Published in final edited form as:
Metabolism. 2014 April ; 63(4): 461–468. doi:10.1016/j.metabol.2013.11.018.

Genome-wide association study identifies common loci influencing circulating glycated hemoglobin (HbA_{1c}) levels in non-diabetic subjects: the Long Life Family Study (LLFS)

Ping An[a,*], **Iva Miljkovic**[b], **Bharat Thyagarajan**[c], **Aldi T. Kraja**[a], **E. Warwick Daw**[a], **James S. Pankow**[d], **Elizabeth Selvin**[e], **W. H. Linda Kao**[e], **Nisa M. Maruthur**[f], **Micahel A. Nalls**[g], **Yongmei Liu**[h], **Tamara B. Harris**[i], **Joseph H. Lee**[j], **Ingrid B. Borecki**[a], **Kaare Christensen**[k,l], **John H. Eckfeldt**[c], **Richard Mayeux**[j], **Thomas T. Perls**[m], **Anne B. Newman**[b], and **Michael A. Province**[a]

[a]Department of Genetics Division of Statistical Genomics, Washington University School of Medicine, St. Louis, Missouri, USA

[b]University of Pittsburgh, Graduate School of Public Health, Department of Epidemiology, Center for Aging and Population Health, Pittsburgh, Pennsylvania, USA

[c]Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota, USA

[d]Division of Epidemiology and Community Health, University of Minnesota, Minneapolis, Minnesota, USA

[e]Department of Epidemiology, Johns Hopkins University, Baltimore, Maryland, USA

[f]Division of General Internal Medicine, Johns Hopkins University, Baltimore, Maryland, USA

[g]Laboratory of Neurogenetics, NIA/NIH, Bethesda, Maryland, USA

[h]Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest University, Winston-Salem, North Carolina, USA

[i]Laboratory of Epidemiology, Demography and Biometry, NIA/NIH, Bethesda, Maryland, USA

[j]Gertrude H. Sergievsky Center and the Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University, New York City, New York, USA

[*]Corresponding author: Ping An, MD, Department of Genetics Division of Statistical Genomics, Washington University School of Medicine (Campus Box 8056), 4444 Forest Park Blvd, St. Louis, Missouri 63108, USA, Tel.: +001 314 3623614, anping@wustl.edu.

Author contributions

P.A. researched the data and wrote the manuscript. I.M., B.T., A.T.K., W.D., J.S.P., E.S., J.H.L., J.H.E., A.B.N. and M.A.P. either wrote paragraphs or sentences or edited the manuscript. All the authors in the LLFS participated in the weekly priority paper teleconference calls where this manuscript was initiated and revised. All the coauthors reviewed and approved submission of the manuscript to Diabetes. P.A. 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. We are grateful to Candace M. Kammerer, Ph.D. and Evan C. Hadley, M.D. for their comments that strengthened this manuscript. Each of the LLFS, ARIC and HABC Publications and Presentations Committees approved submission of this manuscript.

Conflict of interest

The authors declare that there is no duality of interest associated with this manuscript.

[k]Danish Aging Research Center, Epidemiology, University of Southern Denmark, Odense, Denmark

[l]Department of Clinical Biochemistry and Pharmacology and Department of Clinical Genetics, Odense University Hospital, Odense, Denmark

[m]Division of Geriatrics, Department of Medicine, Boston University Medical Center, Boston, Massachusetts, USA

Abstract

Objective—Glycated hemoglobin (HbA_{1c}) is a stable index of chronic glycemic status and hyperglycemia associated with progressive development of insulin resistance and frank diabetes. It is also associated with premature aging and increased mortality. To uncover novel loci for HbA_{1c} that are associated with healthy aging, we conducted a genome-wide association study (GWAS) using non-diabetic participants in the Long Life Family Study (LLFS), a study with familial clustering of exceptional longevity in the US and Denmark.

Methods—A total of 4,088 non-diabetic subjects from the LLFS were used for GWAS discoveries, and a total of 8,231 non-diabetic subjects from the Atherosclerosis Risk in Communities Study (ARIC, in the MAGIC Consortium) and the Health, Aging, and Body Composition Study (HABC) were used for GWAS replications. HbA_{1c} was adjusted for age, sex, centers, 20 principal components, without and with BMI. A linear mixed effects model was used for association testing.

Results—Two known loci at *GCK* rs730497 (or rs2908282) and *HK1* rs17476364 were confirmed ($p < 5e{-}8$). Of 25 suggestive ($5e{-}8 < p < 1e{-}5$) loci, one known (*G6PC2* rs560887, replication $p = 5e{-}5$) and one novel (*OR10R3P/SPTA1*- rs12041363, replication $p = 1e{-}17$) loci were replicated ($p < 0.0019$). Similar findings resulted when HbA_{1c} was further adjusted for BMI. Further validations are crucial for the remaining suggestive loci including the emerged variant near *OR10R3P/SPTA1*.

Conclusions—The analysis reconfirmed two known GWAS loci (*GCK*, *HK1*) and identified 25 suggestive loci including one reconfirmed variant in *G6PC2* and one replicated variant near *OR10R3P/SPTA1*. Future focused survey of sequence elements containing mainly functional and regulatory variants may yield additional findings.

Keywords

Genome-wide association study; Non-enzymatic glycation; Glucose, insulin resistance and diabetes; Premature aging processes

1. Introduction

HbA_{1c} is a form of hemoglobin bound by glucose through the non-enzymatic glycation pathway which provides information about glycation levels. It is associated with premature aging and increased mortality [1–3]. HbA_{1c} is known as a primary measure of prolonged (8–12 weeks) average ambient plasma glucose concentration in the circulation. HbA_{1c} appears to be a superior glycemic marker over fasting plasma glucose in that it is relatively stable at room temperature and can be obtained at almost any time with no requirement of special preparation by subject [4]. It has been used to monitor glycemic control and evaluate risk in development of long-term complications in patients with diabetes [5]. The 2010 American Diabetes Association Standards of Medical Care in Diabetes added $HbA_{1c} \geq 6.5\%$ as another criterion for the diagnosis of diabetes [6]. In addition, non-glycemic factors may also interfere with HbA_{1c} measurement; these include processes altering erythrocyte

turnover (e.g., anemia, hemolysis) and the glycation ability of hemoglobin (e.g., aspirin ingestion) [7].

HbA_{1c} has a confirmed genetic basis with moderate heritability estimates (47–59%) that are slightly higher than those reported for fasting glucose (34–36%, correlation with HbA_{1c} = 0.8) [8–11]. Genetic determinants of variability in HbA_{1c} and glucose homeostasis have been recently identified [11–13]. GWASs including a large study in the MAGIC Consortium have been performed in non-diabetic participants of European decent [14–16]. These studies along with candidate gene studies provided firm evidence of common variants at 15 loci for HbA_{1c} including *FN3K, HFE, TMPRSS6, ANK1, SPTA1, ATPIIA/TUBGCP3, HK1, MTNR1B, GCK, G6PC2/ABCB11, SLC30A8, TCF7L2, SORCS1, BNC2* and *MTNR1B.* These loci may influence HbA_{1c} levels through glycemic and/or non-glycemic pathways.

In the present study, we conducted a GWAS aiming to reconfirm known common loci and uncover additional common loci for HbA_{1c} among non-diabetic participants of the LLFS. For replications, we looked up GWAS of HbA_{1c} data from the ARIC and HABC.

2. Subjects and Methods

2.1. Cohort description and study design

The LLFS is a family-based cohort study designed to enroll families with exceptional longevity in order to identify genetic and environmental factors that promote long healthy lives in these families. The four recruitment centers represent three US field study centers (Boston University Medical Center, Boston, MA; Columbia College of Physicians and Surgeons, New York City, NY; University of Pittsburgh, Pittsburgh, PA) and one non-US field study center (University of Southern Denmark). In the US, the initial contacts included people who: were at least 79 years old on 1/1/2005; had no recorded date of death; were not in the end-stage renal disease or hospice programs; and lived within three hours of driving distance from the three US study centers. Subsequent mailings targeted those age 89 and older. Study participants were also recruited from local communities. The University of Southern Denmark field center identified individuals at least 90 years old during the study recruitment period through the Danish National Register of Persons. Probands were pre-screened for eligibility by phone. The Family Longevity Selection Score (FLoSS) was created [17] to facilitate selection of exceptional longevity families by ranking sibships according to current age or age at death of siblings, the size of the sibship and the number of alive individuals available for study. Each proband's family was required to have a FLoSS score of seven or higher and had to meet the following criteria: each proband was required to have at least one living sibling, and one of their living offspring (minimum family size of 3) that were able to give informed consent, to be interviewed, examined and giving blood sample for serum and DNA extraction. The LLFS study design has been described elsewhere [18]. For the current analysis, a total of 4,088 non-diabetic participants with European ancestry who had measured HbA_{1c} and genotype information were analyzed.

The ARIC is a prospective epidemiologic study conducted in four U.S. communities [19]. Each field center randomly selected and recruited a cohort sample of approximately 4,000 individuals aged 45–64 from a defined population in their community. A total of 15,792 participants received an extensive examination at baseline including medical, social and demographic data. These participants were examined multiple times with the first visit (baseline) occurring in 1987–89 and the second visit in 1990–92. A total of 6,777 non-diabetic participants with European ancestry and complete HbA_{1c} measures (the second visit) and genotypic information were assessed for replication. The ARIC joined 22 other GWAS of HbA_{1c} cohorts in the MAGIC Consortium for a meta-analysis where a total of 46,368 non-diabetic adults of European descent were assessed [16].

The HABC is a longitudinal cohort study designed to investigate relationships among health conditions, body composition, social and behavioral factors and functional decline. The study population included at baseline 3,075 well-functioning black and white men and women aged 70–79 (48% men, 42% Blacks) from Pittsburgh, PA and Memphis, TN. Baseline interview and clinic-based examination occurred between 4/1997–6/1998. The design of the Health ABC Study has been described elsewhere [20]. For this current study, only participants of European ancestry with available baseline HbA_{1c} measures (n = 1,454) were used.

All the study participants provided informed consent. The Institutional Review Board at each study center approved the consent forms and protocols.

2.2. Phenotypic measurements

In the LLFS, HbA_{1c} was measured as samples were collected over four years (2006–09) in EDTA whole blood at the Advanced Research and Diagnostics Laboratory (ARDL), University of Minnesota with the Tosoh 2.2 Plus and after 2007 with the Tosch G7 Glycohemoglobin Analyzer (Tosoh Medics, Inc., San Francisco, CA 94080) which use nearly identical ion exchange-based HPLC (high performance liquid chromatography) methodology. The University of Minnesota served as the central laboratory for the Diabetes Control and Complications Trial (DCCT) and has served as one of the National Glycohemoglobin Standardization Program's (NGSP) secondary reference laboratories since NGSP's inspection. Both Tosch instruments were meticulously calibrated to match the NGSP's BioRex 70 reference method at the University of Missouri that was used to standardize all of the DCCT study's HbA_{1c} measurements. The laboratory CV across a range of HbA_{1c} values ranged 1.4–1.9%.

In the ARIC, frozen whole blood samples were collected after an over-night fast. The samples were frozen, and later thawed and assayed for HbA_{1c} in 2003–04 using the Tosoh 2.2 Plus HPLC instrument and the remaining specimens in 2007–08 using the Tosoh G7 instrument also at the University of Minnesota's ARDL laboratory. In the HABC, HbA_{1c} was measured in approximately 1995 at the University of Vermont using the Biorad Variant (BioRad Laboratories, Hercules, CA 94547) that utilizes a similar ion exchange-based HPLC methodology for separation of HbA_{1c}.

In the LLFS, glucose was measured in serum at the University of Minnesota ARDL using the Roche hexokinase reagent and a Roche Modular P Chemistry Analyzer (Roche Diagnostics, Indianapolis, IN 46250). In this enzymatic method, glucose is converted to glucose-6-phosphate (G6P) by hexokinase in the presence of ATP. G6P dehydrogenase then converts G6P to gluconate-6-P in the presence of NADP. The resulting increase in absorbance at 340 nm as NADP is reduced to NADPH. The method is calibrated against Standard Reference Material 965a from National Institute of Standards and Technology (NIST) which is traceable to the NIST definitive method for glucose (Isotope Dilution Mass Spectroscopy, IDMS). The laboratory CV is 1.6%.

2.3. Genotyping, imputation and quality control

Direct genotyping in the LLFS. Genotypic data from LLFS blood assays using the Illumina Human Omni 2.5 v1 chips were produced by the Center for Inherited Disease Research (CIDR). A second round of QC-process was implemented in the Division of Statistical Genomics of Washington University. Using an acceptance threshold call rate > 98% per marker, we identified 83,774 SNPs with a lower call rate including 1,188 SNPs with high Mendelian error rate. A total of 3,647 SNPs with high Mendelian error rate were dropped. We also identified 18 subjects with a genotype call rate threshold < 97.5% per subject; their

genotypes were set to missing. Additionally, 153,363 Mendelian errors were set to missing in the families in which they occurred. Principal components (PCs) were produced with EIGENSTRAT [21] among 1,522 unrelated individuals using 116,867 tag-SNPs where in advance any SNPs with MAF < 5%, HWE-p < 1e–6 with missing genotypes were excluded. We also had excluded SNPs from some special regions (*2q21, 2q21.1*, HLA1 and HLA on chromosome 6, *8p23.1, 8p23* and *17q21.31*) which included inversions, HLA and other special regions that may drive the PC analysis. PCs produced from unrelated subjects were then expanded within EIGENSTRAT framework to all members of LLFS.

Imputation in the LLFS: Imputations were performed based on the cosmopolitan phased haplotypes of 1000 Human Genome (1000HG, version 2010–11 data freeze, 2012-03-04 haplotypes). Programs used for imputation were MACH (version 1.0.16, for pre-phasing of LLFS data) and MINIMACH (version of May 2012) for performing imputations and ChunkChromosome script for splitting the LLFS data into smaller blocks to speed the process of imputation [22–23]. Imputations were performed in chunks with 5,000 SNPs blocks and 1,000 SNPs overlap from our data. Filters before imputing were: removing markers that had MAF < 1%, HWE-p < 1e–6, if LLFS SNPs alleles mismatched with those of 1000HG, and absent in the 1000HG panel, as well as flipping any SNP when appropriate to the forward strand. A total of 2.23 M SNPs were typed, and a total of 36.02 M SNPs were imputed. For single SNP association testing with imputed dosage, two additional filters were implemented - the MAF > 1% and the r^2 > 0.3 that reduced the analysis to from 38.25 M to 9.25 M variants.

In the ARIC, the Affymetrix SNP 6.0 platform was used for genotyping and MACH v1.0 was used for imputation. SNP QC prior to imputation included using filters of MAF < 1%, HWE-p < 1e–6 and call rate < 95%. Sample QC included using filter of call rate > 95%, and ethnic outliers or other exclusions including gender mismatch, inferred first degree relatives, mismatch of ≥ 10 SNPs with SNPs previously genotyped on other platforms, genetic outlier as assessed by Identity-by-State using PLINK and > 8 SDs along any of the first 10 PCs in EIGENSTRAT with 5 iterations. A total of 5 SNPs were queried for replication.

In the HABC, genotyping was performed by the Center for Inherited Disease Research using the Illumina Human1M-Duo BeadChip system. Samples were excluded from the dataset for the reasons of sample failure, genotypic sex mismatch, and first-degree relative of an included individual based on genotype data. SNPs with MAF ≥ 1%, call rate ≥ 97% and HWE-p ≥ 1e–6 were used for imputation. MACH software (version 1.0.16) was used to impute SNPs on chromosome 1–22 with NCBI build 36 of Phase II HapMap CEU data (release 22) as the reference panel. A total of 5 SNPs were queried for replication.

2.4. Statistical analysis

Association tests in the LLFS. HbA_{1c} was adjusted for age, age^2, age^3, centers and 20 PCs, without and with BMI, within gender. The residuals from a stepwise regression covariate adjustments were standardized (mean zero, SD one) and used as the final phenotype in the linear mixed effects model. The linear mixed effects model was implemented, on an adjusted in advance phenotype for important covariates, in association with SNPs additive genetic fixed effects, using a kinship model to correct for random effects of familial relationship. The kinship matrix was built with “lmekin” and “kinship” R functions [24–25]. The association implemented was single SNP at a time in parallel servers with Linux OS and R version 2.14.1. GWAS in the LLFS was performed using all the assayed and imputed SNPs (n = 9.25 M).

Association tests in the ARIC and HABC. An additive genetic dosage model was assumed in both studies. In the ARIC Study, association tests were performed using the ProbABLE

maximum likelihood regression approach with age, sex, center, without and with BMI as covariates. In the Health ABC Study, analyses of replication were carried out using R v2.14.2 LM procedure with baseline covariates of age, sex, study center, without and with BMI, as well as the first two PCs as a means of controlling for population substructure.

3. Results

3.1 Sample characteristics

In the LLFS, after 328 subjects with clinical diagnosis of diabetes or diabetes treatment and 104 undiagnosed diabetes cases (fasting glucose ≥ 126 mg/dl or HbA_{1c} ≥ 6.5%) were excluded, this analysis included a total of 4,088 family members (1,804 men and 2,284 women) with complete phenotypic and genotypic information (Table 1). Similar exclusions were applied in the replication cohorts. Characteristics of the ARIC (n = 6,777) and HABC (n = 1,454) were also given in Table 1. While significant mean differences in HbA_{1c} were observed across studies, they were non-significant between sexes (Table 1).

3.2. Discovery in LLFS and replication in ARIC (in MAGIC) and HABC

The heritability estimate for HbA1c was 41.6% (standard error = 3.7%). Lambda estimate for GWAS of HbA_{1c} in this analysis was 1.03. Two common (MAF > 1%) SNPs at *GCK-YKT6* (rs730497, rs2908282) and one common SNP at *HK1* (rs17476364) were significantly (p < 5e–8) associated with HbA_{1c} in the LLFS (Table 2, Fig. 1A). When HbA_{1c} was further adjusted for BMI, the association at *HK* (rs17476364 with also rs72805692 and rs10159477, r^2 = 0.6–0.9) remained significant (p = 2e–11 – 3e–8), whereas the association strength at *GCK-YKT6* (rs730497 and rs2908282, r^2 = 1) was slightly attenuated (p = 5e–7 – 6e–7; Fig. 1B). The three associated SNPs or their proxies at *GCK-YKT6* and *HK1* (2 loci) were replicated (p < 0.0019, Bonferroni correction of type 1 error rate of 0.05 for a total of 27 independent tests/loci sought for replication, see below) in the ARIC (p = 1e–13 – 4e–9) but not in the HABC (p = 0.72–0.97; Table 2).

We also observed 25 suggestive (5e–8 < p < 1e–5) common SNPs for HbA_{1c} (without BMI adjustment) in the LLFS (Table 3). With recent availability of the MAGIC GWAS of HbA_{1c} data for download (ftp://ftp.sanger.ac.uk/pub/magic/), we queried all the suggestive signals in both the ARIC and MAGIC Consortium (substituting the HABC for power concern). We obtained replication information for 13 of the 25 SNPs (Table 3); the remaining 12 SNPs (or their proxies) were not available in the replication data (see the listed SNPs in Table 3 footnote only). *OR10R2-/SPTA1*-rs12041363 was replicated in the MAGIC (p = 5.1e–5; borderline in the ARIC, p = 3.7e–3; the less frequent 'C' allele was consistently associated with an increase in HbA_{1c} levels across all studies; Table 3). Additionally, *G6PC2* rs560887 was replicated in both ARIC and MAGIC (p = 1.5e–7 and 1e–17, respectively; the less frequent 'A' or 'T' allele was consistently associated with a decrease in HbA_{1c} levels across all studies; Table 3). The remaining suggestive signals were either not replicated (p > 0.0019) or unavailable for pursuing replication.

In the LLFS, the two significant loci accounted for 0.64% phenotypic variation in HbA_{1c}, the two suggestive loci (*OR10R2/SPTA1, G6PC2*) accounted for 0.47% phenotypic variation in HbA_{1c}, and the four loci collectively accounted for 1.11% phenotypic variation in HbA_{1c}. Because resulting association signals were similar for HbA_{1c} without versus with BMI adjustment (Fig. 1, Table 2), further data analyses were only performed for HbA_{1c} without BMI adjustment (hereinafter referred to as HbA_{1c}).

4. Discussion

The heritability for HbA_{1c} among non-diabetic subjects in the LLFS (42%) was right at the lower end of previously reported estimates (47 – 59%). In general, it is moderate and consistent with findings from other studies including the Framingham Offspring Study [8–13]. Our GWAS of HbA_{1c} values in non-diabetic subjects yielded two known loci at *GCK-YKT6* and *HK1* (same loci but different SNPs) [14–16] along with two suggestive but replicated loci in *G6PC2* and near *OR10R3P*. The association variant (rs560887) in *G6PC2* also was a known locus according to a MAGIC report. [16] And the association variant (rs12041363) near *OR10R3P* was about 110 kb away from a previously reported variant (rs2779116, $r^2 = 0.7$, $D' = 0.9$) near *SPTA1* in the MAGIC for HbA_{1c} [16]. The two SNPs may be correlated pointing to a same causal variant; however, since perfect linkage disequilibrium was not observed between them, they may also simply point to independent causal variants. A nearby variant (rs2142672, 8 kb away from rs12041363) was reportedly associated with LDL cholesterol in eight study populations [26]. Previously, significant and positive correlation between HbA_{1c} and LDL cholesterol was observed [27]; and it was also known that HbA_{1c} was associated with dyslipidemia and atherogenicity [28].

There are two lines of evidence that would indicate modulation of HbA_{1c} levels via the newly identified loci through non-glycemic biological pathways. First, our parallel GWAS of fasting glucose in the LLFS did not yield significant results at these three loci (data not shown). Second, association tests at the three loci after further adjusting for the effect of fasting glucose resulted in unchanged or even strengthened associations (data not shown). Our data also suggested that HbA_{1c} levels are influenced by *GCK* locus mainly through glycemic pathways, whereas the *HK1* locus likely operates through both glycemic and non-glycemic pathways, consistent with a previous observation [16]. We further examined all currently known pathways and processes [20] but did not find significant evidence of enriched gene sets modulating HbA_{1c} levels (data not shown).

Individual GWAS separately by sex was not pursued in this report for the following considerations. No mean differences in the LLFS, ARIC and HABC were observed (Table 1). In the LLFS, male and female participants were expected as non-independent because of family membership. Detection power would be reduced appreciably with variants of heritability or effect size < 2% likely missed, if not unanimously missed. Aiming to identify additional common loci for HbA_{1c} in the overall data, we alternatively performed conditional GWAS with HbA_{1c} pre-adjusted for known loci [14–16], but no new loci emerged (data not shown).

The LLFS data was ascertained through participants with familial clustering of exceptional longevity; some revealed loci for HbA_{1c} might be relevant to healthy aging as well. Nowadays, there is an increasing interest in identifying genetic variants that influence aging and longevity. Hyperglycemia and accumulation of damaging advanced glycation end-products (AGEs) are thought to play an important role in aging and neuro-degeneration [30–31]. It is possible that some of the genetic loci or variants influencing variation in HbA_{1c} levels in the present study might also contribute to long-term health effects of glycemia or glycation processes. In the future, assessments of directly measured toxic AGEs levels may hold profuse promise for understanding genetic contributions to aging and longevity.

As with no exceptions, this study also is not immune to limitations. Sample heterogeneity and thus genetic heterogeneity may exist across our discovery and replication cohorts as demonstrated significant mean differences in HbA_{1c} along with key covariates (see Table 1). Non-steroidal anti-inflammatory drugs (NSAIDs) are known to be influential on insulin and blood sugar levels. In this analysis, medication information in using NSAIDs was self-

reported, and incomplete, and was thus not used as a criterion for excluding subjects in the LLFS. Moreover, the power in the HABC to replicate significant GWAS association signals discovered in the LLFS may not be sufficient. Lastly, it is likely arbitrary to use *p* cutoff to discover suggestive signals for replication; and conclusions may also vary. For instance, if $p < 1e–6$ was instead used, then we would only observe two suggestive signals, with *OR10R2*-locus replicated ($p < 0.0125$ correcting for 4 tests, relative to no replication at $p < 0.0019$ correcting for 27 tests in using $p < 1e–5$ cutoff), *TOX2* locus non-replicated, and *G6PC2* (a known locus) missed in the ARIC. All the suggestive loci, where genes of interest may reside, did not attain stringent GWA significance criterion in the LLFS, possibility of false positive(s) cannot be ruled out. Given all these discussions, cautions may be used in interpreting and generalizing the findings in this report.

In conclusion, our GWAS has reconfirmed two known loci at *GCK* and *HK1*, and yielded 25 suggestive loci including two replicated loci in *G6PC2* and near *OR10R3P or SPTA1*. Further validations of the remaining suggestive loci from independent studies are warranted. Systemic in-depth survey of all sequence elements of the reconfirmed genes/loci mainly functioning and regulatory variants may yield additional new findings.

Acknowledgments

The investigators thank all the LLFS participants and staff for their valuable contributions. We are grateful to Heidi Dubrouillet for her administrative help and effort. The authors thank Lingyun Chang, M.S., Yanan Duan, M.S., Shiow Jiuan Lin, M.S. and Lihua Wang, M.D., M.S. for their assistance in statistical data analysis for this manuscript. Finally, we thank the two anonymous reviewers for their constructive comments that helped to strengthen this manuscript.

Funding Sources

The Long Life Family Study (LLFS): The Long Life Family Study was supported by the National Institute on Aging (NIA) grants U01AG023712, U01AG023744, U01AG023746, U01AG023749, U01AG023755, U19AG023122, K24AG025727, R01AG032319, the Glenn Medical Research Foundation, and the National Heart Lung Blood Institute (NHLBI, R21HL114237). The content is solely the responsibility of the authors, and does not necessarily represent the official views of the NIA or the National Institutes of Health (NIH). The Atherosclerosis Risk in Communities (ARIC) Study: The Atherosclerosis Risk in Communities Study was carried out as a collaborative study supported by the NHLBI contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL59367 and R01HL086694; the National Human Genome Research Institute contract U01HG004402; and the NIH contract HHSN268200625226C. The authors thank the staff and participants of the ARIC Study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the NIH and NIH Roadmap for Medical Research. This research was supported in part by NIH/NIDDK grants R21DK080294 and K01DK076595 (E.S.) and K01067207 (W.H.L.K.). The Health, Aging, and Body Composition (Health ABC) Study: The Health, Aging, and Body Composition Study was supported by the NIA contracts N01AG62101, N01AG62103, and N01AG62106, and in part by the Intramural Research Program of the NIH, NIA. The genome-wide association study was funded by the NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the NIH to the Johns Hopkins University, contract number HHSN268200782096C.

Abbreviations

ARIC	The Atherosclerosis Risk in Communities Study
BMI	Body mass index
GWAS	Genome-wide association study
HABC	The Health, Aging, and Body Composition Study
HbA1c	Glycated hemoglobin

LLFS The Long Life Family Study

REFERENCES

1. O'Sullivan JB, Mahan CM. Mortality related to diabetes and blood glucose levels in a community study. Am J Epidemiol. 1982; 116:678–684. [PubMed: 7137154]
2. Asadollahi K, Beeching N, Gill G. Hyperglycaemia and mortality. J R Soc Med. 2007; 100:503–507. [PubMed: 18048707]
3. DeGroot J. The AGE of the matrix: chemistry, consequence and curve. Current Opinion in Pharmacology. 2004; 4:301–305. [PubMed: 15140424]
4. International Expert Committee. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care. 2009; 32:1327–1334. [PubMed: 19502545]
5. American Diabetes Association. Executive summary: standards of medical care in diabetes – 2009. Diabetes Care. 2009; 32(Suppl 1):S6–S12. [PubMed: 19118288]
6. American Diabetes Association. Executive summary: standards of medical care in diabetes 2010. Diabetes Care. 2010; 33(Suppl 1):S4–S10. [PubMed: 20042774]
7. Hare MJ, Shaw JE, Zimmet PZ. Current controversies in the use of haemoglobin A_{1c}. J Intern Med. 2012; 271:227–236. [PubMed: 22333004]
8. Meigs JB, Panhuysen CI, Myers RH, et al. A genome-wide scan for loci linked to plasma levels of glucose and HbA_{1c} in a community-based sample of Caucasian pedigree: the Framingham Offspring Study. Diabetes. 2002; 51:833–840. [PubMed: 11872688]
9. An P, Freedman BI, Hanis CL, et al. Genome-wide linkage scans for fasting glucose, insulin, and insulin resistance in the National Heart, Lung, and Blood Institute Family Blood Pressure Program. Evidence of linkages to chromosome 7q36 and 19q13 from meta-analysis. Diabetes. 2005; 54:909–914. [PubMed: 15734873]
10. Pilia G, Chen WM, Scuteri A, et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. PLoS Genet. 2006; 2:e132. [PubMed: 16934002]
11. Soranzo N. Genetic determinants of variability in glycated hemoglobin (HbA_{1c}) in humans: review of recent progresses and prospects for use in diabetes care. Curr Diab Rep. 2011; 11:562–569. [PubMed: 21975967]
12. Florez J. A genome-wide association study of treated A_{1c}: A genetic needle in an environmental haystack? Diabetes. 2010; 59:332–334. [PubMed: 20103712]
13. Barker A, Langenberg C, Wareham NJ. Genetic determinants of glucose homeostasis. Best Practice & Research Clinical Endocrinology & Metabolism. 2012; 26:159–170. [PubMed: 22498246]
14. 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 Genetics. 2008; 4:e1000312. [PubMed: 19096518]
15. Paterson AD, Waggott D, Boright AP, et al. A genome-wide association study identifies a novel major locus for glycemic control in type 1 diabetes, as measured by both A_{1c} and glucose. Diabetes. 2010; 59:539–549. [PubMed: 19875614]
16. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A1c levels via glycemic and nonglycemic pathways. Diabetes. 2010; 50:3229–3239. [PubMed: 20858683]
17. Sebastiani P, Hadley EC, Province MA, et al. A family longevity selection score (FLoSS): ranking sibships by their longevity and availability for study. Am J Epidemiol. 2009; 170:1555–1562. [PubMed: 19910380]
18. Newman AB, Glynn NW, Taylor CA, et al. Health and function of participants in the Long Life Family Study: A comparison with other cohorts. Aging. 2011; 3:63–76. [PubMed: 21258136]
19. ARIC Investigators. The Artherosclerosis Risk in Communities (ARIC) Study: design and objectives. Am J Epidemiol. 1989; 129:687–702. [PubMed: 2646917]

20. Newman AB, Haggerty CL, Goodpaster B, et al. Strength and muscle quality in a well-functioning cohort of older adults: the Health, Aging and Body Composition Study. J Am Geriatr Soc. 2003; 51:323–330. [PubMed: 12588575]
21. Price AL, Patterson NJ, Plenge RM, et al. Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics. 2006; 38:904–909. [PubMed: 16862161]
22. Li Y, Willer CJ, Sanna S, et al. Genotype imputation. Annu Rev Genomics Hum Genet. 2009; 10:387–406. [PubMed: 19715440]
23. Li Y, Willer CJ, Ding J, et al. MaCH/l using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol. 2010; 34:816–834. [PubMed: 21058334]
24. Pinheiro, JC.; Bates, DM. Mixed-effects models in S and S-Plus. Springer; 2000.
25. Cortinas Abrahautes J, Burzykowski T. A version of the EM algorithm for proportional hazards models with random effects. Lecture Notes of the ICB Seminars. 2002:15–20.
26. Waterworth DM, Ricketts SL, Song K, et al. Genetic variants influencing circulating lipid levels and risk of coronary artery disease. Arterioscler Thromb Vasc Biol. 2010; 30:2264–2276. [PubMed: 20864672]
27. VinodMahato R, Gyawali P, Raut PP, et al. Association between glycaemic control and serum lipid profile in type 2 diabetic patients: Glycated haemoglobin as a dual biomarker. Biomedical Research. 2011; 22:375–380.
28. Bodhe C, Jankar D, Bhutada T, et al. HbA_{1c}: Predictor of dyslipidemia and atherogenicity in diabetes mellitus. International Journal of Basic Medical Sciences and Pharmacy. 2012; 2:25–27.
29. Segre AV, DIAGRAM Consortium, MAGIC investigators. Groop L, Mootha VK, Daly MJ, Alshuler D. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. PLoS Genet. 2010; 6:e1001058. [PubMed: 20714348]
30. Nawale RB, Mourya VK, Bhise SB. Non-enzymatic glycation of proteins: a cause for complications in diabetes. Indian J Biochphys. 2006; 43:337–344.
31. Takeuchi M, Yamagishi S. TAGE (toxic AGES) hypothesis in various chronic diseases. Medical Hypotheses. 2004; 63:449–452. [PubMed: 15288366]

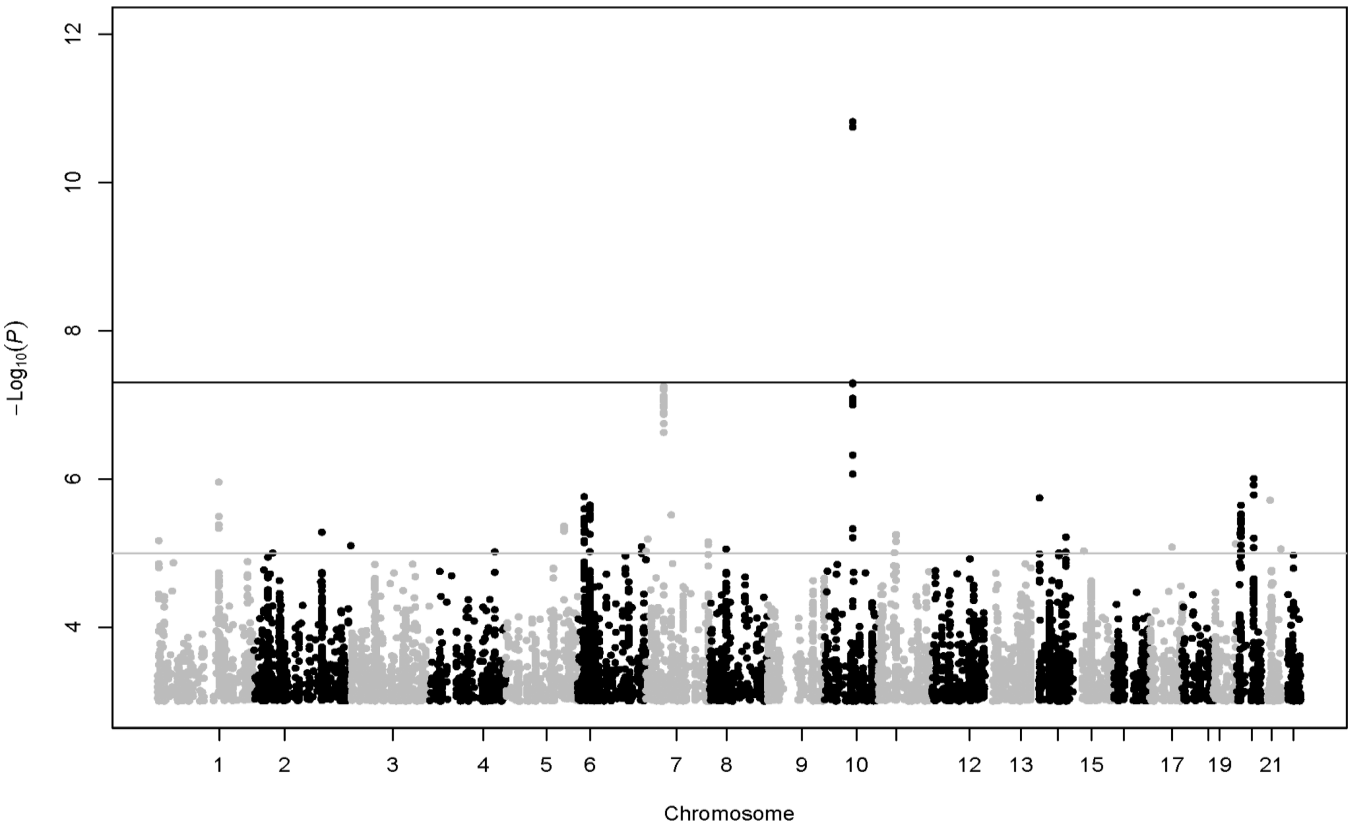

Fig. 1.
Manhattan plot of GWAS results for HbA_{1c} without BMI adjustment (filters, MAF < 0.01, Hardy-Weinberg p < 1e–6). Black horizontal reference line denotes $-\log_{10}(5e{-}8)$ for significant association criterion. Grey horizontal reference line denotes $-\log_{10}(1e{-}5)$ for suggestive association criterion.

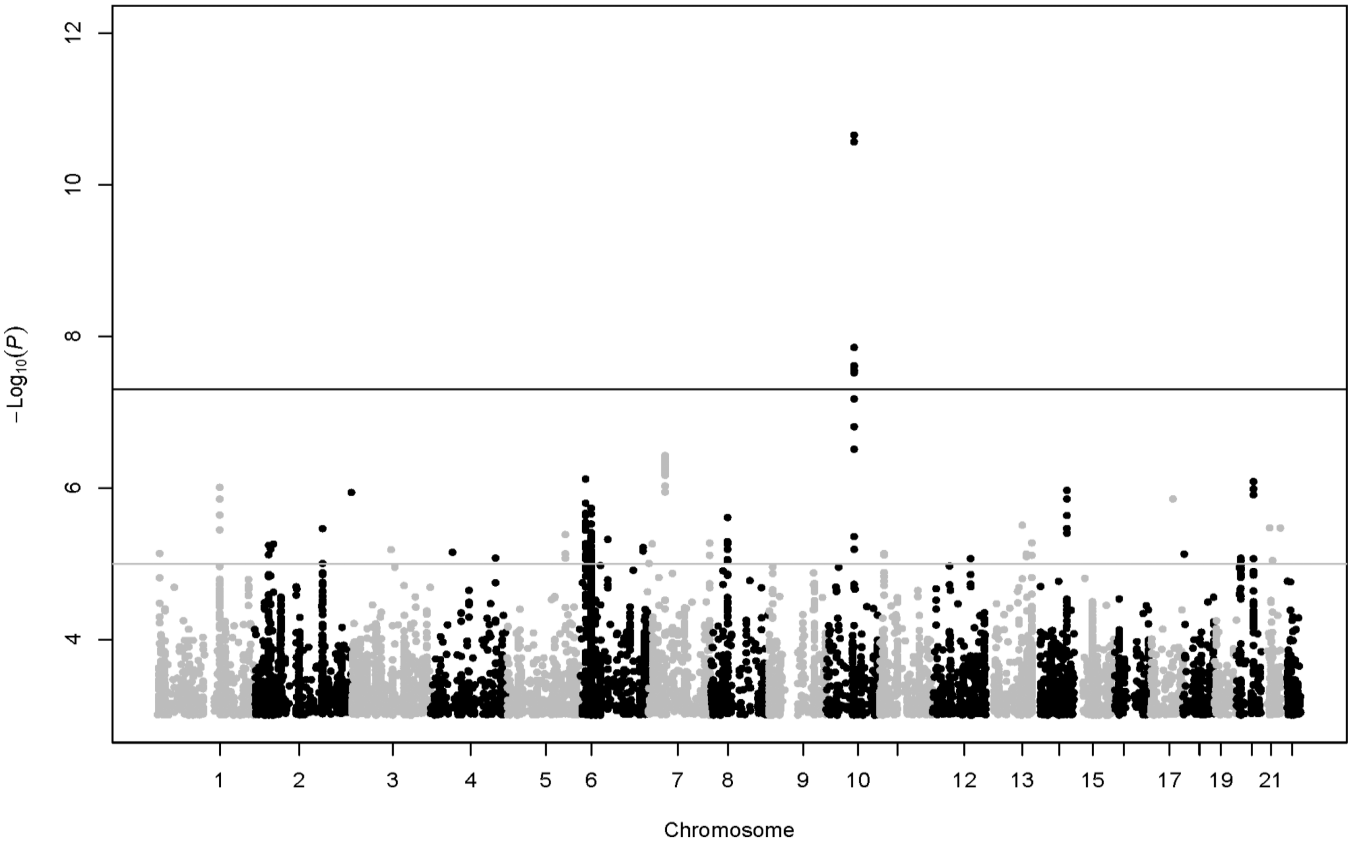

Fig. 2.
Manhattan plot of GWAS results for HbA_{1c} with BMI adjustment.

Metabolism. Author manuscript; available in PMC 2015 April 01.

Table 1

Sample characteristics of the LLFS, ARIC and HABC cohorts.

	Men			Women		
Sample	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
LLFS	**1804**			**2284**		
Age (years)	1804	70.2[†,‡]	15.3	2284	69.5[†,‡]	16.2
BMI (kg/m^2)	1732	27.3 [*,‡]	3.9	2185	26.4 [*]	5.2
FG (mg/dL)	1652	93.5 [*,†,‡]	11.2	2284	89.9 [*,†,‡]	10.9
HbA_{1c} (%)	1804	5.5[†,‡]	0.4	2284	5.5[†,‡]	0.3
HbA_{1c} (mmol/mol)		(37)			(37)	
ARIC	**3106**			**3671**		
Age (years)	3106	57.4[†,§]	5.7	3671	56.7[†,§]	5.6
BMI (kg/m^2)	3106	27.3[§]	3.9	3671	26.6[§]	5.3
FG (mg/dL)	3106	103.5[†,§]	9.0	3671	99.4[†,§]	9.0
HbA_{1c} (%)	3106	5.4[†,§]	0.4	3671	5.4[†,§]	0.4
HbA_{1c} (mmol/mol)		(36)			(36)	
HABC	**746**			**708**		
Age (years)	746	73.9[‡,§]	2.9	708	73.7[‡,§]	2.8
BMI (kg/m^2)	746	26.9[‡,§]	3.6	708	26.0[§]	4.4
FG (mg/dL)	746	99.1[‡,§]	20.6	708	91.7[‡,§]	11.9
HbA_{1c} (%)	746	6.0[‡,§]	0.7	708	5.9[‡,§]	0.5
HbA_{1c} (mmol/mol)		(42)			(41)	

[*] Significant ($p < 0.05$) mean differences between sexes in the LLFS.

[†] Significant ($p < 0.05$) mean differences (within sex) between the LLFS and ARIC.

[‡] Significant ($p < 0.05$) mean differences between the LLFS and HABC, within sex.

[§] Significant ($p < 0.05$) mean differences between the ARIC and HABC.

Table 2

Significant GWAS findings ($p < 5e\text{-}8$) for HbA_{1c} in the LLFS with replication information from the ARIC and HABC.

SNP	Chr	Position	Gene	LLFS AL*	LLFS EAF*	LLFS Beta	LLFS SE	LLFS r^2	LLFS *P*	ARIC AL*	ARIC EAF*	ARIC Beta	ARIC SE	ARIC *P*	HABC AL*	HABC EAF*	HABC Beta	HABC SE	HABC *P*
No BMI-adj.																			
rs730497	7	44223721	*GCK*	A/G	.17	.16	.03	.008	5.9e-8	A/G	.18	.05	.01	3.1e-9	G/A	.84	.01	.03	.81
rs2908282	7	44248828	*YKT6*	A/G	.17	.16	.03	.008	8.4e-9	A/G	.18	.05	.01	2.2e-9	G/A	.84	.003	.03	.93
rs17476364	10	71094504	*HK1*	T/C	.90	.27	.04	.011	1.8e-11	G/A ‡	.88	.07	.01	2.9e-13	T/C	.90	−.01	.04	.73
BMI-Adj.																			
rs730497 †	7	44223721	*GCK*	A/G	.17	.16	.03	.007	4.6e-7	A/G	.18	.05	.01	4.1e-9	G/A	.84	.003	.03	.92
rs2908282 †	7	44248828	*YKT6*	A/G	.17	.16	.03	.007	5.9e-7	A/G	.18	.05	.01	2.9e-9	G/A	.84	−.001	.03	.97
rs17476364	10	71094504	*HK1*	T/C	.90	.28	.04	.012	2.7e-11	G/A ‡	.88	.07	.01	9.6e-14	T/C	.90	−.01	.04	.78
rs72805692	10	71099109	*HK1*	A/G	.89	.25	.04	.013	2.2e-11	G/A ‡	.88	.07	.01	9.6e-14	A/G	.89	−.01	.04	.72
rs10159477	10	71099888	*HK1*	A/G	.16	−.18	.03	.009	3.1e-8	G/A	.88	.07	.01	9.6e-14	G/A	.85	−.01	.03	.78

*AL, effect allele / non-effect allele; EAF, effect allele frequency.

†SNPs significant for HbA_{1c} without BMI adjustment but non-significant for HbA_{1c} with BMI adjustment.

‡Same SNP rs10159477 used as best proxy for rs17476364 and rs72805692 ($r^2 = 0.71$) in the ARIC.

Table 3

Suggestive signals (without BMI adjustment, 5e-8 < p < 1e-5) for HbA_{1c} in LLFS with replication data from ARIC / MAGIC

SNP	Chr	Position	Gene	LLFS AL*	LLFS EAF*	LLFS Beta	LLFS SE	LLFS r^2(%)	LLFS P	ARIC AL*	ARIC EAF*	ARIC Beta	ARIC SE	ARIC P	MAGIC AL*	MAGIC EAF*	MAGIC Beta	MAGIC SE	MAGIC P
rs56305036	1	8283288	*SLC45A1*-	C/A	.95	−.23	.05	.4	6.9e-6	G/C	.95	−.02	.01	.29†	C/G	.05	.004	.008	.59†
rs12041363	1	158475894	*OR10R2/SPTA1*-	C/T	.25	.13	.03	.6	1.1e-6	T/C	.76	−.02	.01	3.7e-3	T/C	.72	−.016	.004	5.1e-5
rs560887	2	169763148	*G6PC2*	A/G	.28	−.12	.03	.5	5.3e-6	T/C	.29	−.04	.01	1.5e-7	T/C	.33	−.032	.004	1e-17
rs72833452	6	16204870	*GMPR*-	C/T	.89	.19	.04	.6	1.8e-6	A/G	.82	.003	.01	.81†	A/G	.80	.004	.006	.49†
rs2516685	6	30361608	*RPP21*-	C/T	.41	−.11	.02	.6	2.3e-6	A/G	.59	.01	.01	.17	A/G	.53	.006	.004	.09
rs11754823	6	159777178	*FNDC1*-	A/C	.94	.23	.05	.5	8.2e-6	A/G	.95	.01	.02	.57	A/C	.96	.006	.018	.73
rs906290	8	41693768	*ANK1*	A/T	.36	.11	.02	.6	8.9e-6	T/A	.62	.004	.007	.52	A/T	.25	3e-4	.004	.95
rs17792599	14	21770681	*RPGRIP1*	G/A	.21	.14	.03	.7	1.8e-6	A/G	.81	−.01	.01	.30	A/G	.79	−.015	.005	2.8e-3
rs1638715	14	87518172	*LOC283585*-	G/T	.66	−.11	.02	.5	6.1e-6	G/A	.61	−.01	.01	.28†	A/G	.42	−.003	.004	.44†
rs9303394	17	55610343	*MSI2*	G/A	.22	.13	.03	.7	8.3e-6	A/G	.77	−.003	.007	.70	A/G	.78	−.005	.004	.26
rs6133998	20	10681095	*JAG1*-	T/C	.12	.17	.04	.8	2.3e-6	C/T	.89	.02	.01	.017	T/C	.12	−.009	.005	.11
rs6130482	20	42557999	*TOX2*	T/C	.87	−.17	.03	.8	9.9e-7	T/C	.85	.005	.01	.57	T/C	.90	.001	.006	.86
rs12172913	21	21208597	*LOC100505973*-	A/G	.57	−.12	.03	.7	2.0e-6	A/G	.62	.004	.007	.59	A/G	.67	−.002	.004	.52

*AL, effect/non-effect alleles; EAF, effect allele frequency.

†Proxy SNPs (rs17032305 for rs56305036, r^2 = 0.89; rs7750030 for rs72833452, r^2 = 0.90; rs1769465 for rs1638715, r^2 = 1) were used in the ARIC and MAGIC.

Twelve additional suggestive signals were not listed in Table 3 that were either not found in the replication data or their proxy SNPs unavailable in using the SNAP Proxy Search (http://www.broadinstitute.org/mpg/snap/index.php). In specific, these included *SNED1* rs139853980, *FSTL5* rs144444769, *PPP2R2B*- rs192137470, *FAM20C* rs28614471, *ZNF733P*- rs5002035, *SHH*- rs56157317, *LRRC4C* rs75879627, *TSPAN18*- rs7111558, *GABRG3* rs73365231, *SSC5D* rs143559825, *PLCB1*- rs6118391, and *DIP2A* rs80186361.