

# Identification of a Novel Gene for Diabetic Traits in Rats, Mice, and Humans

Shirng-Wern Tsaih,\* Katie Holl,\*<sup>†</sup> Shuang Jia,\*<sup>†</sup> Mary Kaldunski,\*<sup>†</sup> Michael Tschannen,\* Hong He,\*<sup>†</sup> Jaime Wendt Andrae,\* Shun-Hua Li,<sup>†</sup> Alex Stoddard,\*<sup>†</sup> Andrew Wiederhold,<sup>†</sup> John Parrington,<sup>‡</sup> Margarida Ruas da Silva,<sup>‡</sup> Antony Galione,<sup>‡</sup> James Meigs,<sup>§</sup> Meta-Analyses of Glucose and Insulin-Related Traits Consortium (MAGIC) Investigators,<sup>1</sup> Raymond G. Hoffmann,\*<sup>†</sup> Pippa Simpson,\*<sup>†</sup> Howard Jacob,\*<sup>†</sup> Martin Hessner,\*<sup>†</sup> and Leah C. Solberg Woods\*<sup>1,2</sup>

\*Human and Molecular Genetics Center and <sup>†</sup>Department of Pediatrics, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, <sup>‡</sup>Department of Pharmacology, University of Oxford, Oxford, OX1 3QT, United Kingdom, and <sup>§</sup>Department of Medicine, Harvard Medical School, General Medicine Division, Massachusetts General Hospital, Boston, Massachusetts 02114

**ABSTRACT** The genetic basis of type 2 diabetes remains incompletely defined despite the use of multiple genetic strategies. Multiparental populations such as heterogeneous stocks (HS) facilitate gene discovery by allowing fine mapping to only a few megabases, significantly decreasing the number of potential candidate genes compared to traditional mapping strategies. In the present work, we employed expression and sequence analysis in HS rats (*Rattus norvegicus*) to identify *Tpcn2* as a likely causal gene underlying a 3.1-Mb locus for glucose and insulin levels. Global gene expression analysis on liver identified *Tpcn2* as the only gene in the region that is differentially expressed between HS rats with glucose intolerance and those with normal glucose regulation. *Tpcn2* also maps as a *cis*-regulating expression QTL and is negatively correlated with fasting glucose levels. We used founder sequence to identify variants within this region and assessed association between 18 variants and diabetic traits by conducting a mixed-model analysis, accounting for the complex family structure of the HS. We found that two variants were significantly associated with fasting glucose levels, including a nonsynonymous coding variant within *Tpcn2*. Studies in *Tpcn2* knockout mice demonstrated a significant decrease in fasting glucose levels and insulin response to a glucose challenge relative to those in wild-type mice. Finally, we identified variants within *Tpcn2* that are associated with fasting insulin in humans. These studies indicate that *Tpcn2* is a likely causal gene that may play a role in human diabetes and demonstrate the utility of multiparental populations for positionally cloning genes within complex loci.

**T**O date, human genome-wide association studies (GWAS) have identified >60 genes involved in type 2 diabetes (T2D) (Zeggini *et al.* 2008; Voight *et al.* 2010; Morris *et al.* 2012; DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium *et al.* 2014) and an additional 53 involved in related metabolic traits such as fasting glucose and insulin levels (Dupuis *et al.* 2010; Scott *et al.* 2012).

Despite the relative success of human GWAS, when combined, these genes explain only a small percentage of the heritable variance (So *et al.* 2011; Morris *et al.* 2012), indicating many more genes have yet to be identified. Improving power with increased sample size, deep sequencing to identify rare variants, and identification of gene–gene and gene–environment interactions will help to identify at least some of this missing heritability in humans (Hu and Daly 2012). Animal models, however, offer an additional, complementary, method for identifying candidate genes and their related pathways. Advantages of animal models include the ability to control environment and genetics and assess gene expression levels in tissues that are not readily available from humans. Importantly, genes found for complex traits in animal models are frequently translated to humans (Swanberg *et al.* 2005; Aitman *et al.* 2006; Samuelson *et al.* 2007; Behmoaras *et al.*

Copyright © 2014 by the Genetics Society of America  
doi: 10.1534/genetics.114.162982

Manuscript received February 14, 2014; accepted for publication May 25, 2014  
Available freely online through the author-supported open access option.

Supporting information is available online at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.162982/-/DC1>.

<sup>1</sup>MAGIC Investigators are listed in Supporting Information, File S1.

<sup>2</sup>Corresponding author: TBRC, C2315, Medical College of Wisconsin, 8701 Watertown Plank Rd., Milwaukee, WI 53226. E-mail: lsolberg@mcw.edu

2008; Petretto *et al.* 2008), demonstrating the utility of animal models for uncovering at least some of the missing heritability in humans.

Particularly useful are multiparental or outbred populations such as heterogeneous stocks (HS). HS animals are derived from crossing eight inbred strains and outbreeding them for many generations (see Flint and Eskin 2012 or Solberg Woods 2014 for a review on HS populations). The genetic makeup of the resulting progeny represents a random mosaic of the founder animals, with each HS rat being phenotypically and genetically distinct. The distance between recombination events decreases with each generation of breeding, such that HS mice and rats have been used to map multiple complex traits to generally <5 Mb (Talbot *et al.* 1999, 2003; Demarest *et al.* 2001; Valdar *et al.* 2006; Solberg Woods *et al.* 2010, 2012; Johnsen *et al.* 2011; Baud *et al.* 2013). This fine-resolution mapping significantly decreases the number of potential candidate genes within each locus relative to those in conventional mapping methods (such as those using an F<sub>2</sub> intercross or backcross). Once narrowed, HS founder sequence can be used to rapidly identify potentially causative genes within at least some of these loci (Keane *et al.* 2011; Baud *et al.* 2013), thus making these populations ideal for positional cloning of genes involved in complex traits.

We have used HS rats to fine-map traits involved in diabetes and related traits. Our work has shown that HS rats harbor alleles for glucose tolerance (Solberg Woods *et al.* 2010), as well as insulin resistance and  $\beta$ -cell dysfunction, the underlying causes of type 2 diabetes (Solberg Woods *et al.* 2012). We have identified a 3.1-Mb region on rat chromosome 1 that maps both fasting and postprandial glucose (Solberg Woods *et al.* 2010, 2012), and this region overlaps a larger 7-Mb region that maps fasting insulin and insulin sensitivity (Solberg Woods *et al.* 2012). Although a great improvement over previous mapping studies in F<sub>2</sub> intercrosses (Galli *et al.* 1996; Gauguier *et al.* 1996; Chung *et al.* 1997; Kanemoto *et al.* 1998; Wei *et al.* 1999; Nobrega *et al.* 2009; Solberg Woods *et al.* 2009; Wallis *et al.* 2009), this region still contains a relatively large number of genes: 86 within the 3.1-Mb glucose region and >200 genes in the insulin region. Although a few plausible genes (including *Kcnq1* and *Igf2*) stand out as likely candidates within the 7-Mb insulin region, none of the 86 genes within the 3.1-Mb region present themselves as obvious potential biological candidate genes.

The goal of the present study was to use expression analysis and next-generation sequencing to evaluate positional candidate genes within this 3.1-Mb glucose locus, located on rat chromosome 1: 204.38–207.48 Mb, a complex region that may harbor multiple causal genes (Granhall *et al.* 2006; Solberg Woods *et al.* 2010, 2012). Expression and sequence analysis in the HS rats identified two-pore segment channel 2 (*Tpcn2*) as a plausible candidate gene within this region. We confirm the role of *Tpcn2* by demonstrating that *Tpcn2* knockout mice exhibit altered fasting glucose and insulin in response to a glucose chal-

lenge. We also demonstrate significant association between SNPs in *Tpcn2* and fasting insulin in humans. Together this work indicates that *Tpcn2* is a strong candidate for being a causal gene that underlies the glucose and insulin phenotypes at this QTL.

## Materials and Methods

### Animals

**Heterogeneous stock colony:** The NMcwi:HS colony, hereafter referred to as HS, was initiated by the National Institutes of Health (NIH) in 1984, using the following eight inbred founder strains: ACI/N, BN/SsN, BUF/N, F344/N, M520/N, MR/N, WKY/N, and WN/N (Hansen and Spuhler 1984). This colony has been maintained at the Medical College of Wisconsin (MCW) since 2006 and has been through >60 generations of breeding (see Solberg Woods *et al.* 2010 for history of the breeding colony).

**Founding inbred substrains:** Expression analysis was conducted in the following substrains (abbreviated names used throughout the article are in parentheses): ACI/Eur (ACI), BN/SsnHsd (BN), BUF/NHsd (BUF), F344/NHsd (F344), M520/N, and WKY/NHsd (WKY). Other than M520/N, which is the original founder of the HS colony, substrains were chosen based on origin similarities to the original founders, as determined by data obtained from the Rat Genome Database: <http://rgd.mcw.edu/strains/>. The M520 and ACI rats are maintained at MCW. All other substrains were ordered from Harlan Sprague Dawley (Indianapolis). We were unable to identify substrains for two of the founders (MR/N and WN/N), so these strains were not tested.

### Cluster analysis of glucose and insulin levels in HS rats

Rats for expression analysis were chosen from a group of 522 HS rats previously used to map glucose tolerance (Solberg Woods *et al.* 2010) (see [Supporting Information, File S2](#)). As previously described, HS rats underwent an intraperitoneal glucose tolerance test (IPGTT) test at 16 weeks of age in which blood glucose and plasma insulin were measured at fasting and at 15, 30, 60, and 90 min after a 1-g/kg glucose injection (Solberg Woods *et al.* 2010, 2012). Glucose area under the curve (glucose\_AUC) and insulin\_AUC were determined using a trapezoidal analysis. The quantitative insulin sensitivity check (QUICKI), a measure of insulin sensitivity that has been validated in both rats and humans (Katz *et al.* 2000; Cacho *et al.* 2008), was calculated as  $1/[\log I_0 + \log G_0]$ , where  $I_0$  is fasting insulin and  $G_0$  is fasting glucose (Solberg Woods *et al.* 2012). HS rats were killed at 17 weeks of age, 1 week after the glucose tolerance test. Fasting cholesterol, triglycerides, and c-peptide were determined from plasma at the time the animals were killed. Livers were immediately removed and frozen in liquid nitrogen for subsequent determination of transcript abundance. Rats were killed by decapitation to avoid gene expression changes that can occur when using inhalant drugs such as CO<sub>2</sub> or isoflurane (Kadar *et al.* 2011; Taylor and Cummins

2011). All protocols were approved by the Institutional Animal Care and Use Committee at MCW.

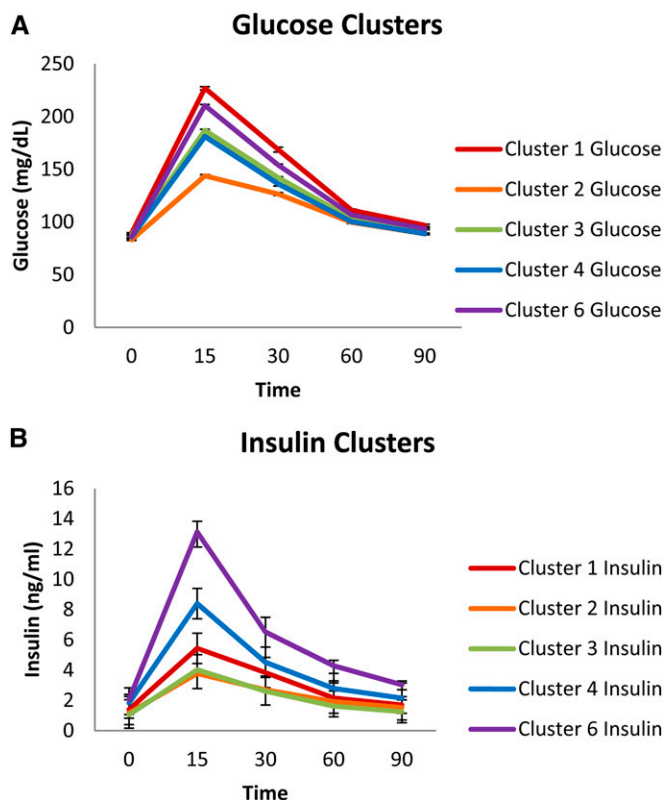
To determine how many phenotypic groups existed within the HS colony, a *K*-means cluster analysis was conducted. Clusters were determined using the following glucose and insulin measurements: glucose\_AUC, insulin\_AUC, peak time, the difference between baseline and the peak value, the slope of the line from right before the peak to peak time, and slope from peak time to right after the peak. The *K*-means clustering has an animal's phenotype values added to a cluster based on minimizing differences from the means of the *K* clusters. Canonical discrimination using SAS 9.1 was used to determine how the clustering related to the phenotype measures.

### Expression analysis

Expression analysis was conducted in 46 male HS rats and 2 males from each of six inbred substrains, for a total of 58 animals. Total RNA was extracted from frozen liver tissue, using TRIzol reagent [Invitrogen (now Life Technologies), Carlsbad, CA]. Purified RNA (50 ng) was amplified using an Affymetrix two-cycle cDNA synthesis kit, and complementary RNA was synthesized, labeled, fragmented, and hybridized to the Gene-Chip rat genome 230\_2.0 array in accordance with standard protocols (Affymetrix, Santa Clara, CA). The rat genome 230\_2.0 array contains >31,000 probe sets representing >17,734 unique UniGenes, containing probes for 57 of the 86 genes (66.3%) within our region of interest [rat chromosome 1: 204.38–207.48 Mb (Solberg Woods *et al.* 2010, 2012)]. After hybridization, arrays were washed and stained with an Affymetrix Wash and Stain kit plus an Affymetrix GeneChip Fluidics Station 450 and scanned with an Affymetrix 7G laser scanner. Data were analyzed with Expression Console 1.1.2 software (Affymetrix) and normalized with robust multichip analysis (RMA) (<http://www.bioconductor.org/>) to determine signal log ratios ( $\log_2R$ ). The statistical significance of differential gene expression was determined through ANOVA, using Partek (St. Louis) Genomics Suite 6.5. Differential expression (DE) was defined as those genes with  $P \leq 0.05$  and  $\log_2R \geq 0.5$ . Because we planned to verify differential expression of genes of interest with quantitative RT-PCR (qRT-PCR), we did not correct for multiple testing (Hessner *et al.* 2004).

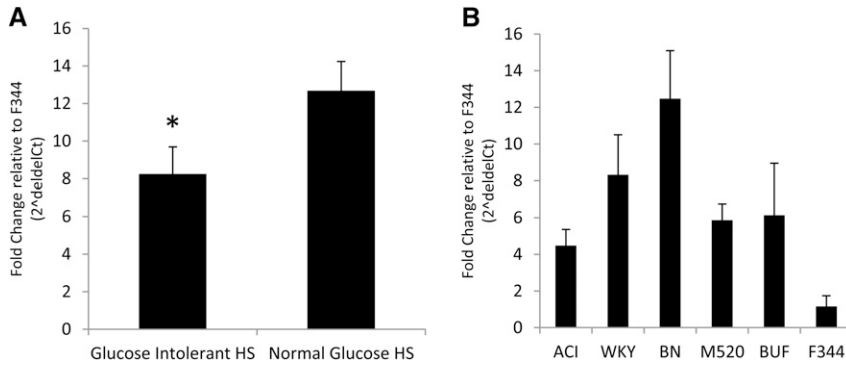
### Real-time qRT-PCR

We initially confirmed differential expression of *Tpcn2* using qRT-PCR in 10 HS rats with glucose intolerance and 10 with normal glucose regulation. qRT-PCR was also conducted in the founder strains, using 2 animals from each inbred strain. Upon confirming *Tpcn2* as a plausible candidate gene, we used qRT-PCR to assess gene expression levels in a total of 120 male HS rats. qRT-PCR was conducted on an ABI Prism 7900HT Sequence Detection System (Applied Biosystems, Foster City, CA). Specific oligonucleotide primers were designed with Primer3 (<http://binoinfo.ut.ee/primer3-0.4.0/primer3/>) (Koressaar and Remm 2007; Untergasser *et al.* 2012). To avoid differences in expression levels based on sequence variation



**Figure 1** (A and B) Cluster analysis of (A) glucose and (B) insulin levels during an intraperitoneal glucose tolerance test in >500 HS rats. Values are expressed in means  $\pm$  SE. We identified six clusters based on glucose and insulin values. Because cluster 5 included only one animal, however, it is not included in the graphs or in subsequent analyses. Note that clusters with similar glucose values (clusters 1 and 6 or 3 and 4) have drastically different insulin values. Based on this cluster analysis, we selected animals for the expression analysis, choosing glucose intolerant HS from cluster 1 and HS with normal glucose from clusters 2 and 4.

(Huang *et al.* 2009), we ensured no sequence variants were present in any of the HS founders at the location of the *Tpcn2* primers. The primers for *Tpcn2* were forward 5'-TATGTGTTTCGC CATGATTGG-3' and reverse 5'-AGCAGCAAAGTCGTCGAAGT (Integrated DNA Technologies). We used primers for GAPDH, forward 5'-CATGGAGAAGGCTGGGGCTC-3' and reverse 5'-AACGGATACATTGGGGGTAG-3', and  $\beta$ -2 microglobulin (B2M), forward 5'-CCGTGATCTTTCTGGTGCTT-3' and reverse 5'-TTTTGGGCTCCTTCAGAGTG-3' (Integrated DNA Technologies), as internal controls. We used iQ SYBR Green Supermix (Bio-Rad, Hercules, CA) according to the manufacturer's instructions. Synthesis of first-strand cDNA from 3  $\mu$ g of RNA per animal was accomplished with random hexamers (Invitrogen, now Life Technologies) and Superscript III (Invitrogen, now Life Technologies), according to the manufacturer's instructions. Triplicate *Tpcn2* and GAPDH or B2M were performed for each sample in 20- $\mu$ l reactions, which included 1  $\mu$ l of cDNA and 10  $\mu$ l of iQ SYBR Green Supermix (Bio-Rad) possessing 1.2  $\mu$ l of *Tpcn2* (10  $\mu$ M), GAPDH, or B2M specific primers and 7.8  $\mu$ l of deionized water. Reactions were cycled as follows: stage 1, 95 $^{\circ}$ , 3 min; stage 2, 55 cycles of 95 $^{\circ}$ ,



**Figure 2** (A and B) Differential expression of *Tpcn2* by qRT-PCR in (A) HS rats and (B) HS founder substrains. Bars show the mean  $\pm$  SE. Glucose intolerant HS rats exhibit an approximately twofold decrease in expression levels of *Tpcn2* relative to HS rats with normal glucose levels (\* $P = 0.05$ ). Large variation is seen in HS founder substrains with BN exhibiting the greatest levels of *Tpcn2* and the F344 strain exhibiting the least. All fold-change differences are relative to the F344 founder strain. Note that the scales in A and B are the same.

30 sec, 55°, 30 sec, 72°, 30 sec; and stage 3, melt curve of 95°, 10 sec, 60°, 1 min, 95°, 1 sec, 40°, 10 sec. Standard curves were created using 1:10, 1:100, and 1:1000 concentrations of *Tpcn2*. Specificity for a qRT-PCR was verified by a melting curve analysis. The data were analyzed with Sequence Detection System 2.3 software (Applied Biosystems), using the cycle threshold (ct) for quantification. Relative gene expression data (fold change) between samples was accomplished using the mathematical model described by Pfaffl (2001), using the mean ct score of 120 animals as the reference.

### Genome-wide genotyping

Tail samples from the original eight inbred founders were obtained from the NIH. We extracted DNA from tail tissue from HS and founder strains, using the QIAGEN (Valencia, CA) DNeasy kit. All eight original founder inbred strains and the 46 HS rats used for expression analysis were genotyped using the Affymetrix GeneChip Targeted Genotyping technology on a custom 10K SNP array panel as previously described (Saar *et al.* 2008). The samples were genotyped by the Vanderbilt Microarray Shared Resource center at Vanderbilt University in Nashville, Tennessee (currently VANTAGE: <http://vantage.vanderbilt.edu>). From the 10,846 SNPs that are on the array, 7634 were informative and produced reliable results in the HS rats. From these final informative markers, the average spacing was 353 kb apart, with an average heterozygosity of 28.8%. These data have been deposited in the Gene Expression Omnibus (GEO) database under accession no. GSE57118 (see: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=sjmjgkwulbebtmd&acc=GSE57118>).

### Genetic mapping of expression QTL

Expression mapping was conducted for all genes on the expression array that were located within the QTL (57 of 86). Genome-wide fine mapping was conducted by linkage disequilibrium mapping, testing for the significance of association between expression levels and inferred haplotype descent at each locus, as previously described (Solberg Woods *et al.* 2010, 2012). Briefly, haplotype descent along each HS rat genome was inferred using the haplotype reconstruction method HAPPY (Mott *et al.* 2000), which applies

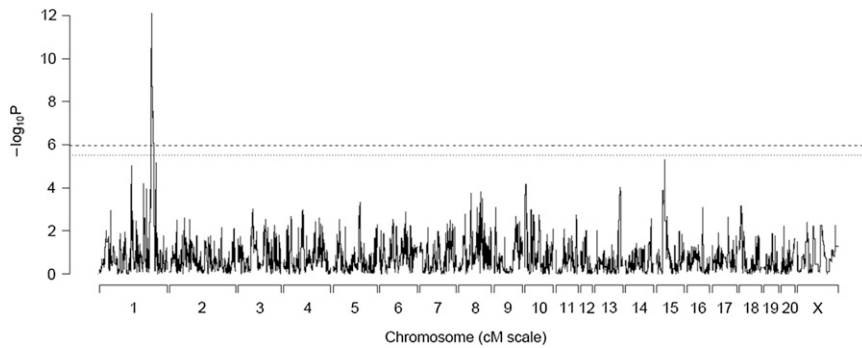
a hidden Markov model simultaneously to the genotypes of the eight founder strains and the 46 HS rats. At each interval  $m = 1, \dots, 7634$  between adjacent pairs of markers, HAPPY produces for each rat  $i$  a vector  $\mathbf{g}_i(m)$  containing the probabilities of descent from each of the  $8(8+1)/2 = 36$  unique haplotype pairs (diplotypes) (Valdar *et al.* 2009). Using Bagpipe software (Valdar *et al.* 2009), this is then used in the following mixed-model regression to predict rat  $i$ 's expression levels  $y_i$ ,

$$f(y_i) = \text{QTL}_i(m) + \text{sibship}_{k[i]} + \text{residual}_i, \quad (1)$$

where  $f(y_i)$  is a transformed version of the expression levels,  $\text{QTL}_i(m) = \boldsymbol{\beta}^T \mathbf{g}_i(m)$  models the putative effect of the QTL,  $\text{sibship}_{k[i]}$  is the (random) effect of the sibship  $k$  to which  $i$  belongs, and  $\text{residual}_i$  accounts for individual variation (further parameters defined as in Solberg Woods *et al.* 2010). Models were fitted by restricted-estimate maximum likelihood, and the significance of association at each interval  $m$  was assessed by comparing the fit of Equation 1 with that of the null model, which is Equation 1 without the QTL term. Genome-wide significance thresholds were estimated by parametric bootstrap from the fitted null model (Valdar *et al.* 2009; Solberg Woods *et al.* 2010). Prior to genetic analysis, expression levels for each gene were normalized using a van der Waerden normal score.

### Sequencing and analysis

We sequenced 5 Mb within rat chromosome 1: 200–208 Mb in the eight inbred founder strains of the HS colony. We designed a custom 385K NimbleGen array that included all coding regions and all regions that are highly conserved. The array was composed of 385,000 long oligonucleotide probes and used a tiling design to target our unique genomic region. Five to eight  $\mu\text{g}$  of genomic DNA from each strain was nebulized and polished, and linkers were added according to the manufacturer's standards (Roche NimbleGen). Samples were then hybridized to the custom NimbleGen 385K array. DNA not specific to these regions were washed off and targeted DNA sequences were eluted and amplified by emulsion PCR. To calculate whether targeted sequences were enriched, both pre- and postcapture DNAs were analyzed by qRT-PCR, using four real-time PCR assays



**Figure 3** Genome-wide eQTL scan of *Tpcn2* expression levels using 10K SNP markers. The x-axis shows chromosome number and position in centimorgans. The y-axis gives the  $-\log_{10}$  *P*-value of association. The dashed lines represent the genome-wide significance thresholds (bottom line represents the suggestive threshold at  $P = 0.1$  and the top line represents significance at  $P = 0.05$ ). *Tpcn2* expression levels map to within 1 Mb of the gene itself on rat chromosome 1, indicating this gene is *cis*-regulated.

developed and optimized by our laboratory. The region was then sequenced on the Roche GS-FLX 454 system (Roche 454, Branford, CT) according to the manufacturer's protocol. Briefly, Roche "A" and "B" adapters were annealed onto captured fragments of template DNA. The library was single stranded and quantified on an Agilent 2100 Bioanalyzer. Library fragments were annealed to beads and emulsion PCR was performed. After clonal amplification by PCR, the emulsion was broken chemically and beads were collected and enriched. Enriched beads were then inserted into the Roche Picotiter Plate, using the appropriate gasket, and loaded onto the machine. Nucleotides flowed over the plate in a specific order and enzymatic cascade reaction caused incorporation of the specific nucleotides to illuminate. An image of every flow was taken with a CCD camera and images were strung together into a flowgram, which was transformed into usable sequence data.

Reads were separated by their multiplex identifiers, using the program *sfutils* (Roche). For each sample, demultiplexed reads were then trimmed and assembled against the rat chromosome 1 reference sequences (Rn4), using the GS Reference Mapper (<http://454.com/products/analysis-software/index.asp>). Variant detection was performed using the high-confidence differences strategies implemented in GS Reference Mapper software. The high-confidence differences (HCDiff) strategy requires the following criteria for a variant to be reported: (1) There must be at least three reads with the difference; (2) there must be both forward and reverse reads showing the difference, unless there are at least five reads with quality scores  $>20$  (or 30 if the difference involves a 5-mer or higher); and (3) if the difference is a single-base overcall or undercall, then the reads with the difference must form the consensus of the sequenced reads.

Since the time of this work, all founder strains have been fully sequenced (Baud *et al.* 2013), and this information is currently available in the Rat Genome Database (<http://rgd.mcw.edu>). We have compared our sequence to what is available online. Some discrepancies were noted between the data sets. To determine the accurate sequence, these regions were sequenced using Sanger sequence technology.

#### Taqman genotyping

We genotyped 18 SNPs within the 3.1-Mb QTL in 508 male HS rats and all eight inbred founders. SNPs were chosen

based on the founder strain distribution pattern (SDP) as a means of defining general haplotype blocks (Yalcin *et al.* 2004a). To ensure most haplotypes were represented, we used a measure of entropy as previously described (Yalcin *et al.* 2005). Briefly, a sliding window was used in which each SNP within the window is prioritized based on its ability to explain the diversity between the founder strains. SNPs were then given a score and those with the highest scores were chosen for genotyping. In addition to SNPs with high entropy scores, we also genotyped nonsynonymous coding variants in the F344 strain, the founder that contributes the susceptibility allele at the QTL (Solberg Woods *et al.* 2010, 2012). SNPs were spaced an average of 164.5 kb apart.

Samples were genotyped using 5'-exonuclease TaqMan technology (Applied Biosystems) with differently fluorescence-labeled probes. Custom TaqMan SNP Genotyping Assays (Applied Biosystems) were used. Ten nanograms of DNA was used in a total volume of 5  $\mu$ l containing 2  $\mu$ l of 5ng/ $\mu$ l DNA, 2  $\mu$ l 1 $\times$  TaqMan Universal PCR Master Mix (Applied Biosystems), 0.125  $\mu$ l probe, and 0.875  $\mu$ l water for each sample. PCR was carried out on a GeneAmp PCR System 9700 (Applied Biosystems) and the post-PCR plate was read on the ViiA 7 real-time PCR system (Applied Biosystems), following the manufacturer's instructions. ViiA 7 RUO version 1.2.1 software was used to assign genotypes, applying the allelic discrimination test. The overall call rate was 98.8%.

#### Glucose tolerance test in *Tpcn2* knockout mice

Wild-type and knockout mice were created using embryonic stem cells from the 129P2 strain carrying a gene trap vector and injected into C57BL/6 blastocysts, as previously described (Calcraft *et al.* 2009). We confirmed that *Tpcn2* is not transcribed in the knockout mice by running qPCR on multiple tissues in knockout and wild-type mice (see Figure S1). PCR was conducted as described above, using primers for GAPDH (forward 3'-CTGAAGGGCATCTTGGGCTA-5' and reverse 3'-GCCGTATTCAATTGTCATACCA-5') and *Tpcn2* (forward 3'-CCAGGCTACCTGTCTACCA-5' and reverse 3'-CAGGAAGCGAAACACAATCA-5'). Heterozygous mice were bred and set up as breeder pairs. Male *Tpcn2* knockout, heterozygote, and wild-type mice from the heterozygote breeders were phenotyped at 9–13 weeks of age ( $n = 5-7$

**Table 1 HS rat SNPs that pass the region-wide significance threshold for fasting glucose and/or *Tpcn2* expression levels**

SNP location (bp)	Gene/functional location	Trait (–log <i>P</i> )	Variant	Founder strain(s) with variant	Amino acid change (Polyphen and SIFT predictions)
205,653,821	<i>Rmt1</i> /intron	Fasting glucose (3.62)	T to C	ACI, BUF, M520, MR, WKY, WN	NA
205,715,459	<i>Tpcn2</i> /exon	<i>Tpcn2</i> expression (6.58) Fasting glucose (3.35) <i>Tpcn2</i> expression (6.17)	G to A	F344	P to L (benign/tolerated)
205,898,738	<i>Cpt1a</i> /exon	<i>Tpcn2</i> expression (4.90)	G to A	ACI, BUF, M520, MR, WN	V to I (benign/tolerated)
206,242,238	<i>RGD1311946</i> /exon	<i>Tpcn2</i> expression (3.69)	G to A	F344	A to T (unknown/tolerated)
206,726,856	<i>Gpr152</i> /exon	<i>Tpcn2</i> expression (4.43)	C to T	F344	Q to * (unknown/unknown)

\* indicates a stop codon.

in each group), using the IPGTT, described above and in Solberg Woods *et al.* (2010, 2012). Blood was collected after an overnight fast and at 15, 30, 60, 90, and 120 min after a 1-g/kg glucose injection. We used the Ascensia Elite system for reading blood glucose values (Bayer, Elkhart, IN). We also collected blood at each time point for subsequent analysis of plasma insulin levels, which was assayed using an ultrasensitive ELISA kit from Alpco Diagnostics (Salem, NH).

### Statistical analysis

A mixed model was used to determine association between 18 SNPs within the chromosome 1 QTL and diabetic phenotypes that map to this region (glucose\_AUC, fasting glucose, fasting insulin, insulin\_AUC, and QUICKI) and *Tpcn2* expression levels (as determined by qPCR). The model included all appropriate covariates (location, injector, collector, and number of glucose injections) and accounted for the complex family relationships of the HS, as previously described (Solberg Woods *et al.* 2010, 2012). The significance threshold was determined using the Bonferroni method to account for multiple comparisons (18 SNPs by six traits). Correlations between *Tpcn2* expression levels (as determined by qRT-PCR) and these traits were assessed using Pearson's correlation coefficient. Pearson's correlation coefficient was also used to determine correlations between all microarray probes within the chromosome 1 region and metabolic traits, using a Bonferroni correction to determine significance. A one-way ANOVA was used to determine whether there were statistical differences between wild-type, heterozygote, and *Tpcn2* knockout mice. If a trend toward significance was noted, a *t*-test was used to assess significance between wild-type and knockout mice only.

### *Tpcn2* lookup in human GWAS

We used data from two large consortiums to determine whether *Tpcn2* may play a role in human diabetes. Diabetes Genetics Replication and Meta-analysis (DIAGRAM) has collected 2.4 million SNPs in >47,000 cases and controls (Zeggini *et al.* 2008; Voight *et al.* 2010), with 34,840 cases and 114,981 controls recently genotyped on the custom Metabochip (Morris *et al.* 2012). The Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) has collected phenotypic information for quantitative glycemic traits such as glucose and insulin

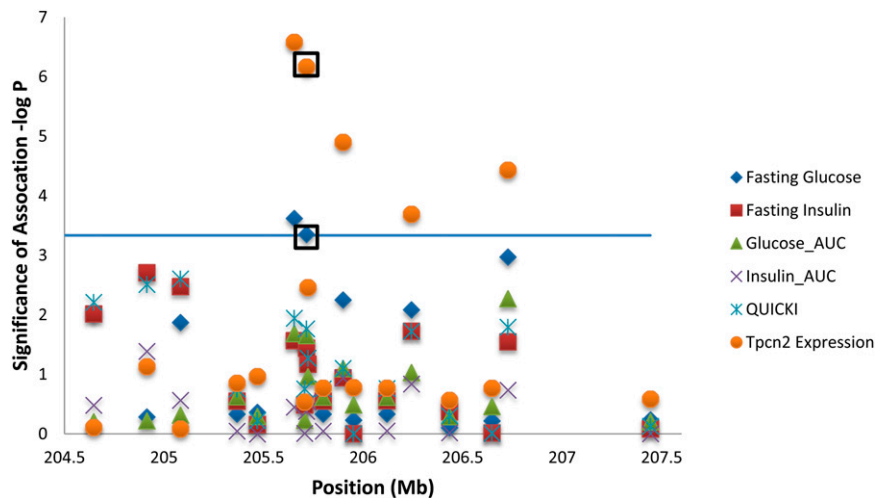
in >40,000 individuals (Dupuis *et al.* 2010), with 133,010 individuals genotyped on the custom Metabochip (Scott *et al.* 2012). Both DIAGRAM and MAGIC cohorts consist of males and females mainly of European descent. Through collaboration with these investigators, we inquired whether a subset of 39 SNPs within *Tpcn2* was associated with diabetic traits in humans. The 39 SNPs were chosen based on information from the HapMap project ([www.hapmap.org](http://www.hapmap.org)), choosing SNPs with an average  $r^2 > 0.8$ . All but 2 SNPs had an average  $r^2 > 0.9$ . Fourteen of the 39 SNPs were upstream and 4 were downstream of the *Tpcn2* gene (see Table S1). For those SNPs that reached an unadjusted level of significance at  $P \leq 0.05$ , we used the Benjamini and Hochberg (1995) false discovery rate to account for multiple comparisons.

## Results

### Glucose and insulin phenotypes in HS rats separate into five main clusters

Although we initially identified six clusters based on glucose and insulin values, one of the clusters (cluster 5) contained a single animal and was excluded from subsequent analysis. There are two dominant phenotypic factors that characterized the clusters. The first (which explained 65% of the variability between clusters) consists of the magnitude of the peak of the insulin curve and the glucose curve, the rate of increase of the glucose from baseline to peak, the rate of increase of insulin from baseline to peak, and glucose\_AUC. The second factor (which explained 31% of the variability between the clusters) consists of the rate of decrease of the insulin curve after the peak and the rate of decrease of the glucose curve after the peak.

The five clusters are shown in Figure 1 and demonstrate the importance of taking into account both glucose and insulin in the clustering. Note that animals in clusters 2 and 3 exhibit very similar plasma insulin levels, but differ significantly in blood glucose levels. Similarly, animals from clusters 3 and 4 exhibit similar blood glucose levels, but differ significantly in plasma insulin levels. For the expression analysis, glucose intolerant HS rats were chosen from cluster 1 (high glucose, moderate insulin) and HS rats with normal glucose regulation were chosen from clusters 2 and 4 (low glucose and insulin and moderate glucose and insulin, respectively). The glucose intolerant group is highly significantly



**Figure 4** Association analysis between 18 SNPs within the 3.1-Mb QTL and all traits that map to this region, including five diabetic traits and *Tpcn2* expression levels. A mixed model, which takes into account the complex family structure of the HS, was used. The nonsynonymous coding variant within *Tpcn2*, which is significant for both fasting glucose and *Tpcn2* expression levels, has been marked with a black square. The x-axis shows the position in megabases, and the y-axis give the  $-\log_{10} P$  of association. The solid line represents the Bonferroni-adjusted region-wide significance threshold (3.33). Expression levels are based on fold change from the qPCR data,  $n = 120$  HS rats; for all other phenotypes,  $n = 508$  HS rats.

different from the normal glucose group (glucose\_AUC = 13,458.9 vs. 9061,  $F_{1,44} = 398.7$ ,  $P < 2e-16$ ). Animals in the glucose intolerant group also exhibit differences in several other phenotypes, including significantly higher levels of body weight (349.1 vs. 314.4 g,  $F_{1,44} = 15.47$ ,  $P = 0.000294$ ), fasting glucose (90.3 vs. 77.5 mg/dl,  $F_{1,44} = 42.65$ ,  $P = 5.59e-08$ ), fasting insulin (1.8 vs. 1.1 ng/ml,  $F_{1,44} = 7.52$ ,  $P = 0.009$ ), insulin\_AUC (311.3 vs. 204.3,  $F_{1,44} = 17.10$ ,  $P = 0.000157$ ), QUICKI (0.48 vs. 0.54,  $F_{1,44} = 7.96$ ,  $P = 0.007$ ) fasting c-peptide (1808.0 vs. 1505.8 pg/ml,  $F_{1,44} = 5.75$ ,  $P = 0.021$ ), and fasting triglycerides (86.2 vs. 72.3 mg/dl,  $F_{1,44} = 8.0$ ,  $P = 0.007$ ). No differences are seen between the groups for fasting cholesterol or fat pad weights (see Table S2).

#### ***Tpcn2* is the only gene measured within the 3.1-Mb QTL that is differentially expressed between HS rats with different glucose profiles**

We identify 48 probes genome-wide that are differentially expressed between HS rats with glucose intolerance and those with normal glucose regulation, 3 of which had an unknown location (see Table S3). These data have been deposited in the GEO database under accession no. GSE57118 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=sjmjgkwulbebtmd&acc=GSE57118>). Of these 48 genes, only one falls within the 3.1-Mb glucose region on rat chromosome 1: *Tpcn2*. Statistics for the other probes within the 3.1-Mb region are listed in Table S4. Using qRT-PCR, we confirm differential expression of *Tpcn2* against both GAPDH ( $F_{1,19} = 4.36$ ,  $P = 0.05$ ; see Figure 2A) and B2M ( $F_{1,18} = 7.03$ ,  $P = 0.017$ ; data not shown). HS founder strains also show different levels of *Tpcn2* expression, with the F344 strain showing the least amount of *Tpcn2* expression and the BN strain showing the greatest amount of *Tpcn2* expression (see Figure 2B). All fold-change differences are expressed relative to the F344 founder substrain.

#### ***Tpcn2* maps as a cis-eQTL to the same region as the physiological QTL**

Using haplotype reconstruction and mixed-model linear regression, which takes into account sibship relationships,

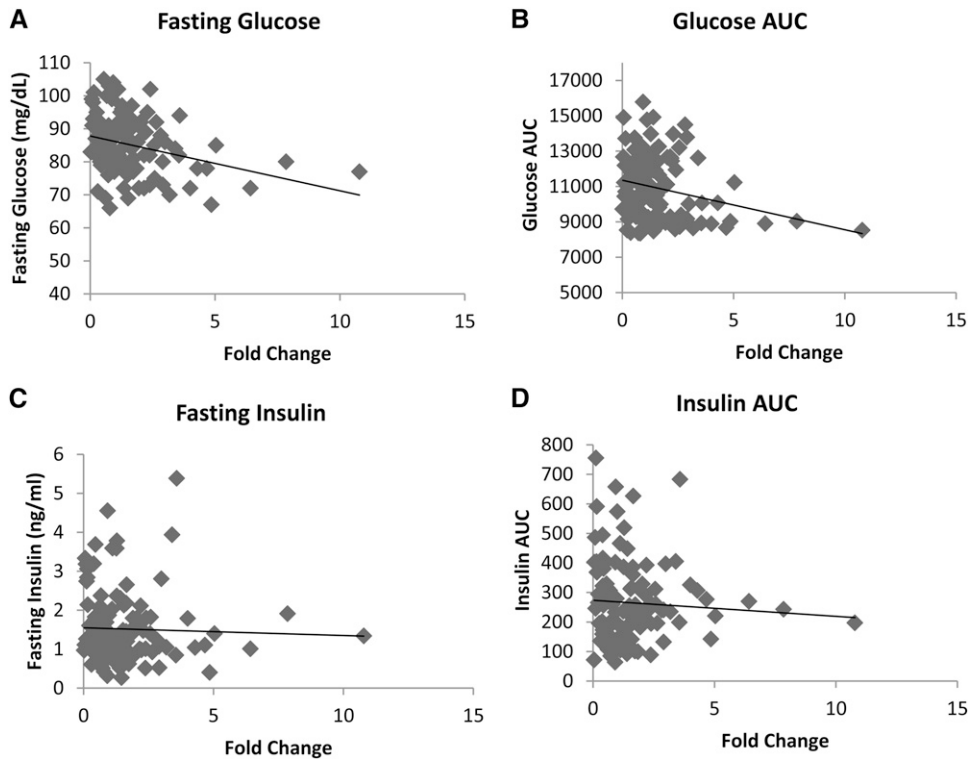
we identify a single sharp peak (marker interval gko.18d13.rp2.b1.157 and Cpn.1207450582, 205.22–207.45 Mb;  $\log P = 12.09$ ) for expression levels of *Tpcn2* (see Figure 3). Only one other gene within this region, *Acy3*, maps as a cis-eQTL, but with a much lower  $\log P$  (marker interval WKYOa01b03.r1.818 and rat106.025.f16.p1ca.195, 204.32–207.50 Mb;  $\log P = 6.28$ ). The genome-wide significance threshold is  $\log P = 5.95$  for both *Tpcn2* and *Acy3* expression levels, where  $-\log P$  is defined as  $-\log_{10}(P\text{-value})$ .

#### **Nonsynonymous coding variants are identified within three genes, including *Tpcn2***

As expected, we identify ~2000–3000 SNPs between each strain and the BN reference genome (Saar *et al.* 2008). Within the 3.1-Mb glucose locus, we identify three genes with a nonsynonymous amino acid change in the F344 founder strain: *Tpcn2*, *RGD1311946*, and *Gpr152* (see Table 1). We focused on variants in the F344 strain because our previous work demonstrates that the F344 strain contributes to the susceptibility allele at the QTL for glucose tolerance (Solberg Woods *et al.* 2010) as well as for fasting glucose and insulin (Solberg Woods *et al.* 2012). It is important to note that loci for fasting glucose and insulin exhibit a complex genetic architecture with several other founder haplotype combinations leading to an increase in fasting glucose or insulin. Because of this, we were unable to run a simple haplotype analysis, as demonstrated by other groups (Svenson *et al.* 2012; Logan *et al.* 2013). The variants within *Tpcn2* and *RGD1311946* are not predicted to be damaging by PolyPhen or SIFT. The variant within *Gpr152*, however, results in a premature stop codon and could potentially lead to changes in protein function.

#### **Variants within and near *Tpcn2* are significantly associated with fasting glucose and *Tpcn2* expression levels**

Using a Bonferroni-adjusted region-wide threshold of 3.33, we found that 2 SNPs, located at 205,653,821 bp ( $-\log P = 3.62$ ) and 205,715,459 bp ( $-\log P = 3.35$ ), are significantly associated with fasting glucose (see Figure 4, Table 1, and Table S5). The SNP at 205,715,459 bp harbors the nonsynonymous



**Figure 5** (A–D) Expression of *Tpcn2* in liver of 120 HS rats as determined by qRT-PCR and correlations with (A) fasting glucose, (B) glucose\_AUC, (C) fasting insulin, and (D) QUICKI. Fold change is relative to the mean ct of all HS animals. The black line represents the best fit. A significant negative correlation is found between *Tpcn2* expression levels and fasting glucose ( $r = -0.302$ ,  $P = 0.001$ ) and glucose\_AUC ( $r = -0.246$ ,  $P = 0.007$ ). When the outliers are removed, fasting glucose remains significant ( $r = -0.232$ ,  $P = 0.013$ ).

coding variant within *Tpcn2*, while the SNP at 205,653,821 bp is located in the intron of the neighboring mammary cancer-associated protein *Rmt1* gene. These SNPs and 3 additional SNPs are also associated with *Tpcn2* expression levels: 205,653,821 bp ( $-\log P = 6.58$ ), 205,715,459 bp ( $-\log P = 6.17$ ), 205,898,738 bp ( $-\log P = 4.90$ ), 206,242,238 bp ( $-\log P = 3.69$ ), and 206,726,856 bp ( $-\log P = 4.43$ ). The additional SNPs include the nonsynonymous coding variants within *RGD1311946* and *Gpr152*, as well as a coding variant within *Cpt1a*. No associations are found between the 18 SNPs and glucose\_AUC, fasting insulin, insulin\_AUC, or QUICKI. Genotypes for the 18 SNPs can be found in [File S3](#).

#### ***Tpcn2* expression levels negatively correlate with fasting glucose**

Using the qRT-PCR data in 120 HS rats, we demonstrate a significant negative correlation between *Tpcn2* expression levels and fasting glucose and glucose\_AUC ( $r = -0.302$ ,  $P = 0.001$  and  $r = -0.246$ ,  $P = 0.007$ , respectively; see [Figure 5](#)). After removing 6 animals that have extremely high expression values ( $>2$  SD from the mean), only fasting glucose remains significantly correlated with expression levels ( $r = -0.232$ ,  $P = 0.013$ ). *Tpcn2* expression levels are not significantly correlated with fasting insulin, insulin\_AUC, and QUICKI ( $r = -0.034$ ,  $-0.063$ , and  $0.078$ , respectively). qRT-PCR data can be found in [File S4](#). To determine whether any other genes in the region are correlated with the diabetic traits, we determined correlation coefficients between expression levels of the microarray probes within the chromosome 1 region and six metabolic traits. We confirm the significant negative correlation between fasting glucose and *Tpcn2* ex-

pression levels ( $r = -0.559$ ,  $P = 0.0000549$ ). Based on a Bonferroni-adjusted significance level of 0.00009, we also find a marginally significant correlation between *RGD1307603*, a gene with no known function, and QUICKI ( $r = 0.544$ ,  $P = 0.00000925$ ). Correlation coefficients and unadjusted  $P$ -values are reported in [Table S6](#).

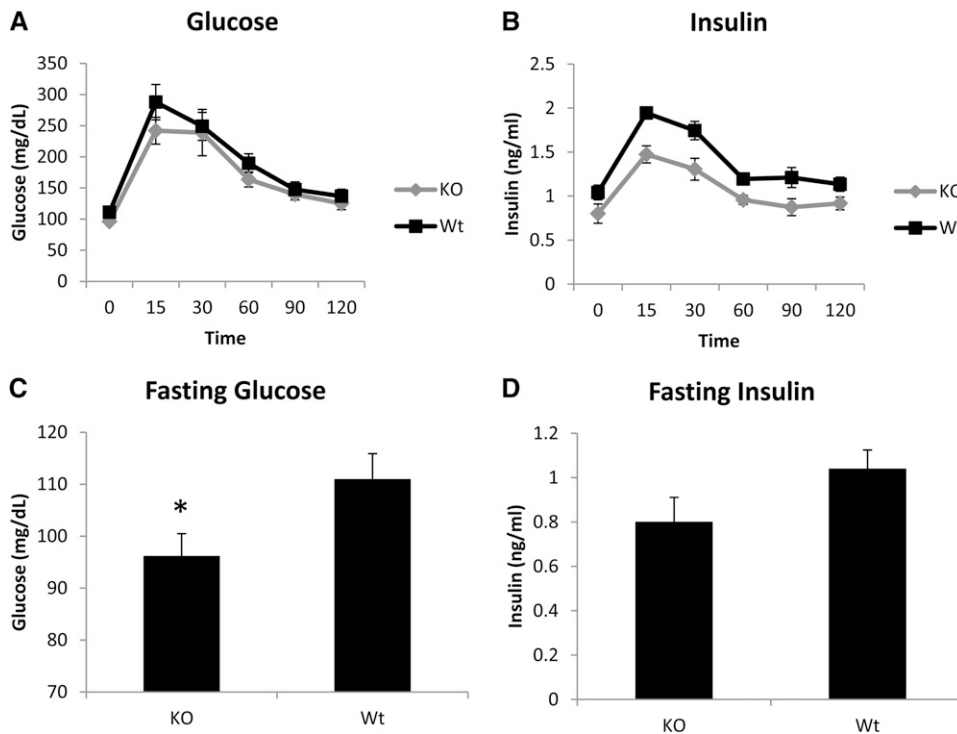
#### ***Tpcn2* knockout mice exhibit decreased fasting glucose and insulin\_AUC**

We find a significant main effect of genotype on insulin\_AUC ( $F_{2,15} = 4.41$ ,  $P = 0.031$ ), with knockout mice exhibiting significantly decreased insulin response relative to wild-type mice (KO = 126.2, Het = 147.2, WT = 165.4,  $P = 0.029$ , Tukey's post-hoc test; see [Figure 6](#) and [Table S7](#)). There is a trend toward significance for fasting glucose (KO = 96.2, Het = 116.3, WT = 111.0,  $F_{2,15} = 3.56$ ,  $P = 0.056$ ) and QUICKI (KO = 0.54, Het = 0.49, WT = 0.49,  $F_{2,15} = 3.32$ ,  $P = 0.066$ ). Upon running a  $t$ -test between knockout and wild-type mice only, we find that *Tpcn2* knockout mice exhibit significantly decreased fasting glucose levels relative to wild-type mice ( $t(8.9) = 2.28$ ,  $P = 0.048$ ). *Tpcn2* knockout mice do not differ significantly from wild-type mice in fasting insulin levels or glucose\_AUC. Means for all traits are reported in [Table S7](#). Raw data for the knockout, wild-type, and heterozygote mice can be accessed in [File S5](#).

#### **Association between *Tpcn2* and fasting insulin and HOMA\_IR in humans**

Although we looked up associations in 39 SNPs, MAGIC was able to provide statistics on only 14 SNPs and DIAGRAM provided statistics on 11 SNPs (see [Table S1](#)). Information





**Figure 6** (A–D) Glucose and insulin levels after a glucose tolerance test (A and B) and at fasting (C and D) in wild-type (WT) and *Tpcn2* knockout (KO) mice. Knockout mice exhibit significantly lower fasting glucose levels (\* $P = 0.048$ ). Although fasting insulin levels do not differ significantly between wild-type and knockout mice, insulin\_AUC is highly significantly different ( $P = 0.031$ ).

was not available for the other SNPs likely because they were either rare variants or failed imputation quality control for other reasons. Of the 14 SNPs from MAGIC, 4 SNPs reach unadjusted levels of significance for fasting insulin, 2 of which remain marginally significant after accounting for multiple comparisons (rs10736671,  $P = 0.051$ , and rs11604251,  $P = 0.051$ ; see Table 2). Three of the 4 SNPs reach unadjusted levels of significance for HOMA\_IR, a measure of insulin sensitivity (Matthews *et al.* 1985), but do not remain significant after adjusting for multiple comparisons. Significance is not found between SNPs within *Tpcn2* and type 2 diabetes or fasting glucose (data not shown).

## Discussion

In our previous work, we used HS rats to identify a 3.1-Mb region that contains a QTL for glucose tolerance and fasting glucose (Solberg Woods *et al.* 2010, 2012) and that overlaps larger QTL for fasting insulin and insulin sensitivity (Solberg Woods *et al.* 2012). In the present study, we use both analysis of gene expression and association analysis to identify *Tpcn2* as a likely causal gene within this locus. These findings are supported by altered glucose and insulin levels in a *Tpcn2* knockout mouse, as well as by an association between variants within *Tpcn2* and fasting insulin levels in humans.

The role of *Tpcn2* in regulating glucose and insulin levels is currently unknown. *Tpcn2* channels localize to the lysosome and are likely receptors for the calcium-mobilizing agent NAADP (Brailoiu *et al.* 2009; Calcraft *et al.* 2009; Schieder *et al.* 2010). Several studies indicate that NAADP

may play a role in insulin signaling of the  $\beta$ -cells (Kim *et al.* 2008; Naylor *et al.* 2009; Shawl *et al.* 2009; Alejandro *et al.* 2010; Arredouani *et al.* 2010), and a recent study suggests NAADP is involved in glucose homeostasis (Park *et al.* 2013). It is also interesting to note that *Tpcn2* and other calcium-handling genes (*Tpcn1* and *IP3R1*) have recently been found to be differentially expressed in patients with heart failure (Garcia-Rua *et al.* 2012). Although previous studies have not looked at the role of *Tpcn2* in liver, alterations of calcium channels such as CaMKII have been shown to lead to impaired liver gluconeogenesis (Ozcan *et al.* 2012), while Cav2.3 calcium channel knockout mice exhibit altered glucose levels, likely as a result of reduced insulin sensitivity (Matsuda *et al.* 2001). Together these studies suggest a possible role for *Tpcn2* in both the  $\beta$ -cell and the liver. Future studies will test the hypothesis that *Tpcn2* is involved in glucose-stimulated insulin secretion of the  $\beta$ -cells as well as probe a potential role of this gene in regulating insulin sensitivity and gluconeogenesis.

Global gene expression analysis in HS rats was used to initially identify *Tpcn2* as a candidate gene within the glucose QTL. We show that *Tpcn2* is differentially expressed in HS rats with glucose intolerance relative to those with normal glucose regulation and demonstrate that this gene maps as a *cis*-eQTL to the same region as the physiological QTL. Several previous studies have found that genes that map as *cis*-eQTL within previously identified physiological QTL can be considered prime candidate genes (Hubner *et al.* 2005; Chen *et al.* 2008; Morrissey *et al.* 2011). Demonstrating a correlation between expression levels and fasting glucose levels further supports that this gene may be playing a causal

**Table 2 Association between *Tpcn2* SNPs and diabetic traits in humans**

SNP	Location	Fasting insulin		HOMA_IR	
		Unadjusted <i>P</i> -value (effect)	BH adjusted <i>P</i> -value	Unadjusted <i>P</i> -value (effect)	BH adjusted <i>P</i> -value
<b>rs11604251</b>	<b>chr11: 68,820,429</b>	<b>0.0073 (+)</b>	<b>0.0511</b>	0.6643 (+)	0.6643
rs7127082	chr11: 68,822,023	0.02475 (–)	0.0937	0.0392 (–)	0.2415
rs753559	chr11: 68,851,605	0.02676 (+)	0.0937	0.03936 (+)	0.2415
<b>rs10736671</b>	<b>chr11: 68,858,907</b>	<b>0.0038 (+)</b>	<b>0.0511</b>	0.05174 (–)	0.2415

Only the 4 most significant SNPs (of 14) are reported. Statistics on the additional SNPs are reported in Table S1. SNPs that remain marginally significant after adjusting for multiple comparisons using the Benjamini and Hochberg FDR method (BH) are in boldface type.

role at the QTL (Farber *et al.* 2009; Leduc *et al.* 2011; Morrissey *et al.* 2011).

Upon assessing association between 18 SNPs within the QTL and traits that map to this region (including five metabolic traits and *Tpcn2* expression levels), we find that the nonsynonymous coding variant within *Tpcn2*, as well as an intronic variant in the neighboring *Rmt1* gene, exhibits the highest level of association with fasting glucose levels and *Tpcn2* expression levels, further supporting a potential causal role for this gene. Despite having a slightly higher level of association, we do not believe that *Rmt1* plays a causal role in diabetes because it is neither correlated with the diabetic traits nor exhibits expression differences. It is interesting to note that three additional SNPs in the region exhibit significant association with *Tpcn2* expression levels, indicating that *Tpcn2* expression may be regulated by other genes in the region. Three of these SNPs, including the *Tpcn2* SNP, are from the F344 founder, the strain that exhibits very low *Tpcn2* expression levels (Figure 2) and that contributes the susceptibility allele at this locus (Solberg Woods *et al.* 2010, 2012). It is unlikely that any of the identified variants are causal, however, because the significance levels of these SNPs do not surpass that of the haplotype mapping for either fasting glucose (Solberg Woods *et al.* 2012) or *Tpcn2* expression levels (Figure 3). A merge analysis, in which genotypes for all SNPs in the region are imputed, may be useful to identify the causative variant (Yalcin *et al.* 2005; Keane *et al.* 2011; Baud *et al.* 2013). However, because this method is most successful when a single causative variant underlies the locus (Baud *et al.* 2013), it may prove less useful for the current region in which multiple genes/variants may be playing a role.

Upon investigating glucose and insulin levels in *Tpcn2* knockout mice, we find that knockout mice exhibit significantly decreased fasting glucose and decreased insulin response to a glucose challenge relative to wild-type mice. These findings contrast with the results in the HS rats, in which decreased *Tpcn2* expression levels are associated with increased, as opposed to decreased, fasting glucose and insulin levels. There are several possible reasons for this discrepancy. Because the mice were created on a mixed C57/129 background, we cannot rule out the possibility that alleles from the 129 strain, as opposed to the gene knockout, are causing the differences in phenotype. To minimize this

effect, however, only siblings from heterozygous breeder pairs were tested. Another possibility is related to the fact that a full gene knockout can have a very different effect on a phenotype than natural allelic variation (Flint and Eskin 2012). Furthermore, several studies have shown that different variants within the same gene, or within the same GWAS location, can have opposite effects on a single phenotype (Cohen *et al.* 2005; Flister *et al.* 2013). To test the role specifically of the F344 variant on glucose and insulin regulation, quantitative complementation can be used (Yalcin *et al.* 2004b, 2010). Alternatively, if we are able to identify the causal variant(s) within this region, we can use genetic techniques to investigate the specific role of this (or these) variant(s) (Katter *et al.* 2013).

To determine whether *Tpcn2* may be playing a role in human diabetes, we collaborated with investigators from the DIAGRAM and MAGIC consortiums. *Tpcn2* has not previously reached significance levels in any of the published genome-wide scans (Morris *et al.* 2012; Scott *et al.* 2012). However, genome-wide thresholds are conservative because of the need to account for the multiple testing across the genome. In the present study, we were able to assess 14 of 39 SNPs within the *Tpcn2* gene and find that two are marginally significantly associated with fasting insulin levels in humans. Although no significant associations were found between these SNPs and diabetes or fasting glucose levels, the majority of SNPs that we investigated were not represented within the MAGIC or DIAGRAM meta-analyses. A potential reason for this is that some of these SNPs may be rare variants and thus could not be properly imputed. It is therefore possible that rare variants within *Tpcn2* may be associated with diabetes and related traits (see Steinthorsdottir *et al.* 2014; Zuk *et al.* 2014). The fact that variants within all three species alter diabetic phenotypes, however, is strong evidence that this gene plays a role in metabolic regulation.

This work demonstrates the power of highly recombinant animal models such as the HS for relative rapid identification of a candidate gene within a QTL. Despite relatively complex genetic architecture at the QTL, we were able to identify a highly plausible causal gene within only a few years. It is important to recognize, however, that we are not able to rule out any other genes within this QTL at this time, particularly the 29 genes that were not assessed on the

expression array. In fact, our mapping studies indicate that more than one gene likely underlies the insulin locus (Solberg Woods *et al.* 2012), and it is reasonable to expect that *Tpcn2* may interact with one or more genes within this region. Of particular interest is *Gpr152*, a G-protein coupled receptor in which the F344 variant encodes a premature stop codon, thus likely affecting protein function. The variant within *Gpr152* was also significantly associated with *Tpcn2* expression levels, suggesting a potential interaction between these genes. Future studies will investigate a potential role for these other genes, as well as potential interactions between genes within this region.

In sum, we have identified *Tpcn2* as a highly plausible causal gene within a QTL for both glucose and insulin traits. We have used expression and sequence analysis in a highly recombinant population to identify this gene in only a few years. A causal role for this gene is supported by alterations in both glucose and insulin in a *Tpcn2* knockout mouse as well as data from clinical cohorts. Future work will identify the causative variant within this gene and begin to unravel the underlying mechanisms involved, as well as investigate a potential role for other genes in the region. The gene expression data, association analysis, and confirmation in both mice and humans demonstrate *Tpcn2* as a new gene contributing to glucose and insulin regulation.

## Acknowledgments

We thank Shawn Levy and Melanie Robinson from Vanderbilt University (currently at Hudson Alpha, Inc.) for running the 10K genotyping arrays. We thank Caitlin O'Meara and Josef Lazar for helpful discussions regarding follow-up studies of *Tpcn2*. This work was funded by grant R01 DK-088975 (to L. C.S.W.), a Basic Research Award from Advancing a Healthier Wisconsin Research (to L.C.S.W.), a Children's Research Institute Pilot and Innovation award (to L.C.S.W.), and funding from the Individualized Medicine Institute at the Medical College of Wisconsin (to L.C.S.W.). MAGIC lookup was funded by grant K24 DK080140 (to J.M.). Knockout mice were created using funding from the Medical Research Council (United Kingdom) programme grant G0901521 (to J.P.).

## Literature Cited

- Aitman, T. J., R. Dong, T. J. Vyse, P. J. Norsworthy, M. D. Johnson *et al.*, 2006 Copy number polymorphism in *Fcgr3* predisposes to glomerulonephritis in rats and humans. *Nature* 439(7078): 851–855.
- Alejandro, E. U., T. B. Kalynyak, F. Taghizadeh, K. S. Gwiazda, E. K. Rawstron *et al.*, 2010 Acute insulin signaling in pancreatic beta-cells is mediated by multiple Raf-1 dependent pathways. *Endocrinology* 151(2): 502–512.
- Arredouani, A., A. M. Evans, J. Ma, J. Parrington, M. X. Zhu *et al.*, 2010 An emerging role for NAADP-mediated Ca<sup>2+</sup> signaling in the pancreatic beta-cell. *Islets* 2(5): 323–330.
- Baud, A., R. Hermsen, V. Guryev, P. Stridh, D. Graham *et al.*, 2013 Combined sequence-based and genetic mapping analysis of complex traits in outbred rats. *Nat. Genet.* 45(7): 767–775.
- Behmoaras, J., G. Bhargal, J. Smith, K. McDonald, B. Mutch *et al.*, 2008 *Jund* is a determinant of macrophage activation and is associated with glomerulonephritis susceptibility. *Nat. Genet.* 40(5): 553–559.
- Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57: 289–300.
- Brailoiu, E., D. Churamani, X. Cai, M. G. Schrlau, G. C. Brailoiu *et al.*, 2009 Essential requirement for two-pore channel 1 in NAADP-mediated calcium signaling. *J. Cell Biol.* 186(2): 201–209.
- Cacho, J., J. Sevellano, J. de Castro, E. Herrera, and M. P. Ramos, 2008 Validation of simple indexes to assess insulin sensitivity during pregnancy in Wistar and Sprague-Dawley rats. *Am. J. Physiol. Endocrinol. Metab.* 295(5): E1269–E1276.
- Calcraft, P. J., M. Ruas, Z. Pan, X. Cheng, A. Arredouani *et al.*, 2009 NAADP mobilizes calcium from acidic organelles through two-pore channels. *Nature* 459(7246): 596–600.
- Chen, Y., J. Zhu, P. Y. Lum, X. Yang, S. Pinto *et al.*, 2008 Variations in DNA elucidate molecular networks that cause disease. *Nature* 452(7186): 429–435.
- Chung, W. K., M. Zheng, M. Chua, E. Kershaw, L. Power-Kehoe *et al.*, 1997 Genetic modifiers of *Lep<sup>rfa</sup>* associated with variability in insulin production and susceptibility to NIDDM. *Genomics* 41(3): 332–344.
- Cohen, J., A. Pertsemlidis, I. K. Kotowski, R. Graham, C. K. Garcia *et al.*, 2005 Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in *PCSK9*. *Nat. Genet.* 37(2): 161–165.
- Demarest, K., J. Koyner, J. McCaughan, Jr., L. Cipp, and R. Hitzemann, 2001 Further characterization and high-resolution mapping of quantitative trait loci for ethanol-induced locomotor activity. *Behav. Genet.* 31(1): 79–91.
- DIAbetes Genetics Replication and Meta-analysis (DIAGRAM) Consortium, Asian Genetic Epidemiology Network Type 2 Diabetes (AGEN-T2D) Consortium, South Asian Type 2 Diabetes (SAT2D) Consortium, Mexican American Type 2 Diabetes (MAT2D) Consortium, Type 2 Diabetes Genetic Exploration by Next-Generation Sequencing in Multi-Ethnic Samples (T2D-GENES) Consortium *et al.*, 2014 Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nat. Genet.* 46(3): 234–244.
- Dupuis, J., C. Langenberg, I. Prokopenko, R. Saxena, N. Soranzo *et al.*, 2010 New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat. Genet.* 42(2): 105–116.
- Farber, C. R., A. van Nas, A. Ghazalpour, J. E. Aten, S. Doss *et al.*, 2009 An integrative genetics approach to identify candidate genes regulating BMD: combining linkage, gene expression, and association. *J. Bone Miner. Res.* 24(1): 105–116.
- Flint, J., and E. Eskin, 2012 Genome-wide association studies in mice. *Nat. Rev. Genet.* 13(11): 807–817.
- Flister, M. J., S. W. Tsaih, C. C. O'Meara, B. Endres, M. J. Hoffman *et al.*, 2013 Identifying multiple causative genes at a single GWAS locus. *Genome Res.* 23(12): 1996–2002.
- Galli, J., L. S. Li, A. Glaser, C. G. Ostenson, H. Jiao *et al.*, 1996 Genetic analysis of non-insulin dependent diabetes mellitus in the GK rat. *Nat. Genet.* 12(1): 31–37.
- Garcia-Rua, V., M. F. Otero, P. V. Lear, D. Rodriguez-Penas, S. Feijoo-Bandin *et al.*, 2012 Increased expression of fatty-acid and calcium metabolism genes in failing human heart. *PLoS ONE* 7(6): e37505.
- Gauguier, D., P. Froguel, V. Parent, C. Bernard, M. T. Bihoreau *et al.*, 1996 Chromosomal mapping of genetic loci associated

- with non-insulin dependent diabetes in the GK rat. *Nat. Genet.* 12(1): 38–43.
- Granhall, C., H. B. Park, H. Fakhrai-Rad, and H. Luthman, 2006 High-resolution quantitative trait locus analysis reveals multiple diabetes susceptibility loci mapped to intervals <800 kb in the species-conserved Niddm1i of the GK rat. *Genetics* 174: 1565–1572.
- Hansen, C., and K. Spuhler, 1984 Development of the National Institutes of Health genetically heterogeneous rat stock. *Alcohol. Clin. Exp. Res.* 8(5): 477–479.
- Hessner, M. J., X. Wang, L. Meyer, R. Geoffrey, S. Jia *et al.*, 2004 Involvement of eotaxin, eosinophils, and pancreatic predisposition in development of type 1 diabetes mellitus in the BioBreeding rat. *J. Immunol.* 173(11): 6993–7002.
- Hu, X., and M. Daly, 2012 What have we learned from six years of GWAS in autoimmune diseases, and what is next? *Curr. Opin. Immunol.* 24(5): 571–575.
- Huang, G. J., S. Shifman, W. Valdar, M. Johannesson, B. Yalcin *et al.*, 2009 High resolution mapping of expression QTLs in heterogeneous stock mice in multiple tissues. *Genome Res.* 19(6): 1133–1140.
- Hubner, N., C. A. Wallace, H. Zimdahl, E. Petretto, H. Schulz *et al.*, 2005 Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. *Nat. Genet.* 37(3): 243–253.
- Johnsen, A. K., W. Valdar, L. Golden, A. Ortiz-Lopez, R. Hitzemann *et al.*, 2011 Genome-wide and species-wide dissection of the genetics of arthritis severity in heterogeneous stock mice. *Arthritis Rheum.* 63(9): 2630–2640.
- Kadar, B., K. Gombos, E. Szele, I. Ember, J. L. Ivanyi *et al.*, 2011 Effects of isoflurane on Nfκappab p65, Gadd45a and Jnk1 expression in the vital organs of CBA/CA mice. *In Vivo* 25(2): 241–244.
- Kanemoto, N., H. Hishigaki, A. Miyakita, K. Oga, S. Okuno *et al.*, 1998 Genetic dissection of “OLETF”, a rat model for non-insulin-dependent diabetes mellitus. *Mamm. Genome* 9(6): 419–425.
- Katter, K., A. M. Geurts, O. Hoffmann, L. Mates, V. Landa *et al.*, 2013 Transposon-mediated transgenesis, transgenic rescue, and tissue-specific gene expression in rodents and rabbits. *FASEB J.* 27(3): 930–941.
- Katz, A., S. S. Nambi, K. Mather, A. D. Baron, D. A. Follmann *et al.*, 2000 Quantitative insulin sensitivity check index: a simple, accurate method for assessing insulin sensitivity in humans. *J. Clin. Endocrinol. Metab.* 85(7): 2402–2410.
- Keane, T. M., L. Goodstadt, P. Danecek, M. A. White, K. Wong *et al.*, 2011 Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature* 477(7364): 289–294.
- Kim, B. J., K. H. Park, C. Y. Yim, S. Takasawa, H. Okamoto *et al.*, 2008 Generation of nicotinic acid adenine dinucleotide phosphate and cyclic ADP-ribose by glucagon-like peptide-1 evokes Ca<sup>2+</sup> signal that is essential for insulin secretion in mouse pancreatic islets. *Diabetes* 57(4): 868–878.
- Koressaar, T., and M. Remm, 2007 Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23(10): 1289–1291.
- Leduc, M. S., R. S. Hageman, R. A. Verdugo, S. W. Tsaih, K. Walsh *et al.*, 2011 Integration of QTL and bioinformatic tools to identify candidate genes for triglycerides in mice. *J. Lipid Res.* 52(9): 1672–1682.
- Logan, R. W., R. F. Robledo, J. M. Recla, V. M. Philip, J. A. Bubier *et al.*, 2013 High-precision genetic mapping of behavioral traits in the diversity outbred mouse population. *Genes Brain Behav.* 12(4): 424–437.
- Matsuda, Y., H. Saegusa, S. Zong, T. Noda, and T. Tanabe, 2001 Mice lacking Ca(v)2.3 (alpha1E) calcium channel exhibit hyperglycemia. *Biochem. Biophys. Res. Commun.* 289(4): 791–795.
- Matthews, D. R., J. P. Hosker, A. S. Rudenski, B. A. Naylor, D. F. Treacher *et al.*, 1985 Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia* 28(7): 412–419.
- Morris, A. P., B. F. Voight, T. M. Teslovich, T. Ferreira, A. V. Segre *et al.*, 2012 Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat. Genet.* 44(9): 981–990.
- Morrissey, C., I. C. Grieve, M. Heinig, S. Atanur, E. Petretto *et al.*, 2011 Integrated genomic approaches to identification of candidate genes underlying metabolic and cardiovascular phenotypes in the spontaneously hypertensive rat. *Physiol. Genomics* 43(21): 1207–1218.
- Mott, R., C. J. Talbot, M. G. Turri, A. C. Collins, and J. Flint, 2000 A method for fine mapping quantitative trait loci in outbred animal stocks. *Proc. Natl. Acad. Sci. USA* 97(23): 12649–12654.
- Naylor, E., A. Arredouani, S. R. Vasudevan, A. M. Lewis, R. Parkesh *et al.*, 2009 Identification of a chemical probe for NAADP by virtual screening. *Nat. Chem. Biol.* 5(4): 220–226.
- Nobrega, M. A., L. C. Woods, S. Fleming, and H. J. Jacob, 2009 Distinct genetic regulation of progression of diabetes and renal disease in the Goto-Kakizaki rat. *Physiol. Genomics* 39(1): 38–46.
- Ozcan, L., C. C. Wong, G. Li, T. Xu, U. Pajvani *et al.*, 2012 Calcium signaling through CaMKII regulates hepatic glucose production in fasting and obesity. *Cell Metab.* 15(5): 739–751.
- Park, K. H., B. J. Kim, A. I. Shawl, M. K. Han, H. C. Lee *et al.*, 2013 Autocrine/paracrine function of nicotinic acid adenine dinucleotide phosphate (NAADP) for glucose homeostasis in pancreatic beta-cells and adipocytes. *J. Biol. Chem.* 288(49): 35548–35558.
- Petretto, E., R. Sarwar, I. Grieve, H. Lu, M. K. Kumaran *et al.*, 2008 Integrated genomic approaches implicate osteoglycin (Ogn) in the regulation of left ventricular mass. *Nat. Genet.* 40(5): 546–552.
- Pfaffl, M. W., 2001 A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29(9): e45.
- Saar, K., A. Beck, M. T. Bihoreau, E. Birney, D. Brocklebank *et al.*, 2008 SNP and haplotype mapping for genetic analysis in the rat. *Nat. Genet.* 40(5): 560–566.
- Samuelson, D. J., S. E. Hesselton, B. A. Aperavich, Y. Zan, J. D. Haag *et al.*, 2007 Rat Mcs5a is a compound quantitative trait locus with orthologous human loci that associate with breast cancer risk. *Proc. Natl. Acad. Sci. USA* 104(15): 6299–6304.
- Schieder, M., K. Rotzer, A. Bruggemann, M. Biel, and C. A. Wahl-Schott, 2010 Characterization of two-pore channel 2 (TPCN2)-mediated Ca<sup>2+</sup> currents in isolated lysosomes. *J. Biol. Chem.* 285(28): 21219–21222.
- Scott, R. A., V. Lagou, R. P. Welch, E. Wheeler, M. E. Montasser *et al.*, 2012 Large-scale association analyses identify new loci influencing glycemic traits and provide insight into the underlying biological pathways. *Nat. Genet.* 44(9): 991–1005.
- Shawl, A. I., K. H. Park, and U. H. Kim, 2009 Insulin receptor signaling for the proliferation of pancreatic beta-cells: involvement of Ca<sup>2+</sup> second messengers, IP3, NAADP and cADPR. *Islets* 1(3): 216–223.
- So, H. C., A. H. Gui, S. S. Cherny, and P. C. Sham, 2011 Evaluating the heritability explained by known susceptibility variants: a survey of ten complex diseases. *Genet. Epidemiol.* 35(5): 310–317.
- Solberg Woods, L. C., 2014 QTL mapping in outbred populations: successes and challenges. *Physiol. Genomics* 46(3): 81–90.

- Solberg Woods, L. C., N. Ahmadiyeh, A. Baum, K. Shimomura, Q. Li *et al.*, 2009 Identification of genetic loci involved in diabetes using a rat model of depression. *Mamm. Genome* 20(8): 486–497.
- Solberg Woods, L. C., K. Holl, M. Tschannen, and W. Valdar, 2010 Fine-mapping a locus for glucose tolerance using heterogeneous stock rats. *Physiol. Genomics* 41(1): 102–108.
- Solberg Woods, L. C., K. L. Holl, D. Oreper, Y. Xie, S. W. Tsaih *et al.*, 2012 Fine-mapping diabetes-related traits, including insulin resistance, in heterogeneous stock rats. *Physiol. Genomics* 44(21): 1013–1026.
- Steinthorsdottir, V., G. Thorleifsson, P. Sulem, H. Helgason, N. Grarup *et al.*, 2014 Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. *Nat. Genet.* 46(3): 294–298.
- Svenson, K. L., D. M. Gatti, W. Valdar, C. E. Welsh, R. Cheng *et al.*, 2012 High-resolution genetic mapping using the Mouse Diversity outbred population. *Genetics* 190: 437–447.
- Swanberg, M., O. Lidman, L. Padyukov, P. Eriksson, E. Akesson *et al.*, 2005 MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat. Genet.* 37(5): 486–494.
- Talbot, C. J., A. Nicod, S. S. Cherny, D. W. Fulker, A. C. Collins *et al.*, 1999 High-resolution mapping of quantitative trait loci in outbred mice. *Nat. Genet.* 21(3): 305–308.
- Talbot, C. J., R. A. Radcliffe, J. Fullerton, R. Hitzemann, J. M. Wehner *et al.*, 2003 Fine scale mapping of a genetic locus for conditioned fear. *Mamm. Genome* 14(4): 223–230.
- Taylor, C. T., and E. P. Cummins, 2011 Regulation of gene expression by carbon dioxide. *J. Physiol.* 589(Pt. 4): 797–803.
- Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth *et al.*, 2012 Primer3–new capabilities and interfaces. *Nucleic Acids Res.* 40(15): e115.
- Valdar, W., L. C. Solberg, D. Gauguier, S. Burnett, P. Klenerman *et al.*, 2006 Genome-wide genetic association of complex traits in heterogeneous stock mice. *Nat. Genet.* 38(8): 879–887.
- Valdar, W., C. C. Holmes, R. Mott, and J. Flint, 2009 Mapping in structured populations by resample model averaging. *Genetics* 182(4): 1263–1277.
- Voight, B. F., L. J. Scott, V. Steinthorsdottir, A. P. Morris, C. Dina *et al.*, 2010 Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat. Genet.* 42(7): 579–589.
- Wallis, R. H., K. Wang, L. Marandi, E. Hsieh, T. Ning *et al.*, 2009 Type 1 diabetes in the BB rat: a polygenic disease. *Diabetes* 58(4): 1007–1017.
- Wei, S., K. Wei, D. H. Moralejo, T. Ogino, G. Koike *et al.*, 1999 Mapping and characterization of quantitative trait loci for non-insulin-dependent diabetes mellitus with an improved genetic map in the Otsuka Long-Evans Tokushima fatty rat. *Mamm. Genome* 10(3): 249–258.
- Yalcin, B., J. Fullerton, S. Miller, D. A. Keays, S. Brady *et al.*, 2004a Unexpected complexity in the haplotypes of commonly used inbred strains of laboratory mice. *Proc. Natl. Acad. Sci. USA* 101(26): 9734–9739.
- Yalcin, B., S. A. Willis-Owen, J. Fullerton, A. Meesaq, R. M. Deacon *et al.*, 2004b Genetic dissection of a behavioral quantitative trait locus shows that *Rgs2* modulates anxiety in mice. *Nat. Genet.* 36(11): 1197–1202.
- Yalcin, B., J. Flint, and R. Mott, 2005 Using progenitor strain information to identify quantitative trait nucleotides in outbred mice. *Genetics* 171: 673–681.
- Yalcin, B., J. Nicod, A. Bhomra, S. Davidson, J. Cleak *et al.*, 2010 Commercially available outbred mice for genome-wide association studies. *PLoS Genet.* 6(9): e1001085.
- Zeggini, E., L. J. Scott, R. Saxena, B. F. Voight, J. L. Marchini *et al.*, 2008 Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* 40(5): 638–645.
- Zuk, O., S. F. Schaffner, K. Samocha, R. Do, E. Hechter *et al.*, 2014 Searching for missing heritability: designing rare variant association studies. *Proc. Natl. Acad. Sci. USA* 111(4): E455–E464.

*Communicating editor: S. F. Chenoweth*

# GENETICS

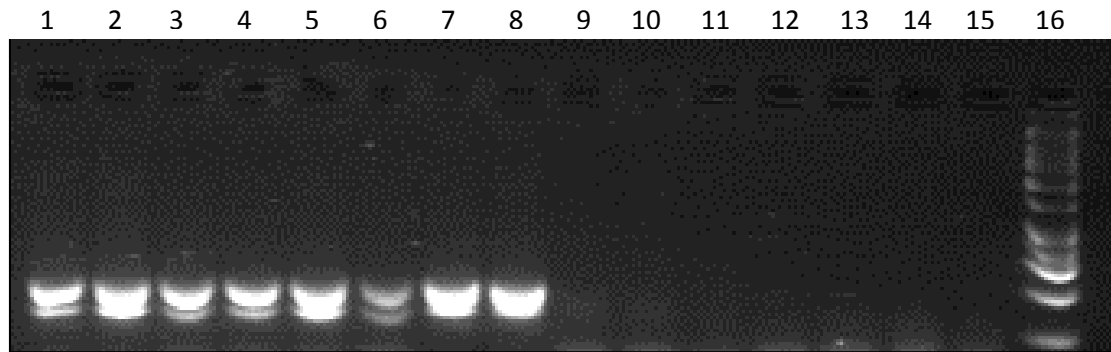
Supporting Information

<http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.162982/-/DC1>

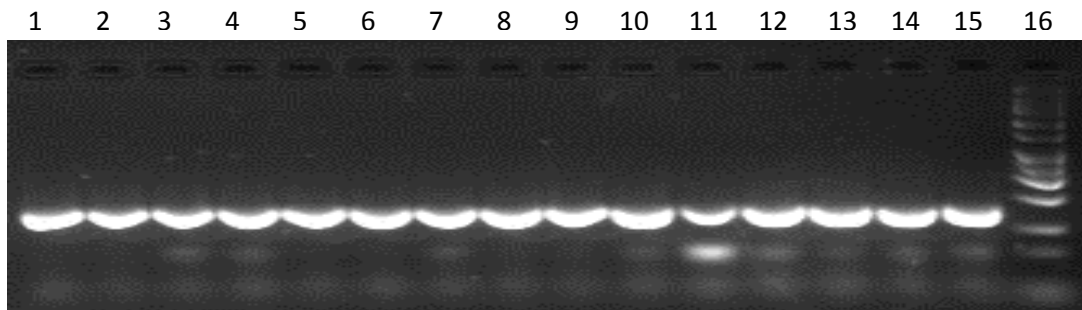
## Identification of a Novel Gene for Diabetic Traits in Rats, Mice, and Humans

Shirng-Wern Tsaih, Katie Holl, Shuang Jia, Mary Kaldunski, Michael Tschannen, Hong He, Jaime Wendt Andrae, Shun-Hua Li, Alex Stoddard, Andrew Wiederhold, John Parrington, Margarida Ruas da Silva, Antony Galione, James Meigs, Meta-Analyses of Glucose and Insulin-Related Traits Consortium (MAGIC) Investigators, Raymond G. Hoffmann, Pippa Simpson, Howard Jacob, Martin Hessner, and Leah C. Solberg Woods

cRNA PCR on *Tpcn2*



cRNA PCR on GAPDH



Lane	Genotype	Tissue cRNA is from
1	Wild Type	Kidney
2	Wild Type	Kidney
3	Wild Type	Retroperitoneal Fat
4	Wild Type	Retroperitoneal Fat
5	Wild Type	Heart
6	Wild Type	Heart
7	Wild Type	Liver
8	Wild Type	Liver
9	Knockout	Kidney
10	Knockout	Kidney
11	Knockout	Retroperitoneal Fat
12	Knockout	Heart
13	Knockout	Heart
14	Knockout	Liver
15	Knockout	Liver
16	Size Standard	

**Figure S1** PCR on RNA from multiple tissues in wild-type and knock-out *Tpcn2* mice. Results confirm that the knock-out mice do not transcribe *Tpcn2*.

## File S1

### **\*\* List of MAGIC (Meta-Analysis of Glucose and Insulin-Related traits) Consortium investigators**

File S1 is available for download as a pdf at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.162982/-/DC1>.

## Files S2-S5

**Available for download as Excel files at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.162982/-/DC1>**

**File S2** Phenotype data used for the cluster analysis

**File S3** HS genotype data for 18 SNPs within the 3.1 Mb QTL on rat chromosome 1

**File S4** Quantitative rt-PCR data for *Tpcn2* and GAPDH in 120 HS rats

**File S5** Glucose and insulin levels during the IPGTT in *Tpcn2* knock-out, heterozygous and wild-type mice



**Table S1 Association between *Tpcn2* SNPs and diabetic traits in humans**

SNP	Location (Mb)	Functional location	Fasting Insulin		HOMA_IR	
			Unadjusted p-value (effect)	BH adjusted p-value	Unadjusted p-value (effect)	BH adjusted p-value
rs2376554	Chr 11:68811445	upstream				
rs749363	Chr. 11:68812024	upstream				
rs7932947	Chr. 11:68812781	upstream				
rs55918435	Chr. 11:68813015	upstream				
rs35678276	Chr. 11:68814544	upstream				
rs59234009	Chr. 11:68814545	upstream				
rs56932042	Chr. 11:68814568	upstream				
rs10792010	Chr. 11:68814887	upstream				
rs3018679	Chr. 11:68814894	upstream				
rs59427626	Chr. 11:68815258	upstream				
rs35443329	Chr. 11:68815437	upstream				
<b>rs10750836</b>	Chr. 11:68815523	upstream	0.5689	0.5689	0.6327	0.6643
rs12790685	Chr. 11:68815948	upstream				
rs12791187	Chr. 11:68815949	upstream				
rs4930642	Chr. 11:68816370	UTR 5'				
rs4453241	Chr. 11:68816825	intron				
rs11823877	Chr. 11:68819011	intron				
rs11228458	Chr. 11:68819204	intron				
<b>rs11604251</b>	Chr. 11:68820429	intron	0.0073	0.0511*	0.6643	0.6643
<b>rs10792012</b>	Chr. 11:68820569	intron	0.5117	0.5689	0.6216	0.6643
rs6591367	Chr. 11:68821947	intron				
<b>rs7127082</b>	Chr. 11:68822023	intron	0.02475	0.0937	0.0392	0.2415
<b>rs7942690</b>	Chr. 11:68823940	intron	0.5004	0.5689	0.6251	0.6643
<b>rs3019776</b>	Chr. 11:68826155	intron	0.5221	0.5689	0.2321	0.6499
rs7930211	Chr. 11:68827717	intron				
<b>rs11228469</b>	Chr. 11:68829638	intron	0.0868	0.2430	0.1142	0.3997
<b>rs3829236</b>	Chr. 11:68838028	intron	0.3042	0.5689	0.3242	0.6643
rs4930263	Chr. 11:68847005	intron				
rs2290418	Chr. 11:68847687	intron				
rs2376558	Chr. 11:68851414	missense				
<b>rs753559</b>	Chr. 11:68851605	intron	0.02676	0.0937	0.03936	0.2415
<b>rs3750967</b>	Chr. 11:68855112	intron	0.4648	0.5689	0.5844	0.6643
<b>rs1060435</b>	Chr. 11:68855595	UTR 3'	0.2497	0.5689	0.507	0.6643

<b>rs4930265</b>	Chr. 11:68855954	UTR 3'	0.5297	0.5689	0.5493	0.6643
rs3018681	Chr. 11:68856224	UTR 3'				
<b>rs3168115</b>	Chr. 11:68858090	downstream	0.5077	0.5689	0.5037	0.6643
<b>rs10736671</b>	Chr. 11:68858907	downstream	0.0038	0.0511*	0.05174	0.2415
rs3018673	Chr. 11:68859608	downstream				
rs7930709	Chr. 11:68860223	downstream				

39 human SNPs within or near *Tpcn2* were used to determine association with diabetes and related traits. DIAGRAM assessed association between 11 SNPs (in italics) and diabetes while MAGIC assessed association between 14 SNPs (in bold) and fasting glucose, fasting insulin and the homeostatic model assessment of insulin resistance (HOMA\_IR). Associations between 14 SNPs and fasting insulin and HOMA\_IR, as assessed by MAGIC, are listed above. Multiple comparisons were adjusted for using the Benjamini and Hochberg FDR method (BH). No significant associations were found between the SNPs and diabetes or fasting glucose. Associations could not be determined for SNPs that are not in bold because of insufficient data. \* p = 0.051 after adjusting for multiple corrections.

**Table S2 Phenotypic differences between glucose intolerant and normal glucose HS rats**

<b>Trait</b>	<b>Normal Glucose HS (mean <math>\pm</math> SD)</b>	<b>Glucose Intolerant HS (mean <math>\pm</math> SD)</b>
<b>Glucose_AUC</b>	9061.0 $\pm$ 209.1	13,458.9 $\pm$ 1,035.5***
<b>Insulin_AUC</b>	204.3 $\pm$ 63.3	311.3 $\pm$ 106.7***
<b>Fasting glucose (mg/dL)</b>	77.5 $\pm$ 6.6	90.3 $\pm$ 6.7***
<b>Fasting insulin (ng/ml)</b>	1.1 $\pm$ 0.5	1.8 $\pm$ 1.2**
<b>Fasting c-peptide (pg/ml)</b>	1505.8 $\pm$ 479.8	1808.0 $\pm$ 367.8*
<b>Fasting cholesterol (mg/dL)</b>	100.1 $\pm$ 46.8	125.5 $\pm$ 56.0
<b>Fasting triglycerides (mg/dL)</b>	72.3 $\pm$ 20.1	86.2 $\pm$ 12.2**
<b>QUICKI (1/(logI0 + logG0))</b>	0.54 $\pm$ 0.06	0.48 $\pm$ 0.08**
<b>Body Weight (g)</b>	314.4 $\pm$ 25.8	349.1 $\pm$ 33.5***
<b>Retroperitoneal fat pad weight (g)</b>	0.014 $\pm$ 0.004	0.014 $\pm$ 0.005
<b>Epidydimal fat pad weight (g)</b>	0.015 $\pm$ 0.005	0.016 $\pm$ 0.005

\*\*\*p < 0.0001, \*\*p < 0.01, \*p < 0.05, see manuscript for full statistics

**Tables S3-S6**

Available for download as .docx files at <http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.162982/-/DC1>

**Table S3** Differentially Expressed Genes between Glucose Intolerant and Normal Glucose HS rats (Genome-wide)

**Table S4** Mean expression levels for all probes within the 3.1 Mb chromosome 1 region

**Table S5** Significance of association ( $-\log_{10}P$ ) between 18 SNPs within the 3.1 Mb QTL on rat chromosome 1 and all traits that map to this locus

**Table S6** Correlations between probes within the chromosome 1 region and six diabetic traits

**Table S7 Phenotypic differences in *Tpcn2* knock-out, heterozygote and wild-type mice**

<b>Trait</b>	<b>Wild-type (mean <math>\pm</math> SD)</b>	<b>Heterozygotes (mean <math>\pm</math> SD)</b>	<b>Knock-out (mean <math>\pm</math> SD)</b>
<b>Glucose_AUC</b>	22,917.5 $\pm$ 4,384.2	21,811.2 $\pm$ 1,609.2	20,700.0 $\pm$ 3,660.1
<b>Insulin_AUC</b>	165.4 $\pm$ 13.4	147.2 $\pm$ 35.8	126.2 $\pm$ 6.4**
<b>Fasting glucose (mg/dL)</b>	111.0 $\pm$ 12.0	116.3 $\pm$ 15.5	96.2 $\pm$ 9.6*
<b>Fasting insulin (ng/ml)</b>	1.04 $\pm$ 0.22	1.03 $\pm$ 0.51	0.80 $\pm$ 0.25
<b>QUICKI (1/(logI0 + logG0))</b>	0.49 $\pm$ 0.02	0.49 $\pm$ 0.04	0.54 $\pm$ 0.04
<b>Body Weight (g)</b>	25.3 $\pm$ 2.1	24.3 $\pm$ 2.4	27.0 $\pm$ 3.6

\*\*p = 0.031 (wild-type mice differ from knock-out mice based on a Tukey-post-hoc based after an ANOVA between all three genotypes), \*p = 0.048 (t-test between wild-type and knock-out mice only)