

# SCIENTIFIC REPORTS

OPEN

## Probing the missing mature $\beta$ -cell proteomic landscape in differentiating patient iPSC-derived cells

Heidrun Vethe<sup>1</sup>, Yngvild Bjørlykke<sup>4</sup>, Luiza M. Ghila<sup>1</sup>, Joao A. Paulo<sup>2</sup>, Hanne Scholz<sup>3</sup>, Steven P. Gygi<sup>2</sup>, Simona Chera<sup>1</sup> & Helge Ræder<sup>1,4</sup>

**MODY1** is a maturity-onset monogenic diabetes, caused by heterozygous mutations of the *HNF4A* gene. To date the cellular and molecular mechanisms leading to disease onset remain largely unknown. In this study, we demonstrate that insulin-positive cells can be generated *in vitro* from human induced pluripotent stem cells (hiPSCs) derived from patients carrying a non-sense *HNF4A* mutation, proving for the first time, that a human *HNF4A* mutation is neither blocking the expression of the insulin genes nor the development of insulin-producing cells *in vitro*. However, regardless of the mutation or diabetes status, these insulin-producing cells are immature, a common downfall off most current  $\beta$ -cell differentiation protocols. To further address the immature state of the cells, *in vitro* differentiated cells and adult human islets were compared by global proteomic analysis. We report the predicted upstream regulators and signalling pathways characterizing the proteome landscape of each entity. Subsequently, we focused on the molecular components absent or misregulated in the *in vitro* differentiated cells, to probe the components involved in the deficient *in vitro* maturation towards fully functional  $\beta$ -cells. This analysis identified the modulation of key developmental signalling pathways representing potential targets for improving the efficiency of the current differentiation protocols.

Monogenic disorders are caused by germline single gene defects where different mutations in the causal gene usually trigger a defined disorder with characteristic clinical features. The identification of the genes and molecular networks underlining monogenic disorders allows for unbiased characterization of the basic mechanisms regulating cell-fate decisions during development and disease onset. This approach also facilitates the understanding of the aetiology of the more prevalent corresponding multifactorial diseases as well as general developmental aspects. One such example is Parkinson disease, where the study of its few monogenic variants tremendously boosted the knowledge of the mechanisms involved in neuronal differentiation, homeostasis and disease initiation<sup>1</sup>. Similarly, MODY (Maturity Onset Diabetes of the Young) represent a distinct group of diabetic disorders characterized by the impairment of pancreatic  $\beta$ -cells (the insulin-producing cells) caused by an autosomal dominantly inherited mutations. Due to their unambiguous and inheritable genetic readout, MODYs are ideal tools for elucidating the molecular and cellular basis involved in  $\beta$ -cell differentiation and failure.

Studies on human patients are extremely challenging and have inherent technical and ethical limitations. As a result, most research on human diseases is based on model organisms and *in vitro* approaches. Moreover, as many of the currently available murine models of MODY fail to accurately replicate the equivalent human conditions<sup>2–5</sup>, the efforts for understanding the dynamic of  $\beta$ -cell failure focuses mostly on *in vitro* setups. Consequently, the past decade has seen the development of several *in vitro* directed differentiation protocols using human pluripotent stem cells (hiPSCs) as a renewable resource for making insulin-producing cells as models for diabetes<sup>6–10</sup>.

The protocols reported in 2014 by *Rezania et al.* and *Pagliuca et al.* shared a similar stepwise directed differentiation strategy attempting to mimic different stages of the embryonic development of  $\beta$ -cells by modulating similar developmental signalling pathways. Both protocols achieved to generate  $\beta$ -like cells on a comparable

<sup>1</sup>KG Jebsen Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway.

<sup>2</sup>Department of Cell Biology, Harvard Medical School, Boston, MA, USA. <sup>3</sup>Department of Transplant Medicine, Oslo University Hospital, Oslo, Norway. <sup>4</sup>Department of Pediatrics, Haukeland University Hospital, Bergen, Norway. Correspondence and requests for materials should be addressed to H.R. (email: [Helge.Rader@uib.no](mailto:Helge.Rader@uib.no))

timescale, employing slightly different compounds and working concentrations. A novel protocol reported by *Russ et al.* in 2015 suggested simplifications of the protocols without compromising differentiation efficacy, i.e. by omitting BMP inhibitors at the pancreas specification stage and by finding that retinoic acid alone was sufficient to generate more than 98% of PDX1+ pancreatic progenitors<sup>8</sup>. To date and to our knowledge there have been no reports of the use of these protocols in the differentiation of hiPSCs from patients with monogenic diabetes (MODY or permanent neonatal diabetes of the young, PNDM). With regards to MODY, differentiation of hiPSCs was recently reported using different protocols: from MODY5 patients towards pancreatic progenitor cells<sup>11</sup>, and from MODY2 patients towards  $\beta$ -like cells<sup>12</sup>.

However, as the current knowledge of  $\beta$ -cell differentiation process during development is incomplete, a common problem of the available *in vitro* differentiation protocols is the production of mostly immature “ $\beta$ -like cells”<sup>13</sup> unable to perform accurate glucose-stimulated insulin secretion unless they are transplanted into mice and allowed to mature *in vivo*<sup>14</sup>. This consistent failure of obtaining mature functional insulin-producing cells *in vitro* signifies the absence of a maturing/differentiating factor or factors present *in vivo*, however absent in the current protocols. Correspondingly, finding and characterizing this developmentally important molecular network required for the final maturation of  $\beta$ -cell *in vitro* is urgently needed in order to generate functional insulin-producing cells. Most current attempts towards characterizing  $\beta$ -cell molecular networks are based on next generation sequencing tools such as RNA-seq. Despite the undeniable power and sensitivity of the transcriptomics methods the progress is slow, hence there is also a need for complementary characterizing methods, such as proteomics methods. An increasing number of studies have reported consistent and biologically relevant differences when comparing transcriptomics and proteomics data<sup>15,16</sup>. These discrepancies are explained by the different dynamics of the RNA and protein products. As an example the ribosome may alter the translational efficiency of mRNA at the initiation and elongation stages<sup>17</sup>. Furthermore, many cellular signals do not activate the transcription of the relevant downstream pathway components, as these proteins have already been synthesized in the cells and are regulated by post-translational modification, such as in the case of insulin signalling. Moreover, the half-lives of transcripts and their respective protein products are different, i.e. with situations where the protein is persistently involved in cellular processes after the disappearance of the corresponding transcript. In any of these cases, transcriptomics tools will fail to detect correctly the changes in gene product abundance or signalling patterns.

Here we used a combination of global proteomics and cellular biology techniques to investigate the differentiation capacity of insulin-producing cells using a seven-step differentiation protocol (as established by *Rezania et al.*), generated from either healthy subjects or MODY1 (i.e. *HNF4A* gene mutation carrying) patients. Next, we compared the *in vitro* stage 7 (S7) cell proteome with human pancreatic islet proteome and identified differentially expressed proteins as well as specific molecular networks distinguishing the end-stage *in vitro* S7 cells from the bona-fide islet cells.

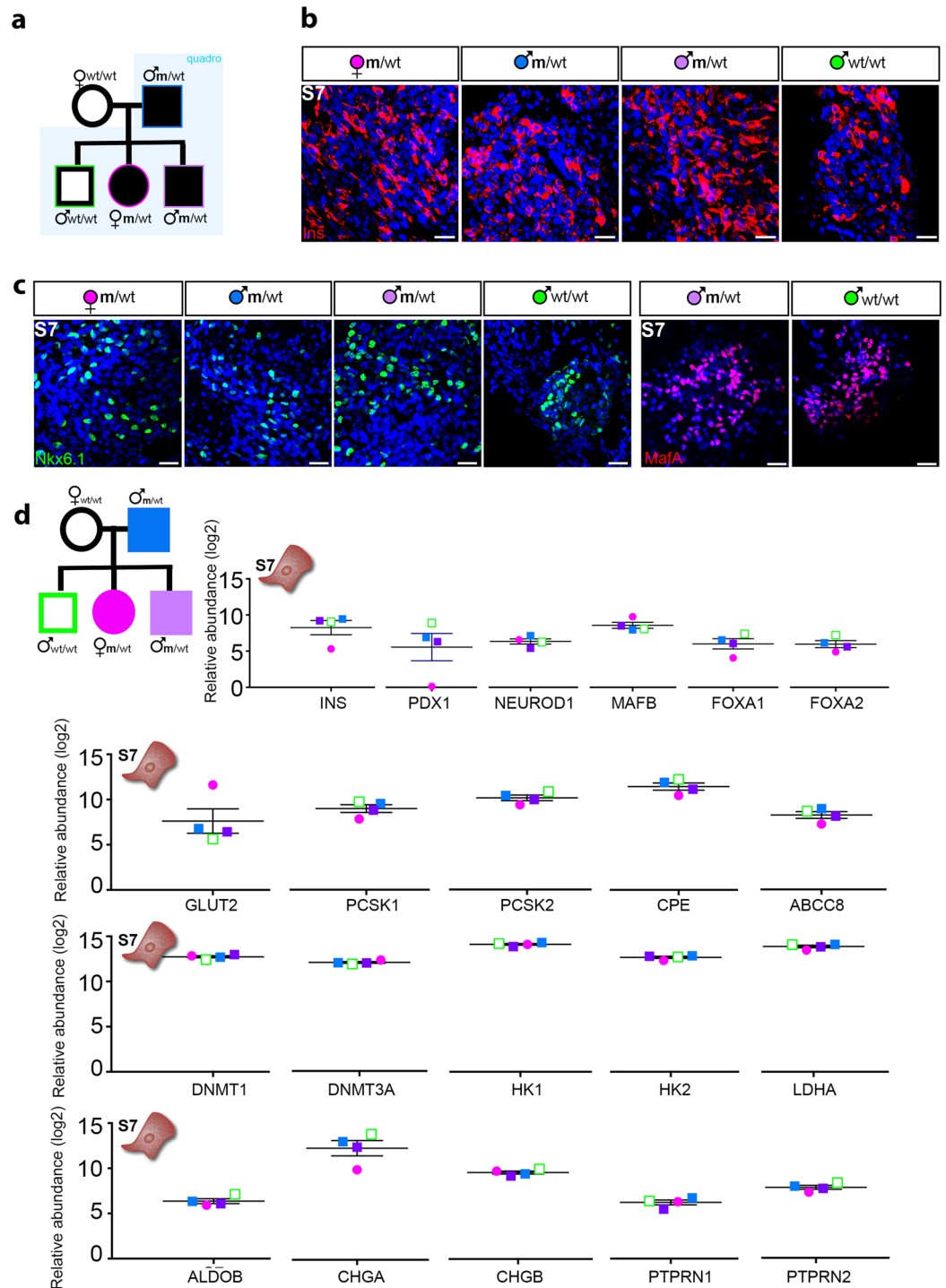
## Results

***HNF4A* mutation (MODY1) does not prevent the formation of insulin<sup>+</sup> cells *in vitro*.** We first inquired if the *HNF4A* mutation or diabetes status prevented the differentiation of insulin<sup>+</sup> cells *in vitro*. As the limited accessibility to human samples and the increased number of variables inherent to any *in vitro* differentiation protocol rule out a differential quantitative analysis, we focused on whether insulin<sup>+</sup> cells are present or, alternatively absent in each sample (qualitative assessment). To answer this question, skin fibroblasts from a healthy family member and mutation carriers before and after the onset of diabetes from a MODY1 cohort quadro (n = 4, see Fig. 1a) were reprogrammed. The “MODY1 cohort quadro” refers here to the four family members, including a parental diabetic mutation carrier, and three offspring: one diabetic mutation carrier, one non-diabetic mutation carrier, and non-diabetic non-mutation carrier (family control). The synchronous differentiation of the four hiPSCs lines towards insulin-producing cells (n = 3 repeated rounds of differentiation) was done following the seven step differentiation protocol established by *Rezania et al.* (Fig. S1) and resulted in the successful generation of insulin<sup>+</sup> cells in all four samples regardless of the mutation or diabetes status (Fig. 1b). The Stage7 cell population (S7 thereafter) also expressed key markers of pancreatic  $\beta$ -cells, including NKX6.1 and MAFA (Fig. 1c). Furthermore, the global proteomic comparison of the final stage (S7) cells derived from MODY1-mutation carriers and their family control detected in all samples similar levels of insulin and most key  $\beta$ -cell specific markers, regardless of the mutation or diabetes status (Fig. 1d), confirming the initial immunofluorescence (IF) observations. Overall, these data suggest that the *HNF4A* mutation is neither blocking the expression of the insulin genes nor the development of insulin-producing cells *in vitro*.

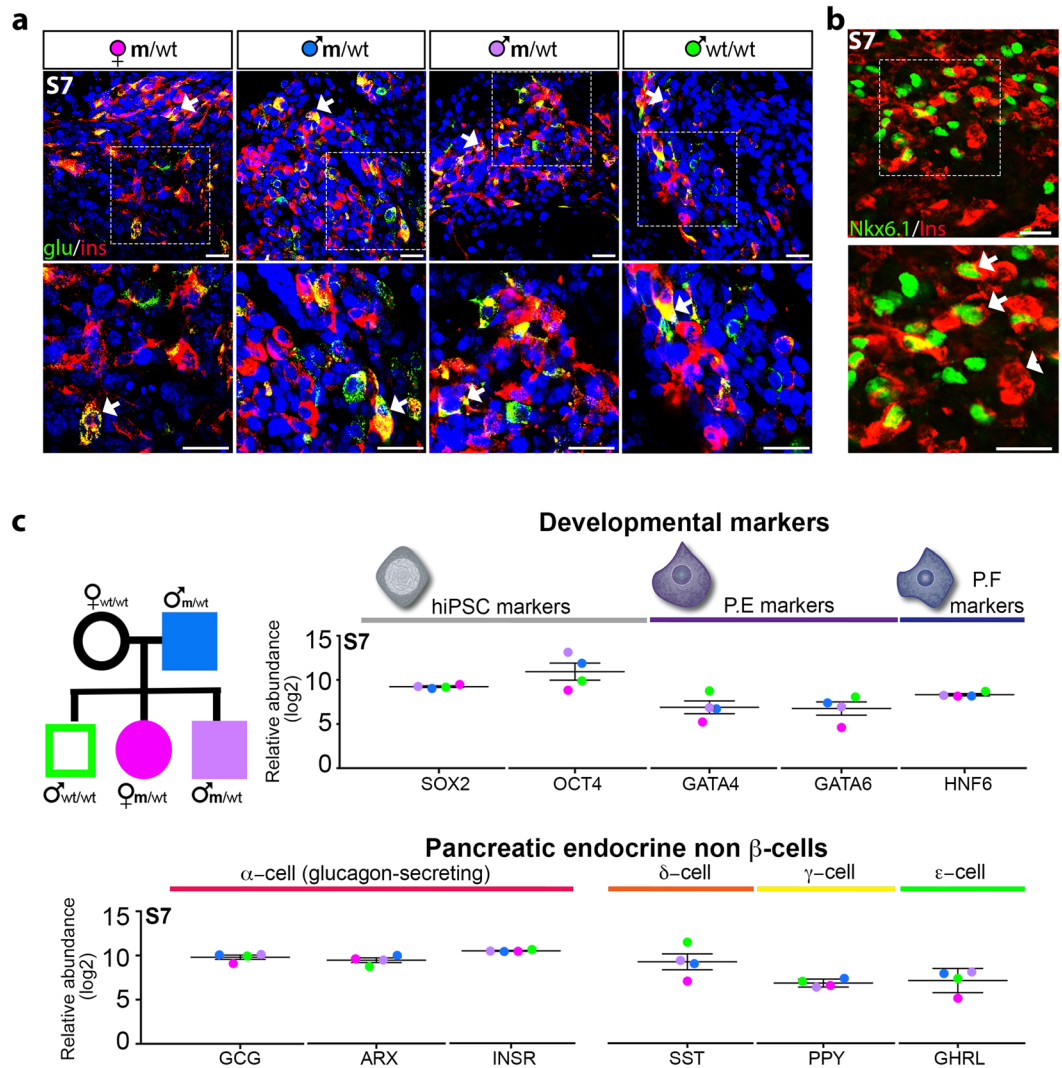
**Differentiated S7 cells are characterized by high heterogeneity regardless of the *HNF4a* mutation and disease status.** As reported previously<sup>9</sup>, we observed the co-expression of other islet hormones in the S7 cell population in all samples analysed regardless of mutation status. The co-expression was observed by IF either in the same cell (bi-, multi-hormonal cells) or in different cells belonging to the same S7 pool (Fig. 2a). Moreover, in all S7 insulin<sup>+</sup> cell populations we observed heterogeneity, with some insulin<sup>+</sup> cells co-expressing the  $\beta$ -cell marker NKX6.1 (Fig. 2b, arrows), while others did not (Fig. 2b, arrowhead).

The global proteomic comparison of the MODY-derived S7 samples and their family control confirmed these results, with significant levels of non- $\beta$  endocrine islet cells markers, progenitor cell markers as well as stem cell markers being detected in all samples (Fig. 2c). In our S7 cell populations we detected specific markers of pancreatic  $\delta$ -cells (somatostatin (SST)), pancreatic  $\alpha$ -cells (glucagon (GCG), the  $\alpha$ -cell specific transcription factor ARX, insulin receptor (IR) and proconvertase 2 (PCSK2)), pancreatic  $\gamma$ -cells (pancreatic polypeptide (PPY)) and  $\epsilon$ -cells (ghrelin (GHR)).

These data suggest that neither *HNF4A* mutation nor the diabetic status impose a regenerative-driven demand promoting an exclusive  $\beta$ -cell fate. However, the limited number of samples included in this study does not qualify



**Figure 1.** Generation of patient-specific insulin-producing cells from a MODY1 quadro. **(a)** Patient-specific iPSCs were used from four family members (quadro) of a MODY1 family, including a parental diabetic mutation carrier (shown as blue filled square), and three offspring; one diabetic mutation carrier (purple filled square), one non-diabetic mutation carrier (pink filled circle), and non-diabetic non-mutation carrier (green unfilled square). **(b)** Differentiation of the four hiPSC lines towards insulin-producing cells were repeated three times ( $n = 3$ ) following the seven-stage protocol (as established by Rezanian *et al.*), and resulted in insulin+ cells (red) from all four hiPSC lines regardless of mutation status. **(c)** The S7 cell population also expressed *NKX6.1* (turquoise) and *MAFA* (magenta). **(d)** Quantitative proteomics detected insulin along with other  $\beta$ -cell specific markers in all subjects regardless of mutation status. Each patient specific data point is shown in the colours as described in **(a)**. Data are shown as mean with SEM. The y-axis shows the relative abundance of each protein (log<sub>2</sub>) compared to all the ten samples of a TMT ten-plex as further explained in Methods. Scale bar: 25  $\mu$ m.



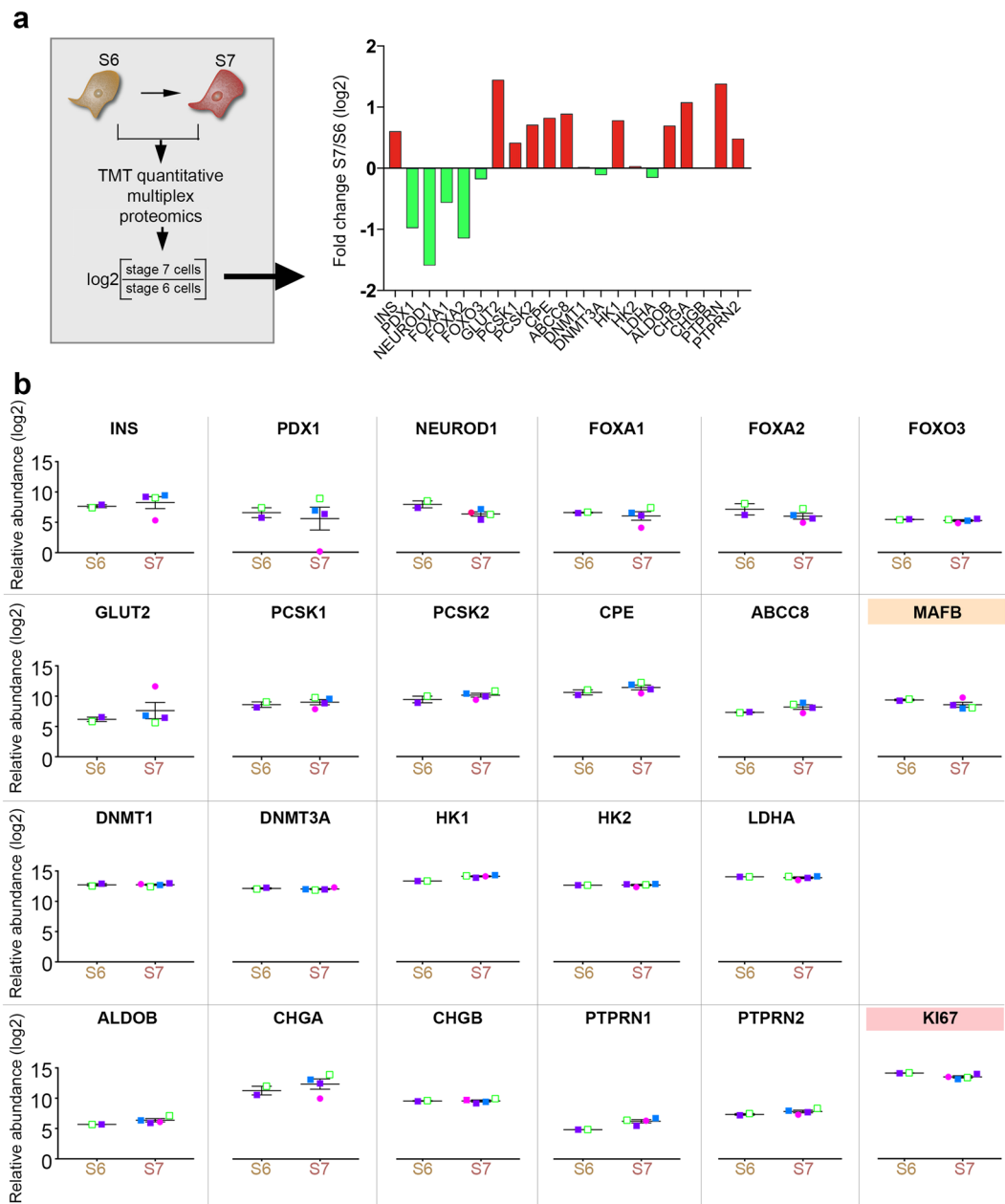
**Figure 2.** Cellular heterogeneity. (a) Co-expression of glucagon (green) and insulin (red) in S7 cells (arrows). (b) Insulin+ S7 cells (red) positive (arrows) or negative (arrowhead) for NKX6.1 transcription factor (green). (c) Relative abundance (log<sub>2</sub>) of non- $\beta$  endocrine islet cells markers in S7 cells compared to all the ten samples of a TMT 10-plex as further explained in Methods. Scale bar: 25  $\mu$ m.

for a reliable quantitative marker analysis between the different S7 samples, hence we cannot exclude that HNF4A mutation or the diabetes status might moderately promote a certain cell fate.

**HNF4A mutation does not interfere with  $\beta$ -cell maturity markers acquisition *in vitro*.** We then investigated if the *HNF4A* mutation or diabetes status impeded the initial *in vitro* maturation steps of the differentiated insulin<sup>+</sup> cells. We addressed this question by performing global proteomics comparison of Stage 6 cell population (S6 thereafter) and S7 cells derived from *MODY1* mutation carriers and their family control without the mutation. The comparison revealed that, in all the samples analysed and regardless of mutation or diabetes status, several but not all  $\beta$ -cell specific markers including those involved in glucose sensing, regulation of glucose stimulated insulin secretion (thereafter GSIS) and insulin biosynthesis, showed an increasing trend at S7 compared to the previous stage (S6) (Fig. 3a and b). The increasing trend for insulin protein abundance was also in accordance with the IF staining data for insulin (Supplementary Fig. 1). Moreover, the levels of MAFB transcription factor decreased between S6 and S7 cells, consistent with previous studies reporting the switch from MAFB to MAFA expression during  $\beta$ -cell maturation<sup>18</sup> (Fig. 3b, orange background).

**Despite gaining certain maturity markers, the S7 insulin+ cells are functionally immature *in vitro*, regardless of the *HNF4a* mutation and diabetic status.** To investigate if the maturation status of the S7 insulin<sup>+</sup> cells is similar to the mature bona-fide  $\beta$ -cells we first performed ultrastructural analysis by transmission electron microscopy (TEM) and identified fields of endocrine granules with different core densities in S7 samples (Fig. 4a–d), consistent with the three types of insulin granules described in mature human  $\beta$ -cells<sup>9</sup>. There were no differences between endocrine granules from mutation-carriers and healthy family control; therefore representative pictures are shown for the healthy control (Fig. 4a–d). These results suggest that despite the S7

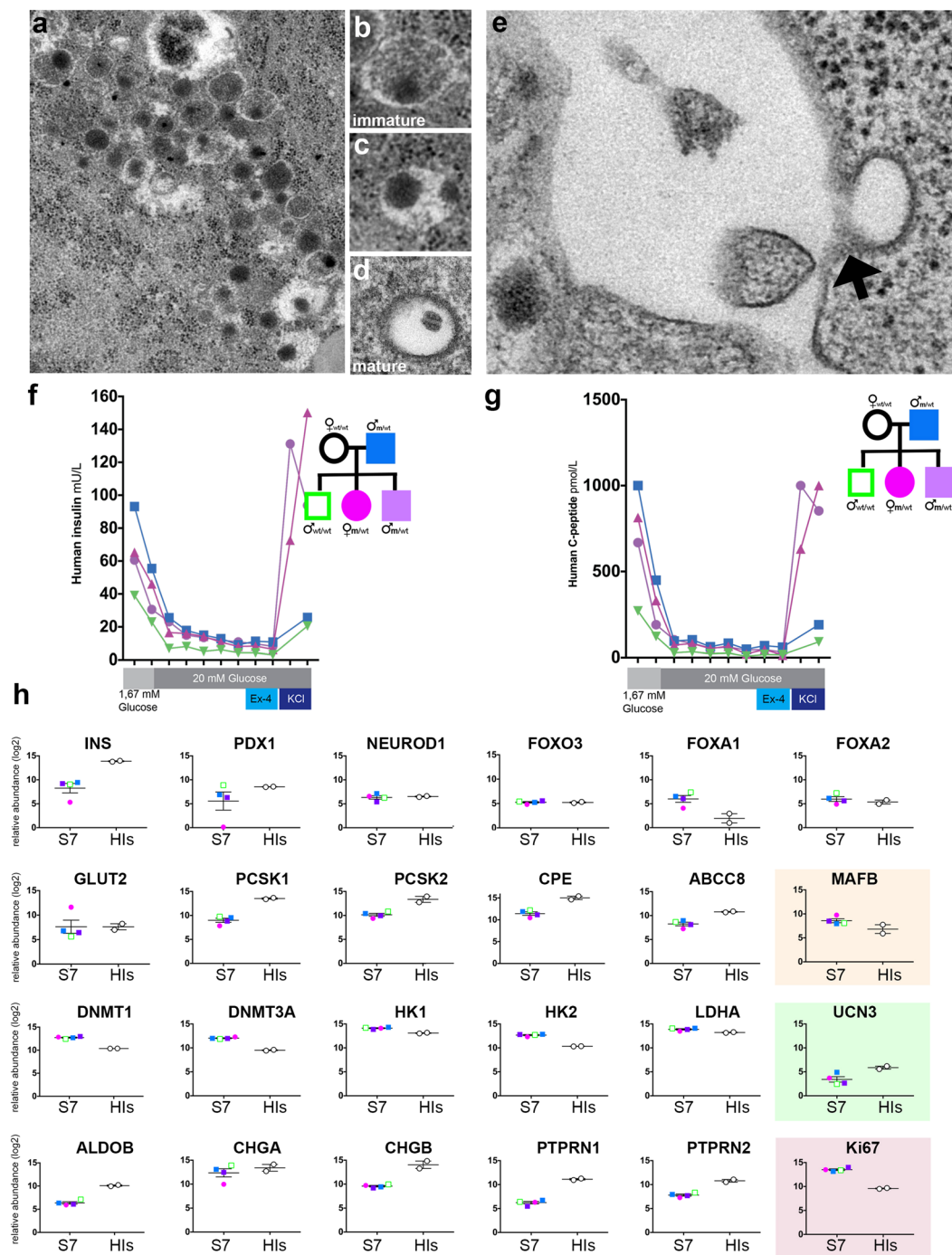




**Figure 3.** Proteomics comparison of S6 and S7 cells. The relative abundance of proteins important for  $\beta$ -cell function are represented between two S6 samples (one mutation carrier and one family control) and four S7 samples (three mutation carriers and one family control without the mutation). Each patient specific data point is shown in the colours as described in 1a). Due to the limited number of samples allowed in one TMT 10-plex ( $n = 10$ ), data from two S6 samples are shown. **a**) Fold difference ( $\log_2$ ) between S6 and S7 **(b)** Relative abundance ( $\log_2$ ) between S6 and S7 cells. All plots show mean protein level and standard error of mean (SEM).

cells containing a heterogeneous population of insulin<sup>+</sup> cells (Fig. 4b and d, respectively), at least some of these insulin<sup>+</sup> cells are capable of insulin processing and secretion (Fig. 4e).

We then evaluated the functionality of the differentiated insulin<sup>+</sup> cells present in each sample by testing their capacity to secrete insulin following glucose stimulation (GSIS). Mature  $\beta$ -cells secrete insulin within minutes of glucose stimulation and will subsequently shut off the secretion to maintain glucose homeostasis. To determine whether our S7 cells contained cells capable of an accurate GSIS response, we performed a static incubation assay, including stimulating the S7 cells with high glucose (20 mM) alone, and supplemented with Exendin-4 (10 nM), a Glucagon-like protein 1 (GLP-1) receptor agonist. Finally, cell depolarization was tested with high glucose supplemented with 30 mM KCl (insulin secretion in Fig. 4f, c-peptide secretion in Fig. 4g). The static incubation assay revealed that our S7 cells are immature, as they show a decreased glucose threshold for insulin secretion<sup>13</sup>, and they release insulin in an uncontrolled manner with surprisingly lower insulin secretion at higher glucose concentrations (Fig. 4f). There were no observed HNF4A mutation-specific differences in the GSIS readout



**Figure 4.** (a) Representative transmission electron microscopy images showing fields of endocrine granules in S7 cells. The S7 cells subjected to TEM analysis were not glucose-stimulated. Scale: 25 k magnification. Insulin granules in human  $\beta$ -cells are typically categorized into three main types: (b) pale, diffuse grey core (immature granule), (c) dense round core without marked edge, and (d) dense round core with marked edge (mature granule) Scale: 25k magnification for (b), (c and d). (e) Exocytosis of endocrine granule (black arrow). Scale: 80k magnification. (f) Human Insulin secretion (mU/L) of S7 cells following GSIS. (g) Human C-peptide secretion (pmol/L) of S7 cells following GSIS. (h) Proteomics comparison of  $\beta$ -cell markers in S7 cells (each patient specific data point is shown in the colours as described in 1a), and human islets (HIs) (shown as white filled circles). All plots show mean protein level with SEM.

(Fig. 4f,g). These data match the insulin secretion observed from hESC-derived S7 cells reported previously<sup>9</sup>. We further confirmed these findings in assays where the cells were kept in low glucose overnight (1.67 mM) followed by extensive washings (glucose-free DPBS and Krebs buffers without glucose) before high glucose stimulation

(20 mM). Moreover, proinsulin was not detected by ELISA confirming the absence of contaminating cells or cell debris in the media. These findings combined, rule out a technical error causing the low insulin and c-peptide release response observed. However, interestingly, insulin and c-peptide secretory response to KCl was increased compared to what have been reported for maturing  $\beta$ -like cells using the same protocol<sup>9</sup>.

To further analyse the S7 cells' maturity status we compared the proteome of S7 cells derived from the three MODY1 mutation carriers and their family control to the proteome of adult human islets. We chose islets as opposed to sorted human  $\beta$ -cells as the pancreatic islet is more similar to S7 cells regarding cell heterogeneity. Moreover, the FACS procedure is known to increase the mortality and stress of sorted  $\beta$ -cells hence potentially bringing unwanted global changes to their proteome distribution. The analysis revealed that, for some selected key pancreatic  $\beta$ -cell-markers, all the S7 cells, regardless of the MODY1 mutation status, exhibited similar abundance levels compared to human islets (HIs) (Fig. 4h, Fig. S2), while other human islets markers had a higher abundance in the human islets. One such example is Urocortin-3 (UCN3), a marker of functionally mature  $\beta$ -cells<sup>19</sup>, which showed increased protein level in human islets as compared to all S7 samples (MODY1 mutation carriers and the family control) (Fig. 4h, green background). Similarly, the abundance level of MAFB was lower in human islets, indicating that S7 cells, regardless of the mutation or diabetes status, are less mature than human islet cells (Fig. 4h, orange background). Moreover, despite the increased abundance of proteins from S6 to S7 involved in granulogenesis (Fig. S3), suggesting that the *in vitro*-generated insulin<sup>+</sup> cells are capable of producing appropriately packed insulin granules and possess exocytosis machinery (Fig. 4a–e), our data also showed lower levels in S7 compared to islets of proconvertase 1 (PCSK1), a key actor in the insulin processing machinery. Compared to the reported expression level of PCSK1 in S7 cells<sup>9</sup>, we find the opposite regulation level as compared to human islets, suggesting an inefficient insulin processing machinery. This, in combination with decreased GSIS, points towards the immaturity status of the S7 cells in all MODY1 mutation carriers as well as in the family control.

Last, the proliferation marker KI67 was increased in the S7 cell population as compared to human islets, showing many of these cells retain their proliferative potential, further supporting the notion of an immature phenotype (Fig. 4h, red background).

**Comparative analysis of S7 cells and human islets reveals missing molecular landscape involved in  $\beta$ -cell fate.** Characterizing the inherent molecular differences between the artificially generated hiPSC-derived insulin<sup>+</sup> cells and bona-fide  $\beta$ -cells is critical for improving the yield of the differentiation protocol and the quality of the end product. To address this critical point using proteomics on individual samples, we first focused on the proteins that were identified in only one of the main compared categories, i.e. in either the S7-cells (from the MODY1 quadro with three MODY1-mutation carriers and their family control) or in the human islets, but not in both (Fig. 5a). All proteins identified in at least one sample were selected for the corresponding condition grouping. We found 996 proteins uniquely identified in human islets and 1874 proteins uniquely identified in the S7 cells (MODY1 quadro S7P).

The Ingenuity Pathway Analysis (IPA) of the human islets specific proteome landscape predicted among the top 5 upstream regulators (i.e. regulators responsible for the observed protein landscape) important transcription factors of  $\beta$ -cell identity and function (Fig. 5b), with HNF1 $\alpha$  (p-value of overlap 2.40E-11) and PPAR $\gamma$  (p-value of overlap 3.22E-11) as top regulators. A similar analysis of the S7-cells-specific proteome landscape revealed among the top 5 upstream regulators, transcription factors involved in cell cycle progression, such as TP53 (p-value of overlap 5.97E-11) and MYC (p-value of overlap 3.79E-07) (Fig. 5c). These results suggest that the S7 cells (from the MODY1 quadro) contain a significant fraction of proliferating cells. Moreover, as indicated by the human islets specific proteome landscape analysis, the S7 cells (from the MODY1 quadro) are also missing critical networks required for  $\beta$ -cell identity and function.

**S7 cells and human islet comparison reveals differentially regulated candidate molecular networks potentially involved in  $\beta$ -cell fate.** We further performed pathway analysis on the proteins present in both S7 cells (from the MODY1 quadro) and human islets, but differentially expressed between the two conditions (FC  $\geq 2$ , hiPSC-derived S7 as compared to human islets, Fig. 5d). The top 5 upstream regulators predicted by the analysis mostly consisted of genes involved in cell-cycle regulation, such as TP53 (predicted inhibition), MYC (predicted activation) and PPARA (Fig. 5e). Moreover, the top canonical pathway identified, the LXR/RXR pathway, which activation is known to block cell proliferation, was predicted as strongly inhibited (p-value 2.28E-13, z-score -4.621). Overall, these results are further confirming our previous observations of increased proliferation in the S7 population (see also KI67 in Fig. 4c). This biological behaviour is in clear contrast to the well-known low proliferation index characterizing the human islet, further confirming the immaturity of the S7 population.

The analysis also predicted the upregulation of key developmentally specific canonical pathways, such as Wnt/ $\beta$ -catenin signalling (p-value 2.7E-06, z-score 0.905) and PCP signalling (p-value 4.75E-05, z-score 2.558) (Fig. 5f). The strong canonical and non-canonical Wnt pathway activation in the S7 cells, as compared to the human islets is expected to affect the differentiation and maturation potential of insulin<sup>+</sup> cells *in vitro*, its modulation being a potential target for improving the efficiency of the current differentiation protocols.

## Discussion

In this study, we demonstrate that, with the experimental set-up described, insulin<sup>+</sup>, MAFA<sup>+</sup>, NKX6.1<sup>+</sup> S7 cells can be generated *in vitro* from hiPSCs regardless of *HNF4A* mutation carrier status or diabetes status. Furthermore, also regardless of *HNF4A* mutation carrier status or diabetes status, the S7 cells displayed the same GSIS kinetics, the same ultrastructural features and the same levels of proteomics-based protein markers of a mature  $\beta$ -cell, including INS, PDX1, NKX6.1, MAFA, GLUT2, UCN3 and others. These data are consistent





with observations from hiPSC-derived  $\beta$ -cells from patients with type 1 diabetes<sup>20</sup>. Milleman *et al.*, reports that using this model system, upon  $\beta$ -cell stress, the T1D-patient derived hiPSCs do not respond differently from non-diabetic hiPSC-derived  $\beta$ -cells. Since reprogramming into iPSC resets aspects of cellular and molecular aging<sup>21</sup>, progerin-induced aging of the *in vitro* generated iPSC-derived patient-specific cells could be one strategy to uncover disease specific traits<sup>22</sup>. Taken together, the above indicate that the trigger of the MODY1 phenotype/ $\beta$ -cell dysfunction acts downstream of  $\beta$ -cell differentiation, probably after the  $\beta$ -cells have been formed, suggesting mechanisms such as cell death upon stress or low clearance of the senescence markers. Nevertheless, as the *HNF4A* mutation described in this manuscript, p.Ile271fs, is not a null mutation, but rather a non-sense mutation generating a truncated HNF4A product, we cannot exclude an even stronger impact on  $\beta$ -cell differentiation in the case of the absence of a full allele also affecting differentiating events. Interestingly, there are no known human HNF4A null mutations known to date. Moreover, the available HNF4 $\alpha$  general knockout murine model is embryonic lethal<sup>23</sup>. Taken together, these two observations suggest that the loss of a full *HNF4A* allele is probably embryonically lethal. However, if the loss of *HNF4A* is occurring in adult  $\beta$ -cells, by using a conditional knockout mouse, then the mice are viable but exhibit impairment of glucose-stimulated insulin secretion<sup>24</sup>.

Importantly, we find that the differentiated S7 cell population is characterized by a high cellular heterogeneity. Regardless of the *HNF4A* mutation and diabetes disease status, the S7 populations were characterized by a mix of different stages of progenitor cells, including mixed islet cell populations, as well as cells expressing different hormone combinations (including bihormonal cells). Mixed islet cell populations could also partly explain our observation that insulin secretion was seemingly reduced at high glucose concentrations compared to low glucose concentration since a subgroup of glucose-stimulated somatostatin secreting cells (given the observed somatostatin expression; Fig. 2c, and in accordance with a previously reported subpopulation of somatostatin+ S7 cells<sup>9</sup>) could potentially inhibit insulin secretion. Moreover, both the proteomic analysis and physiological tests showed that the generated insulin<sup>+</sup> cells are immature and require further maturation to become fully functional  $\beta$ -like cells. Both the heterogeneity of the S7 fraction and the immaturity of the *in vitro* generated insulin-producing cells are not surprising and have been previously reported in most widely used  $\beta$ -cells differentiation protocols. These problems may, however, affect safety issues such as a potential carcinogenesis risk conferred from the multiprogenitor potential in clinical therapeutic studies aimed at transplanting S7 cells as  $\beta$ -cell therapy. Consequently, further efforts are required for refining the current *in vitro* strategies towards generating  $\beta$ -like cells with physiological GSIS kinetics, similar to the mature bona-fide  $\beta$ -cells.

Surprisingly, transplantation of *in vitro* differentiating cells into immunodeficient mice promotes further their *in vivo* differentiation and maturation, and has proven to be successful for curing induced diabetes in mice<sup>25</sup>. This observation suggests the presence of *in vivo* maturation factor(s) missing from the current  $\beta$ -cell differentiation protocols. To reveal these missing molecular components, we performed a global proteome analysis of immature (S6) and maturing (S7) cells derived from MODY1 mutation carriers and a family control and compared to the human islet proteome. Many studies have been performed to generate reference maps of the human pancreatic islet proteome<sup>26</sup>. Previously, the most comprehensive human pancreatic islet study identified 3,365 proteins<sup>27</sup>. By mapping the global human islet cell proteome, we quantified 8,800 proteins. To our knowledge this is the largest number of identified proteins reported for human pancreatic islets, and these data have provided us with unparalleled insights into the specific proteome landscapes of pancreatic islets. There is a potential bias in our analyses introduced by heterogeneity of varying proportions in the cellular mix of the clones vs islets. In addition bias may be introduced by islets have natural variability in the proportions of different subpopulations of cells and by contamination of human islet preparations with acinar and ductal structures as well as bias introduced by stress associated with islet isolation<sup>9</sup>.

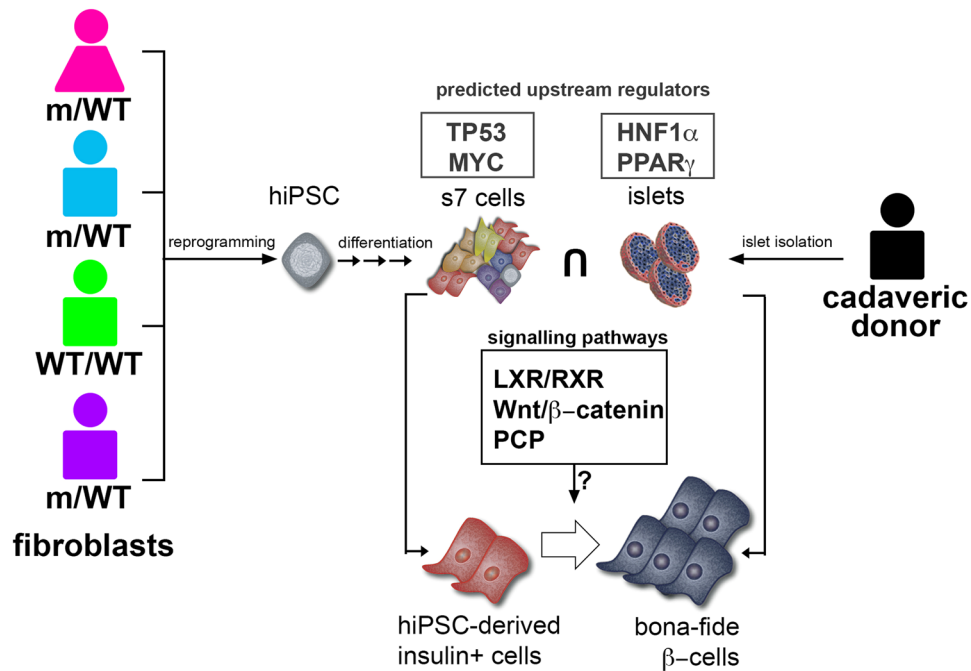
The comparative analysis of S7 cells from the MODY1 quadro to human islets revealed a missing molecular landscape involved in  $\beta$ -cell fate (Fig. 6). Interestingly, we found differences in the proteomic landscape under the presumptive regulation of *HNF1A* (particular to islets). Moreover, we also identified several signalling pathways with known  $\beta$ -cell developmental and differentiation potential in the human islets samples, which exhibited a differentially regulated pattern in the S7-cells fraction, regardless of the MODY1 mutation status. One such example is the predicted upregulation of the canonical and non-canonical Wnt pathways in the S7 cells of all MODY1 quadro samples as compared to the human islets. The modulation of these pathways<sup>28</sup> is expected to affect the differentiation and maturation potential of insulin<sup>+</sup> cells *in vitro*, thereby uncovering potential targets for improving the efficiency of generating  $\beta$ -like cells *in vitro*.

## Materials and Methods

**Cell source.** Human islets (n = 3) were isolated as described previously<sup>29</sup> from pancreata obtained from two female and one male brain-dead deceased donors after informed consent from relatives for organ donation and for use in research at the islet isolation facility of Oslo University Hospital, Oslo, Norway. In this study, we present data from two human islets samples, one of the samples were excluded due to technical issues. All experiments with human islets was approved by the regional committee for medical and health research ethics central in Norway (2011/782).

The human induced pluripotent stem cells iPSCs used in this study were episomal reprogrammed from four family members with MODY1 diabetes, including a healthy family control. The iPSCs cell lines were characterized and confirmed to have normal karyotype, and to be mycoplasma free using a MycoAlert Mycoplasma Detection Kit (Lonza, LT07-418). All hiPS cell lines were subjected to Stage-specific embryonic antigen 4 (SSEA4) enrichment (SSEA4 microbeads, MACS Miltenyi Biotec) before proceeding to differentiation.

All samples used for differentiation and further analyses throughout the manuscript were generated from a MODY1 quadro, i.e. three *HNF4A* mutation carriers (two with overt diabetes and one non-diabetic mutation carrier) and their family control (healthy male sibling without mutation).



**Figure 6.** Comparison of hiPSC-derived insulin-producing cells and human islets suggesting missing molecular landscape involved in  $\beta$ -cell fate. For the differentiation of patient-specific S7 insulin-producing cells, we reprogrammed skin fibroblasts from four donors from the same family ( $n = 4$ ), including a parental diabetic mutation carrier, and three offspring; one diabetic mutation carrier, one non-diabetic mutation carrier, and non-diabetic non-mutation carrier (family control). We found that positive *HNF4A* mutation status does not prevent the formation of insulin positive S7 cells. However, in agreement with previous reports, these S7 cells are immature and require further maturation to become fully functional  $\beta$ -like cells. Specifically addressing this absent maturation of S7 cells, we show that the careful assessment of the proteome of these differentiating cells compared to the human islet proteome can be exploited to detect factors that are deficient or overexpressed, factors that might enable the *in vitro* maturation towards fully functional  $\beta$ -cells. The human islets specific proteome landscape predicted HNF1 $\alpha$  and PPAR $\gamma$  as upstream regulators important for  $\beta$ -cell identity and function. A similar analysis of the S7-cells-specific proteome landscape predicted TP53 and MYC as upstream regulators transcription factors involved in cell cycle progression. These results suggest that the S7 cells contain a significant fraction of proliferating cells. Moreover, as indicated by the human islets specific proteome landscape analysis, the S7 cells are also missing critical networks required for  $\beta$ -cell identity and function.

All experiments with hiPSCs were approved by the Regional Committee of Medical and Health Research Ethics (REK 2010/2295), and all methods were performed in accordance with the Helsinki Declaration.

**In vitro differentiation protocol.** Four iPSC lines (three *MODY1* mutation carriers and their family control) were differentiated according to the protocol<sup>9</sup> on Matrigel-coated plates for the duration of the whole differentiation protocol, without transfer to an air-liquid interface. The cells were grown in stage 6 media and stage 7 media for 7 days, respectively. iPSC-derived cells were harvested for downstream immunofluorescence staining at each of the seven stages of differentiation. We harvested iPSC-derived cells for proteomics analysis at Stage 6 ( $n = 2$ ) and 7 ( $n = 4$ ) corresponding to immature and maturing beta cells, respectively. Stage 7 cells were further subjected to Glucose Stimulated Insulin Secretion (GSIS) and Transmission Electron Microscopy (TEM).

**Static incubation glucose-stimulated insulin secretion Stage 7 cells.** Stage 7 cells were rinsed twice with Krebs buffer (129 mM NaCl, 5 mM NaHCO<sub>3</sub>, 4.8 mM KCl, 2.5 mM CaCl<sub>2</sub>, 1.2 mM MgSO<sub>4</sub>, 1 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.2 mM KH<sub>2</sub>PO<sub>4</sub>, 10 mM HEPES, 0.1% BSA) in deionized water and then sterile filtered. The cells were pre-incubated in Krebs buffer for 30 min. Followed by incubation in Krebs buffer spiked with 1.67 mM glucose (Krebs buffer low glucose) for 40 min. Media was collected after 20 min and 40 min, centrifuged to remove unwanted cells and supernatant was transferred to a new vial. After each incubation, we added the same amount of fresh media to the cells. After low glucose incubation, the cells were incubated in Krebs buffer spiked with 20 mM glucose (Krebs buffer high glucose) for 60 min, and media was collected every 10 min. Followed by incubation in Krebs buffer spiked with 20 mM glucose and 10 nM Exendin-4 for 20 min, with media collection after 10 min and 20 min. As a last stage of stimulation of insulin secretion, cells were incubated in Krebs buffer spiked with 20 mM glucose and 30 mM KCl for 20 min, we collected media after 10 min and 20 min. All supernatants were frozen at  $-80^{\circ}\text{C}$ , and Ultrasensitive C-peptide levels (#10-1141-01; Mercodia), Insulin (#10-113-01; Mercodia) were measured with ELISA. Proinsulin (#10-1118-01; Mercodia) was measured as a negative control, to check that cells were removed from the media.

Primary antibody	Dilution	Company	Catalog #
rabbit anti-human FOXA2	1/200	Abcam	ab23630
goat anti-human Sox17	1/40	R&D Systems	AF1924
mouse anti-human HNF-4 $\alpha$	1/1000	Abcam	ab41898
goat anti-human HNF-1 $\beta$	1/20	R&D Systems	AF3330
Rabbit anti-human MAFA	1/200	Abcam	ab98859
rabbit anti-human NKX6.1	1/100	Novus	NBP1-82553
guinea-pig anti-porcine insulin	1/400	Dako	A056401-2
mouse anti-porcine glucagon	1/1000	Sigma	G2654

**Table 1.** Primary antibodies.

Secondary antibody	Dilution	Company
donkey anti-rabbit A488	1/500	Molecular Probes
goat anti-guinea-pig A488	1/500	Molecular Probes
donkey anti-rabbit A546	1/500	Molecular Probes
donkey anti-mouse A594	1/500	Molecular Probes
donkey anti-goat A647	1/500	Molecular Probes
goat anti-guinea-pig A647	1/500	Molecular Probes

**Table 2.** Secondary antibodies.

**Fixation and embedding of cells for Transmission electron microscopy (TEM).** Cells were fixed in 2% glutaraldehyde in 0.1 M sodium cacodylate buffer for two hours, followed by washing with 0.1 M sodium cacodylate buffer ( $3 \times 15$  min). Post fixation was done in 1% OsO<sub>4</sub> 60 min, followed by washing with 0.1 M sodium cacodylate buffer ( $2 \times 10$  min). Dehydration in a gradient of 30%, 50%, and 70% ethanol for 15–20 min, and 96% followed by 100% ethanol for 20 min. Cells were embedded in 50:50 ethanol/resin and a second round in 100% Resin. Cells were allowed to harden at 60 °C for 48 hours.

**Immunofluorescence.** Cells were cultured on glass coverslips and fixed in 2% PFA for 15 min. The immunofluorescence protocol was performed conform indications provided by the supplier. The antibodies used are given in Tables 1 and 2.

The nuclei were stained with DAPI (D1306, Molecular Probes). The samples mounted in Prolong Diamond Antifade Mountant Media (P36970, Life technologies) and were analyzed with Leica TCS SP2 AOBS or Leica TCS SP5 STED CW confocal microscopes. No specific feature of the original data was obscured, eliminated or misrepresented.

**Global proteomics analysis.** *Cell lysis and protein digestion.* Cell cultures were harvested with Gentle Cell Dissociation Reagent, followed by centrifugation. 200 handpicked equally sized islets were collected in column, washed 1XPBS and islet pellet was stored at  $-70$  °C until lysis. Cells (S6, S7 and islets) were lysed in buffer containing 4% SDS, and boiled at 95 °C for 7 min on a shaker, and sonicated (three rounds a 30 sec, 30% power). The protein concentration was determined using a BCA protein assay kit. Dry aliquots containing an estimated amount of 100  $\mu$ g of proteins were further processed using Filter-Aided Sample Preparation<sup>30</sup>.

*Tandem Mass Tag (TMT) 10-plex labelling.* Our approach using TMT 10-plex labelling allowed for 10 samples to be analysed simultaneously; human islets (n = 3), S6 cells (n = 2), S7 cells (n = 4), and we also included one sample containing a mix of equal amounts from a different MODY family (data not shown). TMT reagents were re-suspended in ACN. Desalted peptides were re-suspended in 100  $\mu$ L of 200 mM HEPES pH 8.5, 30  $\mu$ L of ACN, and 10  $\mu$ L of the TMT reagents were added to the respective peptide samples, gently vortexed, and incubated for 1.5 h at RT. To prevent unwanted labelling, the reaction was quenched by adding 10  $\mu$ L of 5% hydroxylamine and incubated for 15 min at RT. Equal amounts of the TMT-labelled samples were combined and concentrated to near dryness, followed by desalting via C18 solid phase extraction. TMT 10-plex labelling allows for relative quantification of peptides across ten samples. The intensity (abundance) of a labelled peptide in one sample is quantitatively compared to the intensity of the same labelled peptide in ten samples in parallel.

*Off-line basic pH reversed phase fractionation.* The combined labelled peptide samples were pre-fractionated by basic pH reversed phase HPLC as described previously<sup>31</sup>, using an Agilent (P/N 770995-902) 300Extend-C18, 5  $\mu$ m, 250 mm  $\times$  4.6 mm id column, connected to an Agilent Technology off-line LC-system. Solvent A was 5% ACN, 10 mM NH<sub>4</sub>HCO<sub>3</sub> pH8, and solvent B was 90% ACN, NH<sub>4</sub>HCO<sub>3</sub> pH 8. The samples were re-suspended in 500  $\mu$ L solvent A and loaded onto the column. Column flow was set to 0.8 mL/min and the gradient length was 70 min, as follows: from 0–35 min solvent 50% A/ 50% B, and from 35–50 min 100% B, and from 50–70 min 100% A. The labelled peptides were fractionated into 96 fractions, and further combined into a total of 12 fractions.

Each fraction was acidified with 1% formic acid, concentrated by vacuum centrifugation to near dryness, and desalted by StageTip. Each fraction was dissolved in 5% ACN/ 5% formic acid for LC-MS/MS analysis.

**LC-MS3 analysis.** From each of the 12 fractions, ~5 µg was dissolved in 1% aqueous formic acid (FA) prior to LC-MS/MS analysis on an Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA) coupled to a Proxeon EASY-nLC 1000 liquid chromatography (LC) pump (Thermo Fisher Scientific). Peptides were fractionated on a 75-µm inner diameter microcapillary column packed with ~0.5 cm of Magic C4 resin (5 µm, 100 Å, Michrom Bioresources) followed by ~35 cm of GP-18 resin (1.8 µm, 200 Å, Sepax, Newark, DE). For each analysis, we loaded ~1 µg onto the column.

Peptides were separated using a 3 hr gradient of 6 to 26% acetonitrile in 0.125% formic acid at a flow rate of ~350 nL/min. Each analysis used the multi-notch MS3-based TMT method<sup>32</sup> on an Orbitrap Fusion mass spectrometer, which has been shown to reduce ion interference compared to MS2 quantification<sup>31</sup>. The scan sequence began with an MS1 spectrum (Orbitrap analysis; resolution 120,000; mass range 400–1400 m/z; automatic gain control (AGC) target  $2 \times 10^5$ ; maximum injection time 100 ms). Precursors for MS2/MS3 analysis were selected using a TopSpeed of 2 sec. MS2 analysis consisted of collision-induced dissociation (quadrupole ion trap analysis; AGC  $4 \times 10^3$ ; normalized collision energy (NCE) 35; maximum injection time 150 ms). Following acquisition of each MS2 spectrum, we collected an MS3 spectrum using our recently described method<sup>32</sup> in which multiple MS2 fragment ions were captured in the MS3 precursor population using isolation waveforms with multiple frequency notches. MS3 precursors were fragmented by high-energy collision-induced dissociation (HCD) and analysed using the Orbitrap (NCE 55; AGC  $5 \times 10^4$ ; maximum injection time 150 ms, resolution was 60,000 at 400 Th).

**Data analysis.** Mass spectra were processed using a Sequest-based in-house software pipeline<sup>33</sup>, and spectra were converted to mzXML using a modified version of ReAdW.exe. Database searching included all entries from the human uniprot database (March 11, 2014). This database was concatenated with one composed of all protein sequences in the reversed order. Searches were performed using a 50 ppm precursor ion tolerance for total protein level analysis. The product ion tolerance was set to 0.9 Da. These wide mass tolerance windows were chosen to maximize sensitivity in conjunction with Sequest searches and linear discriminant analysis<sup>33,34</sup>. TMT tags on lysine residues and peptide N termini (+229.163 Da) and carbamidomethylation of cysteine residues (+57.021 Da) were set as static modifications, while oxidation of methionine residues (+15.995 Da) was set as a variable modification.

Peptide-spectrum matches (PSMs) were adjusted to a 1% false discovery rate (FDR)<sup>35,36</sup>. PSM filtering was performed using a linear discriminant analysis, as described previously<sup>33</sup>, while considering the following parameters: XCorr, ΔCn, missed cleavages, peptide length, charge state, and precursor mass accuracy. For TMT-based reporter ion quantitation, we extracted the summed signal-to-noise (S/N) ratio for each TMT channel and found the closest matching centroid to the expected mass of the TMT reporter ion.

The search space for each reporter ion was limited to a range of 0.003 Th to prevent overlap between the isobaric reporter ions. For protein-level comparisons, PSMs were identified, quantified, and collapsed to a 1% peptide false discovery rate (FDR) and then collapsed further to a final protein-level FDR of 1%. Moreover, protein assembly was guided by principles of parsimony to produce the smallest set of proteins necessary to account for all observed peptides.

Proteins were quantified by summing reporter ion counts across all matching PSMs using in-house software, as described previously<sup>33</sup>. PSMs with poor quality, MS3 spectra with more than eight TMT reporter ion channels missing, MS3 spectra with TMT reporter summed signal-to-noise ratio that is less than 100, or no MS3 spectra were excluded from quantitation<sup>37</sup>. Protein quantitation values were exported for further analysis in Microsoft Excel or SAS JMP. Each reporter ion channel was summed across all quantified proteins and normalized assuming equal protein loading of all 10 samples.

For all the global proteome comparisons referred to throughout the manuscript we used samples from MODY1 mutation carriers and their family control and compared with human islet samples.

The pathway analyses were generated through the use of QIAGEN's Ingenuity Pathway Analysis (IPA®), QIAGEN Redwood City, www.qiagen.com/ingenuity). Briefly, the analyses were performed with the following settings: Expression Value Type (Exp Log Ration), Reference set (Ingenuity Knowledge Base + Endogenous chemicals), Relationships to consider (Direct and Indirect Relationships), Interaction networks (70 molecules/network; 25 networks/analysis), Data Source (all), Confidence (Experimentally Observed), Species (Human, Mouse, Rat), Tissue & Cell Lines (all), Mutations (all).

We applied non-parametric tests to assess for significant differences in protein abundance in the clones or islets. We chose a significance level of 5% and used GraphPad Prism (GraphPad software, version 7, La Jolla, CA).

## References

- Horowitz, M. P. & Greenamyre, J. T. Gene-environment interactions in Parkinson's disease: the importance of animal modeling. *Clinical pharmacology and therapeutics* **88**, 467–474 (2010).
- Sabatino, D. E. *et al.* Novel hemophilia B mouse models exhibiting a range of mutations in the Factor IX gene. *Blood* **104**, 2767–2774 (2004).
- Magen, I. & Chesselet, M. F. Genetic mouse models of Parkinson's disease The state of the art. *Progress in brain research* **184**, 53–87 (2010).
- Yamagata, K. *et al.* Overexpression of dominant-negative mutant hepatocyte nuclear factor-1 alpha in pancreatic beta-cells causes abnormal islet architecture with decreased expression of E-cadherin, reduced beta-cell proliferation, and diabetes. *Diabetes* **51**, 114–123 (2002).
- Raeder, H. *et al.* Absence of diabetes and pancreatic exocrine dysfunction in a transgenic model of carboxyl-ester lipase-MODY (maturity-onset diabetes of the young). *PLoS One* **8**, e60229 (2013).



6. Reznia, A. *et al.* Production of functional glucagon-secreting alpha-cells from human embryonic stem cells. *Diabetes* **60**, 239–247 (2011).
7. Pagliuca, F. W. *et al.* Generation of functional human pancreatic beta cells *in vitro*. *Cell* **159**, 428–439 (2014).
8. Russ, H. A. *et al.* Controlled induction of human pancreatic progenitors produces functional beta-like cells *in vitro*. *EMBO J* **34**, 1759–1772 (2015).
9. Reznia, A. *et al.* Reversal of diabetes with insulin-producing cells derived *in vitro* from human pluripotent stem cells. *Nat Biotechnol* **32**, 1121–1133 (2014).
10. Agulnick, A. D. *et al.* Insulin-Producing Endocrine Cells Differentiated *In Vitro* From Human Embryonic Stem Cells Function in Macroencapsulation Devices *In Vivo*. *Stem Cells Transl Med* **4**, 1214–1222 (2015).
11. Teo, A. K. *et al.* Early Developmental Perturbations in a Human Stem Cell Model of MODY5/HNF1B Pancreatic Hypoplasia. *Stem Cell Reports* **6**, 357–367 (2016).
12. Hua, H. *et al.* iPSC-derived beta cells model diabetes due to glucokinase deficiency. *The Journal of clinical investigation* **123**, 3146–3153 (2013).
13. Blum, B. *et al.* Functional beta-cell maturation is marked by an increased glucose threshold and by expression of urocortin 3. *Nature biotechnology* **30**, 261–264 (2012).
14. Kroon, E. *et al.* Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells *in vivo*. *Nature biotechnology* **26**, 443–452 (2008).
15. Taylor, R. C. *et al.* Changes in translational efficiency is a dominant regulatory mechanism in the environmental response of bacteria. *Integr Biol (Camb)* **5**, 1393–1406 (2013).
16. Dressaire, C. *et al.* Transcriptome and proteome exploration to model translation efficiency and protein stability in *Lactococcus lactis*. *PLoS Comput Biol* **5**, e1000606 (2009).
17. Rodnina, M. V. The ribosome in action: Tuning of translational efficiency and protein folding. *Protein Sci* **25**, 1390–1406 (2016).
18. Hang, Y. & Stein, R. MafA and MafB activity in pancreatic beta cells. *Trends in endocrinology and metabolism: TEM* **22**, 364–373 (2011).
19. van der Meulen, T. *et al.* Urocortin 3 marks mature human primary and embryonic stem cell-derived pancreatic alpha and beta cells. *PLoS one* **7**, e52181 (2012).
20. Millman, J. R. *et al.* Generation of stem cell-derived beta-cells from patients with type 1 diabetes. *Nature communications* **7**, 11463 (2016).
21. Vera, E. & Studer, L. When rejuvenation is a problem: challenges of modeling late-onset neurodegenerative disease. *Development* **142**, 3085–3089 (2015).
22. Miller, J. D. *et al.* Human iPSC-based modeling of late-onset disease via progerin-induced aging. *Cell Stem Cell* **13**, 691–705 (2013).
23. Chen, W. S. *et al.* Disruption of the HNF-4 gene, expressed in visceral endoderm, leads to cell death in embryonic ectoderm and impaired gastrulation of mouse embryos. *Genes Dev* **8**, 2466–2477 (1994).
24. Miura, A. *et al.* Hepatocyte nuclear factor-4alpha is essential for glucose-stimulated insulin secretion by pancreatic beta-cells. *The Journal of biological chemistry* **281**, 5246–5257 (2006).
25. Reznia, A. *et al.* Maturation of human embryonic stem cell-derived pancreatic progenitors into functional islets capable of treating pre-existing diabetes in mice. *Diabetes* **61**, 2016–2029 (2012).
26. Crevecoeur, I., Rondas, D., Mathieu, C. & Overbergh, L. The beta-cell in type 1 diabetes: What have we learned from proteomic studies? *Proteomics. Clinical applications* **9**, 755–766 (2015).
27. Metz, T. O. *et al.* Characterization of the human pancreatic islet proteome by two-dimensional LC/MS/MS. *Journal of proteome research* **5**, 3345–3354 (2006).
28. Bader, E. *et al.* Identification of proliferative and mature beta-cells in the islets of Langerhans. *Nature* **535**, 430–434 (2016).
29. Friberg, A. S., Stahle, M., Brandhorst, H., Korsgren, O. & Brandhorst, D. Human islet separation utilizing a closed automated purification system. *Cell transplantation* **17**, 1305–1313 (2008).
30. Wisniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nature methods* **6**, 359–362 (2009).
31. Paulo, J. A. & Gygi, S. P. A comprehensive proteomic and phosphoproteomic analysis of yeast deletion mutants of 14-3-3 orthologs and associated effects of rapamycin. *Proteomics* **15**, 474–486 (2015).
32. McAlister, G. C. *et al.* MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes. *Analytical chemistry* **86**, 7150–7158 (2014).
33. Huttlin, E. L. *et al.* A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143**, 1174–1189 (2010).
34. Beausoleil, S. A., Villen, J., Gerber, S. A., Rush, J. & Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nature biotechnology* **24**, 1285–1292 (2006).
35. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol Biol* **604**, 55–71 (2010).
36. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4**, 207–214 (2007).
37. McAlister, G. C. *et al.* Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal Chem* **84**, 7469–7478 (2012).

## Acknowledgements

We thank Anne Hammer Knudsen for technical assistance. We also thank Bergen Stem Cell Consortium (BSCC) for support. The TEM imaging was performed at the Molecular Imaging Centre (MIC) and was thus supported by the Department of Biomedicine and the Faculty of Medicine and Dentistry, at the University of Bergen, and its partners. This work was funded with grants from Bergen Forskningsstiftelse (BFS), the Western Norway Regional Health Authority, Novo Nordisk Foundation and Diabetesforbundet to H.R. S.C. was supported by grants from Research Council of Norway (NFR) and the Novo Nordisk Foundation. This work was funded in part by a National Institute of Health/NIDDK grant K01 DK098285 (J.A.P.).

## Author Contributions

H.V., Y.B., L.G., and S.C. contributed to conception and design of experiments. H.V. and Y.B. were primary responsible for the *in vitro* differentiation, GSIS as well as sample preparation for downstream mass spectrometry analysis. L.G. and S.C. were primary responsible for immunofluorescence. J.A.P. contributed with the TMT-labelling experiment and mass spectrometry analysis. S.P.G. contributed with mass spectrometry analysis. H.S. provided the human islets. H.V., L.G. and S.C. analysed the data and wrote the manuscript. H.R. conceived the experiments, interpreted the observations and wrote the manuscript. All authors approved the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at doi:[10.1038/s41598-017-04979-w](https://doi.org/10.1038/s41598-017-04979-w)

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017