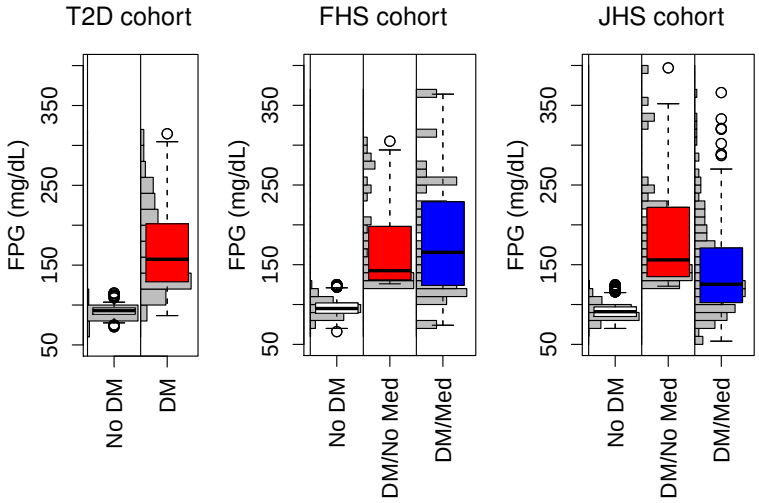
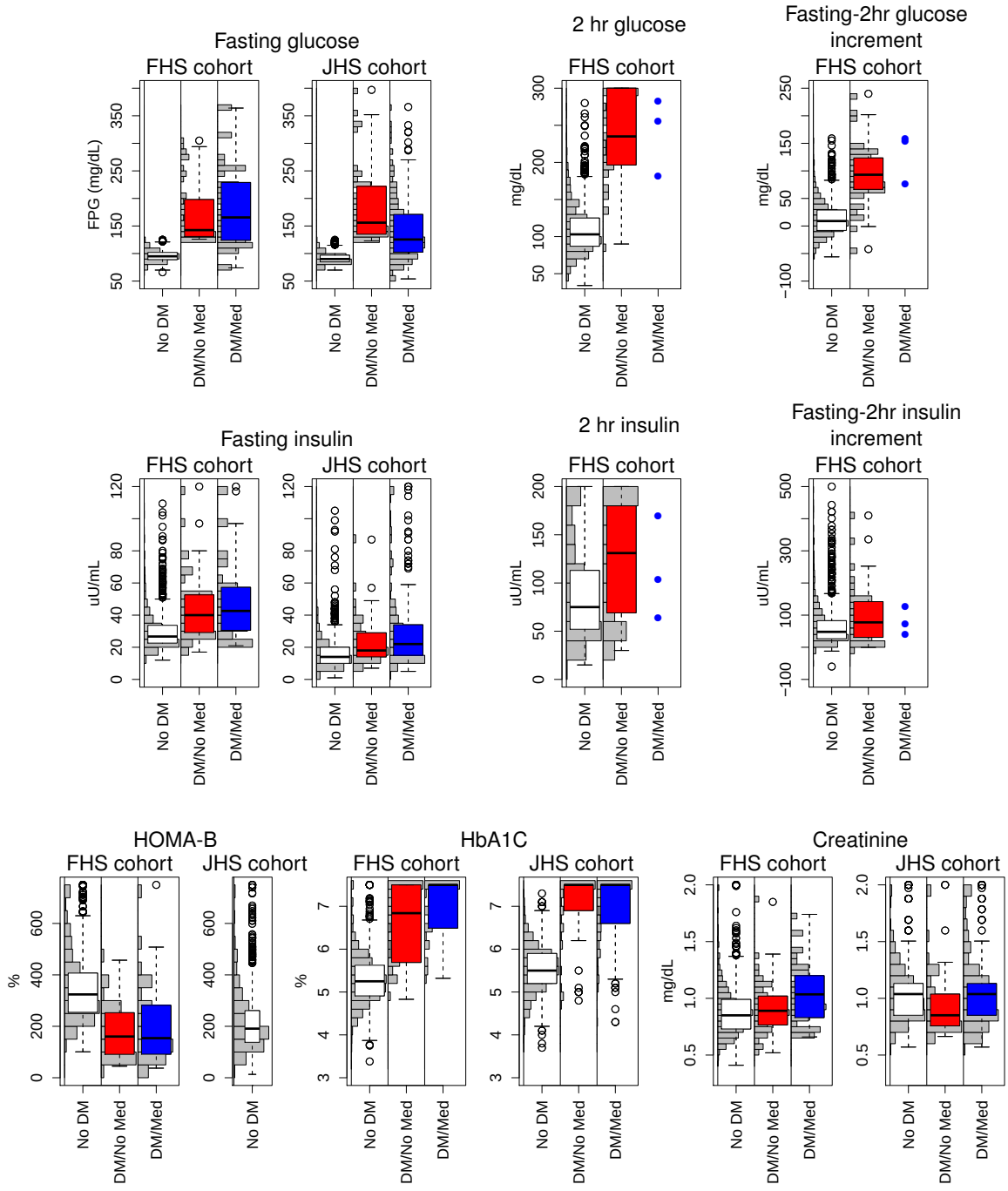
Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes

Supplementary Information

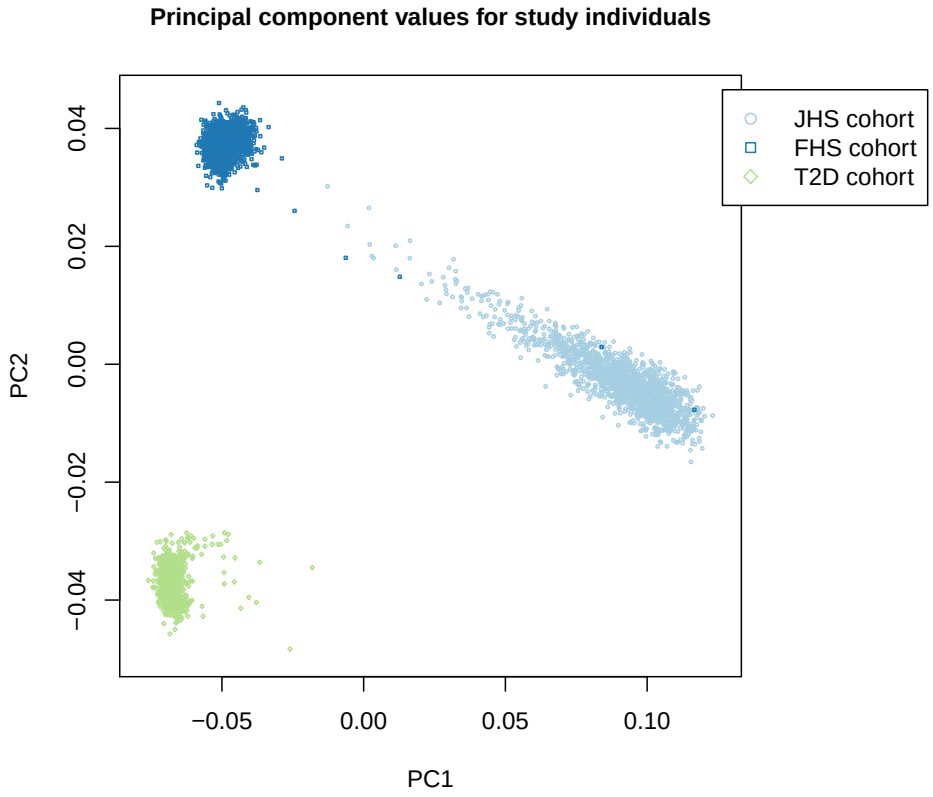
Jason Flannick*, Nicola L Beer*, Alexander G Bick, Vineeta Agarwala, Janne Molnes, Namrata Gupta, Noel P Burtt, Jose C Florez, James B Meigs, Herman Taylor, Valeriya Lyssenko, Henrik Irgens, Ervin Fox, Frank Burslem, Stefan Johansson, Jeff K Trimmer, Christopher Newton-Cheh, Tiinamaija Tuomi, Anders Molven, James G Wilson, Christopher O'Donnell, Sekar Kathiresan, Joel Hirschhorn, Pal R. Njølstad, Tim Rolph, J.G. Seidman, Stacey Gabriel, David R Cox, Christine Seidman, Leif Groop, and David Altshuler



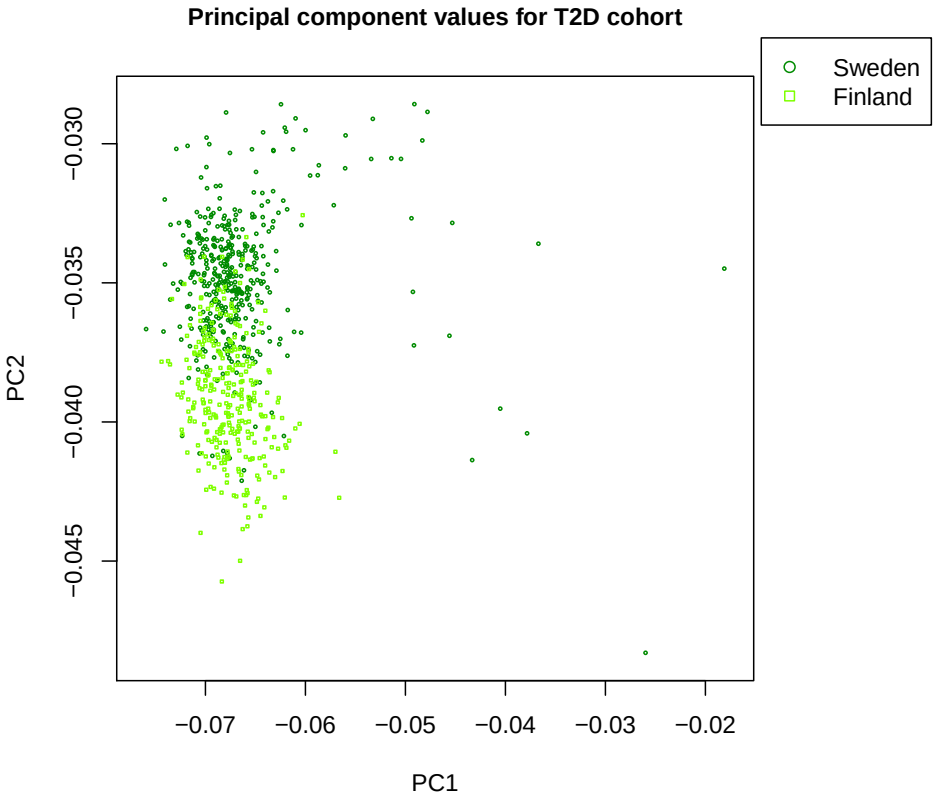
Supplementary Figure 1: **Fasting plasma glucose values for individuals without diabetes, with diabetes minus medication, or with diabetes and receiving medication.** We grouped individuals by diabetes status: without diabetes (No DM, white box), with diabetes but not on glucose-lowering medication (DM/No Med, red box), and on glucose-lowering medication (DM/Med, blue box). The plot shows the distribution of fasting plasma glucose (FPG) values for each group within each cohort; note that these values are impacted by diabetes treatment and are shown for individuals on glucose-lowering medication for completeness only. Medication status was unavailable for the T2D cohort; hence all individuals are labeled as DM.



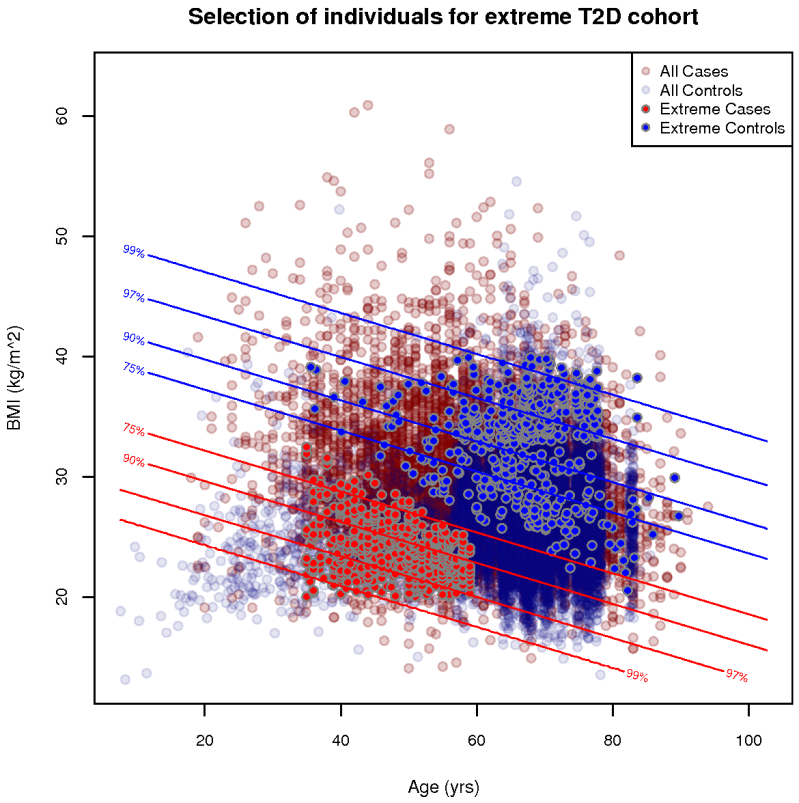
Supplementary Figure 2: **Additional phenotypes of individuals in the phenotypically unselected cohorts.** Plotted are distributions of each trait we examined in the FHS and JHS cohorts, with individuals grouped and colored in the same manner as Supplementary Figure 1. HOMA-B values were unavailable for individuals with diabetes in the JHS.



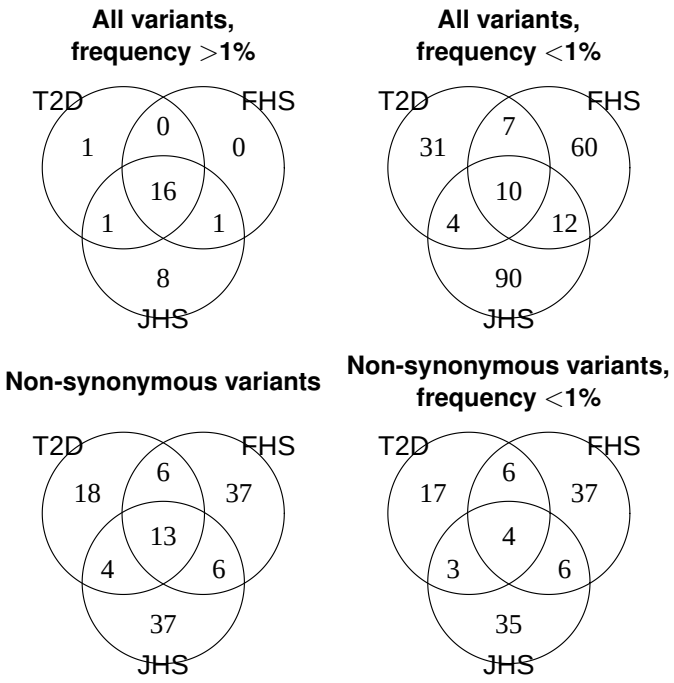
Supplementary Figure 3: **Individual principal component analysis.** To measure the ancestry of each individual, we performed principal component analysis (PCA) using a set of SNPs common to the Affymetrix 500k, Affymetrix 6.0, and the Metabochip. SNPs with MAF <5% or call rate <90% were excluded from analysis, and the remaining SNPs were LD-pruned using the PLINK software package. PCA was performed with the EIGENSTRAT software package. Plotted above are the first (PC1) and second (PC2) principal components for all study individuals, with a point for each individual colored according to cohort membership: green (T2D cohort), dark blue (FHS cohort), light blue (JHS cohort).



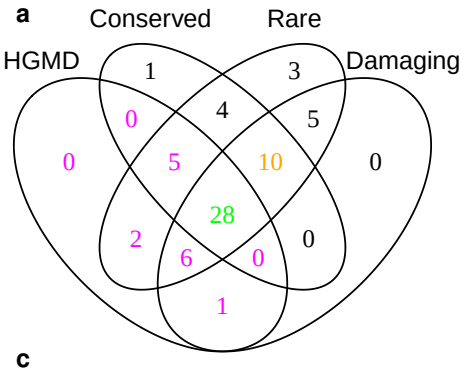
Supplementary Figure 4: **Individual principal component analysis within the T2D cohort.** To measure the ancestry differences between Finnish and Swedish individuals within the T2D cohort, plotted above are the first (PC1) and second (PC2) principal components for these individuals. PCA was performed in an identical manner as that described for Supplementary Figure 3, albeit within only individuals in the T2D cohort. The point for each individual is colored according to country of origin (not necessarily ethnicity): light green (Finland), dark green (Sweden).



Supplementary Figure 5: **Extreme T2D ascertainment.** We modeled diabetes risk based on age, BMI, and gender (see Methods) within a >27,000 individual population-based cohort. For extreme cases, we selected individuals with T2D but of low age and BMI, and for extreme controls we selected individuals without T2D but of high age and BMI. Individuals from the full population cohort are plotted as pale points in the background. Extreme individuals selected for our study are shown as solid points in the foreground. The lines represent contours of percentile for extremity, computed separately for cases (red) and controls (blue).



Supplementary Figure 6: **Variant counts by cohort.** The plot shows the variants observed across the three cohorts. Variants are divided into four (overlapping) classes: all variants with MAF >1%, all variants with MAF <1%, non-synonymous variants (any MAF), and non-synonymous variants with MAF <1%. Minor allele frequency is calculated as the maximum over all cohorts; thus, a variant with MAF 2% in the JHS cohort and MAF 0.5% in the FHS cohort would be counted in the class of all variants with MAF >1%. Most variants observed in multiple cohorts are common (MAF >1%), and more variants are private to the JHS cohort (which contains individuals of African-American ancestry) than to the other cohorts.



b

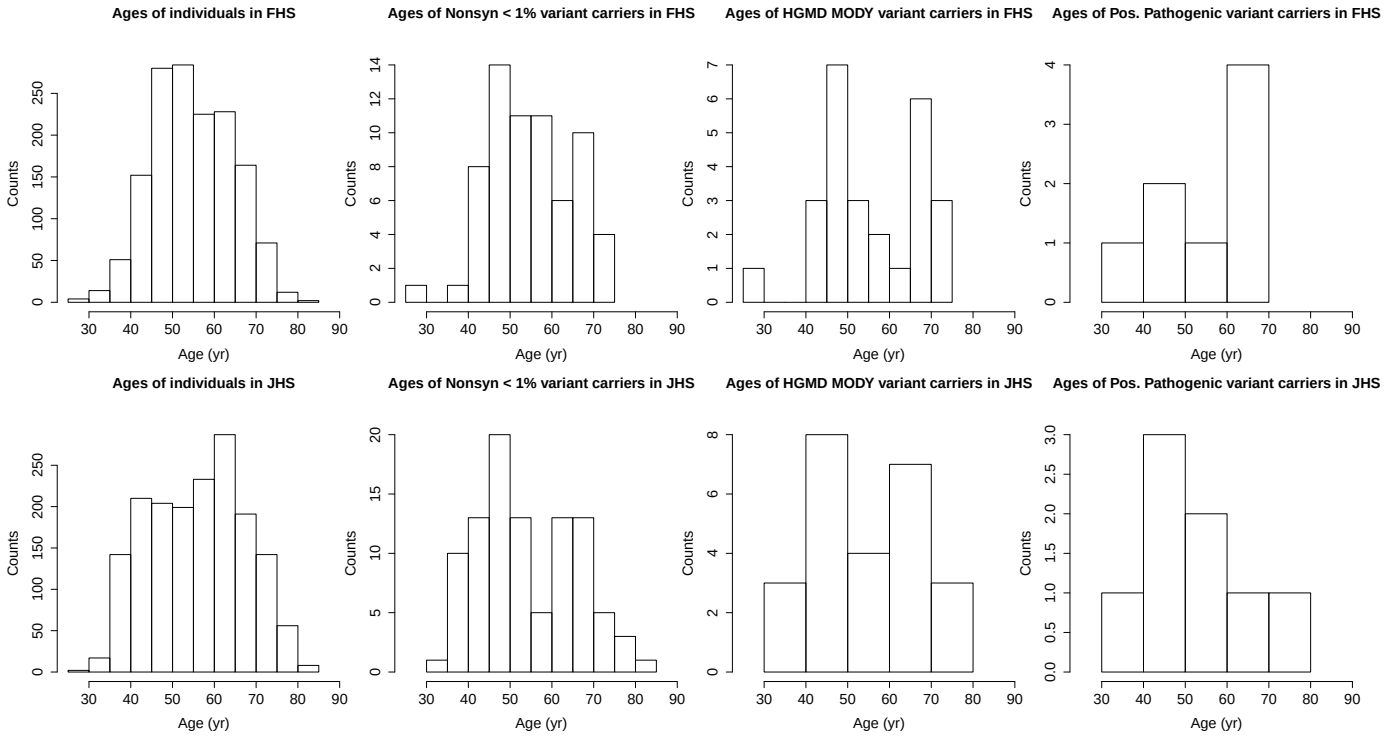
Gene	Individuals Sequenced	Variants Total	HGMD	Cnsrv.	Rare	Dam.
GCK	88	23	18	20	22	19
HNF1A	207	25	18	15	25	18
HNF1B	70	2	1	2	2	2
HNF4A	83	11	3	8	9	7
INS	88	3	1	2	3	3
NEUROD1	31	1	1	0	1	0
PDX1	5	2	0	1	1	1
Total	**250**	**67**	**42**	**48**	**63**	**50**

c

Variant classes

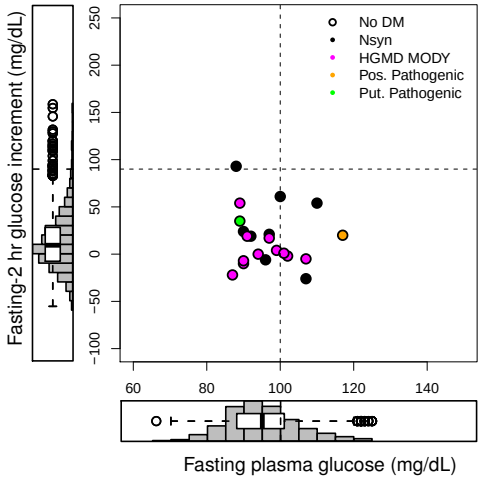
Nonsyn < 1%:	67 variants	Possibly pathogenic:	38 variants	HGMD: MODY	42 variants	Putative pathogenic:	28 variants

Supplementary Figure 7: **Description of low frequency non-synonymous variants in MODY patients.** We annotated variants observed in MODY patients using the same bioinformatics criteria as for the unselected cohorts. Shown are the low frequency non-synonymous variants that fit each annotation, as well as the counts in each variant class (using the same format as in Figure 1). The number of sequenced individuals is also shown, which varies for each gene as individuals screened for MODY did not have all genes sequenced if a mutation was identified (see Methods).

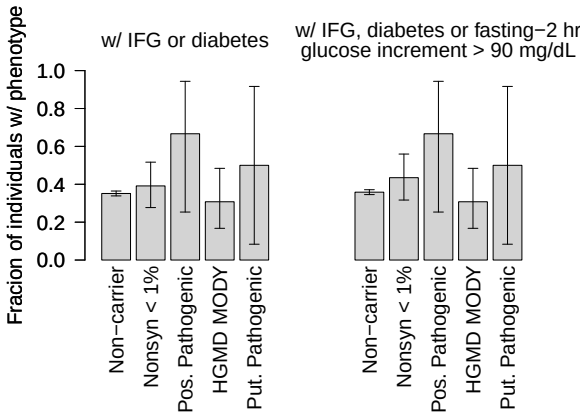


Supplementary Figure 8: **Age distribution of variant carriers.** Shown are the distributions of ages within the FHS and JHS cohorts, as well as those for variant carriers.

a



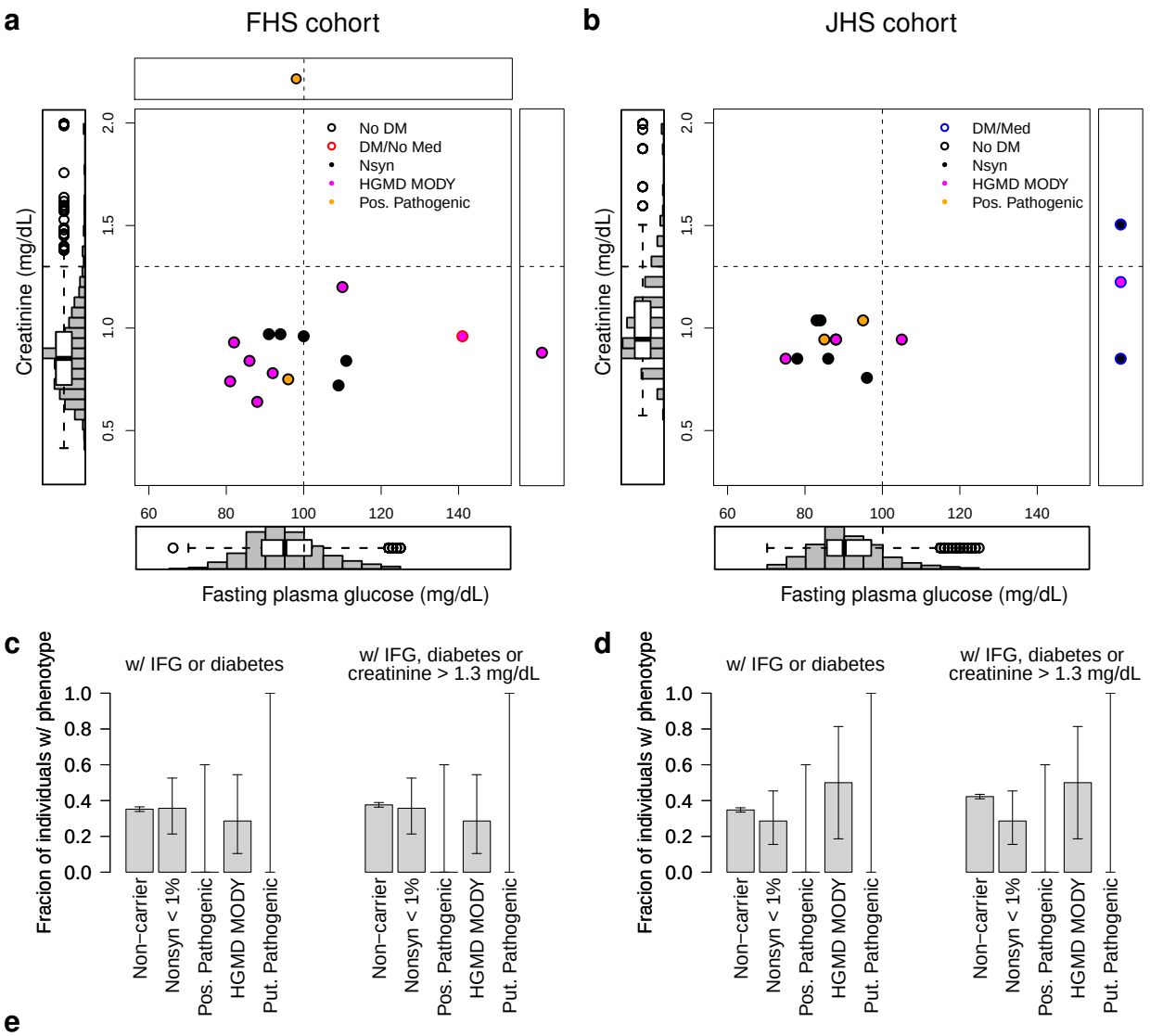
b



c

Class	IFG or diabetes				IFG, diabetes or fasting-2 hr glucose increment $>$ 90 mg/dL			
	Indiv.	**w/ pheno (%)**	**OR (95% CI)**	**P**	**Indiv.**	**w/ pheno (%)**	**OR (95% CI)**	**P**
Non-carrier	1443	507 (35.1%)	-	-	1443	507 (35.1%)	-	-
Nonsyn $<$ 1%	23	9 (39.1%)	1.2 (0.5- 2.8)	0.42	23	9 (39.1%)	1.2 (0.5- 2.8)	0.29
Pos. Pathogenic	3	2 (66.7%)	3.7 (0.3-40.8)	0.29	3	2 (66.7%)	3.7 (0.3-40.8)	0.25
HGMD MODY	13	4 (30.8%)	0.8 (0.3- 2.7)	0.73	13	4 (30.8%)	0.8 (0.3- 2.7)	0.76
Put. Pathogenic	2	1 (50.%)	1.8 (0.1-29.6)	0.59	2	1 (50.%)	1.8 (0.1-29.6)	0.59

Supplementary Figure 9: **Phenotypes of *HNF1A*-variant carriers in the FHS cohort.** We examined FPG and 2 hr glucose levels (post-OGTT) for *HNF1A*-variant carriers, as HNF1A-MODY can present as inappropriately-elevated fasting-2 hr post OGTT increments ($\geq$90 mg/dL) despite fasting normoglycemia. (**a**) The scatter plot is a repeat of Figure 3b. (**b**) The bar charts show the fraction of variant carriers with IFG or diabetes (left), as well as the fraction with IFG or fasting-2 hr glucose increments $\geq$90mg/dL (right); the lines give 65% confidence intervals. (**c**) The bottom panel tabulates the data from the bar charts, with additional columns for estimated odds ratios (OR) and p-values from tests of difference between the proportion of variant carriers and non-variant carriers with the relevant phenotype. Data is only shown for individuals from the FHS cohort, as OGTT measurements were unavailable in the JHS cohort.



e

Phenotype	Class	FHS cohort				JHS cohort				Combined
		Indiv.	w/ pheno (%)	OR (95% CI)	P	Indiv.	w/ pheno (%)	OR (95% CI)	P	P
IFG or diabetes	Non-carrier	1452	511 (35.2%)	-	-	1626	565 (34.7%)	-	-	-
	Nonsyn < 1%	14	5 (35.7%)	1.0 (0.3-3.1)	0.59	14	4 (28.6%)	0.8 (0.2- 2.4)	0.78	0.69
	Pos. Pathogenic	2	0 (0.%)	-	1.	2	0 (0.%)	-	1.	1.
	HGMD MODY	7	2 (28.6%)	0.7 (0.1-3.8)	0.79	4	2 (50.%)	1.9 (0.3-13.4)	0.44	0.59
	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
IFG, diabetes or creatinine > 1.3 mg/dL	Non-carrier	1452	511 (35.2%)	-	-	1626	690 (42.2%)	-	-	-
	Nonsyn < 1%	14	5 (35.7%)	1.0 (0.3-3.1)	0.68	14	4 (28.6%)	0.5 (0.2- 1.7)	0.9	0.85
	Pos. Pathogenic	2	0 (0.%)	-	1.	2	0 (0.%)	-	1.	1.
	HGMD MODY	7	2 (28.6%)	0.7 (0.1-3.8)	0.84	4	2 (50.%)	1.4 (0.2- 9.7)	0.54	0.74
	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.

Supplementary Figure 10: **Phenotypes of *HNF1B*-variant carriers in the FHS and JHS cohorts.** Creatinine levels were examined in addition to FPG for *HNF1B*-variant carriers, as HNF1B-MODY patients can also experience renal dysfunction. The plot is analogous to Supplementary Figure 9 but with serum creatinine plotted as the additional trait. The horizontal dotted line in the scatter plot represents the upper end of the physiological range for serum creatinine (0.7-1.3mg/dL). Data is shown both for individuals from the FHS (**ace**) and JHS (**bde**) cohorts.

Gene	Length (bp)	Bases > 20x (%)	Bases > 10x (%)	Variants
GCK	1485	92.7%	97.9%	14
HNF1A	1893	93.%	98.%	31
HNF1B	1671	97.3%	99.3%	18
HNF4A	1593	97.3%	99.7%	24
INS	330	79.1%	89.6%	6
NEUROD1	1068	100.%	100.%	13
PDX1	849	37.7%	52.5%	15
Total	**8889**	**89.7%**	**94.2%**	**121**

Supplementary Table 1: **Sequencing coverage of each gene.** To estimate the power to detect variants private to an individual for each gene, we counted the fraction of bases, averaged across all individuals, sequenced to >10x coverage (the minimum threshold to call a variant) as well as the fraction sequenced to a more stringent >20x coverage. Shown in the table is the targeted length of each gene, the average fraction of bases with coverage >20x and coverage >10x, and the number of identified non-synonymous variants (of any frequency).

Gene	Nucleotide change	Amino acid change	Type	HGMD	SIFT	PPH	Cons	1000G	FHS cohort		JHS cohort		T2D cohort	
									Count	Freq	Count	Freq	Count	Freq
GCK	c.45+1G>T	NA	Splice	Yes	NA	NA	3.3	-	-	-	-	-	1	0.065%
GCK	c.394G>A	p.Asp132Asn	Mis	Yes	Tol	Benign	2.9	-	-	-	1	0.029%	-	-
GCK	c.1386G>T	p.Met462Ile	Mis	Yes	Tol	Benign	4.0	-	1	0.033%	-	-	-	-
GCK	c.31G>A	p.Ala11Thr	Mis	-	Tol	Benign	-0.3	Yes	-	-	74	2.2%	-	-
GCK	c.39G>T	p.Lys13Asn	Mis	-	Tol	Benign	1.8	-	1	0.032%	-	-	-	-
GCK	c.106C>T	p.Arg36Trp	Mis	-	Del	Prob	0.8	-	-	-	-	-	1	0.066%
GCK	c.334G>A	p.Glu112Lys	Mis	-	Tol	Benign	3.0	-	1	0.032%	-	-	-	-
GCK	c.385T>C	p.Cys129Arg	Mis	-	Del	Prob	4.8	-	1	0.032%	-	-	-	-
GCK	c.493C>A	p.Leu165Ile	Mis	-	Tol	Benign	3.1	-	-	-	-	-	1	0.067%
GCK	c.575G>A	p.Arg192Lys	Mis	-	Tol	Prob	6.2	-	-	-	2	0.059%	-	-
GCK	c.649G>A	p.Asp217Asn	Mis	-	Tol	Prob	6.3	-	1	0.032%	-	-	1	0.065%
GCK	c.835G>C	p.Glu279Gln	Mis	-	Tol	Pos	4.2	Yes	-	-	2	0.059%	-	-
GCK	c.988T>G	p.Phe330Val	Mis	-	Del	Pos	4.7	-	-	-	1	0.029%	-	-
GCK	c.1025C>G	p.Thr342Arg	Mis	-	Tol	Benign	1.6	-	-	-	2	0.059%	-	-
HNF1A	c.92G>A	p.Gly31Asp	Mis	Yes	Tol	Pos	0.8	Yes	5	0.16%	1	0.029%	*1*	*0.066%*
HNF1A	c.142G>A	p.Glu48Lys	Mis	Yes	Tol	Benign	1.5	-	1	0.032%	-	-	-	-
HNF1A	c.185A>G	p.Asn62Ser	Mis	Yes	Del	Benign	2.1	-	1	0.032%	1	0.029%	-	-
HNF1A	c.392G>A	p.Arg131Gln	Mis	Yes	Del	Prob	5.7	-	-	-	-	-	1	0.065%
HNF1A	c.586A>G	p.Thr196Ala	Mis	Yes	Del	Benign	1.8	-	-	-	-	-	1	0.065%
HNF1A	c.827C>G	p.Ala276Gly	Mis	Yes	Del	Pos	5.5	-	1	0.032%	-	-	-	-
HNF1A	c.965A>G	p.Tyr322Cys	Mis	Yes	Del	Prob	4.2	-	-	-	6	0.18%	-	-
HNF1A	c.1165T>G	p.Leu389Val	Mis	Yes	Tol	Pos	0.0	Yes	-	-	18	0.53%	-	-
HNF1A	c.1513C>A	p.His505Asn	Mis	Yes	Tol	Prob	4.6	-	1	0.032%	-	-	-	-
HNF1A	c.1522G>A	p.Glu508Lys	Mis	Yes	Del	Pos	3.9	-	1	0.032%	-	-	-	-
HNF1A	c.1544C>A	p.Thr515Lys	Mis	Yes	Del	Pos	3.8	-	-	-	1	0.029%	-	-
HNF1A	c.1748G>A	p.Arg583Gln	Mis	Yes	Del	Benign	1.4	-	4	0.14%	-	-	-	-
HNF1A	c.79A>C	p.Ile27Leu	Mis	-	Del	Benign	1.6	Yes	1101	35.%	413	12.%	*536*	*35.%*
HNF1A	c.290C>T	p.Ala97Val	Mis	-	Del	Pos	3.3	-	-	-	1	0.029%	-	-
HNF1A	c.293C>T	p.Ala98Val	Mis	-	Del	Benign	5.0	Yes	62	2.%	20	0.59%	*53*	*3.5%*
HNF1A	c.298C>A	p.Gln100Lys	Mis	-	Del	Benign	5.0	-	-	-	1	0.029%	-	-
HNF1A	c.307G>A	p.Val103Met	Mis	-	Tol	Pos	5.0	-	1	0.033%	-	-	-	-
HNF1A	c.341G>A	p.Arg114His	Mis	-	Tol	Benign	2.1	-	-	-	1	0.029%	1	0.065%
HNF1A	c.533C>T	p.Thr178Ile	Mis	-	Del	Benign	1.1	-	1	0.032%	-	-	-	-
HNF1A	c.818_820delAAG	p.Glu274del	Del	-	NA	NA	4.6	-	1	0.032%	-	-	-	-
HNF1A	c.824A>C	p.Glu275Ala	Mis	-	Del	Prob	4.6	Yes	1	0.032%	-	-	-	-
HNF1A	c.854C>T	p.Thr285Met	Mis	-	Del	Benign	3.8	-	1	0.032%	-	-	-	-
HNF1A	c.1360A>G	p.Ser454Gly	Mis	-	Tol	Benign	3.1	-	-	-	-	-	*1*	*0.067%*
HNF1A	c.1405C>T	p.His469Tyr	Mis	-	Del	Prob	3.6	-	2	0.065%	1	0.029%	-	-
HNF1A	c.1460G>A	p.Ser487Asn	Mis	-	Tol	Benign	0.9	Yes	1038	34.%	426	13.%	*476*	*32.%*
HNF1A	c.1469T>C	p.Met490Thr	Mis	-	Del	Prob	4.5	-	-	-	-	-	*1*	*0.068%*
HNF1A	c.1532A>G	p.Gln511Arg	Mis	-	Del	Prob	4.7	-	1	0.032%	-	-	-	-
HNF1A	c.1541A>G	p.His514Arg	Mis	-	Del	Pos	1.4	-	1	0.032%	-	-	-	-
HNF1A	c.1696C>A	p.His566Asn	Mis	-	Del	Benign	3.1	-	-	-	1	0.029%	-	-
HNF1A	c.1720A>G	p.Ser574Gly	Mis	-	Tol	Benign	0.1	Yes	-	-	170	5.%	-	-
HNF1A	c.1729C>G	p.His577Asp	Mis	-	Del	Benign	2.2	-	1	0.035%	-	-	-	-
HNF1B	c.182T>G	p.Val61Gly	Mis	Yes	Tol	Benign	2.8	-	3	0.097%	-	-	*1*	*0.068%*
HNF1B	c.226G>T	p.Gly76Cys	Mis	Yes	Del	Prob	5.7	Yes	3	0.097%	3	0.088%	*2*	*0.14%*
HNF1B	c.1108G>A	p.Gly370Ser	Mis	Yes	Tol	Benign	2.6	Yes	-	-	1	0.029%	-	-
HNF1B	c.1474G>A	p.Gly492Ser	Mis	Yes	Tol	Prob	1.7	Yes	2	0.064%	-	-	*7*	*0.47%*
HNF1B	c.35T>G	p.Leu12Arg	Mis	-	Del	Prob	4.9	-	1	0.032%	-	-	-	-
HNF1B	c.127C>G	p.Leu43Val	Mis	-	Tol	Benign	1.6	-	-	-	2	0.059%	-	-
HNF1B	c.145T>A	p.Ser49Thr	Mis	-	Tol	Benign	2.3	-	1	0.032%	-	-	-	-
HNF1B	c.244G>A	p.Asp82Asn	Mis	-	Del	Pos	5.7	-	2	0.065%	2	0.059%	-	-
HNF1B	c.335G>A	p.Arg112Gln	Mis	-	Tol	Benign	3.1	-	1	0.034%	-	-	*1*	*0.072%*
HNF1B	c.684C>G	p.Asn228Lys	Mis	-	Tol	Benign	3.7	-	-	-	3	0.088%	-	-
HNF1B	c.905A>G	p.Asn302Ser	Mis	-	Del	Benign	4.7	-	-	-	-	-	1	0.065%
HNF1B	c.944T>C	p.Met315Thr	Mis	-	Del	Benign	3.1	-	1	0.032%	-	-	-	-
HNF1B	c.953A>G	p.Tyr318Cys	Mis	-	Del	Prob	3.0	-	-	-	1	0.029%	-	-
HNF1B	c.1561dupC	p.Gln521ArgfsX30	Frame	-	NA	NA	5.1	-	-	-	-	-	31	2.1%
HNF1B	c.1594A>G	p.Met532Val	Mis	-	Del	Prob	5.0	-	1	0.032%	-	-	-	-
HNF1B	c.1598T>C	p.Val533Ala	Mis	-	Del	Prob	5.0	-	-	-	1	0.029%	-	-
HNF1B	c.1621A>C	p.Ser541Arg	Mis	-	Del	Benign	3.1	-	-	-	-	-	1	0.065%
HNF1B	c.1657C>A	p.Pro553Thr	Mis	-	Del	Benign	3.8	-	-	-	1	0.029%	-	-
HNF4A	c.406C>T	p.Arg136Trp	Mis	Yes	Del	Prob	1.2	-	1	0.032%	-	-	-	-
HNF4A	c.488G>A	p.Arg163Gln	Mis	Yes	Tol	Benign	2.5	-	-	-	-	-	1	0.065%
HNF4A	c.790G>A	p.Val264Met	Mis	Yes	Del	Prob	6.0	-	1	0.032%	-	-	-	-
HNF4A	c.929G>A	p.Arg310Gln	Mis	Yes	Tol	Benign	3.3	-	1	0.032%	-	-	-	-
HNF4A	c.1193T>C	p.Met398Thr	Mis	Yes	Tol	Benign	5.2	Yes	1	0.032%	-	-	-	-
HNF4A	c.1333C>T	p.Pro445Ser	Mis	Yes	Tol	Benign	3.9	-	-	-	-	-	1	0.065%
HNF4A	c.134G>A	p.Gly45Asp	Mis	-	Tol	Benign	1.9	-	-	-	1	0.029%	-	-
HNF4A	c.152C>A	p.Pro51His	Mis	-	Del	Benign	4.1	-	-	-	1	0.029%	-	-
HNF4A	c.299G>A	p.Arg100Gln	Mis	-	Del	Prob	6.1	-	-	-	1	0.029%	-	-

HNF4A	c.365G>T	p.Arg122Leu	Mis	-	Del	Prob	6.2	-	-	-	-	-	1	0.065%
HNF4A	c.416C>T	p.Thr139Ile	Mis	-	Tol	Benign	5.9	Yes	92	2.9%	24	0.71%	72	4.7%
HNF4A	c.427A>G	p.Ser143Gly	Mis	-	Tol	Benign	4.9	-	1	0.032%	-	-	-	-
HNF4A	c.505G>A	p.Val169Ile	Mis	-	Tol	Benign	-0.1	Yes	-	-	2	0.059%	-	-
HNF4A	c.530G>A	p.Arg177Gln	Mis	-	Tol	Benign	3.5	-	-	-	1	0.029%	-	-
HNF4A	c.655C>A	p.Leu219Met	Mis	-	Del	Prob	1.6	-	-	-	-	-	1	0.065%
HNF4A	c.748A>G	p.Ile250Val	Mis	-	Tol	Benign	2.5	-	-	-	1	0.029%	-	-
HNF4A	c.928C>T	p.Arg310Trp	Mis	-	Del	Prob	1.1	-	1	0.032%	-	-	-	-
HNF4A	c.1103A>G	p.Asn368Ser	Mis	-	Tol	Benign	1.2	-	1	0.032%	-	-	-	-
HNF4A	c.1161C>A	p.His387Gln	Mis	-	Tol	Pos	1.6	-	-	-	1	0.029%	-	-
HNF4A	c.1210G>A	p.Val404Ile	Mis	-	Tol	Benign	4.6	-	-	-	-	-	1	0.065%
HNF4A	c.1250T>C	p.Met417Thr	Mis	-	Tol	Benign	5.2	-	1	0.032%	-	-	-	-
HNF4A	c.1360G>A	p.Ala454Thr	Mis	-	Tol	Benign	-0.3	-	-	-	-	-	1	0.065%
HNF4A	c.1369G>C	p.Val457Leu	Mis	-	Tol	Benign	1.5	-	-	-	1	0.029%	-	-
HNF4A	c.1387A>G	p.Ile463Val	Mis	-	Del	Benign	4.6	-	1	0.032%	-	-	1	0.065%
INS	c.17G>A	p.Arg6His	Mis	-	Tol	Unknown	0.0	-	1	0.032%	2	0.059%	2	0.28%
INS	c.67G>T	p.Ala23Ser	Mis	-	Tol	Unknown	-0.1	Yes	1	0.032%	1	0.029%	-	-
INS	c.208G>A	p.Gly70Arg	Mis	-	Tol	Benign	1.1	-	1	0.033%	-	-	-	-
INS	c.212G>T	p.Gly71Val	Mis	-	Tol	Benign	0.6	-	1	0.033%	-	-	-	-
INS	c.224G>A	p.Gly75Asp	Mis	-	Tol	Pos	0.5	-	-	-	6	0.18%	1	0.11%
INS	c.227G>A	p.Ser76Asn	Mis	-	Tol	Benign	1.0	-	-	-	6	0.18%	-	-
NEUROD1	c.61A>G	p.Ser21Gly	Mis	-	Tol	Benign	2.5	-	1	0.032%	-	-	-	-
NEUROD1	c.117G>T	p.Lys39Asn	Mis	-	Tol	Benign	0.9	-	-	-	-	-	1	0.065%
NEUROD1	c.133A>G	p.Thr45Ala	Mis	-	Tol	Benign	0.6	Yes	1139	36.%	880	26.%	602	39.%
NEUROD1	c.257C>T	p.Pro86Leu	Mis	-	Tol	Prob	6.1	-	-	-	1	0.029%	-	-
NEUROD1	c.269_271delAGA	p.Lys90del	Del	-	NA	NA	5.0	-	-	-	1	0.029%	-	-
NEUROD1	c.272_274delTGA	p.Met91del	Del	-	NA	NA	6.1	-	-	-	1	0.029%	-	-
NEUROD1	c.454delG	p.Ala152LeufsX110	Frame	-	NA	NA	6.0	-	-	-	-	-	1	0.065%
NEUROD1	c.554C>T	p.Ala185Val	Mis	-	Del	Prob	6.0	-	-	-	1	0.029%	-	-
NEUROD1	c.574C>T	p.Pro192Ser	Mis	-	Tol	Pos	4.4	-	1	0.032%	-	-	-	-
NEUROD1	c.590C>A	p.Pro197His	Mis	-	Tol	Pos	6.0	Yes	96	3.1%	8	0.24%	25	1.6%
NEUROD1	c.610C>T	p.Pro204Ser	Mis	-	Tol	Benign	2.9	Yes	1	0.032%	8	0.24%	-	-
NEUROD1	c.917C>T	p.Ala306Val	Mis	-	Tol	Benign	3.0	-	-	-	3	0.088%	-	-
NEUROD1	c.1031G>A	p.Arg344Gln	Mis	-	Del	Pos	6.2	-	-	-	1	0.029%	-	-
PDX1	c.38T>C	p.Leu13Pro	Mis	-	Del	Pos	4.5	-	1	0.036%	-	-	-	-
PDX1	c.52T>C	p.Cys18Arg	Mis	-	Del	Prob	4.5	-	1	0.037%	1	0.062%	2	0.15%
PDX1	c.54C>A	p.Cys18X	Stop	-	NA	NA	1.5	-	-	-	2	0.13%	1	0.074%
PDX1	c.107T>G	p.Leu36Arg	Mis	-	Tol	Benign	4.6	-	2	0.18%	-	-	-	-
PDX1	c.119G>T	p.Arg40Leu	Mis	-	Del	Pos	3.9	-	-	-	3	0.38%	-	-
PDX1	c.211C>A	p.Pro71Thr	Mis	-	Del	Pos	2.6	-	-	-	3	0.21%	2	1.9%
PDX1	c.226G>A	p.Asp76Asn	Mis	-	Del	Benign	3.7	-	10	0.67%	1	0.063%	*4*	*2.3%*
PDX1	c.244C>T	p.Leu82Phe	Mis	-	Del	Benign	2.1	-	2	0.13%	-	-	-	-
PDX1	c.304G>A	p.Glu102Lys	Mis	-	Tol	Benign	3.7	-	1	0.2%	1	0.034%	-	-
PDX1	c.338A>G	p.Asn113Ser	Mis	-	Tol	Benign	0.7	-	-	-	1	0.03%	-	-
PDX1	c.418G>A	p.Ala140Thr	Mis	-	Tol	Benign	-0.3	Yes	2	0.08%	-	-	*1*	*0.068%*
PDX1	c.679T>C	p.Cys227Arg	Mis	-	Del	Pos	2.4	-	1	0.037%	-	-	-	-
PDX1	c.716C>A	p.Pro239Gln	Mis	-	Tol	Benign	0.0	Yes	6	0.77%	79	3.9%	7	0.73%
PDX1	c.719C>A	p.Pro240Gln	Mis	-	Del	Benign	0.9	-	-	-	4	0.24%	-	-
PDX1	c.725C>T	p.Pro242Leu	Mis	-	Tol	Benign	0.2	-	4	1.3%	2	0.17%	3	0.8%

Supplementary Table 2: **List of variants identified in unselected or extreme cohorts.** Shown is a full catalog of all non-synonymous variants identified in the study. 'Type': impact of variant on protein, where 'Mis' denotes missense variant, 'Splice' denotes variant predicted to affect splicing, and 'Frame' denotes frameshift variant; 'SIFT': SIFT prediction, where 'Tol' denotes tolerated and 'Del' denotes deleterious; 'PPH': PolyPhen2 predictions, where "Pos' denotes possibly damaging and 'Prob' denotes probably damaging; 'Cons': conservation score from 46-way vertebrate alignment, with positive numbers denoting increased conservation; '1000G': whether the variant was identified in the individuals sequenced for the 1000G project; 'Freq': the frequency in each cohort. Variants are sorted first by presence in HGMD and second by transcript position. Variants removed from analysis due to differential call rates between cases and controls are shown in italics.

Gene	Nucleotide change	Amino acid change	FHS cohort				JHS cohort				T2D cohort			
			Count		Frequency		Count		Frequency		Count		Frequency	
			DM	No DM	DM	No DM	DM	No DM	DM	No DM	DM	No DM	DM	No DM
GCK	c.45+1G>T	NA	-	-	-	-	-	-	-	-	1	0	0.14%	0.
GCK	c.394G>A	p.Asp132Asn	-	-	-	-	0	1	0.	0.037%	-	-	-	-
GCK	c.1386G>T	p.Met462Ile	0	1	0.	0.038%	-	-	-	-	-	-	-	-
GCK	c.31G>A	p.Ala11Thr	-	-	-	-	11	63	1.6%	2.4%	-	-	-	-
GCK	c.39G>T	p.Lys13Asn	0	1	0.	0.037%	-	-	-	-	-	-	-	-
GCK	c.106C>T	p.Arg36Trp	-	-	-	-	-	-	-	-	1	0	0.14%	0.
GCK	c.334G>A	p.Glu112Lys	0	1	0.	0.037%	-	-	-	-	-	-	-	-
GCK	c.385T>C	p.Cys129Arg	1	0	0.22%	0.	-	-	-	-	-	-	-	-
GCK	c.493C>A	p.Leu165Ile	-	-	-	-	-	-	-	-	1	0	0.14%	0.
GCK	c.575G>A	p.Arg192Lys	-	-	-	-	1	1	0.14%	0.037%	-	-	-	-
GCK	c.649G>A	p.Asp217Asn	0	1	0.	0.037%	-	-	-	-	0	1	0.	0.12%
GCK	c.835G>C	p.Glu279Gln	-	-	-	-	1	1	0.14%	0.037%	-	-	-	-
GCK	c.988T>G	p.Phe330Val	-	-	-	-	0	1	0.	0.037%	-	-	-	-
GCK	c.1025C>G	p.Thr342Arg	-	-	-	-	0	2	0.	0.074%	-	-	-	-
HNF1A	c.92G>A	p.Gly31Asp	0	5	0.	0.19%	1	0	0.14%	0.	*0*	*1*	*0.*	*0.12%*
HNF1A	c.142G>A	p.Glu48Lys	0	1	0.	0.038%	-	-	-	-	-	-	-	-
HNF1A	c.185A>G	p.Asn62Ser	0	1	0.	0.038%	0	1	0.	0.037%	-	-	-	-
HNF1A	c.392G>A	p.Arg131Gln	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF1A	c.586A>G	p.Thr196Ala	-	-	-	-	-	-	-	-	0	1	0.	0.12%
HNF1A	c.827C>G	p.Ala276Gly	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1A	c.965A>G	p.Tyr322Cys	-	-	-	-	3	3	0.43%	0.11%	-	-	-	-
HNF1A	c.1165T>G	p.Leu389Val	-	-	-	-	4	14	0.57%	0.52%	-	-	-	-
HNF1A	c.1513C>A	p.His505Asn	0	1	0.	0.038%	-	-	-	-	-	-	-	-
HNF1A	c.1522G>A	p.Glu508Lys	1	0	0.23%	0.	-	-	-	-	-	-	-	-
HNF1A	c.1544C>A	p.Thr515Lys	-	-	-	-	1	0	0.14%	0.	-	-	-	-
HNF1A	c.1748G>A	p.Arg583Gln	0	4	0.	0.16%	-	-	-	-	-	-	-	-
HNF1A	c.79A>C	p.Ile27Leu	165	936	37.%	35.%	95	318	14.%	12.%	*259*	*277*	*37.%*	*34.%*
HNF1A	c.290C>T	p.Ala97Val	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF1A	c.293C>T	p.Ala98Val	9	53	2.%	2.%	4	16	0.57%	0.59%	*32*	*21*	*4.6%*	*2.6%*
HNF1A	c.298C>A	p.Gln100Lys	-	-	-	-	1	0	0.14%	0.	-	-	-	-
HNF1A	c.307G>A	p.Val103Met	1	0	0.22%	0.	-	-	-	-	-	-	-	-
HNF1A	c.341G>A	p.Arg114His	-	-	-	-	1	0	0.14%	0.	0	1	0.	0.12%
HNF1A	c.533C>T	p.Thr178Ile	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1A	c.818_820delAAG	p.Glu274del	1	0	0.22%	0.	-	-	-	-	-	-	-	-
HNF1A	c.824A>C	p.Glu275Ala	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1A	c.854C>T	p.Thr285Met	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1A	c.1360A>G	p.Ser454Gly	-	-	-	-	-	-	-	-	*1*	*0*	*0.15%*	*0.*
HNF1A	c.1405C>T	p.His469Tyr	0	2	0.	0.076%	0	1	0.	0.037%	-	-	-	-
HNF1A	c.1460G>A	p.Ser487Asn	148	890	33.%	34.%	90	336	13.%	12.%	*227*	*249*	*34.%*	*31.%*
HNF1A	c.1469T>C	p.Met490Thr	-	-	-	-	-	-	-	-	*0*	*1*	*0.*	*0.12%*
HNF1A	c.1532A>G	p.Gln511Arg	0	1	0.	0.038%	-	-	-	-	-	-	-	-
HNF1A	c.1541A>G	p.His514Arg	1	0	0.23%	0.	-	-	-	-	-	-	-	-
HNF1A	c.1696C>A	p.His566Asn	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF1A	c.1720A>G	p.Ser574Gly	-	-	-	-	35	135	95.%	95.%	-	-	-	-
HNF1A	c.1729C>G	p.His577Asp	0	1	0.	0.041%	-	-	-	-	-	-	-	-
HNF1B	c.182T>G	p.Val61Gly	1	2	0.23%	0.076%	-	-	-	-	*1*	*0*	*0.15%*	*0.*
HNF1B	c.226G>T	p.Gly76Cys	0	3	0.	0.11%	1	2	0.14%	0.074%	*0*	*2*	*0.*	*0.25%*
HNF1B	c.1108G>A	p.Gly370Ser	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF1B	c.1474G>A	p.Gly492Ser	0	2	0.	0.075%	-	-	-	-	*2*	*5*	*0.29%*	*0.61%*
HNF1B	c.35T>G	p.Leu12Arg	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1B	c.127C>G	p.Leu43Val	-	-	-	-	0	2	0.	0.074%	-	-	-	-
HNF1B	c.145T>A	p.Ser49Thr	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1B	c.244G>A	p.Asp82Asn	0	2	0.	0.076%	1	1	0.14%	0.037%	-	-	-	-
HNF1B	c.335G>A	p.Arg112Gln	0	1	0.	0.04%	-	-	-	-	*0*	*1*	*0.*	*0.13%*
HNF1B	c.684C>G	p.Asn228Lys	-	-	-	-	1	2	0.14%	0.074%	-	-	-	-
HNF1B	c.905A>G	p.Asn302Ser	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF1B	c.944T>C	p.Met315Thr	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1B	c.953A>G	p.Tyr318Cys	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF1B	c.1561dupC	p.Gln521ArgfsX30	-	-	-	-	-	-	-	-	18	13	2.6%	1.7%
HNF1B	c.1594A>G	p.Met532Val	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF1B	c.1598T>C	p.Val533Ala	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF1B	c.1621A>C	p.Ser541Arg	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF1B	c.1657C>A	p.Pro553Thr	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF4A	c.406C>T	p.Arg136Trp	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF4A	c.488G>A	p.Arg163Gln	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF4A	c.790G>A	p.Val264Met	1	0	0.22%	0.	-	-	-	-	-	-	-	-
HNF4A	c.929G>A	p.Arg310Gln	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF4A	c.1193T>C	p.Met398Thr	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF4A	c.1333C>T	p.Pro445Ser	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF4A	c.134G>A	p.Gly45Asp	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF4A	c.152C>A	p.Pro51His	-	-	-	-	0	1	0.	0.037%	-	-	-	-

HNF4A	c.299G>A	p.Arg100Gln	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF4A	c.365G>T	p.Arg122Leu	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF4A	c.416C>T	p.Thr139Ile	21	71	4.7%	2.7%	6	18	0.86%	0.67%	35	37	4.9%	4.5%
HNF4A	c.427A>G	p.Ser143Gly	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF4A	c.505G>A	p.Val169Ile	-	-	-	-	0	2	0.	0.074%	-	-	-	-
HNF4A	c.530G>A	p.Arg177Gln	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF4A	c.655C>A	p.Leu219Met	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF4A	c.748A>G	p.Ile250Val	-	-	-	-	1	0	0.14%	0.	-	-	-	-
HNF4A	c.928C>T	p.Arg310Trp	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF4A	c.1103A>G	p.Asn368Ser	1	0	0.22%	0.	-	-	-	-	-	-	-	-
HNF4A	c.1161C>A	p.His387Gln	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF4A	c.1210G>A	p.Val404Ile	-	-	-	-	-	-	-	-	1	0	0.14%	0.
HNF4A	c.1250T>C	p.Met417Thr	0	1	0.	0.037%	-	-	-	-	-	-	-	-
HNF4A	c.1360G>A	p.Ala454Thr	-	-	-	-	-	-	-	-	0	1	0.	0.12%
HNF4A	c.1369G>C	p.Val457Leu	-	-	-	-	0	1	0.	0.037%	-	-	-	-
HNF4A	c.1387A>G	p.Ile463Val	0	1	0.	0.037%	-	-	-	-	0	1	0.	0.12%
INS	c.17G>A	p.Arg6His	0	1	0.	0.038%	0	2	0.	0.074%	1	1	0.3%	0.25%
INS	c.67G>T	p.Ala23Ser	0	1	0.	0.038%	0	1	0.	0.037%	-	-	-	-
INS	c.208G>A	p.Gly70Arg	0	1	0.	0.039%	-	-	-	-	-	-	-	-
INS	c.212G>T	p.Gly71Val	0	1	0.	0.039%	-	-	-	-	-	-	-	-
INS	c.224G>A	p.Gly75Asp	-	-	-	-	0	6	0.	0.22%	1	0	0.23%	0.
INS	c.227G>A	p.Ser76Asn	-	-	-	-	0	6	0.	0.22%	-	-	-	-
NEUROD1	c.61A>G	p.Ser21Gly	0	1	0.	0.037%	-	-	-	-	-	-	-	-
NEUROD1	c.117G>T	p.Lys39Asn	-	-	-	-	-	-	-	-	1	0	0.14%	0.
NEUROD1	c.133A>G	p.Thr45Ala	167	972	63.%	64.%	169	711	76.%	74.%	287	315	60.%	61.%
NEUROD1	c.257C>T	p.Pro86Leu	-	-	-	-	0	1	0.	0.037%	-	-	-	-
NEUROD1	c.269_271delAGA	p.Lys90del	-	-	-	-	0	1	0.	0.037%	-	-	-	-
NEUROD1	c.272_274delTGA	p.Met91del	-	-	-	-	0	1	0.	0.037%	-	-	-	-
NEUROD1	c.454delG	p.Ala152LeufsX110	-	-	-	-	-	-	-	-	1	0	0.14%	0.
NEUROD1	c.554C>T	p.Ala185Val	-	-	-	-	0	1	0.	0.037%	-	-	-	-
NEUROD1	c.574C>T	p.Pro192Ser	1	0	0.22%	0.	-	-	-	-	-	-	-	-
NEUROD1	c.590C>A	p.Pro197His	14	82	3.1%	3.1%	0	8	0.	0.3%	12	13	1.7%	1.6%
NEUROD1	c.610C>T	p.Pro204Ser	0	1	0.	0.037%	1	7	0.14%	0.26%	-	-	-	-
NEUROD1	c.917C>T	p.Ala306Val	-	-	-	-	1	2	0.14%	0.074%	-	-	-	-
NEUROD1	c.1031G>A	p.Arg344Gln	-	-	-	-	0	1	0.	0.037%	-	-	-	-
PDX1	c.38T>C	p.Leu13Pro	0	1	0.	0.042%	-	-	-	-	-	-	-	-
PDX1	c.52T>C	p.Cys18Arg	1	0	0.25%	0.	1	0	0.31%	0.	2	0	0.33%	0.
PDX1	c.54C>A	p.Cys18X	-	-	-	-	0	2	0.	0.16%	1	0	0.17%	0.
PDX1	c.107T>G	p.Leu36Arg	0	2	0.	0.22%	-	-	-	-	-	-	-	-
PDX1	c.119G>T	p.Arg40Leu	-	-	-	-	2	1	1.3%	0.16%	-	-	-	-
PDX1	c.211C>A	p.Pro71Thr	-	-	-	-	0	3	0.	0.26%	0	2	0.	3.1%
PDX1	c.226G>A	p.Asp76Asn	1	9	0.45%	0.71%	0	1	0.	0.078%	*0*	*4*	*0.*	*3.2%*
PDX1	c.244C>T	p.Leu82Phe	1	1	0.4%	0.075%	-	-	-	-	-	-	-	-
PDX1	c.304G>A	p.Glu102Lys	0	1	0.	0.24%	0	1	0.	0.042%	-	-	-	-
PDX1	c.338A>G	p.Asn113Ser	-	-	-	-	0	1	0.	0.037%	-	-	-	-
PDX1	c.418G>A	p.Ala140Thr	1	1	0.27%	0.047%	-	-	-	-	*1*	*0*	*0.15%*	*0.*
PDX1	c.679T>C	p.Cys227Arg	1	0	0.26%	0.	-	-	-	-	-	-	-	-
PDX1	c.716C>A	p.Pro239Gln	1	5	0.78%	0.77%	15	64	3.5%	4.%	3	4	0.72%	0.74%
PDX1	c.719C>A	p.Pro240Gln	-	-	-	-	1	3	0.29%	0.22%	-	-	-	-
PDX1	c.725C>T	p.Pro242Leu	0	4	0.	1.5%	1	1	0.49%	0.1%	2	1	1.2%	0.48%

Supplementary Table 3: **Variant counts in individuals with and without diabetes.** For each variant identified, shown are the counts and frequencies in individuals with (DM) and without (No DM) diabetes.

Gene	Nucleotide change	Amino acid change	Type	HGMD	SIFT	PPH	Cons	1000G	Counts MODY	Frequency FHS	JHS	T2D
GCK	c.1016A>G	p.Glu339Gly	Mis	Yes	Del	Prob	4.7	-	11	-	-	-
GCK	c.1129C>T	p.Arg377Cys	Mis	Yes	Del	Prob	5.9	-	2	-	-	-
GCK	c.1175G>C	p.Arg392Pro	Mis	-	Del	Pos	3.1	-	1	-	-	-
GCK	c.128G>C	p.Arg43Pro	Mis	Yes	Del	Prob	4.1	-	1	-	-	-
GCK	c.130G>A	p.Gly44Ser	Mis	Yes	Del	Prob	5.9	-	1	-	-	-
GCK	c.1322C>A	p.Ser441X	Stop	Yes	NA	NA	4.5	-	1	-	-	-
GCK	c.1340G>A	p.Arg447Gln	Mis	Yes	Del	Prob	3.4	-	1	-	-	-
GCK	c.1361C>A	p.Ala454Glu	Mis	Yes	Del	Prob	6.2	-	1	-	-	-
GCK	c.214G>A	p.Gly72Arg	Mis	Yes	Del	Prob	5.7	-	1	-	-	-
GCK	c.227C>A	p.Ser76Tyr	Mis	Yes	Del	Prob	5.7	-	3	-	-	-
GCK	c.363+2T>A	NA	Splice	Yes	NA	NA	4.6	-	1	-	-	-
GCK	c.469G>A	p.Glu157Lys	Mis	Yes	Tol	Prob	4.3	-	1	-	-	-
GCK	c.485G>A	p.Gly162Asp	Mis	Yes	Del	Prob	6.1	-	1	-	-	-
GCK	c.544G>A	p.Val182Met	Mis	Yes	Del	Prob	6.2	-	1	-	-	-
GCK	c.556C>G	p.Arg186Gly	Mis	-	Del	Prob	0.1	-	1	-	-	-
GCK	c.571C>T	p.Arg191Trp	Mis	-	Del	Prob	0.5	-	2	-	-	-
GCK	c.579+1G>C	NA	Splice	Yes	NA	NA	6.2	-	6	-	-	-
GCK	c.622G>A	p.Ala208Thr	Mis	Yes	Del	Prob	6.2	-	1	-	-	-
GCK	c.645C>G	p.Tyr215X	Stop	Yes	NA	NA	-1.2	Yes	1	-	-	-
GCK	c.658T>C	p.Cys220Arg	Mis	-	Del	Benign	5.2	-	1	-	-	-
GCK	c.766G>A	p.Glu256Lys	Mis	Yes	Del	Prob	5.9	-	1	-	-	-
GCK	c.823C>T	p.Arg275Cys	Mis	Yes	Del	Pos	3.0	-	1	-	-	-
GCK	c.883G>C	p.Gly295Arg	Mis	-	Del	Pos	5.6	-	1	-	-	-
HNF1A	c.1243G>A	p.Gly415Arg	Mis	-	Del	Prob	5.4	-	1	-	-	-
HNF1A	c.1340C>T	p.Pro447Leu	Mis	Yes	Del	Prob	5.4	-	1	-	-	-
HNF1A	c.1460G>A	p.Ser487Asn	Mis	-	Tol	Benign	0.9	Yes	40	34.%	13.%	*32.%*
HNF1A	c.1576G>A	p.Asp526Asn	Mis	Yes	Del	Prob	5.7	-	1	-	-	-
HNF1A	c.1592G>C	p.Ser531Thr	Mis	Yes	Tol	Benign	5.2	-	2	-	-	-
HNF1A	c.1745A>G	p.His582Arg	Mis	-	Del	Benign	2.3	-	1	-	-	-
HNF1A	c.17G>A	p.Ser6Asn	Mis	-	Del	Unknown	1.9	-	2	-	-	-
HNF1A	c.293C>T	p.Ala98Val	Mis	-	Del	Benign	5.0	Yes	4	2.%	0.59%	*3.5%*
HNF1A	c.335C>T	p.Pro112Leu	Mis	Yes	Del	Prob	5.7	-	5	-	-	-
HNF1A	c.346G>A	p.Ala116Thr	Mis	Yes	Del	Prob	3.8	-	1	-	-	-
HNF1A	c.391C>T	p.Arg131Trp	Mis	Yes	Del	Prob	0.4	-	2	-	-	-
HNF1A	c.475C>T	p.Arg159Trp	Mis	Yes	Del	Prob	0.4	-	1	-	-	-
HNF1A	c.523C>G	p.Gln175Glu	Mis	Yes	Tol	Benign	1.8	-	1	-	-	-
HNF1A	c.539C>T	p.Ala180Val	Mis	-	Tol	Benign	5.2	-	3	-	-	-
HNF1A	c.598C>T	p.Arg200Trp	Mis	Yes	Del	Prob	0.1	-	1	-	-	-
HNF1A	c.607C>T	p.Arg203Cys	Mis	Yes	Del	Prob	2.5	-	1	-	-	-
HNF1A	c.608G>A	p.Arg203His	Mis	Yes	Del	Prob	5.0	-	1	-	-	-
HNF1A	c.685C>T	p.Arg229X	Stop	Yes	NA	NA	0.2	-	1	-	-	-
HNF1A	c.686G>A	p.Arg229Gln	Mis	Yes	Del	Pos	5.4	-	3	-	-	-
HNF1A	c.714-1G>A	NA	Splice	Yes	NA	NA	5.5	-	1	-	-	-
HNF1A	c.766T>A	p.Ser256Thr	Mis	Yes	Del	Prob	4.6	-	1	-	-	-
HNF1A	c.788G>A	p.Arg263His	Mis	Yes	Del	Prob	5.5	-	2	-	-	-
HNF1A	c.797A>G	p.Asn266Ser	Mis	-	Del	Pos	4.6	-	2	-	-	-
HNF1A	c.79A>C	p.Ile27Leu	Mis	-	Del	Benign	1.6	Yes	49	35.%	12.%	*35.%*
HNF1A	c.812G>A	p.Arg271Gln	Mis	Yes	Del	Prob	5.5	-	1	-	-	-
HNF1A	c.827C>A	p.Ala276Asp	Mis	Yes	Del	Prob	5.5	-	1	-	-	-
HNF1A	c.835_841delCACAAGCinsTACG	p.His279_Leu281delinsTyrVal	Mis	-	NA	NA	NA	-	1	-	-	-
HNF1A	c.872dupC	p.Gly292ArgfsX25	Frame	-	NA	NA	NA	-	21	-	-	-
HNF1A	c.92G>A	p.Gly31Asp	Mis	Yes	Tol	Pos	0.8	Yes	2	0.16%	0.029%	*0.066%*
HNF1B	c.443C>T	p.Ser148Leu	Mis	Yes	Del	Prob	6.1	-	1	-	-	-
HNF1B	c.860G>T	p.Gly287Val	Mis	-	Del	Prob	5.9	-	1	-	-	-
HNF4A	c.216C>A	p.Tyr72X	Stop	-	NA	NA	0.1	-	1	-	-	-
HNF4A	c.266G>A	p.Arg89Gln	Mis	Yes	Del	Prob	5.9	-	1	-	-	-
HNF4A	c.400C>T	p.Arg134Trp	Mis	Yes	Del	Prob	3.0	-	1	-	-	-
HNF4A	c.416C>T	p.Thr139Ile	Mis	-	Tol	Benign	5.9	Yes	5	-	-	-
HNF4A	c.473C>T	p.Ala158Val	Mis	-	Tol	Pos	5.9	-	1	-	-	-
HNF4A	c.505G>A	p.Val169Ile	Mis	-	Tol	Benign	-0.1	Yes	1	-	-	-
HNF4A	c.523G>A	p.Asp175Asn	Mis	-	Del	Benign	6.2	-	1	-	-	-
HNF4A	c.646C>T	p.Gln216X	Stop,Splice	-	NA	NA	6.2	-	1	-	-	-
HNF4A	c.790G>A	p.Val264Met	Mis	Yes	Del	Prob	6.0	-	1	-	-	-
HNF4A	c.877dupA	p.Ile293AsnfsX4	Frame	-	NA	NA	NA	-	1	-	-	-
HNF4A	c.1033C>G	p.Gln345Glu	Mis	-	Del	Pos	6.0	-	1	-	-	-
INS	c.130G>A	p.Gly44Arg	Mis	-	Del	Pos	3.3	-	1	-	-	-
INS	c.137G>A	p.Arg46Gln	Mis	Yes	Del	Pos	2.5	-	1	-	-	-
INS	c.140G>A	p.Gly47Asp	Mis	-	Del	Prob	3.5	-	1	-	-	-
NEUROD1	c.133A>G	p.Thr45Ala	Mis	-	Tol	Benign	0.6	Yes	27	36.%	26.%	39.%
NEUROD1	c.723C>G	p.His241Gln	Mis	Yes	Tol	Pos	1.7	-	1	-	-	-
PDX1	c.559G>A	p.Glu187Lys	Mis	-	Del	Prob	5.9	-	1	-	-	-
PDX1	c.716C>A	p.Leu239Gln	Mis	-	Tol	Benign	0.0	Yes	2	-	-	-

Supplementary Table 4: **List of variants identified in MODY patients.** Shown is the list of variants observed in the studied MODY patients; the columns are identical to those in Figure 2. The number of MODY patients (out of 250) observed to carry each variant is shown, as well as (for reference) the frequency of each variant in the unselected cohorts.

Class	MODY patients	FHS cohort	JHS cohort	Extreme T2D cohort			
				DM	No DM	OR (95% CI)	P
Low frequency non-synonymous	120 (48%)	68 (4.4%)	97 (5.7%)	17 (4.7%)	6 (1.5%)	3.2 (1.2- 8.2)	**0.011**
Possibly Pathogenic	57 (23%)	8 (0.52%)	8 (0.47%)	4 (1.1%)	0 (0.%)	-	**0.04**
HGMD MODY	69 (28%)	27 (1.8%)	25 (1.5%)	4 (1.1%)	1 (0.24%)	4.7 (0.5-42.0)	0.2
Putative Pathogenic	48 (19%)	3 (0.19%)	1 (0.06%)	2 (0.55%)	0 (0.%)	-	0.25

Supplementary Table 5: **Carrier frequencies.** Shown is the fraction of individuals in each cohort, as well as MODY patients, who carry variants from each of the four variant classes. For the extreme T2D cohort, frequencies are shown separately for individuals with diabetes (DM) and individuals without diabetes (No DM). Based on these frequencies, we computed the odds ratio of carrier to non-carrier between extreme cases and controls, as well as the p-value from a one-sided test of association; these estimates control for potential ethnic differences between individuals from Finland and Sweden (Methods). The table shows 95% confidence intervals for odds ratios and shows p-values that exceed nominal ($p<0.05$) significance in bold. Odds ratios for the putative pathogenic or possibly pathogenic classes were not computed, as no controls carried variants.

Phenotype	Class	FHS cohort				JHS cohort				Combined
		Indiv.	w/ pheno (%)	OR (95% CI)	P	Indiv.	w/ pheno (%)	OR (95% CI)	P	P
Diabetes	Non-carrier	1473	101 (6.9%)	-	-	1594	327 (20.5%)	-	-	-
	Nonsyn $<$ 1%	68	5 (7.3%)	1.1 (0.4- 2.7)	0.51	97	19 (19.6%)	0.9 (0.6-1.6)	0.61	0.53
	Pos. Pathogenic	8	1 (12.5%)	1.9 (0.2-15.9)	0.42	8	1 (12.5%)	0.6 (0.1-4.5)	0.84	0.71
	HGMD MODY	27	2 (7.4%)	1.1 (0.3- 4.7)	0.53	25	6 (24.%)	1.2 (0.5-3.1)	0.44	0.44
	Put. Pathogenic	3	1 (33.3%)	6.8 (0.6-75.5)	0.22	1	1 (100.%)	-	0.21	**0.049**
IFG or diabetes	Non-carrier	1401	487 (34.8%)	-	-	1547	541 (35.%)	-	-	-
	Nonsyn $<$ 1%	65	29 (44.6%)	1.5 (0.9- 2.5)	0.058	93	28 (30.1%)	0.8 (0.5-1.3)	0.86	0.41
	Pos. Pathogenic	8	4 (50.%)	1.9 (0.5- 7.5)	0.29	8	1 (12.5%)	0.3 (0.0-2.2)	0.97	0.7
	HGMD MODY	25	8 (32.%)	0.9 (0.4- 2.1)	0.7	25	8 (32.%)	0.9 (0.4-2.0)	0.7	0.74
	Put. Pathogenic	3	2 (66.7%)	3.8 (0.3-41.5)	0.29	1	1 (100.%)	-	0.37	0.12

Supplementary Table 6: **Phenotypic impact of variants in unselected cohorts** Shown is the fraction of variant carriers in the two unselected cohorts with diabetes as well as the fraction with IFG or diabetes. Figure 2 shows data in graphical form.

Gene	Nucleotide change	Amino acid change	Pos. Pathogenic	HGMD MODY	Put. Pathogenic	DM Count	No DM Count
GCK	c.45+1G>T	NA	Yes	Yes	Yes	1	0
GCK	c.106C>T	p.Arg36Trp	No	No	No	1	0
GCK	c.493C>A	p.Leu165Ile	No	No	No	1	0
GCK	c.649G>A	p.Asp217Asn	No	No	No	0	1
HNF1A	c.392G>A	p.Arg131Gln	Yes	Yes	Yes	1	0
HNF1A	c.586A>G	p.Thr196Ala	No	Yes	No	0	1
HNF1A	c.341G>A	p.Arg114His	No	No	No	0	1
HNF1B	c.905A>G	p.Asn302Ser	No	No	No	1	0
HNF1B	c.1621A>C	p.Ser541Arg	No	No	No	1	0
HNF4A	c.488G>A	p.Arg163Gln	No	Yes	No	1	0
HNF4A	c.1333C>T	p.Pro445Ser	No	Yes	No	1	0
HNF4A	c.365G>T	p.Arg122Leu	Yes	No	No	1	0
HNF4A	c.655C>A	p.Leu219Met	No	No	No	1	0
HNF4A	c.1210G>A	p.Val404Ile	No	No	No	1	0
HNF4A	c.1360G>A	p.Ala454Thr	No	No	No	0	1
HNF4A	c.1387A>G	p.Ile463Val	No	No	No	0	1
INS	c.17G>A	p.Arg6His	No	No	No	1	1
INS	c.224G>A	p.Gly75Asp	No	No	No	1	0
NEUROD1	c.117G>T	p.Lys39Asn	No	No	No	1	0
NEUROD1	c.454delG	p.Ala152LeufsX110	Yes	No	No	1	0
PDX1	c.52T>C	p.Cys18Arg	No	No	No	2	0
PDX1	c.54C>A	p.Cys18X	No	No	No	1	0

Supplementary Table 7: **Variants observed in T2D cohort.** Shown is the list of variants observed in individuals within the T2D cohort. Indicated are the classes to which each variant belongs, as well as the number of observations in individuals with and without type 2 diabetes. One individual carried two variants (c.224G>A in *INS* and c.1621A>C in *HNF1B*); thus the total number of counts is one greater than presented Supplementary Table 5.

Gene	Class	Diabetes								
		FHS cohort				JHS cohort				Combined
		Indiv.	w/ pheno (%)	OR (95% CI)	P	Indiv.	w/ pheno (%)	OR (95% CI)	P	P
-	Non-carrier	1473	101 (6.9%)	-	-	1594	327 (20.5%)	-	-	-
GCK	Nonsyn < 1%	5	0 (0.%)	-	1.	8	2 (25.%)	1.3 (0.3- 6.5)	0.44	0.46
GCK	Pos. Pathogenic	1	0 (0.%)	-	1.	1	0 (0.%)	-	1.	1.
GCK	HGMD MODY	1	0 (0.%)	-	1.	1	0 (0.%)	-	1.	1.
GCK	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
HNF1A	Nonsyn < 1%	24	2 (8.3%)	1.2 (0.3- 5.3)	0.51	26	8 (30.8%)	1.7 (0.8- 4.0)	0.18	0.18
HNF1A	Pos. Pathogenic	3	1 (33.3%)	6.8 (0.6- 75.5)	0.19	2	1 (50.%)	3.9 (0.2-62.9)	0.38	0.11
HNF1A	HGMD MODY	14	1 (7.1%)	1.0 (0.1- 8.1)	0.65	21	6 (28.6%)	1.6 (0.6- 4.1)	0.28	0.26
HNF1A	Put. Pathogenic	2	1 (50.%)	13.6 (0.8-218.8)	0.15	1	1 (100.%)	-	0.2	**0.031**
HNF1B	Nonsyn < 1%	15	1 (6.7%)	1.0 (0.1- 7.4)	0.62	14	3 (21.4%)	1.1 (0.3- 3.8)	0.58	0.58
HNF1B	Pos. Pathogenic	2	0 (0.%)	-	1.	2	0 (0.%)	-	1.	1.
HNF1B	HGMD MODY	8	1 (12.5%)	1.9 (0.2- 15.9)	0.42	4	1 (25.%)	1.3 (0.1-12.5)	0.6	0.4
HNF1B	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
HNF4A	Nonsyn < 1%	9	1 (11.1%)	1.7 (0.2- 13.7)	0.5	9	1 (11.1%)	0.5 (0.1- 3.9)	0.89	0.79
HNF4A	Pos. Pathogenic	1	0 (0.%)	-	1.	1	0 (0.%)	-	1.	1.
HNF4A	HGMD MODY	4	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
HNF4A	Put. Pathogenic	1	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
INS	Nonsyn < 1%	4	0 (0.%)	-	1.	15	0 (0.%)	-	1.	1.
INS	Pos. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
INS	HGMD MODY	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
INS	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
NEUROD1	Nonsyn < 1%	3	1 (33.3%)	6.8 (0.6- 75.9)	0.12	15	2 (13.3%)	0.6 (0.1- 2.7)	0.82	0.63
NEUROD1	Pos. Pathogenic	0	0 (0.%)	-	1.	2	0 (0.%)	-	1.	1.
NEUROD1	HGMD MODY	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
NEUROD1	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
PDX1	Nonsyn < 1%	10	1 (10.%)	1.5 (0.2- 12.0)	0.54	12	4 (33.3%)	2.0 (0.6- 6.5)	0.23	0.21
PDX1	Pos. Pathogenic	1	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
PDX1	HGMD MODY	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
PDX1	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.

Supplementary Table 8: **Impact of variants on diabetes risk in unselected cohorts.** In addition to investigating phenotypic effects across all genes collectively, we estimated variant odds ratios (OR) and association statistics for each gene individually. Shown are columns analogous to those in Figure 2 that measure effect on diabetes, but with separate rows for each gene. Effects on impaired fasting glucose (IFG) or diabetes are shown in Supplementary Table 9. As some individuals carried multiple variants, the sum of carrier numbers across genes is greater than that in Figure 2.

Gene	Class	IFG or diabetes								
		FHS cohort				JHS cohort				Combined
		Indiv.	w/ pheno (%)	OR (95% CI)	P	Indiv.	w/ pheno (%)	OR (95% CI)	P	P
-	Non-carrier	1401	487 (34.8%)	-	-	1547	541 (35.%)	-	-	-
GCK	Nonsyn < 1%	5	3 (60.%)	2.8 (0.5-16.6)	0.24	7	4 (57.1%)	2.5 (0.6-11.3)	0.19	0.092
GCK	Pos. Pathogenic	1	1 (100.%)	-	0.36	1	0 (0.%)	-	1.	0.33
GCK	HGMD MODY	1	0 (0.%)	-	1.	1	1 (100.%)	-	0.34	0.59
GCK	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
HNF1A	Nonsyn < 1%	23	9 (39.1%)	1.2 (0.5- 2.8)	0.42	26	8 (30.8%)	0.8 (0.4- 1.9)	0.74	0.57
HNF1A	Pos. Pathogenic	3	2 (66.7%)	3.7 (0.3-40.8)	0.29	2	1 (50.%)	1.9 (0.1-30.1)	0.57	0.24
HNF1A	HGMD MODY	13	4 (30.8%)	0.8 (0.3- 2.7)	0.72	21	6 (28.6%)	0.8 (0.3- 1.9)	0.82	0.79
HNF1A	Put. Pathogenic	2	1 (50.%)	1.8 (0.1-29.6)	0.57	1	1 (100.%)	-	0.33	0.28
HNF1B	Nonsyn < 1%	14	5 (35.7%)	1.0 (0.3- 3.1)	0.59	14	4 (28.6%)	0.8 (0.2- 2.4)	0.75	0.65
HNF1B	Pos. Pathogenic	2	0 (0.%)	-	1.	2	0 (0.%)	-	1.	1.
HNF1B	HGMD MODY	7	2 (28.6%)	0.7 (0.1- 3.8)	0.79	4	2 (50.%)	1.9 (0.3-13.4)	0.41	0.59
HNF1B	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
HNF4A	Nonsyn < 1%	9	3 (33.3%)	0.9 (0.2- 3.7)	0.66	9	2 (22.2%)	0.5 (0.1- 2.6)	0.9	0.8
HNF4A	Pos. Pathogenic	1	1 (100.%)	-	0.37	1	0 (0.%)	-	1.	0.55
HNF4A	HGMD MODY	4	2 (50.%)	1.8 (0.3-13.1)	0.42	0	0 (0.%)	-	1.	0.44
HNF4A	Put. Pathogenic	1	1 (100.%)	-	0.34	0	0 (0.%)	-	1.	0.37
INS	Nonsyn < 1%	4	3 (75.%)	5.5 (0.6-53.5)	0.13	13	1 (7.7%)	0.2 (0.0- 1.2)	1.	0.91
INS	Pos. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
INS	HGMD MODY	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
INS	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
NEUROD1	Nonsyn < 1%	2	1 (50.%)	1.8 (0.1-29.5)	0.6	14	5 (35.7%)	1.0 (0.3- 3.1)	0.55	0.49
NEUROD1	Pos. Pathogenic	0	0 (0.%)	-	1.	2	0 (0.%)	-	1.	1.
NEUROD1	HGMD MODY	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
NEUROD1	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
PDX1	Nonsyn < 1%	10	6 (60.%)	2.8 (0.8- 9.9)	0.1	12	5 (41.7%)	1.3 (0.4- 4.3)	0.43	0.11
PDX1	Pos. Pathogenic	1	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
PDX1	HGMD MODY	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
PDX1	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.

Supplementary Table 9: **Impact of variants on impaired fasting glucose risk in unselected cohorts.** In addition to investigating phenotypic effects across all genes collectively, we estimated variant odds ratios (OR) and association statistics for each gene individually. Shown are columns analogous to those in Figure 2 that measure effect on impaired fasting glucose (IFG), but with separate rows for each gene. Effects on diabetes are shown in Supplementary Table 8. The number of analyzed individuals is smaller for IFG than T2D because not all studied individuals had glucose measurements available.

Phenotype	Class	FHS cohort				JHS cohort				Combined
		Indiv.	w/ pheno (%)	OR (95% CI)	P	Indiv.	w/ pheno (%)	OR (95% CI)	P	P
Diabetes	Non-carrier	1536	106 (6.9%)	-	-	1683	344 (20.4%)	-	-	-
	Nonsyn < 1%	5	0 (0.%)	-	1.	8	2 (25.%)	1.3 (0.3- 6.5)	0.44	0.48
	Pos. Pathogenic	1	0 (0.%)	-	1.	1	0 (0.%)	-	1.	1.
	HGMD MODY	1	0 (0.%)	-	1.	1	0 (0.%)	-	1.	1.
	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.
IFG, diabetes, or FPG > 99 mg/dL	Non-carrier	1461	513 (35.1%)	-	-	1633	599 (36.7%)	-	-	-
	Nonsyn < 1%	5	3 (60.%)	2.8 (0.5-16.6)	0.27	7	5 (71.4%)	4.3 (0.8-22.3)	0.11	0.054
	Pos. Pathogenic	1	1 (100.%)	-	0.36	1	0 (0.%)	-	1.	0.6
	HGMD MODY	1	0 (0.%)	-	1.	1	1 (100.%)	-	0.34	0.65
	Put. Pathogenic	0	0 (0.%)	-	1.	0	0 (0.%)	-	1.	1.

Supplementary Table 10: **Phenotypes of *GCK*-variant carriers.** GCK-MODY patients have a distinctly mild elevation in fasting plasma glucose levels (99-144 mg/dL). The table contains tabulated data from the graph in Figure 3a. Two additional variants are included in the table, as carriers of these variants have diabetes and are thus not plotted in Figure 3a.