

Soren K. Thomsen,<sup>1</sup> Alessandro Ceroni,<sup>2</sup> Martijn van de Bunt,<sup>1,3</sup> Carla Burrows,<sup>1</sup> Amy Barrett,<sup>1</sup> Raphael Scharfmann,<sup>4</sup> Daniel Ebner,<sup>2</sup> Mark I. McCarthy,<sup>1,3,5</sup> and Anna L. Gloyn<sup>1,3,5</sup>



## Systematic Functional Characterization of Candidate Causal Genes for Type 2 Diabetes Risk Variants



Diabetes 2016;65:3805–3811 | DOI: 10.2337/db16-0361

**Most genetic association signals for type 2 diabetes risk are located in noncoding regions of the genome, hindering translation into molecular mechanisms. Physiological studies have shown a majority of disease-associated variants to exert their effects through pancreatic islet dysfunction. Systematically characterizing the role of regional transcripts in  $\beta$ -cell function could identify the underlying disease-causing genes, but large-scale studies in human cellular models have previously been impractical. We developed a robust and scalable strategy based on arrayed gene silencing in the human  $\beta$ -cell line EndoC- $\beta$ H1. In a screen of 300 positional candidates selected from 75 type 2 diabetes regions, each gene was assayed for effects on multiple disease-relevant phenotypes, including insulin secretion and cellular proliferation. We identified a total of 45 genes involved in  $\beta$ -cell function, pointing to possible causal mechanisms at 37 disease-associated loci. The results showed a strong enrichment for genes implicated in monogenic diabetes. Selected effects were validated in a follow-up study, including several genes (*ARL15*, *ZMIZ1*, and *THADA*) with previously unknown or poorly described roles in  $\beta$ -cell biology. We have demonstrated the feasibility of systematic functional screening in a human  $\beta$ -cell model and successfully prioritized plausible disease-causing genes at more than half of the regions investigated.**

Type 2 diabetes risk is determined by a complex interplay between environmental and genetic factors, with heritability

estimates ranging from 20% to 80% (1). Over the past decade, genome-wide association studies (GWAS) of ever-increasing size have discovered more than 100 regions of the genome (loci) associated with type 2 diabetes risk (2). Studies in individuals with diabetes have demonstrated that a large number of these association signals exert their effects on disease susceptibility through pancreatic islet dysfunction (3).

Despite these advances, progress in translating genetic findings into disease biology has been relatively slow. The majority of risk variants are located in noncoding regions of the genome and pinpointing the underlying causal genes or “effector transcripts” has proved challenging (4). Recent efforts have focused on identifying structural or functional links between association signals and regional genes (5,6). A complementary strategy uses candidate gene biology to prioritize genes located near association signals. High-throughput screening could facilitate the identification of genes implicated in  $\beta$ -cell function and thereby highlight potential effector transcripts at type 2 diabetes GWAS loci. To date, such approaches have been limited by the inadequacies of available human cellular models and the high cost of insulin immunoassays (~\$2 per data point), the gold standard for measuring insulin. To circumvent these issues, previous studies have relied on rodent  $\beta$ -cell models and either used reporter assays as a proxy for insulin measurements or focused on cellular proliferation (7–11).

Recently, the first glucose-responsive human  $\beta$ -cell line, EndoC- $\beta$ H1, was generated (12,13). The line is derived from

<sup>1</sup>Oxford Centre for Diabetes, Endocrinology & Metabolism, Radcliffe Department of Medicine, University of Oxford, Oxford, U.K.

<sup>2</sup>Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford, U.K.

<sup>3</sup>Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, U.K.

<sup>4</sup>INSERM U1016, Institut Cochin, Université Paris Descartes, Paris, France

<sup>5</sup>National Institute for Health Research Oxford Biomedical Research Centre, Churchill Hospital, Oxford, U.K.

Corresponding author: Anna L. Gloyn, [anna.gloyn@drl.ox.ac.uk](mailto:anna.gloyn@drl.ox.ac.uk).

Received 17 March 2016 and accepted 18 August 2016.

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db16-0361/-/DC1>.

© 2016 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

See accompanying article, p. 3541.

fetal pancreatic buds matured in vivo and displays modest but robust induction of insulin secretion in response to glucose and secretagogues. Detailed characterizations have shown the cell line to be an authentic model system for studying stimulus-coupled secretion (14–16).

To accelerate the discovery of causal genes for type 2 diabetes, the current study performed and validated a genetic screen in the EndoC- $\beta$ H1 cell line. We identified genes at half of the type 2 diabetes-associated loci studied (37/75) where small interfering RNA (siRNA)-mediated silencing resulted in  $\beta$ -cell dysfunction. This demonstrates the feasibility of performing systematic screening for insulin secretion in a human  $\beta$ -cell model, with implications for both high-throughput genetic and chemical compound screening. Our results can be integrated with existing lines of evidence to prioritize effector transcripts at GWAS loci and highlight potential roles for *ARL15*, *ZMIZ1*, and *THADA* in the regulation of insulin secretion.

## RESEARCH DESIGN AND METHODS

### RNA-seq

The EndoC- $\beta$ H1 cell line was cultured as previously described and grown to near confluency (12). RNA was then TRIzol extracted and sequenced at the Oxford Genomics Centre (Wellcome Trust Centre for Human Genetics, University of Oxford) (see Supplementary Fig. 4 for details). The raw sequencing data have been deposited at the European Nucleotide Archive (<http://www.ebi.ac.uk/ena>) under accession number PRJEB15283.

### Cellular Assays

Cellular phenotypes were adapted for automated screening on a PerkinElmer Janus liquid handling workstation based on previously described assays (Supplementary Fig. 1A) (17). Briefly, 20,000 cells/well were reverse transfected in 96-well format at final siRNA concentrations of 25 nmol/L preincubated with 0.2  $\mu$ L RNAiMAX in Opti-MEM. Custom libraries of siRNAs (ON-TARGETplus SMARTpools [Dharmacon] for the primary screen and Silencer Select [Thermo Fisher Scientific] for follow-up validation) were designed based on criteria described in Supplementary Table 2. In each case, nontargeting (NT) sequences based on the same chemistries were used as negative controls. Three days after transfection, cells were starved overnight in complete media containing 2.8 mmol/L glucose followed by 1-h starvation in 0 mmol/L media. Static insulin secretion assays were then performed for 1 h in complete media under the indicated conditions, after which cells were counted as described below.

### Sample Analysis

Following secretion assays, supernatants were analyzed for insulin content using AlphaLISA Human Insulin Immunoassays (PerkinElmer) on a PHERAstar FS plate reader (BMG Labtech). Supernatants (50–250 nL) and beads pre-diluted in water (500 nL) were dispensed into 384-shallow well microplates with an Echo 550 (Labcyte) acoustic liquid

handler before manual addition of immunoassay buffer to a final volume of 5  $\mu$ L. Cell counts were measured using the CyQUANT Direct Cell Proliferation kit (Thermo Fisher Scientific) on an EnVision plate reader (PerkinElmer). All responses were normalized as indicated (see relevant figure legends) and expressed as a percentage of NT control for each phenotype. Effect sizes are given as the percentage difference from NT ( $Response_{Gene} - Response_{NT}$ ) and the absolute values hereof ( $|Response_{Gene} - Response_{NT}|$ ).

### Statistical Analysis

Data analysis was performed using R 3.0.2. To identify significant responses, cell counts and normalized insulin secretion measurements for each gene were compared with NT control using Student two-sample *t* test. The false discovery rate (FDR) was controlled at 5% by applying the Benjamini-Hochberg procedure to produce adjusted *P* values (*q* values) for each phenotype. The Z-factor measuring the control response for each phenotype was calculated as

$$Z' = 1 - \frac{3(\sigma_{INS} + \sigma_{NT})}{|\mu_{INS} - \mu_{NT}|}$$

## RESULTS

We first developed an automated assay for disease-relevant phenotypes in the human  $\beta$ -cell line EndoC- $\beta$ H1 (Supplementary Fig. 1A). Selected targets were silenced in a parallel format using RNA interference. Cells were then assessed for effects on cell number and insulin secretion under four different conditions: low glucose (1 mmol/L), high glucose (20 mmol/L), and high glucose with the sulfonylurea tolbutamide (100  $\mu$ mol/L) or with the phosphodiesterase inhibitor IBMX (100  $\mu$ mol/L). Low- and high-glucose conditions were included to provide information on the effect of gene silencing under conditions representing the fasted and fed states in vivo. Tolbutamide and IBMX act on the depolarizing and the potentiating pathways of insulin secretion, respectively, and were included to provide additional mechanistic insights through modulation (e.g., synergy or pharmacological rescue) of any primary defects observed in low or high glucose.

To reduce the cost of sample analysis, we made use of acoustic liquid handling to miniaturize insulin immunoassays. This generalizable method enabled us to maintain high sensitivity for insulin measurements (coefficient of variation <3%) (Supplementary Fig. 2A and B), while obtaining a 10-fold reduction in the cost of sample analysis (\$0.20 per data point). Using the insulin gene (*INS*) as a positive control, we confirmed that we were able to robustly detect effects of gene silencing on the phenotypes of our assay (mean  $Z' = 0.6$  across conditions) (Supplementary Fig. 3).

On the basis of this combined analysis and assay pipeline, we designed a primary screen to assess the role of positional candidate genes for type 2 diabetes GWAS loci in  $\beta$ -cell

function (Supplementary Fig. 1B). For target selection, we considered all protein-coding genes located within 1 Mb of a type 2 diabetes association signal. To exclude genes not expressed in our cellular model, we performed whole-genome RNA sequencing of the EndoC- $\beta$ H1. Our expression data strongly correlated with published sequencing data for enriched primary  $\beta$ -cells ( $\rho = 0.78$ ) (Supplementary Fig. 4) and showed robust expression of key  $\beta$ -cell genes (12,18) (Supplementary Table 1). We included only genes expressed in both EndoC- $\beta$ H1 and primary  $\beta$ -cells (Supplementary Table 2), resulting in inclusion of 300 positional candidates from 75 type 2 diabetes GWAS loci.

We next performed our primary screen in triplicate and derived standardized scores for each phenotype. Knockdown was visibly confirmed using *PLK1*, an essential gene, which caused extensive cell death across all conditions. In a representative subset of 16 genes, we assessed knockdown efficiency at the transcript level and found the median residual expression to be 43% (Supplementary Fig. 5), roughly equivalent to monoallelic loss of function. To account for differences in plating efficiency and proliferation, cell counts were used to normalize insulin secretion data on a per-well basis. Two criteria were then applied to identify robust effects ("hits"): 1) an FDR-adjusted  $q$  value  $<0.05$  and 2) an absolute effect size among the top 5% (Supplementary Fig. 6). This identified a total of 67 hits (15 for cell count and 52 for insulin secretion phenotypes) between 45 genes at 37 loci (Table 1).

For cell numbers, effect sizes for each gene were estimated based on 12 independently plated replicates (four conditions in triplicate) and therefore likely represent true differences in cellular proliferation and/or viability rather than random plating effects. Aside from *KIF11*, a gene with a known role in cell division, the largest effect sizes compared with NT control (coefficient of variation = 4% for cell numbers) were observed for *ZMIZ1* ( $-15.2\%$ ;  $q$  value =  $6.5 \times 10^{-5}$ ) and *PRDX3* ( $+16.6\%$ ;  $q$  value =  $9.2 \times 10^{-5}$ ).

For the insulin secretion data, we first performed an enrichment analysis for genes implicated in maturity-onset diabetes of the young (MODY). MODY describes a collection of monogenic subtypes of diabetes characterized by insufficient release or production of insulin. As would be expected for a set of bona fide regulators of  $\beta$ -cell function, we observed a strong enrichment of MODY genes among the significant hits (Fisher exact test,  $P = 5.5 \times 10^{-9}$ ). Aggregating absolute effect sizes for MODY and non-MODY genes revealed this enrichment to be driven by altered insulin secretion and not by effects on cell numbers (Fig. 1).

Further validating our secretion data, we observed strong positive correlations between the normalized responses across conditions ( $P < 2.2 \times 10^{-16}$ ) (Fig. 2) and found 10 of 35 genes to cause significant effects under two or more conditions. This included four known MODY genes and *ZMIZ1*, which, independently of the effect on cell numbers, was one of the strongest hits for reduced insulin secretion ( $q$  value  $<0.01$  for low and high

glucose). Knockdown of the *ABCC8* gene, which encodes a subunit of the ATP-sensitive potassium channel, was found to significantly increase insulin secretion under low-glucose and IBMX stimulation. As expected, the depolarization caused by this was masked under high glucose (as cells are already partially depolarized) and fully rescued by tolbutamide (due to pharmacological depolarization of the cells). The pattern of modulation by secretion conditions can thus be used to pinpoint specific biological pathways affected by gene silencing. To explore the relationship between conditions in greater detail, we performed clustering analysis on Z-scores derived from the normalized secretion values. This revealed high glucose and tolbutamide to be most similar in terms of modulating knockdown effects, with low glucose and tolbutamide being most dissimilar (Supplementary Fig. 7).

Finally, we assessed the contribution of off-target effects by performing a small-scale validation experiment using siRNAs designed with an alternate algorithm. The sequences were confirmed to be different from those of the primary screen and could thus be used to establish the biological relevance of positive hits. We selected eight target genes, representing hits for both positive and negative defects across the four conditions, and confirmed that knockdown efficiency was satisfactory (median residual expression = 19.3%) (Supplementary Fig. 8). Compared with insulin secretion results from the primary screen, we observed an excellent linear correlation ( $\rho = 0.85$ ,  $P = 6.7 \times 10^{-10}$ ) (Supplementary Fig. 9) and 88% directional consistency in normalized responses. The validated hits included several genes with limited prior evidence of a role in the regulation of  $\beta$ -cell function, including *ARL15* and *ZMIZ1*, which were found to significantly reduce insulin secretion across conditions ( $q$  values  $<0.05$ ) (Fig. 3A and B), and *THADA*, which modestly elevated insulin secretion across three conditions, though the effect under low glucose was not observed in the primary screen ( $q$  value =  $5.6 \times 10^{-3}$ ) (Fig. 3C). Interestingly, gene silencing of the known MODY gene *HNF4A* was confirmed to cause a paradoxical increase in insulin secretion across all four conditions tested ( $q$  values  $<0.001$ ) (Fig. 3D).

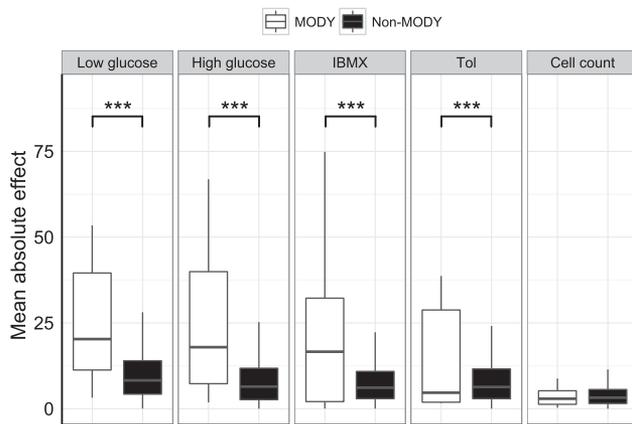
## DISCUSSION

High-throughput screens for  $\beta$ -cell dysfunction offer the opportunity to systematically characterize the role of genes in a disease-relevant tissue for type 2 diabetes. Previous efforts have focused on nonhuman model systems (7–10), reporter-based proxy measurements for insulin (7,8), and/or phenotypes not directly related to insulin production and secretion (10,11). Here, we report a genetic screening strategy for the interrogation of multiple disease-relevant phenotypes in the human  $\beta$ -cell line EndoC- $\beta$ H1. In a primary screen of 300 positional candidates, we successfully identified 15 genes regulating cell number (proliferation and/or viability) and 35 genes regulating insulin secretion. This is, to our knowledge, the first systematic,

**Table 1—Effects of significant hits identified in a primary screen for  $\beta$ -cell dysfunction**

Gene	Locus	Low glucose	High glucose	IBMX	Tolbutamide	Cell count
<i>ABCC8</i>	<i>KCNJ11</i>	48.2*	24.0	26.7*	1.8	-1.4
<i>ADAMTS9</i>	<i>ADAMTS9</i>	6.3	-4.8	2.8	-8.0	12.2*
<i>ADIPOQ</i>	<i>ST64GAL1</i>	87.0*	23.2	23.1	8.8	-7.3
<i>ARL15</i>	<i>ARL15</i>	-5.5	-25.9*	-2.1	-15.5	-1.5
<i>BCAR1</i>	<i>BCAR1</i>	5.9	25.2	28.5*	9.5	-7.0
<i>BCL6</i>	<i>LPP</i>	-20.7*	-8.9	-1.0	-12.0	5.6
<i>BMP8B</i>	<i>MACF1</i>	7.8	16.5	9.5	26.9*	-1.0
<i>CCNT2</i>	<i>TMEM163</i>	-32.6*	1.8	4.7	5.0	-2.8
<i>CDKAL1</i>	<i>CDKAL1</i>	2.4	1.7	-10.6	-23.5*	7.5
<i>DGKQ</i>	<i>MAEA</i>	3.3	17.5	20.4	33.6*	-9.8
<i>DMRTA2</i>	<i>FAF1</i>	32.3*	24.5	13.1	22.7	-0.3
<i>ELAVL4</i>	<i>FAF1</i>	-3.6	9.1	21.2	22.8	-11.4*
<i>ETV5</i>	<i>IGF2BP2</i>	-12.4	-25.6*	-10.9	-12.6	-2.5
<i>FAH</i>	<i>ZFAND6</i>	-23.1*	-20.6*	-16.9	-27.2	-2.3
<i>FBXW7</i>	<i>TMEM154</i>	45.2*	9.6	8.7	11.7	9.9*
<i>GINS4</i>	<i>ANK1</i>	-16.1	-14.2	-9.5	-17.0	8.7*
<i>GLIS3</i>	<i>GLIS3</i>	-13.0	-10.6	6.4	-9.3	-10.3*
<i>HEYL</i>	<i>MACF1</i>	29.2	29.2*	0.4	20.0	-7.2
<i>HMGA2</i>	<i>HMGA2</i>	16.5	4.1	14.2	24.1*	-1.2
<i>HNF1A</i>	<i>HNF1A</i>	21.7	38.3*	23.1	25.5*	5.2
<i>HNF4A</i>	<i>HNF4A</i>	36.7*	66.9*	74.9	92.9*	-8.8
<i>IGF2</i>	<i>DUSP8</i>	-26.3*	-10.4	1.2	0.1	-1.3
<i>INS</i>	<i>DUSP8</i>	-53.5*	-44.8*	-48.7	-38.7*	0.9
<i>KCNK17</i>	<i>KCNK16</i>	-9.4	-17.3	-2.5	-1.2	9.8*
<i>KCTD15</i>	<i>PEPD</i>	21.7	9.4	12.5	0.8	-10.9*
<i>KIF11</i>	<i>HHEX/IDE</i>	45.4*	35.0*	26.9	55.4*	-40.1*
<i>LINGO1</i>	<i>HMG20A</i>	0	19.1	-12.4	-14.9	7.9*
<i>MFGE8</i>	<i>AP3S2</i>	30.0*	3.4	2.8	-1.7	-5.7
<i>MIER3</i>	<i>ANKRD55</i>	5.9	36.5*	5.8	18.2	1.9
<i>NDUFS4</i>	<i>ARL15</i>	1.6	-4.9	3.9	-1.7	10.8*
<i>PABPC1L</i>	<i>HNF4A</i>	-9.7	-12.3	-10.4	-28.8*	-1.2
<i>PHF23</i>	<i>SLC16A11</i>	-25.6*	-1.7	-5.9	3.5	-0.9
<i>PLA2R1</i>	<i>RBMS1</i>	8.6	1.6	10.8	1.6	9.1*
<i>PRDX3</i>	<i>GRK5</i>	24.0	31.7*	23.4	9.6	16.6*
<i>PTHLH</i>	<i>KLHDC5</i>	-2.8	-6.5	-0.5	-25.0*	-5.9
<i>RND3</i>	<i>RND3</i>	-8.7	-6.1	0	-3.0	-14.9*
<i>SLC2A4</i>	<i>SLC16A11</i>	14.5	0	27.2*	18.2	-1.6
<i>SOCS7</i>	<i>HNF1B</i>	3.9	-18.5*	11.2	-14.7	-1.6
<i>SPPL3</i>	<i>HNF1A</i>	-11.9	-21.9*	-6.1	-10.0	-10.8*
<i>STK38L</i>	<i>KLHDC5</i>	15.2	40.9*	4.7	25.2*	-3.0
<i>THADA</i>	<i>THADA</i>	-1.9	6.5	27.5	24.8*	-10.3
<i>TLE1</i>	<i>TLE1</i>	4.3	-5.0	-23.0*	16.2	4.6
<i>TM6SF2</i>	<i>CILP2</i>	-22.6*	-8.3	-0.8	-12.0	-8.7
<i>UPF2</i>	<i>CDC123</i>	7.5	-12.5	4.7	-24.9*	-3.2
<i>ZMIZ1</i>	<i>ZMIZ1</i>	-29.5*	-21.4*	-19.8	-16.8	-15.2*

The table lists effect sizes (% deviation from NT control) for each gene with a least one significant effect across the five phenotypes measured. All insulin secretion measurements were normalized on a per-well basis to cell counts, and the mean percentage deviations from NT control were then calculated for each condition. For cell counts, values were median-normalized for interplate differences, and the mean percentage deviations from NT control were calculated across conditions. \* $q < 0.05$  by Student  $t$  test (FDR-adjusted).



**Figure 1**—Comparing mean absolute effect sizes for MODY and non-MODY genes. Box plots of mean absolute effect sizes for MODY genes and non-MODY genes (excluding controls) across the five phenotypes measured. Effect sizes were calculated as described for Table 1, and the absolute values were then averaged for the two categories of genes. Among 14 identified MODY genes, 8 fulfilled criteria for inclusion in the screen: *HNF4A*, *GCK*, *HNF1A*, *HNF1B*, *PAX4*, *INS*, *ABCC8*, and *KCNJ11*. Tol, tolbutamide. Box plots show median and interquartile ranges for groups of  $n = 8$  and 292 data points. \*\*\* $q$  value  $< 0.001$  by Student  $t$  test (FDR-adjusted).

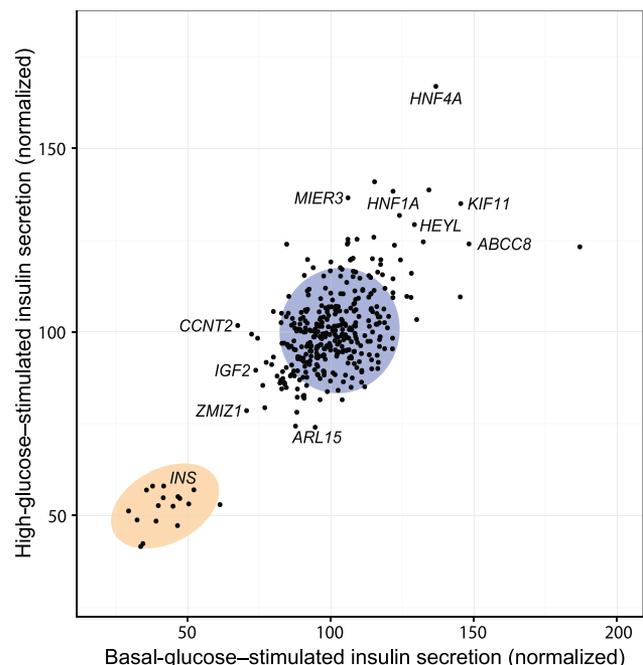
large-scale effort to identify genes involved in insulin secretion. Importantly, the identified hits can be used to prioritize novel effector transcripts for type 2 diabetes GWAS loci and may shed further light on the mechanisms underlying genes previously implicated in  $\beta$ -cell dysfunction.

The known MODY gene *HNF4A* was unexpectedly observed to cause a consistent increase ( $>40\%$ ) in insulin secretion across all conditions. *HNF4A* encodes the transcription factor HNF4 $\alpha$  and is mutated in about 10% of all MODY cases (19). *HNF4A* loss-of-function mutations that cause monogenic diabetes later in life have also been associated with increased birth weight (indicative of increased fetal insulin secretion) and congenital hyperinsulinism in early infancy (20). The underlying reason for this switch from elevated to reduced insulin secretion is unknown, but it has been speculated that gradual  $\beta$ -cell exhaustion or, alternatively, a shift in the modulating cofactors of HNF4 $\alpha$  may underlie this phenomenon (21,22).

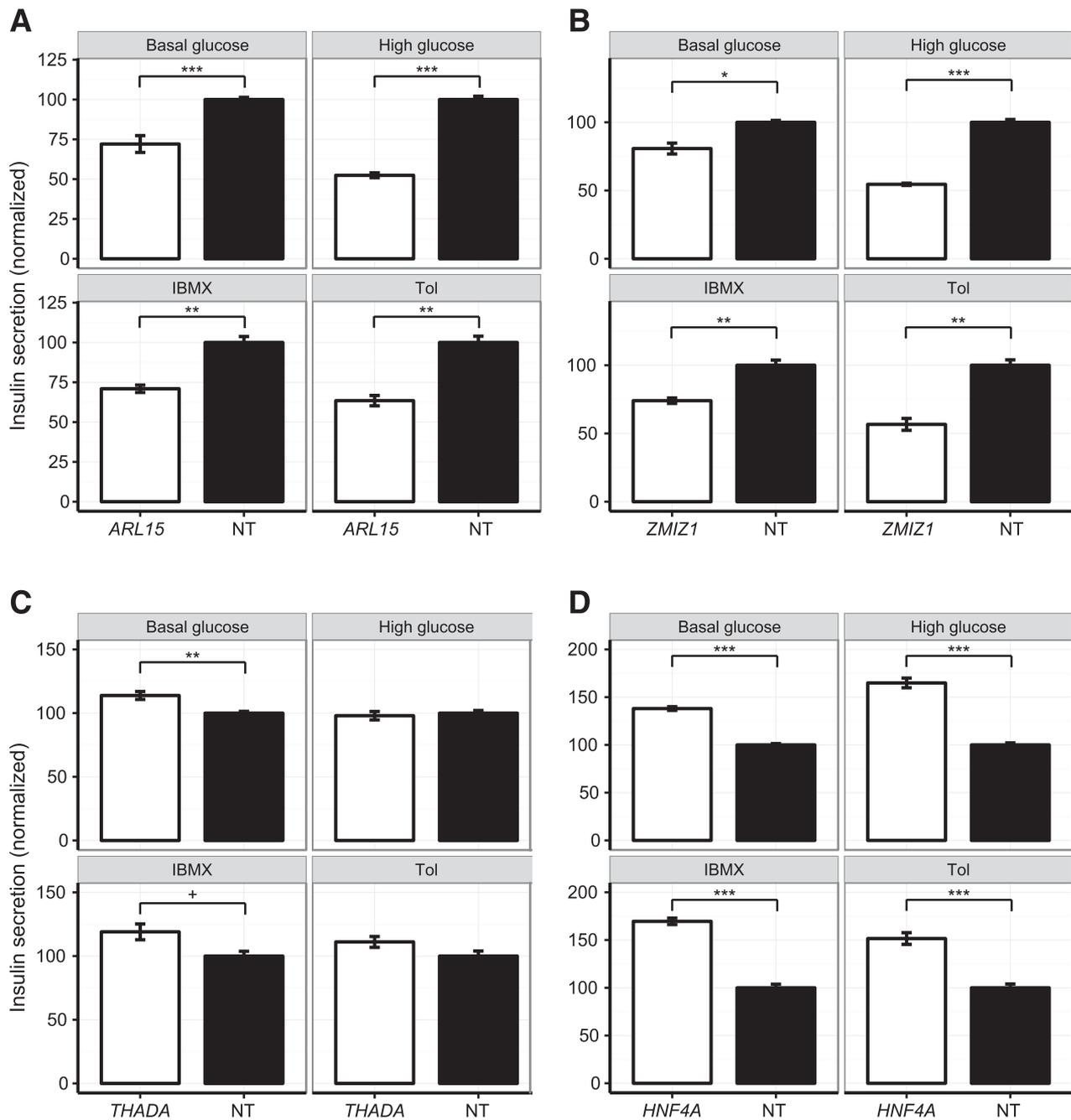
Among the hits with limited prior evidence of a role in  $\beta$ -cell function, we independently validated *ZMIZ1*, *ARL15*, and *THADA*. Overexpression and knockdown of *ZMIZ1*, encoding ZMIZ1, has recently been shown to negatively impact on insulin secretion in primary human islets (6). Moreover, a nearby type 2 diabetes association signal overlaps a *cis*-expression quantitative trait loci for the gene, supporting its candidacy as the regional effector transcript (6). *ARL15* encodes ARL15, a relatively uncharacterized member of the ARF family of proteins involved in regulation of vesicle trafficking and biogenesis. The gene is highly expressed in  $\beta$ -cells and located downstream of an islet-active enhancer bound by key  $\beta$ -cell transcription factors (18,23) (Supplementary Fig. 10A). *THADA* encodes the

protein THADA and contains a coding disease-association signal that has also been associated with reduced  $\beta$ -cell function (24) (Supplementary Fig. 10C). Consistent with the directionality of our findings, expression profiling has shown the gene to be more highly expressed in patients with type 2 diabetes compared with control subjects (25). All three genes thus emerge as strong candidates for future studies.

While successfully enabling unbiased functional characterization, our current screening strategy has a number of limitations. False negatives (i.e., true causal genes not identified as hits) could arise as a result of primary effects of the causal gene on non- $\beta$ -cell tissues or through effects on genes expressed at different developmental stages. Likewise, overexpression or greater knockdown efficiency may in some cases be required to expose a disease-relevant phenotype. Among the targets analyzed for silencing efficiency, a variable range of knockdown was observed (34%–88%), and some genes might remain undetected due to insufficient silencing. Conversely, false-positive effects (i.e., non- $\beta$ -cell regulators identified as hits) also cannot be excluded, and unexpected findings should be further functionally validated (e.g., *SLC2A4* effect on IBMX-stimulated insulin secretion). Though the EndoC- $\beta$ H1 cell line has been found to recapitulate many aspects of  $\beta$ -cell function, it remains a possibility that some findings would not translate directly into human physiology.



**Figure 2**—Comparison of insulin secretion data for high and low glucose. Normalized insulin secretion responses under high glucose vs. low glucose, with selected hits annotated. The blue circle indicates the 95% confidence contour for NT control, and the orange circle indicates the 95% confidence contour for *INS*-positive controls. All measurements were normalized on a per-well basis to cell counts, and averages for each condition were then subsequently normalized to the mean of NT control. Data points are mean of  $n = 3$  and shown as percentage of NT control.



**Figure 3**—Insulin secretion data for selected genes in a follow-up validation experiment. Insulin secretion for *ARL15* (A), *ZMIZ1* (B), *THADA* (C), and *HNF4A* (D) (white bars) vs. NT (black bars) negative control under the indicated conditions. Measurements were processed as described for Fig. 2 and shown as percentage of NT control. Tol, tolbutamide. Data points are the mean of  $n = 6$  for NT and  $n = 3$  for other genes, and error bars are SEM. +q value  $< 0.1$ , \*q value  $< 0.05$ , \*\*q value  $< 0.01$ , \*\*\*q value  $< 0.001$  by Student  $t$  test (FDR-adjusted).

Finally, a subset of the identified hits may represent true  $\beta$ -cell regulators that are independent of any disease risk variants and, though still of biological importance, not genuine effector transcripts for type 2 diabetes. In addition to the possibility of more than a single effector transcript per locus, this phenomenon likely also contributes to the relatively high proportion of multihit loci observed in the primary screen (8/37).

Despite these limitations, our screening strategy successfully replicated well-established biological mechanisms and identified genes involved in  $\beta$ -cell function at half of the loci investigated. This demonstrates, for the first time, the feasibility of performing scalable screens for insulin secretory defects in human pancreatic  $\beta$ -cells and opens up the possibility for not only large-scale genetic manipulations but also compound high-throughput screening to

therapeutically manipulate human  $\beta$ -cells. Insights from this and subsequent functional screens can be integrated with complementary lines of evidence from exome-wide association studies, chromatin conformation capture, and *cis*-expression quantitative trait loci studies to prioritize genes for follow-up studies. Ultimately, this could accelerate the translation of genetic association signals into molecular mechanisms for  $\beta$ -cell dysfunction, insulin insufficiency, and type 2 diabetes.

**Funding.** This study was funded in Oxford by the Wellcome Trust (095101/Z/10/Z and 098381). S.K.T. is a Radcliffe Department of Medicine Scholar. M.I.M. is a Wellcome Trust Senior Investigator. A.L.G. is a Wellcome Trust Senior Fellow in basic biomedical science.

**Duality of Interest.** M.v.d.B. is supported by a Novo Nordisk postdoctoral fellowship run in partnership with the University of Oxford. No other potential conflicts of interest relevant to this article were reported.

**Author Contributions.** S.K.T., A.C., D.E., M.I.M., and A.L.G. conceived and designed the study. S.K.T., M.v.d.B., M.I.M., and A.L.G. analyzed and interpreted the data. S.K.T., A.C., C.B., and A.B. performed the experiments. R.S. provided protocols. S.K.T., M.I.M., and A.L.G. wrote the manuscript. S.K.T., A.C., M.v.d.B., C.B., A.B., R.S., D.E., M.I.M., and A.L.G. edited and approved the manuscript. A.L.G. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Prior Presentation.** Parts of this study were presented at the 52nd European Association for the Study of Diabetes Annual Meeting, Munich, Germany, 12–16 September 2016.

## References

- Ali O. Genetics of type 2 diabetes. *World J Diabetes* 2013;4:114–123
- Prasad RB, Groop L. Genetics of type 2 diabetes-pitfalls and possibilities. *Genes (Basel)* 2015;6:87–123
- Dimas AS, Lagou V, Barker A, et al.; MAGIC Investigators. Impact of type 2 diabetes susceptibility variants on quantitative glycemic traits reveals mechanistic heterogeneity. *Diabetes* 2014;63:2158–2171
- Thomsen SK, Gloyn AL. The pancreatic  $\beta$  cell: recent insights from human genetics. *Trends Endocrinol Metab* 2014;25:425–434
- Fadista J, Vikman P, Laakso EO, et al. Global genomic and transcriptomic analysis of human pancreatic islets reveals novel genes influencing glucose metabolism. *Proc Natl Acad Sci U S A* 2014;111:13924–13929
- van de Bunt M, Manning Fox JE, Dai X, et al. Transcript expression data from human islets links regulatory signals from genome-wide association studies for type 2 diabetes and glycemic traits to their downstream effectors. *PLoS Genet* 2015;11:e1005694
- Burns SM, Vetere A, Walpita D, et al. High-throughput luminescent reporter of insulin secretion for discovering regulators of pancreatic beta-cell function. *Cell Metab* 2015;21:126–137
- Ku GM, Pappalardo Z, Luo CC, German MS, McManus MT. An siRNA screen in pancreatic beta cells reveals a role for Gpr27 in insulin production. *PLoS Genet* 2012;8:e1002449
- Wu W, Shang J, Feng Y, et al. Identification of glucose-dependent insulin secretion targets in pancreatic beta cells by combining defined-mechanism compound library screening and siRNA gene silencing. *J Biomol Screen* 2008;13:128–134
- Lee JA, Chu S, Willard FS, et al. Open innovation for phenotypic drug discovery: the PD2 assay panel. *J Biomol Screen* 2011;16:588–602
- Walpita D, Hasaka T, Spoonamore J, et al. A human islet cell culture system for high-throughput screening. *J Biomol Screen* 2012;17:509–518
- Ravassard P, Hazhouz Y, Pechberty S, et al. A genetically engineered human pancreatic  $\beta$  cell line exhibiting glucose-inducible insulin secretion. *J Clin Invest* 2011;121:3589–3597
- Weir GC, Bonner-Weir S. Finally! A human pancreatic  $\beta$  cell line. *J Clin Invest* 2011;121:3395–3397
- Andersson LE, Valtat B, Bagge A, et al. Characterization of stimulus-secretion coupling in the human pancreatic EndoC- $\beta$ H1 beta cell line. *PLoS One* 2015;10:e0120879
- Krishnan K, Ma Z, Björklund A, Islam MS. Calcium signaling in a genetically engineered human pancreatic  $\beta$ -cell line. *Pancreas* 2015;44:773–777
- Gurgul-Convey E, Kaminski MT, Lenzen S. Physiological characterization of the human EndoC- $\beta$ H1  $\beta$ -cell line. *Biochem Biophys Res Commun* 2015;464:13–19
- Pal A, Potjer TP, Thomsen SK, et al. Loss-of-function mutations in the cell-cycle control gene CDKN2A impact on glucose homeostasis in humans. *Diabetes* 2016;65:527–533
- Nica AC, Ongen H, Irminger JC, et al. Cell-type, allelic, and genetic signatures in the human pancreatic beta cell transcriptome. *Genome Res* 2013;23:1554–1562
- Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT, Ellard S. Maturity-onset diabetes of the young (MODY): how many cases are we missing? *Diabetologia* 2010;53:2504–2508
- Pearson ER, Boj SF, Steele AM, et al. Macrosomia and hyperinsulinaemic hypoglycaemia in patients with heterozygous mutations in the HNF4A gene. *PLoS Med* 2007;4:e118
- Glaser B. Type 2 diabetes: hypoinsulinism, hyperinsulinism, or both? *PLoS Med* 2007;4:e148
- Rahman SA, Nessa A, Hussain K. Molecular mechanisms of congenital hyperinsulinism. *J Mol Endocrinol* 2015;54:R119–R129
- Pasquali L, Gaulton KJ, Rodriguez-Segui SA, et al. Pancreatic islet enhancer clusters enriched in type 2 diabetes risk-associated variants. *Nat Genet* 2014;46:136–143
- Simonis-Bik AM, Nijpels G, van Haefen TW, et al. Gene variants in the novel type 2 diabetes loci CDC123/CAMK1D, THADA, ADAMTS9, BCL11A, and MTNR1B affect different aspects of pancreatic beta-cell function. *Diabetes* 2010;59:293–301
- Parikh H, Lyssenko V, Groop LC. Prioritizing genes for follow-up from genome wide association studies using information on gene expression in tissues relevant for type 2 diabetes mellitus. *BMC Med Genomics* 2009;2:72