

# **A functional genomic approach to identify reference genes for human pancreatic beta cell real-time quantitative RT-PCR analysis**

Maria Inês Alvelos <sup>1\*</sup>, Florian Szymczak <sup>1\*</sup>, Ângela Castela <sup>1</sup>, Sandra Marín-Cañas <sup>1</sup>, Bianca Marmontel de Souza <sup>1</sup>, Ioannis Gkantounas<sup>1</sup>, Maikel Colli<sup>1</sup>, Federica Fantuzzi <sup>1</sup>, Cristina Cosentino <sup>1</sup>, Mariana Igoillo-Esteve <sup>1</sup>, Lorella Marselli <sup>2</sup>, Piero Marchetti <sup>2</sup>, Miriam Cnop <sup>1,3</sup>, Décio L. Eizirik <sup>1,4,5</sup>.

<sup>1</sup> ULB Center for Diabetes Research, Medical Faculty, Université Libre de Bruxelles, Brussels (ULB) B-1070, Belgium

<sup>2</sup> Department of Clinical and Experimental Medicine, Islet Cell Laboratory, University of Pisa, Pisa, Italy

<sup>3</sup> Division of Endocrinology, Erasmus Hospital, Université Libre de Bruxelles, Brussels, Belgium

<sup>4</sup> Welbio, Medical Faculty, Université Libre de Bruxelles, Brussels (ULB) B-1070, Belgium

<sup>5</sup> Indiana Biosciences Research Institute, Indianapolis, IN, USA

\*The first two authors should be regarded as joint First Authors.

Address correspondence and reprint requests to:  
 Maria Inês Alvelos  
 ULB Center for Diabetic Research  
 Medical Faculty, Université Libre de Bruxelles (ULB)  
 Route de Lennik, 808 – CP618  
 B-1070 – Brussels – Belgium  
 Phone: +32 (0) 2 555 6238  
 Fax: +32 (0)2 555 6239  
 Email: mardeoli@ulb.ac.be

Present address:

Cristina Cosentino  
 Département des neurosciences fondamentales.  
 Université de Lausanne, Lausanne, Switzerland.

## Abstract

Exposure of human pancreatic beta cells to pro-inflammatory cytokines or metabolic stressors is used to model events related to type 1 and type 2 diabetes, respectively. Quantitative real-time PCR is commonly used to quantify changes in gene expression. The selection of the most adequate reference gene(s) for gene expression normalization is an important pre-requisite to obtain accurate and reliable results. There are no universally applicable reference genes, and the human beta cell expression of commonly used reference genes can be altered by different stressors. Here we aimed to identify the most stably expressed genes in human beta cells to normalize quantitative real-time PCR gene expression.

We used comprehensive RNA-sequencing data from the human pancreatic beta cell line EndoC- $\beta$ H1, human islets exposed to cytokines or the free fatty acid palmitate in order to identify the most stably expressed genes. Genes were filtered based on their level of significance (adjusted P-value >0.05), fold-change ( $|\text{fold-change}| < 1.5$ ) and a coefficient of variation <10%. Candidate reference genes were validated by quantitative real-time PCR in independent samples.

We identified a total of 264 genes stably expressed in EndoC- $\beta$ H1 cells and human islets following cytokine- or palmitate-induced stress, displaying a low coefficient of variation. Validation by quantitative real-time PCR of the top five genes *ARF1*, *CWC15*, *RAB7A*, *SIAH1* and *VAPA* corroborated their expression stability under most of the tested conditions. Further validation in independent samples indicated that the geometric mean of *ACTB* and *VAPA* expression can be used as a reliable normalizing factor in human beta cells.

## Keywords

Reference genes/ beta cells/ diabetes/ RNA-sequencing/ quantitative real-time PCR

## Introduction

The study of molecular mechanisms involved in pancreatic beta cell dysfunction and death in type 1 (T1D) and type 2 diabetes (T2D) often involves *in vitro* exposure of human beta cells to stressors that may be present *in vivo* in T1D and T2D (1-3). Exposure of beta cells to these stressors, including pro-inflammatory cytokines (as a model of T1D) or palmitate (as a model of metabolic stress in T2D), substantially alters their gene expression (2, 4).

Quantitative real-time polymerase chain reaction (qPCR) is a commonly used technique to measure mRNA transcript levels owing to its sensitivity, specificity and fast execution (5). The accurate quantification of the observed changes relies on the effective normalization to one or more reference gene(s), whose expression should not be altered by the experimental condition(s) under evaluation. A unique and universal reference gene has not yet been identified, and therefore gene expression normalization usually depends on genes classified as “housekeeping genes” which, due to their cellular indispensability, are assumed to have stable expression under different experimental conditions. Commonly used reference genes, such as *beta actin (ACTB)*, *beta-2-microglobulin (B2M)*, the *18 S ribosome small subunit (18S rRNA)* and *glyceraldehyde-3-phosphate dehydrogenase (GAPDH)*, are widely used as normalizers due to their robust expression (6, 7). However, their expression varies widely among conditions and cell types (8-11), including pancreatic islets (12, 13). An inadequate selection of reference genes – that are up- or down-regulated in parallel with the gene under study - could lead to the misinterpretation of qPCR results and obscure genuine changes.

There are few studies on the identification of reference genes in human beta cells, and they are mostly limited to show that the housekeeping gene does not change under the experimental conditions used (14). Rat islets of Langerhans have different expression levels of reference genes during the first 24 hours following isolation (13), and a study in rat insulin-secreting INS-1E cells confirmed that qPCR results are affected by the normalization method and selected reference genes (15).

Selection of the most suitable reference genes should be implemented through the validation of their expression stability in the cell or tissue type under the study. Most studies use bioinformatic tools such as BestKeeper (16), geNorm (17), NormFinder (18) or Global Pattern Recognition (19), which perform a mathematical evaluation of gene expression and rank candidates according to their stability. These tools define the least variable genes from a list of pre-selected candidate genes, but they are not suitable for *de novo* identification of genes with calibrating potential.

Against this background, a genome-wide analysis to identify stable and well-expressed genes in human islets and beta cells represents an essential tool for accurate normalization. To achieve this goal, we used high-depth RNA-sequencing data from the human beta cell line EndoC- $\beta$ H1 and human islets exposed to pro-inflammatory cytokines or palmitate. Genes were validated as putative reference genes by qPCR in EndoC- $\beta$ H1 cells, human islets and induced pluripotent stem cell (iPSC)-derived islets.

## Results

### *In silico* selection of new reference genes

The analysis of high-depth RNA-seq (around 200 million reads) data from insulin-producing EndoC- $\beta$ H1 cells exposed to the proinflammatory cytokines interleukin-1 $\beta$  (IL1 $\beta$ ) plus interferon- $\gamma$  (IL1 $\beta$  + IFN $\gamma$ ) for 48h, or interferon- $\alpha$  (IFN $\alpha$ ) for 2h, 8h, and 24h and human islets exposed to IL1 $\beta$  + IFN $\gamma$  for 48h or IFN $\alpha$  for 2h, 8h, and 18h, or the metabolic stressor palmitate for 48h (**Figure 1A**) (see references in Methods) indicated that  $17,874 \pm 3,751$  genes are not significantly modified, according to the criteria of an adjusted P-value  $>0.05$  and  $|\text{fold-change}| < 1.5$ . The gene expression level quantified as transcripts per million (TPM) was used to evaluate intra- and inter-experimental condition variability, and 264 genes display a coefficient of variation (CV)  $<10\%$ , of which 175 genes have a mean TPM between 10 and 100, and 89 a mean TPM between 100 and 1000 (**Figure 1B-D**) (**Supplementary Table 2**). Genes with mean TPM  $>1000$  (**Figure 1B, C, D**) have the highest mean CV. For this reason, we selected the candidate reference genes among the low ( $10 < \text{mean TPM} < 100$ ) and intermediate TPM ( $100 < \text{mean TPM} < 1000$ ) intervals. Gene Ontology (GO) enrichment analysis on this set of 264 genes revealed an enrichment in pathways associated with central cellular functions such as “intracellular transport”, “mRNA processing”, “RNA splicing”, and “protein catabolic processes” (adjusted P-value  $<0.05$ ) (**Figure 1E**). As a proof of concept, we selected five genes: *ADP-ribosylation factor 1* (*ARF1*), *spliceosome associated protein homolog* (*CWC15*), *member RAS oncogene family* (*RAB7A*), *E3 ubiquitin protein ligase 1* (*SIAH1*) and *vesicle-associated membrane protein-associated protein A* (*VAPA*) from the 264 candidates for subsequent validation. Except for *RAB7A* (20), none of the genes have been previously used as reference genes in human beta cells or other tissues. The mean CV for the selected 5 genes ranged from 7% to 8.4%, and they were more stably expressed in human beta cells as compared to *ACTB* and *GAPDH*, which have CVs of 17.9% and 23.4%, respectively (**Figure 1C, D**). Interestingly, *ARF1*, *CWC15*, *RAB7A*, *SIAH1*, *VAPA* genes have the same mean TPM



distribution and CV in beta cells from T1D and T2D donors and non-diabetic controls, reinforcing our observation that these genes are stably expressed, and may be suitable human beta cell reference genes (**Supplementary Figure 1A, B**).

### **Validation of gene expression stability by quantitative real time PCR (qPCR)**

In order to assess the validity of the RNA-seq findings, we quantified by qPCR the expression of the five selected potential reference genes (*ARF1*, *CWC15*, *RAB7A*, *SIAH1*, *VAPA*) and the commonly used (*ACTB* and *GAPDH*) reference genes in independent samples of EndoC-βH1 cells and human islets exposed to different stress conditions. To allow accurate comparison between samples, all of them were diluted to the same final cDNA concentration, as described in Methods. PCR amplification efficiency was >90% for all genes.

The exposure of EndoC-βH1 cells to IL1β + IFNγ or IL1β + IFNα did not modify the expression of the analyzed reference genes (**Figure 2A, B**). However, IFNα modified expression of several of them, including *ACTB* and *GAPDH* (**Figure 2C, D**). Although *ACTB* and *GAPDH* have higher CV values compared to the novel reference genes in the RNA-seq data (Figures 1C and D), qPCR of these genes indicated stable expression in most conditions (**Figure 2, 3, 4, Supplementary Figure 1**) except for IFNα exposure. A combination of reference genes can provide more accurate correction for mRNA loading (21). We selected *ACTB* and *VAPA* because they are amongst the most well-expressed genes, displaying quantification cycle (Cq) values between 17-28 and 20-28, respectively, do not change following exposure to different stresses (**Figure 1**) and, at least for *ACTB*, are already in use by many islet research laboratories. We calculated the geometric mean between *ACTB* and *VAPA* Cq values and included them for comparison in the figures.

We next tested the effect of thapsigargin (TG) and tunicamycin (TM), two commonly used chemical endoplasmic reticulum stressors. EndoC-βH1 cells exposed to TG for 24h or TM for 48h neither showed differences in Cq values of the five new reference genes nor on *ACTB* and *GAPDH* expression (**Supplementary Figure 1C, D**).

Due to the limited amount of human islet material available, we restricted our analysis to *ACTB* and *VAPA*. IL1β + IFNγ, IFNα or palmitate did not change expression of *ACTB* and *VAPA* genes in human islets (**Figure 3**).

iPSC-derived beta cells represent a new model to study pathophysiologic mechanisms in T1D and T2D (22-27). We analyzed candidate reference genes in iPSC-derived islets exposed to the pro-inflammatory cytokines IL1β + IFNγ and IFNα and palmitate (**Figure 4A, B, C**). None of

these stressors altered the expression of the genes, suggesting that they are suitable for qPCR data normalization in iPSC-derived islets.

### Validation of the necessity for additional reference genes

To validate the normalization strategies for genes of particular interest in the context of islet inflammation, we compared mRNA expression of *IRF1* (a key transcription factor downstream of IFN $\alpha$  signaling in beta cells (28-30)) before and after IRF1 knockdown in EndoC- $\beta$ H1 cells under basal conditions and following IFN $\alpha$  exposure. Normalization of IRF1 expression by *ACTB* alone failed to show *IRF1* silencing following transfection with *IRF1* siRNA (**Figure 5A**). Normalization by the geometric mean of *ACTB* and *VAPA* unveiled a significant *IRF1* inhibition (**Figure 5B**), which is probably due to the reduced CV in all four conditions analyzed (**Figure 5C**). Similarly, mRNA induction of *HLA-ABC* (a key component of islet antigen presentation in the context of T1D (31)) by IFN $\alpha$  (**Figure 5D, E, F**) was detected following normalization by the geometric mean of *VAPA* and *ACTB*, with several fold decreased CV (**Figure 5F**).

### Discussion

We presently combined different strategies to identify new reference genes for qPCR analysis in human pancreatic beta cells. We used RNA-seq data from beta cells or human islets exposed to different stresses and identified 264 top stably expressed genes. These genes are mostly related with cellular housekeeping functions, as indicated by the fact that they are enriched in GO biological pathways related to intracellular (protein) transport, mRNA metabolic processes etc (**Figure 1E**). Following qPCR validation in independent samples, we found that the 5 candidate reference genes *ARF1*, *CWC15*, *RAB7A*, *SLAH1* and *VAPA* fit the criteria of ‘suitable reference genes’ (21). These genes have stable transcript abundance under all biologic contexts tested and are, in general, comparable to the expression of target genes. Since the geometrical mean of two or more reference genes provides a more accurate correction (17), we selected *VAPA* and *ACTB* as a suitable compromise between a novel, highly stable mRNA, and a well-established but more variable reference mRNA.

Our data indicate that the geometric mean of *ACTB* and *VAPA* expression represents a valid normalization strategy, as these gene are stably expressed (present data) and their combined use reduces the CV of target genes such as IRF1 and HLA class I (**Figure 5**). qPCR guidelines suggest to use at least 3 reference genes (21), but taking into account the scarcity of human

islets, and the present data showing that it is possible to obtain reliable data using only two reference genes, we suggest that the geometrical mean of *ACTB* and *VAPA* provides a fair compromise.

*ACTB* encodes one of six different actin isoforms, and has been widely used as a ubiquitously expressed reference gene (32). Actin-B protein is highly conserved, and maintains cell structure, integrity, and motility (33, 34). Its rate of transcription is affected by mitogenic stimuli such as epidermal growth factor, transforming growth factor- $\beta$  (TGF- $\beta$ ) and platelet derived growth factor (35-37). Since human beta cells proliferate very little (38), this should not present a major issue when using *ACTB* as reference gene.

*VAPA* encodes a protein involved in the formation of endoplasmic reticulum contacts with other membranes, such as the Golgi complex and endosomes (39, 40). It participates in fundamental physiological processes such as vesicle trafficking, membrane fusion, protein complex assembly and cell motility (41, 42). The role of *VAPA* is consistent with housekeeping functions and its expression does not change under the presently tested experimental conditions. Of particular relevance, *VAPA* expression is not changed in beta cells from T1D patients or whole islets from T2D patients.

In conclusion, the present study identified a panel of genes that can be used as reference for qPCR studies in human beta cells (Supplementary Table 1). The geometrical mean of two of these, namely *ACTB* and *VAPA*, may provide a robust normalization tool in the study of human beta cells. It will be important to confirm in different experimental conditions that neither *ACTB* nor *VAPA* are changed. Should this be the case, the other presently identified reference genes may be tested as an alternative approach.

## Material and methods

### Culture of human cells, gene silencing and treatments

The human beta cell line EndoC- $\beta$ H1 was provided by Dr R. Scharfmann (Institut Cochin, Université Paris Descartes, Paris, France) (43). EndoC- $\beta$ H1 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 5.6 mmol/L glucose (Gibco, Thermo Fisher Scientific), 2% fatty acid-free bovine serum albumin (BSA) fraction V (Roche), 50  $\mu$ mol/L 2-mercaptoethanol (Sigma-Aldrich), 10 mmol/L nicotinamide (Calbiochem), 5.5  $\mu$ g/mL transferrin (Sigma-Aldrich), 6.7 ng/mL selenite (Sigma-Aldrich), 100 U/mL penicillin + 100  $\mu$ g/mL streptomycin (Lonza) in matrigel–fibronectin–coated plates (44).

Human islets were isolated from non-diabetic organ donors by collagenase digestion and density gradient purification and characterized as previously reported (45), with the approval

of the the local Ethical Committee in Pisa, Italy. After isolation, the islets were cultured in M199 culture medium (5.5 mmol/L glucose) and shipped to our laboratory. On arrival, the islets were cultured in Ham's F-10 medium containing 6.1 mmol/L glucose (Gibco, Thermo-Fisher Scientific), 10% fetal bovine serum (Gibco, Thermo-Fisher Scientific), 2 mmol/L GlutaMAX (Sigma-Aldrich), 50 mmol/L 3-isobutyl-1-methylxanthine (Sigma-Aldrich), 1% fatty acid-free BSA fraction V, 50 U/mL penicillin and 50 mg/mL streptomycin (44).

The iPSC lines HEL46.11 (46) and HEL115.6 (22) were derived from human neonatal foreskin and umbilical cord fibroblasts, respectively. The iPSC line Wolf2010-9 was kindly provided by Dr. Fumihiko Urano (Washington University School of Medicine, St. Louis, MO, USA) and it is derived from patients with Wolfram syndrome (47). iPSCs were cultured in E8 medium (Life Technologies) in matrigel-coated plates (Corning BV, Life Sciences). iPSCs were differentiated into pancreatic beta cells using a 7-step protocol as previously described (22, 46, 48, 49).

EndoC- $\beta$ H1 cells, dispersed human islet cells and iPSC-beta cells were exposed to human pro-inflammatory cytokines IL1 $\beta$  (50 U/mL; R&D Systems) and IFN $\gamma$  (1,000 U/mL; PeproTech) for 24h and/or 48 h, as described (2, 23). EndoC- $\beta$ H1 cells, dispersed pancreatic islets and iPSC-beta cells were exposed to human IFN $\alpha$  (2000 U/mL; PeproTech) alone for 2h/8h/24h, 2h/8h/18h and 24-48h, respectively (23, 50, 51). EndoC- $\beta$ H1 cells were also exposed to a combination of IFN $\alpha$  and IL1 $\beta$  (50 U/mL; R&D Systems) for 48h (50, 51). These conditions are based on previously published time course and dose-response experiments (44, 50).

Dispersed human islets and iPSC-beta cells were exposed to 0.5 mmol/L palmitate (Sigma-Aldrich) for 24 h and 48h, respectively (1, 4, 52).

EndoC- $\beta$ H1 cells were exposed to thapsigargin (1  $\mu$ M; Millipore-Sigma) and tunicamycin (5  $\mu$ g/mL; Sigma-Aldrich) for 24h and 48h, respectively (53, 54).

## RNA-sequencing

The RNA-seq experiments of EndoC- $\beta$ H1 cells, dispersed human islets, and islets from T2D patients were previously published by our group (30, 45, 55). RNA-seq data of FACS-purified beta cells from T1D patients was downloaded via the Gene Expression Omnibus (GEO) repository (56). All RNA-seq data were re-analyzed using our own and up-dated pipeline. Initial quality control of reads was assessed using FastQC (version 0.11.5; FastQC: A quality control tool for high throughput sequence data [Online]. Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and gene expression was

quantified (in TPM) using Salmon v1.3.0(57) with extra parameters “--seqBias --gcBias --validateMappings”. The reference genome (GENCODE gene annotation release 31(58)) was indexed using default parameters. Differentially expressed genes were assessed using DESeq2 version 1.28.1 (59). The Generalized Linear Model was fitted with the formula “design  $\equiv$  pairing + condition” to account for the pair-wise experimental design (control and treatment) whenever possible. Genes were considered differently expressed if they passed a threshold of adjusted P-value  $<0.05$  (Benjamini-Hochberg correction) and  $|\text{fold-change}| > 1.5$ .

### **Identification and selection of reference genes**

From the DESeq2 analysis of RNA-seq datasets, genes not differentially expressed between paired conditions were selected. TPM values were used to evaluate intra- and inter-experimental variability by calculating the coefficient of variation (CV) of each gene for each group and condition. The CV is defined by the ratio of the standard deviation of a gene TPM to the arithmetic mean of the same gene. Next, we calculated the arithmetic mean of CV between different conditions. Genes with a mean CV  $<0.10$  (10%) were selected for further analysis. TPM values from the different conditions were averaged and the mean CV was evaluated within three subgroups of expression, namely genes with TPM between 10-100, 100-1000 and 1000-10000.

### **Functional annotation**

Functional enrichment was performed in R using clusterProfiler (60) and enrichplot (61) packages for Gene Ontology. The genes identified by RNA-seq in EndoC- $\beta$ H1 cells and human islets with TPM  $>0.5$  under control condition or after treatment were considered as expressed and used as background. An adjusted P-value  $<0.05$  (Benjamini-Hochberg correction) was considered statistically significant.

### **mRNA extraction, cDNA synthesis and quantification by qPCR**

Poly(A)<sup>+</sup> mRNA was isolated using Dynabeads mRNA Direct kit (Invitrogen), following the manufacturer’s protocol. mRNA was reverse transcribed using the Reverse Transcriptase Core kit (Eurogentec). In order to reduce variability in mRNA input, cDNA was quantified using the NanoDrop spectrophotometer (NanoDrop ND-1000; Thermo Fisher Scientific). Samples were then diluted to the same concentration of the least-concentrated sample (EndoC- $\beta$ H1 cells were diluted to 800 ng/uL; dispersed human islets to 500 ng/uL; iPSC-beta cells to 300 ng/uL).

Primers were designed using Primer-BLAST software (62), using the following criteria: (I) primers were designed across exon-exon junction whenever possible; (II) PCR amplicon size of 80 -120 base pairs to minimize unwanted effects on the amplification efficiency; (III) primer T<sub>m</sub> (melting temperature) was 56-60°C (with an annealing temperature of approximately 58°C); (IV) primer GC content was 40-65%, with no dimer or hairpin secondary structures. The qPCR amplification was done using IQ SYBR Green Supermix (Bio-Rad) using the CFX Connect Real-Time PCR Detection System (Bio-Rad). The amplification efficiency of each primer pair was evaluated using a standard curve (63), generated from six dilutions of cDNA (10<sup>7</sup> to 10<sup>2</sup> mRNA copies per microliter (copies per µL)). The target gene concentration was expressed as copies per µL (63). The C<sub>q</sub> values specify the number of amplification cycles needed to detect a signal. Because C<sub>q</sub> values are inversely correlated with the amount of target nucleotides, they were used to evaluate gene expression and extrapolate gene expression variability.

The cycling conditions were 95°C for 3 minutes (min), followed by 40 cycles of 95°C for 15 seconds (sec), and 58°C for 20 sec, followed by a final step of 95°C for 1min, 70°C for 5 sec, and 95°C for 50 sec. For each gene, the melting curve was analyzed to confirm amplification of a single PCR product. Primers are listed in **Supplementary table 1**.

### **Statistical analysis**

Significant differences between experimental conditions were determined by Student's paired t-test or by ANOVA followed by Bonferroni correction as indicated. P-values <0.05 were considered statistically significant. Statistical tools used for RNA-seq analysis are described above. Violin plots illustrate kernel probability density (i.e., the width of the shaded area represents the data's density plot). The horizontal line represents the median, boxes quartiles, whiskers most extreme data values, and dots independent experiments.

### **Data accessibility**

All raw RNA-sequencing data are accessible via NCBI Gene Expression Omnibus (GEO), access numbers GSE133218, GSE148058, GSE121863, GSE53949 and GSE159984.



## Figure legends

**Figure 1. Identification of new reference genes for human islets/beta cells.** (A) The approach for the identification of new reference genes in human pancreatic beta cells started with the analysis of RNA-sequencing data from EndoC- $\beta$ H1 cells exposed to IL1 $\beta$  + IFN $\gamma$  (48h, n=5) or IFN $\alpha$  (2h, 8h, and 24h, n=5 per time point); human islets exposed to IL1 $\beta$  + IFN $\gamma$  (48h, n=5) or IFN $\alpha$  (2h, 8h, and 18h, n=6 per time point); human islets exposed to palmitate (48h, n= 5); fluorescence-activated cell sorting (FACS)-purified beta cells from T1D patients (n=4), and islets from T2D patients (n=28). The quantification of expression of gene transcripts was conducted using Salmon v1.3.0 (Genome reference: GENCODE GRCh38.p13) (58). Transcript expression levels, measured in transcripts per million (TPM), were used to calculate the intra- and inter-sample stability, and their respective coefficient of variation (CV). TPM values were subdivided in three categories: 10<TPM<100; 100<TPM<1000; 1000<TPM<10,000. Genes with the lowest mean CV were selected for further validation by qRT-PCR. (B,C,D) Scatterplots of mean CV against mean expression values (TPM) after log2 transformation based on RNA-seq data, for genes with TPM values between 10 and 100 (B), between 100 and 1000 (C), and between 1000 and 10,000 (D). Each circle represents a gene with an adjusted P-value >0.05, |fold-change| <1.5, and blue circles represent the genes with CV <10%. Selected reference genes are shown in the scatterplots, along with the well-established reference genes *ACTB* and *GAPDH*. (E) Significantly enriched Gene Ontology (GO) terms by gene enrichment analysis of the 264 genes with adjusted P-value >0.05, |foldchange| <1.5 and CV <10%. The gene ratio represents the ratio between the genes in the GO term that overlaps with the query gene list (the short list of candidate genes with CV <10%). The top 20 enriched GO terms are represented. Adjusted P-value from hypergeometric distribution (Benjamini-Hochberg).

**Figure 2. Validation of reference genes by qPCR in EndoC- $\beta$ H1 cells.** Distribution of the quantification cycle (Cq) value of five new (*ARF1*, *CWC15*, *RAB7A*, *SIAH1*, *VAPA*) and two commonly used reference genes (*ACTB* and *GAPDH*) based on qRT-PCR, in EndoC- $\beta$ H1 cells treated with IL1 $\beta$  + IFN $\gamma$  for 48h (A), IL1 $\beta$  + IFN $\alpha$  for 48h (B), IFN $\alpha$  for 8h (C) and 24h (D). Geometric means of *ACTB* and *VAPA* Cq values were also calculated (GM *ACTB* *VAPA*). The shape of the violin plots reflects the distribution of the data, and the width of curve corresponds to the frequency of values in each region. The integrated boxplots show the 25<sup>th</sup> and 75<sup>th</sup> percentiles, and the horizontal line represents the median. Results from six to thirteen



independent experiments. \*P <0.05, \*\*P <0.01 *versus* non-treated (NT) (paired Student's t-test).

**Figure 3. Validation of reference genes by qPCR in dispersed human islet cells.**

Distribution of the quantification cycle (Cq) value of five new and two commonly used reference genes (*ACTB* and *GAPDH*) based on qRT-PCR, in human islet cells treated with IL1 $\beta$  + IFN $\gamma$  for 48h (A), palmitate for 48h (B), IFN $\alpha$  for 8h (C) and 24h (D). Geometric means of *ACTB* and *VAPA* Cq values were also calculated (GM *ACTB VAPA*). The violin plots' shape reflects the distribution of the data, and the width of curve corresponds to the frequency of values in each region. The integrated boxplots show the 25<sup>th</sup> and 75<sup>th</sup> percentiles, and the horizontal line represents the median. Results from four independent experiments. There is no statistical difference between the tested conditions (paired Student's t-test, condition *versus* non-treated (NT)).

**Figure 4. Validation of reference genes by qPCR in human iPSC-derived beta cells.**

Distribution of the quantification cycle (Cq) value of five new and two commonly used reference genes (*ACTB* and *GAPDH*) based on qPCR of human iPSC-derived beta cells treated with IL1 $\beta$  + IFN $\gamma$  and IFN $\alpha$  for 24h (A), IL1 $\beta$  + IFN $\gamma$  and IFN $\alpha$  for 48h (B), and the metabolic stressor palmitate for 48h (C). Geometric means of *ACTB* and *VAPA* Cq values were also calculated (GM *ACTB VAPA*). The violin plots' shape reflects the distribution of the data, and the width of curve corresponds to the frequency of values in each region. The integrated boxplots show the 25<sup>th</sup> and 75<sup>th</sup> percentiles, and the horizontal line represents the median. Results from four to six independent experiments. There is no statistical difference between the tested conditions (paired Student's t-test, condition *versus* non-treated (NT)).

**Figure 5. Quantification of *IRF1* and *HLA-ABC* mRNA expression in EndoC- $\beta$ H1 cells.**

EndoC- $\beta$ H1 cells were transfected with control siRNA (siCTL) or siRNA against *IRF1* (siIRF1), for 48h, and then treated with IFN $\alpha$  for 24h. *IRF1* mRNA expression was normalized to *ACTB* (A), and to the geometric mean of *ACTB* and *VAPA* (B). Normalization by the geometric mean of *ACTB* and *VAPA* decreases *IRF1* expression variability (lower CV) (C). EndoC- $\beta$ H1 cells were treated with IFN $\alpha$  for 24h and *HLA-ABC* expression was normalized to *ACTB* (D), and to the geometric mean of *ACTB* and *VAPA* (E). Normalization by the geometric mean of *ACTB* and *VAPA* decreases *HLA-ABC* expression variability (lower CV)

(F). The boxplots show the 25<sup>th</sup> and 75<sup>th</sup> percentiles, and the horizontal line represents the median. Results from four independent experiments. \*P < 0.05, \*\*P < 0.01 *versus* siCTL (paired Student's t-test).

**Supplementary Figure 1: Expression of new reference genes in beta cells from diabetic patients and validation by qPCR in EndoC-βH1 cells.**

The newly identified reference genes *ARF1*, *CWC15*, *RAB7A*, *SIAH1*, *VAPA* are expressed in FACS-purified beta cells isolated from T1D patients (A) and islets from T2D patients (B). Distribution of the quantification cycle (Cq) value of *ARF1*, *CWC15*, *RAB7A*, *SIAH1*, *VAPA*, *ACTB* and *GAPDH* based on qPCR in EndoC-βH1 cells treated with thapsigargin for 48h (C), and *CWC15*, *RAB7A*, *VAPA* and *ACTB* in EndoC-βH1 treated with tunicamycin for 24h (D). For all samples the geometric mean of *ACTB* and *VAPA* Cq values was calculated (GM *ACTB VAPA*). The shape of the violin plots reflects the distribution of the data, and the width of curve corresponds to the frequency of the values in each region. The integrated boxplots show the 25<sup>th</sup> and 75<sup>th</sup> percentiles, and the horizontal line represents the median. Results from six to eight independent experiments. There is no statistical difference between the tested conditions (paired Student's t-test, condition *versus* control (CTL) or non-treated (NT)).

**Supplementary Table 1: Sequence of qPCR primers and siRNAs.**

**Supplementary Table 2: Stably expressed genes in EndoC-βH1 cells and dispersed human islets under pro-inflammatory and metabolic stress conditions.** Coefficient of variance and transcripts per million (TPM) for 264 stably expressed genes with an adjusted P value > 0.05, |fold change < 1.5|.

**Funding information:** D.L.E. is funded by Welbio/FRFS (n° WELBIO-CR-2019C-04), Belgium, by the Dutch Diabetes Research Foundation (project Innovate2CureType1, DDRF; no. 2018.10.002) and by start-up funds provided by the Indiana Biosciences Research Institute; D.L.E. and M.C. are funded by the Brussels Region (INNOVIRIS BRIDGE grant DiaType); DLE, MC and PM are supported by the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement 115797 (INNODIA) and 945268 (INNODIA HARVEST). This Joint Undertaking receives support from the Union's Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations,

JDRF, and the Leona M. and Harry B. Helmsley Charitable Trust; M.C. is funded by the Fonds National de la Recherche Scientifique and the Walloon Region SPW-EERWin2Wal project BetaSource, Belgium.

## ACKNOWLEDGEMENTS

The authors are grateful to Isabelle Millard, Anyishaï Musuaya, Nathalie Pachera, Cai Ying, and Manon Depessemier (ULB Center for Diabetes Research) for providing excellent technical support.

## References

1. Lytrivi M, Ghaddar K, Lopes M, Rosengren V, Piron A, Yi X, et al. Combined transcriptome and proteome profiling of the pancreatic beta-cell response to palmitate unveils key pathways of beta-cell lipotoxicity. *BMC Genomics*. 2020;21(1):590.
2. Eizirik DL, Sammeth M, Bouckennooghe T, Bottu G, Sisino G, Igoillo-Esteve M, et al. The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. *PLoS Genet*. 2012;8(3):e1002552.
3. Eizirik DL, Pasquali L, Cnop M. Pancreatic beta-cells in type 1 and type 2 diabetes mellitus: different pathways to failure. *Nat Rev Endocrinol*. 2020.
4. Cnop M, Abdulkarim B, Bottu G, Cunha DA, Igoillo-Esteve M, Masini M, et al. RNA sequencing identifies dysregulation of the human pancreatic islet transcriptome by the saturated fatty acid palmitate. *Diabetes*. 2014;63(6):1978-93.
5. VanGuilder HD, Vrana KE, Freeman WM. Twenty-five years of quantitative PCR for gene expression analysis. *Biotechniques*. 2008;44(5):619-26.
6. de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, et al. Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. *Lab Invest*. 2005;85(1):154-9.
7. Lee PD, Sladek R, Greenwood CM, Hudson TJ. Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Res*. 2002;12(2):292-7.
8. Jo J, Choi S, Oh J, Lee SG, Choi SY, Kim KK, et al. Conventionally used reference genes are not outstanding for normalization of gene expression in human cancer research. *BMC Bioinformatics*. 2019;20(Suppl 10):245.
9. Sharan RN, Vaiphei ST, Nongrum S, Keppen J, Ksoo M. Consensus reference gene(s) for gene expression studies in human cancers: end of the tunnel visible? *Cell Oncol (Dordr)*. 2015;38(6):419-31.
10. Barber RD, Harmer DW, Coleman RA, Clark BJ. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. *Physiol Genomics*. 2005;21(3):389-95.
11. Derks NM, Muller M, Gaszner B, Tilburg-Ouwens DT, Roubos EW, Kozicz LT. Housekeeping genes revisited: different expressions depending on gender, brain area and stressor. *Neuroscience*. 2008;156(2):305-9.

12. Rodriguez-Mulero S, Montanya E. Selection of a suitable internal control gene for expression studies in pancreatic islet grafts. *Transplantation*. 2005;80(5):650-2.
13. Kosinová L, Cahová M, Fabryová E, Tycová I, Koblas T, Leontovyč I, et al. Unstable Expression of Commonly Used Reference Genes in Rat Pancreatic Islets Early after Isolation Affects Results of Gene Expression Studies. *PLoS One*. 2016;11(4):e0152664.
14. Moore F, Colli ML, Cnop M, Esteve MI, Cardozo AK, Cunha DA, et al. PTPN2, a candidate gene for type 1 diabetes, modulates interferon-gamma-induced pancreatic beta-cell apoptosis. *Diabetes*. 2009;58(6):1283-91.
15. Smidt K, Wogensens L, Brock B, Schmitz O, Rungby J. Real-time PCR: housekeeping genes in the INS-1E beta-cell line. *Horm Metab Res*. 2006;38(1):8-11.
16. Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett*. 2004;26(6):509-15.
17. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol*. 2002;3(7):RESEARCH0034.
18. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res*. 2004;64(15):5245-50.
19. Akilesh S, Shaffer DJ, Roopenian D. Customized molecular phenotyping by quantitative gene expression and pattern recognition analysis. *Genome Res*. 2003;13(7):1719-27.
20. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet*. 2013;29(10):569-74.
21. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem*. 2009;55(4):611-22.
22. Cosentino C, Toivonen S, Diaz Villamil E, Atta M, Ravanat JL, Demine S, et al. Pancreatic beta-cell tRNA hypomethylation and fragmentation link TRMT10A deficiency with diabetes. *Nucleic Acids Res*. 2018;46(19):10302-18.
23. Demine S, Schiavo AA, Marin-Canas S, Marchetti P, Cnop M, Eizirik DL. Pro-inflammatory cytokines induce cell death, inflammatory responses, and endoplasmic reticulum stress in human iPSC-derived beta cells. *Stem Cell Res Ther*. 2020;11(1):7.
24. De Franco E, Lytrivi M, Ibrahim H, Montaser H, Wakeling MN, Fantuzzi F, et al. YIPF5 mutations cause neonatal diabetes and microcephaly through endoplasmic reticulum stress. *J Clin Invest*. 2020;130(12):6338-53.
25. Lytrivi M, Senee V, Salpea P, Fantuzzi F, Philippi A, Abdulkarim B, et al. DNAJC3 deficiency induces beta-cell mitochondrial apoptosis and causes syndromic young-onset diabetes. *Eur J Endocrinol*. 2021.
26. Leite NC, Sintov E, Meissner TB, Brehm MA, Greiner DL, Harlan DM, et al. Modeling Type 1 Diabetes In Vitro Using Human Pluripotent Stem Cells. *Cell Rep*. 2020;32(2):107894.
27. Ellis C, Ramzy A, Kieffer TJ. Regenerative medicine and cell-based approaches to restore pancreatic function. *Nat Rev Gastroenterol Hepatol*. 2017;14(10):612-28.
28. Moore F, Naamane N, Colli ML, Bouckennooghe T, Ortis F, Gurzov EN, et al. STAT1 is a master regulator of pancreatic {beta}-cell apoptosis and islet inflammation. *J Biol Chem*. 2011;286(2):929-41.

29. Colli ML, Hill JLE, Marroqui L, Chaffey J, Dos Santos RS, Leete P, et al. PDL1 is expressed in the islets of people with type 1 diabetes and is up-regulated by interferons-alpha and-gamma via IRF1 induction. *EBioMedicine*. 2018;36:367-75.
30. Colli ML, Ramos-Rodriguez M, Nakayasu ES, Alvelos MI, Lopes M, Hill JLE, et al. An integrated multi-omics approach identifies the landscape of interferon-alpha-mediated responses of human pancreatic beta cells. *Nat Commun*. 2020;11(1):2584.
31. Richardson SJ, Rodriguez-Calvo T, Gerling IC, Mathews CE, Kaddis JS, Russell MA, et al. Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. *Diabetologia*. 2016;59(11):2448-58.
32. Rubenstein PA. The functional importance of multiple actin isoforms. *Bioessays*. 1990;12(7):309-15.
33. Pollard TD, Borisy GG. Cellular motility driven by assembly and disassembly of actin filaments. *Cell*. 2003;112(4):453-65.
34. Bunnell TM, Burbach BJ, Shimizu Y, Ervasti JM. beta-Actin specifically controls cell growth, migration, and the G-actin pool. *Mol Biol Cell*. 2011;22(21):4047-58.
35. Leof EB, Proper JA, Getz MJ, Moses HL. Transforming growth factor type beta regulation of actin mRNA. *J Cell Physiol*. 1986;127(1):83-8.
36. Keski-Oja J, Raghov R, Sawdey M, Loskutoff DJ, Postlethwaite AE, Kang AH, et al. Regulation of mRNAs for type-1 plasminogen activator inhibitor, fibronectin, and type I procollagen by transforming growth factor-beta. Divergent responses in lung fibroblasts and carcinoma cells. *J Biol Chem*. 1988;263(7):3111-5.
37. Elder PK, Schmidt LJ, Ono T, Getz MJ. Specific stimulation of actin gene transcription by epidermal growth factor and cycloheximide. *Proc Natl Acad Sci U S A*. 1984;81(23):7476-80.
38. Cnop M, Hughes SJ, Igoillo-Esteve M, Hoppa MB, Sayyed F, van de Laar L, et al. The long lifespan and low turnover of human islet beta cells estimated by mathematical modelling of lipofuscin accumulation. *Diabetologia*. 2010;53(2):321-30.
39. Alpy F, Rousseau A, Schwab Y, Legueux F, Stoll I, Wendling C, et al. STARD3 or STARD3NL and VAP form a novel molecular tether between late endosomes and the ER. *J Cell Sci*. 2013;126(Pt 23):5500-12.
40. Peretti D, Dahan N, Shimoni E, Hirschberg K, Lev S. Coordinated lipid transfer between the endoplasmic reticulum and the Golgi complex requires the VAP proteins and is essential for Golgi-mediated transport. *Mol Biol Cell*. 2008;19(9):3871-84.
41. Phillips MJ, Voeltz GK. Structure and function of ER membrane contact sites with other organelles. *Nat Rev Mol Cell Biol*. 2016;17(2):69-82.
42. Prinz WA. Bridging the gap: membrane contact sites in signaling, metabolism, and organelle dynamics. *J Cell Biol*. 2014;205(6):759-69.
43. Ravassard P, Hazhouz Y, Pechberty S, Bricout-Neveu E, Armanet M, Czernichow P, et al. A genetically engineered human pancreatic beta cell line exhibiting glucose-inducible insulin secretion. *J Clin Invest*. 2011;121(9):3589-97.
44. Brozzi F, Nardelli TR, Lopes M, Millard I, Barthson J, Igoillo-Esteve M, et al. Cytokines induce endoplasmic reticulum stress in human, rat and mouse beta cells via different mechanisms. *Diabetologia*. 2015;58(10):2307-16.
45. Marselli L, Piron A, Suleiman M, Colli ML, Yi X, Khamis A, et al. Persistent or Transient Human beta Cell Dysfunction Induced by Metabolic Stress: Specific Signatures and Shared Gene Expression with Type 2 Diabetes. *Cell Rep*. 2020;33(9):108466.
46. Saarimäki-Vire J, Balboa D, Russell MA, Saarikettu J, Kinnunen M, Keskitalo S, et al. An Activating STAT3 Mutation Causes Neonatal Diabetes through Premature Induction of Pancreatic Differentiation. *Cell Rep*. 2017;19(2):281-94.

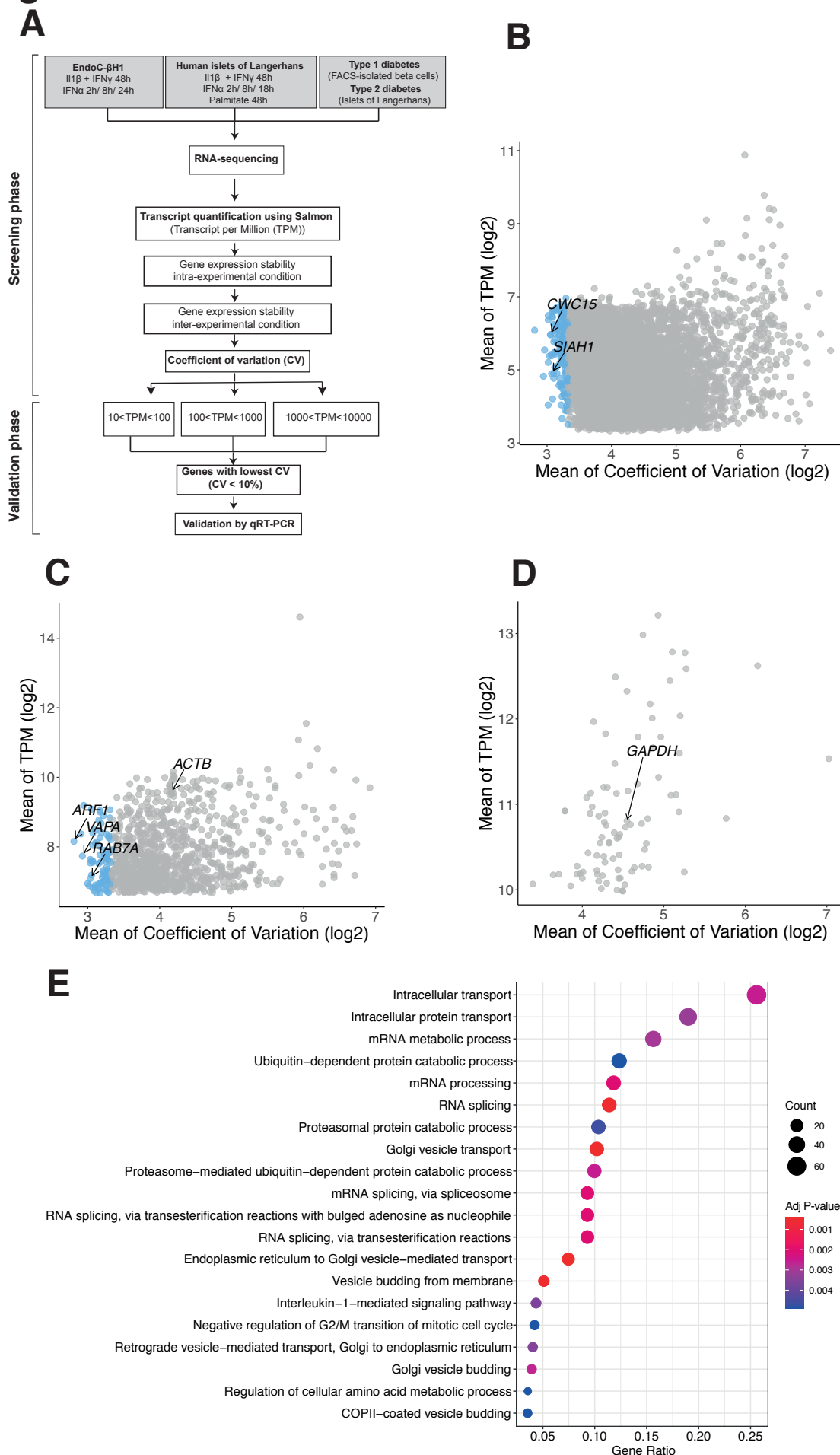


47. Lu S, Kanekura K, Hara T, Mahadevan J, Spears LD, Oslowski CM, et al. A calcium-dependent protease as a potential therapeutic target for Wolfram syndrome. *Proc Natl Acad Sci U S A*. 2014;111(49):E5292-301.
48. Pagliuca FW, Millman JR, Gurtler M, Segel M, Van Dervort A, Ryu JH, et al. Generation of functional human pancreatic beta cells in vitro. *Cell*. 2014;159(2):428-39.
49. Rezania A, Bruin JE, Arora P, Rubin A, Batushansky I, Asadi A, et al. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat Biotechnol*. 2014;32(11):1121-33.
50. Marroqui L, Dos Santos RS, Op de Beeck A, Coomans de Brachene A, Marselli L, Marchetti P, et al. Interferon-alpha mediates human beta cell HLA class I overexpression, endoplasmic reticulum stress and apoptosis, three hallmarks of early human type 1 diabetes. *Diabetologia*. 2017;60(4):656-67.
51. Coomans de Brachene A, Castela A, Op de Beeck A, Mirmira RG, Marselli L, Marchetti P, et al. Preclinical evaluation of tyrosine kinase 2 inhibitors for human beta-cell protection in type 1 diabetes. *Diabetes Obes Metab*. 2020;22(10):1827-36.
52. Bugliani M, Masini M, Liechti R, Marselli L, Xenarios I, Boggi U, et al. The direct effects of tacrolimus and cyclosporin A on isolated human islets: A functional, survival and gene expression study. *Islets*. 2009;1(2):106-10.
53. Paula FMM, Leite NC, Borck PC, Freitas-Dias R, Cnop M, Chacon-Mikahil MPT, et al. Exercise training protects human and rodent beta cells against endoplasmic reticulum stress and apoptosis. *FASEB J*. 2018;32(3):1524-36.
54. Tsonkova VG, Sand FW, Wolf XA, Grunnet LG, Kirstine Ringgaard A, Ingvorsen C, et al. The EndoC-betaH1 cell line is a valid model of human beta cells and applicable for screenings to identify novel drug target candidates. *Mol Metab*. 2018;8:144-57.
55. Ramos-Rodríguez M, Raurell-Vila H, Colli ML, Alvelos MI, Subirana-Granés M, Juan-Mateu J, et al. The impact of proinflammatory cytokines on the beta-cell regulatory landscape provides insights into the genetics of type 1 diabetes. *Nat Genet*. 2019;51(11):1588-95.
56. Russell MA, Redick SD, Blodgett DM, Richardson SJ, Leete P, Krogvold L, et al. HLA Class II Antigen Processing and Presentation Pathway Components Demonstrated by Transcriptome and Protein Analyses of Islet beta-Cells From Donors With Type 1 Diabetes. *Diabetes*. 2019;68(5):988-1001.
57. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14(4):417-9.
58. Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res*. 2019;47(D1):D766-D73.
59. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
60. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284-7.
61. Yu G. enrichplot: Visualization of Functional Enrichment Result. R package version 1.8.1. 2020.
62. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*. 2012;13:134.
63. Overbergh L, Valckx D, Waer M, Mathieu C. Quantification of murine cytokine mRNAs using real time quantitative reverse transcriptase PCR. *Cytokine*. 1999;11(4):305-12.

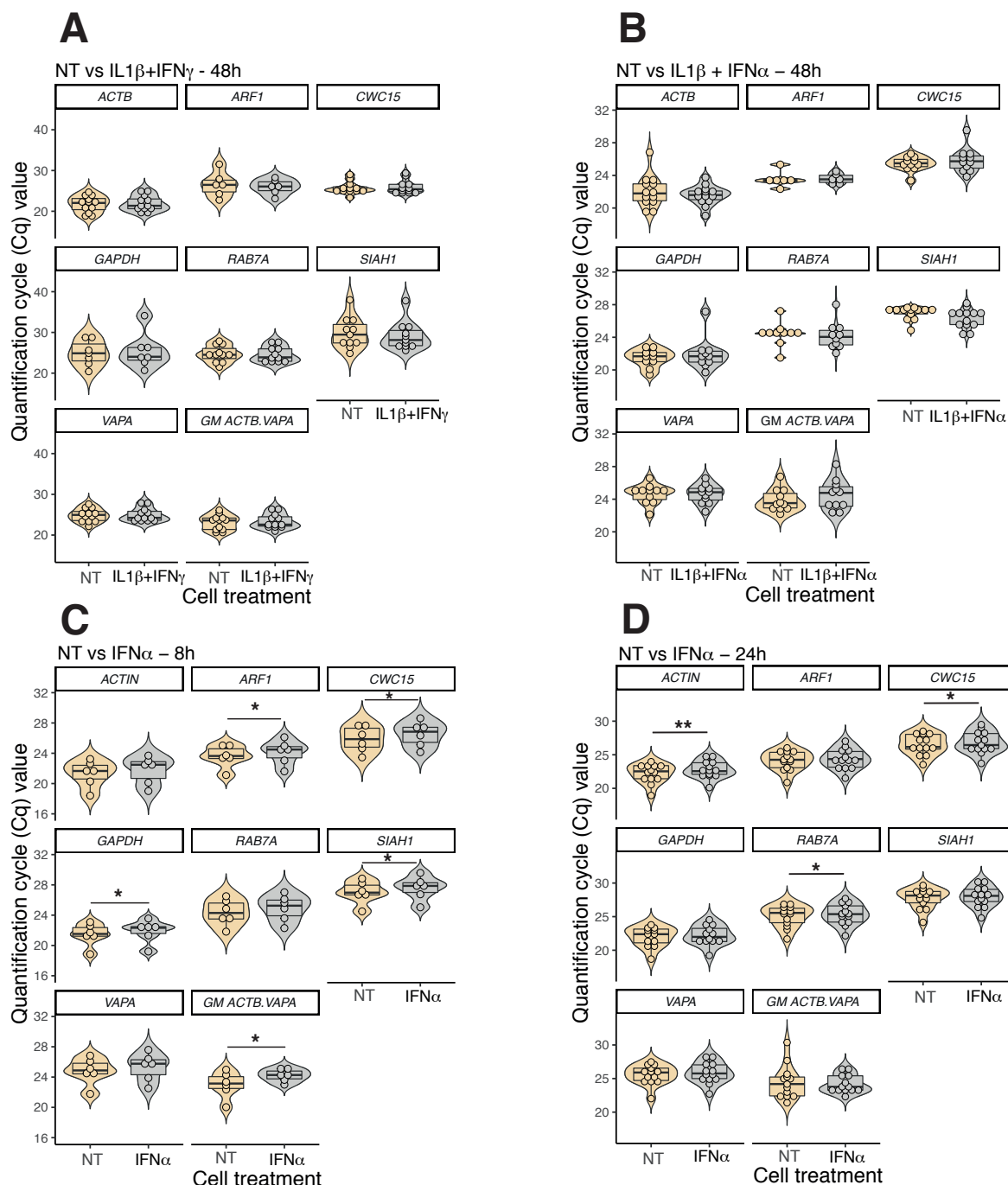




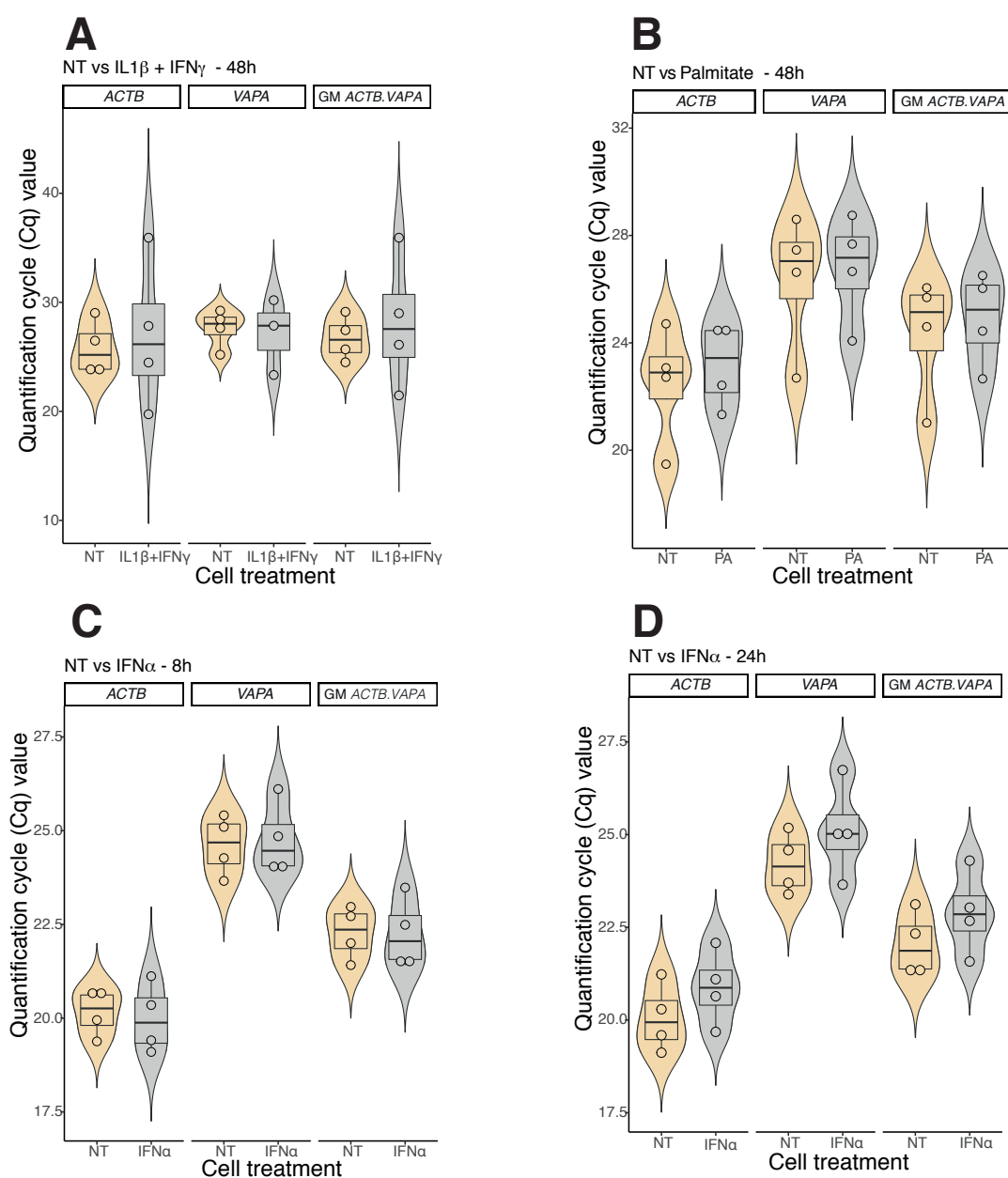
# Fig.01



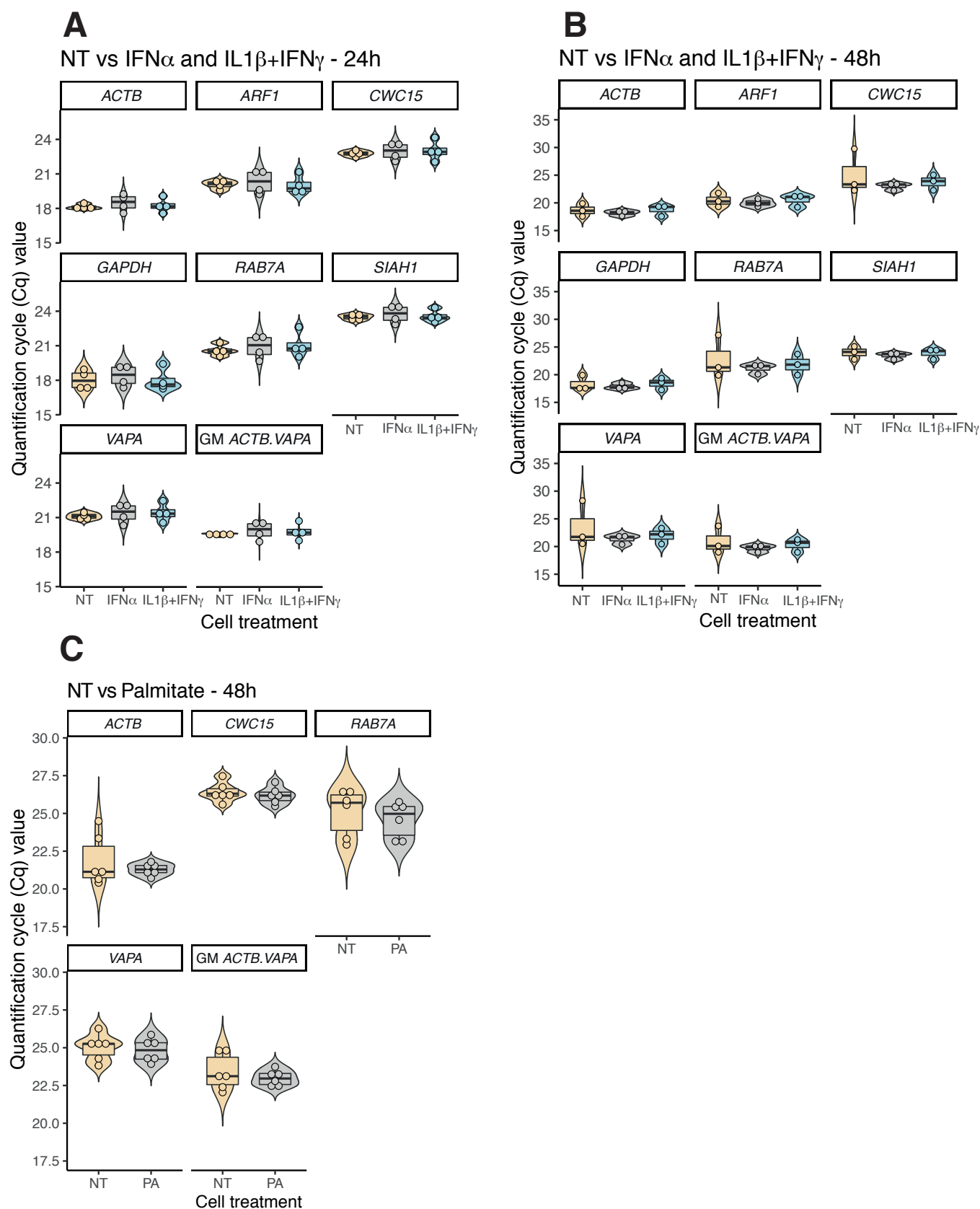
# Fig.02



# Fig.03



# Fig.04



**Fig. 05**

