American Genetic Association

OXFORD

Journal of Heredity, 2019, 793–800
doi:10.1093/jhered/esz059
Original Article
Advance Access publication October 7, 2019

Original Article

An Exploratory Association Analysis of the Insulin Gene Region With Diabetes Mellitus in Two Dog Breeds

Rebecka Hess, Paula Henthorn, Marcella Devoto, Fan Wang, and Rui Feng

From the Department of Clinical Sciences & Advanced Medicine, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104 (Hess and Henthorn); Division of Human Genetics, The Children's Hospital of Philadelphia, Philadelphia, PA 19104 (Devoto); Department of Translational and Precision Medicine, University of Rome Sapienza, Rome, Italy (Devoto); Department of Molecular Cardiology, Cleveland Clinic Lerner Research Institute, Cleveland, OH 44106 (Wang); Department of Biostatistics and Epidemiology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104 (Feng).

Address correspondence to Rebecka Hess at the address above, or e-mail: rhess@vet.upenn.edu.

Received June 27, 2019; First decision August 6 2019; Accepted October 3, 2019.

Corresponding Editor: Ernest Bailey

Abstract

Samoyeds and Australian Terriers are the 2 dog breeds at highest risk (>10-fold) for diabetes mellitus in the United States. It is unknown if the insulin (*INS*) gene is involved in the pathophysiology of diabetes in Samoyeds and Australian Terriers. It was hypothesized that the *INS* gene region provides a common genetic causality for diabetes in Samoyeds and Australian Terriers. We conducted a 2-stage genetic association study involving both breeds. In the discovery stage (Stage 1), Samoyeds with and without diabetes were compared in the frequencies of 447 tagging single-nucleotide polymorphisms (SNPs) within 2.5 megabases (Mb) up- and downstream of the *INS* gene on the Illumina CanineHD BeadChip. SNPs yielding a *P*-value < 0.005 were selected for further follow-up. In the validation stage (Stage 2), Australian Terriers with and without diabetes were compared in the SNPs genotyped by the Affymetrix GeneChip Canine Genome 2.0 Array and within 1 Mb up- and downstream of the selected SNPs from Stage 1. Two SNPs that were in high linkage disequilibrium (LD, r^2 = 0.7) were selected from Stage 1. In Stage 2, among the 76 SNPs examined, 5 were significantly associated with diabetes after Bonferroni's correction for multiple comparisons. Three of these 5 SNPs were in complete LD (r^2 = 1 for all associations) and the 2 remaining SNPs were in moderate LD (r^2 = 0.4). In conclusion, an association between the *INS* gene region and diabetes was suggested in 2 dog breeds of different clades. This region could have importance in diabetes in other breeds or in canine diabetes at large.

Subject areas: Genomics and gene mapping
Keywords: Australian Terrier, genetic, Samoyed, single-nucleotide polymorphism

© The American Genetic Association 2019.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

The etiology of canine diabetes mellitus is, in part, genetic, as demonstrated by the strong disease risk in some pure breed dogs. In the United States, Samoyeds and Australian Terriers are 12 and 32 times more likely to develop diabetes, respectively, compared to mixed breed dogs (Hess et al. 2000; Guptill et al. 2003). Samoyeds are also at risk for the disease in Sweden and the United Kingdom (Fall et al. 2007; Catchpole et al. 2013). Additionally, Australian Terriers are known to be at risk for the disease in Sweden (Fall et al. 2007). Two original studies have investigated the insulin (*INS*) gene region in dogs with diabetes and had a similar design that focused on a number of single-nucleotide polymorphisms (SNPs) within or near the *INS* gene in breed-matched diabetic cases and healthy control dogs (Short et al. 2007, 2014). Both studies included Samoyeds but no Australian Terriers (Short et al. 2007, 2014). Two different *INS* variants were identified as protective against diabetes in Cocker Spaniels and Labrador Retrievers, one of these SNPs increased susceptibility to diabetes in Jack Russell Terriers, and a different *INS* variant increased the risk of diabetes in Cocker Spaniels (Short et al. 2007, 2014). Variation at the insulin-like growth factor 2 gene was found to be protective in Border Terriers (Short et al. 2007). These studies focused on SNPs within, or in close proximity to, a specific candidate gene of interest (mostly within 1.5 Kb of exon 1) (Short et al. 2007, 2014). None of the above *INS* gene associations have been replicated in more than 1 breed and none have been reported in Samoyeds or Australian Terriers with diabetes. The goal of this study was therefore to investigate and replicate an association between a large *INS* gene region and diabetes in Samoyeds and Australian Terriers, 2 breeds from different clades. This association, replicated in Samoyeds and Australian Terriers, is reported here. The *INS* gene was chosen because it is associated with many types of diabetes in humans, and a gene with a major role in the pathogenesis of all forms of diabetes was sought for this first-pass canine study (Bradfield et al. 2011; Saxena et al. 2012; Moritani et al. 2013; Elboudwarej et al. 2016; Huopio et al. 2016; Piccini et al. 2016; Yang and Chan 2016; Mishra et al. 2017). In this study, a large region of 5 megabases (Mb) surrounding the *INS* gene was investigated because linkage disequilibrium (LD) can span several Mb in pure breed dogs (Lindblad-Toh et al. 2005; Hoeppner et al. 2014; Hayward et al. 2016).

Materials and Methods

Dogs were defined as diabetic (cases) if the owner and primary veterinarian confirmed that the dog had insulin-treated diabetes. Dogs were classified as nondiabetic (controls) if the owner and primary veterinarian reported that the dog had no clinical signs suggestive of diabetes and if the dog was not diagnosed with the disease. Owners reported the health status of their dog and other dog-related data on a standardized questionnaire which included questions about the dog's age, sex, neuter status, and if applicable, date of diabetes diagnosis and insulin treatment regimen.

Cases and controls were matched by breed in order to maximize the likelihood that differences between case and control dogs were related to disease status rather than breed differences. Cases were enrolled at any age. However, controls were enrolled only if they were 9 years of age or older to decrease the likelihood that they will develop diabetes later in life. Only dogs residing in the United States were included because geography, population bottlenecks, and intense inbreeding in pure breed dogs can influence genetic risk of disease (Lindblad-Toh et al. 2005; Parker et al. 2017). First-degree relatives were excluded from the same group (case or control), but were included in the study if one had diabetes and the other did not. Demographics of the dogs included in the study are reported in Table 1.

The study protocol and owner consent form were approved by the University of Pennsylvania Privately Owned Animal Protocol Committee. Most blood samples were drawn by the dog's primary care veterinarian, and were shipped overnight to the School of Veterinary Medicine at the University of Pennsylvania in lavender top EDTA glass tubes. Occasionally, blood was collected from the patient population of the School of Veterinary Medicine at the University of Pennsylvania and at Samoyed breed club events.

Standard DNA isolation procedure was performed using QIAamp DNA Blood Midi Kit (QIAGEN) or Puregene DNA Purification Kit (Gentra). Working dilutions with a concentration of 50 µg/mL DNA were made from the DNA stock for each dog. Whole-genome tag SNP genotyping was performed with the Illumina CanineHD BeadChip or Affymetrix GeneChip Canine Genome 2.0 Array in Samoyeds and Australian Terriers, respectively (Awano et al. 2009; Parker et al. 2017). For both arrays, SNPs were eligible for statistical analyses if they had a minor allele frequency greater than 1% and a missing rate less than 5%. Although the whole genome was sequenced, a genome-wide association study was not performed due to small sample size. This study focused only on the sequence of the *INS* gene region. The *INS* gene was chosen because it is associated with numerous types of diabetes in humans and because canine diabetes does not perfectly resemble a single one of these forms of diabetes (Bradfield et al. 2011; Saxena et al. 2012; Moritani et al. 2013; Elboudwarej et al. 2016; Huopio et al. 2016; Piccini et al. 2016; Yang and Chan 2016; Mishra et al. 2017).

Table 1. Demographics of study dogs

	Stage 1: Samoyeds		Stage 2: Australian Terriers	
	Cases (n = 30)	Controls (n = 32)	Cases (n = 26)	Controls (n = 33)
Age [median (range), years]				
At blood collection	9 (5–14)	12 (10–16)	10 (0.8–17)	12 (9–16)
At diabetes onset	8 (4–14)	NA	8 (0.8–16)	NA
Sex [n (%)]				
Neutered female	14 (47%)	17 (53%)	12 (46%)	15 (46%)
Intact female	4 (13%)	1 (3%)	1 (4%)	3 (9%)
Neutered male	12 (40%)	12 (38%)	11 (42%)	7 (21%)
Intact male	0 (0%)	2 (6%)	2 (8%)	8 (24%)

NA, not applicable.

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

The 2-stage strategy is a tactical choice used to maximize the potential power given a limited sample size. Since only 62 Samoyeds and 59 Australian Terriers were enrolled in the study and were genotyped using 2 platforms with little overlap, a modified 2-stage procedure was used (Figure 1) (Satagopan et al. 2004; Christie et al. 2012). In Stage 1 (the discovery stage), all SNP markers on the Illumina assay within the genetic region of interest were evaluated for an association with diabetes on 30 Samoyed cases and 32 Samoyed controls, and the most promising markers ($P < 0.005$) were selected for further evaluation. The Illumina assay was chosen for Stage 1 because it is denser than the Affymetrix assay and therefore more suitable for the discovery stage. A region of 2.5 Mb up- and downstream of the *INS* gene was chosen for Stage 1 because LD in pure breed dogs is high, and r^2 declines to a baseline level of unlinked loci at 5–13 Mb in different breeds (Hayward et al. 2016). For each SNP within ±2.5 Mb of the *INS* gene, its association with diabetes was tested using a generalized linear mixed model implemented in Efficient Mixed-Model Association eXpedited (EMMAX) (Kang et al. 2010). Pairwise coefficients of inbreeding among all dogs were estimated from all genome-wide SNPs using the King software (Manichaikul et al. 2010). Control for relatedness was accomplished by including a random component in EMMAX, which was assumed to follow a normal distribution with the variance–covariance formed by the inbreeding coefficients. In Stage 2 (the validation stage), all SNPs genotyped on the Affymetrix assay within ±1 Mb of the selected SNPs from Stage 1 were tested for an association with diabetes in 26 Australian Terrier cases and 33 Australian Terrier controls. The same quality control procedure and the same linear mixed model employed in Stage 1 were applied in Stage 2. An area of ±1 Mb was chosen in Stage 2 because this area provided achievable power for testing of about 50 independent SNPs. To correct for multiple testing, the effective number of independent tests in the Stage 2 region was determined using genetic type I error correction (http://grass.cgs.hku.hk/gec/) (Li et al. 2012). Bonferroni's correction with the appropriate number of independent tests computed with genetic type I error correction was then used to calculate the significance level for SNPs in this region. Pairwise squared Pearson coefficient of correlation (r^2) for SNPs within the *INS* gene region was estimated using PLINK (Purcell et al. 2007).

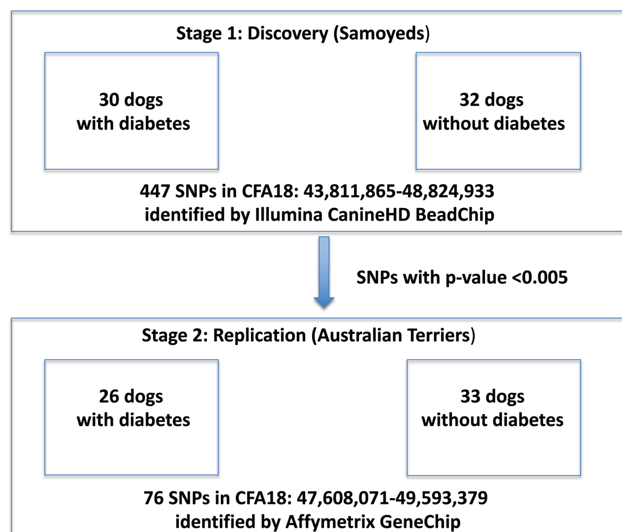

Figure 1. Overview of the study design. SNP, single-nucleotide polymorphism; CFA, *Canis familiaris* chromosome.

Results

The *INS* gene is located on *Canis familiaris* chromosome (CFA)18: 46,311,865–46,324,933 (CanFam 3.1, http://genome.ucsc.edu/) (Lindblad-Toh et al. 2005). In the area within ±2.5 Mb of the *INS* gene (CFA18: 43,811,865–48,824,933) 447 SNPs were present on the Illumina assay used in Samoyeds. Figure 2a represents a regional association plot of $-\log_{10}(P\text{-value})$ in the Stage 1 Samoyed discovery cohort. Two of these 447 SNPs located at CFA18: 48,578,245 and CFA18: 48,632,489 yielded a P-value < 0.005 ($P = 0.0027$ and $P = 0.0033$, respectively), were in high LD ($r^2 = 0.69$), and were carried forward to Stage 2.

For Stage 2, performed in Australian Terriers, all of the SNPs within ±1 Mb of the 2 SNPs carried forward from Stage 1, were tested for an association with diabetes. In this region, which spanned CFA18: 47,608,071–49,593,379 76 SNPs were present in the Affymetrix assay used in Australian Terriers. However, genetic type I error testing determined that there were only 32 independent tests in this region, establishing the Bonferroni-corrected significance level at 0.0016 (=0.05/32). Five of these 76 SNPs, located at CFA18: 47,696,268, CFA18: 47,814,120, CFA18: 48,009,198, CFA18: 48,027,909, and CFA18: 48,044,155 had P-values of 0.0007, 0.0009, 0.0011, 0.0011, and 0.0007, respectively, all significant at the level of 0.0016 (Table 2). Three of these 5 SNPs (CFA18: 48,009,198, CFA18: 48,027,909, and CFA18: 48,044,155) were in complete LD with each other ($r^2 = 1$ for all 3 associations) and the 2 remaining SNPs (CFA18: 47,696,268 and CFA18: 47,814,120) were in moderate LD with each other ($r^2 = 0.4$). Figure 2b represents a regional association plot of $-\log_{10}(P\text{-value})$ in the Stage 2 Australian Terrier validation cohort.

Of the 2 SNPs carried forward from Stage 1 (Samoyeds) to Stage 2 (Australian Terriers), one (CFA18: 48,578,245) was not present on the Affymetrix assay in Australian Terriers and the other (CFA18: 48,632,489) was present on the Affymetrix assay in Australian Terriers but did not yield a significant association with diabetes. Of the 5 SNPs identified as significantly associated with diabetes in Stage 2 (Australian Terriers), one (CFA18: 47,814,120) was not present on the Illumina assay in Samoyeds, and the other 4 SNPs were present on the Illumina assay in Samoyeds but did not yield significant associations with diabetes. The significance and other characteristics of the 2 SNPs carried forward from Stage 1 and the 5 SNPs identified in Stage 2 are reported in Table 2. The entire dog region studied spans about 5.8 Mb on CFA18: 43,811,865–49,593,379 and includes 17 genes which are either associated with a specific type of diabetes in humans or have a proposed mechanism of action that could be involved in the pathogenesis of diabetes in humans (Table 3) (Sladek et al. 2007; Kong et al. 2009; Bradfield et al. 2011; Ketterer et al. 2011; Xu et al. 2011, 2012, 2014; Glisic and Jailwala 2012; Saxena et al. 2012; Almawi et al. 2013; Dayeh et al. 2013; Frederiksen et al. 2013; Hanson et al. 2013; Liu et al. 2013; Moritani et al. 2013; Ng et al. 2014; Roesch et al. 2015; Christoph et al. 2016; Elboudwarej et al. 2016; Fan et al. 2016; Huopio et al. 2016; Piccini et al. 2016; Sakano et al. 2016; Mishra et al. 2017; Qi et al. 2017).

Discussion

An association between the *INS* gene region of about 5.8 Mb on CFA18 and diabetes was identified in Samoyeds and validated in Australian Terriers. The results of this study suggest that 1 or more

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

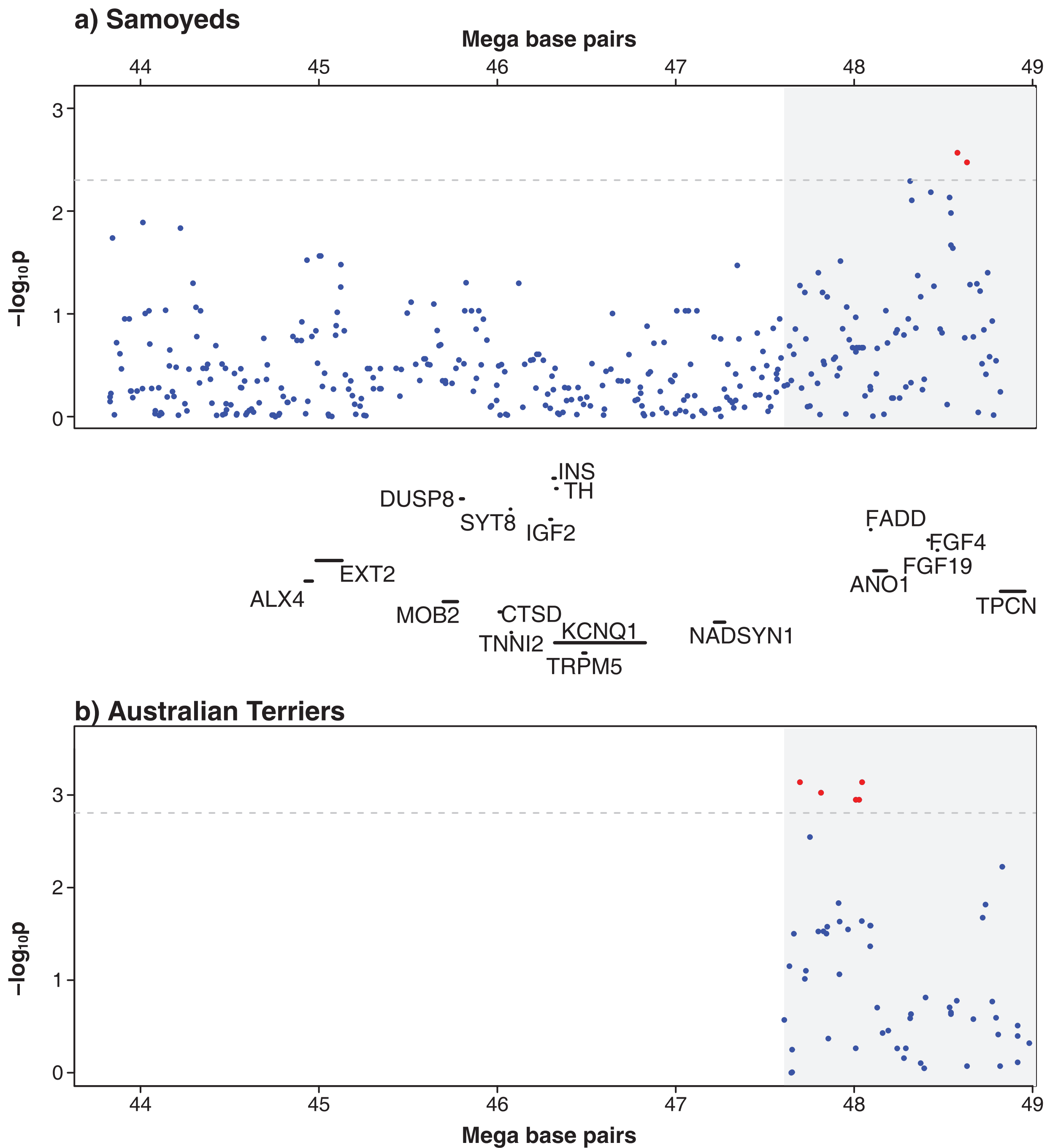

Figure 2. Regional association plots of $-\log_{10}$(*P*-value) on *Canis familiaris* chromosome 18 in the Stage 1 Samoyed discovery cohort (**a**) and Stage 2 Australian Terrier validation cohort (**b**). The horizontal dashed lines are at the value considered significant for each stage, with significant SNPs shown in red. The locations of genes associated with diabetes in humans are shown between the plots. The line above each gene indicates the location of the gene on *C. familiaris* chromosome 18. The figure spans the entire insulin region studied: *C. familiaris* chr18: 43,811,865–49,593,379. The shaded area on the right spans the region studied in Stage 2 (*C. familiaris* chr18: 47,608,071–49,593,379). Please refer to Table 3 for more information about the genes.

genes in this region are associated with diabetes in Samoyeds and Australian Terriers. When interpreting preliminary findings in exploratory dog studies of complex diseases such as diabetes, it is useful to consider candidate genes associated with the disease in humans (Hayward et al. 2016). The neighborhood of the canine *INS* gene region harbors 17 candidate genes related to diabetes in humans (Table 3) (Sladek et al. 2007; Kong et al. 2009; Bradfield et al. 2011; Ketterer et al. 2011; Xu et al. 2011, 2012, 2014; Glisic and Jailwala 2012; Saxena et al. 2012; Almawi et al. 2013; Dayeh et al. 2013; Hanson et al. 2013; Frederiksen et al. 2013; Liu et al. 2013; Moritani et al. 2013; Ng et al. 2014; Roesch et al. 2015; Christoph et al. 2016; Elboudwarej et al. 2016; Fan et al. 2016; Huopio et al. 2016; Piccini et al. 2016; Sakano et al. 2016; Mishra et al. 2017; Qi et al. 2017). Ten of these 17 genes are located within 1.5 Mb of each other on human chromosome 11 (Table 3). However, in humans, 2 of the 17 genes (*ALX4* and *EXT2*) are located about

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

Table 2. Associations between SNPs within the *INS* gene region and diabetes in Samoyeds and Australian Terriers

Breed	Chr	Position[a]	Major/minor allele	MAF in cases[b]	MAF in controls[b]	Odds ratio[c]	*P*-value[d]
Samoyed	18	48578245	G/C	0.55	0.25	3.72	0.0027
Samoyed	18	48632489	A/G	0.57	0.29	3.29	0.0033
Australian Terrier	18	47696268	T/C	0.13	0.44	0.20	0.0007
Australian Terrier	18	47814120	G/C	0.63	0.31	3.67	0.0009
Australian Terrier	18	48009198	T/C	0.13	0.43	0.21	0.0011
Australian Terrier	18	48027909	T/C	0.13	0.43	0.21	0.0011
Australian Terrier	18	48044155	T/C	0.13	0.44	0.20	0.0007

[a]CanFam 3.1, http://genome.ucsc.edu/.
[b]MAF, minor allele frequency.
[c]The genotypic odds ratio for 1 copy increase of minor allele.
[d]Genetic type I error and Bonferroni-corrected *P*-value.

Table 3. Genes implicated in the pathogeneses of human diabetes, in order of their location on *Canis familiaris* chr18: 43,811,865–49,593,379[a]

Gene symbol	Gene name	Gene location on CFA18 (bp)	Type of diabetes or mechanism of action in humans	Gene location on human chr11 (bp)
ALX4	Homeobox protein aristaless-like 4	44,922,150–44,963,383	T2D (Sladek et al. 2007; Almawi et al. 2013)	44,260,444–44,310,166
EXT2	Exostosin glycosyltransferase 2	44,983,793–45,129,935	T2D (Sladek et al. 2007; Liu et al. 2013)	44,095,549–44,244,932
MOB2	Monopolar spindle-one-binder 2	45,696,392–45,776,581	T2D (Kong et al. 2009; Hanson et al. 2013)	1,469,962–1,486,746
DUSP8	Dual-specificity phosphatase 8	45,793,451–45,809,987	T2D (Kong et al. 2009; Dayeh et al. 2013)	1,554,044–1,571,920
CTSD	Cathepsin D	46,009,825–46,019,780	T2D (Christoph et al. 2016)	1,752,752–1,763,992
SYT8	Synaptotagmin 8	46,072,164–46,075,383	Interaction with *INS* gene in human pancreatic islets (Xu et al. 2011, 2012)	1,834,444–1,837,521
TNNI2	Troponin I2, fast skeletal type	46,076,928–46,079,538	Interaction with *INS* gene in human pancreatic islets (Xu et al. 2011)	1,838,989–1,841,678
IGF2	Insulin-like growth factor 2	46,293,985–46,304,805	T1D, T2D (Ng et al. 2014; Mishra et al. 2017)	2,129,112–2,139,389
INS	Insulin	46,311,865–46,324,933	T1D, T1bD, T2D, monogenic, neonatal, LADA, MODY (Bradfield et al. 2011; Saxena et al. 2012; Moritani et al. 2013; Elboudwarej et al. 2016; Huopio et al. 2016; Piccini et al. 2016; Mishra et al. 2017)	2,159,779–2,161,341
KCNQ1	Potassium voltage-gated channel subfamily Q member 1	46,321,292–46,829,537	T2D (Qi et al. 2017)	2,444,991–2,849,109
TH	Tyrosin hydroxylase	46,327,136–46,334,973	Synthesizes dopamine which modulates human beta cell mass (Sakano et al. 2016)	2,166,894–2,168,726
TRPM5	Transient receptor potential cation channel subfamily M member 5	46,478,954–46,496,868	T2D (Ketterer et al. 2011)	2,404,515–2,423,045
NADSYN1	Nicotinamide adenine dinucleotide synthetase 1	47,216,500–47,274,949	T1D (Frederiksen et al. 2013)	71,503,864–71,505,874
FADD	Fas associated via death domain	48,091,329–48,094,388	T1D (Glisic and Jailwala 2012)	70,203,163–70,207,390
ANO1	Anoctamin 1	48,109,817–48,181,867	Regulates insulin secretion in human beta cells (Xu et al. 2014)	70,078,302–70,189,528
FGF19	Fibroblast growth factor 19	48,464,928–48,468,782	T2D (Roesch et al. 2015)	69,698,232–69,704,642
TPCN2	Two pore segment channel 2	48,957,256–48,984,505	T2D (Fan et al. 2016)	69,055,162–69,090,604

CFA, *Canis familiaris* chromosome; T1D, type 1 diabetes; T1bD, type 1b diabetes; T2D, type 2 diabetes; LADA, latent autoimmune diabetes of adults; MODY, maturity-onset diabetes of the young.

[a]The genes are listed in the order found on *C. familiaris* chromosome 18. The locations of the same genes on human chr11 are listed separately and are not in the same order.

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

41 Mb upstream from these 10 genes, and the remaining 5 genes (*NADSYN1*, *FADD*, *ANO1*, *FGF19*, and *TPCN2*) are located about 66 Mb upstream from these 10 genes (Table 3). It is not yet known which of these 17 genes contribute to diabetes in Samoyeds and Australian Terriers. Future studies analyzing the sequence of CFA18: 43,811,865–49,593,379 in Samoyeds and Australian Terriers with and without diabetes will help determine which specific genes within this region are important in the pathophysiology of diabetes in these breeds of dogs. Although the identified SNPs are located in a relatively small area of less than 1 Mb, the entire 5.8 Mb is of interest for future studies because of the extent of LD in pure breed dogs. This study focused on the *INS* gene region because the *INS* gene is associated with several types of diabetes in humans, including type 1 diabetes. While canine diabetes most resembles type 1 diabetes it does not perfectly mimic type 1 diabetes or any other specific type of human diabetes (O'Kell et al. 2017). Therefore, for this investigative study of canine diabetes, examining an area, which is important in the pathophysiology of several forms of human diabetes, was fitting.

The association of the *INS* gene region on CFA18 was replicated in 2 breeds of unrelated clades (Parker et al. 2017). A recent study of 1,346 dogs from 161 breeds determined that most breeds of dogs belong to 1 of 23 distinct clades that were formed before breed clubs and registries (Parker et al. 2017). The medium-sized, white, Samoyed was bred to serve as a working sled dog in the cold climate of Siberia. The first Samoyed was registered with the American Kennel Club in 1906, and was imported from Russia. The Samoyed is considered to be a basal breed, which is a genetically divergent modern breed that has avoided mixture with other modern breeds (Larson et al. 2012). As is the case with other basal breeds, the Samoyed has geographic origins in the Old World (Larson et al. 2012). In contrast to the Samoyed, the Australian Terrier was formed in the late 1800s by mixing a number of European Terriers, was introduced to the United States in the 1940s, and was first registered with the American Kennel Club in 1960. The small-sized, brown, Australian Terrier was bred to serve as a rodent controlling and herding work dog in the warm Australian climate. The fact that the *INS* gene region on CFA18 is associated with diabetes in 2 unrelated breeds of different clades, and that one of these breeds (the Samoyed) is a basal breed, could suggest that the genetic change associated with diabetes in these breeds developed in a common ancestral dog. Alternatively, it is also possible that changes in the *INS* gene region on CFA18 occurred separately in each of the different breeds. Therefore, the association of the *INS* gene region on CFA18 and diabetes could have importance in other breeds related to this common ancestor, or in canine diabetes at large. However, canine diabetes is probably a multifactorial, complex disease that involves several genetic regions, multiple alleles of weak effect, environmental, and other factors, as is the case in humans (Vatanen et al. 2016). Further studies of the *INS* gene region in dogs from a variety of clades are warranted to investigate the role of this region in the pathophysiology of diabetes in different breeds.

Previous studies of the genetics of diabetes in dogs have focused on SNPs within a short distance (about 1.5 Kb) of candidate genes in a variety of breeds. It is possible that the exploratory approach employed in the present study, in which a large genetic region is examined for a replicated association with diabetes in 2 breeds at high risk for the disease, will open up additional paths of research for the genetics of diabetes in dogs. The availability of a new canine genotyping array, which identifies over 670,000 SNPs (Axiom Canine Genotyping Array Set B, https://www.thermofisher.com) will improve the utility of genome-wide association studies for future exploratory investigations in larger cohorts of dogs with diabetes.

One of the previous studies investigating the *INS* gene region in dogs analyzed a number of candidate genes, including the *INS* gene (Short et al. 2007). SNPs were chosen for analysis if they encoded a nonsynonymous amino acid change, were located in exonic regions, or were in the region 1.5 Kb upstream of exon 1. Six *INS* gene SNPs were selected for analysis. This study included 20 Samoyeds with diabetes, 9 control Samoyeds without diabetes, and no Australian Terriers. The Samoyed sample size in this study could have been too small for detection of a SNP significantly associated with diabetes in this breed. The study did identify 2 *INS* SNPs that were protective for diabetes, one in Labrador Retrievers and the other in Cocker Spaniels. However, one of these *INS* SNPs also increased the risk of diabetes in Jack Russell Terriers. The authors acknowledged that it was unexpected to have 1 SNP confer both protection and risk and suggested that false-positive findings could explain this discrepancy. The authors also suggested that a SNP identified as associated with diabetes could be in LD with different alleles that increase the risk or offer protection for diabetes, thus complicating data interpretation. Finally, it was suggested that the etiology of diabetes could differ by breed (Short et al. 2007). A more recent study included a larger number of Samoyeds (40 cases and 74 controls) and no Australian Terriers (Short et al. 2014). This study examined only 1 predetermined synonymous coding *INS* SNP that was not previously studied in dogs, and identified an association between this *INS* SNP and diabetes in Cocker Spaniels but not in Samoyeds or other breeds (Short et al. 2014). It is possible that these studies were not successful in identifying an association in Samoyeds or replicating an association in more than 1 breed because they examined a small number of SNPs in the *INS* gene region (6 SNPs in one study and 1 SNP in the other) and studied a relatively small area in comparison to hundreds of SNPs examined over a 5.8 Mb area in the current study. The approach of the study reported here was different in that all SNPs in a large 5.8 Mb *INS* gene region were examined for a possible association with diabetes. This approach was undertaken because LD spans several Mb in pure breed dogs. The different study design undertaken in the current study, compared to previous ones, could have contributed to the ability to replicate an association of the *INS* gene region with diabetes in 2 different breeds.

The SNPs identified as associated with diabetes in this study are located in an approximately 1 Mb area between CFA18: 47,696,268–48,632,489 which includes the fas associated via death domain, anoctamin 1, and fibroblast growth factor 19 genes. The fas associated via death domain gene expresses a protein associated with apoptosis and is overexpressed in humans with a high genetic risk for type 1 diabetes (Glisic and Jailwala 2012). Anoctamin 1 is a gene which codes for a calcium-activated chloride channel protein. In human islet cell cultures, inhibition of this gene expression decreases insulin secretion (Xu et al. 2014). Fibroblast growth factor 19 is secreted from the small intestine in response to eating, and increases insulin sensitivity. Fibroblast growth factor 19 concentrations are decreased in humans with type 2 diabetes (Roesch et al. 2015). The role of these genes in canine diabetes has yet to be investigated.

One of this study's limitations is the small sample size. While the whole genome was genotyped, the study was not powered to detect whole genome associations, and therefore focused on the *INS* gene region only. The 2-stage study design was chosen to overcome the constraints of a small sample size. With the 2-stage approach all SNPs in the *INS* gene region were evaluated on a subset of cases and controls

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

(Samoyeds) in Stage 1, and only SNPs of interest were further assessed in Stage 2, in Australian Terrier cases and controls (Satagopan et al. 2004). This 2-stage study design has successfully identified genes associated with numerous conditions including acute lung injury risk following major trauma, maternal hyperglycemia and leptin levels in newborns, human brain asymmetry, and human longevity (Christie et al. 2012; Di Cianni et al. 2013; Cote et al. 2016; Tadayon et al. 2016). Another study limitation is that this study was not designed to investigate the mode of inheritance of diabetes in dogs.

It is estimated that about 100 cases and 100 breed-matched controls are needed for genome-wide association studies of complex diseases in pure breed dogs (Hayward et al. 2016), especially when the particular breed is at substantially higher risk of presenting with the disease. Future genetic studies of diabetes will therefore involve larger study populations. Another study limitation is that different SNP arrays were used for each breed, and some SNPs were identified by only 1 array. Furthermore, SNPs identified as significant by 1 array were not necessarily significant on the other array. However, in this study significant SNPs serve as a marker for an association of diabetes with the *INS* gene region rather than identifiers of a specific allelic variation unique to diabetes. This approach in which a SNP serves as a marker for disease association with a several Mb region, is an acceptable exploratory approach in dogs (Hayward et al. 2016). Future studies sequencing and fine mapping the *INS* gene region in Samoyeds and Australian Terriers with and without diabetes will help determine which specific allelic variations and haplotypes within this region are important in the pathophysiology of diabetes in these breeds. Finally, control dogs without diabetes could have carried the genetic makeup for diabetes without exhibiting clinical signs of the disease at the time of enrollment and blood draw. Control dogs were enrolled only if they were 9 years of age or older to minimize enrollment of dogs who might develop diabetes in the future, but some dogs develop diabetes later in life.

The pathophysiology and role of autoimmunity in canine diabetes is incompletely understood (Ahlgren et al. 2014; Holder et al. 2015). Presence of autoantibodies has not been reported in dogs at increased risk for diabetes, prior to disease onset because it is difficult to detect dogs with diabetes before overt clinical signs are present. Early genetic identification of dogs at increased risk for diabetes will enable such studies in the future. Identification of dogs at increased genetic risk for diabetes will also facilitate studies of preventive measures such as vaccination or microbiome manipulation (Insel and Dunne 2016; Vatanen et al. 2016). An improved understanding of the genetic risk of diabetes in pure breed dogs can also guide future breeding practices to decrease the incidence of diabetes in these breeds of dogs.

In conclusion, this preliminary study has identified a replicated association of the canine *INS* gene region, in 2 unrelated breeds of different clades. One of these breeds (the Samoyed) is a basal breed, which is relatively genetically isolated from other modern breeds. Therefore, it is possible that the *INS* gene region association with diabetes developed in a common ancestral dog and that this association affects other breeds of dogs. Future studies of this specific region on CFA18 will determine which of the genes in this area are involved in the pathogenesis of diabetes in Samoyeds and Australian Terriers.

Funding

This work was supported by the American Kennel Club Canine Health Foundation (http://www.akcchf.org, Grant #610); and by a gift from Ms Catherine Adler.

Acknowledgments

Presented as a talk with no proceedings at the Meeting of the Society for Comparative Endocrinology, Alexandria, VA, 5 June 2017. The study protocol and owner consent form were approved by the University of Pennsylvania Privately Owned Animal Protocol Committee.

Data Availability

Whole-genome tag SNP genotyping data are available at https://doi.org/10.5061/dryad.pnvx0k6hm.

References

Ahlgren K, Fall T, Landegren N, Grimelius L, von Euler H, Sundberg K, Lindblad-Toh K, Lobell A, Hedhammar Å, Andersson G, et al. 2014. Lack of evidence for a role of islet autoimmunity in the aetiology of canine diabetes mellitus. *PLoS One*. 9:e105473.

Almawi W, Nemr R, Keleshian S, Echtay A, Saldanha F, AlDoseri F, Racoubian E. 2013. A replication study of 19 GWAS-validated type 2 diabetes at-risk variants in the Lebanese population. *Diabetes Res Clin Pract*. 102:117–122.

Awano T, Johnson GS, Wade CM, Katz ML, Johnson GC, Taylor JF, Perloski M, Biagi T, Baranowska I, Long S, et al. 2009. Genome-wide association analysis reveals a SOD1 mutation in canine degenerative myelopathy that resembles amyotrophic lateral sclerosis. *Proc Natl Acad Sci USA*. 106:2794–2799.

Bradfield J, Qu H-Q, Wang K, Zhang H, Sleiman P, Kim C, Mentch FD, Qiu H, Glessner JT, Thomas KA, et al. 2011. A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet*. 7:e1002293.

Catchpole B, Adams J, Holder A, Short A, Ollier W, Kennedy L. 2013. Genetics of canine diabetes mellitus: are the diabetes susceptibility genes identified in humans involved in breed susceptibility to diabetes mellitus in dogs? *Vet J*. 195:139–147.

Christie JD, Wurfel MM, Feng R, O'Keefe GE, Bradfield J, Ware LB, Christiani DC, Calfee CS, Cohen MJ, Matthay M, *et al*.; Trauma ALI SNP Consortium (TASC) investigators. 2012. Genome wide association identifies PPFIA1 as a candidate gene for acute lung injury risk following major trauma. *PLoS One*. 7:e28268.

Christoph N, Sundström J, Gustafsson S, Giedraitis V, Lind L, Ingelsson E, Fall T. 2016. Protein biomarkers for insulin resistance and type 2 diabetes risk in two large community cohorts *Diabetes*. 65:276–284.

Cote S, Gagne-Ouellet V, Guay SP, Allard C, Houde AA, Perron P, Baillargeon JP, Gaudet D, Guérin R, Brisson D, et al. 2016. PPARGC1α gene DNA methylation variations in human placenta mediate the link between maternal hyperglycemia and leptin levels in newborns. *Clin Epigenetics*. 8:72.

Dayeh T, Olsson A, Volkov P, Almgren P, Rönn T, Ling C. 2013. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia*. 56:1036–1046.

Di Cianni F, Campa D, Tallaro F, Rizzato C, De Rango F, Barale R, Passarino G, Canzian F, Gemignani F, Montesanto A, et al. 2013. MAP3K7 and GSTZ1 are associated with human longevity: a two-stage case–control study using a multilocus genotyping. *Age*. 35:1357–1366.

Elboudwarej E, Cole M, Briggs FB, Fouts A, Fain PR, Quach H, Quach D, Sinclair E, Criswell LA, Lane JA, *et al*. 2016. Hypomethylation within gene promoter regions and type 1 diabetes in discordant monozygotic twins. *J Autoimmun*. 68:23–29.

Fall T, Hamlin HH, Hedhammar A, Kämpe O, Egenvall A. 2007. Diabetes mellitus in a population of 180,000 insured dogs: incidence, survival, and breed distribution. *J Vet Intern Med*. 21:1209–1216.

Fan Y, Li X, Zhang Y, Fan X, Zhang N, Zheng H, Song Y, Shen C, Shen J, Ren F, et al. 2016. Genetic variants of TPCN2 associated with type 2 diabetes risk in the Chinese population. *PLoS One*. 11:e0149614.

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020

Frederiksen BN, Kroehl M, Fingerlin TE, Wong R, Steck AK, Rewers M, Norris JM. 2013. Association between vitamin D metabolism gene polymorphisms and risk of islet autoimmunity and progression to type 1 diabetes: the diabetes autoimmunity study in the young (DAISY). *J Clin Endocrinol Metab*. 98:E1845–E1851.

Glisic S, Jailwala P. 2012. Interaction between Treg apoptosis pathways, Treg function and HLA risk evolves during type 1 diabetes pathogenesis. *PLoS One*. 7:e36040.

Guptill L, Glickman L, Glickman N. 2003. Time trends and risk factors for diabetes mellitus I dogs: analysis of veterinary medical data base records (1970–1999). *Vet J*. 165:240–247.

Hanson R, Guo T, Muller Y, Fleming J, Knowler W, Kobes S, Bogardus C, Baier LJ. 2013. Strong parent-of origin effects in the association of KCNQ1 variants with type 2 diabetes in American Indians. *Diabetes*. 62:2984–2991.

Hayward J, Castelhano M, Oliveira K, Corey E, Balkman C, Baxter T, Casal ML, Center SA, Fang M, Garrison SJ, et al. 2016. Complex disease and phenotype mapping in the domestic dog. *Nat Commun*. 7:10460.

Hess RS, Kass PH, Ward CR. 2000. Breed distribution of dogs with diabetes mellitus admitted to a tertiary care facility. *J Am Vet Med Assoc*. 216:1414–1417.

Hoeppner M, Lundquist A, Pirun M, Meadows J, Zamani N, Johnson J, Sundström G, Cook A, FitzGerald MG, Swofford R, et al. 2014. An improved canine genome and a comprehensive catalogue of coding genes and non-coding transcripts. *PLoS One*. 9:e91172.

Holder A, Kennedy L, Ollier W, Catchpole B. 2015. Breed differences in development of anti-insulin antibodies in diabetic dogs and investigation of the role of dog leukocyte antigen (DLA) genes. *Vet Immunol Immunopathol*. 167:130–138.

Huopio H, Miettinen P, Ilonen J, Nykänen P, Veijola R, Keskinen P, Näntö-Salonen K, Vangipurapu J, Raivo J, Stančáková A, et al. 2016. Clinical, genetic, and biochemical characteristics of early-onset diabetes in the Finnish population. *J Clin Endocrinol Metab*. 101:3018–3026.

Insel R, Dunne JL. 2016. JDRF's vision and strategy for prevention of type 1 diabetes. *Pediatr Diabetes*. 17(Suppl 22):87–92.

Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet*. 42:348–354.

Ketterer C, Müssig K, Heni M, Dudziak K, Randrianarisoa E, Wagner R, Machicao F, Stefan N, Holst JJ, Fritsche A, et al. 2011. Genetic variation within the TRPM5 locus associates with prediabetic phenotypes in subjects at increased risk for type 2 diabetes. *Metabolism*. 60:1325–1333.

Kong A, Steinthorsdottir V, Masson G, Thorleifsson G, Sulem P, Besenbacher S, Jonasdottir A, Sigurdsson A, Kristinsson KT, Jonasdottir A, et al. 2009. Parental origin of sequence variants associated with complex diseases. *Nature*. 462:868–874.

Larson G, Karlsson E, Perri A, Webster M, Ho S, Peters J, Stahl PW, Piper PJ, Lingaas F, Fredholm M, et al. 2012. Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA*. 109:8878–8883.

Li MX, Yeung JM, Cherny SS, Sham PC. 2012. Evaluating the effective numbers of independent tests and significant p-value thresholds in commercial genotyping arrays and public imputation reference datasets. *Hum Genet*. 131:747–756.

Lindblad-Toh K, Wade C, Mikkelsen T, Karlsson E, Jaffe D, Kamal M, Clamp M, Chang JL, Kulbokas III EJ, Zody MC, et al. 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*. 438:803–819.

Liu L, Yang X, Wang H, Cui G, Xu Y, Wang P, Yuan G, Wang X, Ding H, Wang DW. 2013. Association between variants of EXT2 and type 2 diabetes: a replication and meta-analysis. *Hum Genet*. 132:139–145.

Manichaikul A, Mychaleckyj J, Rich S, Daly K, Sale M, Chen W. 2010. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 26:2867–2873.

Mishra R, Chesi A, Cousminer D, Hawa M, Bradfield J, Hodge KM, Guy VC, Hakonarson H; Bone Mineral Density in Childhood Study, Mauricio D, Schloot NC, et al. 2017. Relative contribution of type 1 and type 2 diabetes loci to the genetic etiology of adult-onset, non-insulin-requiring autoimmune diabetes. *BMC Med*. 15:88.

Moritani M, Yokota I, Tsubouchi K, Takaya R, Takemoto K, Minamitani K, Urakami T, Kawamura T, Kikuchi N, Itakura M, *et al*.; Japanese Study Group of Insulin Therapy for Childhood and Adolescent Diabetes (JSGIT). 2013. Identification of INS and KCNJ11 gene mutations in type 1B diabetes in Japanese children with onset of diabetes before 5 years of age. *Pediatr Diabetes*. 14:112–120.

Ng M, Shriner D, Chen B, Li J, Chen W-M, Guo X, Liu J, Bielinski SJ, Yanek LR, Nalls MA, et al. 2014. Meta-analysis of genome-wide association studies in African Americans provides insights into the genetic architecture of type 2 diabetes. *PLoS Genet*. 10:e1004517.

O'Kell A, Wasserfall C, Catchpole B, Davison L, Hess R, Kushner J, Atkinson MA. 2017. Comparative pathogenesis of autoimmune diabetes in humans, NOD mice, and canines – has a valuable animal model of type 1 diabetes been overlooked? *Diabetes*. 66:1443–1452.

Parker HG, Dreger DL, Rimbault M, Davis BW, Mullen AB, Carpintero-Ramirez G, Ostrander EA. 2017. Genomic analyses reveal the influence of geographic origin, migration, and hybridization on modern dog breed development. *Cell Rep*. 19:697–708.

Piccini B, Artuso R, Lenzi L, Guasti M, Braccesi G, Barni F, Casalini E, Giglio S, Toni S. 2016. Clinical and molecular characterization of a novel INS mutation identified in patients with MODY phenotype. *Eur J Med Genet*. 59:590–595.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 81:559–575.

Qi Q, Stilp A, Sofer T, Moon J, Hidalgo B, Szpiro A, Wang T, Ng MCY, Guo X; MEta-analysis of type 2 DIabetes in African Americans (MEDIA) Consortium, Chen YI, et al. 2017. Genetics of type 2 diabetes in U.S. Hispanic/Latino individuals: results from the Hispanic community health study/study of Latinos (HCHS/SOL). *Diabetes*. 66:1419–1425.

Roesch S, Styer A, Wood G, Kosak Z, Seiler J, Benotti P, Petrick AT, Gabrielsen J, Strodel WE, Gerhard GS, et al. 2015. Perturbations of fibroblast growth factors 19 and 21 in type 2 diabetes. *PLoS One*.10:e0116928.

Sakano D, Choi S, Kataoka M, Shiraki N, Uesugi M, Kume K, Kume S. 2016. Dopamine D2 receptor-mediated regulation of pancreatic β cell mass. *Stem Cell Rep*. 7:95–109.

Satagopan JM, Venkatraman ES, Begg CB. 2004. Two-stage designs for gene-disease association studies with sample size constraints. *Biometrics*. 60:589–597.

Saxena R, Elbers CC, Guo Y, Peter I, Gaunt TR, Mega JL, Lanktree MB, Tare A, Castillo BA, Li YR, *et al*.; Look AHEAD Research Group; DIAGRAM Consortium. 2012. Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci. *Am J Hum Genet*. 90:410–425.

Short AD, Catchpole B, Kennedy LJ, Barnes A, Fretwell N, Jones C, Thomson W, Ollier WE. 2007. Analysis of candidate susceptibility genes in canine diabetes. *J Hered*. 98:518–525.

Short AD, Holder A, Rothwell S, Massey J, Scholey R, Kennedy LJ, Catchpole B, Ollier WE. 2014. Searching for "monogenic diabetes" in dogs using a candidate gene approach. *Canine Genet Epidemiol*. 1:8.

Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, *et al*. 2007. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature*. 445:881–885.

Tadayon SH, Vaziri-Pashkam M, Kahali P, Ansari Dezfouli M, Abbassian A. 2016. Common genetic variant in VIT is associated with human brain asymmetry. *Front Hum Neurosci*. 24;10:236.

Vatanen T, Kostic A, d'Hennezel E, Siljander H, Franzosa E, Yassour M, Kolde R, Vlamakis H, Arthur TD, Hämäläinen AM, et al. 2016. Variation in microbiome LPS immunogenicity contributes to autoimmunity in humans. *Cell*. 165:1551.

Xu Z, Lefevre GM, Felsenfeld G. 2012. Chromatin structure, epigenetic mechanisms and long-range interactions in the human insulin locus. *Diabetes Obes Metab*. 14(Suppl 3):1–11.

Xu Z, Lefevre G, Gavrilova O, Foster St Claire M, Riddick G, Felsenfeld G. 2014. Mapping of long-range INS promoter interactions reveals a role for calcium activated chloride channel ANO1 in insulin secretion. *Proc Natl Acad Sci USA*. 111:16760–16765.

Xu Z, Wei G, Chepelev I, Zhao K, Felsenfeld G. 2011. Mapping of INS promoter interactions reveals its role in long-range regulation of SYT8 transcription. *Nat Struct Mol Biol*. 18:372–378.

Yang Y, Chan L. 2016. Monogenic diabetes: what it teaches us on the common forms of type 1 and type 2 diabetes. *Endocr Rev*. 37:190–222.

Downloaded from https://academic.oup.com/jhered/article-abstract/110/7/793/5582035 by guest on 24 April 2020