SUPPLEMENTARY DATA

Supplementary Table S1. Genotype prevalences for the *TBC1D4* variant (rs61736969, a nonsense polymorphism in *TBC1D4* [also called c.2050C>T]) in the Nunavik Inuit and the GOCADAN cohort.

Genotype	Nunavik Inuit	Alaskan Inuit (GOCODAN Study)
CC	72 (69%)	771 (75%)
CT	30 (29%)	241 (23.5%)
TT	2 (2%)	15 (1.5%)
All	104	1027

Supplementary Table S2. Association of the *TBC1D4* variant with metabolic traits in diabetic individuals from the GOCADAN cohort

Results are shown for an additive and a recessive genetic model. For each trait, n is the number of individuals with genotype data for the specific variant and phenotypic data for the specific trait. *β* is the effect size estimated using untransformed values. BMI, body mass index; HDL, high-density lipoprotein;LDL, low-density lipoprotein.$ISI_{0\text{-}120,}$ Gutt insulin sensitivity index

Trait		Additive model			Recessive model		
	n	*β*	*95%CI*	*P*	*β*	*95%CI*	*P*
2-h Plasma glucose (mmol/L)	26	2.10	-0.88;5.08	0.17	13.99	1.98;26.00	**0.02**
Fasting serum insulin (pmol/L)	50	-69.17	-121.35;-16.99	**0.09**	-68.74	-218.88;81.40	0.37
2-h serum insulin (pmol/L)	26	-250.69	-584.93;83.55	0.14	226.61	-1013.53;1466.75	0.72
$ISI_{0,120}$ (standard deviations)	25	-0.11	-0.80;0.57	0.739	-2.56	-5.27;0.16	0.06
Fasting serum HDL-cholesterol (mmol/L)	51	0.31	0.00;0.62	**0.05**	0.26	-0.64;1.15	0.57
Waist circumference (cm)	29	-3.23	-8.41;1.96	0.22	13.65	-9.44;36.75	0.25
Percent fat mass (%)	28	-2.47	-7.06;2.12	0.29	7.39	-13.07;27.85	0.48
BMI (kg/m^2)	27	-5.90	-10.77;-1.02	**0.02**	5.07	-17.47;27.61	0.66
Albuminuria (mg/g)	49	0.18	-0.60;0.96	0.66	0.10	-2.22;2.42	0.93

SUPPLEMENTARY DATA

Supplementary Table S3. Association of the *TBC1D4* variant with metabolic traits in non-diabetic individuals from the GOCADAN cohort, using inverse normalized values for the studied traits.

Trait		Additive model			Recessive model		
	n	*β*	*95%CI*	*P*	*β*	*95%CI*	*P*
Fasting plasma glucose	975	-0.25	-0.38; -0.12	**2.37E-04**	-0.40	-0.94; 0.14	0.14
2-h Plasma glucose	709	0.26	0.09; 0.42	**2.14E-03**	1.65	0.81; 2.48	**1.12E-04**
HbA1C	975	-0.10	-0.24; 0.03	0.12	0.43	-0.10; 0.96	0.11
Fasting serum insulin	974	-0.20	-0.34; -0.07	**3.65E-03**	-0.14	-0.69; 0.42	0.63
2-h serum insulin	706	0.10	-0.06; 0.26	0.21	0.85	0.03; 1.67	**0.04**
Fasting serum HDL-cholesterol	974	0.04	-0.09; 0.17	0.58	0.30	-0.23; 0.83	0.27
Fasting serum LDL-cholesterol	974	0.02	-0.12; 0.15	0.78	0.08	-0.47; 0.62	0.79
Fasting serum triglyceride	974	-0.02	-0.16 0.12;	0.76	0.54	-0.03; 1.10	0.06
Fasting serum total cholesterol	974	0.00	-0.12; 0.13	0.94	0.31	-0.20; 0.82	0.23
Waist circumference	919	-0.21	-0.35; -0.06	**5.21E-03**	0.07	-0.56; 0.70	0.82
Percent fat mass	921	-0.11	-0.20; -0.02	**1.38E-02**	0.00	-0.03; 0.03	0.98
BMI	931	-0.14	-0.28; 0.00	0.06	-0.09	-0.72; 0.54	0.77
Albuminuria	939	-0.05	-0.11; 0.02	0.15	-0.08	-0.30; 0.14	0.47

SUPPLEMENTARY DATA

Supplementary Table S4. Association of the *TBC1D4* variant with metabolic traits in diabetic individuals from the GOCADAN cohort, using inverse normalized values for the studied traits.

Trait		Additive model			Recessive model		
	n	*β*	*95%CI*	*P*	*β*	*95%CI*	*P*
2-h Plasma glucose	26	0.60	-0.15;1.35	0.12	4.34	1.37;7.30	4×10^{-3}
Fasting serum insulin	50	-0.50	-0.94;-0.06	0.02	-0.02	-1.29;-1.26	0.98
2-h serum insulin	26	-0.24	-0.92;0.43	0.48	1.65	-1.09;4.39	0.24
Fasting serum HDL-cholesterol	51	0.50	0.03;0.98	0.04	0.64	-0.73;2.02	0.36
Fasting serum LDL-cholesterol	51	0.00	0.00; 0.00	0.20	0.42	-0.07; 0.90	0.09
Waist circumference	29	-0.34	-0.98;0.31	0.31	2.36	-0.48;5.21	0.10
Percent fat mass	28	-0.43	-0.92;0.06	0.09	0.84	-1.38;3.07	0.46
BMI	27	-0.62	-1.17;-0.06	0.03	0.75	-1.80;3.30	0.57

SUPPLEMENTARY DATA

Supplementary Table S5. Association of the *TBC1D4* variant with type 2 diabetes diagnosis in 1027 individuals from the GOCADAN cohort.

Trait		Additive model			Recessive model		
	n (cases/controls)	*OR*	*95%CI*	*P*	*OR*	*95%CI*	*P*
Type 2 Diabetes diagnosis	1027 (51/976)	1.03	0.76 -1.39	0.86	0.64	0.25-1.67	0.36

Results are shown for an additive and a recessive genetic model. *n* is the number of individuals (cases and controls in parenthesis) with genotype data for the *TBC1D4* variant and phenotypic data for type 2 diabetes. *OR* is the odds ratio of having a diagnosis of type 2 diabetes if one copy of the effect allele is present.

SUPPLEMENTARY DATA

Supplementary Table S6. Classification of 735 GOCADAN participants with available OGTT results according to their type 2 diabetes and prediabetes status and their number of *TBC1D4* variant copies.

TBC1D4 copies	No type 2 diabetes or prediabetes	Type 2 diabetes	Prediabetes	All
0	352	18	203	573
1	188	7	39	156
2	2	1	3	6
All	542	26	245	735

Supplementary Text

Sanger sequencing in the GOCADAN cohort

For Sanger sequencing a 288bp PCR fragment was amplified using the following primers: forward 5' CCT GAT TTC ATT TTC ATA CTG TGG 3' and reverse 5' CCT GAG CCT TGA AAG GAA AGA 3'. Amplification was performed in 25µl reactions according to the manufacturer's protocol using the HotStart PCR Kit, with dNTPs (Kapa Biosystems, Wilmington, MA), using 0.5µM for each primer and 10ng of genomic DNA. PCR conditions were 96ºC/1min; 33x 96ºC/10 sec, 58ºC for 5 sec, 72ºC for 1 sec; 72ºC/30sec; 4ºC.

Sequencing reactions were performed on 384well plates using the BigDye Terminator v3.1 Cycle Sequencing kit (Thermo Fisher Scientific, Grand Island, NY) according to the manufacturer's protocol with the following adjustments. For a 5ul sequencing reaction, 1ul PCR product (diluted 1:5), 0.25ul BigDye Terminator v3.1 ready reaction mix, 1ul Big Dye Terminator 5x sequencing buffer, and 0.14µM sequencing primer (forward primers) were used. The sequencing reaction was performed on a thermocycler for 96ºC/1min; 25x 96ºC/10sec, 50ºC/5sec, 60ºC/4min; 4ºC.

Unincorporated BigDye terminators and salts were removed using the BigDye XTerminator purification kit (Thermo Fisher Scientific) according to the manufacturer's protocol with the following adjustments. 4ul of XTerminator solution and 18ul of SAM solution were added to each well containing 5ul of sequencing product. The plates were vortexed for 30 min at RT before being placed on the 3730xl DNA Analyzer for the analysis using the 3730 DNA Analyzer Data Collection Software v3.0 and Sequence Scanner v1.0 software (Thermo Fisher Scientific) for data collection and analysis.

Exome sequencing in the Quebec Inuit cohort

The detailed exome sequencing methods have been described previously[1]. Briefly, DNA samples were captured using Agilent SureSelect 50 Mb (V4) chip, and then sequenced on Illumina Hiseq 2000 at the McGill University and Génome Québec Innovation Centre. In each lane, three samples were barcoded and sequenced together to reach an average sequence depth of 100 fold. Raw sequence reads were aligned to GRCh37 using Burrow-Wheeler Aligner [2], variants were called by Genome Analysis Toolkit (GATK)[3] , and ANNOVAR[4] was used for variant annotation. For calling rs61736969 variant, quality filters were set at sequencing depth ≥ 20x and variant frequency ≥ 25%, with genotype quality ≥ 10. PLINK 1.9[5] was used for the GATK file conversion and to calculate the genotype frequency. To identify Inuit individuals with excessive admixture, defined as >5% mixture from other ethnicity in their genomes, Principle Component Analysis (PCA) was performed. Specifically PCA was conducted by smartPCA software, implemented in the Eigensoft package (A.L. Price, et al, 2006), using European (CEU), Han Chinese (CHB) and African (YRI) populations from 1000 Genome Project Phase 3[6] as references.

SUPPLEMENTARY DATA

Admixture calculation in the Quebec Inuit cohort

PCA was performed using CHB (Han Chinese) and CEU (European) populations as reference populations to enable clustering. Therefore, the Inuit samples fell outside of the Inuit cluster were defined as outliers (excessive admixture) and were removed. Further, using ADMIXTURE[7], it was calculated that those outliers have more than 5% of other ancestries in their genomes.

1. Zhou, S., L. Xiong, P. Xie, A. Ambalavanan, C.V. Bourassa, A. Dionne-Laporte, D. Spiegelman, M. Turcotte Gauthier, E. Henrion, O. Diallo, P.A. Dion, and G.A. Rouleau, *Increased missense mutation burden of Fatty Acid metabolism related genes in nunavik inuit population.* PLoS One, 2015. **10**(5): p. e0128255.

2. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. **25**(14): p. 1754-60.

3. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Brief Bioinform, 2013. **14**(2): p. 178-92.

4. Wang, K., M. Li, and H. Hakonarson, *ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data.* Nucleic Acids Res, 2010. **38**(16): p. e164.

5. Chang, C.C., C.C. Chow, L.C. Tellier, S. Vattikuti, S.M. Purcell, and J.J. Lee, *Second-generation PLINK: rising to the challenge of larger and richer datasets.* Gigascience, 2015. **4**: p. 7.

6. http://www.1000genomes.org/node/506.

7. Alexander, D.H., J. Novembre, and K. Lange, *Fast model-based estimation of ancestry in unrelated individuals.* Genome Res, 2009. **19**(9): p. 1655-64.