Supplementary Tables - *ONECUT1* mutations and variants in diabetes

Supplementary Tables

Supplementary Table 1. Glucose metabolism in *ONECUT1* heterozygous parents and grand-parents

	Family 1				Family 2		Reference values
	Mother (subject 2)	Father (subject 1)	Grand-father (subject 22)	Grand-mother (subject 25)	Mother (subject 2)	Father (subject 1)	
Age at study (years)	40	38	76	63	23	29	
BMI (kg/m^2)	28.1	28.3	NA	NA	24.5	26.4	
HbA1c (%)	6.6[a]	5.4	6.9[a]	6.0[b]	6.1[b]	6.2[b]	<5.7
Autoantibodies (GAD, IA2, INS)	Negative	Negative	ND	ND	Negative	Negative	
Fasting glucose (mmol/l)	6.3[b]	4.7	8.3[a]	6.6[b]	5.8[b]	5.3	<5.6
Glucose 30 min (mmol/l)	13.9[a]	6.6	ND	ND	ND	ND	<9.7
Glucose 60 min (mmol/l)	19.0[a]	7.4	ND	ND	9.9	8.7	<10.1
Glucose 120 min (mmol/l)	15.8[a]	5.8	ND	ND	9.1[b]	9.2[b]	<7.8
Fasting insulin (pmol/l)	62.5	34.7[b]	74.3	18.7[b]	23.6[b]	31.9[b]	35-174
Fasting C-peptide (pmol/l)	733	766	1490	728	ND	ND	160–1100
Fasting Insulin/C-peptide	0.09	0.05	0.05	0.03	ND	ND	<0.10
Insulin 30 min (mIU/l)	28	45	ND	ND	ND	ND	27-60
Insulin 60 min (mIU/l)	45	79	ND	ND	ND	ND	45-80
Insulin 120 min (mIU/l)	37	55	ND	ND	ND	ND	35-65
ΔI30/ΔG30 (mIU/mmol)	3.7[a]	21.1	ND	ND	ND	ND	16.9 ± 10.7 (NDC); 3.9 ± 2.3 (T2D)

Glucose and insulin values shown are fasting, 30 min, 60 min, and 120 min after the start of oral glucose tolerance test (OGTT). Early insulin secretion (ΔI30/ΔG30) was calculated as the ratio [(insulinemia at 30 min)-(fasting insulinemia)]/[(blood glucose at 30 min)–(fasting blood glucose)]. Reference values shown are American Diabetes Association (ADA) reference threshold as well as laboratory reference values during fasting and OGTT (insulin and C-peptide) and in non-diabetic control (NDC) and Type 2 diabetes (T2D) groups for ΔI30/ΔG30.

[a] Diabetes range; [b] Impaired fasting glucose (IFG)/Impaired glucose tolerance (IGT) range. NA: not available. ND: not done.

Supplementary Table 2. Characteristics of T2D patients from the Ulm Diabetes Cohort

	N	Mean	SD	Min	10^{th} centile	25^{th} centile	Median	75^{th} centile	90^{th} centile	Max
Age at diagnosis	2143	44.68	14.44	10	25	35	45	55	63	85
BMI at recruitment	1143	28.97	7.02	15.6	21.6	24.2	27.9	32.1	37.7	66.7
BMI at diagnosis	660	27.66	6.48	15.6	20.8	23.2	26.4	30.7	35.9	62.8

SD: standard deviation; Min: minimum; Max: maximum.

Supplementary Table 3. Ulm Diabetes Cohort (UDC) screening: *ONECUT1* exonic variant counts and allele frequencies

a. Rare variants

Genomic position (hg19)	Nucleotide change	Amino acid change	T2D patients (N=2165)	Population controls (N=397)	T1D/LADA patients (N=162)	MAF, all (gnomAD)	Max MAF (gnomAD)	MAF, NWE (gnomAD)
53082075	C/G	**A3P**	1	0	0	6.0E-06	4.2E-05	0
53081994	C/T	**G30S**	1	0	0	0.0002	0.0037	0
53081983	G/T	**H33Q**	3	0	0	0.0011	0.0022	0.0020
53081898	C/A	**G62C**	1	0	0	0	0	0
53081840	C/T	**G81D**	1	0	0	0.0003	0.0007	0.0002
53081836	G/A	P82P	1	0	0	0	0	0
53081796	C/A	**G96C**	1	0	0	0	0	0
53081795	C/T	**G96D**	1	0	0	0.0009	0.024	0.0010
53081745	G/C	**P113A**	1	0	0	0	0	0
53081677	C/T	P135P	10	3	0	0.0018	0.0034	0.0030
53081539	G/A	P181P	1	0	0	0	0	0
53081439	G/C	**P215A**	1	0	0	5.2E-05	6.5E-05	0
53081277	G/C	**R269G**	2	0	0	0.0006	0.0011	0.0012
53081221	C/A	G287G	16	1	0	0.021	0.12	0.0044
53081152	G/A	I310I	1	0	0	9.9E-05	0.0011	0
53050010	A/G	D380D	1	0	0	0	0	0
53049812	G/A	D446D	3	0	0	0.0002	0.0003	0
Rare nonsynonymous variants			13	0	0			
Rare synonymous variants			33	4	0			

Counts of heterozygous carriers of rare variants. MAF: minor allele frequency based on the genome aggregate database (gnomAD); Max MAF: maximum MAF across gnomaAD populations; LADA: latent autoimmune diabetes in adult. NWE: North-Western Europeans. Missense variants are shown in bold.

b. Common and low frequency missense variants

Genomic position (hg19)	Nucleotide change	Amino acid change	T2D patients (N=2165)	Population controls (N=397)	T1D/LADA patients (N=162)	MAF, all (gnomAD)	Max MAF (gnomAD)	MAF, NWE (gnomAD)
53081859	G/C	P75A	139 (6.42%)	16 (4.03%)	14 (8.64%)	0.023	0.054	0.032

Counts are the sum of heterozygous and homozygous subjects for the minor allele, excluding subjects carrying rare missense variants.

Supplementary Table 4. *ONECUT1* variants characteristics

a. Rare homozygous variants identified in two neonatal or young onset diabetic patients

						EVS, MAF				gnomAD, MAF				gnomAD, NWE		Deleteriousness prediction programs***						
Chr	Position hg19	dbSNP	Ref	Alt	AA change	EA	AA	All	Nb homo-zygous/N	All	Max MAF	Population	Nb homo-zygous/N (all)	MAF	Nb homo-zygous/N	SIFT	Polyphen2	LRT	MutTaster	FATHMM-MKL	CADD_phred	CONDEL
15	53081391		C	A	E231X	0	0	0	Absent	0	0	-	Absent	0	Absent	Protein-truncating variant						
15	53081389	.	C	G	E231D	0	0	0	Absent	0	0	-	Absent	0	Absent	T	B	D	D	D	N	N

b. Rare heterozygous missense variants identified by screening T2D patients (GnomAD-NWE MAF<0.005)

						EVS, MAF				gnomAD, MAF				gnomAD, NWE		Deleteriousness prediction programs***						
Chr	Position hg19	dbSNP	Ref	Alt	AA change	EA	AA	All	Nb homo-zygous/N	All	Max MAF	Popu-lation	Nb homo-zygous/N (all)	MAF	Nb homo-zygous/N	SIFT	Poly phen 2	LRT	MutT aster	FATHM M-MKL	CADD_p hred	CON DEL
15	53082075	-	C	G	A3P	0	0	0	Absent	6.07E-06	4.22E-05	SAS	0/ 82405	0	0 / 8897	T	B	D	D	D	20.2 (D)	N
15	53081994	rs367600429	C	T	G30S	0	0.0019	0.0006	0 / 5131	4.17E-04	0.0037	AFR	0 / 67127	0	0 / 9525	T	B	D	N	D	17.03 (D)	N
15	53081983	rs202151356	G	T	H33Q	0.0003	0.0018	0.0008	0 / 5164	1.04E-03	0.0022	ASJ	1 / 51947**	3.77E-04	0 / 7952	D	P	D	D	D	23.5 (D)	D
15	53081898	rs201056913	C	A	G62C	0	0	0	Absent	1.60E-05	1.90E-04	EAS	0 / 62373	0	0 / 8002	D	D	D	D	D	26.3 (D)	D
15	53081840	rs142641519	C	T	G81D	0.0003	0	0.0002	0 / 6485	2.09E-04	6.77E-04	AMR	0 / 112424	5.34E-05	0 / 18712	D	D	D	D	D	24.4 (D)	D
15	53081796	-	C	A	G96C	0	0	0	Absent	0	0	0	Absent	0	Absent	D	D	D	D	D	29.6 (D)	D
15	53081795	rs145201345	C	T	G96D	0.0014	0	0.0009	0 / 6487	8.29E-04	0.024	ASJ	6 / 123093*	2.15E-04	0 / 20955	D	D	D	D	D	27.7 (D)	D
15	53081745	-	G	C	P113A	0	0	0	Absent	0	0	0	Absent	0	Absent	T	B	D	D	D	22.9 (D)	N
15	53081439	rs1477846694	G	C	P215A	0	0	0	Absent	3.19E-05	6.49E-05	NFE	0 / 15683	1.16E-04	0 / 4296	T	P	D	D	D	19.07 (D)	N
15	53081277	rs139802359	G	C	R269G	0.0016	0.0002	0.0012	0 / 6487	5.96E-04	0.0011	FIN	0 / 136850	8.12E-04	0 / 25262	D	D	D	D	D	25 (D)	D

c. Rare heterozygous synonymous variants identified by screening T2D patients (GnomAD-NWE MAF<0.005)

						EVS, MAF				gnomAD, MAF				gnomAD, NWE	
Chr	Position hg19	dbSNP	Ref	Alt	AA change	EA	AA	ALL	Nb homozygous/N	All	Max MAF	Population	Nb homozygous/N	MAF	Nb homozygous/N
15	53081836	-	G	A	P82P	0	0	0	Absent	0	0	-	Absent	0	Absent
15	53081677	rs146677141	C	T	P135P	0.0031	0.0007	0.0023	1 / 6486	1.75E-03	0.0034	NFE	0 / 139966	0.003	0 / 24972
15	53081539	-	G	A	P181P	0	0	0	Absent	0	0	-	Absent	0	Absent
15	53081221	rs2075613	C	A	G287G	0.0038	0.003	0.0035	0 / 6487	0.021	0.12	EAS	269 / 138897	0.0044	0 / 25353
15	53081152	rs145547859	G	A	I310I	1.16E-04	0	7.70E-05	0 / 6487	9.91E-05	0.0011	ASJ	0 / 141249	0	0 / 25401
15	53050010	-	A	G	D380D	0	0	0	Absent	0	0	-	Absent	0	Absent
15	53049812	rs141094636	G	A	D446D	0	0	0	Absent	1.95E-04	3.27E-04	NFE	Absent	0	Absent

d. Rare heterozygous missense variants identified as a mutation occurring de novo in a juvenile-onset non-autoimmune diabetic patient

						EVS, MAF				gnomAD, MAF				gnomAD, NWE		Deleteriousness prediction programs***						
Chr	Position hg19	dbSNP	Ref	Alt	AA change	EA	AA	All	Nb homo-zygous/N	All	Max MAF	Population	Nb homo-zygous/N (all)	MAF	Nb homo-zygous/N	SIFT	Polyphen2	LRT	MutTast	FATHMM-MKL	CADD_phred	CONDEL
15	53081438	-	G	C	P215R	0	0	0	Absent	0	0	-	Absent	0	Absent	D	D	D	D	D	24.9(D)	D

* 3 of these homozygous subjects were labeled as "controls" in case/control studies from gnomAD.

** This homozygous subject was labeled as a "control" in a case/control study from gnomAD.

***SIFT, POLYPHEN2(HDIV), LRT and MutationTaster (MutTast) are function prediction tool, and FATHMM, CADD and CONDEL are consensus methods that integrate several prediction tools. CADD_phred scores >15 may be considered deleterious (CADD information).

Ref: reference allele; Alt: alternative allele; AA; amino acid; MAF: minor allele frequency; N: number of subjects sequenced; Nb: number; D: damaging, P: possibly damaging, T: tolerated, N: neutral, B: benign.
Databases: EVS: Exon Variant Server; gnomAD: genome aggregation database. Populations: EA: European American, AA: African American, NFE: Non-Finnish European, FIN: Finnish, ASJ: Ashkenazi Jewish, AMR: Latino/Admixed American, AFR: African/African American, EAS: East Asian, SAS: South Asian, NWE: North-Western European subgroup from NFE population.

Supplementary Table 5. Burden testing for T2D, all *ONECUT1* coding variants

Population	Additive burden test (P-value, OR)	Collapsing burden test (P-value, OR)	SKAT (P-value)	SKAT-O (P-value)	Variable threshold burden test (P-value)
All	0.10, 1.13	0.08, 1.14	0.00026	0.00027	0.005
African-American	0.04, 1.49	0.04, 1.49			
East-Asian	0.17, 0.62	0.17, 0.62			
European	**0.0027, 1.30**	**0.002, 1.31**	NA	NA	NA
Hispanic	0.093, 0.75	0.12, 0.77			
South-Asian	0.063, 0.61	0.063, 0.61			

All coding variants affecting *ONECUT1* RefSeq were selected (N=86 variants). OR: odds ratios.
Burden testing was performed using various programs available on the T2D Knowledge Portal web site (http://www.type2diabetesgenetics.org/; date: 10/2020): Additive burden test (counts variants to generate a significance score that is proportional to the number of variants); Collapsing burden test (test for association between the total number of rare alleles observed per individual and a trait); SKAT (variance component test); SKAT-O (robust test that combines burden and SKAT methods); and Variable threshold burden test (runs successive collapsing burden tests across a range of minor allele frequencies, selecting the optimal statistic adaptively). Additive, collapsing and variable threshold burden tests assume effects that are all predisposing. SKAT and SKAT-O allow variants with opposite effects (predisposing or protecting). SKAT and SKAT-O burden tests are more powerful when combining rare variants that both increase and decrease the risk. We used ancestry stratification for these analyses.
NA: not available.

Supplementary Table 6. Partial extract from Type 2 Diabetes Knowledge Portal association analysis of selected rare *ONECUT1* variants

a. AMP T2D-GENES

Genomic position (hg19)	Nucleotide change	Amino acid change	MAF ALL (gnomAD)	Max MAF (gnomAD)	MAF NWE (gnomAD)	Minor allele count (all)	Minor allele frequency (cases)	Association P-value	Odds ratio	Case minor allele counts (N=19852)	Control minor allele counts (N=23273)
Rare variants shared with the AMP T2D-GENES											
53081994	C/T	G30S	0.0002	0.0037	0	23	0.00044	0.16	9.8	18	5
53081983	G/T	H33Q	0.0011	0.0022	0.002	91	0.0020	0.079	5.0	50	41
53081898	C/A	G62C	0	0	0	3	0.000044	0.54	4.1	3	0
53081840	C/T	G81D	0.0003	0.0007	0.0002	35	0.00040	0.59	0.8	16	19
53081795	C/T	G96D	0.0009	0.0241	0.001	52	0.00056	0.011	0.3	16	36
53081277	G/C	R269G	0.0006	0.0011	0.0012	41	0.00042	0.47	1.3	20	21
Population-specific variant (East-Asian)											
53081357	A/G	V242A*	0.0011	0.0145	0	176	0.0051	0.026	1.4	101	75

These results were extracted from www.type2diabetesgenetics.org/variantSearch/ for ONECUT1 on 12/6/2019, for a population size of 49147 subjects.
*Data for the V242A variant were not available in this specific analysis and data shown were extracted from www.type2diabetesgenetics.org/variantInfo/ on 12/6/2019.

b. DIAMANTE

Amino acid change	P-value	Odds ratio	Effective sample size	Minor allele frequency (cases)
Rare variants shared with DIAMANTE				
G96D	0.40	0.81	107655	0.0018
R269G	0.55	1.06	203107	0.0014

Supplementary Table 7. Distributions of age at diabetes diagnosis and BMI in patients heterozygous for rare ONECUT1 coding variants

a. Age at diagnosis of diabetes

N	Median	Min	10^{th} centile	25^{th} centile	75^{th} centile	90^{th} centile	Max
17	29	12	18.4	23.5	37	44.6	47

b. BMI

N	Median	Min	10^{th} centile	25^{th} centile	75^{th} centile	90^{th} centile	Max
At recruitment							
16	26.5	20.2	22.1	24.2	30.1	36.4	37.3
At diagnosis							
12	26.5	20.2	21	24.6	29.5	30.1	30.1

Patients are *ONECUT1* heterozygous diabetic subjects from the UDC cohort, completed by 4 additional *ONECUT1* heterozygous subjects (three diabetic/IFG/IGT parents of families 1 and 2 and one Lebanese juvenile-onset diabetic patient). N: total number of subjects in each group and subgroup, Min: minimum, Max: maximum.

Supplementary Table 8. Burden testing in selected T2D patients adjusted to the characteristics of *ONECUT1* heterozygous patients identified in our study

Subjects, subgroups			*ONECUT1* heterozygous subjects		
	P-value	OR	%	n	N
Controls, unselected			0.81%	188	23273
Cases, age 12-35	0.41	1.77	0.70%	4	575
Cases, BMI 20-35	0.0018	1.53	1.30%	93	7164
Cases, fasting insulin 18-70	0.0078	2.32	1.36%	12	882
Cases, age 12-35 and BMI 20-35	0.0015	22.3	3.85%	3	78
Cases, age 12-35 and fasting insulin 18-70	0.0020	38.6	11.80%	2	17
Cases, age 12-35 and BMI 20-35 and fasting insulin 18-70	NR*	NR*	20%	2	10

Burden testing was preformed in selected subgroups of patients from the AMP T2D-GENES cohort, using the tool "custom aggregation tests" available online on the AMP-T2D web site (www.type2diabetesgenetics.org; date 10/2020). Patients were selected based on various criteria to best represent T2D patients heterozygous for *ONECUT1* coding variants identified in our screening, depending on age, BMI and fasting insulin. In the absence of age at onset available in this cohort, we used the age at recruitment as a surrogate. In the absence of BMI at onset available in this cohort, we used the BMI at recruitment from our patients as selection values (normal/non-obese range). Range for age is in years, fasting insulin in pmol/l.

The number (n) and percentage (%) of *ONECUT1* heterozygous carriers is shown, and the total number of subjects (N) in the selection criteria. P-values and odds ratios (OR) are in comparison to unselected controls, using all populations, stratified by ancestry. Variants included for analysis are all rare coding variants (MAF<0.01) affecting the *HNF6/ONECUT1* RefSeq canonical transcript (NM_004498.2, 85 variants).

*NR (not reported): not reliable test, due to small sample size.

Supplementary Table 9. Metabolic traits association in *ONECUT1* genomic region

	SNP	rs2456530		rs75332279		rs2440374		rs7178476	
	T2D risk allele frequency(EU, non-Finnish)	0.118		0.0877		0.253		0.260	
Traits	**Study**	**P-value**	**Dir**	**P-value**	**Dir**	**P-value**	**Dir**	**P-value**	**Dir**
T2D, BMI, obesity									
Type 2 diabetes	DIAMANTE (European) T2D GWAS	4.7E-09	↑	5.8E-09	↑	4.3E-08	↑	7.4E-08	↑
Type 2 diabetes adj BMI	DIAMANTE (European) T2D GWAS	1.1E-06	↑	1.4E-07	↑	1.3E-05	↑	2.3E-05	↑
BMI	GIANT UK Biobank GWAS	4.7E-08	↑	0.0036	↑	1.4E-08	↑	1.4E-09	↑
Waist-hip ratio	GIANT-UK Biobank GWAS Meta-analysis	2.7E-10	↑	1.8E-06	↑	1.8E-08	↑	1.6E-08	↑
Waist-hip ratio adj BMI	GIANT-UK Biobank GWAS Meta-analysis	6.9E-06	↑	0.0017	↑	0.00014	↑	0.00050	↑
Body fat percentage	GIANT UK Biobank GWAS	3.8E-10	↑	8.8E-07	↑	1.2E-11	↑	1.4E-11	↑
Childhood obesity	Early Growth Genetics Consortium GWAS	0.00015	↑	ND		0.09	0	0.15	↑
Glycemic traits									
Fasting insulin adj BMI	AAGILE GWAS	0.18	↓	ND		0.02	↓	0.068	0
Insulin secretion rate	IVGTT-Based Insulin Secretion GWAS	0.77	↑	0.70	↓	0.18	↓	0.028	↓
Insulin secretion rate adj BMI	IVGTT-Based Insulin Secretion GWAS	0.71	↑	0.73	↓	0.20	↓	0.029	↓
Two-hour insulin	MAGIC GWAS	0.14	↑	ND		0.013	↑	0.068	0
HOMA-IR	MAGIC GWAS	0.15	↑	ND		0.048	↑	0.067	0
Insulin sensitivity	GENESIS GWAS	0.32	↓	0.048	↓	0.36	↓	0.24	0
Liver traits									
Alkaline phosphatase	Liver function GWAS	0.53	0	ND		0.004	↑	0.57	↑
Bilirubin	BioBank Japan GWAS	0.019	↑	0.53	0	0.024	↑	0.38	↑
Serum albumin	BioBank Japan GWAS	0.026	↑	0.78	0	0.033	↑	0.33	0
Lipid traits									
Hyperlipidemia	UKB SAIGE	5.9E-07	↑	1.2E-06	↑	0.002	↑	0.004	↑
Hypercholesterolemia	UKB SAIGE	6.3E-06	↑	1.5E-05	↑	0.018	↑	0.025	↑
Sleep and circadian traits									
Excessive daytime sleepiness	UK Biobank Sleep Traits GWAS: Self-report	0.009	↑	0.046	↑	0.0029	↑	0.0027	↑
Long sleep duration	UK Biobank Sleep Traits GWAS: Self-report	0.021	↑	0.16	↑	0.019	↑	0.017	↑
Short sleep duration	UK Biobank Sleep Traits GWAS: Self-report	0.18	↑	0.08	↑	0.00016	↑	1.8E-05	↑

P-values for reported GWAS associations in DIAMANTE and other public databases (www.type2diabetesgenetics.org; date 10/2020) are shown for the 4 credible SNPs (posterior probability of association >1%) associated with T2D in the *ONECUT1*-associated genomic region (99% genetic credible SNPs; Mahajan et al.[5]).

The direction of effect (dir), corresponding to the T2D-associated allele, is shown by an arrow (increased risk: up arrow, decreased risk: down arrow). ND: not done. Red: $P<5x10^{-8}$ (genome-wide significant); orange: $P<5x10^{-4}$ (locus-wide significant); light orange: $P<0.05$ (nominal significant) .

Supplementary Table 10. ACMG status of rare *ONECUT1* coding variants identified in the heterozygous status in diabetic patients and control variants

		Population association data			Segregation data	Pathogenicity data		ACMG status	
AA change	Allele frequen-cy	Ulm diabetes cohort: het_cases/ het_controls (2165 cases, 559 controls)	AMP T2D-GENES: OR, P, het_cases/ het_controls (19,852 cases, 23,273 controls)	ACMG code	Adult/Juvenile-onset diabetes	Computational data: number of pathogenic/benign predictions; Other characteristics	In vitro functional data (our study)	Final evidence	Conclusion
Variants causing monogenic recessive syndromic diabetes and monogenic dominant diabetes (Families 1 and 2)									
E231X	PM2	Absent	Absent	-	Dominant transmission, Family 1 (PP1++)	4/0 (PP3); null variant (PVS1)	Deleterious+ (PS3+)	PM2, PP1+, PP3, PVS1, PS3+	P
E231D	PM2	Absent	Absent	-	Dominant transmission, Family 2 (PP1+)	2/9 (BP4); missense intolerant (PP2)	Deleterious+ (PS3+)	PM2, PP1+, BP4, PS3+, PP2	P
Variants identified by population screening (Ulm diabetes cohort)									
A3P	PM2	1/0	Absent	PS4-	Dominant family history (PP4)	3/7 (BP4); missense intolerant (PP2)	ND	PM2, PS4-, PP4, BP4, PP2	L.P
G30S	PM2	1/0	OR=9.8, P=0.16 18/5	PS4	Dominant family history (PP4)	2/8 (BP4); missense intolerant (PP2)	Deleterious (PS3)	PM2, PS4, PP4, BP4, PP2, PS3	P
H33Q	PM2	3/0	OR=4.95,P=0.079 50/41	PS4	Dominant family history (PP4)	7/3 (PP3); missense intolerant (PP2)	Deleterious (PS3)	PM2, PS4, PP4, PP3, PP2, PS3	P
G62C	PM2	1/0	OR=4.14, P=0.54 3/0	PS4	Dominant family history (PP4)	6/4 (PP3); missense intolerant (PP2)	ND	PM2, PS4, PP4, PP3, PP2	P
G81D	PM2	1/0	OR=0.80, P=0.59 16/19	-	Dominant family history (PP4)	5/5 (NE); missense intolerant (PP2)	Deleterious (PS3)	PM2, PP4, PP2, PS3	L.P
G96C	PM2	1/0	Absent	PS4-	Dominant family history (PP4)	8/2 (PP3); missense intolerant (PP2)	ND	PM2, PS4-, PP4, PP3, PP2	L.P
G96D*	PM2	1/0	NA for G96D/P75A*	-	Dominant family history (PP4)	8/2 (PP3); missense intolerant (PP2)	ND	PM2, PP4, PP3, PP2	L.P
P113A	PM2	1/0	Absent	PS4-	Dominant family history (PP4)	5/5 (NE); missense intolerant (PP2)	ND	PM2, PS4-, PP4, PP2	L.P
P215A	PM2	1/0	Absent	PS4-	Dominant family history (PP4)	5/6 (BP4); known pathogenic AA (PM5); missense intolerant (PP2)	ND	PM2, PS4-, PP4, BP4, PM5, PP2	L.P
R269G	PM2	2/0	OR=1.3, P=0.47 20/21***	PS4	Dominant family history (PP4)	6/5 (PP3); missense intolerant (PP2)	ND	PM2, PS4, PP4, PP3, PP2	L.P

Mutation occurring *de novo* in a juvenile-onset insulin-treated diabetic patient (Lebanon)									
P215R	PM2	Absent	Absent	-	-	9/2 (PP3); *de novo* mutation (PS2); missense intolerant (PP2)	Deleterious (PS3)	PM2, PP3, PS2, PP2, PS3	P
T2D risk variant, AMP T2D-GENES									
V242A	PM2	Absent	OR=1.4, P=0.026 101/75	PS4	-	4/6 (BP4); missense intolerant (PP2)	Deleterious (PS3)	PM2, PS4, BP4, PP2, PS3	P
Control variants, selected as likely neutral									
D26E	BS1**	Absent	OR=0.99, P=0.98 66/50	-	-	2/7 (BP4); missense intolerant (PP2)	Neutral (BS3)	BS1, BP4, PP2, BS3	B
K412R	PM2	Absent	OR=1.11, P=0.95 1/1	-	-	4/2 (PP3); missense intolerant (PP2)	Neutral (BS3)	PM2, PP3, PP2, BS3	U.S

The status was defined according to the American College of Medical Genetics and Genomics (ACMG) recommendations using Varsome tool available online (http://varsome.com; date 15/12/2020) on *ONECUT1* RefSeq canonical transcript (NM_004498.2), integrating information from the present study. Criteria were adjusted based on data from the present study and the specificities of the gene. ND: not done; NA: not available; NE: no evidence.

*The patient heterozygous for p.G96D (UDC-T2D Patient-4) was also heterozygous for the low frequency p.P75A variant (G96D rare variant and P75A low-frequency variant in trans); ** Max MAF=0.014 in AFR population, with 2 homozygous subjects (gnomAD). *** Additional association data were available for this variant in DIAMANTE GWAS (19852 cases/23273 controls, OR=1.06, P=0.55).

In vitro functional evidence are from the present study, coded as follows: deleterious+: assessed by our in-depth functional studies; deleterious, neutral: assessed by luciferase assays only. MAF threshold used for Ulm diabetes population screening: 0.005 in Non-Finnish European/North-West European subgroup.

Prediction programs used (Varsome): DANN, DEOGEN2, FATHMM-MKL, M-CAP, Mutation Assessor, Mutation Taster, Primate AI, Polyphen 2, MVP, EIGEN, REVEL, SIFT, GERP (conservation)

ACMG categories: PVS1: null variant (nonsense, frameshift, canonical +- 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease; PS2: De novo (both maternity and paternity confirmed) in a patient with the disease and no family history; PS3: well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product; PS3+: PS3 as a very strong evidence due to our in vitro functional studies supportive of a damaging effect on the gene or gene product); PS4: The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls; PS4-: PS4 as a moderate evidence (overall significant association in the German T2D population, with specific clinical features); PM2: Absent from controls (or at extremely low frequency if recessive) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium; PM5: Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before; PP1: Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease; PP1+/PP1++ (PP1 as strong/very strong evidence due to increasing segregation data); PP2: low rate benign missense (for rare variants, present study; observed/expected ratio for *ONECUT1* missense variants=0.91 [237/260.8, CI:0.82-1.01] compared to observed/expected ratio for *ONECUT1* synonymous variants=1.04 [115/110.2, CI:0.9-1.22], gnomAD); PP3: Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.); PP4: Patient's phenotype or family history is highly specific for a disease with a single genetic etiology; BS1: Allele frequency is greater than expected for disorder; BS3: Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing; BP4: Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.).

Overall conclusion: P: Pathogenic, L.P: Likely pathogenic, U.S: Uncertain significance, B: Benign.