Online Only Supplementary Material

A Novel Intron-Encoded Neuropilin-1 Isoform in Pancreatic Islets Associated With Very Young Age of Onset of Type 1 Diabetes

Running Title: *NRP1* Intronic Isoform and Early Onset Diabetes

Michael J. MacDonald[1], Israr-ul H. Ansari[1], Amy S. Riedemann[1], Scott W. Stoker[1], Jens C. Eickhoff[2], Peter J. Chlebeck[3], Luis A. Fernandez[3,4], and Melissa J. Longacre[1]

[1]Childrens Diabetes Center, [2]Department of Biostatistics & Medical Informatics, and [3]Department of Surgery, Division of Transplantation at University of Wisconsin School of Medicine and Public Health, Madison, WI 53706; [4]Current address, Department of Surgery, Division of Intra-abdominal Transplantation, Loyola University Medical Center and Stritch School of Medicine, Maywood, Illinois, USA

Address for correspondence and reprints: Michael J. MacDonald

Email: mjmacdon@wisc.edu

SUPPLEMENTAL RESEARCH DESIGN AND METHODS

Histochemistry analysis of human pancreas. Pancreases from human donors were stained with antibodies we named antibody 23 or antibody 2733 that we raised against amino acid sequences encoded by the human *NRP1* gene intron 9 and another antibody we named antibody 30 we raised against an amino acid sequence encoded by the human *NRP1* gene exon 4 (S1) as well as with a commercial antibody against human insulin. In addition to the pancreases shown in Figure 1 of the main body of the paper, more human pancreases were analyzed with these antibodies for quantifying the NRP1 intron 9-encoded neuropilin-1 amino acid sequence relative to the normal neuropilin-1 sequence to obtain the data shown in Table 2 of this supplement.

Measurement of *NRP1* mRNA in human pancreatic islets using QPCR. Relative levels of mRNA encoded by exon 9 and readthrough of exon 9 into intron 9 of the *NRP1* gene expressed in pancreatic islets isolated from five different human pancreas donors were measured as previously described (S2). Primers used were Exon 9-142 forward 5' CCGCACCTCATTCCTACATC 3' and Exon 9-246 reverse 5' CCTTGTTCTCTCGGTGCTTC 3', Intron 9-352 forward 5' AGGAGGTCATCCCTTATGTCC 3' and Intron 9-445 reverse 5' TCTAATGTCATGGCTGGAAGG 3', Exon-Intron 9-A 257 forward 5' GAAGTTCAAGATCGGGTACAGC 3' and Exon-Intron 9-A 375 reverse 5' CAAGGACATAAGGGATGACCTC 3', Exon-Intron 9-B 267 forward 5' ATCGGGTACAGCAACAACGG 3' and Exon-Intron 9-B 369 reverse 5' CATAAGGGATGACCTCCTCC 3'. Human pancreatic islets were from the University of Wisconsin Transplant Program at the University of Wisconsin School of Medicine and Public Health, Madison, WI (7 donors), and the Islet Isolation Core Facility, Washington University School of Medicine, St. Louis (3 donors).

Subjects. Medical histories and blood samples were obtained from patients attending the University of Wisconsin School of Medicine and Public Health Pediatrics Diabetes Clinic and their relatives under a protocol approved by the University of Wisconsin Health Sciences Institutional Review Board. Written informed consent was obtained from each subject and/or the subject's legal guardian. Patients with monogenic diabetes, atypical diabetes, cystic fibrosis Down syndrome, neonatal diabetes or suspicion of type 2 diabetes were excluded from analysis. Monozygotic twins were counted as one individual. All type 1 diabetes patients and non-type 1 diabetes controls studied were Caucasian because > 97% of the patients attending the pediatric diabetes clinic were Caucasian. Positive anti-pancreatic islet cell antibodies and/or very high blood glucose levels with severe weight loss at diagnosis, as well as HLA data (Supplemental Table 3) are consistent with the idea that all or essentially all patients had type 1 diabetes. All patients with onset of diabetes before age 18 years, as well as many of their parents with diabetes onset when they were children, were patients of MJM. These patients equaled 91% of the total patients analyzed.

Great care was taken to assure that the patients studied, especially the children with onset of diabetes before age 10 years, had type 1 diabetes. In addition to the above-mentioned exclusion and inclusion factors, there are other indications that all or virtually all of the patients studied had type 1 diabetes. Even if not every single patient with monogenic diabetes was detected and excluded from the study, the incidence of monogenic diabetes in children is less than 1 %. Such an unlikely inadvertent inclusion would not have a significant impact on the data. Type 2 diabetes rarely presents before age 10 years especially in white children. In the 2019 study by the Centers

for Disease Control and Prevention on Diabetes in Youth it was reported that the incidence per 100,000 of type 2 diabetes in youth ages 10 to 19 years was 4.5 in whites, 37.8 in blacks, 20.9 in Hispanics and 32.8 in Native Americans. Since all patients studied were Caucasian, this also makes it highly unlikely that patients with type 2 diabetes were included in the group of patients analyzed, especially those younger than age 10 years at onset of diabetes.

DNA Sequencing. Sequencing of a region of intron 8 to intron 9 of the *NRP1* gene using genomic DNA of cells of the peripheral blood buffy coat was performed as previously described (S3) using the forward primer 5'-AGACCCAAACGACTGAGTGGTATC-3' and the reverse primer 5'-GAGGAAGGTTTGAGGGAGGAAAG-3'. This amplifies a region that spans from part of intron 8 through exon 9 into part of intron 9 which contains the SNP rs2070303 at position 74 of intron 9 of the *NRP1* gene. The coding sequence of the *NRP1* gene is located on the reverse strand of the DNA.

HLA DR typing. HLA DR typing for HLA DR4, DR15 (a split of DR2) and DR17 (a split of DR3) was performed with polymerase chain amplification with sequence specific primers (S4).

Sources of pancreases used for histochemistry studies. The pancreases were from surgical pathology or autopsy specimens. Pancreases shown in Figure 1 of the main paper were from a male age 39 years and a female age 43 years. Other pancreases were from a male age 3 years, a male age 23 years and a male age 55 years. None of the patients had diabetes.

REFERENCES

S1. Hasan NM, Kendrick MA, Druckenbrod NR, Huelsmeyer MK, Warner TF, MacDonald MJ. Genetic association of the neuropilin-1 gene with type 1 diabetes in children: Neuropilin-1 expression in pancreatic islets. Diabetes Res Clin Pract 2010;87:e29-32

S2. Ansari IH, Longacre MJ, Paulusma CC, Stoker SW, Kendrick MA, MacDonald MJ. Characterization of P4 ATPase phospholipid translocases (flippases) in human and rat pancreatic islets. J Biol Chem 2015;290:23110-23128

S3. MacDonald MJ, Hasan NM, Ansari IH, Longacre MJ, Kendrick MA, Stoker SW. Discovery of a genetic metabolic cause for Mauriac syndrome in type 1 diabetes. Diabetes 2016;65:2051-2059

S4. Olerup O, Zetterquist H. HLA-DR typing by PCR amplification with sequence-specific primers (PCR-SSP) in 2 hours: An alternative to serological DR typing in clinical practice including donor-recipient matching in cadaveric transplantation. Tissue Antigens 1992;39: 225-235

ANALYSES OF ADDITIONAL DATA

The NRP1 intron 9 rs2070303 minor allele is significantly higher among patients with onset of diabetes before 4 years among the 844 unrelated type 1 diabetes patients. An additional separate comparison of the 844 unrelated diabetes patients from within the total 1004 patients with type 1 diabetes was made. These data that are not displayed in Table 1 of the main paper showed that the incidences of the minor allele in intron 9 of the *NRP1* gene in the various age groups were similar to those among the 1004 patients. Unrelated patients with onset of diabetes before age 4 years (N = 176 patients) had a significantly higher frequency of the minor allele (33%) than patients with onset of type 1 diabetes between ages 4 and 8 years (N = 247 patients) (21.5%, p = 0.008) or between ages 8 and 12 years (N = 238 patients) (22.7%, p = 0.02) or control subjects without type 1 diabetes (16.0%) (p < 0.0001). The frequency of the minor *NRP1* allele among patients with onset of diabetes after age 16 years (16.7%) was not significantly different than the frequency of the minor allele among the control subjects without type 1 diabetes (16.0%). The incidence of the minor allele among the unrelated patients with onset of diabetes between ages 12 and 16 years and all unrelated type 1 diabetes patients of 28.3% and 25.4%, respectively, were significantly higher than in control subjects without type 1 diabetes (16.0%) (p = 0.004 and 0.03, respectively).

Higher incidence of the NRP1 intron 9 rs2070303 minor allele among patients with onset of type 1 diabetes between ages 0.67 and 2 years and between 2 and 4 years vs. onset at older years and vs. non-type 1 diabetes controls. Not shown in Table 1 of the main body of the paper are additional analyses of the group of 1004 patients that includes children, their parents and siblings diagnosed with type 1 diabetes between 0.67 years (the youngest age of a patient at diagnosis of type 1 diabetes) and 2 years (N = 69 patients) and between 2 and 4 years (N = 123 patients). Among the group with the onset of diabetes between 0.67 and 2 years the percentage of patients with the minor *NRP1* allele (31.9%) was significantly higher than patients with the minor *NRP1* allele with diabetes onset between ages 4 and 8 years (20.8%) (p = 0.05) and non-type 1 diabetic subjects (16.0%) (p = 0.004). Among the patients with the onset of diabetes between 2 and 4 years the percentage with the minor *NRP1* allele (31.7%) was significantly higher than among patients with the minor allele with the onset of diabetes between ages 4 and 8 years (20.8%) (p = 0.02) and among patients with the onset of diabetes between 8 and 12 years (22.6%) (p = 0.05). The percentage of patients with the minor *NRP1* allele among those with the onset of diabetes between 2 and 4 years (31.7%) was also significantly higher than among patients with the onset of diabetes after age 16 years (16.1%) (p = 0.01) and control subjects without type 1 diabetes (16.0%) (p = 0.008).

Also not shown in the main body of the paper are additional analyses of the 844 unrelated children with the onset of type 1 diabetes between 0.67 and 2 years and between 2 and 4 years. The frequencies of children with the onset of diabetes between 0.67 and 2 years (34.4% of 61 patients) and between 2 and 4 years (32.2% of 115 patients) who carried the minor *NRP1* allele were significantly higher than those who carried minor the allele with the onset of diabetes between 4 and 8 years (20.8% of 247 patients) (p = 0.036 and p = 0.029, respectively). The higher percentage of the minor *NRP1* allele among children with the onset of diabetes between 0.67 and 2 years and between 2 and 4 years just missed statistical significance when compared to patients with the onset of diabetes between ages 8 to 12 years (22.6% of 238 patients) (p = 0.06

and p = 0.05, respectively). The incidences of the minor *NRP1* allele in children with the onset of diabetes between 0.67 and 2 years and between 2 and 4 years were much higher than the minor allele among the control subjects without type 1 diabetes (16.0%) (p = 0.0018 and p = 0.0007, respectively).

Supplemental Table 1. *NRP1* intron 9-encoded amino acid sequences in various primates are homologous to the human sequence that can result from inframe translational read-through of exon 9 into intron 9. Each of the sequences terminates with a stop codon. Red letters indicate a difference in sequence vs. human sequence. When present in humans the minor allele of SNP rs2070303 will change the amino acid at position 25 of the 59 amino acid intron 9-encoded sequence from arginine to glutamine. Data are from the U.S. National Center for Biological Information (NCBI).

Primate	NCBI Accession No.	Length of Intron 9-Encoded Amino Acid Sequence	Intron 9-Encoded Sequence (Percent Identity of Primate Sequence to Human Sequence)
Human *Homo sapiens*	NG_030328.1	59	VRAGDWEEVIPYVLDFNQGPLAFSRGRASYFLPAMTLEEV C DGVSKTLDSRKTGIWHEK
Sumatran Orangutan *Pongo abelli*	PNJ_39711.7	59	VRAGDWEEVIPYVLDFNQGPLAFSRGRASYFLPAMTLEEV C DGVSKTLDSRKTGIWHEK (100%)
Western Lowland Gorilla *Gorilla gorilla*	NC_044612.1	59	VRAGDWEEVIPYVLDFNQGPLAFSRGRASYFLPAMTLEEV C DGVSKTLDSKKTGIWHEK (98%)
Chimpanzee *Pan troglodytes*	AC186441.3	51	_____–VIPYALDFNQGPLAFSRGRASYFLPAMTLEEC DGISKTLDSRKTGIWHGK (96%)
Tufted Capuchin *Sapajus apella*	NW_022437051.1	64	VRAGDWEEVVPQVLDFSHRPPAFSRGRASSYFLPAVISEEV C DGVSKTLDSRKTGIWHGKRCVR (86%)
Rhesus Monkey *Macaca mulatta*	NC_041762.1	74	VRAGDWQEVVPCVLNFHQGLWHSPEEGLLFFFFFFLRRS F RSCCPGWSAMARSQLIAPSASRVQVILLPQPPE (74% vs first 19 amino acids)
Golden Snub-nosed Monkey *Rhinopithecus roxellana*	NC_044559.1	24	VRAGDWQEVVPRVLNFNQGPLAFS (83%)
Green Monkey *Chlorocebus sabaeus*	NC_023650.1	24	VRAGDWQEVVPCVLNFHQGPLAFS (79%)
Olive Baboon *Papio anubis*	NC_044986.1	24	VRAGDWQEVVPRVLNFHQGPLAFS (79%)

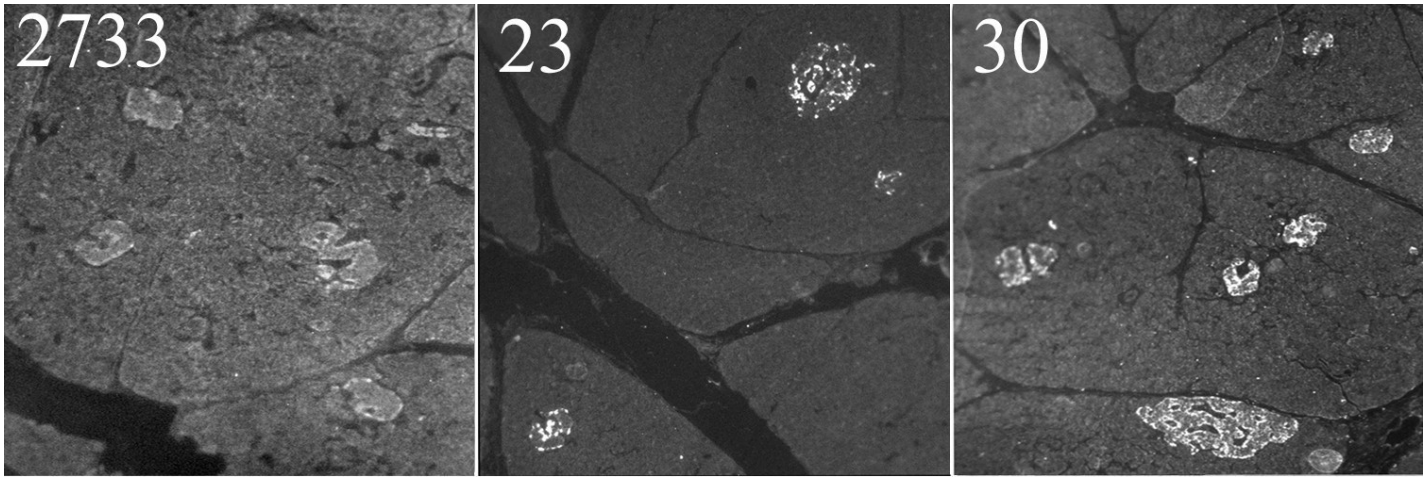
Supplemental Table 2. Variability of ratios in human pancreases of islet cells that contain the truncated neuropilin-1 protein sequence and do not contain insulin to islet cells that contain both the full- length neuropilin-1 protein sequence and insulin. Cells in pancreatic islets that stain positive for both insulin and NRP1 exon 4 encoded amino acid sequence contain the full length neuropilin-1 (Indicated by merged colors as in Figure 1 main paper). The number of these cells were compared to the number of cells in these same islets that contain NRP1 exon 4 encoded amino acid sequence and do not contain insulin or to cells that contain NRP1 intron 9-encoded amino acid sequence and do not contain insulin (separate colors and not merged as in Figure 1 main paper). Because NRP1 intron 9 possesses a stop codon any cell that expresses NRP1 intron 9 will contain a truncated neuropilin-1 protein. Cells that stained positive for glucagon were not used in the calculation. Islets in pancreases of five non-diabetic adult humans were studied.

Islets	**Cells with full length neuropilin-1** (%)	**Cells with truncated neuropilin-1** (%)
1	58	42
2	85	15
3	62	38
4	60	40
5	69	31
6	68	32
7	68	32
8	76	24
9	67	33
10	66	34
11	100	0
12	77	23
13	82	18
14	98	2
15	93	7
16	99	1
17	61	29
18	54	46
19	63	37
20	69	31
21	75	26

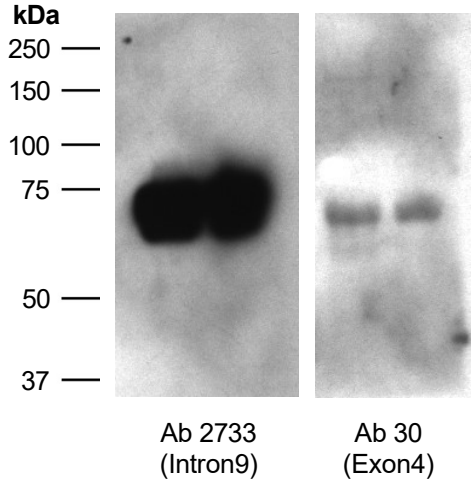
Supplemental Table 3. Percentages of HLA DR 4, 17, 15 and combinations of these HLA alleles in 1004 type 1 diabetes patients with the C (wildtype) or T (minor) allele of SNP rs2070303 with onset of type 1 diabetes at various age intervals. DR15 and DR17 are, respectively, splits of previous DR2 and DR3.

	IDDM Pts	Pts Dx 0.7 - < 2	Pts Dx 2 - < 4	Pts Dx 4 - < 6	Pts Dx 6 - < 8	Pts Dx 8 - < 10	Pts Dx 10 - < 12	Pts Dx 12 - < 14	Pts Dx 14 - < 16	Pts Dx ≥ 16
Total Patients	1004	69	123	135	144	134	132	116	64	87
% of Total C/C (HLA ≠ 4, 15, 17)	6.55	8.51	4.76	9.09	8.11	6.60	5.00	5.95	4.17	6.85
% of Total T/C + T/T (HLA ≠ 4, 15, 17)	5.81	9.09	2.56	8.00	6.06	7.14	9.38	3.13	0.00	7.14
% of Total C/C (HLA = 4)	66.06	61.70	70.24	60.91	71.17	61.32	58.00	73.81	70.83	69.86
% of Total T/C + T/T (HLA = 4)	65.56	50.00	66.67	68.00	63.64	53.57	75.00	75.00	68.75	64.29
% of Total C/C (HLA = 17)	54.52	59.57	57.14	62.73	44.14	59.43	57.00	48.81	56.25	46.58
% of Total T/C + T/T (HLA = 17)	55.60	63.64	69.23	60.00	54.55	53.57	43.75	50.00	62.50	35.71
% of Total C/C (HLA = 4 + 17)	28.05	29.79	33.33	33.64	24.32	28.30	22.00	28.57	31.25	23.29
% of Total T/C + T/T (HLA = 4 + 17)	30.71	27.27	41.03	36.00	27.27	21.43	31.25	28.13	43.75	14.29
% of Total C/C (HLA = 15)	3.28	0.00	3.57	6.36	2.70	0.94	4.00	2.38	4.17	4.11
% of Total T/C + T/T (HLA = 15)	2.49	0.00	2.56	0.00	3.03	3.57	0.00	3.13	6.25	7.14
% of Total C/C (HLA = 4 + 15)	1.31	0.00	1.19	1.82	0.90	0.00	1.00	2.38	4.17	1.37
% of Total T/C + T/T (HLA = 4 + 15)	0.41	0.00	0.00	0.00	0.00	3.57	0.00	0.00	0.00	0.00
% of Total C/C (HLA = 15 + 17)	1.05	0.00	1.19	3.64	0.90	0.00	1.00	0.00	0.00	1.37
% of Total T/C + T/T (HLA = 15 + 17)	1.66	0.00	2.56	0.00	0.00	0.00	0.00	3.13	6.25	7.14

Supplemental Figure 1. Low power micrograph of human pancreas showing neuropilin-1 is located only in pancreatic islet cells. Slides of a human pancreas were stained with antibodies that we made. Antibody 30 against *NRP1* exon 4-encoded protein will react with both the normal full length neuropilin-1 and with truncated neuropilin-1. *NRP1* intron 9 contains a stop codon. Antibodies 2733 and 23 against *NRP1* intron 9-encoded protein sequence will react only with the truncated neuropilin-1 protein.



Supplemental Figure 2. Immunoblots of abnormal 67 kDa neuropilin-1 truncated protein caused by a stop codon within intron 9. The blots show a protein from an insulinoma tumor from a human infant with hypoglycemia probed with antibody 2733 raised against *NRP1* intron 9-encoded amino acid sequence. This blot was stripped and probed again with antibody30 raised against *NRP1* exon 4-encoded amino acid sequence. The size of the predicted 597 amino acid protein is consistent with the 538 amino acid sequence encoded by *NRP1* gene exons 1-9 plus the *NRP1* gene intron 9-encoded sequence terminated by a stop codon after 59 amino acids.



Supplemental Figure 3. Schematic of *NRP1* Exon 9 (332 nt) and the first 180 nt of the intron 9 region. The shaded boxes represent approximate location of qPCR regions in exon 9, intron 9 and exon-intron junctions. The numbers in parentheses indicate the expected sizes of the PCR products. Two independent primer sets were used to amplify exon-intron junctions: Exon-Intron9A and Exon-Intron9B. The hypothetical terminal codon for Exon 9 readthrough is highlighted in bold and underlined.