       | 5332      | Phospholipase C beta 4                         | 3    | −1.64      | .003                  |
| 30 | RBP4        | 5950      | Retinol binding protein 4                      | 3    | −1.62      | .002                  |
| 31 | ELMO1       | 9844      | Engulfment and cell motility 1                 | 3    | −1.62      | .129                  |
| 32 | PLCLD3      | 345557    | Phosphatidylinositol specific phospholipase C X domain containing 3 | 3    | −1.62      | <.001                 |
| 33 | ELAVL4      | 1996      | ELAV like RNA binding protein 4                | 3    | −1.61      | <.001                 |
| 34 | VAT1L       | 57687     | Vesicle amine transport 1 like                 | 3    | −1.59      | <.001                 |
| 35 | ROBO2       | 6092      | Roundabout guidance receptor 2                 | 3    | −1.59      | .012                  |
| 36 | SLC4A8      | 9498      | Solute carrier family 4 member 8               | 3    | −1.53      | .001                  |
| 37 | ALDOB       | 229       | Aldolase, fructose-bisphosphate B              | 3    | 6.07       | .041                  |
| 38 | PNLIPRP1    | 5407      | Pancreatic lipase related protein 1            | 3    | 2.72       | .01                   |
| 39 | GAD1        | 2571      | Glutamate decarboxylase 1                      | 3    | 2.10       | .023                  |
| 40 | MMP3        | 4314      | Matrix metallopeptidase 3                      | 3    | 2.09       | .071                  |
| 41 | IL6         | 3569      | Interleukin 6                                  | 3    | 2.02       | .321                  |
| 42 | AADAC       | 13        | Arylacetamide deacetylase                     | 3    | 1.99       | <.001                 |
| 43 | IL33        | 90865     | Interleukin 33                                 | 3    | 1.91       | <.001                 |
| 44 | CCDC69      | 26112     | Coiled-coil domain containing 69               | 3    | 1.85       | .028                  |

(Continues)
3.3 Analysis of relevant DEGs in pancreatic islets from PD cases

Based on the identification of the 48 potentially relevant DEGs which could play a role in the pathogenesis of T2D, we attempt thereafter to evaluate whether they could also be manifested at an early stage of the disease, that is, PD. For this purpose, we repeated the previous search using the same methodology described above but related to PD without success. However, we found that in the previous search related to T2D, two of the selected studies (GSE50397 and GSE76895) had included pancreatic islets from people with PD (Table 1); therefore, we used them for the pertinent analysis. This analysis demonstrated that 18 of the 48 selected DEGs (38%) showed altered expression in islets from people with PD in at least one of the two available studies (Figure 1B, column PD). These results indicate that some of the transcriptional alterations observed in people with T2D were already present in the PD stage. In fact, data from the GSE50397 study showed significantly reduced expressions of SLC2A2, CHL1, GLRA1, PFKFB2, RASGRP1, CAPN13, TMED6, GLP1R, G6PC2, ROBO2, PLCXD3, RBP4, VATL1 and SLC4A8. Conversely, SV2B expression was significantly increased in islets from PD compared to ND people. Additionally, in the GSE76895 study, we found significantly reduced expressions of PLA1A, LINC01116 and TMED6 and significantly increased expression of CD44. Relative expressions of DEGs not previously reported as dysregulated in PD are shown in Figure 3. Supplementary Table 2 shows a statistic of differentially expressed genes identified in islets from people with PD in at least one of the two available studies.

4 DISCUSSION

Using the previously explained methodology and a slightly restrictive cutoff value (P value < 0.05 and FC > 1.5), we have currently identified 55 genes differentially expressed in islets from people with T2D that exhibited consistent transcriptional alterations among the different datasets. Then, we have performed the meta-analysis using combined data from different studies, thus attaining greater statistical power. As a result, 48 genes were confirmed as having statistically significant up- or down-modulation in T2D vs ND. This integrative approach allowed the identification of some transcriptional alterations among these 48 genes that not being reported as relevant in the original studies.

Despite RNA sequencing in recent years has become an important technology for transcriptomic analysis, at the time we performed the dataset search, most of the studies found corresponded to microarray assays. In the next few years however, new studies based on RNA sequencing would allow the identification of novel biomarkers, such as new alternative splicing isoforms, that cannot be identified by closed platforms like microarrays.

Since the studies selected were carried out under variable conditions (different operators, people from different populations, diverse islet extraction and analysis techniques), the signature found represents the more reproducible transcriptional changes. Thus, we assumed that this signature could play an effective active role in the pathogenesis of T2D.

As expected, the pathway analysis of these relevant DEGs showed several over-represented pathways related to carbohydrate metabolism and the regulation of β-cell development and gene expression. Specifically, we found that the FOXA2 and FOXA3 transcription factor network was the most enriched pathway. In this regard, Blodgett DM et al. have shown that several genes involved in early development are highly expressed in fetal islet cells, mainly those that are associated with inflammatory function.26 Further, FOXA proteins are expressed early in embryonic endoderm playing an important role in the regulation of gene expression in liver and pancreas and in the regulation of several pancreas-specific genes.27 These genes include Pdx-1, a transcription factor that plays a pivotal role in
FIGURE 2  Functional enrichment analysis of relevant DEGs in T2D. A, Visualization of pathway enrichment analysis of relevant DEGs in T2D. B, Protein–protein interaction network of relevant DEGs in T2D. Nodes represent the genes/proteins inputs as seeds (Green and red nodes), as well as protein added by the platform to obtain a minimum network (Grey nodes). Green indicates down-regulated genes and red indicates up-regulated genes. The colour intensity is proportional to the fold-change. Edges indicate interactions between the proteins. This network has 36 seeds, 64 nodes and 95 edges.
pancreas development and islet cell ontogeny, being a major regulator of β-cell identity and function.  

Thirty-six of the 48 (75%) relevant DEGs currently described have been previously reported as being dysregulated in islets from people with T2D in some of the original microarray studies (SLC2A2, CHL1, PPP1R1A, ARG2, GLRA1, FFAR4, PPM1E, CAPN13, HHATL, EDN3, ABCC8, RASGRP1, TAGLN3, TMEM37, GLP1R, SCD, HADH, G6PC2, PLCB4, PLCXD3, ELAVL4, ALDOB, CD44, TMED6, NR0B1, RBP4, VAT1L, SLC4A8, IL7R, MYCN, PLA1A, H565T2, PFKFB2, IAPP, GAD1). Another three genes (6%) were also reported as dysregulated in pancreatic islets from T2D in other experimental studies (IL33, NNMT, SV2B). Altogether, their previous reported identification lends validity support to the methodology currently employed and the results obtained.

On the other hand, as far as we know, 9 genes (19%) of the signature have not been previously reported as dysregulated in islets from people with T2D, namely, LINC01933, LOC101929550, ROBO2, PNLIPRP1, AADAC, CCDC69, TPDSL1, ITIH4 and LINC01116.

**FIGURE 3** Relative expression of novel relevant DEGs in PD. Expression levels were extracted from the GSE50397 or GSE76895 studies. Data are presented as the mean ± SE of the mean (SEM). Statistical analyses of two groups were done using one-tailed t test. *P < .05, **P < .01, ***P < .001
We discuss some of these genes, as well as other genes of the signature poorly discussed in literature. For that purpose, we will consider them separately according to their dysregulation (either down- or up-regulated) and the mechanisms associated with islet mass and function.

### 4.1 DEGs downregulated in islets from people with T2D

The protein encoded by ROBO2 gene is a transmembrane receptor for the slit homologue 2 protein and its presence becomes essential for endocrine cell type sorting and mature architecture in mice islets.\(^{29}\) It has also been shown that SLIT-ROBO signalling potentiates insulin secretion and is required for \(\beta\)-cell survival.\(^{30}\)

MYCN, a proto-oncogene that encodes a bHLH transcription factor, has been associated with \(\beta\)-cell mass expansion during pregnancy.\(^{31}\) Thus, its downregulation might be involved in the long-term impairments of the offspring.\(^{32}\)

Some other DEGs might be associated with the process of \(\beta\)-cell membrane depolarization which opens the voltage-gated calcium channel raising the cytoplasmic Ca\(^{2+}\) concentration that finally triggers exocytosis of insulin-containing granules.\(^{33}\) For example, the protein encoded by SLC4A8 is a solute carrier that mediates sodium- and carbonate-dependent chloride-HCO3-exchange, an important process for intracellular pH regulation\(^{34}\) which could control membrane polarization/depolarization process. Therefore, its down regulation may play an active role in the impaired secretion of insulin in T2D.

On the other hand, HS6ST2 (Heparan Sulfate 6-O-Sulfotransferase 2) is related to transferase activity and glycosaminoglycan metabolism. HS6st2 knockout mice show increased body weight, impaired glucose metabolism and insulin resistance.\(^{35}\)

Since all these genes are related to metabolic and physical processes that promote \(\beta\)-cell function and mass, their downregulation might play a critical role in the pathogenesis of T2D.

### 4.2 DEGs upregulated in islets from people with T2D

SV2B is one of the three homologous isoforms of synaptic vesicle protein 2, which participates in exocytosis process in a Ca\(^{2+}\)-dependent manner.\(^{36}\) Perhaps its upregulation is a compensatory response of \(\beta\)-cells to the higher hormone demand (insulin resistance) in people with T2D.\(^{37}\)

TPD52L1 gene which encodes a member of a family of proteins (DS2-like proteins) has been identified as a cell cycle-regulated protein whose impairment affects the cell-mitosis process.\(^{38}\) Immune system dysregulation and inflammation have been strongly associated with T2D.\(^{39,40}\) In our case, some of the DEGs identified in islets from people with T2D are related to immune response: ITIH4 and ILR7. It has been demonstrated that the protein encoded by ITIH4 gene was dramatically elevated in poorly controlled T1D patients.\(^{41}\) IL7R, is a key regulator of T lymphocyte development and homeostasis,\(^{42}\) associated with adipogenesis and insulin resistance.\(^{43,45}\) Moreover, based on the results of the protein-protein interaction analysis, IL7R was the molecule with the largest number of interactions, reflecting its crucial role in the regulation of biological systems implicated in T2D pathogenesis. Therefore, IL7R could be considered a novel therapeutic target. Altogether, this evidence shows that upregulation of the genes described strongly suggest that they might play an important role in the pathogenesis of development and progression of T2D.

Other DEGs are associated with varied molecular function: lipid metabolism (PLA1A, PNPLP1 and AADAC), transcription activity (NROB1), retinoic binding (RBP4). Besides, little is known about the molecular function of the rest of the identified genes: CCDC69 (coiled-coil domain containing 69), VATL1 (vesicle amine transport 1 like) and 3 non-coding RNAs (LOC101929550, LINCO1933 and LINCO11116). Therefore, future studies are necessary to demonstrate their potential role in the \(\beta\)-cell dysfunction dysregulation associated with T2D.

Throughout the identification of relevant DEGs in T2D islets, we aimed to assess whether these dysfunctional genes could be also identified at an early stage of this disease: i.e. prediabetes (PD). Data from GSE50397 and GSE76895 studies - which include analysis of prediabetic pancreatic islets - showed that 2 (SV2B and CD44) of the 13 upregulated genes (15%) and 16 (SLC2A2, CHL1, GLRA1, PKF/K2R, RASGRP1, CAPN13, TMED6, GLPIR1, G6PC2, ROBO2, PLCXD3, RBP4, VATL1, PLA1A, LINCO1933 and SLC4A8) of the 35 downregulated genes (46%), were already present in people with PD.

Some of these DEGs found in PD have not been previously associated with this disease stage in pancreatic islets. Among them, ROBO2, VATL1, LINCO1933, SV2B, SLC4A8, PLA1A and RBP4 are discussed above. Otherwise, CD44, PKF/K2R, GLPIR1 and CAPN13 have been reported to be altered in islets from patients with T2D, but not in islets from people with PD. Since they were not reported in PD islets but their effects on \(\beta\)-cell mass and function were already reported, we will not discuss them further.\(^{19,46-48}\) As for TMED6, it must be stressed that it was down-regulated in both studies performed with islets from people with PD.

Altogether, the results above described suggest that the transcriptional alterations shown in all these genes could have an important role in the development of PD and also perhaps in its progression to T2D.

In summary, we have identified several differentially expressed genes, not previously reported in islets from people with T2D that could potentially play a relevant role in the pathogenesis of the disease. Some of them were also identified in islets from people with an early stage of the disease (PD). It remains to be demonstrated whether their early expression in people with PD can be used as a marker to identify people with higher risk of its fast progression to T2D. Anyhow, these results provide new evidence to interpret T2D pathogenesis and the transition from PD to T2D. Further population studies are necessary to validate the latter hypothesis and its potential use for the development of new strategies to improve the prevention, diagnosis/prognosis and treatment of T2D.
ACKNOWLEDGEMENTS
The authors thank Mrs. Susan Hale Rogers for careful manuscript edition/correction. LEF, JIG, MCA and BM are members of the Research Career of CONICET while MVM is a postdoctoral fellow from CONICET. This study was partially supported by a grant (PUE 2017 N° 22920170100082CO) provided by the National Scientific and Technical Research Council.

CONFLICT OF INTEREST
The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS
Bárbara Maiztegui, Martín C. Abba, and Juan J. Gagliardino conceived and designed the study. Drs. María V. Mencucci, Bárbara Maiztegui, Martín C. Abba, Luis E. Flores, and Juan J. Gagliardino carried out the integrative transcriptomic analysis and the identification of differential expressed genes. All the authors contributed to develop and approve the final manuscript.

ORCID
Bárbara Maiztegui https://orcid.org/0000-0002-9871-943X

REFERENCES
1. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO consultation. Diabet Med. 1998;15(7):539-553.
2. Williams R, Van Gaal L, Lucioni C. Assessing the impact of complications on the costs of type II diabetes. Diabetologia. 2014;45(S1):S13-S17. https://doi.org/10.1007/s00125-012-2085-9.
3. Sereday M, Damiano M, Lapertosa S. Complicaciones crónicas en personas con diabetes mellitus tipo 2 de reciente diagnóstico. Endocrinol y Nutr. 2008;55:64-68. https://doi.org/10.1016/S1575-0922(08)70638-7.
4. ADA (American Diabetes Association). Standards of medical care in diabetes: classification and diagnosis of diabetes. Diabetes Care. 2018; 41(Supplement 1):S13-S27. https://doi.org/10.2337/dc18-S002.
5. Pan XR, Li GW, Hu YH, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and diabetes study. Diabetes Care. 1997;20:537-544. https://doi.org/10.2323/diacare.20.4.537.
6. Li R, Zhang P, Barker LE, Chowdhury FM, Zhang X. Cost-effectiveness of interventions to prevent and control diabetes mellitus: a systematic review. Diabetes Care. 2010;33:1872-1894. https://doi.org/10.2323/diacare.20.4.537.
7. Cuscieri S. Type 2 diabetes – An unresolved disease across centuries contributing to a public health emergency. Diabetes Metab Syndr Clin Res Rev. 2019;13(1):450-453. https://doi.org/10.1016/j.dsx.2018.11.010.
8. Stumvoll M, Goldstein BJ, Van Haften TW. Type 2 diabetes: principles of pathogenesis and therapy. Lancet. 2005;365:1333-1346. https://doi.org/10.1016/S0140-6736(05)1032-X.
9. Marchetti P, Bugliani M, De Tata V, Marselli L, Suleiman M. Pancreatic Beta cell identity in humans and the role of type 2 diabetes. Front Cell Dev Biol. 2017;5(May):1-8. https://doi.org/10.3389/fcelle.2017.00055.
10. Fadista J, Vikman P, Laakso EO, et al. Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. Proc Natl Acad Sci. 2014;111:13924-13929. https://doi.org/10.1073/pnas.1402665111.
11. Taneera J, Fadista J, Ahlqvist E, et al. Identification of novel genes for glucose metabolism based upon expression pattern in human islets and effect on insulin secretion and glycermia. Hum Mol Genet. 2014;24:1945-1955. https://doi.org/10.1093/hmg/ddu610.
12. Taneera J, Mohammed AK, Dhaiban S, et al. RORB and RORC associate with human islet dysfunction and inhibit insulin secretion in INS-1 cells. Islets. 2019;11:10-20. https://doi.org/10.1080/19382014.2019.1566684.
13. Taneera J, Lang S, Sharma A, et al. A systems genetics approach identifies genes and pathways for type 2 diabetes in human islets. Cell Metab. 2012;16:122-134. https://doi.org/10.1016/j.cmet.2012.06.006.
14. Taneera J, Fadista J, Ahlqvist E, et al. Expression profiling of cell cycle genes in human pancreatic islets with and without type 2 diabetes. Mol Cell Endocrinol. 2013;375:35-42. https://doi.org/10.1016/j.mce.2013.05.003.
15. Kanatsuna N, Taneera J, Vaziri-Sani F, et al. Autoimmunity against INS-IGF2 protein expressed in human pancreatic islets. J Biol Chem. 2013;288:29013-29023. https://doi.org/10.1074/jbc.M113.478222.
16. Domínguez V, Raimondi C, Somanath S, et al. Class II phosphoinositide 3-kinase regulates exocytosis of insulin granules in pancreatic β cells. J Biol Chem. 2011;286:4216-4225. https://doi.org/10.1074/jbc.M111.200295.
17. Marselli L, Thorne J, Dahiya S, et al. Gene expression profiles of beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. PLoS One. 2010;5:e11499. https://doi.org/10.1371/journal.pone.0011499.
18. Gunton JE, Kulkami RN, Yim SH, et al. Loss of ARNT/HIF1α mediates altered gene expression and pancreatic-islet dysfunction in human type 2 diabetes. Cell. 2005;122:337-349. https://doi.org/10.1016/j.cell.2005.05.027.
19. Solimena M, Schulte AM, Marselli L, et al. Systems biology of the IMIDIA biobank from organ donors and pancreatectomised patients defines a novel transcriptomic signature of islets from individuals with type 2 diabetes. Diabetologia. 2018;61:641-657. https://doi.org/10.1007/s00125-017-4500-3.
20. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics. 2010;26:2363-2367. https://doi.org/10.1093/bioinformatics/btq431.
21. Gautier L, Cope L, Bolstad BM, Irizarry RA. Affy - analysis of Affymetrix GeneChip data at the probe level. Bioinformatics. 2004;20:307-315. https://doi.org/10.1093/bioinformatics/btg405.
22. Silver JD, Ritchie ME, Smyth GK. Microarray background correction: maximum likelihood estimation for the normal-exponential convolution. Biostatistics. 2009;10:352-363. https://doi.org/10.1093/biostatistics/kxn042.
23. Smyth GK. Limma: linear models for microarray data BT - bioinformatics and computational biology solutions using R and bioconductor. Statistics for Biology and Health. New York, NY: Springer; 2005. https://doi.org/10.1007/0-387-29362-0_23.
24. Breuer K, Foroushani AK, Laird MR, et al. InnateDB: systems biology of innate immunity and beyond-recent updates and continuing curation. Nucleic Acids Res. 2013;41:D1228-D1233. https://doi.org/10.1093/nar/gks1147.
25. Xia J, Benner MJ. Hancock REW. NetworkAnalyst - integrative approaches for protein-protein interaction network analysis and visual exploration. Nucleic Acids Res. 2014;42:W167-W174. https://doi.org/10.1093/nar/gku443.
26. Blodgett DM, Nowosielska A, Afik S, et al. Novel observations from next-generation RNA sequencing of highly purified human adult and fetal islet cell subsets. Diabetes. 2015;64:3172-3181. https://doi.org/10.2337/db15-0039.
27. Friedman JR, Kaestner KH. The Foxa family of transcription factors in development and metabolism. Cell Mol Life Sci. 2006;63:2317-2328. https://doi.org/10.1007/s00018-006-6095-6.
28. McKinnon CM, Docherty K. Pancreatic duodenal homebox-1, PDX-1, a major regulator of beta cell identity and function. Diabetologia. 2001;44:1203-1214. https://doi.org/10.1007/s001250100628.

29. Adams MT, Gilbert JM, Hinojosa Paiz J, Bowman FM, Blum B. Endocrine cell type sorting and mature architecture in the islets of Langerhans require expression of roundabout receptors in β-cells. Sci Rep 2018;8(1):1-12. https://doi.org/10.1038/s41598-018-29118-x.

30. Yang YHC, Manning Fox JE, Zhang KL, MacDonald PE, Johnson JD. Intraislet SLIT-ROBO signaling is required for beta-cell survival and potentiates insulin secretion. Proc Natl Acad Sci. 2013;110(41):16480-16485. https://doi.org/10.1073/pnas.1214312110.

31. Horn S, Kirkegaard JS, Hoelper S, et al. Research resource: a dual proteomic approach identifies regulated islet proteins during β-cell mass expansion in vivo. Mol Endocrinol. 2016;30(1):133-143. https://doi.org/10.1210/me.2015-1208.

32. Obesity, impaired glucose tolerance (GAD) and diabetes in adolescence or early adulthood (Prevention CDC). National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. Department of Health and Human Services CDCaP, Editor. 2011.

33. Trefler AJ, Taraska JW. Regulation of insulin exocytosis by calcium-dependent protein kinase C in beta cells. Cell Calcium. 2017;67:1-10. https://doi.org/10.1016/j.ceca.2017.07.008.

34. Aalkjaer C, Boedtkjer E, Choi I, Lee S. Cation-coupled bicarbonate transporters. Compr Physiol. 2014;4(4):1605-1637. https://doi.org/10.1002/cphy.c130005.

35. Nagai N, Habuchi H, Sugaya N, et al. Involvement of heparan sulfate 6-O-sulfation in the regulation of energy metabolism and the alteration of thyroid hormone levels in male mice. Glycobiology. 2013;23(8):980-992. https://doi.org/10.1093/glycob/cwt037.

36. Iezzi M, Theander S, Janz R, Loze C, Wollheim CB. SV2A and SV2C are not vesicular Ca2+ transporters but control glucose-evoked granule recruitment. J Cell Sci. 2005;118(23):5647-5660. https://doi.org/10.1242/jcs.02658.

37. Weyer C, Bogardus C, Mott DM, Pratley RE. The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus. J Clin Invest. 1999;104:787-794. https://doi.org/10.1172/JCI7231.

38. Boutros R, Byrne JA. D53 (TPD52L1) is a cell cycle-regulated protein maximally expressed at the G2-M transition in breast cancer cells. Exp Cell Res. 2005;310:152-165. https://doi.org/10.1016/j.yexcr.2005.07.009.

39. Grant RW, Dixit VD. Mechanisms of disease: inflammasome activation and the development of type 2 diabetes. Front Immunol. 2013;4:4-50. https://doi.org/10.3389/fimmu.2013.00050.

40. Imai Y, Dobrian AD, Morris MA, Nadler JL. Islet inflammation: a unifying target for diabetes treatment? Trends Endocrinol Metab. 2013;24:351-360. https://doi.org/10.1016/j.tem.2013.01.007.

41. Gourgari E, Ma J, Playford MP, et al. Proteomic alterations of HDL in youth with type 1 diabetes and their associations with glycemic control: a case–control study. Cardiovasc Diabetol. 2019;18(1):43. https://doi.org/10.1186/s12933-019-0846-9.

42. Oliveira ML, Akkapeddi P, Ribeiro D, Melão A, Barata JT. IL-7R-mediated signaling in T-cell acute lymphoblastic leukemia: an update. Adv Biol Regul. 2019;71:88-96. https://doi.org/10.1016/j.jbior.2018.09.012.

43. Moreno-Viedma V, Amor M, Sarabi A, et al. Common dysregulated pathways in obese adipose tissue and atherosclerosis. Cardiovasc Diabetol. 2016;15(1):120. https://doi.org/10.1186/s12933-016-0441-2.

44. Lee M, Song SJ, Choi M-S, Yu R, Park T. IL-7 receptor deletion ameliorates diet-induced obesity and insulin resistance in mice. Diabetologia. 2015;58(10):2361-2370. https://doi.org/10.1007/s00125-015-3684-7.

45. Poitou C, Perret C, Mathieu F, et al. Bariatric surgery induces disruption in inflammatory signaling pathways mediated by immune cells in adipose tissue with type 2 diabetes. Proc Natl Acad Sci. 2012;109:7049-7054. https://doi.org/10.1073/pnas.1114513109.

46. Kodama K, Horikoshi M, Toda K, et al. Expression-based genomewide association study links the receptor CD44 in adipose tissue with insulin secretion in pancreatic β-cells. Proc Natl Acad Sci. 2012;109:7049-7054. https://doi.org/10.1073/pnas.1114513109.

47. Arden C, Hampson LJ, Huang GC, et al. A role for PFK-2/FBPase-2, as distinct from fructose 2,6-bisphosphate, in regulation of insulin secretion in diabetic and nondiabetic humans. Diabetologia. 2001;44:1203-1214. https://doi.org/10.1002/dmrr.3359.