

Association Analysis of European-Derived Type 2 Diabetes SNPs from Whole Genome Association Studies in African Americans

Joshua P. Lewis^{1,2}, Nicholette D. Palmer^{1,3}, Pamela J. Hicks^{1,3}, Michele M. Sale^{4,5}, Carl D. Langefeld⁶, Barry I. Freedman⁷, Jasmin Divers⁶, Donald W. Bowden^{1,3,7}.

¹Center for Human Genomics, ²Program in Molecular Genetics and Genomics, ³Department of Biochemistry, Wake Forest University School of Medicine, Winston-Salem, NC; ⁴Department of Medicine, ⁵Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA; ⁶Division of Public Health Sciences, ⁷Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, NC.

Please address correspondence to:
Donald W. Bowden, Ph.D.
Center for Human Genomics
Wake Forest University School of Medicine
Medical Center Blvd
Winston-Salem, NC 27157
Email: dbowden@wfubmc.edu

Received for publication 14 September 2007 and accepted in revised form 21 April 2008

Additional information for this article can be found in an online appendix at
<http://diabetes.diabetesjournals.org>

Objective: Several whole genome association (WGA) studies have reported identification of type 2 diabetes mellitus (T2DM) susceptibility genes in various European-derived populations. Little investigation of these loci has been reported in other ethnic groups, specifically African Americans (AAs). Striking differences exist between these populations, suggesting they may not share identical genetic risk factors. Our objective was to examine the influence of T2DM genes identified in WGA studies in a large AA case-control population.

Research Design and Methods: SNPs in 12 loci (e.g. *TCF7L2*, *IDE/KIF11/HHEX*, *SLC30A8*, *CDKAL1*, *PKN2*, *IGF2BP2*, *FLJ39370*, and *EXT2/ALX4*) associated with T2DM in European-derived populations were genotyped in 993 T2DM AA cases and 1054 AA controls. Additionally, 68 ancestry-informative markers (AIMs) were genotyped to account for the impact of admixture on association results.

Results: Except for *TCF7L2*, little evidence of association was observed between SNPs and T2DM in AAs. One *TCF7L2* SNP (rs7903146) showed compelling evidence of association with T2DM (admixture adjusted $P_a=1.59 \times 10^{-6}$). Only the intragenic SNP on 11p12 (rs9300039, $P_d=0.029$) was also associated with T2DM after admixture adjustments. Interestingly, 4 of the SNPs are monomorphic in the Yoruba population of the HAPMAP project with only the “risk” allele from the populations of European descent present.

Conclusions: Results suggest these variants do not significantly contribute to inter-individual susceptibility to T2DM in AAs. Consequently, genes contributing to T2DM in AAs may, in part, be different from those in European-derived populations. High frequency of risk alleles in several of these genes may, however, contribute to the increased prevalence of T2DM in AAs.

Recently, several whole genome association (WGA) studies have reported evidence for the existence of multiple type 2 diabetes mellitus (T2DM) susceptibility genes. These results were primarily observed in different European-derived populations. Included in these loci are polymorphisms in *TCF7L2* [1-5], *HHEX* [1-6], *SLC30A8* [1-4, 6], *CDKAL1* [2-6], *IGF2BP2* [3-6], *FTO* [4-7], *PKN2* [3, 4], *FLJ39370* [3], *EXT2/ALX4* [1], and *LOC387761* [1]. Despite the compelling evidence of association of many of these genes in European-derived populations little or no investigation has been reported in African-Americans (AAs).

Differences between European and African-derived populations (e.g. haplotype block structure and allele frequency discrepancies) suggest genetic risk factors may not be identical between them. Previously, we have reported that common genetic variants in *TCF7L2* and *HNF4A* contribute to T2DM in AAs while polymorphisms in *CAPN10*, *TCF1*, and *PPARG* have little or no evidence of association [8]. Lack of association with other T2DM susceptibility loci in AAs has also been observed [9]. In addition, differences in allele frequencies of risk polymorphisms between ethnicities have been shown (e.g. Chandalia, *et al.* [10]). These substantial discrepancies could contribute to the differences in overall prevalence of T2DM within these ethnic groups. Consequently, we examined the influence of SNPs recently identified in WGA studies on T2DM susceptibility in a large AA case-control population.

Research Design and Methods

Subjects

Recruitment of patients and controls has been previously described [11]. Briefly, 993 unrelated AA patients with T2DM born in

Virginia, Tennessee, North Carolina, South Carolina, or Georgia were recruited from dialysis clinics. Type 2 diabetes was diagnosed if patients reported an initial diagnosis of diabetes mellitus after 35 years of age, received dietary therapy or received hypoglycemic agents in the absence of insulin alone for at least one year after initial diagnosis, and were currently receiving diabetes medications. Cases had severe diabetes accompanied by nephropathy, and T2DM was diagnosed > 5 years prior to the start of renal replacement therapy, with background or greater diabetic retinopathy and/or > 3+ proteinuria on urinalysis. Non-diabetic controls were recruited from the same geographic region: 1054 were AA and 36 European-American (EA) without prior diagnosis of T2DM or renal disease. DNA extraction was performed using the PureGene system (Gentra Systems, Minneapolis, MN). DNA was obtained from 44 Yoruba Nigerians from the National Institute of General Medical Sciences Human Variation Collection (Coriell Cell Repositories, Camden, NJ). This study was approved by the Institutional Review Board of Wake Forest University School of Medicine and was in accordance with the principles of the Declaration of Helsinki.

Candidate gene genotyping

SNPs *PKN2* rs6698181; *IGF2BP2* rs4402960; *FLJ39370* rs17044137; *CDKAL1* rs10946398, rs7754840; *SLC30A8* rs13266634; *CDKN2A/CDKN2B* rs564398, rs10811661; *HHEX/IDE/KIF11* rs1111875, rs5015480, rs7923837; *EXT2/ALX4* rs3740878, rs11037909, rs1113132; *FTO* rs8050136; *LOC387761* rs7480010; intragenic rs9300039; *TCF7L2* rs7903146 were genotyped using iPLEX technology (Sequenom Inc., San Diego, CA) [12]. Primer sequences are available upon request. SNP genotyping success rates were >98.6% in cases and

>98.4% in controls. Concordance between blind duplicate samples included in the genotyping was >99.5%.

Admixture analyses genotyping

Sixty eight biallelic AIMs were genotyped using iPlex technology (Sequenom Inc., San Diego, CA) in 993 AA cases, 1054 AA controls, 44 Yoruba Nigerians and 36 EAs. Both the Yoruba Nigerian and EA individuals were utilized only for admixture analyses. Primer Sequences are available upon request. Genotyping success rates for AIMs were >96.9% in AA cases and >95.2% in AA controls.

Statistical analyses

Hardy Weinberg equilibrium (HWE) values were determined by calculating a χ^2 statistic and corresponding P-value. Structures of the haplotype blocks for the SNPs genotyped in *CDKAL1*, *EXT2/ALX4*, and *IDE/KIF11/HHEX* were ascertained using Haploview 3.2 [13] using the criteria outlined in Gabriel *et al.* [14].

Measures of linkage disequilibrium (LD) and association were calculated using the program SNP-GWA [15]. For each tandem pair of SNPs the LD statistics D' and r^2 were computed. Multiple tests of association including the overall 2-degree of freedom (genotypic), additive (Cochran-Armitage trend test), and corresponding lack of fit to additivity were calculated. Tests of association under the dominant and recessive genetic models are also reported. Haplotypic association was calculated using the expectation-maximization algorithm implemented in this program. Initially, tests were performed using 1,000 permutations. Where association tests indicated possible significance (empirical P-value<0.10), permutations were increased to 10,000. The quantitative trait (BMI) association analysis for *FTO* SNP rs8050136 was performed using

a module of SNP-GWA (QSNP-GWA). Association values adjusted for age or sex were calculated using the program SNP-ADDMIX. Estimates of case and control haplotype frequency were obtained using Dandelion 1.26 [16].

To provide context for these tests of association power analyses were computed using QUANTO [17]. Specifically, power estimates to detect the range of reported odds ratios from the European-derived samples were computed assuming a T2DM population prevalence of 0.15 (consistent with AA populations) and polymorphism-specific type 1 error rates of $\alpha=0.05$ and $\alpha=0.10$. Power was calculated based on the risk allele frequency observed in our African American sample.

Individual ancestral proportions were calculated using an expectation maximization algorithm (FRAPPE) [18] under a two population model. Estimates of “ancestral” allele frequencies were obtained from the genotyped African and EA samples. Logistic regression tests of additive, dominant, and recessive genetic models included adjustments for individual estimates of African ancestry [19].

Results

Eighteen SNPs in 12 T2DM loci previously associated in European-derived populations were genotyped in a sample of 993 AA T2DM cases enriched for diabetic nephropathy and 1054 AA controls. Characteristics of the study population are shown in Table 1. Within this cohort we observed a greater percentage of females in the cases (63%) compared to the controls (55%). Controls were younger than cases but older than the mean age at which T2DM was diagnosed. The ages of 26% of the controls were unavailable. The mean age of ESRD diagnosis was 58.4 ± 10.5 years old. In the cases, the estimated mean proportion of

African ancestry was 0.821, while in the controls it was 0.795.

All SNPs were consistent with Hardy-Weinberg proportions in the entire population (Supplementary Table 1). rs9300039, which is located in an intragenic region of chromosome 11p12, was inconsistent with HWE in the case population ($P=0.028$) due to an excess of homozygotes. All SNPs examined were in HWE in the control population.

High levels of linkage disequilibrium were observed between the 2 *CDKALI* SNPs rs10946398 and rs7754840 ($D'=1.000$, $r^2=0.996$). The three SNPs within the *IDE/KIF11/HHEX* region were in high to moderate LD with pairwise D' values of 0.679 ($r^2=0.244$) between rs1111875 and rs5015480, $D'=0.945$ ($r^2=0.140$) between rs5015480 and rs7923837, and $D'=0.894$ ($r^2=0.236$) between rs1111875 and rs7923837. Likewise, high LD was also observed within the three SNPs in the *EXT2/ALX4* region with pairwise values of $D'=1.000$ ($r^2=0.569$) between rs1113132 and rs11037909, $D'=0.988$ ($r^2=0.601$) between rs11037909 and rs3740878, and $D'=1.000$ ($r^2=0.935$) between rs1113132 and rs3740878.

Results of the single SNP association analyses are shown in Table 2. Genotype frequencies and counts for each SNP are provided in Supplementary Table 1. Evidence of association with T2DM was observed with rs10946398 (additive P [P_a] =0.029, OR 1.15) and rs7754840 ($P_a=0.039$, OR 1.14), both of which are located in intron 5 of *CDKALI*. However, after adjusting for admixture neither SNP remained nominally significant. A significant association with T2DM was also seen with LOC387761 SNP rs7480010 ($P_a=0.002$, OR 1.33) but this marker also did not show evidence of association after adjusting for admixture ($P_a=0.084$). Rs9300039, located in an intragenic region of chromosome 11p12, was

also associated under the admixture adjusted dominant model ($P_a=0.029$) but deviated from Hardy Weinberg proportions in the case population (HWE $P=0.028$). In addition, the minor allele homozygote count for rs9300039 is small (21 in the cases and 9 in the controls, Supplementary Table 1), so association values for this marker should be interpreted with caution. The most significant association with T2DM was observed with *TCF7L2* SNP rs7903146 ($P_a=1.73 \times 10^{-6}$, OR 1.39) and this SNP remained statistically significant after admixture adjustment (admixture-adjusted $P_a=1.59 \times 10^{-6}$). All other SNPs examined failed to show evidence of association with T2DM in this AA cohort under an additive model (Table 2). Tests of association under the dominant and recessive genetic models are shown in Supplementary Table 2. We also tested whether *FTO* SNP rs8050136 was associated with BMI as previously reported in European-derived populations [4, 6]. We found no evidence of association with this polymorphism and BMI ($P_a=0.480$). Association with T2DM adjusted for age or sex was calculated for each SNP (data not shown), but did not meaningfully change the results of this study.

It is interesting to note that 5 of the 18 SNPs had odds ratios in the opposite direction and inconsistent with the previously reported odds ratio. In fact, the 95% confidence interval for 7 of the 18 SNPs did not include the previously reported odds ratio from the EA samples.

With this very limited evidence of association in our AA population, the question arises whether adequate power was available in this sample to detect association. This has been addressed by calculating the power to detect association for each SNP based on the estimated effect size from studies in European-derived populations and the observed allele frequency of the risk allele in the African American population. Power was calculated for $\alpha=0.05$ (nominal evidence

of association) and $\alpha=0.10$ (for the ability to detect loci that are trending towards association). The power estimates using an intermediate measure of effect size (between the highest and lowest reported in European-derived populations) under the additive model are shown in Table 2. More extensive power estimates are shown in Supplementary Table 3 which provides estimates for high and low observed effect sizes also. For some SNPs the power is quite low, e.g. rs7480010, but for other SNPs rs3740878, rs11037909, rs8050136 power is over 70% for $\alpha=0.05$.

All three SNPs examined in the *EXT2/ALX4* region reside in a single LD block but showed no haplotypic association with T2DM (data not shown). Of the three SNPs genotyped in the *IDE/KIF11/HHEX* region two (rs5015480 and rs7923837) were located in an LD block but also showed no evidence of association. Haplotype analysis of the two *CDKALI* SNPs (rs10946398 and rs7754840) showed significant association with T2DM (global empirical $P=0.036$; data not shown) but provides little information beyond that obtained with single SNP analysis due to high measures of linkage disequilibrium ($D'=1.000$, $r^2=0.996$).

Discussion

We examined eighteen SNPs in twelve loci identified in recent genome-wide association studies in European-derived populations that were associated with T2DM or BMI (*FTO*) for association with T2DM in a large AA population. There is very limited evidence that any of these European-derived “GWAS SNPs” contribute to diabetes in AAs. In contrast, significant association with T2DM was observed with the previously described *TCF7L2* SNP rs7903146. *TCF7L2* has been the focus of multiple investigations and is one of the most highly replicated T2DM candidate genes in several populations [1-5]. Few publications have examined the influence of *TCF7L2* on T2DM risk in AAs. In this study

we expand on the previous investigation from our center by Sale, *et al.* which described compelling evidence of association with T2DM for polymorphisms in this gene [8]. From the results of these studies, as well as the convincing replication of association with T2DM in European-derived populations, a more comprehensive investigation of *TCF7L2* polymorphisms in AAs is warranted. Parenthetically, this observation of strong association gives us confidence that the ascertainment and diagnoses used to recruit these AA subjects are accurate and the analyses of the European-derived T2DM loci presented here are valid.

Of all the other loci examined, the only nominally significant association with T2DM in AAs before admixture adjustment was observed with the SNPs located in *CDKALI* and LOC387761. If one takes into consideration that multiple SNPs were genotyped, any standard multiple comparisons correction suggests the evidence for association is very modest. Both rs10946398 and rs7754840 have been previously associated with T2DM in European-derived populations [3, 4, 6]. In addition, other publications have also reported evidence of association in this region making *CDKALI* one of the most highly replicated genes identified from recent WGA studies [2, 5]. Measures of linkage disequilibrium between rs10946398 and rs7754840 in AAs ($D'=1.000$, $r^2=0.996$) were similar to those reported in Europeans [6]. In contrast, the frequencies of the rs10946398 and rs7754840 risk allele (C for both) differed substantially between AA (0.60, 0.60), YRI (0.67, 0.67), and CEU (0.31, 0.31) populations. In addition, *FTO* SNP rs8050136 was previously associated with BMI [7], however we found no evidence of association with this trait in any genetic model.

The genotyping data have been evaluated for association with T2DM taking into consideration the European-African

admixture of African Americans. We have adjusted association statistics using estimates of individual proportions of European and African ancestry for each subject. This approach may under- or over-estimate the true association at each locus since the difference in allele frequencies between the ancestral populations might vary at each marker, and thus the adjustment for an individual genome may overestimate or underestimate admixture at a specific site. Thus it is possible in this dataset that stronger evidence of association might be revealed by a locus specific analysis. Given the magnitude of the great majority of the P-values observed here (Table 2), this influence would have to be very substantial to adjust the P-values into the range suggesting association.

Even though all of the SNPs examined in this study showed highly significant association with T2DM in European-derived populations, we found limited evidence of association with these markers in African Americans. For some of these SNPs we have limited power to detect evidence of association given the sample sizes, allele frequencies, and effect sizes (Table 2, Supplementary Table 3). There is, however, reasonable power to detect association for a number of the SNPs. It is important to note that these conventional power analyses reflect the power to detect association at a single locus. If one takes into consideration for the overall study, we would have expected to see evidence of association with 4-5 loci at $\alpha=0.05$.

Another caveat to this study is that we have tested only 1-3 SNPs that were highly associated in European-derived populations per gene. Ongoing extensive analyses of individual genes, a pending genome-wide association study in these samples, and genotyping in additional samples of cases and controls will further clarify the role of these loci in African American diabetes. The

absence of association which was observed in this study may, however, be accountable to differences in genetic risk factors that exist between African- and European-derived populations. Moreover, multiple other T2DM susceptibility genes reported in populations with European descent fail to show association in African-derived samples [9].

Disparities in allele frequencies between different ethnic groups may impair our ability to observe true associations if they exist in African Americans. For example, the T2DM risk allele reported by recent WGA studies for *CDKN2A/CDKN2B* SNPs rs10811661 and rs564398, *IDE/KIF11/HHEX* SNP rs7923837, and *LOC387761* SNP rs7480010 have frequencies of 1.00, i.e. are monomorphic, in the YRI population of the HAPMAP project [20] (Table 3). This observation presumes that Yoruba are representative of all African populations. We have compared these sequences to chimpanzee and other available primate sequences, but no common pattern of sequence variation is apparent. Presumably the non-African alleles for these loci are the result of European admixture, which is not great enough to provide an observable level of protection from T2DM. The observation that these loci are nonpolymorphic in Africans and that Africans solely have the risk allele, suggests that these loci may, in part, contribute to the increased overall prevalence of T2DM in African-derived populations compared to Europeans. Thus the appropriate conclusion from this study is that these European T2DM susceptibility loci do not measurably contribute to differential susceptibility in the African American population.

Acknowledgements

This work was supported by grants R01 DK066358, R01 DK53591, R01 DK70941 and the Wake Forest University General Clinical Research Center M01 RR07122

References:

1. Sladek, R., et al., *A genome-wide association study identifies novel risk loci for type 2 diabetes*. *Nature*, 2007. **445**(7130): p. 881-5.
2. Steinthorsdottir, V., et al., *A variant in CDKAL1 influences insulin response and risk of type 2 diabetes*. *Nat Genet*, 2007. **39**(6): p. 770-5.
3. Saxena, R., et al., *Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels*. *Science*, 2007. **316**(5829): p. 1331-6.
4. Scott, L.J., et al., *A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants*. *Science*, 2007. **316**(5829): p. 1341-5.
5. *Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls*. *Nature*, 2007. **447**(7145): p. 661-78.
6. Zeggini, E., et al., *Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes*. *Science*, 2007. **316**(5829): p. 1336-41.
7. Frayling, T.M., et al., *A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity*. *Science*, 2007. **316**(5826): p. 889-94.
8. Sale, M.M., et al., *Variants of the Transcription Factor 7-Like 2 (TCF7L2) Gene are Associated with Type 2 Diabetes in an African American Population Enriched for Nephropathy*. Submitted.
9. Elbein, S.C., *Evaluation of polymorphisms known to contribute to risk for diabetes in African and African-American populations*. *Curr Opin Clin Nutr Metab Care*, 2007. **10**(4): p. 415-9.
10. Chandalia, M., et al., *Ethnic differences in the frequency of ENPP1/PC1 121Q genetic variant in the Dallas Heart Study cohort*. *J Diabetes Complications*, 2007. **21**(3): p. 143-8.
11. Freedman, B.I., et al., *Genetic linkage analysis of growth factor loci and end-stage renal disease in African Americans*. *Kidney Int*, 1997. **51**(3): p. 819-25.
12. Buetow, K.H., et al., *High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry*. *Proc Natl Acad Sci U S A*, 2001. **98**(2): p. 581-4.
13. Barrett, J.C., et al., *Haploview: analysis and visualization of LD and haplotype maps*. *Bioinformatics*, 2005. **21**(2): p. 263-5.
14. Gabriel, S.B., et al., *The structure of haplotype blocks in the human genome*. *Science*, 2002. **296**(5576): p. 2225-9.
15. Matarin, M., et al., *A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release*. *Lancet Neurol*, 2007. **6**(5): p. 414-20.
16. Green, L., Lange, EM, Langefeld, CD, *Power comparison for phase-known versus phase-unknown haplotype analyses for case-control designs*. *Am J Hum Genet*, 2001. **69**: p. 1948a.
17. Gauderman, W.J., *Sample size requirements for matched case-control studies of gene-environment interaction*. *Stat Med*, 2002. **21**(1): p. 35-50.
18. Tang, H., et al., *Estimation of individual admixture: analytical and study design considerations*. *Genet Epidemiol*, 2005. **28**(4): p. 289-301.

19. Redden, D.T., et al., *Regional admixture mapping and structured association testing: conceptual unification and an extensible general linear model*. PLoS Genet, 2006. **2**(8): p. e137.
20. *A haplotype map of the human genome*. Nature, 2005. **437**(7063): p. 1299-320.

Table 1: Characteristics of African American Study Participant

Trait	T2DM-ESRD Cases		Controls	
	N*	Mean \pm Std. Deviation or %	N*	Mean \pm Std. Deviation or %
Female %(N)	993	63% (622)	1054	55% (582)
Age at exam (years)	957	61.8 \pm 10.2	781	51.4 \pm 11.4
Age at T2DM diagnosis (years)	965	41.3 \pm 12.2	-	N/A
Age at ESRD diagnosis (years)	990	58.4 \pm 10.5	-	N/A

N* = number with data available

Table 2. Single-SNP tests of association with T2DM-ESRD

			African American Data					Reported European Data			Power to Detect Association in African Americans	
Gene	SNP	European Reported Risk Allele	Risk Allele Frequency Controls	Risk Allele Frequency Cases	Additive P-value	Admixture-Adjusted Additive P-value	Admixture-Adjusted OR (95% CI)	Reported Risk Allele Frequency Controls	Reported Risk Allele Frequency Cases	Reported OR (95% CI)	$\alpha=0.05$	$\alpha=0.10$
PKN2	rs6698181	T	0.153	0.156	0.829	0.388	1.08 (0.91-1.29)	0.290	0.320	1.11 (1.05-1.16)	0.237	0.345
IGF2BP2	rs4402960	T	0.525	0.528	0.865	0.803	0.98 (0.87-1.11)	0.304	0.341	1.18 (1.08-1.28)	0.555	0.675
FLJ39370	rs17044137	A	0.329	0.326	0.854	0.747	0.98 (0.86-1.12)	0.230	0.270	1.13 (1.06-1.19)	0.060	0.115
CDKAL1	rs10946398	C	0.582	0.615	0.029	0.110	1.11 (0.98-1.26)	0.319	0.361	1.16 (1.10-1.22)	0.427	0.522
CDKAL1	rs7754840	C	0.585	0.616	0.039	0.136	1.10 (0.97-1.25)	0.360	0.387	1.12 (1.03-1.22)	0.427	0.552
SLC30A8	rs13266634	C	0.914	0.916	0.861	0.543†	1.46 (0.43-4.89)	0.609	0.649	1.18 (1.09-1.29)	0.169	0.263
CDKN2B/CDKN2A	rs564398	T	0.934	0.943	0.196	0.320†	2.99 (0.34-25.98)	0.558	0.595	1.13 (1.08-1.19)	0.140	0.225
CDKN2B/CDKN2A	rs10811661	T	0.933	0.927	0.412	0.128†	0.18 (0.02-1.64)	0.850	0.872	1.20 (1.07-1.36)	0.304	0.422
IDE/KIF11/HHEX	rs1111875	C	0.766	0.774	0.547	0.767	1.02 (0.88-1.19)	0.522	0.546	1.10 (1.01-1.19)	0.371	0.495
IDE/KIF11/HHEX	rs5015480	C	0.633	0.621	0.412	0.400	0.95 (0.83-1.08)	0.425	0.379	1.13 (1.08-1.17)	0.470	0.595
IDE/KIF11/HHEX	rs7923837	G	0.917	0.929	0.143	0.303†	1.87 (0.57-6.12)	0.597	0.622	1.11 (1.02-1.20)	0.143	0.229
Intragenic	rs9300039	C	0.889	0.884	0.618	0.029†	0.42 (0.19-0.91)	0.892	0.924	1.48 (1.28-1.71)	0.584	0.701
LOC387761	rs7480010	G	0.858	0.890	0.002	0.084	1.18 (0.98-1.44)	0.301	0.336	1.14 (1.01-1.27)	0.062	0.117
EXT2/ALX4	rs1113132	C	0.915	0.920	0.579	0.221†	0.47 (0.14-1.57)	0.733	0.763	1.15 (0.88-1.42)	0.475	0.600
EXT2/ALX4	rs11037909	T	0.862	0.859	0.768	0.511	0.94 (0.79-1.13)	0.729	0.760	1.27 (0.97-1.57)	0.913	0.953
EXT2/ALX4	rs3740878	A	0.907	0.914	0.415	0.129†	0.46 (0.17-1.26)	0.728	0.760	1.26 (0.97-1.55)	0.760	0.846
FTO	rs8050136	A	0.446	0.452	0.686	0.783	1.02 (0.90-1.15)	0.398	0.455	1.23 (1.18-1.32)	0.711	0.808
TCF7L2*	rs7903146	T	0.284	0.354	1.73x10⁻⁶	1.59x10⁻⁶	1.39 (1.21-1.60)	0.181	0.227	1.37 (1.31-1.43)	0.997	0.999

†SNPs have minor allele homozygote counts <10 in case or control population and dominant model P-value and OR are reported. *Power analysis for *TCF7L2* was calculated using a population of 960 cases and 1000 controls. **Bold:** P-values <0.05 and corresponding OR. European risk allele frequencies and odds ratios were obtained from recent WGA studies [1-6].

Table 3. Minor allele Frequencies for AA, YRI, and CEU populations

SNP	Alleles	Reported risk allele	Reported risk allele frequency		
			AA frequency	HAPMAP frequency (YRI)	HAPMAP frequency (CEU)
rs6698181	C/T	T	0.155	0.050	0.367
rs4402960	T/G	T	0.527	0.550	0.292
rs17044137	T/A	A	0.327	0.400	0.258
rs10946398	C/A	C	0.598	0.667	0.308
rs7754840	C/G	C	0.600	0.667	0.308
rs13266634	C/T	C	0.915	0.942	0.750
rs564398	T/C	T	0.939	1.000	0.625
rs10811661	T/C	T	0.930	1.000	0.792
rs1111875	C/T	C	0.770	0.858	0.558
rs5015480	C/T	C	0.627	0.568	0.552
rs7923837	G/A	G	0.923	1.000	0.625
rs9300039	C/A	C	0.886	0.831	0.892
rs7480010	G/A	G	0.873	1.000	0.246
rs1113132	C/G	C	0.918	0.925	0.700
rs11037909	T/C	T	0.860	0.842	0.700
rs3740878	A/G	A	0.911	0.924	0.698
rs8050136	C/A	A	0.449	0.467	0.450
rs7903146	C/T	T	0.319	0.292	0.250