1. Clinical description of the patients and their families

1.1. Family 1: patients 1 and 2

Patient 1, a boy, and patient 2, his elder sister, were born to healthy French 2nd cousin consanguineous parents. The couple had two miscarriages in addition to their two children. At the time of the study patient 1 was alive (32 years old) and was included in the genetic study and patient 2 was dead (death more than 25 years ago) with no biological material available.

Patient 1 was born at term following an uncomplicated pregnancy, with a birth weight of 3,600 g. His development was normal, with a normal adult stature (180 cm, adult) and BMI (22.8 kg/m^2 at 24 years). Diabetes was diagnosed at 28 years when he was displaying symptomatic hyperglycemia including polyuria, polydipsia and an 8 kg weight loss. Initially, he was treated with oral antidiabetic drugs only (a combination of metformin and gliclazide). GAD and IA2 autoantibodies were negative at the onset of diabetes. Four years after diabetes onset, insulin secretion and insulin action were assessed using an 75 g Oral Glucose Tolerance Test (OGTT) and a short intravenous insulin tolerance test (ITT) (1) respectively. Early insulin response to oral glucose was dramatically decreased compared to both control subjects and T2D patients, and insulin sensitivity was decreased to the range of T2D patients (Supplementary Table 5). Because glucose control was not optimal, with HbA1c value between 7 and 7.5% (N<5.7%) and post-prandial capillary glucose level between 7.3 and 9.6 mmol/l (recommended value for T2D patients <8.8 mmol/l), a bed time insulin injection was started with the continuation of oral antidiabetic drugs (2 g of metformin in combination with 120 mg of gliclazide per day). In addition, he was noted to have macrocytosis since birth (mean erythrocyte volume: 100-116 fl), without anemia except during 2 infectious episodes: at 5 years old, he had pneumonia due to *Mycoplasma pneumoniae* infection, with regenerative

anemia associated with hemolysis; at 6 years old, he had lymphocytic meningitis with decreased hemoglobin at 11 g/l. He never received any blood transfusion. Plasma concentrations of vitamin B12 and folic acid were always normal. He also had multiple naevi, which were resected. His general health is excellent at 32 years old, under diabetes treatment.

Patient 2 was born at 36 weeks of gestational age with a birth weight of 2,480 g (50^{th} percentile), following a pregnancy that was complicated by two threatened abortions. At birth she had macrocytic anemia, with mean erythrocyte volume of 139 fl and hemoglobin at 12.2 g/l, falling to 7.9 g/l 11 days after birth, requiring transfusion. Anemia was highly regenerative (185,000 reticulocytes at 2 days of life), bone marrow aspirate was normal in the neonatal period, with no megaloblastosis and a rich bone marrow, with normality of the three lines. Vitamin B12 and folic acid were normal. Fetomaternal erythrocyte incompatibility was excluded. At 5 years old, a routine complete blood count revealed thrombopenia, in addition to macrocytic anemia and leucopenia (Table 1), despite no clinical complaint. Vitamin B12 and folic acid levels remained normal. Fetal hemoglobin was high. She also had multiple naevi. Shortly after, hyperglycemia was found and diabetes diagnosed and insulin treatment was started. Thrombopenia progressively worsened, and transfusions were started at 6 years old. Physical and psychomotor development was normal and remained so thereafter. At 8 years old, she was treated by androgens, allowing the temporary suspension of blood transfusion. However, at 10 years old her general health condition degraded, including severe hepatitis with cholestasis, pericarditis and increased pancytopenia; androgen therapy was then discontinued, resulting in the remission of the liver and cardiac manifestations. She died at 11 years old following brain hemorrhage. Autopsy showed important signs of hemochromatosis particularly in the liver, lung and exocrine pancreas. Pancreas histology showed a fibrotic pancreas with rare and hypertrophic Langerhans islets and partial fat involution.

Additional explorations of patients 1 and 2

Bone marrow aspirates and bone marrow biopsies were performed repeatedly in patient 2 during the course of her disease, and bone marrow aspirate twice in patient 1.

In patient 1, bone marrow aspirates performed at 6 and 22 years old showed increased erythroblasts (>40%; N: 25±5%), that were macroblastic and with some binucleated cells, indicating dyserythropoiesis, while the other lines appeared normal.

In patient 2, bone marrow aspirate performed at 5 years old showed clear erythroblastosis (50-55% ; N: 25±5%) and decreased megacaryocytes; erythroblasts were megaloblastic, with important abnormalities (binucleated nuclei, abnormal shape), indicating dyserythropoiesis; bone marrow biopsy showed a high cellularity. At 9 years old, the erythrocytic line was at 86% (N : 25±5%), the granulocytic line was normal and megacaryocytes were absent; bone marrow biopsy showed a medium cellularity. At autopsy (11 years old), bone marrow cellularity was poor, evidencing medullar aplasia. Erythrokinetic studies performed in patient 2 at 5 years showed a shortened erythrocyte lifespan (66 days, N=100-120 days) and an active erythropoiesis, with a total number of erythroblasts estimated at >3 times the normal reference values.

Additional tests were also done in both patients at age 6 years (patient 1) and 9 years (patient 2): 1) karyotype was done several times in both patients, on peripheral lymphocytes and on bone marrow cells, with and without induction of breakage by antimitotic or alkylating agents. In all instances, the karyotype was normal, with no translocations and no increased sensitivity to chromosome breakage, excluding Fanconi anemia diagnosis; 2) Levels of vitamin B12, serum and intra-erythrocyte folates and of cobalamine were normal; 3) Metabolic status and blood amino acid chromatography were normal; 4) Urinary amino acids,

including urinary orotic acid, were normal, thus excluding Uridine Monophosphate Synthase (UMPS) deficiency; 5) Autoimmune status: antiglobulin test (Coombs test) was negative, anti-polynuclear antibodies, anti-platelet antibodies, antinuclear antibodies, anti-tissue antibodies were all absent and IgG levels were normal.

1.2. Family 2: patients 3 and 4

Patient 3 was an Egyptian boy born to healthy consanguineous parents; he was resident in France at the time of the study. T1D was diagnosed at age 17 years. GAD and IA2 autoantibodies were negative when tested four years later. He presented at age 21 years with medullar aplasia. Based on his medical records, pancytopenia had been diagnosed in infancy and he received blood transfusions at age 10 and 14 years, each time in a context of viral infection. Following a vaccination at age 19, his condition degraded to medullar aplasia and he received regular transfusions, initially every 3 weeks, then weekly from age 20. At age 21, he had severe pancytopenia, with absence of reticulocytes (Table 1). Karyotype was normal, with no increase of spontaneous and induced chromosome breakage. Bone marrow biopsy done at age 21 years showed high cellularity, moderate erythroblastosis with signs of dyserythropoiesis. At 22 years old, he developed acute lymphoblastic leukemia, which was treated by chemotherapy and bone marrow allograft. He then had a severe graft versus host disease, affecting the skin, digestive tract and liver, and died shortly after. Autopsy was not performed.

Patient 3 had two healthy siblings and one brother (patient 4) who also had medullar aplasia that had been successfully treated by a bone marrow allograft at the age of 10 years and who had T1D diagnosed at 12 years old. Patient 4 was alive at 24 years at the time of the study. Patient 3's parents and siblings were all living in Egypt and were not available for genetic and clinical investigations.

2. Additional patients studied

To search for additional patients with *DUT* mutations and possibly extend the phenotype spectrum of the syndrome, we selected patients from three independent groups:

2.1. Patients with diabetes associated with bone marrow failure: 4 selected subjects

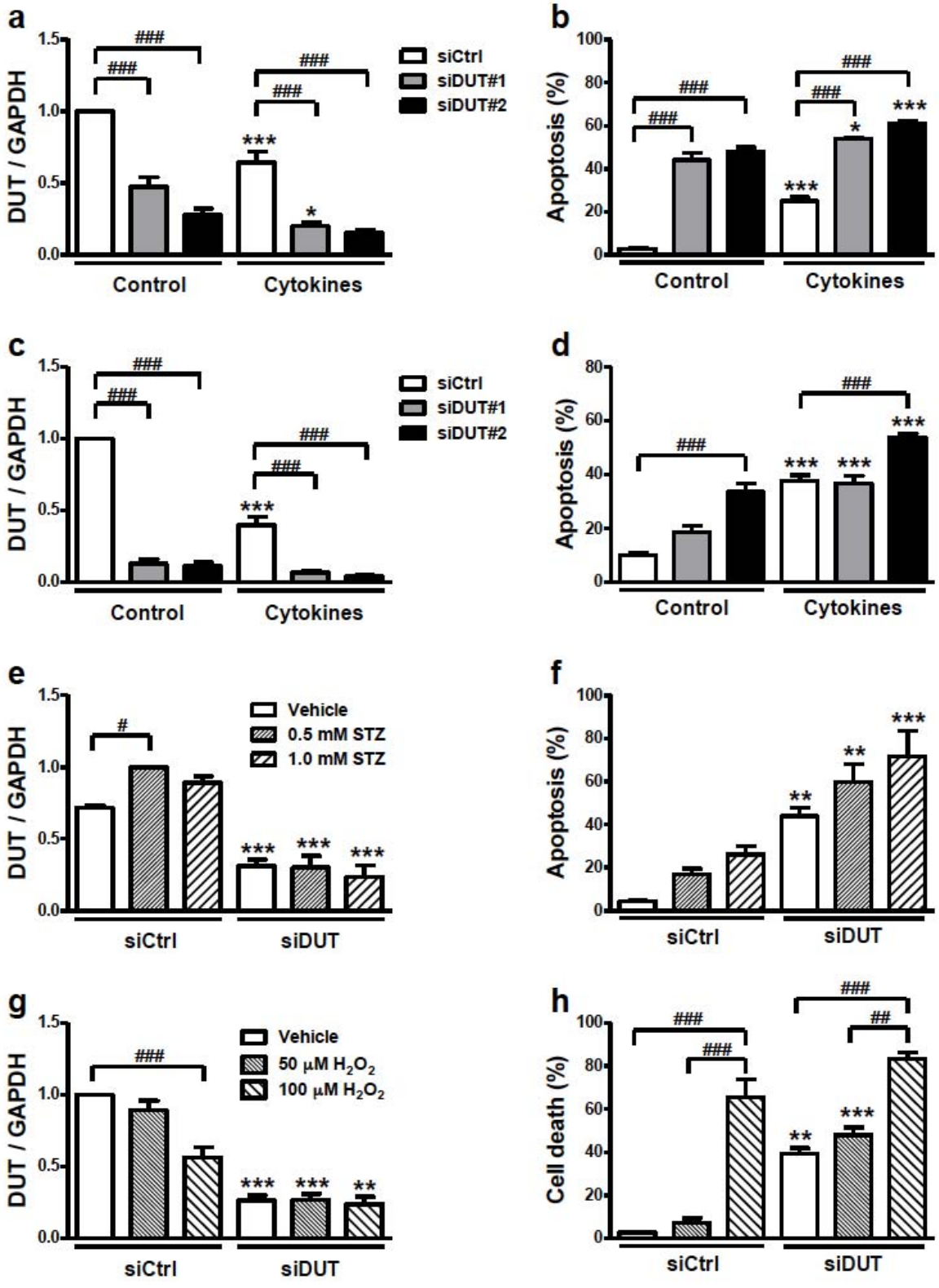
These patients were identified through several pediatric and adult diabetes clinics or by the French Reference Center for Constitutional Bone Marrow Failure (Saint-Louis Hospital, Paris, France). Four patients had bone marrow failure with increased chromosome breakage (Fanconi Anemia or similar syndromes) (2) and had a mutation in a known disease-causing gene and were excluded for *DUT* mutation screening. Four patients did not have increased chromosome breakage and had not been genetically characterized. Diabetes in these patients ranged from insulin-treated neonatal diabetes to young-onset T2D. Bone marrow deficit ranged from dyserythropoiesis to complete medullar aplasia. They all had additional clinical manifestations (short stature, intellectual impairment and/or dysmorphology) that were absent from the patients of the two index families studied.

2.2. Juvenile-onset diabetic patients that were likely to be monogenic and whose family was compatible with genetic mapping to *DUT* chromosome region: 14 selected subjects. We selected patients and their families that were likely to have monogenic diabetes based on family structure (consanguinity or multiplex family) or atypical clinical presentation, that were compatible with linkage to *DUT* chromosome region based on parametric linkage analysis (rare recessive model, no phenocopy) performed on genome-wide genotype data. These patients were identified through diabetes clinics. In total, we identified 14 unrelated diabetic patients and their families corresponding to these criteria.

2.3. Patients with bone marrow failure or myelodysplastic syndrome, with increased risk of monogenic cause based on a young age at onset, physical abnormalities and/or familial history: 95 selected subjects. The patients were followed up by the French

Reference Center for Constitutional Bone Marrow Failure (Saint-Louis and Robert Debré Hospitals, Paris, France). Patients diagnosed with Fanconi Anemia based on chromosomal breakage test (2) were excluded. We selected 95 unrelated patients for genetic screening of *DUT* mutations. These patients were not known to have diabetes based on their clinical records.

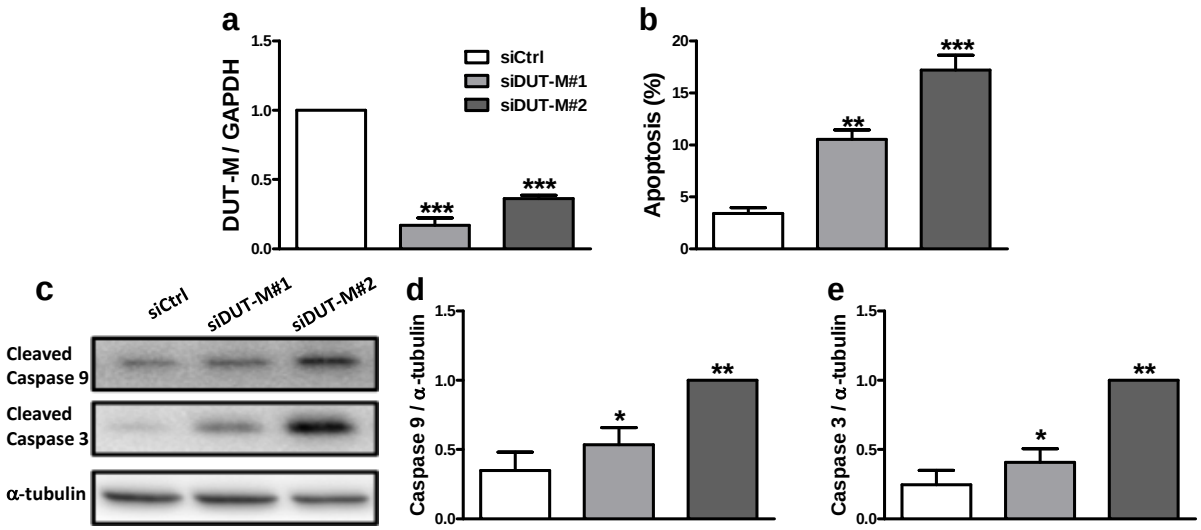
Supplementary Figure 1



Supplementary Figure 1. Beta cell death in DUT-silenced cells upon different stimuli.

(**a-d**) INS-1E cells (**a,b**) or primary rat β-cells (**c,d**) were transfected with siCtrl or with two different siRNAs targeting rat *DUT* (siDUT#1 and siDUT#2). After 24 h of recovery cells were left untreated or treated with IL-1β + IFNγ (10 and 100 U/mL, respectively) for 24 h (**a,b**), and with IL1β + IFNγ (50 and 500 U/mL, respectively) for 48 h (**c,d**). To confirm knockdown, DUT mRNA expression was evaluated by RT-qPCR in INS-1E (**a**, $n = 4$) and primary rat β-cells (**c**, $n = 5$). DUT expression was normalized by the housekeeping gene GAPDH and then by the highest value of each experiment considered as 1. Results are means ± SEM (*n* as indicated for each figure). Beta cell apoptosis was evaluated in INS-1E (**b**, $n = 4$) and primary rat β-cells (**d**, $n = 5$) using Hoechst 33342/PI staining. Results are means ± SEM (*n* as indicated for each figure). (**e,f**) INS-1E cells were transfected with siCtrl or with two different siRNAs targeting rat *DUT* (siDUT#1 and siDUT#2). After 24 h of recovery, cells were left untreated or treated with streptozotocin (STZ, 0.5 or 1.0 mM) for 30 min, washed with PBS and incubated for 24 h. (**e**) Knockdown of DUT was confirmed by RT-qPCR as measured in (**a,c**). (**f**) Beta cell apoptosis was evaluated in INS-1E using Hoechst 33342/PI staining. Results are means ± SEM of 5 independent experiments. (**g,h**) INS-1E cells were transfected with siCtrl or with two different siRNAs targeting rat *DUT* (siDUT#1 and siDUT#2). After 24 h of recovery, cells were left untreated or treated with H_2O_2 (50 or 100 μM) for 24 h. (**g**) Knockdown of DUT was confirmed by RT-qPCR as in (**a,c**). (**h**) Beta cell death was evaluated in INS-1E using Hoechst 33342/PI staining and is the result of summing the percentage of apoptotic and necrotic cells. Results are means ± SEM of 4 independent experiments; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$ compared to *siCtrl*. $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$ as indicated by bars. ANOVA + Bonferroni correction.

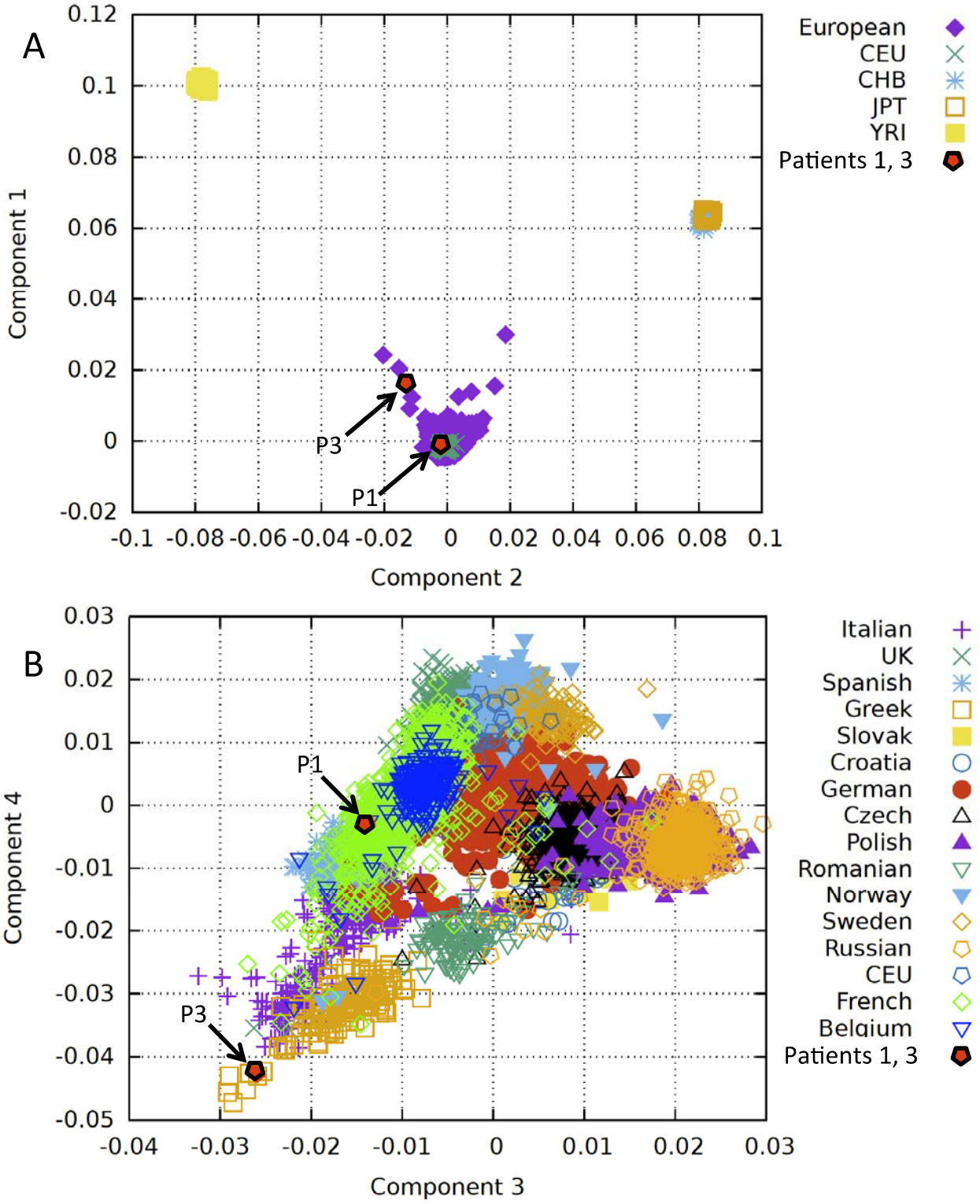
Supplementary Figure 2



Supplementary Figure 2. DUT-M inhibition exacerbates apoptosis in beta cells.

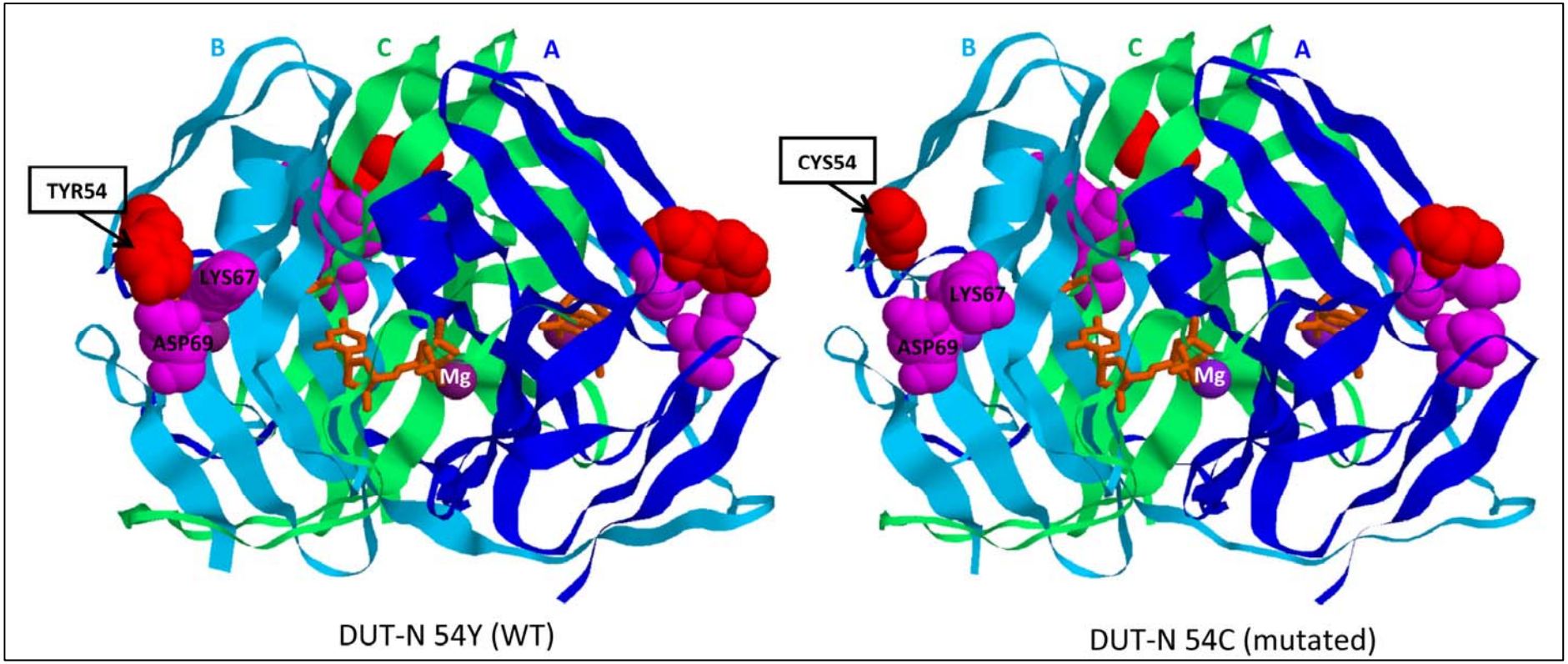
INS-1E cells were transfected with siCtrl or with two different siRNAs targeting rat *DUT-M* (siDUT-M#1 and siDUT-M#2) and left to recover for 48 h. (**a**) To confirm knockdown, DUT-M mRNA expression was evaluated by RT-qPCR, normalized by the housekeeping genes GAPDH and then by the highest value of each experiment (considered as 1). Results are means ± SEM of 5 independent experiments. (**b**) Beta cell apoptosis was evaluated using Hoechst 33342/PI staining. Results are means ± SEM of 5 independent experiments. (**c**) Western blot showing expression of cleaved caspase 9, cleaved caspase 3 and α-tubulin (used as loading control) in INS-1E cells transfected with siCtrl or siDUT-M (siDUT-M#1 and siDUT-M#2). The figure is representative of 4 independent experiments. (**d,e**) Densitometry analysis of the Western blots (**c**) for cleaved caspases 9 (**d**) and 3 (**e**) after normalization for the protein content using α-tubulin and then by the highest value of each experiment (considered as 1). Results are means ± SEM of 4 independent experiments. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 compared to *siCtrl*. Paired Student's t-Test.

Supplementary Figure 3



Supplementary Figure 3. Principal Component Analysis of high-density SNP genotyping array of Patients 1 and 3 and various control population groups. A: components 1 and 2, showing that both patients belong to the extended European cluster; B: components 3 and 4, dissociating patient 1 (French) and patient 3 (Egyptian) within the extended European group. P1 and P3: patients 1 and 3. CEU: Caucasian European (CEPH), CHB: Chinese, JPT: Japanese, YRI: Yoruban.

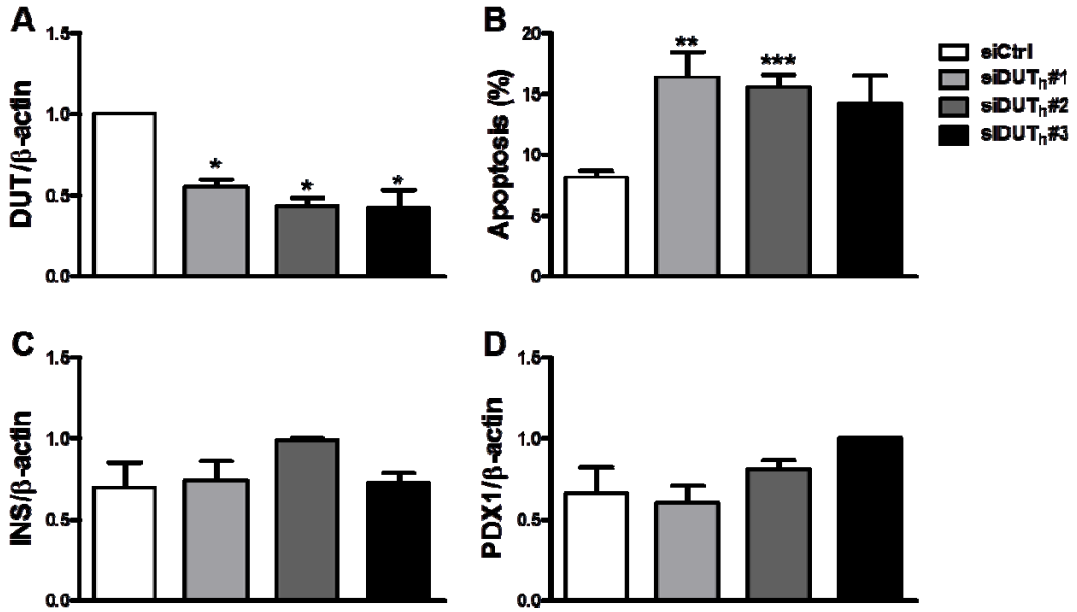
Supplementary Figure 4



Supplementary Figure 4. Structure of the homotrimeric DUT complex and predicted consequences of the DUT Tyr54Cys (DUT-N)/ Tyr142Cys (DUT-M) mutation (left: wild type (WT), right: mutated).

DUT structure (WT) is according to PDB (3EHW) (3), and corresponds to the C-terminal part of the protein (last 141 amino-acids), which is common to the DUT-N and DUT-M isoforms. As it corresponds to the quasi-totality of the DUT-N isoform (excluding the first 23 amino-acid), we have named the residues with respect to this isoform. In the mutated model (DUT-N Cys54), the interaction of the Tyr54 residue (red) with several residues was predicted to be lost, including the Lys67 and Asp69 residues (magenta) and other residues (Supplementary Table 11). The three subunits are color-coded (cyan, green, blue), the α,β-imido-dUTP substrate is colored in orange, and the Mg2+ ion in purple. DUT-N Tyr54 (WT) structure is according to PDB (3EHW). DUT-N Cys54 (mutated) structure is modeled based on 3EHW using MODELLER package.

Supplementary Figure 5



Supplementary Figure 5. DUT inhibition exacerbates basal apoptosis in dispersed human islets, but does not affect insulin or PDX1 mRNA expression. Dispersed human islet cells were transfected with siCtrl or siDUT (siDUTh#1, siDUTh#2 and siDUTh#3). A: To confirm knockdown, DUT mRNA expression was evaluated by RT-qPCR, normalized by the housekeeping gene β-actin, and then by the highest value of each experiment considered as 1. B: Apoptosis was evaluated using Hoechst 33342/PI staining. C and D: Insulin and PDX1 mRNA expression was evaluated by RT-qPCR, normalized by the housekeeping gene β-actin, and then by the highest value of each experiment considered as 1. Results are means ± SEM of 3 independent experiments; *P < 0.05, **P < 0.01, ***P < 0.001 compared to siCtrl. Paired Student's t-Test.

Supplementary Table 1. PCR amplification and sequencing primers used for Sanger sequencing of *DUT*

Region amplified	Chr15 map position[a]		PCR amplification and sequencing primer (5'-3')		Internal sequencing primers	Product Size (bp)
	Start	End	Forward	Reverse		
Promoter, DUT-M	48621559	48622232	AAAGCACTTAGAATACACCTGACA	AAAGAGAAACGGCCATAATCTA		674
Promoter, DUT-M	48622174	48622528	TGCTTACTGGCTCAGAATCC	GTAGTGACCTCTGCCCCATC		355
Promoter, DUT-M	48622737	48623242	GCTCGGATCGTGTTGCTA	TTTCAGCACCGAGTGTCAG		506
Exon 1*, DUT-M	48623142	48624039	CTGCGTTTCTGTTGGGTTATC	CGTGGATTCCTTCCAGACG	AGGTAAATCAGTCCAGGAGCAG TTTCCTTACGTCTCTGCTTCG	898
Promoter, DUT-N	48623686	48624439	ACCCAACCAGCCCAAACTC	AGAAGGCGAGCGAGGAGAC		754
Exons 2a (DUT-M) and exon 2b* (DUT-N)	48623915	48624960	GCGGAGCCAGTACAGTCG	TTTCACGCTCACATCTAGGC	AGCCAGAACTGTGGACTCGT	1046
Exon 3	48626418	48626813	AAATAGCCACCTCTACCTTCTTCA	AGGACACAGCCCAAGCAATA		396
Exon 4	48627929	48628706	TTGCTCACCTTTCAGCTCATT	GGGAAACAGGTAACACAGAGG		778
Exons 5 and 6	48633353	48634239	CCAGGGCAAGATTCACTTATTC	CCTTTCGGTGTCATCCAAG		887
Exon 7	48634143	48635139	TCTTAGGAGCTGGAGAGGAAA	CTGCAATGAAATGCTGCTTATG	AAGTTGCTGAAGGGAGATCA	997
Exon 7	48635019	48635882	AGTGAAGTAGCAATAGGCTGTAATC	ACCATCTCAGGTATAGCAGCAC		864

 Map position on hg19.

Exons are named with respect to DUT-M isoform. The first exon of each isoform is indicated by a « * » . Exon 1 (5'UTR and coding) is specific to DUT-M; exon 2b (5'UTR and coding) is the first exon on DUT-N and includes exon 2a (2^{nd} exon of DUT-M, coding); exons 3 (coding), 4 (coding), 5 (coding), 6 (coding) and 7 (coding and 3'UTR) are common to DUT-M and DUT-N isoforms.

Supplementary Table 2. Regions covered by targeted next generation sequencing of *DUT*

Region amplified	Chromosome 15 map position[a]	
	Start	End
Exon 1*, DUT-M	48623676	48624169
Exon 2a (DUT-M) and exon 2b* (DUT-N)	48624410	48624673
Exon 3	48626523	48626791
Exon 4	48628204	48628432
Exon 5	48633349	48633616
Exon 6	48633543	48633788
Exon 7	48634158	48634384

 Map position on hg19.

Nomenclature of exons as in Supplementary Table 1.

Supplementary Table 3. Small interfering RNAs used in the study

Name	Description	Distributor	Sequence
siCTRL	Allstars Negative Control siRNA	Qiagen, Verlo, Netherlands	Sequence not provided
siDUT#1 (rat)	Stealth Selected siRNAi siRNA Duplex Oligonucleotides	Invitrogen, Pasley, UK	5'-CAGUGCCUAUGAUUAUACAAUUCCA-3' 5'-UGGAAUUGUAUAAUCAUAGGCACUG-3'
siDUT#2 (rat)	Stealth Selected siRNAi siRNA Duplex Oligonucleotides	Invitrogen, Pasley, UK	5'-GAGGAAAUGUCGGCGUCGUGCUGUU-3' 5'-AACAGCACGACGCCGACAUUUCCUC-3'
siDUT$_h$#1 (human)	Stealth Selected siRNAi siRNA Duplex Oligonucleotides	Invitrogen, Pasley, UK	5'-GGACAUUCAGAUAGCGCUCCCUUCU-3' 5'-AGAAGGGAGCGCUAUCUGAAUGUCC-3'
siDUT$_h$#2 (human)	Stealth Selected siRNAi siRNA Duplex Oligonucleotides	Invitrogen, Pasley, UK	5'-CAGAAAUAGAAGAAGUUCAAGCCUU-3' 5'-AAGGCUUGAACUUCUUCUAUUUCUG-3'
siDUT-M#1 (rat)	Stealth Selected siRNAi siRNA Duplex Oligonucleotides	Invitrogen, Pasley, UK	5'-CAACCUUCUCUGCUCCGCUCGCUUA-3' 5'-UAAGCGAGCGGAGCAGAGAAGGUUG-3'
siDUT-M#2 (rat)	Dharmacon® Option A4 Custom siRNA	Thermo Scientific, Chicago, IL., USA	5'-CGGCGGAAGCAGAGGAGUCUU-3' 5'-AAGACUCCUCUCGCUUCCGCCG-3'

Supplementary Table 4. Primers used for expression studies

Primers	Forward (5'-3')	Reverse (5'-3')
Rat GAPDH	AGTTCAACGGCACAGTCAAG	TACTCAGCACCAGCATCACC
Human β-actin	TCTACGCCAACACAGTCGT	GCTCAGGAGGAGCAATGATC
Rat DUT	GCCCTTGTGAAGACAGACAT	ACACCGGCTCCTACATCTAT
Human DUT	CCCTCAGTGAAGTAGCAATAGG	AAACCTGGTGGCTGAAGTATTA
Rat DUT-M	CGACTCCAACCTTCTCTGCT	GCTCGAGCCCTCTTAGAGAC

Supplementary Table 5. Metabolic evaluation of Patient 1 prior to insulin treatment

	Patient 1	Non diabetic controls Mean ± SD	T2D patients Mean ± SD
Age (years)	32	26.8 ± 6.6	44.4 ± 10.7
BMI (kg/m^2)	21.4	22.9 ± 3.3	27.2 ± 4.1
Fasting glucose (mmol/l)	4.3	4.6 ± 0.3	6.2 ± 1.3
Fasting insulin (mUI/l)	1.0	5.5 ± 3.8	10.8 ± 8.4
Early insulin secretion during OGTT (mUI/mmol)	0.1	16.9 ± 10.7	3.9 ± 2.3
K_{ITT} (%/min)	2.1	5.6 ± 0.9	2.1 ± 0.9

Early insulin secretion in response to an oral glucose load was calculated as the ratio [insulinemia at 30 min (I30) - fasting insulinemia (I0)]/[blood glucose at 30 min (G30)–fasting blood glucose (G0)]. Insulin sentitivity was evaluated using a short insulin intravenous tolerance test (ITT) by the constant rate plasma glucose disappearance parameter (K_{ITT}) as described (1). Mean and SD were estimated in 18 non diabetic controls (4) and 15 T2D patients (JF Gautier, unpublished data).

Supplementary Table 6. Linkage regions identified in family 1

Chromosome	Start	End	Region length
5	105,425,352	116,752,778	11,327,427
8	1	5,981,454	5,981,454
15	38,986,368	71,722,819	32,736,452

The maximum predicted LOD score (1.78) in family 1 was reached in three chromosome regions, indicating continuous homozygosity within defined boundaries (linkage regions). Linkage regions are defined by the map positions (hg19) of the flanking heterozygous SNPs (start and end) of regions of homozygosity in patient 1 identified by linkage analysis.

Supplementary Table 7. Filtering of variants identified by whole-exome sequencing

Filter	Whole-exome variants, patient 1
All variants	52,506
Homozygous variants[a]	21,145
Coding variants[b]	3,811
Rare variants[c]	1

Counts are the number of variants (SNVs and insertion/deletion variants (Indels)) identified by whole-exome sequencing of patient 1 compared to the Human Reference Genome on UCSC build hg19. The consecutive filters applied after quality control are shown: [a]homozygous variants, [b]coding variants including missense and nonsense, splice-site variants and Indels (frameshift and inframe), [c]variants that were absent in the homozygous status and with MAF<0.005 in public (EVS, ExAC, dbSNP) and in-house databases. A unique rare homozygous variant was identified in patient 1 after applying these consecutive filters. This variant is located on chr15:g.48626619A>G (hg19) within the 33 Mb linkage region on chr15 and affects the dUTPase (*DUT*) gene (NCBI Gene ID: 1854).

Supplementary Table 8. Mutation description in the dUTPase gene (*DUT*)

Genomic position (dbSNP rsID)	DUT isoform	cDNA		Protein		Genotype counts AA/GA/GG and MAF in EVS	Genotype counts AA/GA/GG and MAF in ExAC
		Refseq	Nucleotide exchange	Refseq	Protein effect		
chr15: g.48626619A>G (rs373184762)	DUT-M	NM_001025248	c.425A>G	NP_001020419	p.Tyr142Cys	EA : 4295/2/0, 0.000233 AA : 2198/0/0, 0 ALL: 6493/2/0, 0.000154	CEU : 33309/8/0, 0.000120 OTH: 27298/0/0, 0 ALL: 60607/8/0, 0.000066
	DUT-N	NM_001948	c.161A>G	NP_001939	p.Tyr54Cys		

Mutation description and cDNA and protein consequences are shown for DUT-M (mitochondrial) and DUT-N (nuclear) isoforms. Genotype counts (three genotypes) and rare allele frequencies (MAF) provided by Exome Variant Server (EVS) are shown for European American (EA), American American (AA) and overall (ALL). Genotype counts and rare allele frequencies provided by the Exome Aggregation Consortium (ExAC) (5) are shown for European [non-Finnish] (CEU), all other populations [European (Finnish), African, East Asian, South Asian, Latino, Other] (OTH) and overall (ALL). Because of partial inclusion of EVS dataset in ExAC, the CEU MAF from ExAC is the most accurate estimate of the Caucasian European MAF. Genomic position is on hg19. Mutation nomenclature follows the Human Genome Variation Society (HGVS) recommendations.

Supplementary Table 9

Population Code	Population Description	Super Population Code	N	DUT haplotype frequency estimate
CHB	Han Chinese in Bejing, China	East Asian	103	0
JPT	Japanese in Tokyo, Japan	East Asian	104	0
CHS	Southern Han Chinese	East Asian	105	0
CDX	Chinese Dai in Xishuangbanna, China	East Asian	93	0
KHV	Kinh in Ho Chi Minh City, Vietnam	East Asian	99	0
CEU	Utah Residents (CEPH) with Northern and Western Ancestry	European	99	$5.1\ 10^{-21}$
TSI	Toscani in Italia	European	107	$1.1\ 10^{-15}$
FIN	Finnish in Finland	European	99	$4.3\ 10^{-17}$
GBR	British in England and Scotland	European	91	$2.3\ 10^{-15}$
IBS	Iberian Population in Spain	European	107	$1.2\ 10^{-15}$
YRI	Yoruba in Ibadan, Nigeria	African	108	0
LWK	Luhya in Webuye, Kenya	African	99	0
GWD	Gambian in Western Divisions, Gambia	African	113	0
MSL	Mende in Sierra Leone	African	85	0
ESN	Esan in Nigeria	African	99	0
ASW	Americans of African Ancestry in SW USA	African	61	0
ACB	African Caribbeans in Barbados	African	96	0
MXL	Mexican Ancestry from Los Angeles USA	Admixed American	64	$3.4\ 10^{-18}$
PUR	Puerto Ricans from Puerto Rico	Admixed American	104	$1.2\ 10^{-15}$
CLM	Colombians from Medellin, Colombia	Admixed American	94	$8.9\ 10^{-15}$
PEL	Peruvians from Lima, Peru	Admixed American	85	0
GIH	Gujarati Indian from Houston, Texas	South Asian	103	0
PJL	Punjabi from Lahore, Pakistan	South Asian	96	0
BEB	Bengali from Bangladesh	South Asian	86	0
STU	Sri Lankan Tamil from the UK	South Asian	102	0
ITU	Indian Telugu from the UK	South Asian	102	0

The frequency of the homozygous DUT haplotype shared between Patients 1 and 3 was estimated on the 1000 Genomes data (Phase 3; ref (6)) using PLINK software. In order to estimate non-null frequencies, the haplotype was broken in independent consecutive segments, and the overall frequencies were obtained in each population group by multiplying the consecutive sub-haplotypes' frequencies. N: number of subjects in each population group. The shared DUT haplotype was observed only in population groups with European ancestry (European of admixed American populations).

Supplementary Table 10. Predicted consequences of DUT mutation according to prediction programs

DUT isoform	Protein	Provean	SIFT	Polyphen-2[a]	MutationTaster
DUT-M	p.Tyr142Cys	Deleterious	Damaging	Possibly damaging	Disease causing
DUT-N	p.Tyr54Cys	Deleterious	Tolerated	Possibly damaging	Disease causing

In silico prediction of the impact and severity of mutation on protein function was performed using Polyphen-2 (7), SIFT (8), Provean (8) and MutationTaster (9), using recommended parameters. using recommended parameters. [a]Polyphen-2 predictions were made based on the HumDiv model.

Supplementary Table 11. Predicted consequences of the Tyr54Cys (DUT-N)/ Tyr142Cys (DUT-M) mutation on intraprotein interactions

Type of interaction	Tyr54/Cys54 interactions lost in the mutated compared to the WT model
Hydrophobic	Ala53, Tyr56
Hydrogen Bonds	Asp69
Aromatic-Aromatic	Tyr56
Cation-Pi	Lys67

DUT residues are given with respect to the DUT-N isoform (see Figure 1C legend). Interactions with the Tyr54/Cys54 residue were estimated on the trimeric DUT-N WT (Tyr54, PDB: 3EHW) and the mutated (Cys54, modeled) structures as described (online methods). The loss of hydrophobic and aromatic-aromatic interaction is a direct consequence of the nature of the amino acid change. There was no novel interaction identified with Cys54 residue compared to Tyr54 under these models.

REFERENCES

1. Gautier J-F, Mourier A, de Kerviler E, Tarentola A, Bigard AX, Villette JM, Guezennec CY, Cathelineau G: Evaluation of abdominal fat distribution in noninsulin-dependent diabetes mellitus: relationship to insulin resistance. *The Journal of Clinical Endocrinology and Metabolism* 83:1306-1311, 1998

2. Soulier J: Fanconi anemia. *Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program* 2011:492-497, 2011

3. Varga B, Barabás O, Kovári J, Tóth J, Hunyadi-Gulyás É, Klement É, Medzihradszky KF, Tölgyesi F, Fidy J, Vértessy BG: Active site closure facilitates juxtaposition of reactant atoms for initiation of catalysis by human dUTPase. *FEBS letters* 581:4783-4788, 2007

4. Meur G, Simon A, Harun N, Virally M, Dechaume A, Bonnefond A, Fetita S, Tarasov AI, Guillausseau PJ, Boesgaard TW, Pedersen O, Hansen T, Polak M, Gautier J-F, Froguel P, Rutter GA, Vaxillaire M: Insulin Gene Mutations Resulting in Early-Onset Diabetes: Marked Differences in Clinical Presentation, Metabolic Status, and Pathogenic Effect Through Endoplasmic Reticulum Retention. *Diabetes* 59:653-661, 2010

5. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won H-H, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, Mcpherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J,

Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG: Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536:285-291, 2016

6. Sudmant PH, Rausch T, Gardner EJ, Handsaker RE, Abyzov A, Huddleston J, Zhang Y, Ye K, Jun G, Hsi-Yang Fritz M, Konkel MK, Malhotra A, Stütz AM, Shi X, Paolo Casale F, Chen J, Hormozdiari F, Dayama G, Chen K, Malig M, Chaisson MJP, Walter K, Meiers S, Kashin S, Garrison E, Auton A, Lam HYK, Jasmine Mu X, Alkan C, Antaki D, Bae T, Cerveira E, Chines P, Chong Z, Clarke L, Dal E, Ding L, Emery S, Fan X, Gujral M, Kahveci F, Kidd JM, Kong Y, Lameijer E-W, McCarthy S, Flicek P, Gibbs RA, Marth G, Mason CE, Menelaou A, Muzny DM, Nelson BJ, Noor A, Parrish NF, Pendleton M, Quitadamo A, Raeder B, Schadt EE, Romanovitch M, Schlattl A, Sebra R, Shabalin AA, Untergasser A, Walker JA, Wang M, Yu F, Zhang C, Zhang J, Zheng-Bradley X, Zhou W, Zichner T, Sebat J, Batzer MA, Mccarroll SA, Mills RE, Gerstein MB, Bashir A, Stegle O, Devine SE, Lee C, Eichler EE, Korbel JO: An integrated map of structural variation in 2,504 human genomes. *Nature* 526:75-81, 2015

7. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR: A method and server for predicting damaging missense mutations. *Nat Methods* 7:248-249, 2010

8. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP: Predicting the Functional Effect of Amino Acid Substitutions and Indels. *PLoS ONE* 7:e46688, 2012

9. Schwarz JM, Rödelsperger C, Schuelke M, Seelow D: MutationTaster evaluates disease-causing potential of sequence alterations. *Nature Methods* 7:575-576, 2010