

The *CTRB1/2* Locus Affects Diabetes Susceptibility and Treatment via the Incretin Pathway

Leen M. 't Hart,^{1,2} Andreas Fritsche,³ Giel Nijpels,^{4,5} Nienke van Leeuwen,¹ Louise A. Donnelly,⁶ Jacqueline M. Dekker,^{4,7} Marjan Alsema,⁴ Joao Fadista,⁸ Françoise Carlotti,⁹ Anette P. Gjesing,¹⁰ Colin N.A. Palmer,⁶ Timon W. van Haefen,¹¹ Silke A. Herzberg-Schäfer,³ Annemarie M.C. Simonis-Bik,¹² Jeanine J. Houwing-Duistermaat,¹³ Quinta Helmer,¹³ Joris Deelen,² Bruno Guigas,^{1,14} Torben Hansen,^{10,15} Fausto Machicao,³ Gonneke Willemsen,¹⁶ Robert J. Heine,^{4,17} Mark H.H. Kramer,¹⁸ Jens J. Holst,¹⁹ Eelco J.P. de Koning,^{9,20,21} Hans-Ulrich Häring,³ Oluf Pedersen,^{10,22} Leif Groop,⁸ Eco J.C. de Geus,^{4,16} P. Eline Slagboom,^{2,23} Dorret I. Boomsma,¹⁶ Elisabeth M.W. Eekhoff,¹² Ewan R. Pearson,⁶ and Michaela Diamant¹²

The incretin hormone glucagon-like peptide 1 (GLP-1) promotes glucose homeostasis and enhances β -cell function. GLP-1 receptor agonists (GLP-1 RAs) and dipeptidyl peptidase-4 (DPP-4) inhibitors, which inhibit the physiological inactivation of endogenous GLP-1, are used for the treatment of type 2 diabetes. Using the Metabochip, we identified three novel genetic loci with large effects (30–40%) on GLP-1-stimulated insulin secretion during hyperglycemic clamps in

nondiabetic Caucasian individuals (*TMEM114*; *CHST3* and *CTRB1/2*; $n = 232$; all $P \leq 8.8 \times 10^{-7}$). rs7202877 near *CTRB1/2*, a known diabetes risk locus, also associated with an absolute $0.51 \pm 0.16\%$ (5.6 ± 1.7 mmol/mol) lower A1C response to DPP-4 inhibitor treatment in G-allele carriers, but there was no effect on GLP-1 RA treatment in type 2 diabetic patients ($n = 527$). Furthermore, in pancreatic tissue, we show that rs7202877 acts as expression quantitative trait locus for *CTRB1* and *CTRB2*, encoding chymotrypsinogen, and increases fecal chymotrypsin activity in healthy carriers. Chymotrypsin is one of the most abundant digestive enzymes in the gut where it cleaves food proteins into smaller peptide fragments. Our data identify chymotrypsin in the regulation of the incretin pathway, development of diabetes, and response to DPP-4 inhibitor treatment. *Diabetes* 62:3275–3281, 2013

From the ¹Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, the Netherlands; the ²Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, the Netherlands; the ³Division of Endocrinology, Diabetology, Angiology, Nephrology, and Clinical Chemistry, Department of Internal Medicine, Eberhard Karls University Tübingen, Member of the German Centre for Diabetes Research (DZD), Tübingen, Germany; ⁴The EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, the Netherlands; the ⁵Department of General Practice, VU University Medical Center, Amsterdam, the Netherlands; the ⁶Division of Cardiovascular and Diabetes Medicine, Medical Research Institute, University of Dundee, Dundee, U.K.; the ⁷Department of Epidemiology and Statistics, VU University Medical Center, Amsterdam, the Netherlands; the ⁸Lund University Diabetes Center, Department of Clinical Sciences, Diabetes, and Endocrinology, Skåne University Hospital Malmö, Lund University, Malmö, Sweden; the ⁹Department of Nephrology, Leiden University Medical Center, Leiden, the Netherlands; ¹⁰The Novo Nordisk Foundation Center for Basic Metabolic Research, Section of Metabolic Genetics, University of Copenhagen, Copenhagen, Denmark; the ¹¹Department of Internal Medicine, Utrecht University Medical Center, Utrecht, the Netherlands; the ¹²Diabetes Center, Department of Internal Medicine, VU University Medical Center, Amsterdam, the Netherlands; the ¹³Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands; the ¹⁴Department of Parasitology, Leiden University Medical Center, Leiden, the Netherlands; the ¹⁵Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark; the ¹⁶Department of Biological Psychology, VU University, Amsterdam, the Netherlands; ¹⁷Eli Lilly and Company, Indianapolis, Indiana; the ¹⁸Department of Internal Medicine, VU University Medical Center, Amsterdam, the Netherlands; ¹⁹The Novo Nordisk Foundation Center for Basic Metabolic Research, Department of Biomedical Sciences, Panum Institute, University of Copenhagen, Copenhagen, Denmark; the ²⁰Department of Endocrinology, Leiden University Medical Center, Leiden, the Netherlands; the ²¹Hubrecht Institute, Utrecht, the Netherlands; the ²²Faculty of Health Sciences, University of Aarhus, Aarhus, Denmark; and the ²³Netherlands Consortium for Healthy Aging, Leiden, the Netherlands.

Corresponding author: Leen M. 't Hart, l.m.t_hart@lumc.nl.

Received 8 February 2013 and accepted 30 April 2013.

DOI: 10.2337/db13-0227

This article contains Supplementary Data online at <http://diabetes.diabetesjournals.org/lookup/suppl/doi:10.2337/db13-0227/-/DC1>.

A.F., G.N., N.v.L., D.I.B., E.M.W.E., E.R.P., and M.D. contributed equally to this study.

M.A. is currently affiliated with Unilever Research and Development, Vlaardingen, the Netherlands; and R.J.H. is currently affiliated with Eli Lilly and Company, Indianapolis, Indiana.

© 2013 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. See <http://creativecommons.org/licenses/by-nc-nd/3.0/> for details.

See accompanying commentary, p. 3019.

The incretin hormone glucagon-like peptide 1 (GLP-1) is released after a meal by L cells in the distal parts of the gastrointestinal tract, and it potentiates glucose-dependent insulin secretion by the pancreas, which is known as the incretin effect. Furthermore, it reduces glucagon secretion and exerts beneficial effects on gut motility, satiety, and food intake (reviewed in 1,2). Based on these functions and the impaired incretin effect in type 2 diabetic patients, two novel incretin-based classes of glucose-lowering drugs have been developed for the treatment of type 2 diabetes (1). These include injectable GLP-1 receptor agonists (GLP-1 RAs), which provide pharmacological levels of GLP-1, and dipeptidyl peptidase-4 (DPP-4) inhibitors, which inhibit the physiological inactivation of endogenous GLP-1. The enzyme DPP-4 rapidly inactivates endogenous GLP-1 by cleavage of the two NH_2 -terminal amino acids (1–4). In addition to their direct insulinotropic effects on the pancreas, injectable GLP-1 RAs, like endogenous GLP-1, lower blood glucose by inhibiting glucagon secretion and slow down gastric emptying in a glucose-dependent manner. Also, GLP-1 RAs promote satiety and weight loss. All these actions contribute to their therapeutic efficacy (1–4). Similarly, oral DPP-4 inhibitors lower blood glucose by stimulating insulin and inhibiting glucagon secretion but are considered weight neutral (1,4). Importantly, the treatment response to compounds of either drug class varies widely, and it is conceivable that both environmental and genetic factors may underlie these differences.

Previously, we showed that during hyperglycemic clamps, the insulin response to intravenous GLP-1 stimulation was

affected by genetic factors (5) with an estimated heritability of 0.53 (0.33–0.70) (6). These findings indicate that genetic factors exert substantial effects on GLP-1–induced insulin response and, as a consequence, may affect an individual's response to the GLP-1–based therapies.

In the current study, the MetaboChip (200k single nucleotide polymorphisms [SNPs]) was used to identify genetic variants affecting GLP-1–induced insulin secretion during hyperglycemic clamps in 232 nondiabetic participants from two independent populations, the Netherlands Twin Register (NTR) and the German Tübingen cohort (7,8). Subsequently, the influence of associated SNPs was tested on response to GLP-1 RA and DPP-4 inhibitor treatment in type 2 diabetic patients ($n = 527$) from the Netherlands (Diabetes Care System [DCS] West-Friesland) (9) and the U.K. (GoDARTS) (10). Finally, gene expression in the pancreas and pancreatic islets and functional tests in healthy volunteers were performed to further elucidate potential underlying mechanisms.

Our studies identify rs7202877 near *CTRB1* and *CTRB2*, which both encode the digestive enzyme chymotrypsin, in the regulation of the incretin pathway, development of diabetes, and response to DPP-4 inhibitor treatment.

RESEARCH DESIGN AND METHODS

Study subjects and experimental procedures

Hyperglycemic clamp cohort. We included 232 Caucasian subjects from two independent studies in the Netherlands and Germany. There were 126 subjects from the NTR, including 69 monozygotic (MZ) and 29 dizygotic twins as well as 28 of their nontwin siblings recruited from 54 families (120 normal glucose tolerant [NGT] and 6 impaired glucose tolerant [IGT]) (7). The German cohort consisted of 73 participants with NGT and 33 participants with IGT who were all unrelated (8). Clinical characteristics of both cohorts are given in Supplementary Table 1. All studies were approved by the local medical ethics review board, all subjects gave written informed consent, and the studies were performed according to the principles of the declaration of Helsinki.

All subjects underwent a modified hyperglycemic clamp with glucose and GLP-1 as previously described (11). In brief, subjects received a priming infusion of glucose to acutely raise blood glucose levels to 10 mmol/L. Variable glucose infusions were used to keep glucose levels constant throughout the whole procedure. After the 2nd hour, insulin secretion was further stimulated using intravenous GLP-1 infusion (4.5 pmol/kg bolus for 1 min at $t = 120$ followed by a continuous infusion of $1.5 \text{ pmol} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$ for 1 h). In the Dutch NTR twin cohort, slightly lower GLP-1 concentrations were used (1.5 pmol/kg and $0.5 \text{ pmol} \cdot \text{kg}^{-1} \cdot \text{min}^{-1}$, respectively). The near maximal insulin response was tested by injecting a bolus of 5 g arginine hydrochloride at $t = 180$ min. First and second phase glucose-stimulated insulin secretion (GSIS), insulin sensitivity index (ISI), and disposition index (DI) were calculated as described previously (7). GLP-1–stimulated insulin secretion was measured as the incremental area under the curve during the last 20 min of the glucose + GLP-1 stimulation ($t = 160$ – 180). In order to increase the power and robustness of the analysis of GSIS in carriers of SNPs associated with GLP-1–stimulated insulin secretion, we included another 216 subjects from two Dutch cohorts (Hoorn and Utrecht study cohorts) who underwent an identical hyperglycemic clamp but without the additional GLP-1 stimulation. Details of these cohorts can be found in Hart et al. (7) and Supplementary Table 1.

Pharmacogenetic study cohort. Type 2 diabetic patients participating in this study were from the Dutch DCS West-Friesland ($n = 7,515$) (9) and the U.K. Genetics of Diabetes Audit and Research Tayside Scotland cohort (GoDARTS, $n = 8,000+$) (10). Patients were included in the current study if they were Caucasian and treated with either a GLP-1 RA or a DPP-4 inhibitor for at least 3 months ($n = 527$). In DCS, 22 patients were treated with a GLP-1 RA and another 49 patients with a DPP-4 inhibitor. In GoDARTS, 151 patients were treated with a GLP-1 RA and another 305 patients were treated with a DPP-4 inhibitor. Patient characteristics of both study cohorts are depicted in Supplementary Table 2. Baseline A1C was defined as the measurement closest to the initiation of treatment (maximum 12 [DCS] or 6 [GoDARTS] months before). On treatment, A1C was defined as the minimum A1C measure between 3 and 24 [DCS] or 18 [GoDARTS] months after therapy initiation.

MetaboChip genotyping and quality control. All subjects participating in the hyperglycemic clamp study were genotyped using the MetaboChip (www.illumina.com). The MetaboChip was designed for replication ($n = \sim 66,000$

SNPs) and fine mapping of regions on the genome ($n = \sim 120,000$ SNPs) that have previously been associated with cardiovascular and/or metabolic diseases, including type 2 diabetes (12). Our analyses focused on the 66,000 SNPs for replication/follow-up. GenomeStudio software (Illumina, Inc., Eindhoven, the Netherlands) was used to visualize the data and SNP calling. For quality control (QC), we used a cutoff for the genotyping call rate of 95%, a minor allele frequency (MAF) in the combined cohort of 0.05, and the P value cutoff for Hardy-Weinberg equilibrium was set at 10^{-4} . In total, 13.0% (8,465) of the SNPs on the MetaboChip were monomorphic or had a MAF < 0.05 , and another 5.4% (3,526) of the SNPs failed the other QC criteria. In total, 53,354 SNPs passed QC, which is comparable to previous reports (12). We computed the genomic inflation factor (λ) as previously described (12,13). Given the low λ of 1.014 (Supplementary Fig. 1), the test statistics were not further adjusted.

The top loci ($P \leq 1 \times 10^{-4}$) were visually inspected for SNPs with a low number of homozygous mutant alleles ($n < 5$), and these SNPs were subsequently reanalyzed using a carrier model (11 vs. 12 + 22). The Sequenom MassARRAY iPLEX Gold platform (Sequenom, Inc., San Diego, CA) was used for validation genotyping of our top SNPs and one proxy for each of them ($r^2 > 0.9$) and for two samples that failed QC on the MetaboChip (concordance $> 99.9\%$). Taqman SNP genotyping assays (Life Technologies, Bleiswijk, the Netherlands) were used to genotype subjects participating in the other studies except for the GoDARTS pharmacogenetic study where data were extracted from a genome-wide association study and MetaboChip dataset (14).

Gene expression studies. Human cadaveric donor pancreata ($n = 35$) were procured through a multiorgan donor program. Tissue samples from the pancreata were used in the study if research consent was present and pancreata could not be used in clinical β -cell therapy programs according to national laws. Tissue samples were snap frozen in liquid nitrogen and stored at -80°C . Tissues were first incubated overnight in RNAlater-ICE (Life Technologies), and subsequently RNA was isolated using Trizol reagent (Life Technologies). Gene expression was quantified using Taqman gene expression assays (Life Technologies). The β -2-microglobulin (*B2M*, Hs00984230_m1) gene was used as reference, but similar data were obtained when using other housekeeping genes such as *GAPDH* or *ACTB*. *CTRB* gene expression was tested using a probe detecting both copies of the gene (Hs00157181_m1), and *BCAR1* was detected with probe Hs01547079_m1. Expression levels were given as arbitrary units relative to the expression of *B2M*.

Islets from cadaver donors ($n = 45$) were provided by the Nordic Islet Transplantation Program (<http://www.nordicislets.org>). Purity of islets was assessed as described previously (15). RNA-seq libraries were prepared using the standard Illumina mRNA-Seq protocol, and sequencing was performed in an Illumina HiSeq 2000 machine. The FASTQ output files generated were then aligned to the human reference genome (UCSC hg19) with TopHat (16). The annotated GTF transcript and fasta genome files were from University of California, Santa Cruz (Santa Cruz, California) and were downloaded from <http://cufflinks.cbc.umd.edu/igenomes.html>. Gene expression was then considered as the normalized sum of expression of all its exons. The htseq count script (<http://www-huber.embl.de/users/anders/HTSeq/doc/count.html>) was used by counting uniquely mapped reads in each exon using the mode “intersection-nonempty.” Gene expression normalization was then performed using the trimmed mean of M values (TMM) method (17), and further normalization was applied by adjusting the expression to gene length.

Chymotrypsin activity study. In order to quantify chymotrypsin activity in carriers and noncarriers of rs7202877 near *CTRB1/2* we used stool samples from 40 carriers of the G allele and 40 age-, sex-, and BMI-matched noncarriers, which were all unrelated. They were nondiabetic German subjects at increased risk for type 2 diabetes (positive family history and overweight) participating in an ongoing intervention study using different types of lifestyle interventions. Clinical characteristics of this study group are given in Supplementary Table 3. A fresh stool sample was collected prior to participation in the intervention study and immediately frozen at -80°C . Fecal chymotrypsin activity was measured using a standard procedure first described by Kaspar et al. (18). Chymotrypsin activity is expressed as units per gram stool at 25°C . Sensitivity of chymotrypsin activity for proteolysis by human DPP-4 has been tested after a 30-min preincubation of the stool suspension (10 μL) with or without 0.1 units of DPP-4 (BPS Biosciences, San Diego, CA) in a total volume of 40 μL of 50 mmol/L Tris buffer (pH 7.5) at 37°C .

Statistics. If necessary, outcome variables were log transformed prior to analysis to obtain normality. To take into account the family relatedness in the twin sample, we used linear generalized estimating equations (GEEs) assuming additive models in Bioconductor R (GEEpack, www.bioconductor.org). The analyses of glucose and GLP-1–stimulated insulin secretion were adjusted for age, sex, BMI, study center, glucose tolerance status (NGT/IGT), and ISI. For the analysis of ISI and DI, ISI was removed from the covariates. Additional adjustment for the prevailing insulin levels did not alter our results (data not shown). For the MZ twin pairs, we used the mean of the individual β -cell responses and covariates for both twins in a pair. This approach yields similar

TABLE 1
Significantly associated SNPs for GLP-1-stimulated insulin secretion during hyperglycemic clamps

SNP	Chr	Nearest gene	Alleles (MAF) ^a	Estimated mean insulin secretion (pmol/L)			<i>P</i> _{add}	<i>P</i> _{dom}	<i>P</i> _{rec}
				11	12	22			
rs4148941	10	<i>CHST3</i>	A/C (0.39)	70 2,253 (2,026–2,506)	104 1,481 (1,324–1,657)	25 1,659 (1,373–2,004)	4.9×10^{-5}	3.9×10^{-8}	0.57
rs7202633	16	<i>TMEM114</i>	A/T (0.44)	55 1,324 (1,138–1,541)	110 1,823 (1,639–2,027)	32 2,410 (2,058–2,822)	2.0×10^{-7}	3.5×10^{-5}	3.7×10^{-5}
rs7202877	16	<i>CTRB1/CTRB2</i>	T/G (0.11)	156 1,612 (1,473–1,765)	36 2,424 (2,073–2,834)	4 2,638 (1,876–3,710)	8.8×10^{-7}	1.9×10^{-6}	0.017

Insulin secretion data are given as estimated means calculated with linear regression analysis (GEE) adjusted for age, sex, BMI, glucose tolerance status, study center, and ISI. *P* values are calculated using GEE assuming different genetic models with adjustment as described above. Genotype counts for carriers of the homozygous wild-type (11), heterozygous (12), and the mutant (22) allele are given. ^aAlleles at each locus are given as major allele/minor allele.

results compared with the inclusion of one randomly chosen twin from an MZ pair. GEE (SPSS, Inc., Amsterdam, the Netherlands) was also used to calculate the estimated means adjusted for the above-mentioned covariates. We computed the genomic inflation factor (Supplementary Fig. 1) using SNPs included for the QT interval as previously described (12,13). Quantile-quantile plots were constructed using Bioconductor R (SNPmatrix, www.bioconductor.org). LocusZoom was used to generate regional association plots of the loci of interest (19). The study-wide significance threshold was set at $P < 9 \times 10^{-7}$ based on the number of SNPs tested in this study (53,000).

Efficacy of GLP-1-based treatment in type 2 diabetic patients was tested by linear regression analysis. The difference between baseline and the lowest on treatment A1C measurement or the lowest on treatment A1C itself was used as the dependent variable. Age, sex, BMI, and estimated glomerular filtration rate were obtained at baseline. In order to take into account the possible effects of other glucose-lowering medication, which in some cases was discontinued at initiation of GLP-1-based therapy, we used “stopped medication” (yes/no) as additional covariate. In addition to the above-mentioned covariates, time between baseline and treatment A1C, diabetes duration, and dose of GLP-1-based drugs were considered as other potential covariates. For DPP-4 inhibitor users, a stepwise regression analysis showed significant effects of baseline A1C, time between baseline and treatment A1C, stopped other medication, and BMI as significant covariates in the model, and these were subsequently included together with the genotype. BMI was no longer a significant predictor in those using a GLP-1 RA and was thus dropped from the model. For meta-analysis of the results obtained in the two separate cohorts, we used CMA version 2 software (www.meta-analysis.com).

A paired Student *t* test was used to compare differences in fecal chymotrypsin activity between age-, sex-, and BMI-matched carriers and noncarriers of rs7202877. Fecal chymotrypsin activity was logarithmically transformed prior to analysis to obtain normality. For the above statistical analyses, SPSS version 20 (SPSS, Inc.) was used, and a $P \leq 0.05$ was regarded significant unless stated otherwise.

RESULTS

GLP-1-stimulated insulin secretion. Three loci showed genome or study-wide significance for association with the magnitude of GLP-1-stimulated insulin secretion (all $P \leq 8.8 \times 10^{-7}$) (Table 1). The regional plots for each locus are given in Supplementary Fig. 2. To check the robustness of the associations, we also analyzed the data from Dutch and German subjects separately and found that the effect sizes and directions were comparable (all $P \leq 1 \times 10^{-3}$).

At chromosome 10q22.1, several SNPs showed evidence for association, with rs4148941 being genome-wide significant. rs4148941 is located in the 3' untranslated region of the second exon of the *CHST3* gene. However, the linkage disequilibrium (LD) block also includes the nearby genes *SPOCK2* and *ASCC1*. Carriers of the C allele at rs4148941 (MAF = 0.39) had a 32% reduced GLP-1-stimulated insulin secretion ($P = 3.9 \times 10^{-8}$). Our second signal is represented by rs7202633, which is located on chromosome 16p13.2, 84 kb upstream of *TMEM114* ($P = 2.0 \times 10^{-7}$). Homozygous carriers of the minor allele of rs7202633 (MAF = 0.44) had an almost twofold increased response during stimulation with GLP-1. Finally the third most significant locus was tagged by rs7202877 near *CTRB1*, *CTRB2*, and *BCAR1* (MAF = 0.11) at chromosome 16q23.1, which has recently been identified as a locus that protects from type 2 diabetes (13). Subjects carrying the G allele showed a 33% increased GLP-1-stimulated insulin secretion ($P = 1.9 \times 10^{-6}$). To check whether these SNPs also associate with other measures of β -cell function, we also examined the effect of these SNPs on response to glucose- and arginine-stimulated insulin secretion. rs7202633 near *TMEM114* also showed an increased response to arginine stimulation during hyperglycemia ($P = 5.6 \times 10^{-4}$), but none of these SNPs were significantly associated with GSI (all $P > 0.01$) (Supplementary Table 4).

GLP-1 RA and DPP-4 inhibitor treatment study in type 2 diabetic patients. In total, 527 patients were

included, from the Dutch DCS West-Friesland and the U.K. GoDARTS studies (Supplementary Table 2). In both cohorts, patients carrying the G allele of rs7202877 near *CTRB1/2* showed a significantly smaller decrease in A1C levels after DPP-4 inhibitor treatment compared with TT carriers ($n = 354$) (Table 2 and Supplementary Tables 5–7). Meta-analysis of both cohorts showed that carriers of the G allele had an absolute $0.51 \pm 0.16\%$ (5.6 ± 1.7 mmol/mol) smaller decrease in A1C after adjustment for potential confounders ($P = 0.0015$) (Table 2). None of the other two loci showed significant association with DPP-4 inhibitor treatment response. In type 2 diabetic patients treated with a GLP-1 RA ($n = 173$), none of the SNPs showed a significant association; although rs4148941, located in *CHST3*, showed a trend toward lower treatment response in C-allele carriers ($0.24 \pm 0.13\%$ [2.6 ± 1.4 mmol/mol]; $P = 0.08$) (Table 2 and Supplementary Tables 5 and 6). In contrast to those treated with a DPP-4 inhibitor, carriers of the rs7202877 G allele treated with a GLP-1 RA showed a response comparable to carriers of the TT genotype ($P = 0.58$) (Table 2). There were no significant SNP \times drug interactions.

Given the results described above and the overlap with genetic loci for type 1 (20) and type 2 (13) diabetes, we focused our next steps on rs7202877 at chromosome 16q23.1. rs7202877 is located ~ 6 kb from the start of both *CTRB1* and *CTRB2*, which both encode for the digestive enzyme chymotrypsinogen, and ~ 54 kb from *BCAR1* encoding breast cancer antiestrogen resistance-1 (Supplementary Fig. 3).

Gene expression in pancreas and islets. In order to further elucidate the potential working mechanism, we examined whether rs7202877 acted as an expression quantitative trait locus (eQTL) by examining if it was associated with gene expression in RNA isolated from whole pancreas and isolated pancreatic islets. In 35 whole pancreata, the minor G allele at rs7202877 increases mRNA expression of *CTRB1/2* ($P = 0.01$) but not *BCAR1* (Fig. 1A). We next tested whether this SNP also acts as an eQTL in isolated islets from normal ($n = 24$) and hyperglycemic ($n = 21$) subjects. Using RNAseq data, we confirmed the observation from whole pancreatic tissue showing that the SNP also acts as an eQTL for *CTRB1* and *CTRB2* but not *BCAR1* in islets, which was independent from A1C level, age, sex, and BMI ($P \leq 0.05$) (Fig. 1B and C).

Chymotrypsin activity measurements. Given the enhanced expression of both *CTRB* genes, we tested whether rs7202877 was associated with altered chymotrypsin activity. In an analysis of 40 carriers of the G allele, compared with 40 age-, sex-, and BMI-matched noncarriers, the rs7202877 G allele was associated with an increased chymotrypsin activity in stool samples ($P = 0.023$) (Fig. 2A). Furthermore, chymotrypsin has been identified as a target for DPP-4 in mice (21). A 30-min preincubation of the stool samples with 0.1 units of DPP-4 resulted in a small but significantly lower chymotrypsin activity in TT carriers but not in carriers of the G allele (-5.3% [$P = 0.01$] and -2.1% [$P = 0.26$], respectively) (Fig. 2B).

DISCUSSION

In a unique sample of deeply phenotyped subjects who received GLP-1 during a hyperglycemic clamp to measure insulin secretion, we identified three loci that influence GLP-1-stimulated insulin secretion. The loci on chromosomes 10q22.1 (in *CHST3* and near *ASCC1* and *SPOCK2*)

TABLE 2
Beta values for linear regression of SNP effects on treatment A1C in type 2 diabetic patients treated with DPP-4 inhibitors or GLP1-RA

SNP; nearest gene(s)	DCS West-Friesland		GoDARTS		Meta-analysis	
	GLP-1 RA (n = 22)	DPP-4 inhibitors (n = 49)	GLP-1 RA (n = 151)	DPP-4 inhibitors (n = 305)	GLP-1 RA (n = 173)	DPP-4 inhibitors (n = 354)
rs4148941; <i>CHST3</i>	0.41 \pm 0.45%	-0.08 \pm 0.21%	0.22 \pm 0.14%	-0.07 \pm 0.10%	0.24 \pm 0.13%	-0.07 \pm 0.09%
	4.5 \pm 4.9	-1.0 \pm 2.3	2.4 \pm 1.5	-0.8 \pm 1.1	2.6 \pm 1.4	-0.8 \pm 1.0
rs7202633; <i>TMEM114</i>	0.43 \pm 0.45%	0.09 \pm 0.24%	0.03 \pm 0.15%	0.00 \pm 0.11%	0.07 \pm 0.14%	0.02 \pm 0.10%
	4.7 \pm 4.9	1.0 \pm 2.6	0.3 \pm 1.6	0.0 \pm 1.2	0.8 \pm 1.5	0.2 \pm 1.1
rs7202877; <i>CTRB1/2</i>	-1.33 \pm 0.74%	0.77 \pm 0.37%*	0.03 \pm 0.29%	0.45 \pm 0.18%**	-0.15 \pm 0.27%	0.51 \pm 0.16%
	-14.5 \pm 8.1	8.4 \pm 4.0	0.3 \pm 3.2	4.9 \pm 2.0	-1.6 \pm 3.0	5.6 \pm 1.7
					MAF	
					A/C (0.38)	
					A/T (0.43)	
					T/G (0.09)	
					P value	
					0.08	0.43
					0.62	0.87
					0.58	0.0015

Data are presented as A1C in % or mmol/mol adjusted for pretreatment A1C, stopped other baseline medication, BMI, and time between measures or pretreatment A1C, stopped other baseline medication, and time between measures for DPP-4 inhibitors and GLP-1 RA, respectively. The effect allele is in bold. * $P \leq 0.05$. ** $P \leq 0.01$.

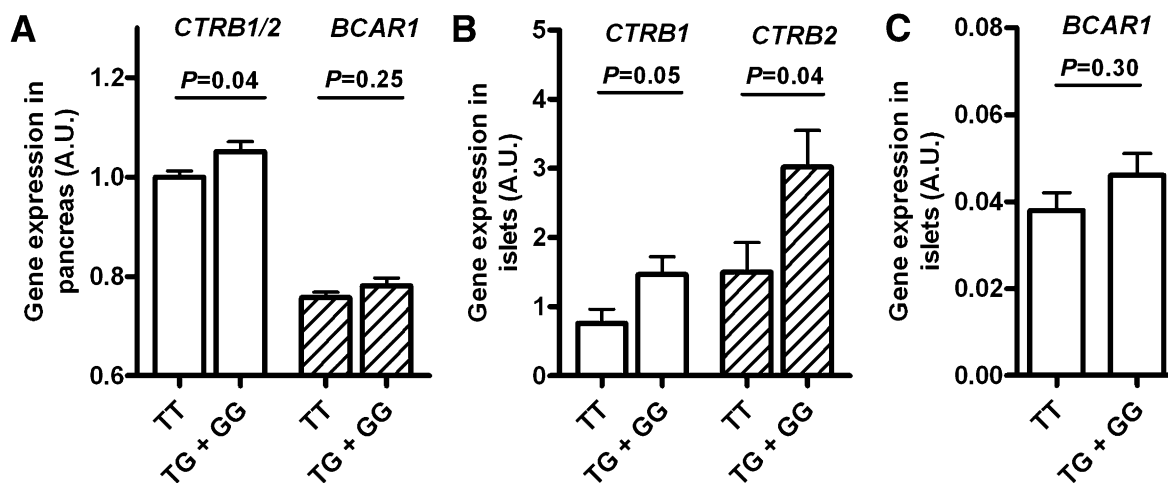


FIG. 1. Gene expression according to rs7202877 genotype in 35 whole pancreata and 45 human islet samples. Gene expression of *CTRB1/2* and *BCAR1* was measured using Taqman gene expression assays (pancreas) or RNAseq (islets). A: *CTRB1/2* and *BCAR1* gene expression in whole pancreas in TT ($n = 26$) and TG+GG ($n = 9$) carriers adjusted for age, sex, and BMI. *CTRB1/2* (B) or *BCAR1* (C) gene expression in islets from 27 TT carriers and 18 TG+GG carriers adjusted for age, sex, BMI, and A1C. A.U., arbitrary units.

and 16p13.2 (near *TMEM114*) have not previously been implicated in β -cell function, type 2 diabetes susceptibility, or related phenotypes. However, in publically available gene expression data from the MuTHER consortium, rs4148941 acts as eQTL for *CHST3* in lymphoblast cell lines ($P = 5 \times 10^{-51}$) and *SPOCK2* in both adipose tissue ($P = 1 \times 10^{-21}$) and lymphoblast cell line ($P = 3 \times 10^{-4}$) (22). Given the additional trend toward association with GLP-1 RA treatment response in diabetic patients, further research is warranted to establish its role in diabetes development and treatment.

Previously, the G allele of rs7202877 near *CTRB1/2* on chromosome 16q23.1 has been identified as a risk factor in a genome-wide association study for type 1 diabetes (20) and as a protective factor for development of type 2 diabetes (13). We now report positive effects of the G allele of rs7202877 on GLP-1-stimulated insulin secretion, which likely explains the protective effect of this allele on type 2 diabetes susceptibility. Furthermore, we observed decreased

glucose reduction in patients treated with DPP-4 inhibitors. Our observations are further supported by the fact that the SNP increases gene expression of *CTRB1/2* in islets and pancreata and enzyme activity in stool, placing *CTRB1/2* in a key role of the enteroendocrine system.

rs7202877 is located in an intergenic region between *CTRB1* and *CTRB2*, encoding chymotrypsinogen B1 and B2, respectively (Supplementary Fig. 3). In an in silico functional analysis using the regulome database, there is no known functional effect of rs7202877, indicated by a regulomeDB score of 5 (www.regulomedb.org) (23). At least 50 SNPs are in high LD ($r^2 > 0.8$) with rs7202877, some of which might affect gene expression but require further detailed study. The linkage disequilibrium block also encompasses the breast cancer antiestrogen resistance protein 1 gene (*BCAR1*); for which it has been shown that rs7202877 acts as an eQTL in blood (13), although this was not confirmed in the MuTHER consortium data (22). *BCAR1* is ubiquitously expressed in many tissues, including

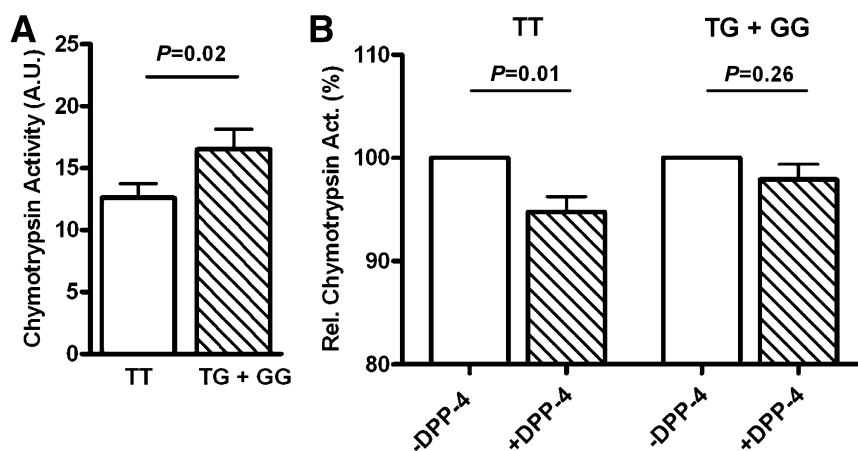


FIG. 2. Fecal chymotrypsin activity measurements according to rs7202877 genotype. Chymotrypsin activity in stool samples was measured with the method of Kaspar et al. (18). A: Mean chymotrypsin activity (Act. \pm SEM) of healthy volunteers with TT ($n = 40$) or TG+GG genotype ($n = 40$) at rs7202877 near *CTRB1/2*. B: Relative chymotrypsin activity (% Act. \pm SEM) in stool samples preincubated with or without human DPP-4 (0.1 units) for 30 min at 37°C. The samples without DPP-4 preincubation were used a reference. TT genotype, $n = 15$; TG+GG genotype, $n = 14$. A.U., arbitrary units.

the pancreas. However, given that the SNP acts as eQTL for *CTRB1/2* but not *BCAR1* in the pancreas and alters chymotrypsin activity in the gut, *CTRB1* and *CTRB2* are the most likely candidate genes in this region, although a role of *BCAR1* cannot be excluded (24). The chymotrypsinogens (EC 3.4.21.1) belong to a family of serine proteases that are secreted by the pancreas into the gastrointestinal tract as inactive precursors, which are then activated by proteolytic cleavage with trypsin. Chymotrypsin cleaves food proteins into smaller peptides suitable for further digestion. Chymotrypsinogens are highly expressed in the exocrine pancreas, but expression in β -cells has also been shown (25), which is confirmed in this study and in INS1E cells. Interestingly, in a study in mice, chymotrypsin has been identified as a target for DPP-4 (21). A peptide fragment (VPAIQPVLGTG) containing the DPP-4 H₂N-Xaa-Pro consensus motif was found at a 10-fold higher level in gut tissue from DPP-4^{-/-} versus DPP-4^{+/+} mice. Our in vitro findings, showing reduced chymotrypsin activity after preincubation with DPP-4, confirm this observation in humans and may add to the observed differences in chymotrypsin activity.

As previously shown, altered exposure of intestinal L cells to nutrients, resulting from changes in food composition, point of entry (after gastric bypass), rate of gastric emptying, and gut motility, may affect GLP-1 secretion (1,4,26). It could be hypothesized that the observed increase in fecal chymotrypsin activity may be associated with alterations in L-cell stimulation and/or gastric emptying as a result of increased nutrient digestion. However, incretin secretion during oral glucose tolerance test in nondiabetic subjects was not significantly different ($n = 457$) (Supplementary Table 8). It has been shown that oral treatment with chymotrypsin inhibitors slows gastric emptying and acutely reduces glucose and insulin levels after a meal in type 2 diabetic patients (27). Prolonged use of these inhibitors in a phase IIa cancer chemoprevention trial significantly reduced glucose levels in nondiabetic cancer patients (28). Interestingly, various vegetables, including potatoes and soybeans, contain high levels of trypsin and chymotrypsin inhibitors, and high intake of soy has been associated with reduced risk of diabetes (29). These results support the observation that rs7202877 genotype, enhancing pancreatic chymotrypsin expression and enzyme activity in the gut, alters diabetes risk.

The divergent effect of rs7202877 on treatment response to DPP-4 inhibitors and GLP-1 RA is intriguing. At present, our study design and sample size, especially regarding GLP-1 RA-treated subjects (~50% power to detect an effect with similar size to the DPP-4 inhibitor group), does not allow us to establish the robustness of the difference in treatment response between GLP-1 RA and DPP-4 inhibitors and/or whether these findings have clinical consequences. To this end, a well-powered genotype-selected prospective clinical trial would be required. It is, however, of note that the G allele at rs7202877 has opposing effects on type 1 and type 2 diabetes risk (13,20), which may hint at divergent effects at the level of the pancreas and other tissues, most likely the gut, differently affecting type 1 and type 2 diabetes susceptibility and GLP-1-based treatment responses, but this requires further exploration.

Type 2 diabetes, of which pancreatic endocrine dysfunction is a hallmark, has also been associated with pancreatic exocrine dysfunction (30). Furthermore, GLP-1-based therapies exert insulinotropic effects, but their use has also been linked to pancreatic exocrine disturbances, including pancreatitis and pancreatic cancer (31–33). In

this respect, the present finding is of interest as it identifies a genetic factor encoding a pancreatic digestive enzyme at the crossroad of pancreatic endocrine function and the response to GLP-1-based therapies. It is of further note that there is suggestive evidence from the PanScan consortium that SNPs, including rs7202877, near the *CTRB1/CTRB2* region are linked to pancreatic neoplasms (34). Studies into the interrelationship of these genetic factors and phenotypic defects are mandatory to help increase our insight into the complex pathophysiology of diabetes and its comorbidities as well as our understanding of the regulation and physiological function of the incretin pathway.

Limitations of this study include the following. Due to the unique and comprehensive deep phenotypes, the sample sizes are relatively small, although results are replicated and size and directionally consistent in all independent cohorts. A further limitation is that most studies presented here only involve nondiabetic individuals, and thus our results do not necessarily represent the situation in diabetic states. The power in our pharmacogenetic studies is limited, and larger, better-powered studies are needed to check the robustness and clinical utility of our associations. Finally, the Metachip focuses on loci previously implicated in metabolic and cardiovascular phenotypes, and although this increases power, it is likely that other loci exist that are not captured with this chip.

In conclusion, several loci are associated with GLP-1-stimulated insulin secretion in healthy nondiabetic volunteers. One of the loci, *CTRB1/2* on chromosome 16q23.1, was further associated with a reduced response to treatment with DPP-4 inhibitors in type 2 diabetic patients. Carriers of the G allele at rs7202877 additionally showed increased chymotrypsin mRNA expression and activity in the pancreas and feces, respectively. These results highlight the importance of chymotrypsin in the incretin response and regulation of glucose homeostasis and may provide novel targets for preventive and therapeutic strategies in the combat of diabetes.

ACKNOWLEDGMENTS

This research was supported by grants from the Dutch hub of the Biobanking and Biomolecular Research Infrastructure (BBMRI-NL, NWO 184.021.007, Grant CP1), the Dutch Diabetes Research Foundation (Grant 2006.00.060), an unrestricted investigator-initiated grant of Merck Sharp & Dohme, and the German Federal Ministry of Education and Research (BMBF 01GI0925 and 0315381B). The Wellcome Trust provides support for Wellcome Trust U.K. Type 2 Diabetes Case Control Collection (GoDARTS, 099177/Z/12/Z). Metachip genotyping of GoDARTS was funded by the U.K. Medical Research Council (G0601261), and the informatics support was provided by the Chief Scientist Office and the Wellcome Trust-funded Scottish Health Informatics Programme.

A.F., G.N., J.M.D., E.J.P.d.K., E.R.P., and M.D. have indicated that they have a potential duality of interest as they serve on advisory boards or act as a consultant or speaker or have received grants from (pharmaceutical) companies who may either gain or lose from this publication. These companies include Abbott, Amylin, Astra-BMS, Boehringer Ingelheim, Eli Lilly and Company, Merck Sharp & Dohme, Novo Nordisk, Sanofi, and Poxel Pharma. None of these companies were involved in this work or during the writing process. No other potential conflicts of interest relevant to this article were reported.

L.M.t.H., A.F., G.N., N.v.L., E.J.C.d.G., D.I.B., E.M.W.E., E.R.P., and M.D. researched data and wrote the manuscript. L.A.D., M.A., J.F., F.C., A.P.G., T.W.v.H., S.A.H.-S., A.M.C.S.-B., J.J.H.-D., Q.H., J.D., B.G., F.M., G.W., and J.J.H. researched data. J.M.D., C.N.A.P., T.H., R.J.H., M.H.H.K., E.J.P.d.K., H.-U.H., O.P., L.G., and P.E.S. contributed to the discussion and reviewed and edited the manuscript. L.M.t.H. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

This study has been submitted for presentation at the 49th Annual Meeting of the European Association for the Study of Diabetes, Barcelona, Spain, 23–27 September 2013.

The authors thank all participants for their cooperation.

REFERENCES

- Baggio LL, Drucker DJ. Biology of incretins: GLP-1 and GIP. *Gastroenterology* 2007;132:2131–2157
- Holst JJ. The physiology of glucagon-like peptide 1. *Physiol Rev* 2007;87:1409–1439
- DeFronzo RA, Okerson T, Viswanathan P, Guan X, Holcombe JH, MacConell L. Effects of exenatide versus sitagliptin on postprandial glucose, insulin and glucagon secretion, gastric emptying, and caloric intake: a randomized, cross-over study. *Curr Med Res Opin* 2008;24:2943–2952
- Nauck MA. Incretin-based therapies for type 2 diabetes mellitus: properties, functions, and clinical implications. *Am J Med* 2011;124(Suppl.):S3–S18
- Müssig K, Staiger H, Machicao F, Häring HU, Fritsche A. Genetic variants affecting incretin sensitivity and incretin secretion. *Diabetologia* 2010;53:2289–2297
- Simonis-Bik AM, Eekhoff EM, de Moor MH, et al. Genetic influences on the insulin response of the beta cell to different secretagogues. *Diabetologia* 2009;52:2570–2577
- 't Hart LM, Simonis-Bik AM, Nijpels G, et al. Combined risk allele score of eight type 2 diabetes genes is associated with reduced first-phase glucose-stimulated insulin secretion during hyperglycemic clamps. *Diabetes* 2010;59:287–292
- Fritsche A, Madaus A, Renn W, et al. The prevalent Gly1057Asp polymorphism in the insulin receptor substrate-2 gene is not associated with impaired insulin secretion. *J Clin Endocrinol Metab* 2001;86:4822–4825
- Zavrelouva H, Hoekstra T, Alsema M, et al. Progression and regression: distinct developmental patterns of diabetic retinopathy in patients with type 2 diabetes treated in the Diabetes Care System West-Friesland, the Netherlands. *Diabetes Care* 2011;34:867–872
- Morris AD, Boyle DIR, MacAlpine R, et al.; DARTS/MEMO Collaboration. The diabetes audit and research in Tayside Scotland (DARTS) study: electronic record linkage to create a diabetes register. *BMJ* 1997;315:524–528
- Fritsche A, Stefan N, Hardt E, Schützenauer S, Häring H, Stumvoll M. A novel hyperglycaemic clamp for characterization of islet function in humans: assessment of three different secretagogues, maximal insulin response and reproducibility. *Eur J Clin Invest* 2000;30:411–418
- Voight BF, Kang HM, Ding J, et al. The metabochip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. *PLoS Genet* 2012;8:e1002793
- Morris AP, Voight BF, Teslovich TM, et al.; Wellcome Trust Case Control Consortium; Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) Investigators; Genetic Investigation of ANthropometric Traits (GIANT) Consortium; Asian Genetic Epidemiology Network–Type 2 Diabetes (AGEN-T2D) Consortium; South Asian Type 2 Diabetes (SAT2D) Consortium; Diabetes Genetics Replication And Meta-analysis (DIAGRAM) Consortium. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nat Genet* 2012;44:981–990
- Zhou K, Bellenguez C, Spencer CC, et al.; GoDARTS and UKPDS Diabetes Pharmacogenetics Study Group; Wellcome Trust Case Control Consortium 2; MAGIC investigators. Common variants near ATM are associated with glycemic response to metformin in type 2 diabetes. *Nat Genet* 2011;43:117–120
- Taneera J, Lang S, Sharma A, et al. A systems genetics approach identifies genes and pathways for type 2 diabetes in human islets. *Cell Metab* 2012;16:122–134
- Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 2009;25:1105–1111
- Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 2010;11:R25
- Kaspar P, Moller G, Wahlefeld A, Staehler F. A new photometric method for the determination of chymotrypsin in stool. *Fresenius Z Anal Chem* 1982;311:391–392
- Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336–2337
- Barrett JC, Clayton DG, Concannon P, et al.; Type 1 Diabetes Genetics Consortium. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet* 2009;41:703–707
- Tinoco AD, Tagore DM, Saghatelian A. Expanding the dipeptidyl peptidase 4-regulated peptidome via an optimized peptidomics platform. *J Am Chem Soc* 2010;132:3819–3830
- Grundberg E, Small KS, Hedman AK, et al.; Multiple Tissue Human Expression Resource (MuTHER) Consortium. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* 2012;44:1084–1089
- Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* 2012;22:1790–1797
- Mercader JM, Puiggros M, Segrè AV, et al.; DIAGRAM Consortium; MITIN Consortium. Identification of novel type 2 diabetes candidate genes involved in the crosstalk between the mitochondrial and the insulin signaling systems. *PLoS Genet* 2012;8:e1003046
- Eizirik DL, Sammeth M, Bouckenouoghe T, et al. The human pancreatic islet transcriptome: expression of candidate genes for type 1 diabetes and the impact of pro-inflammatory cytokines. *PLoS Genet* 2012;8:e1002552
- Pournaras DJ, Aasheim ET, Bueter M, et al. Effect of bypassing the proximal gut on gut hormones involved with glycemic control and weight loss. *Surg Obes Relat Dis* 2012;8:371–374
- Schwartz JG, Guan D, Green GM, Phillips WT. Treatment with an oral proteinase inhibitor slows gastric emptying and acutely reduces glucose and insulin levels after a liquid meal in type II diabetic patients. *Diabetes Care* 1994;17:255–262
- Armstrong WB, Kennedy AR, Wan XS, et al. Clinical modulation of oral leukoplakia and protease activity by Bowman-Birk inhibitor concentrate in a phase IIa chemoprevention trial. *Clin Cancer Res* 2000;6:4684–4691
- Villegas R, Gao YT, Yang G, et al. Legume and soy food intake and the incidence of type 2 diabetes in the Shanghai Women's Health Study. *Am J Clin Nutr* 2008;87:162–167
- Larger E, Philippe MF, Barbot-Trystram L, et al. Pancreatic exocrine function in patients with diabetes. *Diabet Med* 2012;29:1047–1054
- Cure P, Pileggi A, Alejandro R. Exenatide and rare adverse events. *N Engl J Med* 2008;358:1969–1970; discussion 1971–1972
- Gale EA. GLP-1-based therapies and the exocrine pancreas: more light, or just more heat? *Diabetes* 2012;61:986–988
- Singh S, Chang HY, Richards TM, Weiner JP, Clark JM, Segal JB. Glucagonlike peptide 1-based therapies and risk of hospitalization for acute pancreatitis in type 2 diabetes mellitus: a population-based matched case-control study. *JAMA Intern Med* 2013;173:534–539
- PanScan Consortium; Pancreatic Cancer Cohort Consortium and Pancreatic Cancer Case-Control Consortium. PanScan, 2013. Available from http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/analysis.cgi?study_id=phs000206.v3.p2&pha=2889. Accessed 1 April 2013