



Universiteit
Leiden
The Netherlands

Long RNA sequencing and ribosome profiling of inflamed beta-cells reveal an extensive translome landscape

Thomaidou, S.; Slieker, R.C.; Slik, A.R. van der; Boom, J.; Mulder, F.; Munoz-Garcia, A.; ... ; Zaldumbide, A.

Citation

Thomaidou, S., Slieker, R. C., Slik, A. R. van der, Boom, J., Mulder, F., Munoz-Garcia, A., ... Zaldumbide, A. (2021). Long RNA sequencing and ribosome profiling of inflamed beta-cells reveal an extensive translome landscape. *Diabetes*, 70(10), 2299-2312.
doi:10.2337/db20-1122

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3274076>

Note: To cite this publication please use the final published version (if applicable).



Long RNA Sequencing and Ribosome Profiling of Inflamed β -Cells Reveal an Extensive Translatome Landscape

Sofia Thomaidou,¹ Roderick C. Sliker,^{1,2} Arno R. van der Slik,^{1,3} Jasper Boom,⁴ Flip Mulder,⁵ Amadeo Munoz-Garcia,⁶ Leen M. 't Hart,¹ Bobby Koeleman,⁵ Françoise Carlotti,⁶ Rob C. Hoeven,¹ Bart O. Roep,^{6,7} Hailiang Mei,⁴ and Arnaud Zaldumbide¹

Diabetes 2021;70:2299–2312 | <https://doi.org/10.2337/db20-1122>

Type 1 diabetes (T1D) is an autoimmune disease characterized by autoreactive T cell-mediated destruction of the insulin-producing pancreatic β -cells. Increasing evidence suggest that the β -cells themselves contribute to their own destruction by generating neoantigens through the production of aberrant or modified proteins that escape central tolerance. We recently demonstrated that ribosomal infidelity amplified by stress could lead to the generation of neoantigens in human β -cells, emphasizing the participation of nonconventional translation events in autoimmunity, as occurring in cancer or virus-infected tissues. Using a transcriptome-wide profiling approach to map translation initiation start sites in human β -cells under standard and inflammatory conditions, we identify a completely new set of polypeptides derived from noncanonical start sites and translation initiation within long noncoding RNA. Our data underline the extreme diversity of the β -cell translatome and may reveal new functional biomarkers for β -cell distress, disease prediction and progression, and therapeutic intervention in T1D.

Type 1 diabetes (T1D) is an autoimmune disease characterized by the selective and progressive destruction of the

insulin-producing pancreatic β -cells by the invading immune cells (1–3). T cells of high affinity are normally eliminated in the thymus by negative selection in order to establish central tolerance. In autoimmune disease, impaired thymic education or low-affinity T cells are believed to be responsible for the immune attack directed against native self-proteins (4). Yet, there is increasing evidence that local inflammation or other forms of stress combined with genetic predisposition lead to the generation and accumulation of aberrant or modified proteins to which central tolerance is lacking and increased visibility of β -cells to immune cells (5–7). Metabolic and inflammatory stress may perturb the cellular equilibrium and affect high-fidelity transcriptional and translational processes during conversion of the genetic information into proteins. We and others have shown increased RNA splicing in β -cells maintained under pathophysiological conditions (8,9), as well as T-cell autoreactivity against deamidated autoantigens and fusion epitopes illustrating that transcriptional, posttranscriptional, and post-translational processes can be affected during insulinitis, supporting the participation of unconventional antigens in T1D pathology (10–13). While transcription regulation is a slow process, translation and ribosome position along RNA molecules can change upon stress in a matter of seconds

¹Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, the Netherlands

²Department of Epidemiology and Biostatistics, Amsterdam Public Health Institute, Amsterdam UMC, location VUMC, Amsterdam, the Netherlands

³Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, the Netherlands

⁴Sequencing Analysis Support Core, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

⁵Center for Molecular Medicine, Utrecht Medical Center, Utrecht, the Netherlands

⁶Department of Internal Medicine, Leiden University Medical Center, Leiden, the Netherlands

⁷Department of Diabetes Immunology, Arthur Riggs Diabetes & Metabolism Research Institute, City of Hope, Duarte, CA

Corresponding author: Arnaud Zaldumbide, a.zaldumbide@lumc.nl

Received 4 November 2020 and accepted 11 June 2021

This article contains supplementary material online at <https://doi.org/10.2337/figshare.14771181>.

© 2021 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

See accompanying article, p. 2185.

(14). Many genes involved in the unfolded protein response to stress are regulated at a translational level. This swift cell adaptation is illustrated by the regulation of stress-responsive genes (i.e., *ATF4*, *ATF5*, *GADD34*, or *CHOP*), where the activation of the PERK sensor within the endoplasmic reticulum triggers phosphorylation of the translation initiation factor eIF2 and the translational activation of ATF4 controlling the unfolded protein response (15). In this classical model of translation initiation, the RNA precursor is converted into mature RNA by addition of 7-methylguanosine at its 5', addition of a poly-A tail in 3' end, and splicing of intronic regions. Canonically, 5'capping of the mRNA and binding of the 40s ribosomal subunit via eIF4 represents the earliest stage of translation initiation from which the mRNA scanning process continues until recognition of a start codon where the complete 80s ribosome complex starts translation upon recognition of an AUG by the first methionine-charged initiator, tRNA. Based on this model, the identification of proteins expressed by a cell can be predicted from transcriptomics data analysis. However, the translation process is not flawless and noncanonical initiation can occur (16). In addition, environmental perturbation may affect translation fidelity and cellular stress aggravates translation initiation at alternative start codons with possible consequences at the immunological level (17). In fact, changes in reading frame will necessarily affect the amino acid sequence encoded by RNA molecules. Recognized as a danger signal or just as a translational error, these so-called "alternative open reading frames products" may trigger cell destruction by immune surveillance. These nonconventional translational events, arising from translation of normally untranslated regions (UTR), ribosomal frame shifting, or alternative translation initiation (upstream translation initiation sites [uTIS] or downstream translation initiation sites), represent a unique pool of antigenic peptides that rapidly undergo degradation to the proteasome (18,19). Such epitopes have been described frequently in tumor cells (i.e., ARFs, MELOE, NY-ESO-1, and LAGE-1) and are currently explored as tumor biomarkers or as therapeutic targets to promote a specific antitumor response (13). Surprisingly, in the context of autoimmunity, these neoepitopes have rarely been investigated (20), in spite of reports estimating that up to 70% of peptides presented at the cell surface in MHC class I may be derived from defective ribosomal products (21). We recently showed that translation of the human insulin mRNA can generate translational errors that participate in β -cell destruction, providing the first evidence for the implication of translational "junk" products in triggering autoimmunity (22). However, current transcriptomic, proteomic, and peptidomic approaches aiming at extending our view on the β -cell ligandome and improving our understanding of potential involvement of β -cells in T1D pathogenesis (23) are not compatible with the detection of out-of-frame rapidly degraded polypeptide products, thereby masking important features of the β -cell ligandome.

We used a novel unbiased proteogenomic approach to identify all possible open reading frames (ORFs) generated by β -cells maintained in normal and T1D pathophysiological conditions. Our results uncover evidence for translation initiation in novel RNA transcripts as well as long noncoding (lnc)RNA. Moreover, we show that cytokine treatment leads to profound changes in genes involved in MHC class I processing and a significantly increased number of ORFs per transcript and in particular an increased ribosome density within 5'-UTR regions. Altogether, our data present new insights into the dynamic regulation of transcription and translation, revealing an extended β -cell translatome.

RESEARCH DESIGN AND METHODS

Cells and Reagents

EndoC- β H1 cells, kindly provided by Dr. Raphael Scharfmann (Paris Descartes University, Paris, France) (24), were maintained in low-glucose DMEM supplemented with 5.5 μ g/mL human transferrin, 10 mmol/L nicotinamide, 6.7 ng/mL Selenit, 50 μ mol/L β -mercaptoethanol, 2% human albumin, 100 units/mL penicillin, and 100 μ g/mL streptomycin. Cells were seeded in extracellular matrix- and fibronectin-precoated culture plates. Inflammatory stress was induced by a mixture of 1,000 units/mL IFN γ and 2 ng/mL IL1 β for 24 h. Upon incubation, cells were treated with 2 μ g/mL harringtonine for 30 min at 37°C as previously described (16). Cycloheximide (100 μ g/mL) was added for 30 s. Ten million cells were used per condition. Subsequent cell lysis was performed without freezing as described with 900 μ L lysis buffer. The clarified lysates were stored overnight at -80°C .

Nuclease Footprinting and Ribosome Recovery

The materials and methods used were adapted from a previously described protocol (25). Following cell lysis (20 mmol/L Tris Cl [pH 7.4], 150 mmol/L NaCl, 5mmol/L MgCl $_2$, 1 mmol/L dithiothreitol supplemented with 100 μ g/mL cycloheximide, 1% Triton X-100, and 25 units/mL Turbo DNaseI), ribosome complexes were purified by addition of three times the volume of sucrose cushion to the clarified lysate solution in Thickwall Polycarbonate Tube, 13 $\times$ 64 mm (part no. 344645; Beckman Coulter Life Sciences). Ribosomes were pelleted by centrifugation using Delrin Tube Adapters (part no. 303313; Beckman Coulter Life Sciences) in a Type 70.1 Ti Fixed-Angle Titanium Rotor (part no. 337922; Beckman Coulter Life Sciences). The ribosomal pellets were further lysed using 300 μ L Buffer ML from the NucleoSpin miRNA Kit (product code 740971; BIOKÉ). Lysates were stored overnight at -80°C and RNA purified using NucleoSpin miRNA according to the manufacturer's guidelines. The RNA was precipitated using Glycogen, RNA grade (cat. no. R0551; Thermo Fisher Scientific) and resuspended in 10 μ L DEPC-Treated Water (cat. no. AM9920; Thermo Fisher Scientific).

Footprint Fragment Purification

Purified RNA was denaturated (98% [v/v] formamide, 10 mmol/L EDTA, 300 μ g/mL bromophenol blue) and loaded on 18% (w/v) polyacrylamide TBE-urea gel (200 V for 90 min in 1 $\times$ Tris-Borate, EDTA buffer [TBE]. Molecular weight ladder was prepared by mixing 5 μ L microRNA Marker (cat. no. N2102; New England Biolabs) and 0.5 μ L Small RNA Marker (cat. no. R0007; Abnova). The gel with SYBR Gold staining (cat. no. S11494; Thermo Fisher Scientific) was visualized with a ChemiDoc MP Imaging System (Bio-Rad Laboratories). RNA was extracted from excised gel slices spanning the 26–34 nt region with use of the ZR small-RNA PAGE Recovery Kit (cat. no. R1070; Zymo Research) according to the manufacturer's protocol. The size-selected RNA was stored at -80°C .

Dephosphorylation and Linker Ligation

Dephosphorylation reaction and linker ligation were performed with the Universal miRNA Cloning Linker (cat. no. S1315S; New England Biolabs) in a T100 Thermal Cycler (product no. 1861096; Bio-Rad Laboratories). RNA was precipitated with Glycogen, RNA grade (cat. no. R0551; Thermo Fisher Scientific), and resuspended in 10 μ L DEPC-Treated Water. Size separation and gel extraction of the ligation reaction were performed as described in the footprint fragment purification section.

Reverse Transcription and Circularization

The reverse transcription reaction was performed in a T100 Thermal Cycler using RT primers). Gel extraction were performed as described in the footprint fragment purification section and cDNA circularized using CirLigase following guidelines from the manufacturer.

rRNA Depletion and Barcoding

The rRNA depletion was performed with a T100 Thermal Cycler and an Eppendorf Thermomixer 5350 Mixer using depletion primers (Table 1). PCR was performed with use of library preparation primers. PCR strips were removed at the end of the extension step after 7, 8, 9, 10, 11, and 12 cycles. The samples were separated by electrophoresis using a 9% (w/v) polyacrylamide TBE-urea gel at 180 V for 45 min in 1 $\times$ TBE. As molecular weight marker, 5 μ L of 50 bp DNA Ladder (cat. no. B7025; New England Biolabs) was included. The SYBR Gold-stained gel was visualized with a ChemiDoc MP. Excised gel slices containing the amplicons were crushed and soaked overnight at room temperature under gentle mixing, with 400 μ L diffusion buffer for PAGE gels (500 mmol/L ammonium, pH 8.0; 0.1% sodium dodecyl sulfate; 1 mmol/L EDTA; and 10 mmol/L magnesium acetate) and NucleoSpin Gel & PCR Clean-up (product code 740609; BIOKÉ). Subsequently, the libraries were isolated from the extraction mix according to the manufacturer's recommendations and eluted in 30 μ L Elution Buffer NE. The libraries were quantified and characterized with the Experion DNA 1K Analysis Kit (product number 7007107; Bio-Rad Laboratories) on an

Experion Automated Electrophoresis System (Bio-Rad Laboratories) following the manufacturer's protocol.

Data and Resource Availability

All data needed to evaluate the conclusions here are present in the article and Supplementary Material. Raw data sets are accessible through Gene Expression Omnibus (GEO) series accession no. GSE167223 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE167223>).

RESULTS

Precise Mapping of the Translation Initiation Start Sites in β -Cells by Ribosome Profiling

To evaluate translational changes induced by proinflammatory cytokines during the early events of T1D, we treated the human β -cell line EndoC- β H1 with proinflammatory cytokines (IFN γ and IL1 β) and isolated ribosome-protected fragments (RPFs) to map translation initiation sites (TIS) ($n = 3$ independent experiments). Using a combined treatment with harringtonine that incorporates at the A-site of the large 60S ribosomal subunit, thereby blocking the entry of a second tRNA, and cycloheximide, which displays a high affinity for the uncharged tRNA E-pocket of the 80S complex, thereby preventing further elongation, we froze ribosomes at the initiation of translation (16,25,26). Following RNaseI treatment, polysomes were purified and RPFs isolated to generate ribosome footprint libraries containing the TIS (25) (Fig. 1A). As anticipated from the stoichiometric hindrance of ribosomes on RNA molecules, the length of fragments sequenced ranged from 26 to 31 nucleotides independently of the treatment (Fig. 1B). Alignment of the RPFs to the human transcriptome revealed an accumulation of reads at the beginning of the translated region of the transcripts, a consequence of the high ribosome occupancy at the translation start (Supplementary Fig. 1A). For reads of the same length, P-site offsets were computed to precisely map the position of the ribosome P-site, corresponding to the start codon, on the RNA footprint (27). Shorter incubation with harringtonine failed to stop the ribosome complex at the start codon, leading to noise during elongation and accumulation of reads at 3'-UTR resulting from ribosome stalling before completing peptide synthesis (Supplementary Fig. 1B). The occupancy profiles of the RPFs 5' and 3' end extremities around the referenced start codons showed accumulation of read ends at locations ± 12 (Supplementary Fig. 1C), validating both ribosomal freezing and computational calculation and allowing mapping of the start site (Supplementary Fig. 1D and E). In order to eliminate false positives and minimize background artifacts resulting from the speed of transcript translation, incomplete blockade by translation inhibitors, or possible nonribosomal protein complexes interacting with the mRNA, we

Table 1—Primer sequences used to generate the RPF library

Primers sequences (5' to >3')	
Size markers	
Upper size marker	AUGUACACGGAGUCGAGCUCAACCCGCAACGCGA-(Phos)
Lower size marker	AUGUACACGGAGUCGACCCAACGCGA-(Phos)
Universal miRNA cloning linker	rAppCTGTAGGCACCATCAAT-NH2
cDNA	
Reverse transcription	(Phos)-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC-(SpC18)- CACTCA-(SpC18)-TTCAGACGTGTGCTCTTCCGATCTATTGATGGTGCCTACAG
Purification	
rRNA depletion	GGGGGGATGCGTGCATTTATCAGATCA TTGGTGA CTCTAGATAACCTCGGGCCGATCGCACG GAGCCGCCTGGATACCGCAGCTAGGAATAATGGAAT TCGTGGGGGGCCCAAGTCCTTCTGATCGAGGCC GCACTCGCCGAATCCCGGGGCGAGGGAGCGA GGGGCCGGGCGCCCTCCACGCGCG GGGGCCGGGCGCCCTCCACGCGCG CCCACTGCGCCCGGGCGTCTGCGCCGTGCGGTCCCGG TCCGCGAGGGCGCACCAACGGCCGTCTCGCC AGGGGCTCTCGCTTCTGGCGCCAAGCGT GAGCCTCGGTTGGCCCGGATAGCCGGTCCCCGT GAGCCTCGGTTGGCCTCGGATAGCCGGTCCCCCG TCGCTGCGATCTATTGAAAGTCAGCCCTCGACACA TCCTCCGGGGCTACGCCTGTCTGAGCGTCTGCT
Library preparation	
Forward library PCR	AATGATACGGCGACCACCGAGATCTACAC
Index reverse library PCR	CAAGCAGAAGACGGCATACGAGATAGTCGTGTGACTGGAGTTCAGACGTGTGCTCTTCCG CAAGCAGAAGACGGCATACGAGATGATGTGACTGGAGTTCAGACGTGTGCTCTTCCG CAAGCAGAAGACGGCATACGAGATATGCTGGTGAAGTTCAGACGTGTGCTCTTCCG CAAGCAGAAGACGGCATACGAGATACGTCGGTGAAGTTCAGACGTGTGCTCTTCCG

developed an algorithm based on the read distribution surrounding the suspected P-site to identify the most relevant translation initiation starts. Reasoning that the accumulation of RPFs at a start position is normally distributed, and the read spread in flanking position is determined by the number of reads of the peak, we defined an enrichment method using a dynamic sliding window encompassing each P-site position (Supplementary Material and Supplementary Fig. 2A and B). Significance testing was performed using a negative binomial regression model, and samples belonging to different library preparations and treatment were statistically paired. Using this enrichment method, we identified 15,014 start sites within 5,529 genes in both untreated and cytokine-treated samples (Fig. 1C and Supplementary Material). Codon occupancy analysis at the E, P, and A pockets indicated a conserved AUG motif at the P-site within a Kozak consensus sequence both in nontreated and cytokine-treated samples (Fig. 1D).

Direct RNA-Sequencing and Ribosomal Sequencing Technology Reveals Active Translation in Novel Transcript Variants and lncRNA

To evaluate the impact of cytokine stimulation on translation, we performed a negative binomial regression on the read numbers at the P-site position. As expected, pathways related to inflammatory response, such as antigen

processing, presentation, and apoptosis, were induced upon treatment (Fig. 2A). More specifically, translation initiation within transcripts of these inflammatory pathways was significantly upregulated (P value <0.05 and $\log_2$ fold change >1.5) (Fig. 2B). Increased expression of key proteins for secretory pathway (CHGA), IFN signaling (STAT1, STAT2), and antigen processing (PSMB8, PSMB9, and PSMB10) was validated by Western blot in EndoC- β H1 cells and confirmed in primary human islets (Supplementary Fig. 3). While these data do not provide proofs for expression of the alternative ORFs, they contribute to technical validation of the ribosome profiling approach. Despite a stringent detection method, 17% of P-sites identified corresponded to the referenced annotated start codon (Fig. 2C). Importantly, we identified several misannotated translational start sites (i.e., TAP1 and PSMD8) experimentally validated in previous ribosome profiling assays (Translation Initiation Site Database [TISdb]) (28) or extra start sites previously confirmed by mass spectrometry (i.e., YBX1 and PGRMC1) (Supplementary Fig. 4A–D). Through reconstituting the 3-base periodicity move of the ribosomes on the RNA molecule from the identified start sites, our analysis revealed a wide range of novel or modified ORFs leading to translation rearrangements (upstream initiation, uORF, that ends before the coding sequence [CDS]; N-terminal extension of annotated protein; overlapping uORF spanning 5'-UTRs and CDS on an alternative frame/translation

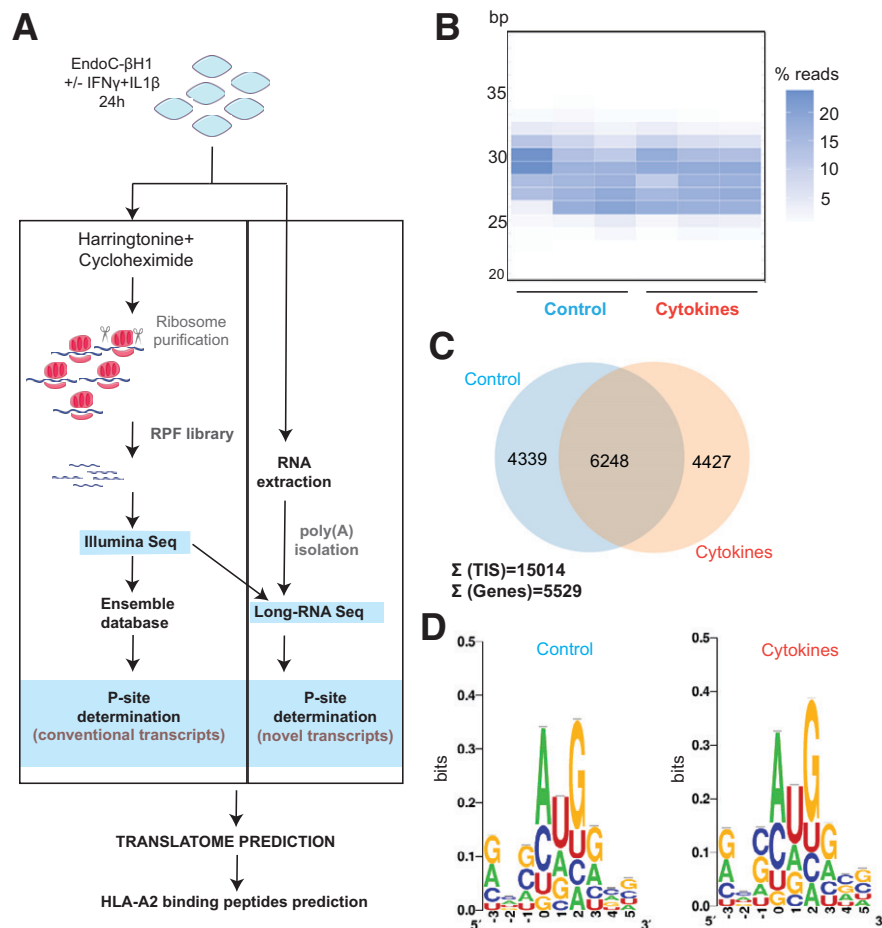


Figure 1—Ribosome profiling of EndoC-βH1. **A**: Schematic representation of experimental workflow. **B**: Length distribution of ribosomal footprints of nontreated (control) and IFN γ +IL1 β -treated samples (cytokines) ($n = 3$). **C**: Venn diagram of P-sites identified in control and cytokines-treated samples. The total number of genes carrying a TIS and the total number of TIS are indicated. **D**: Qualitative sequence logo of codons composition corresponding to E- (-3 to 0), P- (0 to 3) and A- (3 to 6) pockets of the ribosome.

frameshift, internal off frame; and downstream initiation, initiation on the 3'-UTR (Fig. 2D). Moreover, we detected a considerable fraction (7%) of the RPFs mapped to non-coding RNA and preferentially in lncRNA (Fig. 2E and F). The polypeptide lengths of ORFs initiated at alternative start sites in coding ($n = 25,466$ in nontreated and $n = 25,984$ in cytokine conditions) or noncoding RNA ($n = 2,634$ in nontreated and $n = 2,381$ in cytokine) were shorter than for canonical ORFs (Fig. 2G). These differences in peptide sizes observed are in line with previous studies conducted on embryonic stem cells and cancer cell lines (29,30).

Moreover, nearly 20% of the purified ribosome footprints did not align to the referenced transcriptome, indicating the presence of TIS on novel transcripts (Fig. 3A and Supplementary Material). To further explore the β -cell transcriptome and translome, we stimulated EndoC-βH1 cells with cytokines and performed direct long RNA sequencing (RNA-seq) in search of novel transcript variants (Supplementary Material and Fig. 3B). Following a differential expression analysis, we identified a

classical transcriptional inflammatory signature on T1D susceptibility genes (31,32) (Fig. 3C and Supplementary Material). Long RNA-seq identified 6,892 new transcripts, among which 37 were specifically present under inflammatory conditions (Fig. 3D). To validate our results, we aligned reads from our de novo transcriptome to RNA-seq data sets from pancreatic β -cells (33). As described in Supplementary Fig. 5A and B and Supplementary Material, 74% of the newly identified transcripts spanning over splicing junctions in the β -cell line (77% of the T1D susceptible genes) are supported by RNA-seq reads from the primary β -cell-enriched fraction. In addition, we evaluated the expression pattern of two novel cytokines-specific transcripts (i.e., PSMB9 TALONT000709680 and HLA-C TALONT000700063) in EndoC-βH1 and primary human islets. As shown in Supplementary Fig. 5, classical RT-PCR analysis confirmed the cytokine specificity in EndoC-βH1. The expression in islets may illustrate the difference between EndoC-βH1 and primary β -cell but more likely reflect the presence of other islet cell types in the RNA preparation (Supplementary Fig. 5C–F).

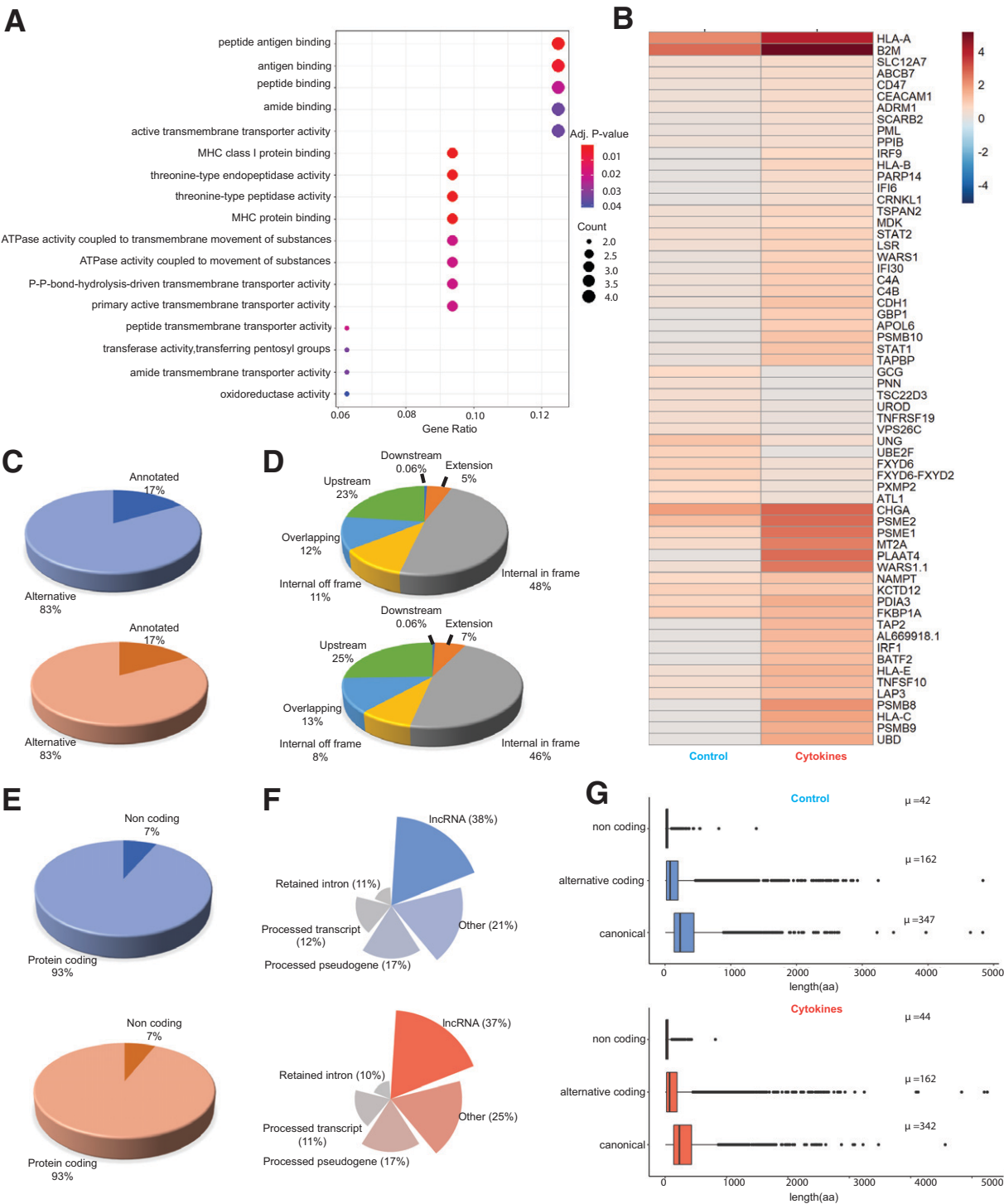


Figure 2—Mapping of TIS on the transcriptome. **A**: Gene ontology plot of canonical TIS (cpm > 0.5), showing significant log fold change compared with the control (log fold change > 1.5). The plot was generated with clusterProfiler. **B**: Heat map of the canonical TIS (cpm > 0.5) significantly up- or downregulated in cytokines-treated samples. The heat map was generated with pheatmap. **C**: Pie charts representing the percentage of alternative and annotated TIS in experimental conditions. **D**: Pie charts representing the abundance percentage of the different types of ORFs derived from the alternative TIS represented in **C**. **E**: Pie charts of TIS identified in coding and noncoding RNA. **F**: Polar plots of noncoding transcripts subtypes, showing one or more TIS. **G**: Length of ORFs potentially generated by the identified TIS. The average length of the ORFs is indicated as μ . aa, amino acids; Adj., adjusted.

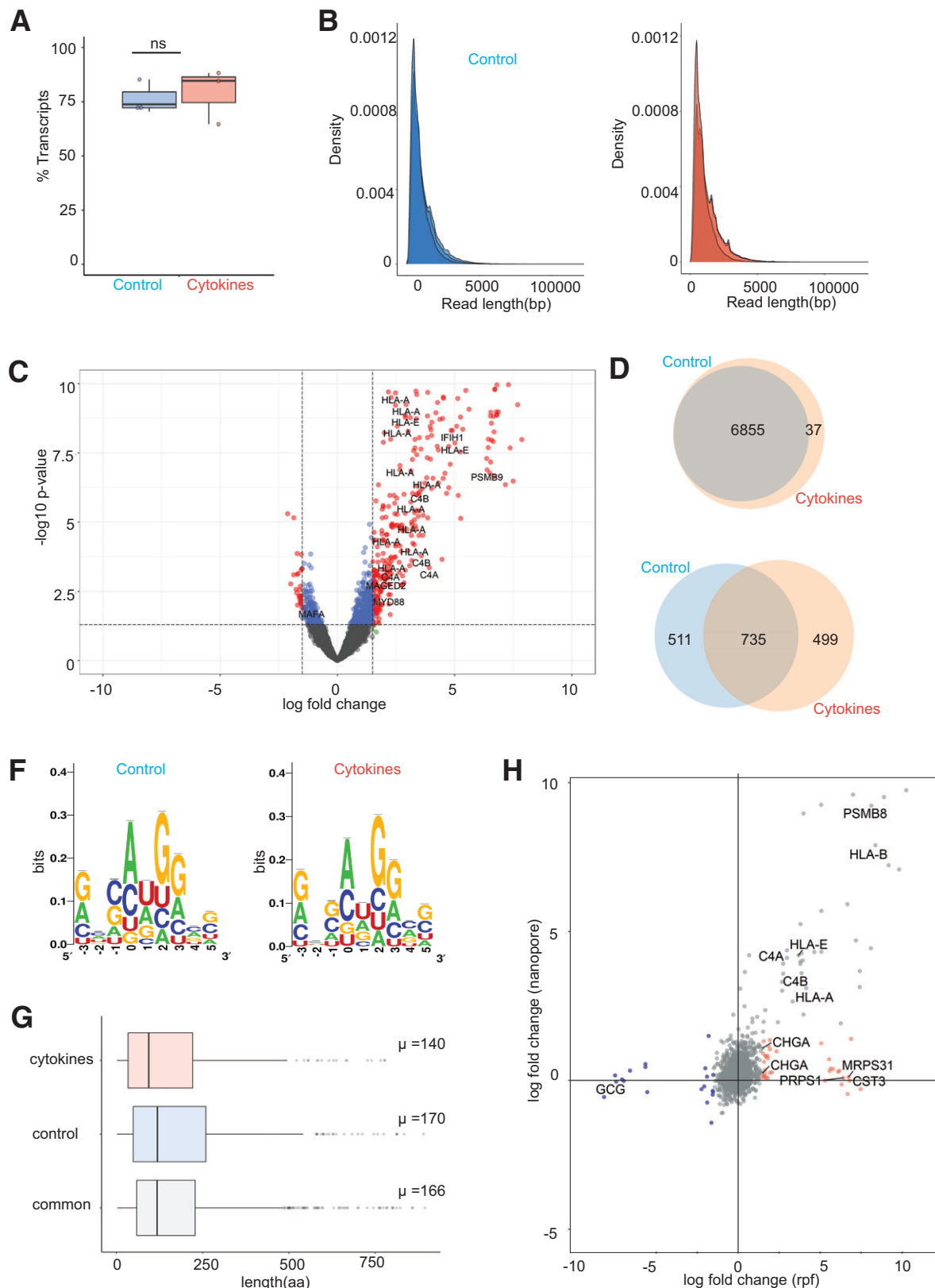


Figure 3—Integration of long RNA-seq and ribosome profiling. **A:** Mean percentages of ribosomal footprints aligned to the transcriptome ($n = 3$). **B:** Nanopore QC analysis showing the RNA read length. The cleaned reads were aligned to the transcriptome (Ensembl build GRCh38.99) for the controls ($n = 3$) and cytokine-treated samples ($n = 3$) with use of STAR, version 2.7.3a. **C:** Volcano plot of transcripts expression detected in long RNA-seq ($n = 3$, read cutoff: mean of replicates = 10). T1D susceptibility genes are depicted. **D:** Venn diagram representing the novel transcripts detected in control and cytokine-treated conditions. **E:** Venn diagram representing the number of

Alignment of the RPFs to the newly defined β -cell transcriptome revealed >1,745 TIS located on the novel transcripts, including 499 TIS specifically detected upon treatment (Fig. 3E). Importantly, these TIS were surrounded by Kozak-like sequences and could potentially give rise to ORFs ~150 amino acids long (Fig. 3F and G and Supplementary Material). By integrating transcriptome and translatome data sets, we evaluated the transcriptional and translational adaptive response to inflammatory cytokines in β -cells. CST3, PRPS1, and MRPS31 in particular appear regulated at the translational level rather than transcriptional level (Fig. 3H and Supplementary Material).

Cytokine Treatment Distorts Translation Fidelity, Triggering Generation of Alternative Translation Initiation Sites

Metagene distribution of the TIS throughout the whole transcriptome indicated that cytokine stimulation led to a significant decrease in the number of transcripts presenting a unique start site (P value <0.0001), and an increased proportion of transcripts with multiple TIS (P value <0.0001), suggesting that inflammation may actively trigger the generation of intratranscript initiation events (Fig. 4A). Moreover, we observed an increased number of P-sites within 5'-UTR regions (~4% increase, P value <0.0001) in the cytokine condition in comparisons with nontreated samples, which is in line with studies showing a slower progression of the ribosome complex on RNA under stress conditions (34,35) (Fig. 4B). These changes were associated with a modification of the initiation start codon motif (CUG) (Fig. 4C). Differential read expression at the alternative TIS showed increased translation efficiency within transcripts involved in the inflammation response, in particular in the 5'-UTR and CDS region of HLA, TAP1, TAP2, and IFIH1 (Fig. 4D). A similar inflammatory response was obtained within newly identified HLA-A and HLA-E transcripts and within transcripts encoding for immunoproteasome subunits (e.g., PSMB9) (Supplementary Fig. 6A). Within noncoding RNA a higher translation initiation was also detected in noncoding isoforms of PSMB9 and ARF4 transcripts (Supplementary Fig. 6B). A differential ribosomal occupancy was found specifically in lncRNA, in particular within EBLN3P, AL627171.4, or AC112487.1, suggesting an increased production of those small ORFs during inflammation (Supplementary Material).

Combining information on translation efficiency at different start sites, we observed a poor correlation between differential read expression at the alternative P-site location and changes at the canonical counterparts ($R = 0.32$),

suggesting specific regulatory mechanisms for translation initiation at each position. Cytokine treatment prevented or favored ribosome docking to alternative positions ($n = 189$ and $n = 135$, respectively) (Fig. 4E and Supplementary Material). Pathway analysis performed on these transcripts pointed to an active regulatory mechanism driven by inflammatory signals to adjust RNA splicing and translation machinery in response to stress (Supplementary Fig. 7). Alternative translation initiation within MAGED2, VEGFA, and CHGA transcripts was induced by cytokine stimulation. While no polypeptide could be expected from MAGED2 alternative start (AAT_{Start}TGA_{Stop}), we observed an increased ribosome density during inflammation at an alternative position within the 5'-UTR and 3'-UTR region of the CHGA transcript that could generate short polypeptides. Using a GFP fusion construct, in which the GFP ORF is positioned in frame with the CGG-103, CCG-40, and CTG-22 start codons identified by ribosome profiling, we validated that the 5'-UTR region can be used to initiate protein translation in a +2 frame (Supplementary Fig. 8). Of note, translation initiation within 3'-UTR at +1512 could generate a potentially immunogenic polypeptide carrying a strong HLA-A2 epitope. Similarly, within VEGFA transcripts, a novel upstream TIS was identified located upstream the canonical CUG (VEGFA₁₁₁), upregulated upon treatment, that can potentially generate two strong HLA-A2 binders (KVSDLLLG and LLLGVTAGA) (Fig. 4E and F). Of note, we detected an extra in-frame start site, in position +3 (data not shown), in addition to the canonical CUG start codon (position 0) that carries the SRFGGAVVR epitope, approximately sevenfold upregulated upon cytokine treatment, which could drive increased presentation of previously identified cryptic VEGFA epitope (36).

Regarding insulin gene products, a high ribosome density was detected at the canonical start site at position 0, as expected. Additionally, we identified alternative starts within the 5'-UTR region at position -44 and -34 that generate overlapping ORFs (upstream ORF in +1 and +2 reading frame, respectively) (Fig. 5A and Supplementary Material). Moreover, we detected an accumulation of ribosomes downstream the canonical start site at position 286, confirming that this AUG may serve as docking site for translation initiation of the immunogenic out-of-frame insulin defective ribosomal product (INS-DRiP) that we described earlier (22). With use of long RNA-seq, five novel INS transcripts were identified resulting from size variations within exon 1 and 3, intron 1 retention, or 3' extension. While TALONT000499163, TALONT000499165, TALONT000499178, and TALONT000499193 were in low abundance, the INS-IGF2 transcript (TALONT000499153),

TIS detected on novel transcripts (shown in D) upon alignment of ribosome profiling with long RNA-seq data. F: Codon usage of TIS on novel transcripts. Sequence logo, corresponding to E- (-3 to 0), P- (0 to 3), and A- (3 to 6) sites of the ribosome pockets. G: Length of ORFs potentially generated by the identified TIS on novel transcripts. The average length of the ORFs is indicated as μ . H: Scatterplot representing the differential expression of the RPF at TIS position (x-axis) vs. differential expression of the long RNA-seq transcripts (y-axis). T1D susceptibility genes are depicted.

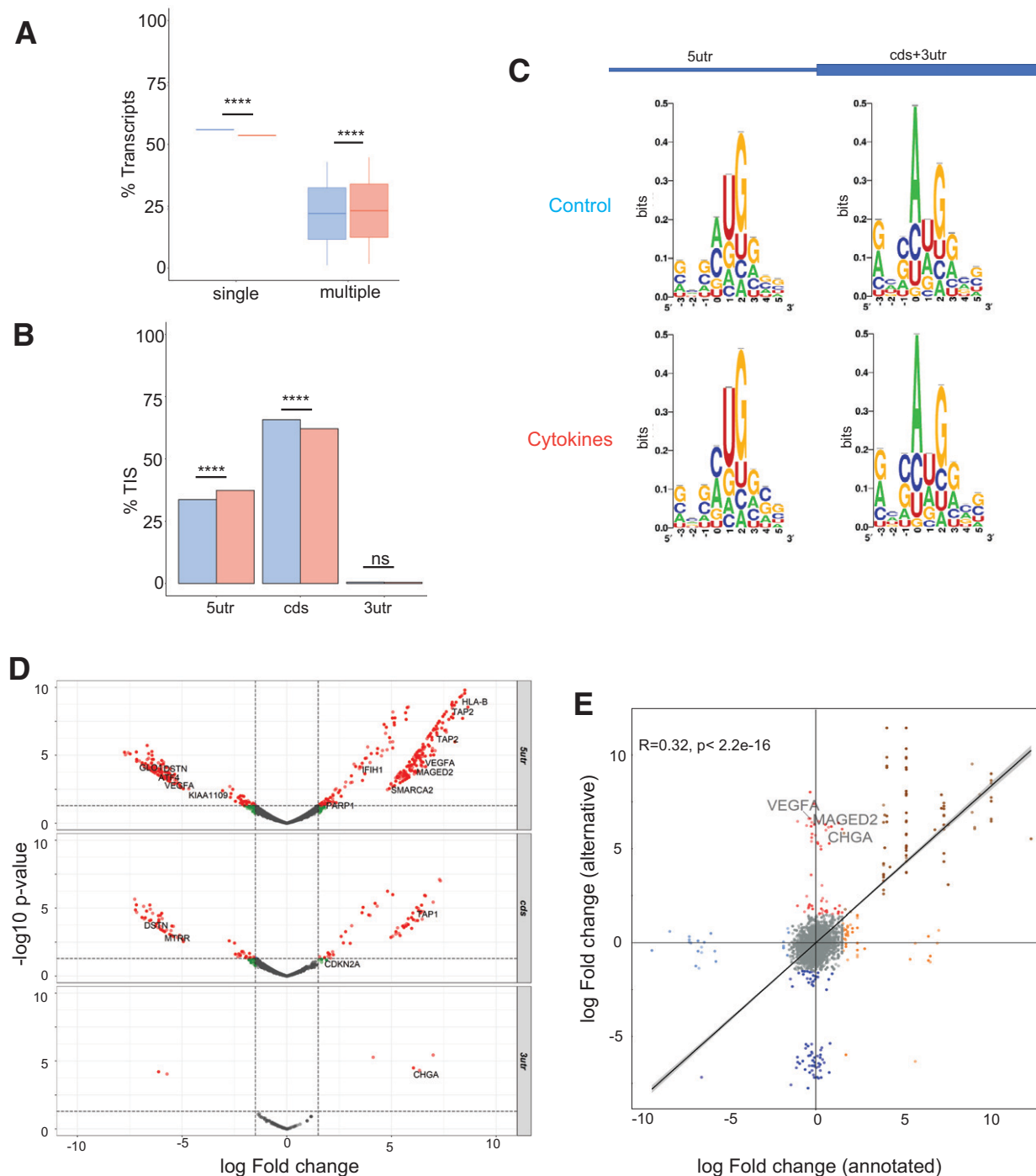


Figure 4—Inflammation disturbs the translational process and fidelity. **A:** Percentage of transcripts showing a unique or multiple TIS (two or more). Statistical analysis was performed using the z-score test. **B:** Region distribution of the TIS identified on the transcriptome in control or cytokine treated conditions. Statistical analysis was performed using the z score test. **C:** Qualitative codon usage profiles of upstream and downstream the canonical start codon. Sequence logo, corresponding to E- (−3 to 0), P- (0 to 3), and A- (3 to 6) sites of the ribosome pockets. **D:** Volcano plot of the out-of-frame TIS differential expression per transcript regions. **E:** Scatterplot representing the differential expression of the RPF at canonical TIS position vs. differential expression of the out-of-frame alternative TIS. Spearman correlation test was used to calculate the R factor and the corresponding *P* value. T1D susceptibility genes are labeled.

deriving from the fusion between exon 3 of the *INS* gene and the *INS-IGF2* intronic region, is the third most abundant *INS* transcript expressed by β -cells, representing 3%

of the total *INS* reads. Using a primer set spanning the end of the classical *INS* transcript, located in *INS* exon 3 and in the *INS-IGF2* intergenic region, we validated the presence

of the TALONT000499153 transcript in EndoC- β H1 cells and primary islets and validated the similar transcript abundance, measured by long RNA-seq, in normal and cytokine conditions (Fig. 5B). Ribosome density on this region identified a potent TIS that would lead to the generation of a 116-amino acid polypeptide carrying the insulin DRiP epitope and bearing another potential HLA-A2 binder (Fig. 5C). To validate alternative translation initiation within INS transcripts in EndoC- β H1 cells, we tested the MLYQHLLPL epitope presentation to DRiP-specific CD8 T-cell clone. Using JY cells pulsed with insulin-derived peptides (PPI, ALWGPDPAAG; B-chain, HLVEALYLV; DRiP, MLYQHLLPL), we confirmed DRiP T-cell specificity by detecting T-cell activation only in the presence of the cognate peptide (Fig. 5D). Following genetic modification to express HLA-A2 in EndoC- β H1 cells, we validated that the exogenous HLA-A2 expression in the presence or absence of cytokines was similar to the endogenous HLA class I expression (Fig. 5E). After coculture with the DRiP-specific CD8 T cells, we detected an increase CD8 activation, measured by CD107a staining and MIP1 β release in the presence of cytokine stimulation (Fig. 5F), confirming the importance of an HLA hyperexpression in the β -cell destruction process but also providing proof for processing and presentation of the N-term fragment of the INS alternative ORF in EndoC- β H1 cells.

DISCUSSION

The defective ribosomal protein (DRiP) theory arose in 1996 (21) when discrepancies between the ligandome and proteome were found in virally infected cells as a plausible explanation for T-cell reactivity against flu-derived peptides detected prior to synthesis of their source proteins. Originally identified as prematurely terminated polypeptides and misfolded polypeptides produced from translation of bona fide mRNAs in the proper reading frame (21), DRiPs include all alternative translation initiation products, including UTR translation and ribosomal frameshifts products, each of these susceptible to generation of a new pool of antigenic peptides that would share not any, or little, amino acid similarities with genuine mRNA translation products. Such DRiPs are widely studied in viral infections and cancer and are considered the main target of effective immunosurveillance. In human pancreatic islets, we previously showed that cryptic TIS can be used during translation of insulin mRNA to generate highly immunogenic insulin gene-derived polypeptide targeted by T-cell autoreactivity in T1D patients, providing proof of principle for the participation of translational errors in autoimmune disease (22). Here, our data reinforce our findings and confirm that DRiPs can be generated from human insulin mRNA, even though no significant effect of cytokines could be detected on this particular start site. The specific killing observed in inflammatory milieu despite similar amounts of DRiP polypeptide in cytokine-treated and untreated

β -cells illustrates the importance of protein stability, degradation, and presentation in β -cell destruction (37). Moreover, combining long RNA-seq and ribosome profiling, we propose another possible explanation for the origin of this immunogenic polypeptide, through the presence of a novel INS-IGF2 transcript that could serve as a template for protein synthesis. In line with these findings, the INS-IGF2 locus has been correlated with autoimmunity and loss of β -cell functionality in T1D (38,39). Therefore, deciphering its transcriptional and translational profiling in human islets may be crucial for a better understanding of disease onset and progression.

The large data set of alternative translatome revealed in this study begs for revision of transcriptome or mass spectrometry-based approaches for epitope discovery in autoimmunity, viral infection, and cancer. Our data showing ribosome docking at an alternative position within the INS transcript and cytotoxic T lymphocyte (CTL) immunoreactivity against the N-terminal part of the INS-DRiP polypeptide, while this epitope could not be eluted from previous peptide elution study, emphasize the limitation of the current immunopeptidomic studies (23). Though the human β -cell line are different (ECN90 vs. EndoC- β H1/A2) these results emphasize the need to combine transcriptomic/ribosome profiling approaches with immunopeptidomic for epitope discovery. Despite the fact that translation initiation can be regulated by the secondary RNA structure (favoring some transcripts) (40), the results presented here show that β -cells may generate up to 22,670 novel ORFs potentially immunogenic out of the 15,014 start sites identified, assuming a minimal 9-mer residue core for HLA class I peptide. Of note, nine-amino acid ORFs can be presented without further processing, as shown for the METAATAVAA peptide derived from a 5'-UTR ORF of ARAF protein, eluted as an intact peptide from HLA-B45 in previous study (41) and identified in our database. Although our data largely expand the pool of peptides generated by β -cells and potentially presented at the cell surface, their pathological significance remains to be determined. Certainly, the extreme sensitivity and specificity of CD8 cells to recognize peptide-HLA complexes (42) could explain how the expression of rare peptides can drive an immune response and break tolerance in T1D. This process still remains poorly understood.

Despite remarkable improvement regarding sensitivity, mass spectrometry may not suffice for the identification of short-lived unstable proteins or for proteins containing many cleavage sites for trypsin that is commonly used in sample preparation. Moreover, biases created by the input database and low Mascot scores of out-of-frame sequences may hamper DRiP discovery. Nearly 40% of the overall initiation footprints derived from events of noncanonical translation may generate neopolypeptides (not in frame sequences), even in resting β -cells. These findings are consistent with other studies on mouse embryonic stem cells and healthy tissues deriving from hepatocellular carcinoma

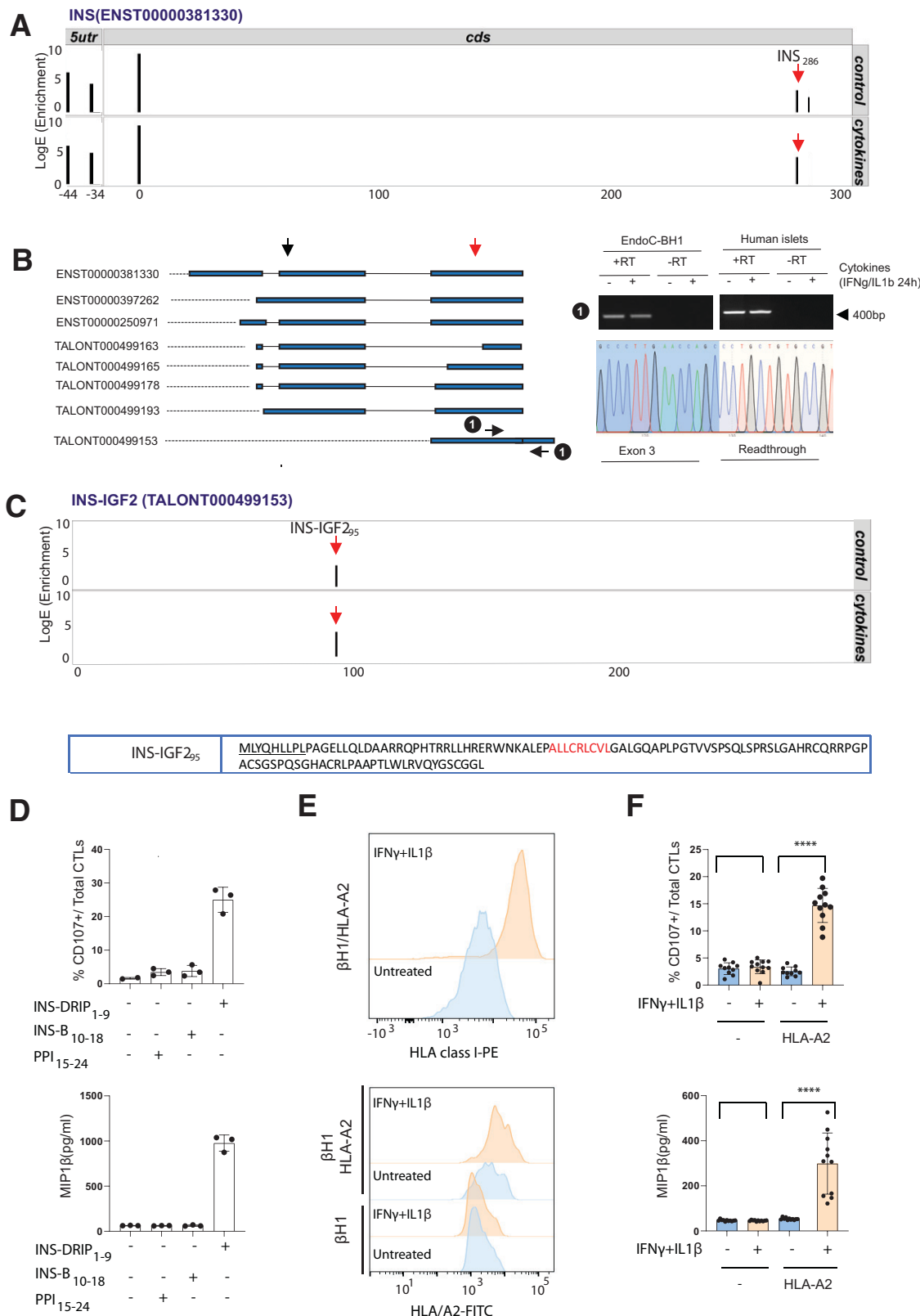


Figure 5—Translation initiation sites (TIS) and alternative translation initiation sites (aTIS) within INS transcripts. **A:** TIS positioned on the INS transcript (ENST00000381330). The red arrow indicates the position of INS₂₈₆. Primers used for RT-PCR are indicated as primer set 1. **B:** Schematic representation of the INS and INS-IGF2 transcripts detected by long RNA-seq. Newly identified transcripts are labeled as “TALONT.” The arrows indicate the canonical start site (black) and INS₂₈₆ (red) (left panel). RT-PCR performed on EndoC-βH1 RNA and primary islet isolated RNA, in presence or absence of reverse transcriptase (+/– RT), using TALON-specific primers and sequence analysis of the PCR amplicon. **C:** TIS mapping of the INS-IGF2₉₅ TIS on the novel INS-IGF2 transcript, TALONT000499153. The INS-DRIP

patients (29,43). Our data demonstrate that an inflammatory environment led to significant increases in transcripts bearing more than one translation start in β -cells, alike during tumorigenesis (43). The impact of inflammation on alternative start sites within transcripts encoding for ribosomal subunit proteins (RPL17, RPL35A, RPL37A) may suggest a role in the regulation of the translational process or in the assembly of the ribosome complex that could impact on the generation of immunoribosomes and production of DRiPs (44,45). In line with other ribosome profiling studies, the 5'-UTR entailed a high proportion of footprints (16,29,30,46). Inflammation induced further upstream translation, highlighting this region as a potentially interesting source for the generation of neoantigens or small regulatory peptides. Besides the difference in amount of uTIS, treatment with inflammatory cytokines caused a shift versus CUG translation compared with the control. Previous studies have shown that CUG translation, mediated by a dedicated Leu tRNA and EIF2a (47), has been widely correlated with a variety of stressors, like proteotoxic and oxidative stress, viral infection, and treatment with type I and type II interferons (17,48). Moreover, CUG-initiated polypeptides gave rise to highly immunogenic epitopes (e.g., cryptic antigens originating from VEGFA, bearing alternative CTG ORFs, have already been reported in the Ensembl database and studied in the context of renal cell carcinoma) (17,36,49).

While in our data the classical AUG represent a minority of TIS, these results are in line with previous ribosome profiling studies. Yet, the conservation of a G in -3 and $+1$ position underlines the importance of the Kozak consensus sequence in initiating the translation. Recent work aiming at evaluating the strength of start codons has demonstrated that a non-AUG codon located within a Kozak-like surrounding sequence may lead to higher translation rates than a non-Kozak AUG (50). Polypeptides arising from these non-AUG alternative starts may not necessarily initiate with non-methionine amino acids, since eIF2a methionine carrying tRNA loading on a non-AUG starting codon has been provided by *in vitro* assays using [35 S]-labeled Met-tRNA^{Met} (51,52). Also, some findings from ribosome profiling experiments have been confirmed by mass spectrometry (53), in particular N-terminal extension of annotated proteins.

Noncanonical TIS were not only present on coding transcripts; we observed that 7% of the footprints originated from various forms of presumably noncoding RNAs, especially lncRNAs. Several studies have confirmed the

presence of ribosome attached to lncRNA, suggesting the synthesis of small polypeptides (<https://www.sorfs.org/database>). The absence of epitopes in these polypeptides suggests a regulatory function (54).

The study relies on the use of EndoC- β H1 cells as an *in vitro* model for human β -cells. While those cells have been described to represent the best alternative so far to primary human β -cells and, as such, constitute a relevant "humanized" model to study β -cell function (in particular, EndoC- β H1 cells and human islets have very similar responses to cytokines and other stressors [55–58]), recent in-depth multiomics analyses have highlighted some differences between EndoC- β H1 cells and primary β -cells (59) (e.g., the absence of iNOS expression in response to cytokines [60]), inherent to both the fetal/embryonic origin of the cell line and/or its transformed state. The presence of ribosomes bound to glucagon transcripts in our study may be the reflection of this observation. Another limitation is intrinsic to the ribosome profiling assay, in which a number of identified start sites may be the consequence of monosome footprint or background artifacts. Yet, despite very stringent detection methods (that exclude robust interpretation of ribosome reads at IAPP, GAD65, ZnT8, and IGRP start position), future follow-up studies to validate the identified polypeptides, presented here, are required to define the missing fraction of the β -cell proteome. Altogether, our results reveal that the translation process plays a significant role in proteome diversity. While some of these peptides synthesized by β -cells may prove unstable and lacking biological functions, they may represent by-products of the translation process that can contribute to the β -cell translatome. The fact that translation is highly dynamic and affected by environmental modifications suggests that perturbations occurring on β -cells at the onset of disease or throughout the course of T1D may change the β -cell ligandome, generating neoepitopes provoking immune surveillance. Currently, drugs are being developed to inhibit noncanonical translation initiation (61), which may offer strategies to control β -cell homeostasis, immunogenicity, and preservation.

Novel noncanonical ORFs warrant further functional and immunological studies. We contend that understanding mechanisms of neoantigen generation is critical for the development of novel immune-targeted therapies.

Acknowledgments. The authors thank Dr. Fabio Lauria (Institute of Biophysics, CNR Unit at Trento, Trento, Italy) and Ruben van 't Slot (UMC Utrecht) for technical assistance.

Funding. This work is supported by JDRF, DON, and the Dutch Diabetes Research Foundation and by the IMI2-JU under grant agreement no. 115797

epitope is underlined. The sequence in red shows the new potential HLA-A2 epitope predicted from the INS-IGF2 polypeptide. *D*: Killing assay performed on JY cells loaded with preproinsulin-derived peptide (PPI), insulin B chain peptide, and DRiP peptide using DRiP-specific CD8 T cells. T-cell activation is shown as CD107a staining (left panel) and MIP1 β release (right panel). *E*: HLA-A/B/C expression in EndoC- β H1/HLA-A2 upon treatment (upper panel) and HLA-A2 expression in EndoC- β H1 and EndoC- β H1/HLA-A2 cells (lower panel). *F*: EndoC- β H1 and EndoC- β H1/HLA-A2 coculture with DRiP-specific CD8 T cells after incubation with cytokines (IFN γ and IL1 β) for 24 h ($n = 3$ independent experiments). T-cell activation is shown as CD107a staining (left panel) and MIP1 β release (right panel). CTLs, cytotoxic T lymphocytes.

(INNODIA) and no. 945268 (INNODIA HARVEST). This joint undertaking receives support from the European Union's Horizon 2020 research and innovation program and European Federation of Pharmaceutical Industries and Associations (EFPIA), JDRF, and The Leona M. and Harry B. Helmsley Charitable Trust. B.O.R. is supported by the Wanek Family Project for Type 1 Diabetes.

Duality of Interest. No potential conflicts of interest relevant to this article were reported.

Author Contributions. S.T., R.C.S., A.R.v.d.S., A.M.-G., and F.C. performed the experiments and wrote the manuscript. S.T., R.C.S., J.B., F.M., L.M.T.H., and B.K. analyzed the data. R.C.H. and B.O.R. wrote the manuscript. H.M. supervised data analysis. A.Z. supervised the project and wrote the manuscript. A.Z. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

- Coppieters KT, Dotta F, Amiran N, et al. Demonstration of islet-autoreactive CD8 T cells in insulinitic lesions from recent onset and long-term type 1 diabetes patients. *J Exp Med* 2012;209:51–60
- Pinkse GG, Tysma OH, Bergen CA, et al. Autoreactive CD8 T cells associated with beta cell destruction in type 1 diabetes. *Proc Natl Acad Sci USA* 2005;102:18425–18430
- Skowera A, Ellis RJ, Varela-Calviño R, et al. CTLs are targeted to kill beta cells in patients with type 1 diabetes through recognition of a glucose-regulated preproinsulin epitope. *J Clin Invest* 2008;118:3390–3402
- Abreu JR, Martina S, Verrijn Stuart AA, et al. CD8 T cell autoreactivity to preproinsulin epitopes with very low human leucocyte antigen class I binding affinity. *Clin Exp Immunol* 2012;170:57–65
- Eizirik DL, Colli ML, Ortis F. The role of inflammation in insulinitis and beta-cell loss in type 1 diabetes. *Nat Rev Endocrinol* 2009;5:219–226
- O'Sullivan-Murphy B, Urano F. ER stress as a trigger for β -cell dysfunction and autoimmunity in type 1 diabetes. *Diabetes* 2012;61:780–781
- Thomaidou S, Kracht MJL, van der Slik A, et al. β -cell stress shapes CTL immune recognition of preproinsulin signal peptide by posttranscriptional regulation of endoplasmic reticulum aminopeptidase 1. *Diabetes* 2020;69:670–680
- de Jong VM, Zaldumbide A, van der Slik AR, Persengiev SP, Roep BO, Koelman BP. Post-transcriptional control of candidate risk genes for type 1 diabetes by rare genetic variants. *Genes Immun* 2013;14:58–61
- Eizirik DL, Sammeth M, Bouckenoghe T, 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:e1002552
- DeLong T, Wiles TA, Baker RL, et al. Pathogenic CD4 T cells in type 1 diabetes recognize epitopes formed by peptide fusion. *Science* 2016;351:711–714
- McLaughlin RJ, de Haan A, Zaldumbide A, et al. Human islets and dendritic cells generate post-translationally modified islet autoantigens. *Clin Exp Immunol* 2016;185:133–140
- Roep BO, Kracht MJ, van Lummel M, Zaldumbide A. A roadmap of the generation of neoantigens as targets of the immune system in type 1 diabetes. *Curr Opin Immunol* 2016;43:67–73
- Kracht MJL, Zaldumbide A, Roep BO. Neoantigens and microenvironment in Type 1 diabetes: lessons from antitumor immunity. *Trends Endocrinol Metab* 2016;27:353–362
- Gerashchenko MV, Lobanov AV, Gladyshev VN. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc Natl Acad Sci USA* 2012;109:17394–17399
- Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci USA* 2004;101:11269–11274
- Lee S, Liu B, Lee S, Huang SX, Shen B, Qian SB. Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A* 2012;109:E2424–E2432
- Prasad S, Starck SR, Shastri N. Presentation of cryptic peptides by MHC class I is enhanced by inflammatory stimuli. *J Immunol* 2016;197:2981–2991
- Yewdell JW. DRiPs solidify: progress in understanding endogenous MHC class I antigen processing. *Trends Immunol* 2011;32:548–558
- Starck SR, Shastri N. Non-conventional sources of peptides presented by MHC class I. *Cell Mol Life Sci* 2011;68:1471–1479
- Saulquin X, Scotet E, Trautmann L, et al. +1 frameshifting as a novel mechanism to generate a cryptic cytotoxic T lymphocyte epitope derived from human interleukin 10. *J Exp Med* 2002;195:353–358
- Yewdell JW, Antón LC, Bennink JR. Defective ribosomal products (DRiPs): a major source of antigenic peptides for MHC class I molecules? *J Immunol* 1996;157:1823–1826
- Kracht MJ, van Lummel M, Nikolic T, et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat Med* 2017;23:501–507
- Gonzalez-Duque S, Azoury ME, Colli ML, et al. Conventional and neo-antigenic peptides presented by β cells are targeted by circulating naïve CD8+ T cells in type 1 diabetic and healthy donors. *Cell Metab* 2018;28:946–960.e6
- Ravassard P, Hazhouz Y, Pechberty S, et al. A genetically engineered human pancreatic β cell line exhibiting glucose-inducible insulin secretion. *J Clin Invest* 2011;121:3589–3597
- Ingolia NT, Brar GA, Rouskin S, McGeachy AM, Weissman JS. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* 2012;7:1534–1550
- Fresno M, Jiménez A, Vázquez D. Inhibition of translation in eukaryotic systems by harringtonine. *Eur J Biochem* 1977;72:323–330
- Lauria F, Tebaldi T, Bernabò P, Groen EJN, Gillingwater TH, Viero G. riboWaltz: optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* 2018;14:e1006169
- Wan J, Qian SB. TISdb: a database for alternative translation initiation in mammalian cells. *Nucleic Acids Res* 2014;42:D845–D850
- Ingolia NT, Lareau LF, Weissman JS. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 2011;147:789–802
- Ji Z, Song R, Regev A, Struhl K. Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* 2015;4:e08890
- Lu G, Hao X, Chen WH, Mu S. GAAD: a gene and autoimmune disease association database. *Genomics Proteomics Bioinformatics* 2018;16:252–261
- Martens GA, Jiang L, Hellemans KH, et al. Clusters of conserved beta cell marker genes for assessment of beta cell phenotype. *PLoS One* 2011;6:e24134
- Russell MA, Redick SD, Blodgett DM, et al. HLA class II antigen processing and presentation pathway components demonstrated by transcriptome and protein analyses of islet β -cells from donors with type 1 diabetes. *Diabetes* 2019;68:988–1001
- Gerashchenko MV, Lobanov AV, Gladyshev VN. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc Natl Acad Sci U S A* 2012;109:17394–17399
- Jiang Z, Yang J, Dai A, Wang Y, Li W, Xie Z. Ribosome profiling reveals translational regulation of mammalian cells in response to hypoxic stress. *BMC Genomics* 2017;18:638
- Weinzierl AO, Maurer D, Altenberend F, et al. A cryptic vascular endothelial growth factor T-cell epitope: identification and characterization by mass spectrometry and T-cell assays. *Cancer Res* 2008;68:2447–2454
- Thomaidou S, Zaldumbide A, Roep BO. Islet stress, degradation and autoimmunity. *Diabetes Obes Metab* 2018;20(Suppl. 2):88–94
- 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
- Casellas A, Mallol C, Salavert A, et al. Insulin-like growth factor 2 overexpression induces β -cell dysfunction and increases beta-cell susceptibility to damage. *J Biol Chem* 2015;290:16772–16785

40. Mustoe AM, Busan S, Rice GM, et al. Pervasive regulatory functions of mRNA structure revealed by high-resolution SHAPE probing. *Cell* 2018;173:181–195.e18
41. Ouspenskaia T, Law T, Clauser KR, et al. Thousands of novel unannotated proteins expand the MHC I immunopeptidome in cancer bioRxiv. 5 March 2020 [preprint]. DOI: 10.1101/2020.02.12.945840v2
42. George AJ, Stark J, Chan C. Understanding specificity and sensitivity of T-cell recognition. *Trends Immunol* 2005;26:653–659
43. Zou Q, Xiao Z, Huang R, et al. Survey of the translation shifts in hepatocellular carcinoma with ribosome profiling. *Theranostics* 2019;9:4141–4155
44. Wei J, Yewdell JW. Immunoribosomes: where's there's fire, there's fire. *Mol Immunol* 2019;113:38–42
45. Wei J, Kishton RJ, Angel M, et al. Ribosomal proteins regulate MHC class I peptide generation for immunosurveillance. *Mol Cell* 2019;73:1162–1173.e5
46. Machkovech HM, Bloom JD, Subramaniam AR. Comprehensive profiling of translation initiation in influenza virus infected cells. *PLoS Pathog* 2019;15:e1007518
47. Starck SR, Jiang V, Pavon-Eternod M, et al. Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* 2012;336:1719–1723
48. Yewdell JW, Holly J. DRiPs get molecular. *Curr Opin Immunol* 2020;64:130–136
49. Schwab SR, Shugart JA, Horng T, Malarkannan S, Shastri N. Unanticipated antigens: translation initiation at CUG with leucine. *PLoS Biol* 2004;2:e366
50. Diaz de Arce AJ, Noderer WL, Wang CL. Complete motif analysis of sequence requirements for translation initiation at non-AUG start codons. *Nucleic Acids Res* 2018;46:985–994
51. Peabody DS. Translation initiation at an ACG triplet in mammalian cells. *J Biol Chem* 1987;262:11847–11851
52. Liang H, Chen X, Yin Q, et al. PTEN β is an alternatively translated isoform of PTEN that regulates rDNA transcription. *Nat Commun* 2017;8:14771
53. Na CH, Barbhuiya MA, Kim MS, et al. Discovery of noncanonical translation initiation sites through mass spectrometric analysis of protein N termini. *Genome Res* 2018;28:25–36
54. Jackson R, Kroehling L, Khitun A, et al. The translation of non-canonical open reading frames controls mucosal immunity. *Nature* 2018;564:434–438
55. Brozzi F, Nardelli TR, Lopes M, et al. Cytokines induce endoplasmic reticulum stress in human, rat and mouse beta cells via different mechanisms. *Diabetologia* 2015;58:2307–2316
56. Coomans de Brachène A, Dos Santos RS, Marroqui L, et al. IFN- α induces a preferential long-lasting expression of MHC class I in human pancreatic beta cells. *Diabetologia* 2018;61:636–640
57. Marroqui L, Dos Santos RS, Op de Beeck A, et al. Interferon- α mediates human beta cell HLA class I overexpression, endoplasmic reticulum stress and apoptosis, three hallmarks of early human type 1 diabetes. *Diabetologia* 2017;60:656–667
58. Tsonkova VG, Sand FW, Wolf XA, et al. The EndoC- β H1 cell line is a valid model of human beta cells and applicable for screenings to identify novel drug target candidates. *Mol Metab* 2018;8:144–157
59. Lawlor N, Márquez EJ, Orchard P, et al.; NIH Intramural Sequencing Center. Multiomic profiling identifies cis-regulatory networks underlying human pancreatic β cell identity and function. *Cell Rep* 2019;26:788–801.e6
60. Oleson BJ, McGraw JA, Broniowska KA, et al. Distinct differences in the responses of the human pancreatic β -cell line EndoC- β H1 and human islets to proinflammatory cytokines. *Am J Physiol Regul Integr Comp Physiol* 2015;309:R525–R534
61. Kearse MG, Wilusz JE. Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev* 2017;31:1717–1731