A putative long noncoding RNA-encoded micropeptide maintains cellular homeostasis in pancreatic β-cells

Mark Li[1,2, 3#], Fan Shao[4,6#], Qingwen Qian[1,2, 3], Wenjie Yu[3,7], Zeyuan Zhang[1,2, 3], Biyi Chen[2,7], Dan Su[1,4], Yuwei Guo[4], An-Vi Phan[4], Long-sheng Song[2,7], Samuel B. Stephens[2,3,7], Julien Sebag[2,3,8], Yumi Imai[2,3,7], Ling Yang[1,2,3*], Huojun Cao[2,4,5,6*]

[1]Department of Anatomy and Cell Biology; [2]Fraternal Order of Eagles Diabetes Research Center; [3]Pappajohn Biomedical Institute; [4]Iowa Institute for Oral Health Research, [5]Division of Biostatistics and Computational Biology, [6]Department of Endodontics; [7]Department of Internal Medicine; [8]Department of Molecular Physiology and Biophysics, University of Iowa Carver College of Medicine (1,2,3,7,8) and University of Iowa College of Dentistry (4,5,6), Iowa City, IA 52242

Co-first author

* Co-responding author

*Address all correspondence to:

Huojun Cao, Ph.D.

Iowa Institute for Oral Health Research

Division of Biostatistics and Computational Biology

Department of Endodontics

University of Iowa College of Dentistry, Iowa City, 52242

Email: huojun-cao@uiowa.edu

Or *Address all correspondence to:

Ling Yang, Ph.D.

Department of Anatomy and Cell Biology

Fraternal Order of Eagles Diabetes Research Center

Pappajohn Biomedical Institute

University of Iowa Carver College of Medicine, Iowa City, IA 52242

Email: ling-yang@uiowa.edu

ABSTRACT

Micropeptides (microproteins) encoded by transcripts previously annotated as long noncoding RNA (lncRNAs) are emerging as important mediators of fundamental biological processes in health and disease. Here we applied two computational tools to identify putative micropeptides encoded by lncRNAs that are expressed in the human pancreas. We experimentally verified one such micropeptide encoded by a β-cell- and neural cell-enriched lncRNA *TUNAR (also known as TUNA, HI-LNC78 or LINC00617)*. We named this highly conserved 48-amino-acid micropeptide Beta cell- and Neural cell-regulin (BNLN). BNLN contains a single-pass transmembrane domain and localized at the endoplasmic reticulum in pancreatic β-cells. Overexpression of BNLN lowered ER calcium levels, increased cytosolic calcium levels, and maintained ER homeostasis in response to high glucose challenge. To determine the physiological and pathological roles of BNLN, we assessed the BNLN expression in islets from mice fed with a high-fat diet and a regular diet, and found that *BNLN* is suppressed by diet-induced obesity (DIO). Conversely, overexpression of BNLN elevated glucose-stimulated insulin secretion in INS-1 cells. Lastly, BNLN overexpression enhanced insulin secretion in islets from lean and obese mice as well as from humans. Taken together, our study provides the first evidence that lncRNA-encoded micropeptides play a critical role in pancreatic β-cell function and provides a foundation for future comprehensive analyses of micropeptide function and pathophysiological impact on diabetes.

Abbreviations:

ATF6: activating transcription factor-6; BNLN: beta cell and neuron regulin; Cre: Cre recombinase; DIO: diet-induced obesity; ER: endoplasmic reticulum; GAPDH: glyceraldehyde-3-phosphate dehydrogenase; GFP: green fluorescent protein; Glc: glucose; HFD: high-fat diet; lncRNA: long noncoding RNA; RD: regular diet; sORF: small open reading frame; T2D: type 2 diabetes; TUNAR: TCL1 Upstream Neural Differentiation-Associated RNA.

26 **INTRODUCTION:**

27 Large-scale projects such as ENCODE and FANTOM have found thousands of long noncoding RNAs
28 (lncRNAs) in the human genome[1-4]. lncRNAs play an essential role in regulating islet function in health and
29 diseases including diabetes[5]. It has been demonstrated that lncRNAs are key components of the islet
30 regulatome, with more than 1000 lncRNAs identified in human islets and mouse β-cells[6,7]. These lncRNAs
31 play critical roles in islet and β-cell survival, cell development, and physiological function through regulation
32 of the islet transcriptome[5-7]. Of note, a genome-wide association study that examined T2D susceptibility
33 loci demonstrated that dysregulation of lncRNAs is involved in pancreatic pathologies in humans[6]. In
34 addition, dysregulation of lncRNAs has been implicated in diabetes complications, such as diabetic
35 nephropathy and retinopathy[5,8]. Recently, several circulating lncRNAs have been identified as biomarkers
36 of diabetes and diabetic complications[9,10]. However, although much knowledge has accumulated
37 correlating dysregulation of lncRNA with islet pathology, the functional roles and molecular mechanisms of
38 these lncRNAs remain poorly understood.

39

40 By definition, lncRNAs do not contain protein-coding open reading frames (ORF), which comprise a start
41 codon, in-frame codons, and a stop codon. To accurately identify *bona fide* protein-coding ORFs, most
42 ORF-finding computational algorithms have historically used a cutoff of 100 amino acids as the minimum
43 size for detection. This filtering strategy dramatically reduces the likelihood of false-positive detection, but
44 may also result in some genes that contain small open reading frames (sORFs) being mis-annotated as
45 lncRNAs[11]. Indeed, recent studies have revealed that some of lncRNAs do contain sORFs that encode
46 micropeptides (or microprotein, protein less than 100 amino acids in length)[12-20]. There is emerging
47 evidence that micropeptides are important regulators in fundamental biological processes such as
48 embryonic development[21] and metabolism[22,23]. However, whether and how micropeptides modulate islet
49 biology is largely unknown.

50

51 In this study, we employed two computational algorithms, PhyloCSF and RNAcode, to comprehensively
52 survey the protein-coding potential of lncRNAs expressed in the human pancreatic islet[24,25]. Among the

53 novel micropeptides encoding lncRNAs transcripts we identified, we characterized one putative
54 micropeptide encoded by *TUNAR (also known as TUNA, HI-LNC78 or LINC00617),* which is expressed
55 highly in human β-cells and neural cells. We then named this highly conserved micropeptide beta cell
56 and neural cell regulin (*BNLN*). Pancreatic β-cells synthesize and secrete insulin to govern systemic
57 glucose homeostasis[26]. We found BNLN localizes at the endoplasmic reticulum (ER) in pancreatic β-cells.
58 Robust and highly functional endoplasmic reticulum (ER) dynamics are required for β-cells to synthesize
59 insulin to maintain systemic glucose metabolic homeostasis[27]. The ER also plays an important role in
60 coordinating intracellular Ca^{2+} signaling which regulates insulin secretion[28,29]. In both rodent models and
61 humans, obesity and diabetes are characterized by ER dysfunction which leads to β-cell death and
62 impaired islet function[30,31]. We further showed that *BNLN* is downregulated in islets from obese mice, and
63 gain of function of *BNLN* increases glucose-induced insulin secretion in islets from obese mice as well as
64 from human donors. These findings provide first evidence of the biological function of pancreatic
65 micropeptides in health and disease.

66

67 **RESULTS**

68 **An lncRNA, *TUNAR*, encodes a conserved 48-amino-acid micropeptide in the human pancreas**
69
70 Recent studies have discovered that a few lncRNAs harbor small open reading frames (sORFs) that
71 encode sORF-encoded peptides (SEPs) or micropeptides[12-20]. We hypothesized that there are additional
72 transcripts currently annotated as lncRNAs that may encode micropeptides. To prioritize the search for
73 potential micropeptide genes, we used two computational tools, PhyloCSF and RNAcode, to
74 comprehensively evaluate the coding potential of lncRNAs that are expressed in human pancreatic
75 islets[24,25]. We found several lncRNAs that are highly expressed in human pancreatic islets and contained
76 putative sORFs (Fig. 1A). TCL1 Upstream Neural Differentiation-Associated RNA, *TUNAR (also known as*
77 *TUNA, HI-LNC78 or LINC00617)* ranked at the top in both the expression level in human islets and the
78 coding potential (Fig. 1A). Of note, its PhyloCSF and RNAcode scores were higher than many recently
79 identified micropeptides (Fig. 1B). In addition to the pancreas, *TUNAR* is highly expressed in the brain,
80 pituitary, testis, fallopian tube and the uterus (Supplemental Fig. 1). Knockdown of *TUNAR* has been

81 associated with reduced insulin content and impaired glucose-stimulated insulin secretion in a human β-
82 cell line, as well as impaired neural differentiation of mouse embryonic stem cells (mESCs)[32,33]. Genomic
83 annotation indicated that *TUNAR* spans ~ 49.2kb and has 3 transcript isoforms: NR_132399.1,
84 NR_132398.1 and NR_038861.1 (Fig. 1C). Further analysis of RNA-Seq read splice junctions visualized
85 with sashimi plot revealed that NR_038861.1 is the major isoform in the pancreas (data not shown). Our
86 computational pipeline predicted an sORF at the beginning of the last exon (which is shared among all
87 three transcript isoforms) that is highly conserved among 100 vertebrate genomes assessed by UCSC
88 genome browser PhyloP track (Fig. 1C). Moreover, ribosomal profiling (Ribo-seq) datasets suggest that
89 this region is actively translated (Supplemental Fig. 2A). We then extracted other species' DNA sequences
90 (multiz100way) corresponding to predicted sORF and translated them to protein sequences. The predicted
91 micropeptide sequence was highly conserved (Fig. 1D, Supplemental Fig. 3 and Supplemental Table1)
92 and could be found in 88 species out of the 100 vertebrate species (conserved down to fishes).
93
94 To experimentally validate the translation of the predicted sORF within *TUNAR*, we inserted three reporters
95 (GFP, hRluc and FLAG) CDS without ATG start codon either in-frame or out-of-frame before the stop codon
96 of the full length *TUNAR* transcript (NR_038861.1; Fig. 2A). As expected, GFP and Renilla luciferase were
97 translated when inserted in-frame (Fig.2B&C). However, no translation was observed when GFP or hRluc
98 were inserted out-of-frame by deletion of one or two bases between the predicted sORF and GFP/hRluc
99 (Fig. 2B&C). Consistent with these results, western blot analysis showed a band ~ 10-kDa only when FLAG
00 was inserted in-frame (Fig. 2D). There are 3 potential ATG initiating codons in-frame with the predicted
01 reading frame in *TUNAR* (Fig. 1D). Multispecies alignment showed that more than half of all the species
02 contain the 1st ATG, all 88 species have the 2nd, and 87 species have 3rd ATG of *TUNAR* (Supplemental
03 Fig. 3). Therefore, we generated ATG deletion constructs for each of these potential initiating codons and
04 found that removing the 2nd ATG abolished translation of *TUNAR* (Fig. 2E&F). Hence, *TUNAR* contains an
05 sORF from the 2nd ATG start codon to the stop codon that translates to a 48-amino-acid micropeptide. To
06 keep consistency with the previous nomenclature, we named this micropeptide Beta cell- and Neural cell-
07 regulin (BNLN).

08

09 **BNLN localizes at the endoplasmic reticulum (ER) in pancreatic β cells**

10 *BNLN / TUNAR* is highly expressed in the human islet (Fig. 2A). *In situ* hybridization analysis in islets from
11 human donors showed that *BNLN* expression is enriched in the insulin producing β-cells (Fig. 3A), which
12 is in agreement with the analysis of previous RNA-Seq data[32,34]. Our *in silico* sequence analysis revealed
13 that BNLN harbors a single-pass transmembrane domain at its C-terminus (Fig. 1D, Supplemental Fig. 2B).
14 The multispecies alignment showed that this transmembrane domain is extremely conserved and
15 presented in all 88 species (Fig. 1D, Supplemental Fig. 3). To determine the sub-cellular localization of
16 BNLN, we generated N-terminal and C-terminal GFP-FLAG tagged *BNLN* constructs (Fig. 3B), then
17 performed immunofluorescence analysis in INS-1 cells. As shown in figure 3C, in contrast to the GFP
18 control, which was dispersed throughout cells, BNLN mainly localized at the endoplasmic reticulum (ER)
19 in the INS-1 cell. Much weaker signal was detected in lysosomes, mitochondria, peroxisomes, or Golgi in
20 INS-1 cells (Fig. 3D). To further assess whether BNLN is secreted or expressed extracellularly, we
21 generated a BNLN construct fused with HiBiT (Fig. 3E). In the Nano-Glo HiBiT detection system, HiBiT-
22 tagged protein is recognized by the complementing polypeptide LgBiT, the bound complex then
23 reconstitutes a luminescent signal. As shown in figure 3F, BNLN mainly localized intracellularly in INS-1
24 cells. Taken together, these data show that BNLN is primarily located at the ER of pancreatic β-cells.

25

26 **BNLN modulates ER calcium homeostasis and ER function in the β cell**

27 ER calcium homeostasis has a pivotal role in insulin production and secretion in β-cells[26]. To determine
28 the functional effect of BNLN in the ER, we first assessed the ER Ca^{2+} ($[Ca^{2+}]^{ER}$) level by using an ER Ca^{2+}
29 biosensor (ER-GCaMP6-150) in INS-1 cells overexpressing BNLN or the control construct. Overexpression
30 of BNLN lowered $[Ca^{2+}]^{ER}$ level at both the basal (2.5mM) and high glucose (17mM) conditions (Fig. 4A&B),
31 indicating a potential role of BNLN in modulating ER calcium influx. Glucose-stimulated β-cells display
32 oscillations of the interrelated cytosolic Ca^{2+} ($[Ca^{2+}]^{c}$) and $[Ca^{2+}]^{ER}$. To evaluate the influence of the BNLN
33 on $[Ca^{2+}]^{c}$ homeostasis, we generated FLAG- or V5-tagged *BNLN* constructs and measured $[Ca^{2+}]^{c}$ by a
34 FLIPR Tetra system. Overexpression of BNLN increased $[Ca^{2+}]^{c}$ in INS-1 cells under both basal and high

35 glucose conditions (Fig. 4C&D). We observed this effect with both FLAG- and V5-tagged BNLN. To further
36 assess the effect of BNLN on depolarization-induced calcium dynamics, we measured $[Ca^{2+}]^c$ in INS-1 in
37 the absence or presence of KCl. As shown in figure 4E&F while BNLN overexpression increased high
38 glucose-induced $[Ca^{2+}]^c$, it did not alter KCl-mediated $[Ca^{2+}]^c$ flux, indicating a regulatory role of BNLN
39 independent from depolarization dependent $[Ca^{2+}]^c$ flux that is primarily driven by voltage-dependent
40 calcium channel.
41
42 Calcium level is tightly regulated to maintain ER chaperone activity and protein folding[35]. We found that
43 BNLN reduced $[Ca^{2+}]$ in the ER of INS-1 cells (Fig. 4A&B). To determine whether the BNLN-mediated
44 lowering $[Ca^{2+}]^{ER}$ effect alters ER function, we examined the functional impact of BNLN on ER homeostasis,
45 using an ATF6 luminal domain (ATF6LD) *cypridina noctiluca (Cluc)* secretion reporter assay in INS-1 cells.
46 In this assay, cells with abundant ER chaperones retain the ATF6LD-Cluc in the ER and secrete minimal
47 luciferase, while cells with limited ER chaperone availability release the chimeric protein, promoting
48 luciferase secretion into the medium[36]. We found that the high glucose challenge (17mM) led to a
49 decreased ER chaperone availability (increased ATF6LD-cluc secretion) in INS-1 cells, which was
50 ameliorated by BNLN overexpression (decreased ATF6LD-cluc secretion) (Fig. 4G). These data indicate
51 that BNLN-mediated calcium dynamic in the ER did not impair ER function in INS-1 cells. In contrast, BNLN
52 primed the β-cell ER against the high glucose-induced ER stress.
53
54 It has been demonstrated that micropetide exert its functions through interaction with protein
55 complex[12,13,15,20]. To identify potential BNLN interacting protein partners in pancreatic β-cells, we performed
56 a protein pull-down assay followed by mass spectrometry analysis, using lysates from INS-1 cells with
57 BNLN-FLAG overexpression. GFP-FLAG overexpression was used as a control construct. One protein
58 enriched in BNLN-FLAG pull down compared to GFP-FLAG pull-down was sarco-endoplasmic reticulum
59 Ca^{2+}-ATPases 3 (SERCA3; Supplemental Table 2), a major calcium transporter in the ER. We further
60 validated this interaction by Co-IP-western blot analysis in HEK 293 cells with BNLN and SERCA3
61 overexpression (Fig. 4H). To determine the impact of glucose on the interaction of BNLN with SERCA3,

62 we further performed a proximity ligation assay (PLA) in INS-1 cells in the presence of low (2.5mM) or high
63 (17mM) glucose. As shown in figure 4I, a low level of interaction between BNLN and SERCA3 occurred in
64 low glucose conditions (2.5mM), and this was elevated in the presence of high glucose (17mM). Taken
65 together, these data indicate an important role of BNLN on modulating calcium dynamics in the pancreatic
66 β-cells.
67
68 **BNLN enhances glucose-stimulated insulin secretion**
69 Ca^{2+} oscillations in pancreatic β-cells are critical for their response to changes in glucose level[37]. To
70 establish the functional role of BNLN in the pancreatic-β cell, we first measured glucose-stimulated insulin
71 secretion (GSIS) in INS-1 cells with BNLN overexpression using adenovirus-mediated gene delivery. As
72 shown in figure 5A, BNLN overexpression significantly increased GSIS in the INS-1 cells compared to cells
73 overexpressed with control vector. To verify that this effect was mediated by the translated micropeptide
74 and not the RNA transcript, we transfected INS-1 cells with the 2nd ATG-deletion BNLN construct (which
75 does not produce the micropeptide), and found that this construct did not affect GSIS (Fig. 5B). To further
76 examine the BNLN-mediated insulin secretion, we transduced islets isolated from lean mice with the
77 adenovirus-mediated BNLN construct. As shown in figure 5C, GSIS was significantly enhanced in Islets
78 overexpressing BNLN.
79
80 Obesity results in dysregulated GSIS in pancreatic islets. To establish the pathophysiological relevance of
81 BNLN expression, we next assessed the expression level of *BNLN* transcripts in islets from mouse islets
82 from lean mice and mice with diet-induced obesity (DIO). *BNLN* expression was significantly
83 downregulated in islets from obese mice (Fig. 5D). To test whether BNLN could rescue obesity-associated
84 islet dysfunction, we further examined GSIS in lean and obese mouse islets following *BNLN*
85 overexpression. As shown in figure 5E, overexpression of *BNLN* significantly improved GSIS in islets from
86 mice with DIO. Finally, to establish the functional relevance of BNLN in humans, we measured GSIS in
87 human islets with BNLN overexpression (Supplemental Table 3). Overexpression of BNLN significantly

88 increased GSIS in islets from non-diabetic humans with minimal effect on insulin content (Fig. 5F-I). These
89 data strongly implicate that BNLN functions as a small peptide to modulate insulin secretion in β-cells.
90

91 **DISSCUSION**

92 Obesity and diabetes are characterized by dysregulation of insulin secretion from pancreatic β-cells.
93 Although lncRNAs have been implicated in this process[5-7], thus far there is no evidence of regulatory role
94 of lncRNA-encoded micropeptides in pancreatic β-cell function. Our study offers the first insight into the
95 impact of micropeptides on β-cell physiology. We showed that the lncRNA *TUNAR (also known as TUNA,*
96 *HI-LNC78 or LINC00617)* encodes BNLN, a 48aa micropeptide. We further demonstrate that BNLN
97 regulates glucose-induced insulin secretion in β-cells and intact islets in part through modulating ER
98 homeostasis.
99

100 LncRNAs regulate key biological processes through regulation of transcriptional activation,
101 heterochromatin formation, interaction with miRNAs and other mechanisms[38,39]. However, recent studies
102 have revealed that some transcripts identified as lncRNAs actually do contain sORFs that encode
103 micropeptide (microprotein)[12-20]. In this study, we applied two computational programs PhyloCSF and
104 RNAcode to prioritize the search for additional micropeptide encoding lncRNAs expressed in human
105 pancreas. It should be noted that both PhyloCSF and RNAcode depend on the quality of whole genome
106 sequence alignment. False negatives could be resulted from poor or incorrect whole genome sequence
107 alignment. Among the predicted micropeptide encoding lncRNA, we choose one lncRNA, TCL1 Upstream
108 Neural Differentiation-Associated RNA (*TUNAR*), for further characterization because it ranked at the top
109 of both the expression level in human islets and the coding potential. *TUNAR* has been shown to be
110 important for ESC pluripotency and neural differentiation[40,41]. Dysregulation of *TUNAR* has been implicated
111 in the progression of Huntington's Disease[42] and breast cancer[5]. Notably, a recent study showed that
112 knocking down *TUNAR* suppresses insulin secretion in a human β-cell line[5]. Multiple molecular
113 mechanisms have been proposed underlying functions of *TUNAR* in the context of diverse biology
114 processes. For example, it has been suggested that *TUNAR* forms an RNA–multiprotein complex to

activate pluripotency genes in a β-cell line[40]. Our computational and experimental studies identified an sORF encoding BNLN, a 48aa micropeptide within the last exon of *TUNAR*. Overexpression of this micropeptide regulated GSIS, while deletion of start codon (ATG) abolished its function (Fig. 5B). However, we could not completely rule out a role for *TUNAR* lncRNA in this process. Therefore, future studies will focus on dissecting the contributions of TUNAR RNA and BNLN micropeptide on the β-cell transcriptome development and differentiation, and metabolic function.

Accumulating evidence demonstrates that transmembrane micropeptides tightly modulate intracellular and extracellular signaling cascades[43]. We found that BNLN contains a transmembrane domain at the C-terminus. To determine the sub-cellular localization, cellular function, and molecular mechanisms of BNLN in β-cells, we generated a set of epitope-tagged BNLN constructs. It is possible that epitope tags altered the signal sequences and function of BNLN, however, we did not observe functional differences between the differently tagged BNLN constructs (Fig. 4C). Further studies with specific anti-BNLN antibodies will be necessary to confirm subcellular localization of the endogenous micropeptide.

ER calcium homeostasis is required for insulin production and secretion in β-cells[26]. Ca^{2+} serves as a critical co-factor for protein chaperones and foldases needed to support high levels of protein translation and pro-insulin processing[35]. Moreover, the ER of β-cells contributes to a tight regulation of cytosolic Ca^{2+} by taking up Ca^{2+} when the cell is depolarized (cytosolic Ca^{2+} is elevated), and slowly releasing Ca^{2+} when the cell is repolarized[44]. In pancreatic islets, the ER primarily takes up cytosolic Ca^{2+} by two SERCAs: ubiquitously expressed SERCA2b, and SERCA3, which is expressed only in the islet β-cell[45,46]. While SERCA2b regulates basal $[Ca^{2+}]^{c}$[46], it has been suggested that SERCA3 is responsible for replenishing ER Ca^{2+} when the cytosolic Ca^{2+} level is elevated by depolarization[47], thus prolonging the Ca^{2} oscillation. We showed that overexpression of BNLN did not alter KCl-induced $[Ca^{2+}]^{c}$ peak, but decreased $[Ca^{2+}]^{ER}$ and increased $[Ca^{2+}]^{c}$ in response to stimulatory glucose concentrations (Fig. 4E). This is in line with a study from Ravier *et al.*, which showed that SERCA3-null islets exhibited decreased ER Ca^{2+} influx, leading to an increased cytosolic Ca^{2+} oscillatory amplitude and enhanced insulin secretion[48]. Our study

242 mainly relied on INS-1 cells as a surrogate for β-cells, and determining how BNLN modulates islet
243 Ca^{2+} oscillation/kinetics in health and obesity/diabetes requires further studies.
244
245 We found that, in β-cells, BNLN interacts with SERCA3A (Fig. 4H&I), and decreases Ca^{2+} buffer capacity
246 of the ER in response to high glucose (Fig. 4A&B), indicating a potential inhibitory role of BNLN on SERCA3.
247 However, whether this is primarily the result of reduction of SERCA3A enzyme activity requires further
248 investigation. Furthermore, sustained high-glucose elicits ER stress in the β-cell[35]. Disturbed ER
249 homeostasis leads to impaired ER Ca^{2+} efflux, Ca^{2+}-induced Ca^{2+} release (CICR), as well as dysfunction
250 of insulin production in the β-cell. We found that the high glucose-induced ER stress was ameliorated by
251 BNLN overexpression, while BNLN decreased ER $Ca^{2"}$ level. Further study is needed to comprehensively
252 delineate the impact of BNLN on ER homeostasis and ER $Ca^{2"}$ leakage in β-cells under diverse ER stress
253 conditions. Notably, glucose and amino acids promote ER Ca^{2+} uptake in the absence of external Ca^{2+} [37,49],
254 suggesting that cytosolic Ca^{2+}-binding proteins and other intracellular Ca^{2+} storage represent an important
255 source of Ca^{2+} that can be pumped by the ER. Future study will be required to establish the functional
256 impact of BNLN on other protein binding partners (Supplemental Table 2).
257
258 In summary, our study demonstrates that *TUNAR,* a previously annotated lncRNA, encodes a conserved
259 48aa micropeptide BNLN, which modulates Ca^{2+} homeostasis in β-cells, consequently regulating GSIS.
260 This study indicates that dysregulation of BNLN in the pancreas might contribute to the obesity-associated
261 malfunction of insulin secretion. Our study provides important insights into the molecular mechanisms that
262 underlie the fine-tuning of glucose-induced insulin secretion by translational function of lncRNA. In addition,
263 our study will also have a broader impact on understanding of the interactions between lncRNA biology
264 and dysfunction of islets relevant to obesity and type 2 diabetes. It remains an intriguing possibility that
265 insights from this study could be exploited for therapeutic interventions in obesity and type 2 diabetes.
266
267 **AUTHOR CONTRIBUTIONS**

M.L. and F.S performed the experiments, analyzed the data. QW.Q, WJ.Y., ZY.Z., BY.C, D.S., YW.G., and AV.P. performed the experiments. J.S., Y.I, S.S. and LS.S. provided critical reagents and scientific suggestions on this study. HJ.C and L.Y conceived and supervised the study and wrote the manuscript. L.Y. is supported by NIH (DK108835-01A1) and American Diabetes Association (1-18-IBS-149); Z.Z. is supported by American Heart Association (19PRE34380258).

ACKNOWLEDGEMENT

We are grateful to the Alberta Diabetes Institute Islet Core (Canada) for providing us the human islets, and the contributions of the donors whose islets were used, as well as the physicians, nurses, and researchers who procured the specimens. We also thank Dr. Michael Welsh (University of Iowa) for providing technical supports in RNA imaging; Dr. Monica Nagendran and Dr. Tushar Desai (Stanford University School of Medicine) for sharing the Proximity ligation in situ hybridization protocol with Welsh lab. We thank Dr. Timothy Ryan (Weill Cornell Medicine) for donating the ER-GCaMP6-15 to Addgene and Alyona Li for technical support with graphic illustrations. We would like thank members of the Yang, Cao, and Amendt laboratories for helpful discussion and suggestions.

MATERIAL AND METHODS

<u>Prioritize putative micropeptide encoding lncRNAs in human pancreas</u>

RNA-seq data of human pancreas islets were downloaded from EBI arrayexpress database (E-MTAB-1294[34]). RNA-seq reads were quality checked using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Low-quality and adapter sequences were trimmed using the Trimmomatic[50]. Expression of transcripts was quantified using the Salmon[51], and estimates of transcript abundance for gene-level analysis were imported and summarized using the tximport[52] function of the R/Bioconductor software suite[53]. Transcripts that have TPM above 1 were evaluated for their coding potential with PhyloCSF and RNAcode. We downloaded 100-species whole genome alignments (multiz100way) from the UCSC database. A custom python script was used to extract

294 and prepare multispecies DNA sequence alignments for each transcript of protein coding (<100aa) genes
295 and lncRNAs. We filtered for transcripts that had at least 2 species that have more than 60-nucleotide (nt)
296 DNA sequence in the multispecies alignments. PhyloCSF and RNAcode were run with these multispecies
297 DNA sequence alignments as input. To prioritize search of putative micropeptide encoded by lncRNAs, we
298 ranked genes with both expression level in human islets and coding potential.
299

300 Animals

301 Animal care and experimental procedures were performed with approval from the University of Iowa's
302 Institutional Animal Care and Use Committee. C57BL/6J mice (The Jackson Laboratory, 000664) were
303 kept on a 12 hr light/dark cycle, and were fed the RD (7319 Teklad global diet). Mice used in generating
304 the DIO model were placed on a 60% kcal high-fat diet (HFD; Research Diets, D12492) immediately after
305 weaning, at 3 weeks of age.
306

307 Human islets

308 Human islet study was reviewed by IRB at University of Iowa and deemed non-human study. Human islets
309 were received from the Alberta Diabetes Institute Islet Core with reported viability and purity above 80%.
310 Islet were cultured in CMRL1066 (Thermo Fisher Scientific, 11530037) medium with 1% human serum
311 albumin, 1% Pen-Strep (Thermo Fisher Scientific, 15140163), and 1% L-Glutamine (Thermo Fisher
312 Scientific, 25030164) overnight at 37°C upon arrival for recovery upon shipping. The eight donors were:
313 age, 20-69 years; BMI, 19-33.9; Hba1c, 4.9-5.8; males; cause of death, neurological.
314

315 Cell culture, transfection and Luciferase Assay

316 The rat insulinoma cells (INS-1, 832/3)[54] were cultured in RPMI-1640 medium (Thermo Fisher Scientific,
317 11875119) supplemented with 10% FBS, 10mmol/L HEPES (Thermo Fisher Scientific, 15630130),
318 2mmol/L L-glutamine, 1mmol/L sodium pyruvate (Sigma-Aldrich, P2256), and 50μmol/L 2-mercaptoethanol
319 (VWR, 97064-880). Cells were transfected with DNA constructs (1ug/well in 12-well plate) using
320 Lipofectamine 2000 reagent (Invitrogen, 11668019) in Opti-MEM medium (Thermo Fisher Scientific,

21 31985088). At 48 hr post transfection, the transfected cells were trypsinized (Trypsin-EDTA 0.05%; Thermo
22 Fisher Scientific, 25300054), mixed, and re-plated in an appropriate plate in order to achieve homogeneous
23 population across wells subsequent analyses. The HEK 293T cells (ATCC) were cultured in DMEM
24 medium with 10% FBS, 1% penicillin-streptomycin. Cells were transfected with DNA constructs (0.5 ug/well
25 of 24-well plate) using Lipofectamine 3000 reagent (Thermo Fisher, L3000015) in Opti-MEM medium
26 (Thermo Fisher Scientific, 31985-062). At 48 hr post transfection, the activities of firefly luciferase and
27 Renilla luciferase were measured using the Luciferase Assay System (Promega, E1500) and Renilla
28 Luciferase Assay System (Promega, E2810), separately.
29
30 Constructs, adenovirus preparation and transduction
31 The BNLN constructs were generated by Sequence- and ligation-independent cloning (SLIC) method[55].
32 The detailed information of constructs used in each figure is list in supplementary Table2. Briefly, tRFP and
33 Renilla luciferase-GFP expressing sequences were fused to C-terminus of human BNLN (in-frame and out-
34 of-frame; fs1, fs2). The recombined fragments were then inserted into pCAGEN vector (Addgene, 11179).
35 To determine transcription initial site and BNLN localization, full length human BNLN, ATGs- and CTG-
36 deleted BNLN (dATG1, dATG2, dATG3, dCTG) were fused with Flag and cloned into pcDNA3.1(-)
37 (Invitrogen, V79520) vector. KDEL plasmid was obtained from Addgene (114177). The HiBiT-BNLN and
38 V5-BNLN constructs were generated by fusing HiBiT vector (Promega) and V5 sequence, respectively, in-
39 frame with N-terminus of BNLN. The fused fragments were then inserted into pQCXIP vector (Addgene,
40 631516), by using NEBuilder® HiFi DNA Assembly Cloning Kit (New England Biolabs, E5520S).
41 ER-GCaMP6-15[56] was provided by Addgene. Adenoviruses carrying mouse *TUNA BNLN* were generated
42 using the ViraPowerTM adenoviral expression system (Invitrogen). Adenoviruses were amplified in
43 HEK293A cells (ATCC) and purified by CsCl gradient centrifugation. The viruses were titered using Adeno-
44 X Rapid Titer Kit (Takaba Bio, 632250) and transduced into the cells as previously described [57]. Cultured
45 intact mouse or human islets were resuspended in 0.1mmol/L EGTA (Sigma-Adlrich, 234626) in serum
46 free CRMLfollowed by transduction of indicated adenoviral constructs at 10,000 plaque-forming unit per
47 islet equivalent (pfu/IEQ) for 1hr at room temperature before transferring to CMRL medium with 10% heat-

inactivated FBS for culture overnight[]. The culture medium was replaced next day and the islets were cultured for additional 24hr before subjecting them to glucose stimulated insulin secretion (GSIS) assay.

Nano-Glo HiBiT extracellular detection system

Ins-1 cells were transfected with FLAG-GFP or HiBiT-BNLN as described above and seeded in an opaque 96-well plate(Fisher Scientific, 08-771-26) at ~15,000 cells/well in the culture medium. The experiment was carried out according to the manufacturer's protocol (Promega, N2422). Briefly, the transfected cells were incubated with Nano-Glo HiBiT extracellular buffer, LgBiT protein, and Nano-Glo HiBiT extracellular substrate, and luminescence was detected after 5-10min using a luminescence microplate reader (SpectraMax, Molecular Devices, USA). The expression of extracellular HiBiT-BNLN was normalized to the total intracellular pool of HiBiT-BNLN by lysing the cells and measuring luminescence with Nano-Glo HiBiT lytic detection system (Promega, N2422), following the protocol.

Western Blotting and immunoprecipitation

Cells were lysed with a lysis buffer containing 0.4% of NP-40, protease inhibitors cocktail (Sigma-Aldrich, P8340), and Na_3VPO_4 (Sigma-Aldrich). Protein concentration was determined by Pierce BCA kit (Thermo Fisher Scientific, 23225) and samples were subjected to SDS–polyacrylamide gel electrophoresis, as previously described (48). For immunoprecipitation assay, cells were lysed in radioimmunoprecipitation assay (RIPA) buffer (50mM Tris, 150mM NaCl, 1% NP40, 0.25% sodium deoxycholate). 1mg proteins were immunoprecipitated with protein A/G magnetic Beads (ThermoFisher, 88802) conjugated with anti-IgG (Cell Signaling Technology, 2729), anti-Flag (Thermo Fisher, A36797) and eluted by 4× lituium dodecyl sulfate (LDS; Invitrogen, NP0008). Membranes were immunoblotted with anti-V5 (ab15528), anti-FLAG and anti-beta Actin (ab8227) at a 1:1000 dilution. Secondary antibodies were horseradish peroxidase-conjugated goat-anti-mouse-IgG (Santa Cruz Biotechnology, sc-2005), horseradish peroxidase-conjugated mouse-anti-rabbit-IgG (Santa Cruz Biotechnology, sc-2357) at a 1:10000 dilution. Signal was detected using the ChemiDoc Touch Imaging System (BioRad), and densitometry analyses of western blot images were performed by using the Image Lab software (BioRad).

Proximity ligation *in situ* hybridization (PLISH)

In situ hybridization was performed as previously described[58] with some modifications. Briefly, human islets were fixed and sectioned following the typical formalin-fixed paraffin embedded tissue processing protocol. The sections were then incubated in citrate-based target unmasking solution (Vector laboratories, H-3300) with 0.05% lithium dodecyl sulfate (Sigma, L4632) at 65°C for 30 min. The slides were then incubated with 0.05 mg/ml pepsin (Sigma-Aldrich, P6887) in 0.1M HCl for 10 min at 37 °C, and processed with the typical proximity ligation *in situ* hybridization[58]. RNA sequences that were targeted by short-paired hybridization probes were: Human INS-1: CTGGTGGAAGCTCTCTACCTAGTGTGCGGGGAACGAGGCT;Human INS-2: GCTGGAGAACTACTGCAACTAGACGCAGCCCGCAGGCAGC; HumanGCG-1: TGGACT CCAGGCGTGCCCAAGATTTTGTGCAGTGGTTGAT; Human GCG-2: AGATGAACACCATTCTTGATAA TCTTGCCGCCAGGGACTT; Human GCG-3: TCTTCACAACATCACCTGCTAGCCACGTGGGATGTTTG AA; Human BNLN: TCTTGGCAATGCTGGGGATTATCGGGACCATTCTGAACCT; Human BNLN: AGCA GGCTTGACCCGCACATACCACCCAATCAAATGCACC'; Human BNLN: ATTCCAGATCGCTGACAGAT ATCACATATTTGAAAAGATG; Human BNLN: CTTCACTTGACGAGCTATTTAGTGGAAAAACCACAGG CGC; Human BNLN: GCAGACCTTAGATGCACCCTATCTTTACTGAGAATTATGC; Human BNLN: TTTGATTTTAGCGGTCATGTACCGCGAGAGTTGGGAAGAA.

Immunocytochemistry and Confocal microscopy

INS-1 cells were then fixed in 4% paraformaldehyde for 10min at room temperature, blocked in 5% goat serum, incubated with Hoechst33342 (Thermo Fisher Scientific, H3570), and mounted with a glass coverslip. For subcellular organelle localization, Lystroacker™ Red DND-99 (Thermo Fisher Scientific, L7528), Mitotracker™ Red CMXRos (Thermo Fisher Scientific, M7512), ACAA1 antibody (Gene Tex, GTX114229), RFP-KDEL plasmid, and TGN38 antibody (Novus, NB300-575) were used to detect lysosomes, mitochondria, peroxisomes, the ER, and Golgi respectively. Images were acquired using a Zeiss 700-point scanning confocal microscope with a 63x lens with oil (Zeiss, Oberkochen, Germany). Images were analyzed by Fiji software. During analysis, single Z-stacks were processed and images from

different channels were merged. For PLISH assays, stack images were taken using the 60x lens of Olympus microscopy (FV3000, confocal laser scanning microscope).

Quantitative real-time RT-PCR

Total RNA was isolated using the Trizol reagent (Invitrogen, 15596026) and reverse transcribed into cDNA using a PrimeScript RT RT Master Mix (Takara RR036A). Quantitative real-time RT-PCR analysis was performed using TB Green Premix Ex Taq (Clontech, RR420A) and following primers: *BNLN*-F: 5'-TTC TCC TCG CCT TCC TGC-3';*BNLN*-R: 5'-CAT CTT GGT TGC AAA TGT CT-3'; *ACTB*-F: 5'-CTC TTC CAG CCT TCC TTC-3'; *ACTB*-R: 5'-ATC TCC TTC TGC ATC CTG TC-3'; *Hprt*-F: 5'-CAG TCC CAG CGT CGT GAT TA-3'; *Hprt*-R: 5'-GGC CTC CCA TCT CCT TCA TG-3'; *Bnln*-F: 5'-GGA GAA TGA GGC AGG AAC CC-3'; *Bnln*-R: 5'ACA CGC TCT CTT CCT TGC TC-3'.

Calcium imaging

$[Ca^{2+}]^c$ measurement: INS-1 832/3 cells were transfected with FLAG-GFP or FLAG-GFP-BNLN. 24hr after transfection, the Ins-1 cells were plated at ~25,000 cells/well in a HTB9-coated Greiner Sensoplate™ glass bottom 96-well plate (Millipore Sigma). Prior to imaging, the cells were loaded with Screen Quest™ Fluo-8 no wash calcium red cye (AAT Bioquest, 36314) indicator in the DMEM (Thermo Fisher Scientific, A1443001) culture medium supplemented with 2.5mM glucose and 0.1% BSA for 1hr at 37°C. A fluorescent microplate reader (CLARIOstar PLUS, BMG Labtech, Germany) was used to inject glucose (final concentration of 20mM) to cells, and to acquire fluorescent signal every 0.6 sec for 10min.

For measurement of $[Ca^{2+}]^c$ using Fluorescent Imaging Plate Reader(R) (FLIPR) Tetra® System (Molecular Devices, USA), INS-1 cells were transfected with pQCXIP, FLAG-BNLN, or V5-BNLN (1ug DNA/well) in a 12-well plate (3 wells/each DNA construct) with Lipofectamine 2000, as described above. A suspension of Ins-1 cells was pipetted in a 384-well culture black microplate with optic bottom plate (Greiner-Bio, 781090) at ~17,000 cells/well in 25uL of the low glucose DMEM culture medium. 25uL of Fluo-8 calcium dye was used added to the cell suspension. The plate was incubated at 37° for 30min, followed by another 30min incubation at room temperature in dark. Meanwhile, a 384-well compound plate (Costar, 3657) was loaded

with low glucose (27.5mM), high glucose (181.5mM), and potassium chloride as a positive control (220mM). The FLIPR was utilized to transfer the solutions from the compound plate into the black plate and to read the Ca^{2+} signal. The basal signal was measured for 20 seconds every second, then the solutions were added (5uL/each), and the response signal was read for 10min every 0.8sec.

$[Ca^{2+}]^{ER}$ measurement: INS-1 cells were co-transfected with ER-GCaMP6s and pQCXIP or V5-BNLN using Lipofectamine 2000, as described above. Next day, the cells were plated at ~25,000 cells/well in a HTB-9 coated plate, as described above. 24hours later, the cells were subjected to treatment with DMEM containing 2.5mM glucose at 37° for 1hr. Basal signal was acquired in the GFP excitation/emission spectrum using the fluorescent microplate reader as described above every 0.8sec for 1min. Then a high glucose solution (17mM final concentration) was injected and the fluorescent signal was collected for 5min.

Murine islet isolation

The islets were isolated as previously described[59]. Briefly, lean C57BL/6J male mice (12-20 weeks of age) and mice with diet-induced obesity (12-16 weeks of HFD) were anesthetized with isoflurane, and the pancreas was perfused with HBSS containing type V collagenase (0.8mg/mL) via the common bile duct. The pancreas was then removed and digested at 37°C for 10-12 min with physical agitation to release islets. Islets were further washed with RPMI-1640 with 1% FBS and purified on Histopaque 1077 and 1119 gradients Sigma-Aldrich, 10771, 11191). The islets were then collected, washed, and cultured in RPMI-1640 supplemented with 10% FBS, 1% pen-strep, and 1% HEPES in 30mm dishes at 37°C.

Glucose stimulated insulin secretion (GSIS)

For GSIS in murine islets, 30-40 islets of similar size were picked and placed in separate wells in a plastic 12-well plate (Fisher Scientific, 07-200-82). The islets were rinsed with 1mL of Secretion Assay Buffer (SAB) adjusted to pH7.2, containing 0.114mol/L sodium chloride (Sigma Aldrich, S9888), 0.47mmol/L potassium chloride (Sigma Aldrich, P9333), 0.12mmol/L potassium dihydrogen phosphate (Sigma Aldrich, 1.04871), 0.116mmol/L magnesium sulfate (Sigma Aldrich, M2643), 20mmol/L HEPES (Thermo Fisher Scientific, 15630130), 2.5mmol/L calcium chloride (Sigma Aldrich, 449709), 0.2% bovine serum albumin, and 0.2%

56 sodium bicarbonate (Sigma Aldrich, S5761)[59]. The islets were then equilibrated in SAB containing low
57 glucose (2.5mmol/L) for 1hr at 37°C. Next, the islets were incubated in fresh SAB with low glucose for
58 additional 1-2 hours at 37°C. The supernatant was collected for measuring basal insulin secretion. The
59 islets were then incubated with SAB containing high glucose (16.7mmol/L) for 1-2 hours at 37°C. The
60 supernatant was collected for measuring glucose-stimulated insulin secretion. The islets were lysed in
61 acidified ethanol and insulin content was measured. Secreted insulin and insulin content were analyzed
62 with STELLUX® Chemi Rodent Insulin ELISA (ALPCO, 80-INSMR-CH01), according to the manufacturer's
63 protocol.
64
65 *For GSIS in Human islets*, 30-40 islets of similar size were picked and subjected to GSIS similar to murine
66 islets. Secreted insulin and insulin content were analyzed with STELLUX® Chemi Human Insulin ELISA
67 (ALPCO, 80-INSHU-CH01), according to the manufacturer's protocol.
68
69 *GSIS in Ins-1 cells:* 100% confluent cells were rinsed with SAB and equilibrated with low-glucose-SAB for
70 1-2 hours at 37°C. Fresh SAB containing low glucose was added to the cells for additional 1-2 hours at
71 37°C. The supernatant was then collected for measuring basal insulin secretion. SAB containing high
72 glucose was added to the cells and incubated for additional 1-2 hours at 37°C before supernatant collection.
73 The cells were lysed and protein was extracted for insulin content. Secreted insulin and insulin content
74 were analyzed with STELLUX® Chemi Rodent Insulin ELISA (ALPCO), according to the manufacturer's
75 protocol.
76
77 Statistics
78 The data are presented as mean ± standard error of the mean or standard deviation as noted; *n* represents
79 the number of individual experiments; *N* represents the number of human donors or mice. Statistical
80 differences of numeric parameters between two groups were determined with Student's *t*-test. For
81 statistical analysis of differences between multiple groups, one way and two-way ANOVA were applied,
82 followed by post-hoc Tukey's test. The statistical analyses were performed in Prism.

REFERENCES

1. Forrest, A.R.R., *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462-470 (2014).
2. Davis, C.A., *et al.* The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* **46**, D794-D801 (2018).
3. Dunham, I., *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).
4. Hon, C.-C., *et al.* An atlas of human long non-coding RNAs with accurate 5′ ends. *Nature* **543**, 199-204 (2017).
5. Akerman, I., *et al.* Human Pancreatic beta Cell lncRNAs Control Cell-Specific Regulatory Networks. *Cell Metab* **25**, 400-411 (2017).
6. Moran, I., *et al.* Human beta cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. *Cell Metab* **16**, 435-448 (2012).
7. Wong, W.K.M., Sorensen, A.E., Joglekar, M.V., Hardikar, A.A. & Dalgaard, L.T. Non-Coding RNA in Pancreas and beta-Cell Development. *Noncoding RNA* **4**(2018).
8. Alvarez, M.L. & DiStefano, J.K. Functional characterization of the plasmacytoma variant translocation 1 gene (PVT1) in diabetic nephropathy. *PLoS One* **6**, e18671 (2011).
9. Carter, G., *et al.* Circulating long noncoding RNA GAS5 levels are correlated to prevalence of type 2 diabetes mellitus. *BBA Clin* **4**, 102-107 (2015).
10. de Gonzalo-Calvo, D., *et al.* Circulating long-non coding RNAs as biomarkers of left ventricular diastolic function and remodelling in patients with well-controlled type 2 diabetes. *Sci Rep* **6**, 37354 (2016).
11. Makarewich, C.A. & Olson, E.N. Mining for Micropeptides. *Trends in Cell Biology* **27**, 685-696 (2017).
12. Anderson, D.M., *et al.* A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance. *Cell* **160**, 595-606 (2015).
13. Anderson, D.M., *et al.* Widespread control of calcium signaling by a family of SERCA-inhibiting micropeptides. *Sci. Signal.* **9**, ra119-ra119 (2016).
14. Cai, B., *et al.* LncRNA-Six1 Encodes a Micropeptide to Activate Six1 in Cis and Is Involved in Cell Proliferation and Muscle Growth. *Front. Physiol.* **8**(2017).
15. D'Lima, N.G., *et al.* A human microprotein that interacts with the mRNA decapping complex. *Nature Chemical Biology* **13**, 174-180 (2017).
16. Makarewich, C.A., *et al.* MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid β-Oxidation. *Cell Reports* **23**, 3701-3709 (2018).
17. Matsumoto, A., *et al.* mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* **541**, 228-232 (2017).
18. Nelson, B.R., *et al.* A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* **351**, 271-275 (2016).
19. Pauli, A., *et al.* Toddler: An Embryonic Signal That Promotes Cell Movement via Apelin Receptors. *Science* **343**, 1248636 (2014).
20. Stein, C.S., *et al.* Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency. *Cell Reports* **23**, 3710-3720.e3718 (2018).
21. Pauli, A., *et al.* Toddler: an embryonic signal that promotes cell movement via Apelin receptors. *Science* **343**, 1248636 (2014).
22. Lee, C., *et al.* The mitochondrial-derived peptide MOTS-c promotes metabolic homeostasis and reduces obesity and insulin resistance. *Cell Metab* **21**, 443-454 (2015).
23. Makarewich, C.A., *et al.* MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid beta-Oxidation. *Cell Rep* **23**, 3701-3709 (2018).

24. Lin, M.F., Jungreis, I. & Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **27**, i275-282 (2011).
25. Washietl, S.*, et al.* RNAcode: robust discrimination of coding and noncoding regions in comparative sequence data. *RNA* **17**, 578-594 (2011).
26. Fu, Z., Gilbert, E.R. & Liu, D. Regulation of insulin synthesis and secretion and pancreatic Beta-cell dysfunction in diabetes. *Curr Diabetes Rev* **9**, 25-53 (2013).
27. Araki, E., Oyadomari, S. & Mori, M. Impact of endoplasmic reticulum stress pathway on pancreatic beta-cells and diabetes mellitus. *Exp Biol Med (Maywood)* **228**, 1213-1217 (2003).
28. Berridge, M.J. The endoplasmic reticulum: a multifunctional signaling organelle. *Cell Calcium* **32**, 235-249 (2002).
29. Zhang, I.X., Raghavan, M. & Satin, L.S. The Endoplasmic Reticulum and Calcium Homeostasis in Pancreatic Beta Cells. *Endocrinology* **161**(2020).
30. Engin, F., Nguyen, T., Yermalovich, A. & Hotamisligil, G.S. Aberrant islet unfolded protein response in type 2 diabetes. *Sci Rep* **4**, 4054 (2014).
31. Safra, M., Ben-Hamo, S., Kenyon, C. & Henis-Korenblit, S. The ire-1 ER stress-response pathway is required for normal secretory-protein metabolism in C. elegans. *Journal of cell science* **126**, 4136-4146 (2013).
32. Akerman, I.*, et al.* Human Pancreatic β Cell lncRNAs Control Cell-Specific Regulatory Networks. *Cell Metabolism* **25**, 400-411 (2017).
33. Lin, N.*, et al.* An Evolutionarily Conserved Long Noncoding RNA TUNA Controls Pluripotency and Neural Lineage Commitment. *Molecular Cell* **53**, 1005-1019 (2014).
34. Morán, I.*, et al.* Human β Cell Transcriptome Analysis Uncovers lncRNAs That Are Tissue-Specific, Dynamically Regulated, and Abnormally Expressed in Type 2 Diabetes. *Cell Metabolism* **16**, 435-448 (2012).
35. Arruda, A.P. & Hotamisligil, G.S. Calcium Homeostasis and Organelle Function in the Pathogenesis of Obesity and Diabetes. *Cell Metab* **22**, 381-397 (2015).
36. Fu, S.*, et al.* Phenotypic assays identify azoramide as a small-molecule modulator of the unfolded protein response with antidiabetic activity. *Science translational medicine* **7**, 292ra298 (2015).
37. Gilon, P., Chae, H.Y., Rutter, G.A. & Ravier, M.A. Calcium signaling in pancreatic beta-cells in health and in Type 2 diabetes. *Cell Calcium* **56**, 340-361 (2014).
38. Batista, P.J. & Chang, H.Y. Long Noncoding RNAs: Cellular Address Codes in Development and Disease. *Cell* **152**, 1298-1307 (2013).
39. Sarropoulos, I., Marin, R., Cardoso-Moreira, M. & Kaessmann, H. Developmental dynamics of lncRNAs across mammalian organs and species. *Nature* **571**, 510-514 (2019).
40. Angcajas, A.B.*, et al.* Diversity of amino acid signaling pathways on autophagy regulation: a novel pathway for arginine. *Biochemical and biophysical research communications* **446**, 8-14 (2014).
41. Ulitsky, I., Shkumatava, A., Jan, C.H., Sive, H. & Bartel, D.P. Conserved function of lincRNAs in vertebrate embryonic development despite rapid sequence evolution. *Cell* **147**, 1537-1550 (2011).
42. Dai, J., Ma, J., Yu, B., Zhu, Z. & Hu, Y. Long Noncoding RNA TUNAR Represses Growth, Migration, and Invasion of Human Glioma Cells Through Regulating miR-200a and Rac1. *Oncol Res* **27**, 107-115 (2018).
43. Makarewich, C.A. The hidden world of membrane microproteins. *Exp Cell Res* **388**, 111853 (2020).
44. Bertram, R. & Sherman, A. Filtering of calcium transients by the endoplasmic reticulum in pancreatic beta-cells. *Biophys J* **87**, 3775-3785 (2004).
45. Misquitta, C.M., Mack, D.P. & Grover, A.K. Sarco/endoplasmic reticulum Ca2+ (SERCA)-pumps: link to heart beats and calcium waves. *Cell Calcium* **25**, 277-290 (1999).
46. Klionsky, D.J.*, et al.* Guidelines for the use and interpretation of assays for monitoring autophagy (3rd edition). *Autophagy* **12**, 1-222 (2016).

47. Arredouani, A., *et al.* SERCA3 ablation does not impair insulin secretion but suggests distinct roles of different sarcoendoplasmic reticulum Ca(2+) pumps for Ca(2+) homeostasis in pancreatic beta-cells. *Diabetes* **51**, 3245-3253 (2002).
48. Ravier, M.A., *et al.* Mechanisms of control of the free Ca2+ concentration in the endoplasmic reticulum of mouse pancreatic beta-cells: interplay with cell metabolism and [Ca2+]c and role of SERCA2b and SERCA3. *Diabetes* **60**, 2533-2545 (2011).
49. Zarain-Herzberg, A., Garcia-Rivas, G. & Estrada-Aviles, R. Regulation of SERCA pumps expression in diabetes. *Cell Calcium* **56**, 302-310 (2014).
50. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120 (2014).
51. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).
52. Soneson, C., Love, M.I. & Robinson, M.D. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015).
53. Gentleman, R.C., *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5**, R80 (2004).
54. Hohmeier, H.E., *et al.* Isolation of INS-1-derived cell lines with robust ATP-sensitive K+ channel-dependent and -independent glucose-stimulated insulin secretion. *Diabetes* **49**, 424-430 (2000).
55. Beigi, F., *et al.* Dynamic denitrosylation via S-nitrosoglutathione reductase regulates cardiovascular function. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 4314-4319 (2012).
56. de Juan-Sanz, J., *et al.* Axonal Endoplasmic Reticulum Ca(2+) Content Controls Release Probability in CNS Nerve Terminals. *Neuron* **93**, 867-881 e866 (2017).
57. Yang, L., *et al.* METABOLISM. S-Nitrosylation links obesity-associated inflammation to endoplasmic reticulum dysfunction. *Science* **349**, 500-506 (2015).
58. Nagendran, M., Riordan, D.P., Harbury, P.B. & Desai, T.J. Automated cell-type classification in intact tissues by single-cell molecular profiling. *Elife* **7**(2018).
59. Stephens, S.B., *et al.* A VGF-derived peptide attenuates development of type 2 diabetes via enhancement of islet beta-cell survival and function. *Cell Metab* **16**, 33-43 (2012).

FIGURE LEGENDS

Figure 1. Identification of lncRNAs that have coding potential in the human pancreas. A. Scott plot of lncRNA expression level in human islets and coding potential. Expression level was measured by log2(TPM+1). Coding potential was measured by log2(PhyloCSF score + 1). **B.** PhyloCSF score and RNAcode score for 12 lncRNAs that have recently been found to encode micropeptides. *TUNAR* is the micropeptide encoding lncRNA identified in this study. **C.** Genomic architecture of the *TUNAR* locus. 100 vertebrates sequence conservation shows a conserved region at the beginning of last exon. **D.** Multi-species alignment of predicted *TUNAR* encoding micropeptide BNLN. A conserved transmembrane domain is located at the C-terminus. Three potential start codons were presented.

620 **Figure 2. *TUNAR* encodes a novel micropeptide, BNLN. A**. Schematic representation of *BNLN* reporter
621 constructs. *BNLN* coding sequence (CDS) is shown in pink block. **B.** Representative images (20X) and **C.**
622 Activity of Renilla luciferase in HEK 293T cells transfected with the indicated BNLN constructs as in (A)
623 for 48 hr. Scale bar: 50μm. The data were normalized to firefly luciferase. * indicates statistical significance
624 compared to in-frame constructs (n=4, biological replicates). AU, arbitrary units. **D.** Representative western
625 blots of BNLN expression in HEK 293T cells transfected with the indicated *BNLN* constructs. CT: non-
626 transfected cell control. **E.** Representative western blots of BNLN expression in cells transfected with the
627 indicated *BNLN* constructs for 48hrs. **F.** Levels of mRNAs encoding for BNLN in HEK293T cells expressing
628 full length, ATG-deleted and CTG-deleted *BNLN* CDS. The data were normalized to *ACTB* expression
629 level. * indicates statistical significance compared to the full *BNLN* expressing cells (n=6, biological
630 replicates). Data are shown as means ± SEM, statistical significance was determined by ANOVA followed
631 by Sidak's multiple comparisons test in C&F. $p<0.05$.
632
633 **Figure 3. BNLN localizes at the endoplasmic reticulum in β-cells. A.** Representative confocal images
634 (60X) of *TUNAR* in human islets. Small molecular *in situ* hybridization of *BNLN*, and *INSULIN* (β-cell
635 marker). Scale bar: 10μm. **B.** Representative western blots of BNLN expression in INS-1 cells transfected
636 with the indicated *BNLN* constructs for 48hrs. **C.** Representative confocal images (63X) of BNLN in INS-1
637 cells co-transfected with FLAG-GFP-BNLN and KDEL-RFP (24hrs post transfection). CT vevtor: FLAG-
638 GFP construct. Scale bar: 10μm. **D**. Representative confocal images (63X) of BNLN in INS-1 cells
639 transfected with BNLN-GFP and stained with lysotracker (lysosomes), mitotracker (mitochondria), Acaa1
640 antibody (peroxisomes), and TGN38 antibody (Golgi). Scale bar: 10μm. **E**. Model of HiBiT protein tagging
641 system in monitoring BNLN localization. INS-1 cells were transfected with HiBiT-tagged BNLN, the cell
642 surface HiBiT was recognized by recombinant LgBiT and NanoBiT substrate in live cells. **F**. Expression of
643 cell surface HiBiT-BNLN level in INS-1 cells 24hrs after transfecting with the HiBiT-BNLN construct. The
644 level of cell surface HiBiT-BNLN was measured by using LgBiT and the extracellular substrate, and
645 normalized to the total expression of the construct by lysing the cells with the lytic substrate in the presence

646 of LgBiT. Data are shown as means ± SEM, and statistical significance was determined by Student's t-test
647 ($p<0.05$, n=3, biological replicates). AU, arbitrary units.
648
649 **Figure 4. BNLN regulates glucose-induced calcium dynamics in pancreatic β-cells. A.**
650 Representative ER Ca^{2+} recording of INS-1 cells cotransfected with V5-BNLN and an ER calcium probe
651 (GCaMP6s) after stimulation with high glucose (17mM) with a microplate reader at 37°C. CT; control V5
652 vector. **B.** Quantification of the area under a curve in (C) (n=3; experimental replicates). **C.** Representative
653 cytoplasmic Ca^{2+} recording of INS-1 cells transfected with V5-BNLN or FLAG-BNLN after stimulation with
654 high glucose (17mM). The fluorescent signal was measured in the emission/excitation spectrum of Fluo-8
655 with FLIPR Tetra System at 37°C. **D.** Quantification of the area under a curve in (E) (n=6-7; experimental
656 replicates). **E.** Representative cytoplasmic Ca^{2+} recording of INS-1 cells overexpressed with FLAG-GFP-
657 BNLN before (2.5 mM glucose) and after stimulation with high glucose (17mM) in the absence or presence
658 of KCl (40mM). The fluorescent signal was measured in the emission/excitation spectrum of Fluo-8 with a
659 microplate reader at 37°C. **F.** Quantification of the area under a curve in (E) (n=8-9, experimental
660 replicates). **G.** ATF6LD-cluc secretion measured in INS-1 cells with BNLN overexpression in the presence
661 of low glucose (2.5mM) or high glucose (16.7mM) for 1-2hr at 37°C. Data were normalized to Gluc secretion
662 (n=13-15; experimental replicates). **H.** Representative western blots of BNLN and SERCA3 in HEK293T
663 cells transfected with the indicated constructs for 24hrs. **I.** PLA assay for SERCA3 interaction with BNLN
664 in INS-1 cells cultured in low glucose (2.5 mM) or challenged with high glucose (17 mM). Scale bar: 10μm.
665 -Ab: no antibody control**.** A model of PLA assay is shown on top of the panel. V5-SERCA3 and FLAG-GFP-
666 BNLN were recognized by primary antibodies and secondary antibodies coupled with connector oligos.
667 Data are shown as means ± SEM. * indicates statistical significance compared to the CT vector group in
668 (B&D), and compared to treatment of low glucose in cells with same construct in (G); # indicates statistical
669 significance compared to the CT group in cells treated with the same concentrations of glucose in (G).
670 Statistical significance was determined by student's t-test in (B&F), and ANOVA followed by Tukey's
671 multiple comparisons test in (D&G), $p<0.05$. AU: arbitrary unit.
672

Figure 5. BNLN regulates glucose-stimulated insulin secretion. A. Glucose stimulated insulin secretion in INS-1 cells overexpressed with FLAG-GFP-BNLN in static incubation in media containing low glucose (2.5 mM) or high glucose (17 mM) for 1 hr each (n=3-4, experimental replicates). The data were normalized to insulin content. Representative western blots of BNLN expression is shown on the top of pannel. **B.** GSIS in INS-1 cells overexpressed with the indicated constructs for 48hrs in static incubation in media containing low glucose (2.5mM) or high glucose (17mM) for 1hr each (n=3-4, experimental replicates). The data were normalized to insulin content. **C.** GSIS measured in isolated islets from mice fed a RD (n = 5-10, biological replicates; 12 weeks) followed transduction of adeno-GFP (CT) or adeno-BNLN constructs (48 hrs), in static incubation in media containing low glucose (2.5mM) or high glucose (17mM) for 1hr each. The data were normalized to insulin content. **D.** Levels of mRNAs encoding for *BNLN* in islets isolated from mice fed with a RD or a HFD (12 weeks). The data were normalized to *Hprt* expression in the islets. n=3-4, biological replicates. **E.** GSIS measured in isolated islets from mice fed a HFD (n=3-4, biological replicates; 12 weeks on HFD) followed transduction of adeno-GFP (CT) or adeno-BNLN constructs (48hrs), in static incubation in media containing low glucose (2.5mM) or high glucose (17mM) for 1hr each. The data were normalized to insulin content. **F-H.** Representative GSIS measured in islets from human donors followed transduction of adeno-GFP (CT) or adeno-BNLN constructs (48hrs), in static incubation in media containing low glucose (2.5mM) or high glucose (17mM) for 1hr each. The data were normalized to insulin content. In (F), GSIS was performed on a male non-diabetic human donor of 69 years of age, 27.2 BMI, in (G), GSIS was performed on a male non-diabetic human donor of 18 years of age, 19 BMI. **I.** Insulin content measured in islets from human donors followed transduction of adeno-GFP (CT) or adeno-BNLN constructs (48hrs). Data are represented as a difference in intracellular insulin content after performing GSIS between CT and BNLN human islets. Data are shown as means ± SEM. *indicates statistical significance compared to treatment of low glucose in cells/islets with same type of construct in (A,C,E&H), to the CT vector group in (B); # indicates statistical significance compared to the CT group in cells/islets incubated with the same concentrations of glucose. Statistical significance was determined by student's t-test in (D&H), and ANOVA followed by Tukey's multiple comparisons test in (A&B&C&E), $p<0.05$. AU: arbitrary unit.

'00

'01 **Supplementary Material:**

'02 **Supplemental Figure 1. Expression of TUNAR in various human tissues.** Violin plot of *TUNAR* RNA

'03 expression across various human tissues. Expression values are measured by log10(TPM+1), data from

'04 GTEX database.

'05

'06 **Supplemental Figure 2. Characterization of BNLN. A.** Ribosomal (Ribo-seq) profiling track from GWIPS-

'07 viz genome browser indicating translation of sORF at the begining of the last exon. **B.** TMHMM algorithms

'08 predict a transmembrane domain at the C-terminus of BNLN.

'09

'10 **Supplemental Figure 3. Conservation of BNLN across vertebrates species.** 88 species out of 100

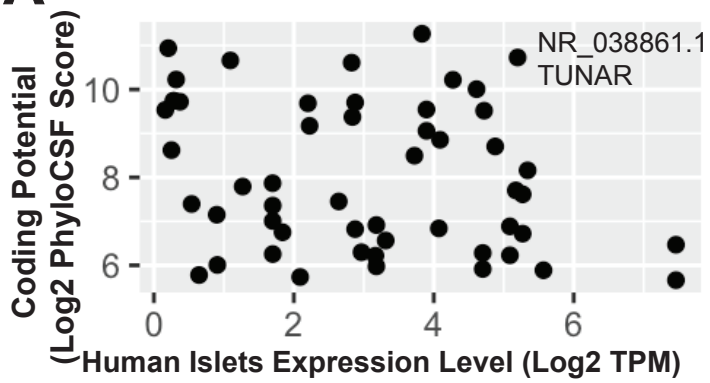
'11 vertebrate species (UCSC multiz100way) contain homologous sORFs. Multi-species alignment of *in silico*

'12 translated putative BNLN protein sequence.

'13

Figure 1

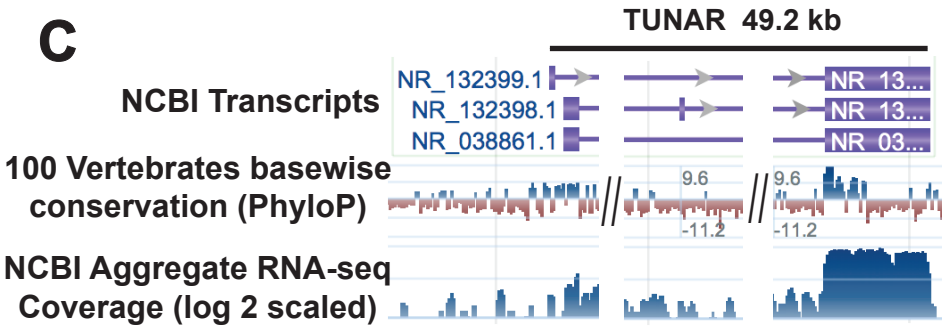
A



B

Gene	PhyloCSF Score	RNAcode Score
SPAAR(SPAR)	857	43
NBDY	386	32
STRIT1(DWORF)	406	23
MYMX	1344	29
C4orf3(ALN)	738	24
SMIM6(ELN)	58	18
MRLN	265	9
MTLN	1192	19
APELA	262	18
C12orf75	678	17
CYREN	605	17
HOXB-AS3	-188	28
TUNAR*	1696	23

C



D

Figure 2

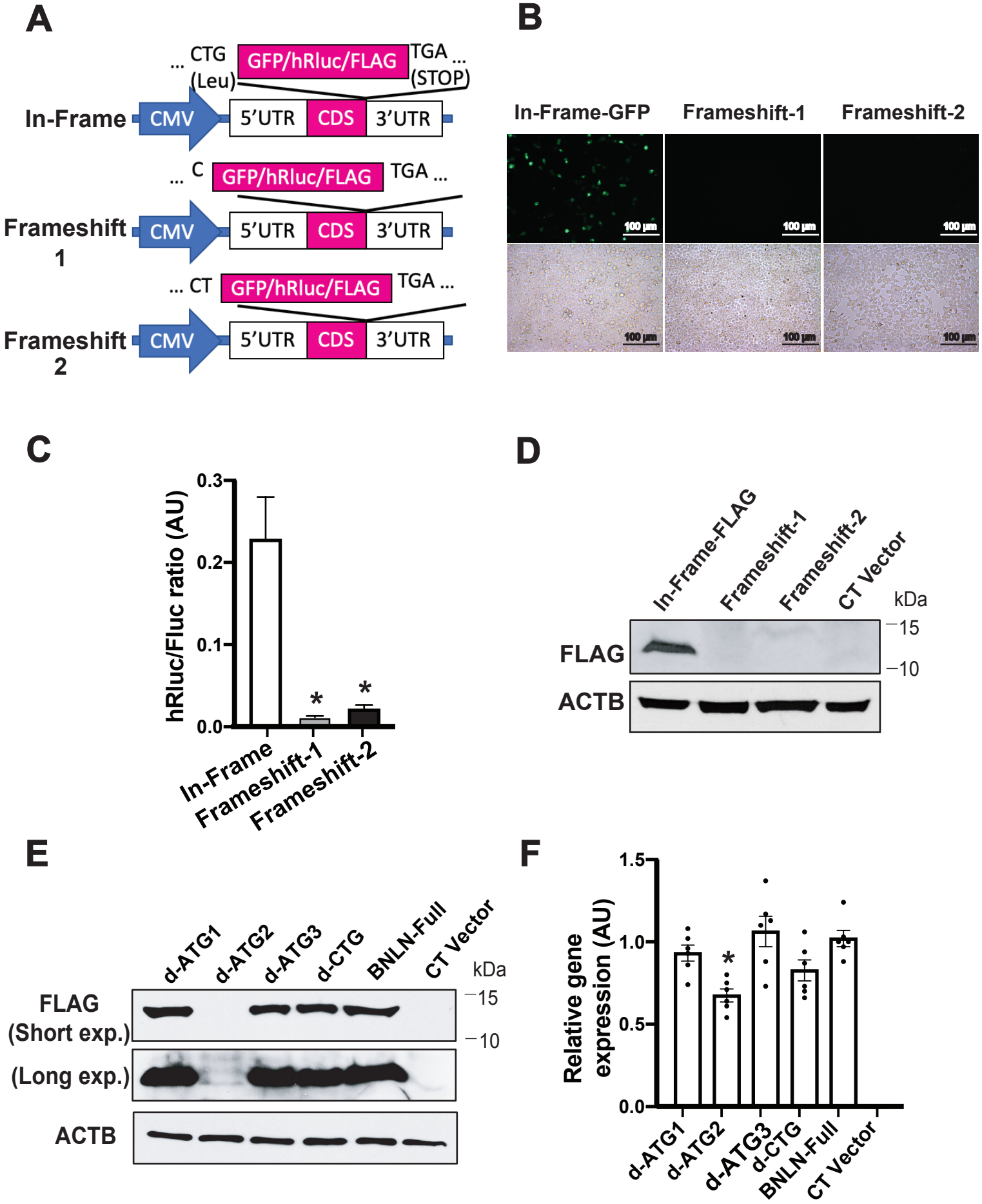

A
... CTG (Leu)
GFP/hRluc/FLAG
TGA ... (STOP)
In-Frame
CMV
5'UTR
CDS
3'UTR
... C
GFP/hRluc/FLAG
TGA ...
Frameshift 1
CMV
5'UTR
CDS
3'UTR
... CT
GFP/hRluc/FLAG
TGA ...
Frameshift 2
CMV
5'UTR
CDS
3'UTR
B
In-Frame-GFP
Frameshift-1
Frameshift-2
100 µm
C
hRluc/Fluc ratio (AU)
0.3
0.2
0.1
0.0
*
*
In-Frame
Frameshift-1
Frameshift-2
D
In-Frame-FLAG
Frameshift-1
Frameshift-2
CT Vector
kDa
FLAG
ACTB
15
10
E
d-ATG1
d-ATG2
d-ATG3
d-CTG
BNLN-Full
CT Vector
kDa
FLAG (Short exp.)
(Long exp.)
ACTB
15
10
F
Relative gene expression (AU)
1.5
1.0
0.5
0.0
*
d-ATG1
d-ATG2
d-ATG3
d-CTG
BNLN-Full
CT Vector

Figure 3

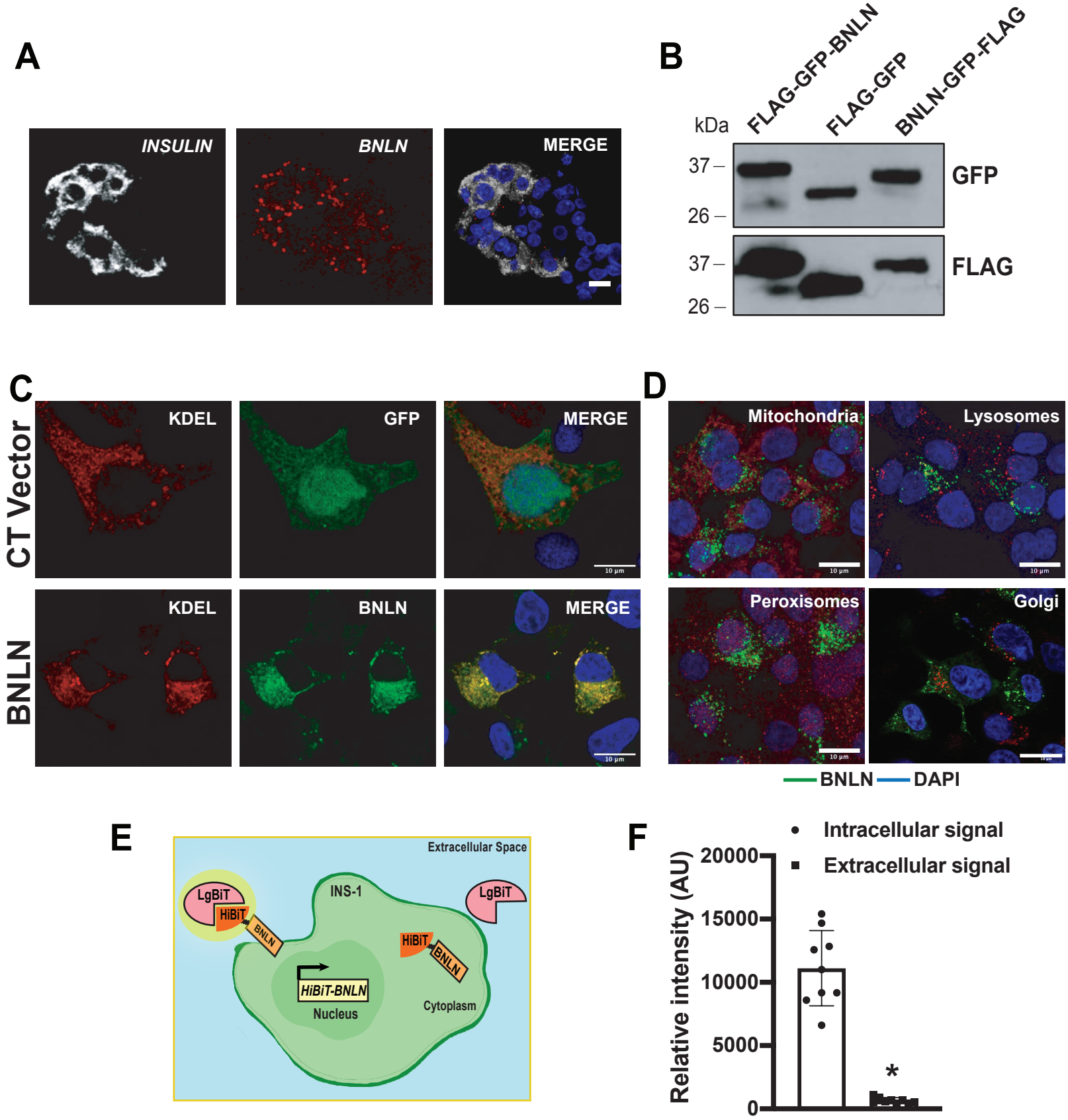

A
INSULIN
BNLN
MERGE
B
FLAG-GFP-BNLN
FLAG-GFP
BNLN-GFP-FLAG
kDa
37
26
GFP
37
26
FLAG
C
CT Vector
KDEL
GFP
MERGE
BNLN
KDEL
BNLN
MERGE
D
Mitochondria
Lysosomes
Peroxisomes
Golgi
BNLN
DAPI
E
Extracellular Space
LgBiT
HiBiT
BNLN
INS-1
LgBiT
HiBiT
BNLN
HiBiT-BNLN
Nucleus
Cytoplasm
F
Intracellular signal
Extracellular signal
Relative intensity (AU)
20000
15000
10000
5000
0
*

Figure 4

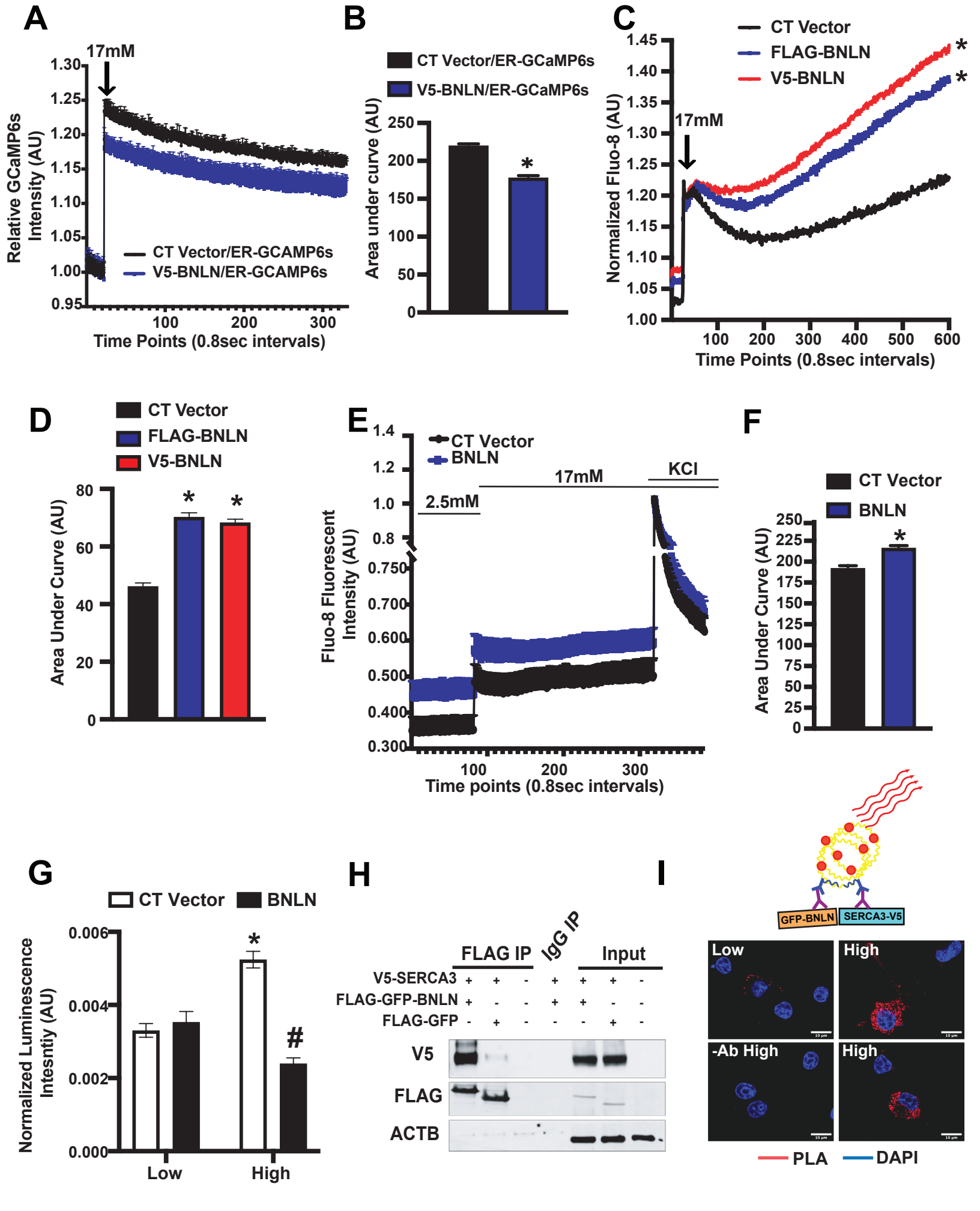
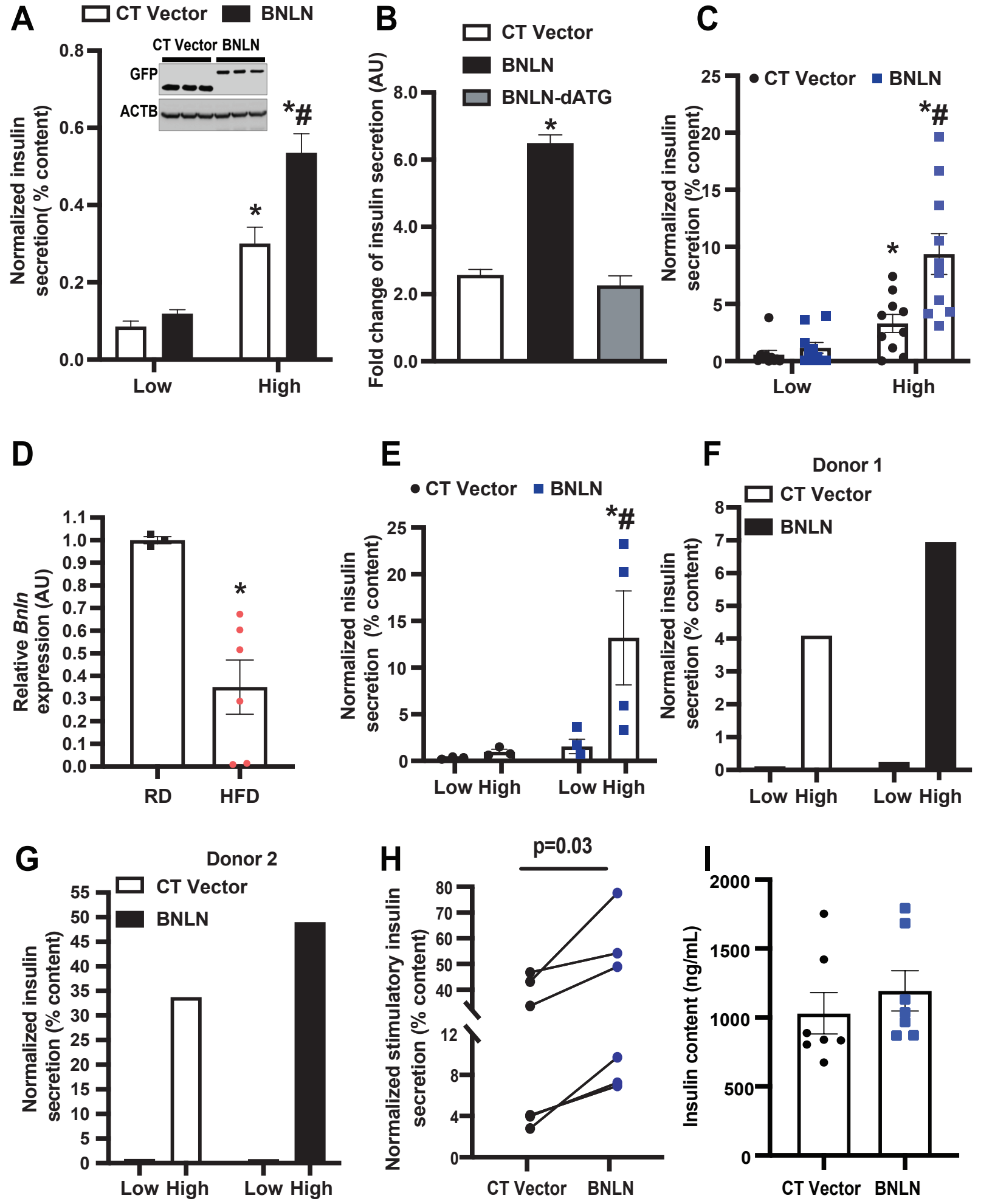

A
17mM
Relative GCaMP6s Intensity (AU)
CT Vector/ER-GCAMP6s
V5-BNLN/ER-GCAMP6s
Time Points (0.8sec intervals)
B
CT Vector/ER-GCaMP6s
V5-BNLN/ER-GCaMP6s
Area under curve (AU)
C
CT Vector
FLAG-BNLN
V5-BNLN
17mM
Normalized Fluo-8 (AU)
Time Points (0.8sec intervals)
D
CT Vector
FLAG-BNLN
V5-BNLN
Area Under Curve (AU)
E
CT Vector
BNLN
2.5mM
17mM
KCl
Fluo-8 Fluorescent Intensity (AU)
Time points (0.8sec intervals)
F
CT Vector
BNLN
Area Under Curve (AU)
G
CT Vector
BNLN
Normalized Luminescence Intesntiy (AU)
Low
High
H
FLAG IP
IgG IP
Input
V5-SERCA3
FLAG-GFP-BNLN
FLAG-GFP
V5
FLAG
ACTB
I
GFP-BNLN
SERCA3-V5
Low
High
-Ab High
High
PLA
DAPI

Figure 5


A
CT Vector
BNLN
CT Vector BNLN
GFP
ACTB
Normalized insulin secretion(% content)
0.8
0.6
0.4
0.2
0.0
*
*#
Low
High
B
CT Vector
BNLN
BNLN-dATG
Fold change of insulin secretion (AU)
8.0
6.0
4.0
2.0
0.0
*
C
CT Vector
BNLN
Normalized insulin secretion (% conent)
25
20
15
10
5
0
*
*#
Low
High
D
Relative Bnln expression (AU)
1.1
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
*
RD
HFD
E
CT Vector
BNLN
Normalized nisulin secretion (% content)
25
20
15
10
5
0
*#
Low High
Low High
F
Donor 1
CT Vector
BNLN
Normalized insulin secretion (% content)
8
7
6
5
4
3
2
1
0
Low High
Low High
G
Donor 2
CT Vector
BNLN
Normalized insulin secretion (% content)
55
50
45
40
35
30
25
20
15
10
5
0
Low High
Low High
H
p=0.03
Normalized stimulatory insulin secretion (% content)
80
70
60
50
40
12
8
4
0
CT Vector
BNLN
I
Insulin content (ng/mL)
2000
1500
1000
500
0
CT Vector
BNLN