



Derivation and molecular characterization of pancreatic differentiated MODY1-iPSCs

Carmel Braverman-Gross^a, Neta Nudel^a, Daniel Ronen^a, Nicola L. Beer^b, Mark I. McCarthy^b, Nissim Benvenisty^{a,*}

^a The Azrieli Center for Stem Cells and Genetic Research, Department of Genetics, Silberman Institute of Life Sciences, The Hebrew University, Jerusalem 91904, Israel

^b Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, United Kingdom

ARTICLE INFO

Keywords:

MODY1
HNF4α
iPS cells
Differentiation
Transcription factor
Binding sites

ABSTRACT

Maturity onset diabetes of the young (MODY) is a hereditary form of diabetes mellitus presenting at childhood or adolescence, which eventually leads to pancreatic β-cells dysfunction. The underlying genetic basis of MODY disorders is haploinsufficiency, where loss-of-function mutations in a single allele cause the diabetic phenotype in heterozygous patients. MODY1 is a type of MODY disorder resulting from a mutation in the transcription factor hepatocyte nuclear factor 4 alpha (HNF4α). In order to establish a human based model to study MODY1, we generated patient-derived induced pluripotent stem cells (iPSCs). Differentiation of these pluripotent cells towards the pancreatic lineage enabled to evaluate the effects of the MODY1 mutation and its impact on endodermal and pancreatic cells. Analyzing the gene expression profiles of differentiated MODY1 cells, revealed the outcome of HNF4α haploinsufficiency on its targets. This molecular analysis suggests that the differential expression of HNF4α target genes in MODY1 is affected by the number of HNF4α binding sites, their distance from the transcription start site, and the number of other transcription factor binding sites. These features may help explain the molecular manifestations of haploinsufficiency in MODY1 disease.

1. Introduction

Maturity-onset diabetes of the young (MODY), the most common form of monogenic diabetes (approximately 2–5% of all diabetic cases), is a group of clinically heterogeneous metabolic disorders that are characterized by impairment of pancreatic β cell function (Henzen, 2012; Siddiqui et al., 2015). The typical features of MODY include an autosomal dominant inheritance pattern, early manifestation usually during childhood or adolescence (before 25 years of age), non-obesity, and hyperglycemia with absence of pancreatic autoantibodies (Hattersley, 1998; Henzen, 2012; Nyunt et al., 2009; Siddiqui et al., 2015; Thanabalasingham and Owen, 2011). To date, fourteen distinct subtypes of MODY have been reported; all caused by mutations in genes involved in pancreatic β cell development, regulation and function (Yang and Chan, 2017). These gene products include mostly transcription factors such as HNF4α, HNF1α, PDX1, HNF1β, NEUROD1, KLF11 and PAX4, causing MODY1, 3, 4, 5, 6, 7 and 9, respectively. In addition, MODY2, 8 and 11 are caused by mutations in the GCK, CEL and BLK enzymes, MODY10 by a mutation in the insulin hormone and MODY12, 13 and 14 are caused by aberrations in the *SUR1*, *KCNJ11* and *APPL1* genes respectively (Bonnefond et al., 2012; Prudente et al.,

2015; Servitja and Ferrer, 2004; Stride and Hattersley, 2002).

All types of MODY are characterized by a heterozygous mutation and many of them are considered haploinsufficiency disorders. Haploinsufficiency is a phenomenon by which loss of function mutation in one allele results in insufficient dosage of the gene product, leading to failure in displaying the normal cellular phenotype. Haploinsufficiency disorders are relatively rare, since for most heterozygous mutations, the remaining unmutated allele is sufficient to preserve the function of the gene and present a normal phenotype. Genes characterized by haploinsufficiency tend to have conserved coding regions and promoter sequences and are usually expressed during early development in a tissue specific manner. (Huang et al., 2010). Haploinsufficiency genes are involved in processes related to development, cell cycle, nucleic acid metabolism and transcription (Dang et al., 2008).

MODY1, that accounts for approximately 5–10% of MODY cases, was first identified in a family known as the RW family, with a 70 years old proband who was diagnosed with diabetes at the age of 41 (Fajans et al., 1978). Further examination of this family revealed young members with abnormal glucose tolerance and hyperglycemia, leading to the conception of the term “MODY” (Fajans and Bell, 2011; Fajans

* Corresponding author at: The Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel.
E-mail address: nissimb@mail.huji.ac.il (N. Benvenisty).

et al., 1994).

MODY1 patients suffer from impairments in β cell function with normal insulin sensitivity (Fajans et al., 2001; Winter, 2003). Besides their β -cell associated phenotypes, MODY1 patients show impairments in glucagon and pancreatic polypeptide secretion, and decreased triglyceride and apolipoprotein biosynthesis (Fajans et al., 2001; Herman et al., 1997; Ilag et al., 2000; Lehto et al., 1999; Shih et al., 2000). Although *HNF4A* mutations in adults are associated with reduced insulin secretion, neonates sometimes present with macrosomia and hypoglycemia, suggesting an increased insulin secretion in utero (Fajans and Bell, 2007; Flanagan et al., 2010; McDonald and Ellard, 2013; Pearson et al., 2007). Treatment of MODY1 patients often aims in raising insulin levels and includes hypoglycemic drugs, mainly sulphonylurea and in more severe cases insulin (Fajans and Bell, 2011; Gardner and Tai, 2012).

The *HNF4A* gene is located on chromosome 20q13. It is alternatively spliced to create 12 different isoforms that are transcribed from one of two promoters (Colclough et al., 2013; Furuta et al., 1997; Huang et al., 2009c). The HNF4 α protein is a member of the nuclear hormone receptor family. It consists of a DNA binding domain, a dimerization domain (HNF4 α binds the DNA as a homodimer) and a transactivation domain (Hadzopoulou-Cladaras et al., 1997; Jiang and Sladek, 1997). HNF4 α is expressed mainly in the liver, gut, kidney and pancreas (Taraviras et al., 1994). It functions as a master regulator of hepatic gene transcription and is required for hepatic functions such as glycogen accumulation and control of lipid metabolism. It regulates pancreatic islet function and is required for insulin secretion and glucose metabolism (Gupta and Kaestner, 2004; Odom et al., 2004). Recent high-throughput transcriptomic analyses revealed that the number and variety of HNF4 α targets are much broader than previously postulated, suggesting that HNF4 α regulates over 1000 genes in hepatocytes and pancreatic islets (Bolotin et al., 2010; Odom et al., 2004, 2006; Rada-Iglesias et al., 2005).

Previous attempts to generate rodent models carrying mutations in MODY genes do not always recapitulate the human phenotype, possibly due to differences in human and rodent pancreas development (Piper et al., 2004; Teo et al., 2013a). Furthermore, human and rodent pancreatic islets have different architecture, insulin secretion properties and response to injury (Dai et al., 2012). This is mostly prominent in the case of MODY1, where homozygous knockout of *Hnf4a* causes embryonic lethality, while heterozygote mice have normal glucose tolerance and do not show any diabetic features (Shih et al., 2002; Stoffel and Duncan, 1997).

An alternative to mouse models are models generated using human induced pluripotent stem cells (iPSCs) (Avior et al., 2016; Teo et al., 2013a). Reprogramming of patient derived somatic cells allows the generation of pluripotent stem cell lines from patients with a known clinical phenotype. Furthermore, the development of protocols for direct differentiation of iPSCs to the pancreatic lineage enabled the derivation of patient-derived pancreatic progenitors carrying disease causing mutations (Melton, 2016; Lu et al., 2016; Madsen and Serup, 2006). iPSCs were derived from different MODY patients (Teo et al., 2013b; Wang et al., 2016), and in some of these cases, iPSCs were further differentiated towards the pancreatic lineage (Stepniwski et al., 2015; Teo et al., 2016; Yabe et al., 2015; Vethe et al., 2017).

Here, we demonstrate both the derivation of MODY1 iPSCs as well as their differentiation into the pancreatic lineage, alongside additional analyses of the affected targets of the *HNF4A* gene. We also suggest a model for the haploinsufficiency mode-of-action of HNF4 α .

2. Materials and methods

2.1. Cell culture

Fibroblasts of MODY1 patients were received from NIGMS Human Genetic Cell Repository at the Coriell Institute for Medical Research.

Two lines of fibroblast, GM01430 and GM01122, were derived from male and female members of the RW family, diagnosed with MODY1 at the ages of 27 and 16, respectively. The fibroblasts were maintained in medium containing Minimum Essential Medium Eagle M5650 (Sigma) supplemented with 15% fetal bovine serum (Sigma), 2 mM L-glutamine (Biological Industries) and Pen-Strep: 50 units/ml of penicillin and 50 μ g/ml streptomycin (Biological Industries). CSE7, IPSO and MODY1-iPS cells were cultured on mouse embryonic fibroblast (MEF) feeder layer, in ESC medium consists of Knockout Dulbecco's Modified Eagle's Medium (Life Technologies), containing 15% Knockout SR (Invitrogen), 2 mM L-glutamine, 0.1 mM Non-essential amino acids, 50 units/ml of penicillin and 50 μ g/ml streptomycin (Pen-Strep) (Biological Industries), 0.1 mM β -mercaptoethanol (Sigma) and 8 ng/ml basic fibroblast growth factor (R&D Systems). Cells were passaged using trypsin solution A (Biological Industries) and maintained in a humidified incubator at 37 °C and 5% CO₂.

2.2. Reprogramming of MODY1 fibroblast to MODY1 iPSCs

293T cells were maintained in Dulbecco's Modified Eagle's Medium - high glucose (Sigma) and transfected using PEI (Sigma) with 4.5 μ g PMXs retroviral vectors containing human OCT4, SOX2, KLF4 or C-MYC and 4.5 μ g of PCL-ampho plasmid. 24 and 48 h post-transfection, the supernatant was collected, filtered through a 0.45- μ m cellulose acetate filter (Whatman) and supplemented with 4 μ g/ml polybrene (Sigma). 1.2×10^6 fibroblasts (GM01430 and GM01122) were infected twice (every 24 h) with media containing the viruses. 4 days post the first infection cells were seeded on mitomycin treated MEF feeder plates with ESC medium. Colonies with iPSC morphology appeared approximately 21 days post-infection and thereafter picked and cultured. Eight iPSC lines were characterized, one from GM01122 fibroblasts (MODY1-iPS-A-3), and seven from GM01430 fibroblast (MODY1-iPS-B-1, MODY1-iPS-B-5, MODY1-iPS-B-6, MODY1-iPS-B-7, MODY1-iPS-B-8, MODY1-iPS-B-9, MODY1-iPS-B-13).

2.3. Differentiation of iPSCs towards pancreatic lineage

MODY1-iPS-A-3, MODY1-iPS-B-6, CSE7 and IPSO Cells were maintained on matrigel (BD Biosciences) plates in mTeSR1 media (Stemcell Technologies) for 3 days prior to differentiation. We conducted differentiation to two stages of pancreatic differentiation: primitive gut tube (PGT) and pancreatic progenitors (PP). A schematic diagram of the differentiation protocols is shown in Supplementary Fig. 1. During differentiation, media was changed daily: in the priming stage (1 day), cells were cultured in RPMI1640 media (Life Technologies) supplemented with 0.1% Pen-Strep (Biological Industries) and 3 μ M CHIR99021/ GSK-3 Inhibitor XVI (Calbiochem). On the first day of definitive endoderm stage (1 day), cells were cultured in RPMI1640 media (Life Technologies) supplemented with 0.1% Pen-Strep (Biological Industries) and 100 ng/ml Activin A (Peprotech). On the next days of definitive endoderm stage (2 days in the PGT differentiation and 3 days in PP differentiation), cells were cultured in RPMI1640 media (Life Technologies) supplemented with 0.1% Pen-Strep (Biological Industries), 2% B27 (Life Technologies) and 100 ng/ml Activin A (Peprotech). At the primitive gut tube stage (3 days in PGT and 2 days in PP differentiation), cells were treated with Dulbecco's Modified Eagle's Medium/Nutrient F-12 Ham (Sigma) supplemented with 0.1% Pen-Strep (Biological Industries), 2% fetal bovine serum (Sigma) and 50 ng/ml recombinant human KGF (Peprotech). At the posterior foregut stage (4 days in PP differentiation), cells were treated with Dulbecco's Modified Eagle's Medium - high glucose (Sigma) supplemented with 1% B27 (Life Technologies), 2 mM L-glutamine (Biological Industries), 0.1% Pen-Strep (Biological Industries), 2 μ M retinoic acid (Sigma), 0.25 μ M sant-1 (Sigma) and 100 ng/ml noggin (Peprotech). At the pancreatic progenitors stage (4 days in PP differentiation), cells were treated with Dulbecco's Modified Eagle's Medium

- high glucose (Sigma) supplemented with 1% B27 (Life Technologies), 2 mM L-glutamine (Biological Industries), 0.1% Pen-Strep (Biological Industries), 100 ng/ml noggin (Peprotech), 1 μ M Alk5 inhibitor II / TGF- β RI kinase inhibitor II (Calbiochem) and 50 nM TPB/ α -Amyloid Precursor Protein Modulator (Calbiochem).

2.4. Alkaline phosphatase staining and immunostaining

MODY1-iPS-A-3, MODY1-iPS-B-1, MODY1-iPS-B-7, MODY1-iPS-B-8 and MODY1-iPS-B-9 cells were stained using Leukocyte Alkaline Phosphatase kit (Sigma), according to the manufacturer's instructions. For immunostaining, cells were washed twice with PBS, fixed with 4% paraformaldehyde solution for 10 min, washed 3 times with PBS and blocked for 1 h with PBS containing 2% bovine serum albumin (BSA), and 0.1% Triton X-100 (Sigma). Cells were incubated with primary antibodies diluted in blocking solution for 1 h, washed 3 times in PBS and incubated with secondary antibody diluted in blocking solution for 1 h. Primary antibodies: rabbit anti human OCT3/4 (1:200, Santa Cruz) was used for staining of MODY1-iPS-A-3, MODY1-iPS-B-1, MODY1-iPS-B-5, MODY1-iPS-B-6 and MODY1-iPS-B-13 cells. Mouse anti human TRA-1-60 (1:200, Santa Cruz) was used for staining of MODY1-iPS-A-3, MODY1-iPS-B-1, MODY1-iPS-B-6 and MODY1-iPS-B-13 cells. Goat anti human HNF4 α (1:200, Santa Cruz) was used for staining MODY1-iPS-A-3, MODY1-iPS-B-6 and CSES7 PGT differentiated cells. Rabbit anti human PDX1 (1:100, Sigma) was used for staining MODY1-iPS-A-3, MODY1-iPS-B-6 and CSES7 PP differentiated cells. Secondary antibodies: donkey anti rabbit cy2 (1:100, Jackson ImmunoResearch), goat anti mouse cy3 (1:100, Jackson ImmunoResearch), donkey anti goat cy3 (1:100, Jackson ImmunoResearch) and donkey anti rabbit cy3 (1:100, Jackson ImmunoResearch). Quantification of PDX1 expression in MODY1-iPS-B-6 and CSES7 PP differentiated cells was done using ImageJ software (Abramoff et al., 2004).

2.5. Teratoma formation

MODY1 undifferentiated iPSCs were trypsinized, harvested and resuspended in a 200 μ l matrigel. The cells were injected subcutaneously into NOD-SCID *IL2rg*^{−/−} immunodeficient mice (Jackson Laboratory). About 8 weeks after injection teratomas were dissected and cryosections were prepared for hematoxylin and eosin staining. All experimental procedures in animals were approved by the ethics committee of the Hebrew University.

2.6. Karyotype analysis

Cells were incubated in 100 ng/ml colcemid (Biological Industries) solution for 40 min in 37 °C, washed with PBS, trypsinized, centrifuged, treated with hypotonic solution for 20 min and fixed. Metaphases were stained with Giemsa dyes (Sigma).

2.7. RNA extraction and high throughput-RNA sequencing

RNA was extracted using PerfectPure RNA Cultured Cell kit (5-Prime) kit and NucleoSpin RNA (Machery Nagel) kit, according to the manufacturer's instructions. Paired-end RNA sequencing was conducted using illumina HiSeq2000.

2.8. RNA sequencing analysis

FASTQC (Andrews, 2010) was used to evaluate RNA reads' quality and TRIMOMATIC (Bolger et al., 2014) removal of outliers was performed. Alignment to the GRCh37 human genome was performed using TopHat and initial RPKM (reads per kilobase per million mapped reads) values were generated using cufflinks (Anders et al., 2014). The SeqMonk tool (<http://www.bioinformatics.babraham.ac.uk/projects/seqmonk>) was used for normalization, batch effect removal (caused

due to three independent RNA sequencing analyses), final RPKM values calculation and also for performing differential expression tests. Promoter sequences were defined as (−2000) to (+200) bp from the TSS. Sequences were obtained from ENSEMBLE BioMart (Kinsella et al., 2011), then subjected to FIMO (Grant et al., 2011) in order to identify HNF4 α targets using HNF4 α PFM MA0114.2 from Jaspar, (Mathelier et al., 2016). Consensus of this PFM was obtained by using RSAT (Medina-Rivera et al., 2015). Targets were selected based on *P* value < 1×10^{-4} . Putative transcription factor regulating HNF4 α targets were obtained using PRIMA (Sharan et al., 2003). Following transcription factors identification, number of motifs per gene in the unchanged gene list were normalized to the number of downregulated genes in order to match the difference in gene numbers.

2.9. Flow cytometry

Pluripotent and PGT differentiated MODY1-iPS-B-6 and IPSO cells were dissociated using TrypLE Select (Life technologies) into single cell, fixed with 90% ice cold methanol for 20 min on ice. Samples were washed twice with 1% BSA-PBS, permeabilized for 20 min with 0.1% Tween20- PBS, blocked for 10 min with 10% FBS / 0.3 M Glycin/ PBS. Cells were incubated with HNF4 α primary antibody (Abcam K9218, 2 μ g/ 1×10^6 cells) for 1 h diluted in 10% FBS/0.1% Tween20/PBS, washed, and incubated with secondary antibody (donkey anti mouse IgG H&L Alexa Fluor 488, ab 150,109) at 1/500 dilution for 1 h. Cells were filtered through a 70 μ m cell strainer and analyzed by flow cytometry (BD Biosciences FACSARIA III) and Flowjo software (FlowJo LLC).

3. Results and discussion

3.1. Reprogramming of MODY1 cells

To establish a human based MODY1 model we utilized fibroblast cells from two MODY1 patients that belong to the RW family (Fajans et al., 1994). These patients harbor a nonsense mutation in the *HNF4A* gene, caused by a C to T substitution at amino acid 268 (Q268X) in exon 7, affecting all *HNF4A* transcripts and impairs the dimerization and transactivation domains of HNF4 α protein. While this mutation leads to loss-of-function, there was no dominant negative effect on the normal HNF4 α allele product, verifying that the MODY1 phenotype is caused due to the reduction in the levels of HNF4 α (Ellard and Colclough, 2006; Stoffel and Duncan, 1997; Yamagata et al., 1996).

The fibroblasts were reprogrammed by retroviral infection constructs containing *OCT4*, *c-MYC*, *SOX2* and *KLF4* genes, to generate eight MODY1-iPSC lines. These cell lines were positive for pluripotency markers including alkaline phosphatase, TRA-1-60 and OCT4 (Fig. 1A–B). In addition, teratoma formation assay confirmed that the cells are pluripotent (Fig. 1C). Chromosomal integrity of the cells was confirmed by karyotyping (Fig. 1D).

3.2. Differentiation of MODY1 iPSCs to pancreatic cells

Two MODY1 iPSC lines (MODY1-iPS-A-3 and MODY1-iPS-B-6, generated from different patients) as well as wild-type CSES7 and IPSO were differentiated towards the pancreatic lineage in two stages: (1) primitive gut tube (PGT) - the stage in which HNF4 α is expressed (Fig. 2A), and (2) the successive stage of pancreatic progenitors (PP) - defined by the expression of PDX1 (Fig. 2B). The protocols used for differentiation were based on a previously published protocol for human pancreatic progenitor differentiation (Rezania et al., 2012) (see Materials and Methods, and (Supplementary Fig. 1)). Differentiation efficiency was assessed by FACS analysis of HNF4 α expression and immunostaining for PDX1, indicating > 80% of cells differentiated towards PGT cells (Supplementary Fig. 3) and about 60% of cells differentiated to PP cells in both control and MODY1 iPSCs. We did not detect

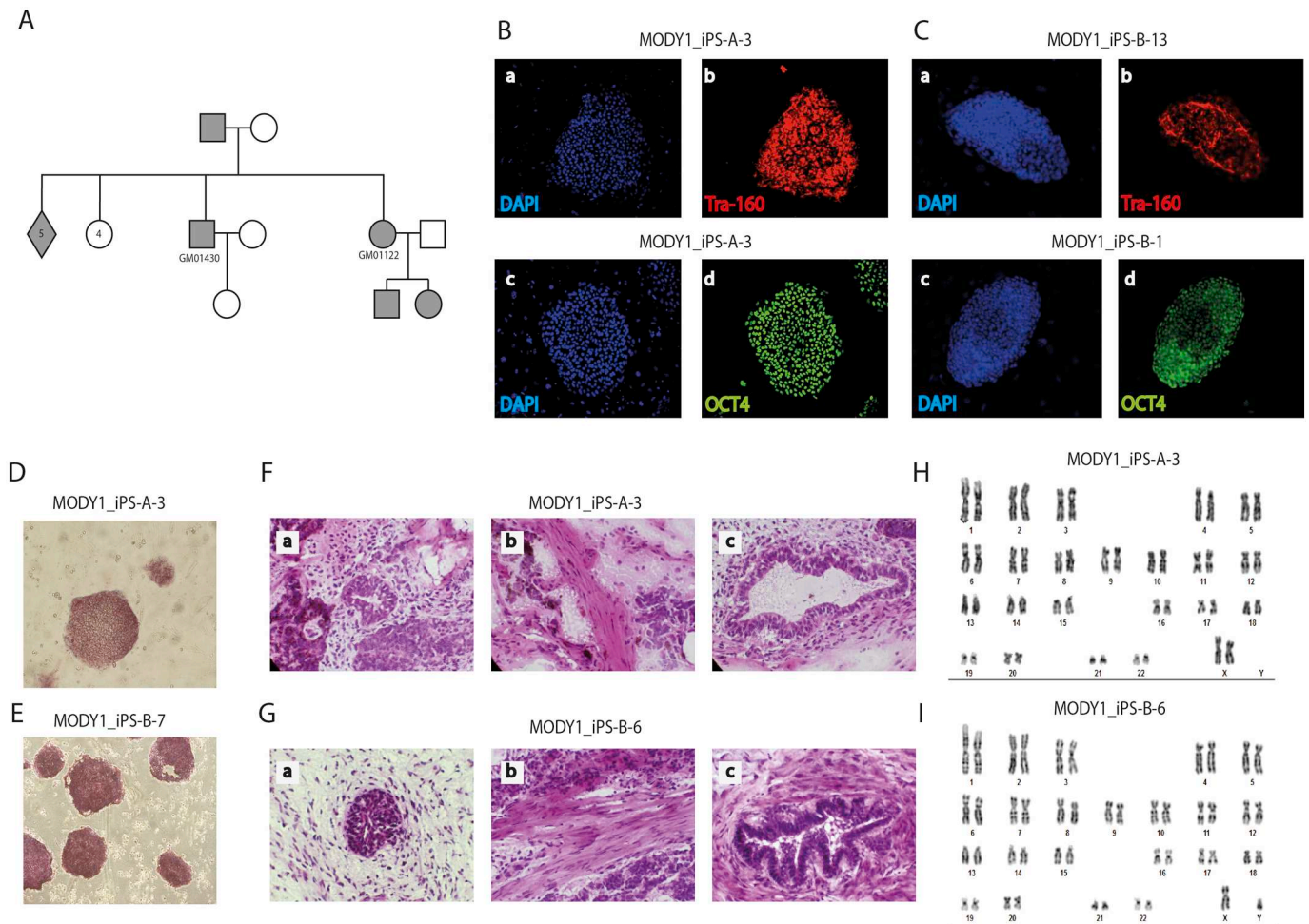


Fig. 1. Characterization of MODY1 iPSCs.

A: RW family pedigree (modified from [Fajans & Bell, 2011](#); [Yamagata et al., 1996](#)). Squares denote males, circles denote females, and diamond shape denotes unspecified gender. Grey symbols denote patients with MODY1. Numbers within symbols denote number of healthy/MODY1 patient siblings. Repository numbers under symbols denote the fibroblasts line stored at the Coriell Institute for Medical Research used for reprogramming. B and C: Immunostaining of a representative MODY1 iPSC clones derived from GM01122 and GM01430 patients, respectively. a, c: nuclear staining (blue); b: staining for the pluripotent marker TRA-1-60 (red); d: staining for the pluripotent marker OCT4 (green). D and E: Alkaline phosphatase staining of representative MODY1 iPSC clones derived from GM01122 and GM01430 patients, respectively. F and G: Sections of teratomas generated from MODY1 iPSCs from GM01122 (MODY1-iPS-A-3) and GM01430 (MODY1-iPS-B-6) patients, respectively, presenting derivatives of all three embryonic germ layers; ectoderm (a), mesoderm (b) and endoderm (c). H and I: Karyotype analysis of MODY1 iPSCs from GM01122 (MODY1-iPS-A-3) and GM01430 (MODY1-iPS-B-6) patients, respectively, indicating normal 46,XY and 46,XX karyotype.

any significant difference between MODY1 and control cells differentiation capacity, suggesting that HNF4 α mutation in one allele has no significant effect on the ability of the iPSCs to differentiate in vitro, as was previously suggested ([Vethe et al., 2017](#)).

The differentiated MODY1 and control cells were analyzed by high-throughput RNA-sequencing. Hierarchical clustering and principal component analysis (PCA) of average gene expression values from the different procedures exhibited strong similarities between cells from the same differentiation stage, indicating that the difference in global gene expression between MODY1 and control cells is smaller than the difference between differentiation stages ([Fig. 3](#) and Supplementary Fig. 2).

3.3. Characterizing MODY1 differentiated cells

In order to verify the efficiency of pancreatic lineage differentiation, we analyzed the average levels of expression of cell type specific markers in MODY1 and control cells ([Fig. 4A](#)). For the pluripotent stage, we analyzed the levels of OCT4, DNMT3B, DPPA4 and NANOG. All these markers were highly expressed in the undifferentiated cells and were significantly downregulated at the PGT and PP cell stages ([Fig. 4B](#)). For

the following stage, we analyzed the expression of the definitive endoderm markers CXCR4, SOX17, FOXA2 and GATA4, and the primitive gut tube markers HNF4A and HNF1B. These markers were markedly increased in the PGT cell stage and downregulated in the PP cell stage ([Fig. 4C](#)), in accordance with their expected developmental stage-dependent expression. For the PP cell stage, we analyzed the levels of the posterior foregut markers PDX1, PAX6, PAX4, ONECUT1 and SOX9, and the pancreatic endoderm markers NKX2-2, NEUROD1, NEUROG3 and NKX6-1. As expected, these markers showed low levels of expression in the pluripotent and PGT stages and were dramatically upregulated in the PP cell stage.

Interestingly, we found that some of the transcription factors, which serve as markers for the PP cell stage, were expressed at higher levels in the MODY1 cells compared to normal cells ([Fig. 4D](#)). Our data shows that the levels of MODY1 PP markers have increased in a limited significance (for example, the significance *p* values of increase in PAX6, NEUROD1, NEUROG3 and NKX6-1 expression in MODY1 vs. control PP cells were 0.08, 0.18, 0.12 and 0.15, respectively). However, it may indicate a compensatory mechanism utilized by MODY1 cells to overcome the decrease in HNF4 α levels. A similar mechanism was reported in another type of MODY disease, MODY5, where upregulation of

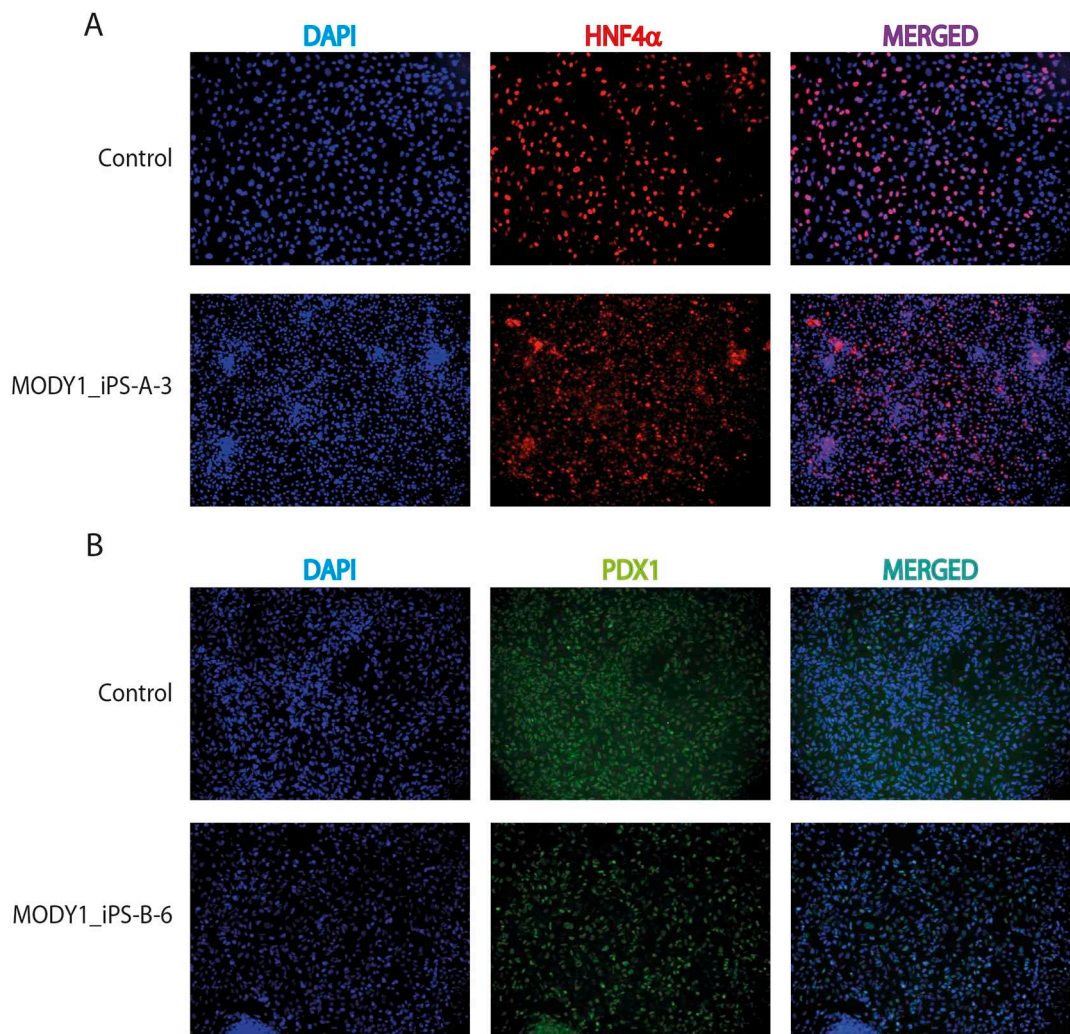


Fig. 2. Expression of primitive gut tube and pancreatic progenitors markers in differentiated iPSCs. A: Immunostaining for HNF4α in control cells (CSES7) and MODY1 (MODY1-iPS-A-3) cells differentiated into primitive gut tube cells. B: Immunostaining for PDX1 in control (CSES7) cells and MODY1 (MODY1-iPS-B-6) cells differentiated to pancreatic progenitor cells.

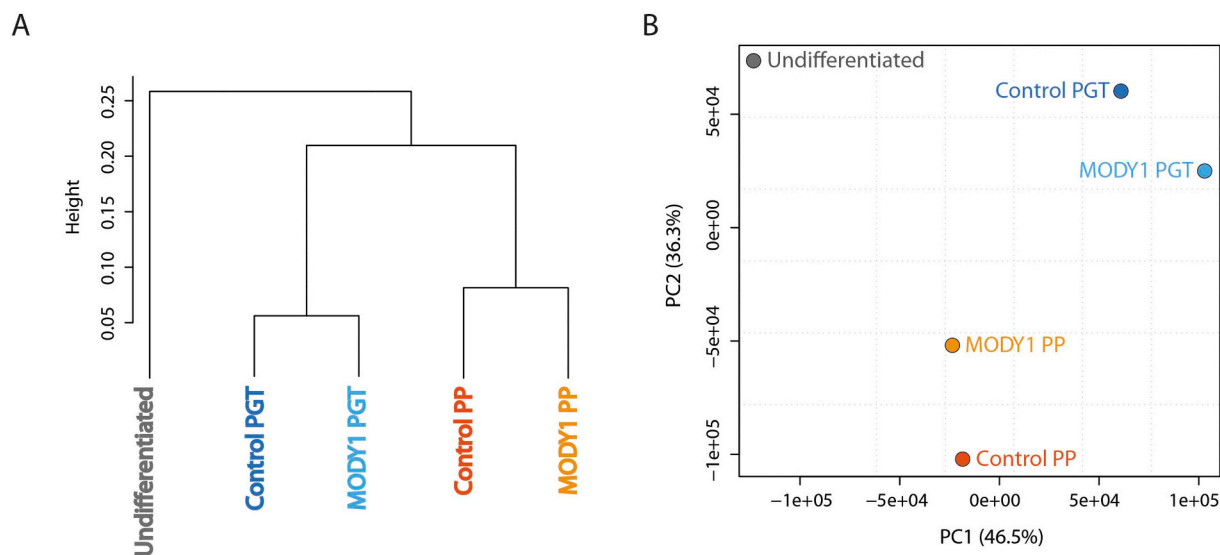
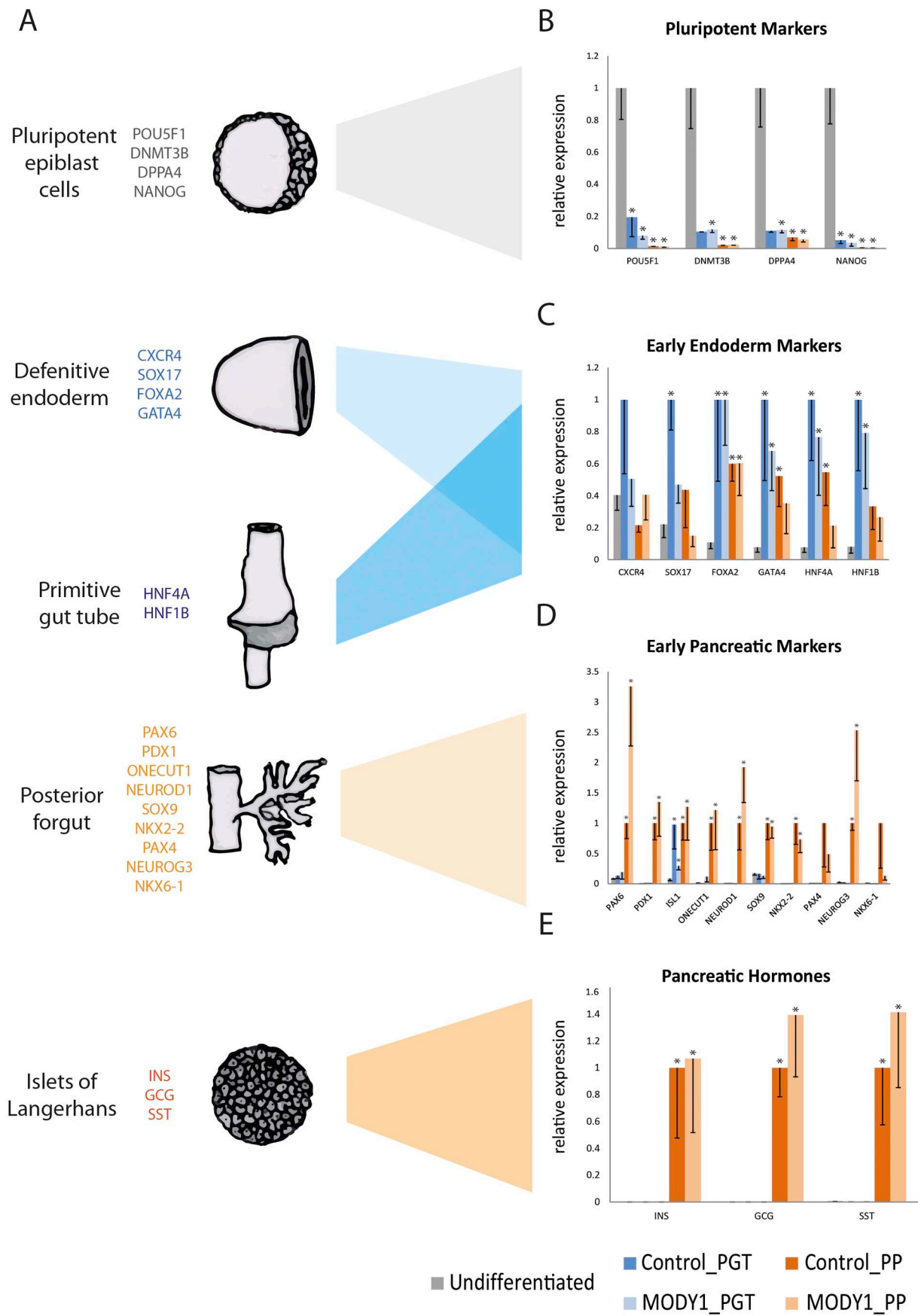


Fig. 3. Gene expression levels in MODY1 and control cells reveal clustering according to differentiation state. A: Hierarchical clustering and B: Principle component analysis (PCA) of high-throughput RNA sequencing gene expression data obtained and averaged from three independent differentiation repeats using control and MODY1 cells. The differentiation included three stages: undifferentiated, primitive gut tube (PGT), and pancreatic progenitors (PP) cells.



(caption on next page)

Fig. 4. Analysis of stage specific gene expression of MODY1 iPSCs differentiated towards pancreatic lineage cells.

A: Schematic presentation of key stages in pancreas development and representative genes expressed in each stage, modified from (Madsen and Serup, 2006). B-E: Expression of specific marker genes in control and MODY1 derived cells from all three differentiation stages: undifferentiated (average of control and MODY1 undifferentiated cells, denoted in grey), PGT (average of either control or MODY1 cells, in dark and light blue, respectively) and PP (average of either control or MODY1 cells, in dark and light orange, respectively). B: Relative expression of pluripotent markers C: Relative expression of early endodermal markers - corresponding to definitive endoderm and primitive gut tube stage, D: Relative expression of early pancreatic markers - corresponding to posterior foregut stage, and E: Expression of pancreatic hormones secreted from islet of Langerhans. Error bars represent SEM. * - P value < .05, by one-tailed student t -test between undifferentiated to PGT or to PP cells.

pancreatic development regulators was suggested to overcome the heterozygosity of *HNF1B* (Teo et al., 2016).

As the PP cell stage of differentiation may also lead to a subpopulation of pancreatic-hormone-expressing cells, we evaluated the expression of the insulin (*INS*), glucagon (*GCK*) and somatostatin (*SST*) genes. The expression of these hormones was highly elevated in the PP cell stage and absent in former stages of differentiation (Fig. 4E). Here again, the levels of the hormones were higher in the MODY1 cells compared to the control cells in a limited significance, which may relate to the hyperinsulinemia seen in MODY1 patients (Flanagan et al., 2010; Pearson et al., 2007) and to the increased insulin secretion in *HNF4α* mutated fetal EndoC-βH1 cells (Thomsen et al., 2016). This analysis confirmed the different stages of pancreatic lineage differentiation of both MODY1 and normal PSCs, assuring that *HNF4α* mutation does not prevent expression of pancreatic markers in heterozygous cells.

3.4. Analysis of MODY1 differentially expressed genes related to pancreatic differentiation

In order to explore the direct effects of *HNF4A* heterozygous mutation, we focused on the primitive gut tube stage of differentiation, in which *HNF4α* is highly expressed.

The differentiation towards PGT stage was robust and resulted in a relatively homogenous population (Supplementary Fig. 3), thus enabling the analysis of the transcriptome at this stage. As *HNF4α* acts primarily as a transcriptional activator (Odom et al., 2004), we looked for any downregulated genes in MODY1-iPSCs, aiming to reveal targets of *HNF4α* that are affected by its decrease. We compared the averaged RPKM values of normal (CSES7 and IPSO cells) versus MODY1 (MODY1-iPSC-A-3 and MODY1-iPSC-B-6) PGT cells, and looked for the genes that showed significant expression differences (P value < .05). To assure that this group of genes was relevant for the specific differentiation stage of PGT, and mainly affected by *HNF4α* haploinsufficiency, we only used genes which were upregulated by at least 1.5 folds in the control PGT compared with the control undifferentiated stages, and downregulated by at least 2 folds in the MODY1 differentiated cells compared with the control. By these means, we defined a specific list of early pancreatic genes that were downregulated in the MODY1-iPSCs (Supplementary Table 1).

Functional annotation analyses of the downregulated genes in MODY1 using DAVID Bioinformatics Resources (Huang et al., 2009a,b) and the Molecular Signature Database (MSigDB) (Subramanian et al., 2005) revealed significant enrichment for clusters including “apolipoproteins”, “triglyceride catabolic process”, “lipoprotein metabolic process”, “lipid metabolic process”, “hormone biosynthesis” and “secretion”. These clusters are known to be associated with established pancreatic and hepatic roles of *HNF4α* in insulin and glucagon secretion, lipoproteins, triglyceride biosynthesis and lipid metabolism (Fajans et al., 2001; Herman et al., 1997; Shih et al., 2000) and are in accordance with the dyslipidemia observed in MODY1 patients (Fajans et al., 2001) (Supplementary Table 2).

As a control, we generated a list of genes upregulated during PGT differentiation (expression ratio of 1.5 or higher between the PGT and the undifferentiated stages) that were similarly expressed in both WT and MODY1-derived cells (Supplementary Table 1). Molecular Signature Database (MSigDB) (Subramanian et al., 2005) analysis of

these genes revealed mostly generalized clusters, including signal transduction and development (Supplementary Table 2).

3.5. Identifying *HNF4α* putative targets

We used the FIMO software in order to define which of the genes in the downregulated and control (unchanged) lists may be direct targets of *HNF4α* regulation. We scanned the promoter region, defined as −2000 to +200 around the TSS of the genes from both lists and searched for appearance of *HNF4α* consensus (cTGmCTTTGmCyy), according to *HNF4α* position frequency matrices (PFM) from JASPAR database. More than half of the genes, both from the downregulated (55%) and the unchanged (59%) lists were predicted to harbor an *HNF4α* motif in their promoter, and were thus considered as *HNF4α* putative targets (Supplementary Table 3).

3.6. Molecular features characterizing *HNF4α* targets that alter their expression levels in MODY1 cells

3.6.1. Number of *HNF4α* binding sites affects *HNF4α* targets expression levels

Gene promoter regions often harbor more than one binding site for a specific transcription factor. The occurrence of such binding site clusters may influence the expression levels of a gene (Ezer et al., 2014). Several models have been proposed for explaining the relations between transcription factor binding sites at the promoter region and gene expression levels. One of these models corresponds to the “AND logic gate”, suggesting that all binding sites must be bound by the transcription factor in order to affect the gene. A second model corresponding to the “OR logic gate”, requires binding of a single transcription factor binding site (even if more than one site is found in the promoter) in order to affect gene expression. According to the latter model, if there are many binding sites for a transcription factor, it is more likely that the transcription factor will activate the target even if the transcription factor is expressed at low levels (Ezer et al., 2014). Thus, a “AND logic” transcriptional activator that is downregulated, is expected to cause downregulation in targets that harbor many binding sites, since the chances that the transcription factor will occupy all the binding sites is lower. Conversely, a “OR logic” transcriptional activator that is downregulated, is expected to lead for the downregulation of targets containing a single binding site, because the chances of binding a single site is decreased when the transcription factor levels are low. Because of the *HNF4α* haploinsufficiency in MODY1 cells, we speculated that there would be a difference in the number of binding sites for *HNF4α* in its targets that reflects the model that fits *HNF4α* activity. In order to test this, we examined the correlation between *HNF4α* depletion and the number of *HNF4α* binding sites in the promoters of its targets. A high number of binding sites for *HNF4α* in the downregulated genes in MODY1 iPSCs compared to the unchanged genes, would indicate an “AND logic” behavior of *HNF4α* activity, since targets that have many binding sites for *HNF4α* will be more affected by the decrease in its levels. The opposite scenario, where a higher number of *HNF4α* binding sites is detected in the unchanged genes in MODY1 cells compared with the downregulated genes, fits the “OR logic” model, where low levels of *HNF4α* will bind more efficiently to targets having many binding sites.

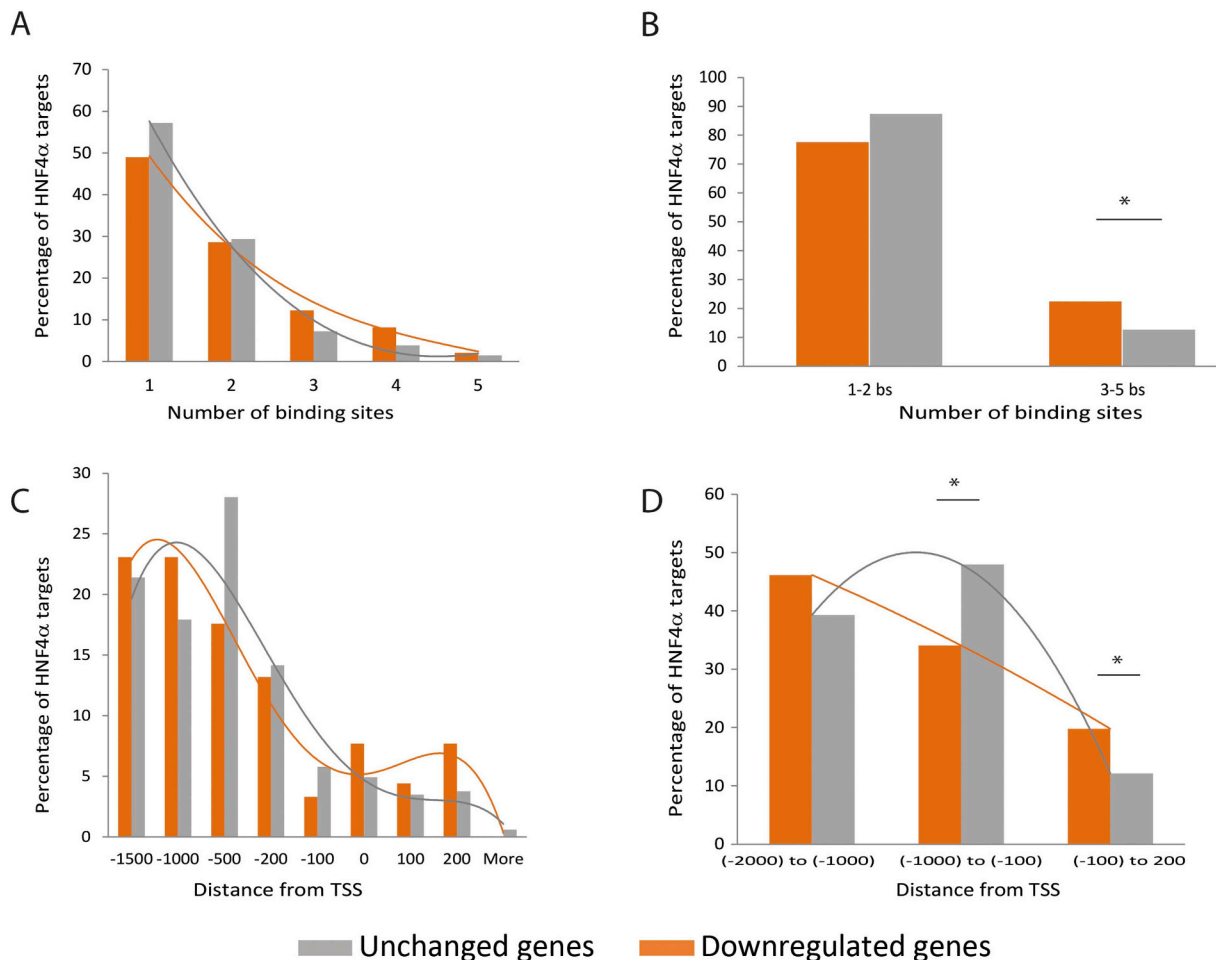


Fig. 5. Number and distribution of HNF4α binding sites in HNF4α targets.

A and B: Number of HNF4α binding sites in HNF4α target genes that are downregulated (orange) or unchanged (grey) in MODY1 cells, shown as detailed (A) or grouped to bins (B). C and D: Distribution of distances of HNF4α binding sites from the transcription start site (TSS) in HNF4α targets that are downregulated (orange) or unchanged (grey) in MODY1 cells, shown as detailed (C) or grouped to bins (D). * - P value < .05, by one-sided Z score test for two population proportion.

Our results showed that downregulated genes in MODY1 contain a higher number of HNF4α binding sites, compared to the set of unchanged genes (Fig. 5A). This difference was found to be significant (P value = .037, one-tail Z score test for two population proportion) among genes having 3–5 binding sites (Fig. 5B). This result suggests that genes that are relevant for pancreatic development and harbor multiple binding sites for HNF4α are more sensitive to the levels of HNF4α. Thus, in MODY1 cells, where HNF4α levels are decreased, genes having > 3 binding sites for HNF4α in their promoter will be more prone to downregulation, and genes with a single or a pair of binding sites will be less affected. This observation supports the “AND logic gate” model for HNF4α activity on its targets.

3.6.2. Distance of binding sites from TSS affects HNF4α targets expression levels

Various studies examined the correlation between gene expression and the distribution of transcription factor binding sites along the regulatory region of the target genes. These studies suggested that transcription factors that have different roles (such as activators vs. repressors), or that participate in different processes in a cell type or tissue specific manner, bind at different locations in the regulatory regions of their targets (Koudritsky and Domany, 2008; Tabach et al., 2007; Whitfield et al., 2012; Wu and Lai, 2015). Chip-Chip and Chip-seq analysis for HNF4α in human intestinal and liver cells revealed HNF4α binding sites enrichment in active genes around the TSS (Boyd et al., 2009; Wallerman et al., 2009). Analysis of HNF4α binding site

distribution in MODY1 cells revealed 3 ranges of distances from the transcription start site: –2000 to –1000; –1000 to –100 and –100 to +200 bp from the transcription start site (Fig. 5C). Comparison of the regulatory sites revealed a significant difference between downregulated and unchanged genes. Specifically, unchanged genes in MODY1 cells were enriched for HNF4α binding sites in the region of –1000 to –100 bp upstream to the TSS (P value = .01, one-tail Z score test for two population proportion), while downregulated genes in MODY1 cells were enriched for HNF4α binding sites in the region of –100 to 200 bp surrounding the transcription start site (P value = .03, one-tail Z score test for two population proportion) (Fig. 5D). This may indicate a higher sensitivity of HNF4α targets for its expression decrease if HNF4α binding sites are found in proximity to the transcription start site.

3.6.3. The number of alternative transcription factor binding sites affects the expression levels of HNF4α target

Gene regulatory regions contain specific sequences, or motifs, recognized by distinct transcription factors that regulate gene expression. Phenomena such as transcriptional synergy and transcription factor cooperation, require the existence of several motifs within a specific gene regulatory region (Carey, 1998; Todeschini et al., 2014). In our analysis, we searched for differences in the number of putative transcription factor motifs found in the promoter region of HNF4α target genes. Using PRIMA software (Sharan et al., 2003), we scanned for predicted transcription factor motifs in our downregulated and

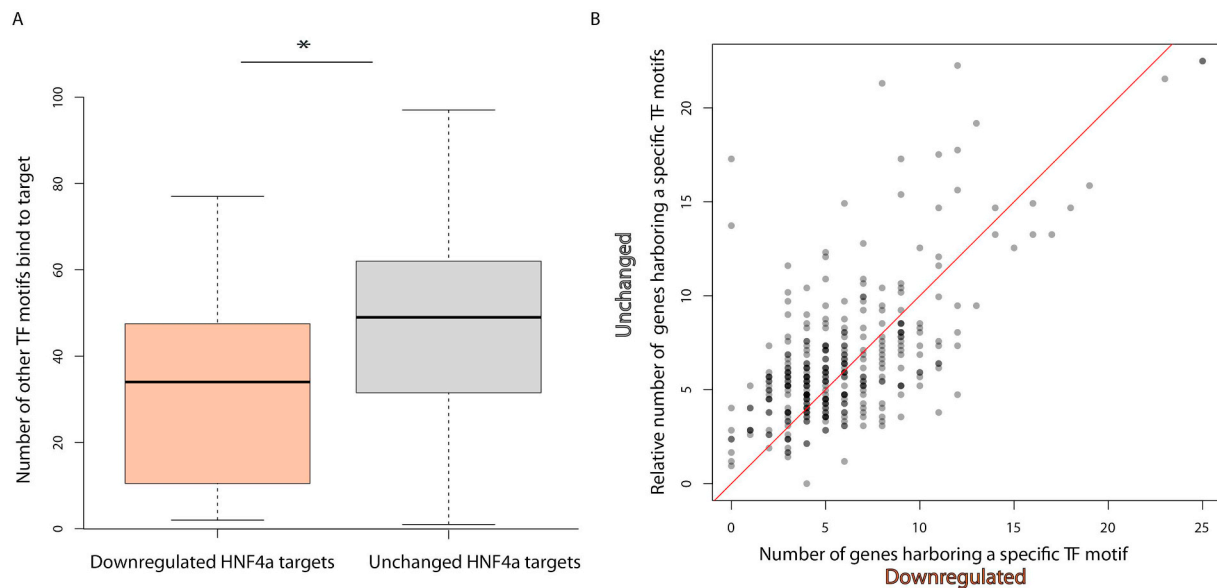


Fig. 6. Regulation of HNF4 α targets by other transcription factors.

A: Boxplot for number of different transcription factor motifs predicted by PRIMA to appear in the promoters of the two groups of HNF4 α targets. B: Number of downregulated and relative number of unchanged HNF4 α targets harboring each of the transcription factor motifs. * - P value < .05, by one-tailed student t-test.

unchanged HNF4 α targets (Supplementary Table 4). We examined the correlation between the number of other transcription factor motifs found in HNF4 α targets and the expression level of these targets in MODY1 cells.

As expected, PRIMA output showed that both downregulated and unchanged genes are predicted to harbor motifs for different transcription factors. However, there was a significant difference in the number of motifs for other transcription factors found in the genes of the two lists (P value = 3.9×10^{-5} , one-tailed two sample t-test), where HNF4 α targets that were unchanged in MODY1 cells harbor more transcription factor motifs compared to the downregulated genes (Fig. 6A). This observation supports the notion that some genes could be redundant for HNF4 α activity, and that other transcription factors may compensate for HNF4 α downregulation. Hence, if a variety of transcription factors binds to HNF4 α targets, they may compensate for the decrease in HNF4 α , and explain why the expression level of a subpopulation of target genes will not change in MODY1 cells.

Transcription factors predicted by PRIMA software to have binding sites at only targets from the unchanged group are: ETF, ZF5, STAT1, E2F, Alx-4 and GZF1, and the transcription factor Elk-1 have binding sites only at genes from the MODY1 downregulated group. Some of these transcription factors, such as ETF, E2F and ELK-1, were previously shown to have binding sites in HNF4 α target genes, mainly in intestinal and liver cells, (Boyd et al., 2009; Godoy et al., 2016; Wang et al., 2011) and their roles in pancreatic development are yet to be determined. The majority of transcription factor motifs which were found by PRIMA software were present in both the downregulated and unchanged genes in MODY1 cells. However, comparing the number of genes in either the unchanged or downregulated lists that are regulated by each specific motif, identified significant differences (P value = .0002, one-tailed paired t-test), where most motifs are predicted to be found in higher numbers in promoters of genes that are unchanged in MODY1 compared to downregulated genes (Fig. 6B). This observation emphasizes the possible redundancy effect of other transcription factors for the downregulation of HNF4 α .

4. Conclusions

In this study, we generated a model for MODY1 disease in human iPSCs. iPSC lines from different patients were characterized and differentiated towards pancreatic progenitor cells. MODY1 cells differentiated to the PP cell stage, exhibited an expression increase of stage specific transcription factors, as previously described in MODY5 (Teo et al., 2016). Expression of hormone genes was also elevated in MODY1 cells compared to control cells, which may relate to the hyperinsulinemia found in part of neonatal MODY1 patients. The upregulation of pancreatic transcription factors and pancreatic hormones in MODY1 cells may indicate the presence of a compensatory mechanism aimed to ensure pancreatic development and overcoming the HNF4A heterozygosity.

Since MODY1 is characterized by haploinsufficiency, this model may also be relevant for the research of haploinsufficiency mechanism of action. Our analyses suggested that HNF4 α targets (that contain HNF4 α binding motif) are more affected by the decrease in HNF4 α levels if: (1) they contain multiple binding sites for HNF4 α ; (2) the distribution of binding sites is in proximity to the transcription start site of the target gene, and (3) they harbor fewer motifs for other transcription factors in their promoters (Fig. 7). We suggest that these molecular features influence the sensitivity of target genes to the haploinsufficiency of HNF4 α which is the cause for MODY1, while the influence of these features in other haploinsufficiency disorders remains to be investigated.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scr.2018.06.013>.

Acknowledgements

We thank all members of The Azrieli Center for Stem Cells and Genetic Research for their help and support and for critical reading of the manuscript. N.B. is the Herbert Cohn Chair in Cancer Research. This work was partially supported by the Rosetrees Trust, and by the Azrieli

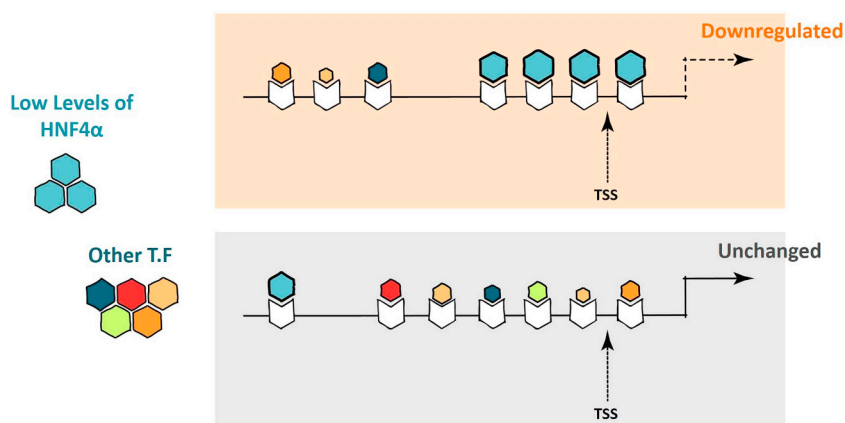


Fig. 7. Model for HNF4 α targets features in MODY1 cells. Schematic model for the promoter region of HNF4 α targets in MODY1 cells. This model represents the difference in the number of HNF4 α binding sites, the binding sites distribution, and the number of other transcription factor motifs between downregulated and unchanged HNF4 α targets.

Foundation. The research leading to these results has also received partial funding from the Innovative Medicines Initiative Joint Undertaking under grant agreement number 115439 (StemBANCC), resources of which are composed of financial contributions from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies in kind contribution. This publication reflects only the author's views and neither the IMI JU nor EFPIA nor the European Commission are liable for any use that may be made of the information contained therein. For more information, go to <http://www.stembancc.org/>.

References

- Abramoff, M.D., Magalhães, P.J., Ram, S.J., 2004. Image processing with ImageJ. *Biophotonics Int.* 11, 36–41.
- Anders, S., McCarthy, D.D., Chen, Y.Y., Okoniewski, M., Smyth, G.K.G., Huber, W., Robinson, M.D.M., Axtell, M.J., Baker, M., Bartel, D.P., et al., 2014. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *PLoS One* 8, 1–13.
- Andrews, S., 2010. FastQC: A quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Avior, Y., Sagi, I., Benvenisty, N., 2016. Pluripotent stem cells in disease modelling and drug discovery. *Nat. Rev. Mol. Cell Biol.* 17, 170–182.
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120.
- Bolotin, E., Liao, H., Ta, T.C., Yang, C., Hwang-Verslues, W., Evans, J.R., Jiang, T., Sladek, F.M., 2010. Integrated approach for the identification of human hepatocyte nuclear factor 4 α target genes using protein binding microarrays. *Hepatology* 51, 642–653.
- Bonnefond, A., Philippe, J., Durand, E., Dechaume, A., Huyvaert, M., Montagne, L., Marre, M., Balkau, B., Fajard, I., Vambergue, A., et al., 2012. Whole-exome sequencing and high throughput genotyping identified KCNJ11 as the thirteenth MODY gene. *PLoS one* 7.
- Boyd, M., Bressendorff, S., Möller, J., Olsen, J., Troelsen, J.T., 2009. Mapping of HNF4 α target genes in intestinal epithelial cells. *BMC Gastroenterol.* 9, 68.
- Carey, M., 1998. The enhancosome and transcriptional synergy. *Cell* 92, 5–8.
- Colclough, K., Bellanè-Chantelot, C., Saint-Martin, C., Flanagan, S.E., Ellard, S., 2013. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 Alpha and 4 Alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia. *Hum. Mutat.* 34, 669–685.
- Dai, C., Brissova, M., Hang, Y., Thompson, C., Poffenberger, G., Shostak, A., Chen, Z., Stein, R., Powers, A.C., 2012. Islet-enriched gene expression and glucose-induced insulin secretion in human and mouse islets. *Diabetologia* 55, 707–718.
- Dang, V.T., Kassahn, K.S., Marcos, A.E., Ragan, M.A., 2008. Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *Eur. J. Hum. Genet.* 16(11), 1350–1357.
- Ellard, S., Colclough, K., 2006. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. *Hum. Mutat.* 27, 854–869.
- Ezer, D., Zabet, N.R., Adryan, B., 2014. Homotypic clusters of transcription factor binding sites: a model system for understanding the physical mechanics of gene expression. *Comput. Struct. Biotechnol. J.* 10, 63–69.
- Fajans, S.S., Bell, G.I., 2007. Macrosomia and neonatal hypoglycaemia in RW pedigree subjects with a mutation (Q268X) in the gene encoding hepatocyte nuclear factor 4 α (HNF4A) [1]. *Diabetologia* 50, 2600–2601.
- Fajans, S.S., Bell, G.I., 2011. MODY: history, genetics, pathophysiology, and clinical decision making. *Diabetes Care* 34, 1878–1884.
- Fajans, S.S., Cloutier, M.C., Crowther, R.L., Arbor, A., 1978. The banting memorial lecture 1978 clinical and etiologic heterogeneity of idiopathic diabetes mellitus. *Diabetes* 27, 1112–1125.
- Fajans, S.S., Bell, G.I., Bowden, D.W., Halter, J.B., Polonsky, K.S., 1994. Maturity-onset diabetes of the young. *Life Sci.* 55, 413–422.
- Fajans, S.S., Bell, G.I., Polonsky, K.S., 2001. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N. Engl. J. Med.* 345, 971–980.
- Flanagan, S.E., Kapoor, R.R., Mali, G., Cody, D., Murphy, N., Schwahn, B., Sihanidou, T., Banerjee, I., Akcay, T., Rubio-Cabezas, O., et al., 2010. Diazoxide-responsive hyperinsulinemic hypoglycemia caused by HNF4A gene mutations. *Eur. J. Endocrinol.* 162, 987–992.
- Furuta, H., Iwasaki, N., Oda, N., Hinokio, Y., Horikawa, Y., Yamagata, K., Yano, N., Sugahiro, J., Ogata, M., Ohgawara, H., et al., 1997. Organization and partial sequence of the hepatocyte nuclear factor-4 alpha/MODY1 gene and identification of a missense mutation, R127W, in a Japanese family with MODY. *Diabetes* 46, 1652–1657.
- Gardner, D.S., Tai, E.S., 2012. Clinical features and treatment of maturity onset diabetes of the young (MODY). *Diabetes. Metab. Syndr. Obes.* 5, 101–108.
- Godoy, P., Wiedera, A., Schmidt-Heck, W., Campos, G., Meyer, C., Cadenas, C., Reif, R., 2016. Gene network activity in cultivated primary hepatocytes is highly similar to diseased mammalian liver tissue. *Arch. Toxicol.* 90, 2513–2529.
- Grant, C.E., Bailey, T.L., Noble, W.S., 2011. FIMO: scanning for occurrences of a given motif. *Bioinformatics* 27, 1017–1018.
- Gupta, R.K., Kaestner, K.H., 2004. HNF-4 α : from MODY to late-onset type 2 diabetes. *Trends Mol. Med.* 10, 521–524.
- Hadzopoulou-Cladaras, M., Kistanova, E., Evagelopoulou, C., Zeng, S., Cladaras, C., Ladias, J.A., 1997. Functional domains of the nuclear receptor hepatocyte nuclear factor 4. *J. Biol. Chem.* 272, 539–550.
- Hattersley, A.T., 1998. Maturity-onset diabetes of the young: clinical heterogeneity explained by genetic heterogeneity. *Diabet. Med.* 15, 15–24.
- Henzen, C., 2012. Monogenic diabetes mellitus due to defects in insulin secretion. *Swiss Med. Wkly.* 142, 1–9.
- Herman, W.H., Fajans, S.S., Smith, M.J., Polonsky, K.S., Bell, G.I., Halter, J.B., 1997. Diminished insulin and glucagon secretory responses to arginine in nondiabetic subjects with a mutation in the hepatocyte nuclear factor-4 /MODY1 gene. *Diabetes* 46, 1749–1754.
- Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57.
- Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13.
- Huang, J., Levitsky, L.L., Rhoads, D.B., 2009c. Novel P2 promoter-derived HNF4A isoforms with different N-terminus generated by alternate exon insertion. *Exp. Cell Res.* 315, 1200–1211.
- Huang, N., Lee, I., Marcotte, E.M., Hurler, M.E., 2010. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* 6, 1–11.
- Ilag, L.L., Tabaei, B.P., Herman, W.H., Zawacki, C.M., D'Souza, E., Bell, G.I., Fajans, S.S., 2000. Reduced pancreatic polypeptide response to hypoglycemia and amylin response to arginine in subjects with a mutation in the HNF-4 α /MODY1 gene. *Diabetes* 49, 961–968.
- Jiang, G., Sladek, F.M., 1997. The DNA binding domain of hepatocyte nuclear factor 4 mediates cooperative, specific binding to DNA and heterodimerization with the retinoid X receptor alpha. *J. Biol. Chem.* 272, 1218–1225.
- Kinsella, R.J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., et al., 2011. Ensembl BioMarts: A hub for data retrieval across taxonomic space. *Database* 2011.
- Koudritsky, M., Domany, E., 2008. Positional distribution of human transcription factor binding sites. *Nucleic Acids Res.* 36, 6795–6805.
- Lehto, M., Bitzén, P.O., Isomaa, B., Wipemo, C., Wessman, Y., Forsblom, C., Tuomi, T., Taskinen, M.R., Groop, L., 1999. Mutation in the HNF-4alpha gene affects insulin secretion and triglyceride metabolism. *Diabetes* 48, 423–425.
- Lu, J., Xia, Q., Zhou, Q., 2016. How to make insulin-producing pancreatic β cells for diabetes treatment. *Sci. China Life Sci.* 60, 239–248.
- Madsen, O.D., Serup, P., 2006. Towards cell therapy for diabetes. *Nat. Biotechnol.* 24, 1481–1483.
- Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R., et al., 2016. JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids*

- Res. 44, D110–D115.
- McDonald, T.J., Ellard, S., 2013. Maturity onset diabetes of the young: identification and diagnosis. *Ann. Clin. Biochem.* 50, 403–415.
- Medina-Rivera, A., Defrance, M., Sand, O., Herrmann, C., Castro-Mondragon, J.A., Delerce, J., Blanchet, C., Vincens, P., Caron, C., Staines, D.M., et al., 2015. RSAT 2015: Regulatory sequence analysis tools. *Nucleic Acids Res.* 43, 50–56.
- Melton, D.A., 2016. Applied developmental biology: making human pancreatic beta cells for diabetics. *Curr. Top. Dev. Biol.* 117, 65–73.
- Nyunt, O., Wu, J.Y., McGown, I.N., Harris, M., Huynh, T., Leong, G.M., Cowley, M., Cotterill, A.M., 2009. Investigating maturity onset diabetes of the young. *Clin. Biochem. Rev.* 30, 67–74.
- Odom, D.T., Zizlsperger, N., Gordon, D.B., Bell, G.W., Rinaldi, N.J., Murray, H.L., Volkert, T.L., Schreiber, J., Rolfe, P.A., Gifford, D.K., et al., 2004. Control of pancreas and liver gene expression by HNF transcription factors. *Science* 303, 1378–1381.
- Odom, D.T., Dowell, R.D., Jacobsen, E.S., Neklodova, L., Rolfe, P.A., Danford, T.W., Gifford, D.K., Fraenkel, E., Bell, G.I., Young, R.A., 2006. Core transcriptional regulatory circuitry in human hepatocytes. *Mol. Syst. Biol.* 2, 2006.0017.
- Pearson, E.R., Boj, S.F., Steele, A.M., Barrett, T., Stals, K., Shield, J.P., Ellard, S., Ferrer, J., Hattersley, A.T., 2007. Macrosomia and hyperinsulinaemic hypoglycaemia in patients with heterozygous mutations in the HNF4A gene. *PLoS Med.* 4, 760–769.
- Piper, K., Brickwood, S., Turnpenny, L.W., Cameron, I.T., Ball, S.G., Wilson, D.I., Hanley, N.A., 2004. Beta cell differentiation during early human pancreas development. *J. Endocrinol.* 181, 11–23.
- Prudente, S., Jungtrakoon, P., Marucci, A., Ludovico, O., Buranasupkajorn, P., Mazza, T., Hastings, T., Milano, T., Morini, E., Mercuri, L., et al., 2015. Loss-of-function mutations in APPL1 in familial diabetes mellitus. *Am. J. Hum. Genet.* 97, 177–185.
- Rada-Iglesias, A., Wallerman, O., Koch, C., Ameur, A., Enroth, S., Clelland, G., Wester, K., Wilcox, S., Dovey, O.M., Ellis, P.D., et al., 2005. Binding sites for metabolic disease related transcription factors inferred at base pair resolution by chromatin immunoprecipitation and genomic microarrays. *Hum. Mol. Genet.* 14, 3435–3447.
- Rezania, A., Bruin, J.E., Riedel, M.J., Mojibian, M., Asadi, A., Xu, J., Gauvin, R., Narayan, K., Karanu, F., O'Neil, J.J., et al., 2012. Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. *Diabetes* 61, 2016–2029.
- Servitja, J.M., Ferrer, J., 2004. Transcriptional networks controlling pancreatic development and beta cell function. *Diabetologia* 47, 597–613.
- Sharan, R., Maron-Katz, A., Shamir, R., 2003. CLICK and EXPANDER: a system for clustering and visualizing gene expression data. *Bioinformatics* 19, 1787–1799.
- Shih, D.Q., Dansky, H.M., Fleisher, M., Assmann, G., Fajans, S.S., Stoffel, M., 2000. Genotype/phenotype relationships in HNF-4alpha/MODY1: haploinsufficiency is associated with reduced apolipoprotein (AII), apolipoprotein (CIII), lipoprotein(a), and triglyceride levels. *Diabetes* 49, 832–837.
- Shih, D.Q., Heimesaat, M., Kuwajima, S., Stein, R., Wright, C.V.E., Stoffel, M., 2002. Profound defects in pancreatic beta-cell function in mice with combined heterozygous mutations in Pdx-1, Hnf-1alpha, and Hnf-3beta. *Proc. Natl. Acad. Sci. U. S. A.* 99, 3818–3823.
- Siddiqui, K., Musambil, M., Nazir, N., 2015. Maturity onset diabetes of the young (MODY) — History, first case reports and recent advances. *Gene* 555, 66–71.
- Stepniowski, J., Kachamakova-Trojanowska, N., Ogrocki, D., Szopa, M., Matlok, M., Beilharz, M., Dyduch, G., Malecki, M.T., Jozkowicz, A., Dulak, J., 2015. Induced pluripotent stem cells as a model for diabetes investigation. *Sci. Rep.* 5, 8597.
- Stoffel, M., Duncan, S.A., 1997. The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc. Natl. Acad. Sci. U. S. A.* 94, 13209–13214.
- Stride, A., Hattersley, A.T., 2002. Different genes, different diabetes: lessons from maturity-onset diabetes of the young. *Ann. Med.* 34, 207–216.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550.
- Tabach, Y., Brosh, R., Buganim, Y., Reiner, A., Zuk, O., Yitzhaky, A., Koudritsky, M., Rotter, V., Domany, E., 2007. Wide-scale analysis of human functional transcription factor binding reveals a strong bias towards the transcription start site. *PLoS One* 2.
- Taraviras, S., Paula Monaghan, A., Schütz, G., Kelsey, G., 1994. Characterization of the mouse HNF-4 gene and its expression during mouse embryogenesis. *Mech. Dev.* 48, 67–79.
- Teo, A.K.K., Wagers, A.J., Kulkarni, R.N., 2013a. New opportunities: harnessing induced pluripotency for discovery in diabetes and metabolism. *Cell Metab.* 18, 775–791.
- Teo, A.K.K., Windmueller, R., Johansson, B.B., Dirice, E., Njolstad, P.R., Tjora, E., Raeder, H., Kulkarni, R.N., 2013b. Derivation of human induced pluripotent stem cells from patients with maturity onset diabetes of the young. *J. Biol. Chem.* 288, 5353–5356.
- Teo, A.K.K., Lau, H.H., Valdez, I.A., Dirice, E., Tjora, E., Raeder, H., Kulkarni, R.N., 2016. Early developmental perturbations in a human stem cell model of MODY5/HNF1B pancreatic hypoplasia. *Stem Cell Rep.* 6, 357–367.
- Thanabalasingham, G., Owen, K.R., 2011. Diagnosis and management of maturity onset diabetes of the young (MODY). *BMJ* 343, d6044.
- Thomsen, S.K., Ceroni, A., Van De Bunt, M., Burrows, C., Barrett, A., Scharfmann, R., Ebner, D., McCarthy, M.I., 2016. Systematic functional characterization of candidate causal genes for type 2 diabetes risk variants. *Diabetes* 65, 3805–3811.
- Todeschini, A., Georges, A., Veitia, R.A., 2014. Transcription factors: specific DNA binding and specific gene regulation. *Trends Genet.* 30, 211–219.
- Vethe, H., Bjørlykke, Y., Ghila, L.M., Paulo, J.A., Scholz, H., Gygi, S.P., Chera, S., Ræder, H., 2017. Probing the missing mature β -cell proteomic landscape in differentiating patient iPSC-derived cells. *Sci. Rep.* 7, 4780.
- Wallerman, O., Motallebipour, M., Enroth, S., Patra, K., Sudhan, M., Bysani, R., Komorowski, J., Wadelius, C., 2009. Molecular interactions between HNF4 α , FOXA2 and GABP identified at regulatory DNA elements through ChIP-sequencing. *Nucleic Acid Res.* 37, 7498–7508.
- Wang, Z., Bishop, E.P., Burke, P.A., 2011. Expression profile analysis of the inflammatory response regulated by hepatocyte nuclear factor 4 α . *BMC Genom.* 12.
- Wang, X., Chen, S., Burtscher, I., Sterr, M., Hieronimus, A., Machicao, F., Staiger, H., Häring, H.-U., Lederer, G., Meitinger, T., et al., 2016. Generation of a human induced pluripotent stem cell (iPSC) line from a patient carrying a P33T mutation in the PDX1 gene. *Stem Cell Res.* 17, 273–276.
- Whitfield, T.W., Wang, J., Collins, P.J., Partridge, E.C., Aldred, S.F., Trinklein, N.D., Myers, R.M., Weng, Z., 2012. Functional analysis of transcription factor binding sites in human promoters. *Genome Biol.* 13.
- Winter, W.E., 2003. Newly defined genetic diabetes syndromes: maturity onset diabetes of the young. *Rev. Endocr. Metab. Disord.* 4, 43–51.
- Wu, W.-S., Lai, F.-J., 2015. Functional redundancy of transcription factors explains why most binding targets of a transcription factor are not affected when the transcription factor is knocked out. *BMC Syst. Biol.* 9, S2.
- Yabe, S.G., Iwasaki, N., Yasuda, K., Hamazaki, T.S., Konno, M., Fukuda, S., Takeda, F., Kasuga, M., Okochi, H., 2015. Establishment of maturity-onset diabetes of the young-induced pluripotent stem cells from a Japanese patient. *J. Diabetes Investig.* 6, 543–547.
- Yamagata, K., Furuta, H., Oda, N., Kaisaki, P.J., Menzel, S., Cox, N.J., Fajans, S.S., Signorini, S., Stoffel, M., Bell, G.I., 1996. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). *Nature* 384, 458–460.
- Yang, Y., Chan, L., 2017. Monogenic diabetes: what it teaches us on the common forms of type 1 and type 2 diabetes I. Introduction 1–35.