

SUPPLEMENTARY DATA

Functional Investigations of *HNF1A* Identify Rare Variants as Risk Factors for Type 2 Diabetes in the General Population

Laeya Abdoli Najmi,^{1,2,3} Ingvild Aukrust,^{1,2} Jason Flannick,⁴ Janne Molnes,^{1,5} Noel Burrtt,⁴ Anders Molven,^{1,6,7} Leif Groop,⁸ David Altshuler,^{4,9,10} * Stefan Johansson,^{1,2} Lise Bjørkhaug,^{1,11} † and Pål Rasmus Njølstad^{1,5} †‡

Supplementary Figure S1. Analysis of HNF-1A protein variants on DNA binding by electrophoretic mobility shift analysis (EMSA).

Supplementary Figure S2. Measurement of HNF-1A protein level expression.

Supplementary Figure S3. Total sample sizes needed to obtain 80% power.

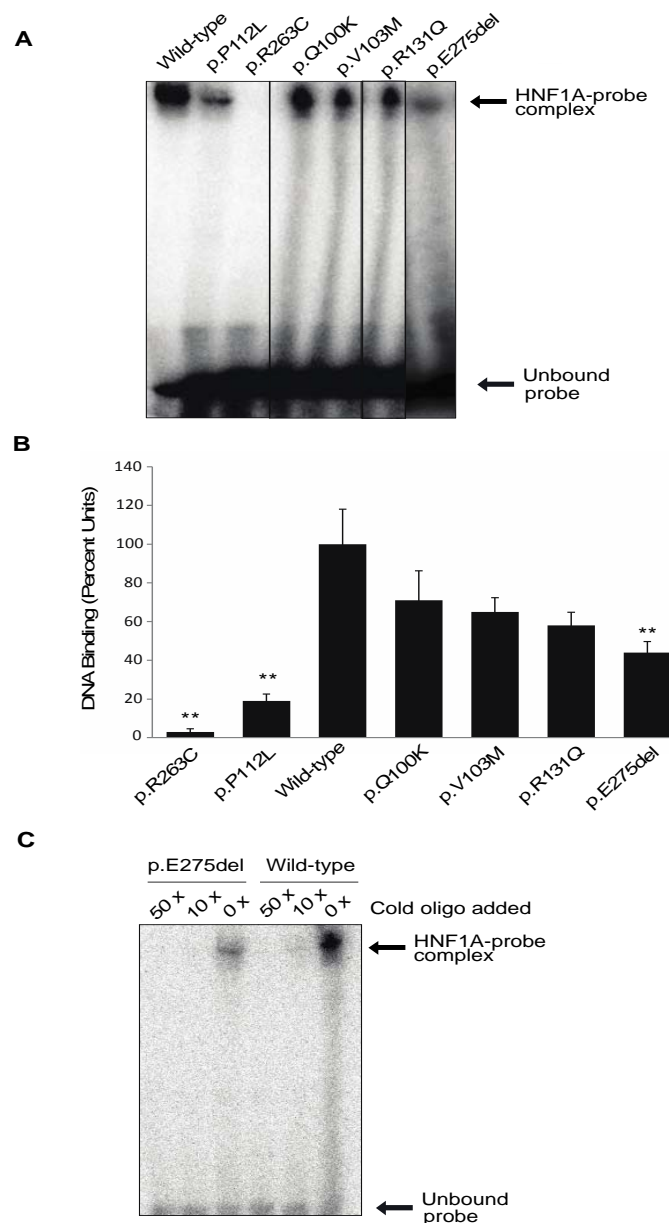
Supplementary Figure S4. Effect sizes (OR) needed to obtain 80% power.

Supplementary Table S1. Allele frequencies of *HNF1A* variants found in the type 2 diabetes genetics database.

SUPPLEMENTARY DATA

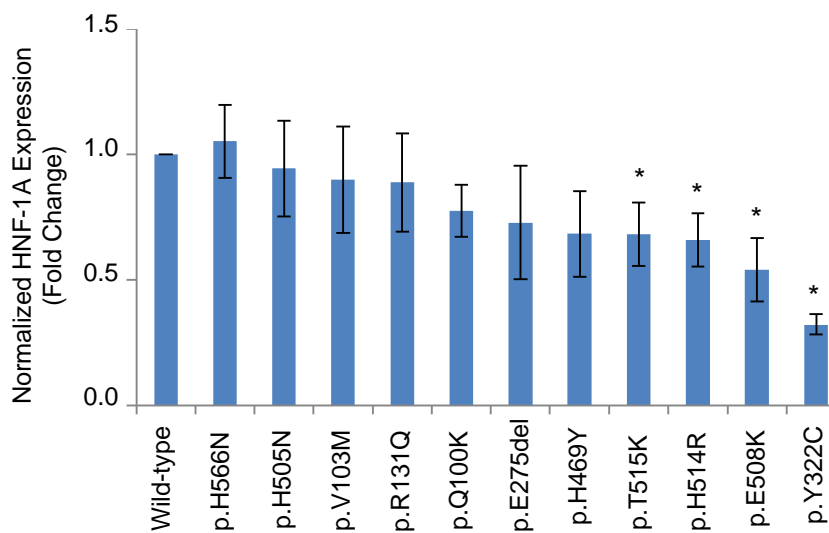
Supplementary Figure S1. Analysis of HNF-1A protein variants on DNA binding by electrophoretic mobility shift analysis (EMSA).

A: Xpress-epitope tagged wild-type and HNF-1A variant proteins were synthesized in vitro in a coupled transcription/translation system and incubated with a radio-labeled DNA fragment containing the HNF-1A-binding site in the rat albumin promoter. Bound complexes were analyzed by EMSA. Two variants with severely reduced DNA binding, known to cause MODY3 (p.P112L and p.R263C), were included for comparison. B: The quantification of bound signals. Each point represents the mean of three biological replicates (based on two separate protein expression reactions, three separate DNA binding reactions and gel analyses)(n = 3). C: Competition assays performed by adding increasing amounts of non-labeled DNA fragment to the binding reaction.



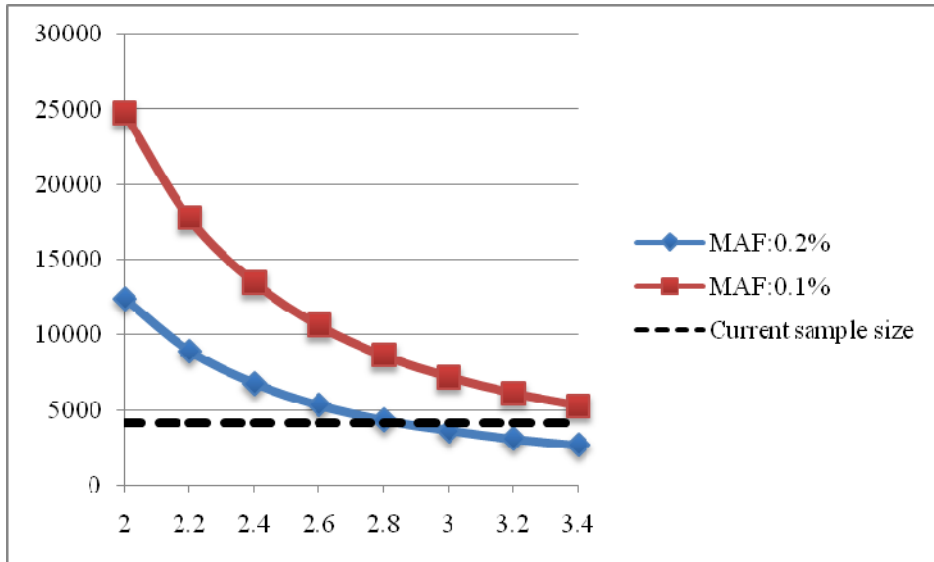
SUPPLEMENTARY DATA

Supplementary Figure S2. Measurement of HNF-1A protein level expression. Protein level expression measured in cell lysate from transfected HeLa cells. Samples were analyzed by SDS-PAGE and immunoblotting, and HNF-1A protein quantitated by densitometric analysis. Each bar represent of mean of 3 readings \pm SD. *P* values were calculated in comparison to wild-type protein expression using paired t test. **P* < 0.05.



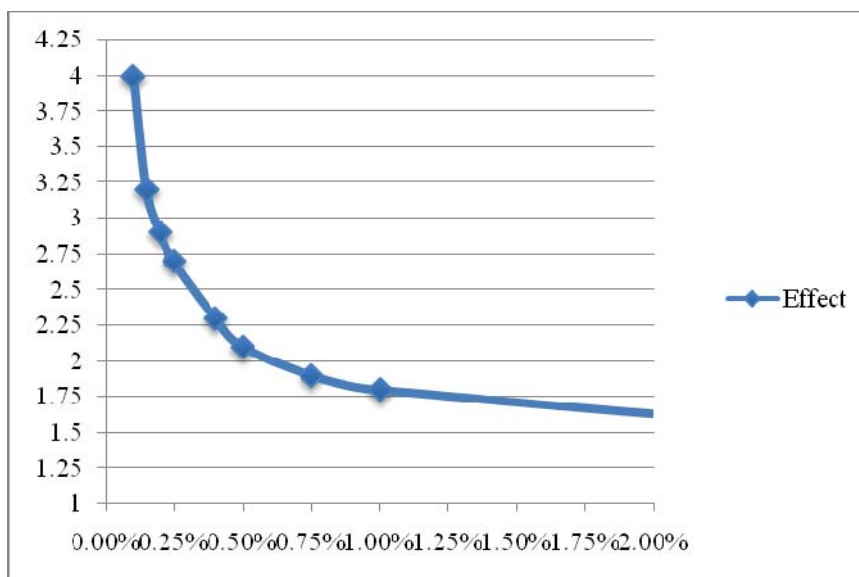
SUPPLEMENTARY DATA

Supplementary Figure S3. Total sample sizes needed to obtain 80% power. Graph shows sample size at $\alpha=0.05$ for various ORs and minor allele frequencies (MAF) 0.2% (blue) and 0.1% (red), given that 22.5% of the total sample have diabetes (as in the present study). Sample size in the current study is indicated with the black dotted line ($n=926$ cases and $n=3189$ controls).



SUPPLEMENTARY DATA

Supplementary Figure S4. Effect sizes (OR) needed to obtain 80% power. Graph shows OR at $\alpha=0.05$ for various minor allele frequencies (MAF), given the current sample size ($n=926$ cases and $n=3189$ controls).



SUPPLEMENTARY DATA

Supplementary Table S1. Allele frequencies of *HNF1A* variants found in the type 2 diabetes genetics database.

	Amino Acid Change	Functional evaluation (this study)			T2D genetics database*	
		Transcriptional activity (%)	DNA Binding (%)	Nuclear Localization (%)	DM	No DM
					(n=8379) (n/MAF)	(n=8478) (n/MAF)
c.965A>G	Y322C	30	-	-	6 (0.036%)	4 (0.024%)
c.818_820del	E275del	30	44	-	-	-
c.392G>A	R131Q	35	58	-	-	-
c.1522G>A	E508K	40	-	15	44 (0.23%)	8 (0.047%)
c.1405C>T	H469Y	41	-	63	1 (0.006%)	0
c.1544C>A	T515K	46	-	68	2 (0.012%)	4 (0.024%)
c.1541A>G	H514R	51	-	59	14 (0.084%)	6 (0.035%)
c.1513C>A	H505N	54	-	80	0	1 (0.006%)
c. 298C>A	Q100K	57	71	59	1 (0.006%)	0
c.307G>A	V103M	59	65	74	-	-
c.1696C>A	H566N	59	-	57	-	-
c.142G>A	E48K	61	-	86	1 (0.006%)	1 (0.006%)
c.1360A>G	S454G	64	-	64	-	-
c.1469T>C	M490T	63	-	78	-	-
c.1748G>A	R583Q	63	-	73	2 (0.012%)	5 (0.029%)
c.92G>A	G31D	65	-	75	7 (0.042%)	8 (0.047%)
c.854C>T	T285M	64	-	78	-	-
c.185A>G	N62S	65	-	73	2 (0.012%)	1 (0.006%)
c.1532A>G	Q511R	66	-	78	-	-
c.1729C>G	H577D	69	-	81	4 (0.024%)	1 (0.006%)
c.533C>T	T178I	67	-	87	-	-
c.827C>G	A276G	69	-	70	-	-
c.1165T>G	L389V	68	-	73	17 (0.10%)	16 (0.09%)
c.290C>T	A97V	71	-	74	1 (0.006%)	1 (0.006%)

SUPPLEMENTARY DATA

c.824A>C	E275A	78	-	79	-	-
c.341G>A	R114H	83	-	70	0	2 (0.012%)
c.586A>G	T196A	101	-	79	1 (0.006%)	4 (0.024%)

The variants are ranged based on their functional effect as described in Table 1 (starting with the most impaired variant). *Web site address: <http://www.type2diabetesgenetics.org>. Data obtained from the website September 2016.

Abbreviations: MAF = Minor Allele Frequency.