Glycated hemoglobin (HbA1c) is an important measure of glycemia in diabetes. HbA1c is influenced by environmental and genetic factors both in people with and in people without diabetes. We performed a genome-wide association study (GWAS) for HbA1c in a Finnish type 1 diabetes (T1D) cohort, FinnDiane. Top results were examined for replication in T1D cohorts DCCT/EDIC, WESDR, CACTI, EDC, and RASS, and a meta-analysis was performed. Three SNPs in high linkage disequilibrium on chromosome 13 near relaxin family peptide receptor 2 (RXFP2) were associated with HbA1c in FinnDiane at genome-wide significance ($P < 5 \times 10^{-8}$). The minor alleles of rs2085277 and rs1360072 were associated with higher HbA1c also in the meta-analysis with RASS ($P < 5 \times 10^{-8}$), where these variants had minor allele frequencies $\geq 1\%$. Furthermore, these SNPs were associated with HbA1c in an East Asian population without diabetes ($P \leq 0.013$). A weighted genetic risk score created from 55 HbA1c-associated variants from the literature was associated with HbA1c in FinnDiane but explained only a small amount of variation. Understanding the genetic basis of glycemic control and HbA1c may lead to better prevention of diabetes complications.

Glycated hemoglobin (HbA1c) reflects the level of blood glycemia and is commonly used to monitor glycemic control in diabetes and to diagnose type 2 diabetes. Genome-wide association studies (GWAS) have revealed $>60$ genetic variants to be associated with HbA1c in individuals without diabetes (1–6). A recent GWAS of multiple ethnic groups confirmed 18 variants previously associated with HbA1c and added 42 new loci (6). Some of the variants affect biological pathways maintaining glucose homeostasis. For example, single nucleotide polymorphisms (SNPs) in or near glucokinase (GCK) and melatonin receptor 1B (MTNR1B) are known to be associated with HbA1c, fasting plasma glucose, and $\beta$-cell function (3). HbA1c is further influenced by genes and pathways that affect erythrocytes. In particular, SNPs in hemochromatosis (HFE) and hexokinase 1 (HK1) (3) have an effect on HbA1c.
through nonglycemic pathways that affect erythrocyte physiology and survival.

HbA1c reflects the glycemic load over ~3 months (7), and traits such as hemoglobinopathies, alterations in intracellular glucose metabolism, or defects of glucose transport into erythrocytes affect HbA1c. In addition, diet, drugs, and insulin administration affect HbA1c in individuals with type 1 diabetes (T1D) who have little to no endogenous insulin production and, thus, rely on an external source of insulin to ensure glucose uptake in cells. Despite a strong environmental stimulus, HbA1c levels are correlated between both dizygotic twins and monozygotic twins concordant or discordant for T1D (8). Heritability estimates for HbA1c in individuals without diabetes range from 27% to 62% (8,9), thus providing evidence for a genetic component for the trait.

To date, one GWAS on HbA1c has been conducted in individuals with T1D in the Diabetes Control and Complications Trial (DCCT) (10). The authors found that rs1358030 near sortilin-related VPS10 domain receptor 1 (SORC1) was associated with HbA1c (P = 2.2 × 10⁻⁶). The current study aims to define more genetic variants associated with HbA1c using GWAS of 4,622 individuals with T1D from the Finnish Diabetic Nephropathy (FinnDiane) Study and attempts to replicate GWAS results in five additional cohorts including, altogether, 3,097 individuals with T1D.

RESEARCH DESIGN AND METHODS

Phenotypes

HbA1c

FinnDiane is a nationwide, comprehensive multicenter study launched in 1997. Local ethics committees approved the study protocol, and participants gave informed consent before participation. The study was performed in accordance with the Declaration of Helsinki. HbA1c was measured with standardized laboratory methods locally from a blood sample taken during the visits to the FinnDiane study centers. Additional HbA1c values were collected from laboratory records both before and after the FinnDiane visit. A shift from high-performance liquid chromatography method to immunoassays was seen in many laboratories during the HbA1c collection time (1 January 1991–1 June 2016). In Finland, the HbA1c measurements show a high correlation with the DCCT reference method (11). HbA1c measurements from <1 year of diabetes diagnosis and the measurements after end-stage renal disease (ESRD) diagnosis were excluded. A median of 19 HbA1c measurements per person (range 1–129) were collected (Supplementary Fig. 1). Phenotype in the GWAS was an intrapersonal mean HbA1c, and it roughly followed normal distribution (Supplementary Fig. 2). In the replication cohorts, a mean of natural log-transformed HbA1c was used. Similar phenotypic values were subsequently calculated in FinnDiane to enable comparison of effect sizes (β) in the meta-analysis.

Clinical Parameters

At the FinnDiane visit, each patient underwent a thorough clinical investigation as previously described (12). The GWAS covariates age and diabetes duration were calculated at the intrapersonal mean date of HbA1c measurements. Micro- or macroalbuminuria was defined based on the urinary albumin excretion rate (AER) in two out of three consecutive timed overnight or 24-h urine collections as previously described (12). ESRD was defined as ongoing dialysis or the patient having received a kidney transplant. ESRD diagnosis year was verified from hospital records. Hematological measurements were taken from whole blood samples of 429 individuals using flow cytometry in the Helsinki University Central Hospital Laboratory.

Statistical Methods and Analyses

Patient Selection

Individuals with age at diabetes onset of <40 years, insulin treatment initiated within 1 year of the diabetes diagnosis, and one or more HbA1c measurements were selected from FinnDiane. Relatedness of individuals was calculated with KING 1.40 (13). Among first-degree family members, the individual with the least number of HbA1c measurements (nHbA1c) was excluded, and altogether 4,622 individuals were included in the GWAS.

After also excluding more distantly related individuals according to the estimated genetic relationship matrix (≥0.05), we calculated heritability in 4,319 individuals. Narrow-sense heritability of HbA1c (h²) was the proportion of the phenotypic variance explained by the additive effects of the genotyped SNPs with or without covariates sex, age, duration, genotyping batch, and nHbA1c (GCTA, version 1.24.4) (14).

Power Calculations

Power calculations were performed with Quanto, version 1.2.4 (http://biostats.usc.edu/Quanto.html), using α-level of 5 × 10⁻⁸ (genome-wide significance). In the GWAS, we had 94% power to detect variants with a minor allele frequency (MAF) ≥5% and an effect size (β) 0.40 (non-transformed HbA1c [%]) [Supplementary Table 1]. In the replication cohort (n = 3,097), we had >57% and 99% power (two-sided α = 0.05, β = 0.30) to replicate variants with MAFs 1–5% or MAFs ≥5%, respectively.

Genotyping and GWAS

SNP genotyping was performed with HumanCoreExome-12 v. 1.0, -12 v. 1.1, or -24 v. 1.0 BeadChip (Illumina, San Diego, CA). SNPs with low genotyping call rate, deviation from Hardy-Weinberg equilibrium, MAF <1%, Mendelian inconsistency, allele frequency difference >20% (MAF ≥5%), or allele frequency difference >5% (MAF <5%) with 1000 Genomes European (EUR) population were filtered.

After quality control, phased haplotypes (SHAPEIT v2.r837) (15) of 316,899 SNPs were used for imputation with 1000 Genomes EUR phase 3, version 5 (16), using
Minimac3/Minimac3-omp v1.0.14 (17). A total of 6,019 individuals passed quality control when samples with genotyping rate <0.95 or extreme heterozygosity, sample mix-ups, and genetic outliers were removed.

GWAS was performed with dosages of 8,354,674 variants with imputation quality $r^2 \geq 0.7$ with PLINK v1.9b3.26 (18) using linear regression with covariates sex, age, diabetes duration, nHbA1c, genotyping batch, and 10 first principal components that were calculated with the EIGENSTRAT, version 3.0 (19). A genomic inflation factor $\lambda$ was 1.003 (Supplementary Fig. 3).

Altogether, 1,489 and 1,863 individuals were genotyped with the TaqMan technology and ABI Prism 7000/7900HT Sequence Detection Systems and SDS 2.3 software (Life Technologies, Foster City, CA) to verify the imputed genotypes of rs2085277 and rs17076364, respectively. Of the TaqMan-genotyped individuals, 1,259 of 1,489 (84.6%) for rs2085277 and 1,545 of 1,863 (82.9%) for rs17076364 were included in the HbA1c GWAS.

**Longitudinal HbA1c**

Longitudinal HbA1c data were studied with linear mixed models (nlme package in R, version 3.3.3) in 4,153 individuals with three or more HbA1c measurements. The top SNPs were analyzed with continuous time including diabetes duration (at first HbA1c), sex, age, genotyping batch, and 10 first principal components as covariates assuming random intercept and slope. The best model (lowest Akaike information criterion) had an autocorrelation structure of order 1 with continuous time covariate and a combined covariance structure with exponential variances for fitted model values and separate variances for HbA1c, measurements from different decades of diabetes duration.

**Phenotypic Comparisons**

Groups were compared with Student $t$ test, Welch $t$ test, or Mann-Whitney $U$ test in R. Histograms were used for evaluating normality. Linear regression was used to study phenotypic associations with covariates.

**Replication**

**T1D Cohorts.** Replication of the top SNPs was conducted in DCCT/Epidemiology of Diabetes Interventions and Complications (EDIC), the Wisconsin Epidemiologic Study of Diabetic Retinopathy (WESDR), the Coronary Artery Calcification in Type 1 Diabetes (CACTI) study, the Pittsburgh Epidemiology of Diabetes Complications (EDC) study, and the Renin Angiotensin System Study (RASS) comprising altogether 3,097 individuals with T1D (Fig. 1 and Supplementary Table 2). Individuals with conventional ($n = 667$) or intensive ($n = 637$) treatment of diabetes during DCCT were analyzed separately. CACTI was also divided (clinic, $n = 265$, and nonclinic, $n = 253$). SNPs with imputation quality $r^2 \geq 0.6$ and MAF $\geq 1\%$ in the individual replication cohorts were included in the meta-analysis. These SNPs were replicated in one to seven cohorts comprising 253–3,097 individuals. A mean of natural log-transformed HbA1c was the phenotype used in the replication cohorts, and to get comparable effect sizes ($\beta$) the association of the top SNPs and similarly transformed HbA1c was calculated in the discovery cohort. Thereafter, a fixed-effects meta-analysis based on $\beta$s and SEs was performed with GWAMA (20).

**Populations Without Diabetes and Literature Search.** The top nine independent variants were compared with publicly available data from GWAS meta-analyses of up to 123,517 individuals without diabetes of European ancestry (6). For variants not included in the GWAS by Wheeler et al. (6), we searched for proxy variants ($r^2 \geq 0.3$) in the 1000 Genomes EUR population using LDlink (21) (https://analysistools.nci.nih.gov/LDlink). Seven out of nine SNPs or proxy variants had HbA1c association data available. In addition, replication of three genome-wide significant SNPs on chromosome 13 was performed in the GWAS data from individuals with East Asian ancestry ($n = 23,461$), since these SNPs were not available in the Europeans, likely due to their low MAFs. Thereafter, we compared 86 previously reported variants for HbA1c with our GWAS results: 11 variants from the DCCT GWAS (10) and 75 variants from GWAS of individuals without diabetes.

**Genetic Risk Scores**

Genome-wide significant autosomal SNPs in the largest HbA1c GWAS (6), which had MAF $\geq 1\%$ and $P < 0.01$
also in the Europeans, were selected for genetic risk score (GRS). GRS-ALL comprised all 55 SNPs, and like Wheeler et al. (6), we calculated separate GRS for erythrocytic SNPs (n = 19) (GRS-E) and glycemic SNPs (n = 18) (GRS-G). GRS were the sum of expected number of risk alleles weighted by the effect size (β) of the SNP in the Europeans, multiplied by the number of SNPs and divided by the sum of βs. Association of the GRS with Hba1c was studied using linear regression with covariates age, sex, diabetes duration, and nHba1c.

After the FinnDiane GWAS, a weighted GRS, GRSFD, was created from nine noncorrelated SNPs showing significant or suggestive association with Hba1c (P < 5 × 10^{-6}). GRSFD was tested for associations with diabetes duration, diabetic nephropathy (DN), and laser-treated diabetic retinopathy.

**In Silico Analyses**

The closest genes were searched (Ensembl transcripts downloaded from https://genome.ucsc.edu/cgi-bin/hg Tables [July 2017]) for variants with P < 0.001 in the GWAS. A total of 9,742 SNPs were mapped to 4,470 unique transcripts and were then connected with the HUGO Gene IDs. A total of 1,175 genes were located <150 kb from the GWAS SNPs. PANTHER database (22) (http://www.pantherdb.org) was used to identify the underlying biological pathways, and STRING, version 10.5 (https://string-db.org), was used to study protein interactions.

For the nearest genes of the top variants, gene expression (Expression Atlas [http://www.ebi.ac.uk/gxa/home]) and protein expression (ProteomicsDB [https://www.proteomicsdb.org]) were studied. The ENCODE data were accessed through UCSC Genome Browser (http://www.genome.ucsc.edu) and RegulomeDB v1.1 (23) (http://regulome.stanford.edu/ [July 2017]). Expression quantitative trait loci (eQTL) of the top variants were studied using the Genotype-Tissue Expression project portal (GTEx V7) (https://www.gtexportal.org/home/). Phenoscanner V2 database (24) (http://www.phenoscanner.medschl.cam.ac.uk/phenoscanner) was used to find associations with metabolites.

**RESULTS**

Mean ± SD Hba1c was 8.5 ± 1.2% (69.5 ± 13.1 mmol/mol) in FinnDiane (Table 1). Narrow-sense heritability of Hba1c was 17.0% (P = 0.002) (unadjusted) or 15.3% (P = 0.003) when adjusted for covariates. Minor alleles of three SNPs on chromosome 13q12.3 showed genome-wide significant association with higher Hba1c values in the GWAS (lead SNP rs2085277: β = 0.42 [95% CI 0.27, 0.56], P = 1.52 × 10^{-6}) (Fig. 2). The three SNPs rs2085277, rs1360072, and rs17076364 are in high linkage disequilibrium (LD) in the 1000 Genomes EUR population and are located between β 3-glucosyltransferase (B3GLCT) (SNPs 232–261 kb downstream) and 146–175 kb upstream of the closest gene, relaxin family peptide receptor 2 (RXFP2) (Supplementary Fig. 4).

Additional genotyping confirmed the imputed rs2085277 and rs17076364 genotypes (Supplementary Table 3). The lead SNP rs2085277 was not significantly associated with insulin dose (IU/kg), coefficient of variation of Hba1c (Supplementary Table 4), or hematocrit or other hematological parameters that were measured in a small subcohort of 429 individuals with T1D (Supplementary Table 5). Altogether, 68 variants had P < 5 × 10^{-6} in the GWAS. Of these, five tightly correlated SNPs from chr13 locus and eight other independent SNPs with P < 5 × 10^{-6} were selected for replication and subsequent analyses. Notably, these variants were associated with Hba1c also when longitudinal Hba1c measurements were analyzed in a subcohort of 4,153 individuals with three or more Hba1c measurements (Supplementary Table 6).
Replication

**T1D Cohorts**

The five T1D cohorts all used the mean of the natural log-transformed HbA1c values as the end phenotype. Two SNPs showed significant replication with ln(HbA1c): rs373883321 in CACTI clinic (chr6:165,403,975, $P = 0.006$) and rs10459356 near the lead SNP rs2085277 on chromosome 13 in RASS ($P = 0.046$ [95% CI 0.030, 0.063]) and $P = 4.2 \times 10^{-8}$, respectively (Supplementary Table 7).

In a meta-analysis with the discovery cohort, variants rs2085277 and rs17076364 on chromosome 13 locus were similar to the top variant, rs2085277: $P = 3.7 \times 10^{-8}$ and $P = 4.2 \times 10^{-8}$, respectively ($P = 6.5 \times 10^{-8}$, respectively).

Of note, the MAFs for rs2085277 and rs1360072 on chromosome 13 were 4% in FinnDiane and 1% in RASS but only 0.2%–0.5% in the DCCT/EDIC and WESDR cohorts that were thus not included in the meta-analysis for these SNPs (Supplementary Fig. 5). Because of low imputation quality of the chr13 top variants ($r^2 < 0.60$), CACTI and EDC cohorts were also excluded from the meta-analysis of these SNPs. Low MAFs reduced the statistical power to replicate the findings, but in RASS the effect was in the same direction as in FinnDiane: lead SNP rs2085277, $\beta = 0.059$ [95% CI –0.079, 0.198] for ln(HbA1c), $P = 0.402$. The MAF for rs2085277 is similar to our study in the 1000 Genomes populations: 5.6% in FIN (Finnish in Finland), 0.5% in GBR (British in England and Scotland).

Table 2—Top loci in the FinnDiane HbA1c GWAS and results from replication in other T1D cohorts

| Variant | Position | Closest gene | FinnDiane (n = 4,622) | T1D replication cohorts | Meta-analysis |
|---------|----------|--------------|-----------------------|-------------------------|--------------|
|         |          |              | EA/OA | EAF (%) | $\beta$ (95% CI)* | $P$ | n cohorts; subjects | $P$ | $P$ |
| rs7514675 | 1:99,782,478 | LPPR4 | G/A | 10.1 | 0.23 (0.13, 0.33) | $2.7 \times 10^{-8}$ | 7; 3,097 | 12.4 | $1.5 \times 10^{-2}$ |
| rs144918527 | 3:12,740,810 | RAFT | T/TTG | 53.2 | 0.13 (0.08, 0.19) | $4.4 \times 10^{-6}$ | 5; 2,248 | 51.2 | $2.5 \times 10^{-4}$ |
| rs373883321 | 6:165,403,975 | MEAT6 | T/- | 34.4 | 0.13 (0.07, 0.18) | $3.9 \times 10^{-6}$ | 5; 2,248 | 36.8 | $2.1 \times 10^{-2}$ |
| rs478414 | 7:154,526,318 | DPP6 | C/T | 56.0 | 0.13 (0.07, 0.18) | $3.1 \times 10^{-6}$ | 7; 3,097 | 50.5 | $8.3 \times 10^{-3}$ |
| rs113241624 | 8:109,658,463 | TME74 | A/G | 1.2 | 0.60 (0.35, 0.85) | $3.2 \times 10^{-6}$ | 1; 596 | 1.0 | $8.5 \times 10^{-5}$ |
| rs4744017 | 9:93,615,717 | SYK | A/G | 47.9 | 0.12 (0.07, 0.17) | $8.8 \times 10^{-6}$ | 7; 3,097 | 46.8 | $8.3 \times 10^{-4}$ |
| rs2085277 | 13:32,167,717 | RXFP2 | T/A | 4.1 | 0.42 (0.27, 0.56) | $1.5 \times 10^{-8}$ | 1; 253 | 1.0 | $3.7 \times 10^{-8}$ |
| rs552894079 | 15:89,840,012 | FANCI | A/G | 97.6 | 0.40 (0.23, 0.56) | $8.1 \times 10^{-6}$ | NA | NA | NA |
| rs2595441 | 19:43,989,088 | PHLD3 | G/A | 72.8 | 0.14 (0.08, 0.19) | $2.0 \times 10^{-6}$ | 7; 3,097 | 76.2 | $3.2 \times 10^{-3}$ |

EA, effect allele; EAF, effect allele frequency; OA, other allele. *$\beta$ is the effect size for the mean HbA1c phenotype in the discovery cohort. Replication cohorts are two DCCT/EDIC cohorts (conventional [n = 667] and intensive [n = 637] treatment of HbA1c in the DCCT separately), WESDR (n = 596), RASS (n = 253), CACTI divided into two cohorts (nonclinic [n = 253] and clinic [n = 265]), and EDC (n = 253). Since the mean of natural log-transformed HbA1c was the phenotype used in the T1D replication cohorts, top SNP association with ln(HbA1c) was further calculated in the discovery cohort to get comparable $\beta$s and SEs for the meta-analysis. Meta-analysis $P$ value is from fixed-effects meta-analysis of FinnDiane and all replication cohorts restricted to variants with MAF ≥1% in the individual replication cohorts. Results for the variants rs1360072 and rs17076364 on chromosome 13 locus were similar to the top variant, rs2085277: FinnDiane, $P = 1.8 \times 10^{-8}$ and $P = 2.6 \times 10^{-8}$, and meta-analysis, $P = 4.2 \times 10^{-8}$ and $P = 6.5 \times 10^{-8}$, respectively. rs113241624 on chromosome 8 had MAF ≥1% only in WESDR and rs2085277 on chromosome 13 only in RASS. SNP rs552894079 was not available in the T1D replication cohorts.
and 0% (monomorphic) in the CEU (Northern Europeans from Utah). The variant showed a markedly higher frequency in East Asians and South Asians with MAFs of 38.5% and 9.2%, respectively.

**Cohorts Without Diabetes**
The association of the thirteen SNPs with HbA1c was then queried in GWAS meta-analysis data on multiple cohorts without diabetes. As expected due to low MAF in Europeans, no data were available for the chr13 top SNPs or for any proxy SNPs with $r^2 \geq 0.3$ in the European cohorts. Therefore, replication of the chr13 variants was sought in the meta-analysis results of individuals of Asian ancestry. The SNPs were missing from the South Asian cohorts, but the chr13 top SNPs rs2085277 and rs1360072 were associated with HbA1c (%) in East Asians with $\beta = 0.02$ (95% CI 0.006, 0.03), $P = 0.005$, and $\beta = 0.02$ (0.003, 0.03), $P = 0.013$, respectively. (Table 3). None of the other top SNPs or the proxy variants were associated with HbA1c in the Europeans (Supplementary Table 9).

**In Silico Analyses**
Interestingly, minor T allele for rs2085277 was associated with higher levels of ketone bodies acetoacetate and 3-hydroxybutyrate (25) (PhenoScanner) ($P = 0.003$ and $P = 0.020$, respectively). No direct ENCODE data were available for the chr13 top SNPs or for any proxy SNPs with $r^2$ $= 0.003$ and $P = 0.005$, respectively. (Table 3). None of the other top SNPs or the proxy variants were associated with HbA1c in the Europeans (Supplementary Table 9).

$RXFP2$ belongs to the positive regulation of cAMP biosynthesis process pathway (Gene Ontology term in the STRING protein network analysis [Fig. 3]). Interestingly, when the search was restricted to databases as opposed to all interaction sources (databases, text mining, experiments, coexpression, neighborhood, gene fusion, and co-occurrence), gastric inhibitory polypeptide receptor (GIPPR) and glucagon (GCG)—proteins relevant for glucose homeostasis and diabetes—occurred in the same protein interaction network with $RXFP2$.

Genes involved in cell-cell adhesion processes and developmental processes were overrepresented in GWAS results (PANTHER, $P < 0.05$, corrected for multiple testing [Supplementary Table 10]). PANTHER pathway analysis pointed out cadherin signaling, endogenous cannabinoid signaling, and Wnt signaling pathways ($P < 0.05$, corrected for 160 independent classes).

**Duration of Diabetes and Chronic Diabetes Complications**
Our GWAS top variants had a consistent effect on HbA1c irrespective of diabetes duration. In all 10-year diabetes duration bins, the GRSFD (Supplementary Table 11 and Supplementary Fig. 6) was significantly associated with HbA1c (Supplementary Table 12). We confirmed this constant association between GRSFD and HbA1c in a small group of individuals ($n = 271$) having HbA1c measurements from three consecutive diabetes duration bins: 1–9.9, 10–19.9, and 20–29.9 years.

Subjects with T1D with DN had higher mean HbA1c compared with individuals with normal AER (8.9 ± 1.3% vs. 8.2 ± 1.0%, $P < 10^{-16}$). Similarly, individuals with diabetic retinopathy had higher HbA1c compared with non–laser-treated subjects with at least 15 years of diabetes duration (8.8 ± 1.2% vs. 8.4 ± 1.2%, $P < 10^{-16}$). As expected, individuals with DN had higher GRSFD (6.5 vs. 6.4, $P = 0.038$), while the subjects with diabetic retinopathy did not have higher GRSFD ($P = 0.205$).

**Previous HbA1c Studies**
rs1358030 was previously associated with HbA1c in the GWAS of 1,307 individuals with T1D (10). However, this association was not replicated in FinnDiane, although the MAFs were similar (MAFDCCT = 36% vs. MAFFinnDiane = 40%), imputation quality was $r^2 = 0.99$, and the effect was

| Table 3—Three genome-wide significant variants on chromosome 13: LD structures and HbA1c association in a population without diabetes |
| Variant | EA < OA | MAF (%) | $r^2$ | HbA1c association in a Finnish T1D cohort | HbA1c association in an East Asian population without diabetes |
|---------|--------|---------|------|-----------------------------------------------|----------------------------------------------------------|
|         | FD     | EAS     | SAS  | EUR  | $eta$ (95% CI) | $P$ | $\beta$ (95% CI) | $P$ |
| rs2085277 | T < A  | 4.2     | 38.5 | 9.2  | 1.2  | 0.42 (0.27, 0.56) | $1.5 \times 10^{-8}$ | 0.02 (0.006, 0.03) | 0.005 |
| rs1360072 | C < T  | 4.1     | 34.6 | 9.0  | 1.2  | 0.75 (0.65, 0.85) | 1.0 | 0.41 (0.27, 0.56) | $1.8 \times 10^{-8}$ | 0.02 (0.003, 0.03) | 0.013 |
| rs17076364 | T < C  | 3.1     | 2.5  | 0.2  | 1.2  | 0.97 (0.84, 1.1)  | 0.97 | 0.04 (0.02) | 0.02 | 1.0 | 0.45 (0.29, 0.61) | $2.6 \times 10^{-8}$ | 0.02 (0.003, 0.03) | 0.02 |
| n | 4,622 | 4,622 | 4,622 | 4,622 | 4,622 | 4,622 | 4,622 | 4,622 | 23,461 | 23,461 |

EA is the effect allele in the FinnDiane GWAS and OA is the other allele. *MAF frequency in the FinnDiane cohort (FD) and in East Asian (EAS), South Asian (SAS), and EUR populations in the 1000 Genomes phase 3 data. †LD ($r^2$) with rs2085277 in the FinnDiane and in the 1000 Genomes phase 3 populations. Position of the SNPs in the human genome build GRCh37 (hg19) was rs2085277, chr13:32,167,717; rs1360072, chr13:32,138,331; and rs17076364: chr13:32,163,399. HbA1c association statistics for these variants were not available for the European population without diabetes ($n = 123,517$) or South Asian population without diabetes ($n = 9,908$) (Wheeler et al. [6]).
in the same direction ($\beta = 0.034$ [95% CI $-0.02$, $0.09$], $P = 0.192$) (Supplementary Table 13).

Of the SNPs identified in populations without diabetes, four variants were associated with HbA1c in FinnDiane: the minor A allele of the missense variant rs1800562 in HFE was associated with lower HbA1c ($P = 0.007$) with higher effects size than in the population without diabetes (Fig. 4). The HbA1c values for the genotypes AA ($n = 3$), AG ($n = 250$), and GG ($n = 4,369$) were $7.6 \pm 1.1\%$, $8.3 \pm 1.2\%$, and $8.5 \pm 1.2\%$, respectively. Variants rs906216 and rs7072268 in HK1 and rs12621844 near FOXN2 were also directionally consistently associated with HbA1c in FinnDiane ($P < 0.05$) (Supplementary Tables 13 and 14). Of note, 16 of 19 SNPs classified as erythrocytic variants by Wheeler et al. (6) showed similar direction of effect in FinnDiane compared with the original study (binomial test for enrichment, $P = 0.002$ [Fig. 4]).

GRS

Next, we tested the combined genetic effect by creating GRS from significant variants found in the largest HbA1c GWAS (6) (Supplementary Fig. 7). A GRS-ALL comprising 55 SNPs was associated with HbA1c in FinnDiane ($\beta = 0.010$ [95% CI $0.002$, $0.018$], $P = 0.014$) and, along with covariates, explained 2.5% of the variation in HbA1c. However, age, diabetes duration, sex, and $n$HbA1c alone explained 2.4% of the variation, whereas GRS-ALL explained only 0.1% of the variation in HbA1c. A similar amount of HbA1c variance was also explained by GRS-E including only 19 erythrocytic SNPs ($\beta = 0.021$ [95% CI $0.005$, $0.037$], $P = 0.01$). In contrast, GRS-G created from the 18 glucemic variants was not associated with HbA1c ($\beta = 0.004$ [95% CI $-0.01$, $0.018$], $P = 0.573$).

**DISCUSSION**

We identified a locus on chromosome 13 that show genome-wide significant association with HbA1c in 4,622 Finnish individuals with T1D. The three associated SNPs in the locus are located <30 kb from each other and are in high LD, therefore indicating the same association signal. Notably, the association with HbA1c at rs2085277 and rs1360072 remained genome-wide significant in a meta-analysis including individuals with T1D from RASS. In DCCT/EDIC and WESDR, the MAFs for the chr13 top SNPs were $\leq 0.5\%$. The low MAFs ($\leq 1\%$) might explain the lack of significant replication in the T1D cohorts, since the power to detect significant association with MAF $\leq 1\%$ is limited. On the contrary, both SNP associations were replicated ($P \leq 0.013$) in the East Asian population without diabetes, where the SNPs were present with high allelic frequency of 38.5% and 34.5% for rs2085277 and rs1360072, respectively.

rs2085277 is an intergenic SNP surrounded by B3GLCT (261 kb 5’) and RXFP2 (146 kb 3’). More distant B3GLCT is a glucosyltransferase that adds glucose molecules to multiple proteins. RXFP2 seems a more plausible candidate gene, since many SNPs in close vicinity to rs2085277 (although not in LD) have cis-eQTLs for RXFP2 in adrenal gland. RXFP2 is a receptor for insulin-like peptide 3 (INSL3) and relaxin, and RXFP2 is expressed at low levels in many tissues. ProteomicsDB shows RXFP2 expression in adrenal glands and in pancreatic islets. Although the pharmacology and biology of relaxin receptors are relatively well-known (26), we are not aware of any previous studies on RXFP2 in relation to glucose homeostasis. Of note, protein interaction analysis restricted to databases shows that RXFP2 is linked to the glucagon

**Figure 3**—Protein networks around RXFP2 (red). A: Two most enriched Gene Ontology (GO) biological pathways were “cellular response to glucagon stimulus” (GNB1, GNB4, GNB2, and GNB3) and “positive regulation of cAMP biosynthetic process” (RXFP2, RLN2 [relaxin], AVPR2, and INSL3), with false discovery rates $4.6 \times 10^{-5}$ and $2.8 \times 10^{-5}$, respectively. Protein network was generated with STRING, version 10.5, selecting all active interaction sources, minimum required interaction score 0.90 (highest confidence), and no more than 10 interactions. B: Only database interaction sources were selected with minimum required interaction score 0.90 (highest confidence) and no more than 10 interactions. Six out of 11 proteins in this network belong to “positive regulation of cAMP biosynthetic process” (GO biological process), which was the most enriched pathway in this network, connecting proteins RXFP2, GIPR, GCG, ADCYAP1, MC1R, and FSHR (false discovery rate $1.3 \times 10^{-5}$). Of note is that gastric inhibitory polypeptide receptor (GIPR) and glucagon (GCG) present in this protein network have a high relevance in glucose homeostasis and diabetes.
response pathway and to relevant proteins for glucose homeostasis, like GIPR and GCG. Interestingly, adenylate cyclase 5 (ADCY5), which has previously been linked to HbA1c and T2D (6,27), is found in the same protein interaction network. However, there is yet no experimental evidence that would link these proteins. Thus, the biological mechanisms by which RXFP2 possibly affects HbA1c remain to be elucidated.

RXFP2 stimulates adenylate cyclase to increase intracellular second messenger cAMP concentrations. The main ligand for RXFP2 is relaxin hormone, which is naturally increased in pregnancy, and it has both vasodilative and antifibrinolytic actions (26). Interestingly, relaxin might act along with insulin to control blood glucose homeostasis (28). This was also suggested by a study in mice, where relaxin treatment ameliorated the insulin sensitivity of mice on a high-fat diet (29). The main receptor for human relaxin is relaxin family peptide receptor 1 (RXFP1) but, to a smaller extent, relaxin is also a ligand for RXFP2.

Our main phenotype was mean HbA1c, which was based on a median of 19 measurements per individual. One limitation of the study is that some individuals had only one HbA1c measurement. A single HbA1c measurement is not optimal for T1D studies, since HbA1c values tend to vary within a person over time. When a smaller subcohort of FinnDiane patients with three or more HbA1c measurements was studied with another method, linear mixed models, all top variants were consistently associated with HbA1c, although the effect sizes tended to be smaller. This was maybe due to a smaller cohort (n = 4,153) and deviation from the linear trend over time (individual slopes) in the HbA1c values.

In diabetes, HbA1c values are used as an indicator of glycemic control achieved with insulin treatment or with drugs improving insulin sensitivity, the latter mainly used in patients with type 2 diabetes. The heritability estimate for HbA1c in our study was modest but significant (h² = 15%) and confirmed that HbA1c in T1D can hold a heritable component (30). In the HbA1c GWAS in DCCT, a variant near SORCS1 was found. We were, however, unable to replicate the finding, although the effect was in the same direction. The lack of replication may be due to the winner’s curse, which causes the effect in the discovery cohort to be overestimated, and therefore the power for replicating the true smaller genetic effect is lower. Here, we had 100% power to replicate rs1358030 with a β = 0.045 [for ln(HbA1c)] and only 62% power with a β = 0.003 corrected for winner’s curse with a bootstrap method at two-sided α = 0.05 (31). On the other hand, four variants in or near HFE, HK1, and FOXN2 originally found in cohorts without diabetes were also associated with HbA1c (P < 0.05) in FinnDiane.

Variants associated with HbA1c have mostly been discovered in large GWAS meta-analyses and have relatively small effect sizes. Therefore, the power to replicate these variants in our study was only ~33%. However, a GRS comprising 55 SNPs associated with HbA1c in our T1D cohort, although it explained only a small part of the variation of HbA1c. Interestingly, when divided into erythrocytic and glycemic risk scores, GRS-E including only 19 variants classified as erythrocytic SNPs was also associated with HbA1c, while no association was seen for the glycemic risk score. This is expected, as the erythrocyte biology can be expected to be similar in individuals with and without diabetes, but the mechanisms affecting the level of glycemia are presumably different in individuals with T1D compared with the individuals without diabetes. As discussed by Wheeler et al. (6), the genetic factors primarily affecting erythrocyte biology may result in falsely high or low HbA1c levels no longer fully reflecting the level of glycemia and, thus, in false interpretation of HbA1c.
Allele frequencies of risk variants may differ between populations, and therefore it is important to conduct genetic analyses in different populations. Finns represent a population that has been genetically isolated from the rest of Europe, thus presenting an interesting opportunity to explore genetic variants affecting HbA1c in T1D. Indeed, the lead SNP was observed with a markedly higher MAF in the Finnish than in non-Finnish Europeans.

To conclude, we identified a locus on chromosome 13 near RXFP2 that is associated with HbA1c in individuals with T1D. Further studies are needed to find out whether relaxin hormone mediates the effect of RXFP2 on HbA1c. Since effect sizes are expected to be smaller in the population without diabetes, additional replication of the top locus as well as the search for additional genetic factors affecting HbA1c in other T1D cohorts is warranted. As HbA1c has a major influence on the development of diabetes complications, it is important to continue the search for genetic variants affecting HbA1c.

Acknowledgments. The authors thank the staff and all participants in the FinnDiane, DCCT/EDIC, CACTI, EDC, RASS, and WESDR study cohorts. Data on glycemic traits have been contributed by MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium) investigators and have been downloaded from www.magicinvestigators.org.

Funding. The FinnDiane Study and the work reported here were supported by grants from the Folkhalsan Research Foundation, Wilhelm och Else Stockmann Foundation, Novo Nordisk Foundation, Liv och Hälso Society, Helsinki University Central Hospital Research Funds (EVO) and Academy of Finland (299200 and 275614), European Foundation for the Study of Diabetes (to N.S.), and the Orion Research Foundation (to A.S.). JDRF supported the genotyping of the FinnDiane subjects (grant 17-2013-7). CACTI was performed at the Barbara Davis Center for Childhood Diabetes in Denver, CO; the Colorado Heart Imaging Center in Denver, CO; and the University of Colorado Hospital. Support for CACTI was provided by the National Heart, Lung, and Blood Institute, National Institutes of Health (NIH), grants R01-HL-61753, R01-HL-079611, and HL-113029, and JDRF grants 17-2013-313 and 17-2013-9. The study was performed at the Clinical Translational Research Center at the University of Colorado Denver supported by NIH MO1-RR-000051 and NIH/National Center for Advancing Translational Sciences Colorado CTSA grant UL1-TR-001082. RASS was partly supported by research grants from the NIH, the National Institute of Diabetes and Digestive and Kidney Diseases (DK-51975), and the Canadian Institutes of Health Research (DCT 14281). RASS was also supported in part by a grant from the National Center for Research Resources of the NIH to the University of Minnesota General Clinical Research Center (M01-RR00400). WESDR was supported by grants from Research to Prevent Blindness and by National Eye Institute grant EY016378 (B.E.K.K. and R.K.).

Duality of Interest. RASS was partly supported by research grants from Merck (in the U.S.) and Merck Frosst (in Canada). P.-H.G. has received investigator-initiated research grants from Eli Lilly and Boehringer Ingelheim, Boehringer Ingelheim, Ciba, Eli Lilly, Janssen, Medscape, Merck Sharp & Dohme, MSD, Mundipharma, Novartis, Novo Nordisk, and Sanofi; and has received lecture fees from AstraZeneca, Boehringer Ingelheim, Eli Lilly, Elow Water, Genzyme, MSD, Medscape, Novartis, Novo Nordisk, PeerVoice, and Sanofi. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. A.S. contributed to study design, researched FinnDiane data, performed meta-analysis, and wrote and edited the manuscript. N.S. prepared genotype data, contributed to study design, and wrote and reviewed the manuscript. J.C. prepared and analyzed data from the T1D replication cohorts. I.T. prepared the genotype data, contributed to data analysis, and reviewed the manuscript. D.M.M. and M.J.R. contributed to the CACTI study. J.K.S.-B. contributed to the CACTI study and reviewed the manuscript. T.C. contributed to the EDC study and reviewed and edited the manuscript. T.J.O. contributed to the EDC study and reviewed the manuscript. M.L.C. and M.M. contributed to RASS and reviewed the manuscript. B.E.K.K. and R.K. contributed to the WESDR study and reviewed the manuscript. E.V. reviewed the manuscript. M.P. contributed to the genome-wide genotyping and TaqMan SNP genotyping. C.F. and V.H. contributed to the collection of clinical data (FinnDiane) and reviewed the manuscript. A.D.P. contributed to study design and reviewed and edited the manuscript. P.-H.G. contributed to study design and reviewed and edited the manuscript. P.-H.G. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at the 54th Annual Meeting of the European Association for the Study of Diabetes, Berlin, Germany, 1–5 October 2018.

References
1. Paré G, Chasman DI, Parker AN, et al. Novel association of HK1 with glycated hemoglobin in a non-diabetic population: a genome-wide evaluation of 14,618 participants in the Women’s Genome Health Study. PLoS Genet 2008;4:e1000312
2. Franklin CS, Aulchenko YS, Huffman JE, et al. The TCF7L2 diabetes risk variant is associated with HbA1c levels: a genome-wide association meta-analysis. Ann Hum Genet 2010;74:471–478
3. Soranzo N, Sanna S, Wheeler E, et al.; WTCCC. Common variants at 10 genomic loci influence hemoglobin A1c levels via glycemic and nonglycemic pathways. Diabetes 2010;59:3229–3239
4. An P, Miljkovic I, Thyagarajan B, et al. Genome-wide association study identifies common loci influencing circulating glycated hemoglobin (HbA1c) levels in non-diabetic subjects: the Long Life Family Study (LLFS). Metabolism 2014;63:461–468
5. Chen P, Takeuchi F, Lee JY, et al.; CHARGE Hematology Working Group. Multiple nonglycemic genomic loci are newly associated with blood level of glycated hemoglobin in East Asians. Diabetes 2014;63:2551–2562
6. Wheeler E, Leong A, Lu CT, et al.; EPIC-CVD Consortium; EPIC-InterAct Consortium; Lifelines Cohort Study. Impact of common genetic determinants of hemoglobin A1c on type 2 diabetes risk and diagnosis in ancestrally diverse populations: a transethnic genome-wide meta-analysis. PLoS Med 2017;14:e1002383
7. Nathan DM, Turgeon H, Regan S. Relationship between glycated hemoglobin levels and mean glucose levels over time. Diabetologia 2007;50:2239–2244
8. Snieider H, Sawtell PA, Ross L, Walker J, Spector TD, Leslie RD. HbA1c levels are genetically determined even in type 1 diabetes: evidence from healthy and diabetic twins. Diabetes 2001;50:2858–2863
9. Meigs JB, Panhuysen CI, Myers RH, Wilson PW, Cupples LA. A genome-wide scan for loci linked to plasma levels of glucose and HbA1c in a community-based sample of Caucasian pedigrees: the Framingham Offspring Study. Diabetes 2002;51:833–840
10. Paterson AD, Waggott D, Boright AP, et al.; MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium); Diabetes Control and Complications Trial/Epidemiology of Diabetes Interventions and Complications Research Group. A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both A1C and glucose. Diabetes 2010;59:539–549
11. Petttilä IM, Halonen T, Punnonen K, Tiikkainen U. Best use of the recommended IFCC reference method, material and values in HbA1c analyses. Scand J Clin Lab Invest 2005;65:453–462
12. Syrjänen A, El-Osta A, Forsslom C, et al.; FinnDiane Study Group. Genetic examination of SETD7 and SUV39H1/H2 methyltransferases and the risk of diabetes complications in patients with type 1 diabetes. Diabetes 2011;60:3073–3080
13. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010;26:2867–2873
14. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011;88:76–82
15. Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods 2011;9:179–181
16. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. Nature 2015;526:68–74
17. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48:1284–1287
18. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81:559–575
19. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 2006;38:904–909
20. Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics 2010;11:288
21. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics 2015;31:3555–3557
22. Mi H, Huang X, Muruganujan A, et al. PANTHER version 11: Expanded annotation data from gene ontology and reactome pathways, and data analysis tool enhancements. Nucleic Acids Res 2017;45:D183–D189
23. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 2012;22:1790–1797
24. Staley JR, Blackshaw J, Kamat MA, et al. PhenoScanner: a database of human genotype-phenotype associations. Bioinformatics 2016;32:3207–3209
25. Kettunen J, Demirkan A, Würtz P, et al. Genome-wide study for circulating metabolites identifies 62 loci and reveals novel systemic effects of LPA. Nat Commun 2016;7:11122
26. Halls ML, Bathgate RA, Sutton SW, Dschietzig TB, Summers RJ. International Union of Basic and Clinical Pharmacology. XCV. Recent advances in the understanding of the pharmacology and biological roles of relaxin family peptide receptors 1-4, the receptors for relaxin family peptides. Pharmacol Rev 2015;67:389–440
27. Dupuis J, Langenberg C, Prokopenko I, et al.; DIAGRAM Consortium; GIANT Consortium; Global BPGen Consortium; Anders Hamsten on behalf of Procardis Consortium; MAGIC investigators. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42:105–116
28. Bani D, Pini A, Yue SK. Relaxin, insulin and diabetes: an intriguing connection. Curr Diabetes Rev 2012;8:329–335
29. Bonner JS, Lantier L, Hocking KM, et al. Relaxin treatment reverses insulin resistance in mice fed a high-fat diet. Diabetes 2013;62:3251–3260
30. Florez JC. A genome-wide association study of treated A1C: a genetic needle in an environmental haystack? Diabetes 2010;59:332–334
31. Sun L, Dimitromanolakis A, Faye LL, Paterson AD, Waggott D, Bull SB; DCCT/EDIC Research Group. BR-squared: a practical solution to the winner’s curse in genome-wide scans. Hum Genet 2011;129:545–552