Human Molecular Genetics, 2010, Vol. 19, No. 14 **2907–2916**
doi:10.1093/hmg/ddq178
Advance Access published on April 29, 2010

Fine mapping of the association with obesity at the *FTO* locus in African-derived populations

Mohamed T. Hassanein[1,†,‡], Helen N. Lyon[2,3,†,‡], Thutrang T. Nguyen[2,3], Ermeg L. Akylbekova[4], Kevin Waters[1], Guillaume Lettre[5,6], Bamidele Tayo[7], Terrence Forrester[8], Daniel F. Sarpong[4], Dan O. Stram[1], Johannah L. Butler[2,3], Rainford Wilks[9], Jiankang Liu[4], Loic Le Marchand[10], Laurence N. Kolonel[10], Xiaofeng Zhu[11], Brian Henderson[1], Richard Cooper[7], Colin McKenzie[8], Herman A. Taylor Jr[4,12], Christopher A. Haiman[1,*,‡] and Joel N. Hirschhorn[2,3,13,*,‡]

[1]Department of Preventive Medicine, Keck School of Medicine and Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90089, USA, [2]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA, [3]Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA, [4]Jackson Heart Study, Jackson State University, Jackson, MS 39213, USA, [5]Faculté de médecine, Université de Montréal, Montréal, Québec, Canada, H1C 3J7, [6]Institut de cardiologie de Montréal, Montréal, Québec, Canada H1T 1C8, [7]Department of Preventive Medicine and Epidemiology, Stritch School of Medicine, Loyola University Chicago, Maywood, IL 60153, USA, [8]Tropical Metabolism Research Unit and [9]Epidemiology Research Unit, Tropical Medicine Research Institute, University of the West Indies, Mona, Kingston 7, Jamaica, [10]Epidemiology Program, Cancer Research Center, University of Hawaii, Honolulu, HI 96813, USA, [11]Department of Biostatistics and Epidemiology, Case Western Reserve University, Cleveland, OH 44106, USA, [12]Jackson Heart Study, University of Mississippi Medical Center, Jackson, MS 39213, USA and [13]Program in Genomics and Divisions of Genetics and Endocrinology, Children's Hospital, Boston, MA 02115, USA

Received February 25, 2010; Revised and Accepted April 26, 2010

Genome-wide association studies have identified many common genetic variants that are associated with polygenic traits, and have typically been performed with individuals of recent European ancestry. In these populations, many common variants are tightly correlated, with the perfect or near-perfect proxies for the functional or true variant showing equivalent evidence of association, considerably limiting the resolution of fine mapping. Populations with recent African ancestry often have less extensive and/or different patterns of linkage disequilibrium (LD), and have been proposed to be useful in fine-mapping studies. Here, we strongly replicate and fine map in populations of predominantly African ancestry the association between variation at the *FTO* locus and body mass index (BMI) that is well established in populations of European ancestry. We genotyped single nucleotide polymorphisms that are correlated with the signal of association in individuals of European ancestry but that have varying degrees of correlation in African-derived individuals. Most of the variants, including one previously proposed as functionally important, have no significant association with BMI, but two variants, rs3751812 and rs9941349, show strong evidence of association ($P = 2.58 \times 10^{-6}$ and 3.61×10^{-6} in a meta-analysis of 9881 individuals). Thus, we have both strongly replicated this association in African-ancestry populations and narrowed the list of potentially causal variants to those that are correlated with rs3751812 and rs9941349 in African-derived populations. This study illustrates the potential of using populations with different LD patterns to fine map associations and helps pave the way for genetically guided functional studies at the *FTO* locus.

*To whom correspondence should be addressed. Tel: +1 3234427755 (C.A.H.)/+1 6179192129 (J.N.H); Fax: +1 3234247749 (C.A.H.)/+1 6177300253 (J.N.H.); Email: haiman@use.edu (C.A.H.)/joelh@broadinstitute.org (J.N.H.)
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.
‡These authors contributed equally to the work.

© *The Author 2010. Published by Oxford University Press. All rights reserved.*
For Permissions, please email: journals.permissions@oxfordjournals.org

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

INTRODUCTION

Susceptibility or resistance to obesity is partially determined by inherited factors, but most of the genetic contributors for this polygenic trait remain unknown (1,2). Recently, genome-wide association (GWA) studies have been performed to identify common genetic variants that influence body mass index (BMI) and other heritable anthropometric measures of obesity (1,3–10). Although genetic loci identified by GWA studies can provide insights into novel biology, these studies usually highlight a number of common variants that are typically tightly correlated [in linkage disequilibrium (LD)] with many other nearby variants, effectively precluding further localization or fine mapping of the association signal (11). Furthermore, GWA studies have largely been performed using samples from European-derived populations, so it is not clear whether the associated loci are also relevant in populations of different continental ancestries.

At least two purposes can be served by studying, in African-ancestry populations, the loci that were highlighted by GWA studies performed in European-ancestry populations (12). First, sufficiently large studies in populations of African ancestry can test whether the loci identified in European-derived samples are also relevant in individuals of African ancestry. Second, because of the longer time to common ancestors between individuals of African ancestry, the degree of LD in African-ancestry populations is less than in European populations (13). Thus, fewer common variants are expected to be tightly correlated with the biologically relevant variant, potentially enabling fine-mapping efforts that decrease the set of variants that need to be tested in laborious functional or mechanistic studies.

The first GWA study of BMI reported an association in Europeans between BMI and rs9939609, an intronic variant in the fat mass and obesity associated (*FTO)* gene (9). This association has been replicated in many other studies among men and women with recent European ancestry (5–7,14). To date, this variant has the strongest known effect of any single nucleotide polymorphism (SNP) on BMI in European-derived populations. Functional studies have suggested that rs8050136, tightly correlated with rs9939609, could be the biologically relevant variant at this locus (15).

Initial attempts to study this variant in African-ancestry populations have yielded inconsistent results. Scuteri *et al*. (7) were the first to test variants in the *FTO* gene in African-American samples, and found no evidence of association. Grant *et al*. (16) have shown that the variant highlighted in the first GWA study, rs9939609, was not associated with obesity in African-American children. They identified a nearby SNP, rs3751812, that was in strong LD with rs9939609 in Europeans ($r^2 = 0.9$ in the HapMap CEU reference panel) but not in Africans ($r^2 = 0.06$ in the HapMap YRI reference panel) and provided evidence that this variant was associated with obesity in African-American children (OR = 1.31; $P = 0.017$). However, further analyses of rs3751812 on BMI as a quantitative trait in African-American children in this same study were not significant. In addition, no signal of association was seen with rs9939609 in a multiethnic study of women (17), or with SNPs in the *FTO* region in a Gambian population (18).

In an attempt to clarify whether variants in the *FTO* region influence BMI in African-ancestry populations, and to begin to fine map any association signals, we performed extensive genotyping of SNPs in the region in two African-American populations, and genotyped the best-associated SNPs in additional populations with predominantly African ancestry. We used a dense fine-mapping strategy including ancestry-adjusted analyses to minimize possible confounding by stratification. We strongly confirmed an association at rs3751812, providing definitive evidence that the *FTO* locus influences BMI in African Americans. We also identified an additional variant, rs9941349, which shows similar levels of association with BMI as seen for rs3751812. We are unable to distinguish between rs3751812 and rs9941349 through conditional and haplotype analyses, and suggest that one of these SNPs, one of the variants known to be in strong LD with these SNPs, or an as yet undetermined variant in strong LD with one or both of these SNPs, could plausibly be the biologically relevant variant at this locus. We also conclude that other well-studied variants such as rs9939609 or rs8050136 are highly unlikely to be responsible for the signal of association seen in Europeans.

RESULTS

We sought to fine map and extend in African Americans the association of variation at *FTO* with BMI. In this study, different fine-mapping approaches were taken in parallel in the Multiethnic Cohort (MEC; $n = 3482$) and the Jackson Heart Study (JHS; $n = 4217$) samples (see Materials and Methods; baseline characteristics of the samples are described in Table 1), so the sets of SNPs genotyped in each sample are slightly different. In the JHS, we chose a set of SNPs designed to capture variation in African Americans across a 245 kb region (chr16: 52,181,324 to 52,406,062) which includes the 51 kb LD block in intron 1 of *FTO* (HapMap CEU) containing the index signal (chr16: 52,355,066 to 52,406,062) as well as the neighboring *RPGRIP1L* gene (Fig. 1). In JHS, we focused on variants that are in LD with the associated SNPs in European ancestry reference panels (HapMap CEU) but not in LD with each other in an African-ancestry panel (HapMap YRI). In the MEC, fine mapping was limited to the defined LD block containing the index signal and common alleles in the YRI HapMap panel that had previously been genotyped in an African-American population and had not been associated with BMI (7).

We did not replicate the initial observed association from studies conducted in European populations between BMI and the lead SNPs rs9939609 and rs8050136 for JHS and MEC, respectively (Table 2). In the JHS sample, we observed strong evidence of association at rs3751812, and its proxy rs1558902 ($r^2 = 0.96$ with rs3751812 in JHS; Table 2). Slightly weaker associations were observed at two other proxies of rs3751812 (Table 2): rs1421085 and rs17817964 ($r^2 = 0.95$ and $r^2 = 0.94$ in JHS with rs3751812, respectively). We also observed a strong association at rs9941349, a SNP with a higher minor allele frequency (MAF) but only in moderate LD with rs3751812 ($r^2 = 0.50$ in JHS), and at rs9931494, a nearly perfect proxy for rs9941349 ($r^2 = 0.97$ in JHS; Table 2). The rs3751812 and rs9941349 SNPs are

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

Table 1. Characteristics of the African-derived populations

	N	Women	Age	BMI (kg/m^2)	Height (cm)	Weight (kg)	Normal weight (BMI <25)	Overweight (BMI 25–30)	Obese (BMI >30)
JHS	4217	2631 (62.4%)	54.20 (12.56)	31.85 (7.21)	169.10 (9.28)	91.02 (21.43)	586 (13.9%)	1359 (32.2%)	2272 (53.9%)
MEC	3482	1254 (36.0%)	61.95 (8.09)	28.78 (5.33)	172.77 (9.81)	85.79 (16.49)	747 (21.5%)	1583 (45.5%)	1152 (33.1%)
SPT	1407	864 (61.4%)	46.45 (13.81)	26.87 (6.01)	165.44 (8.55)	73.33 (16.24)	596 (42.4%)	436 (31.0%)	375 (26.7%)
G×E	938	698 (74.4%)	39.66 (8.16)	30.82 (6.91)	166.39 (8.04)	85.20 (19.62)	216 (23.0%)	121 (12.9%)	601 (64.1%)
Maywood	775	453 (58.5%)	38.34 (10.94)	29.88 (8.20)	169.05 (11.32)	85.43 (23.39)	249 (32.1%)	207 (26.7%)	319 (41.2%)

For age, BMI, height and weight, mean and standard deviation are shown. JHS, African-American sample from Jackson Heart Study; MEC, African-American sample from the Multiethnic Cohort; SPT, Spanish Town, Jamaica population-based sample; G×E, gene by environment study from Kingston, Jamaica; Maywood, African-American sample from Maywood, IL, USA.

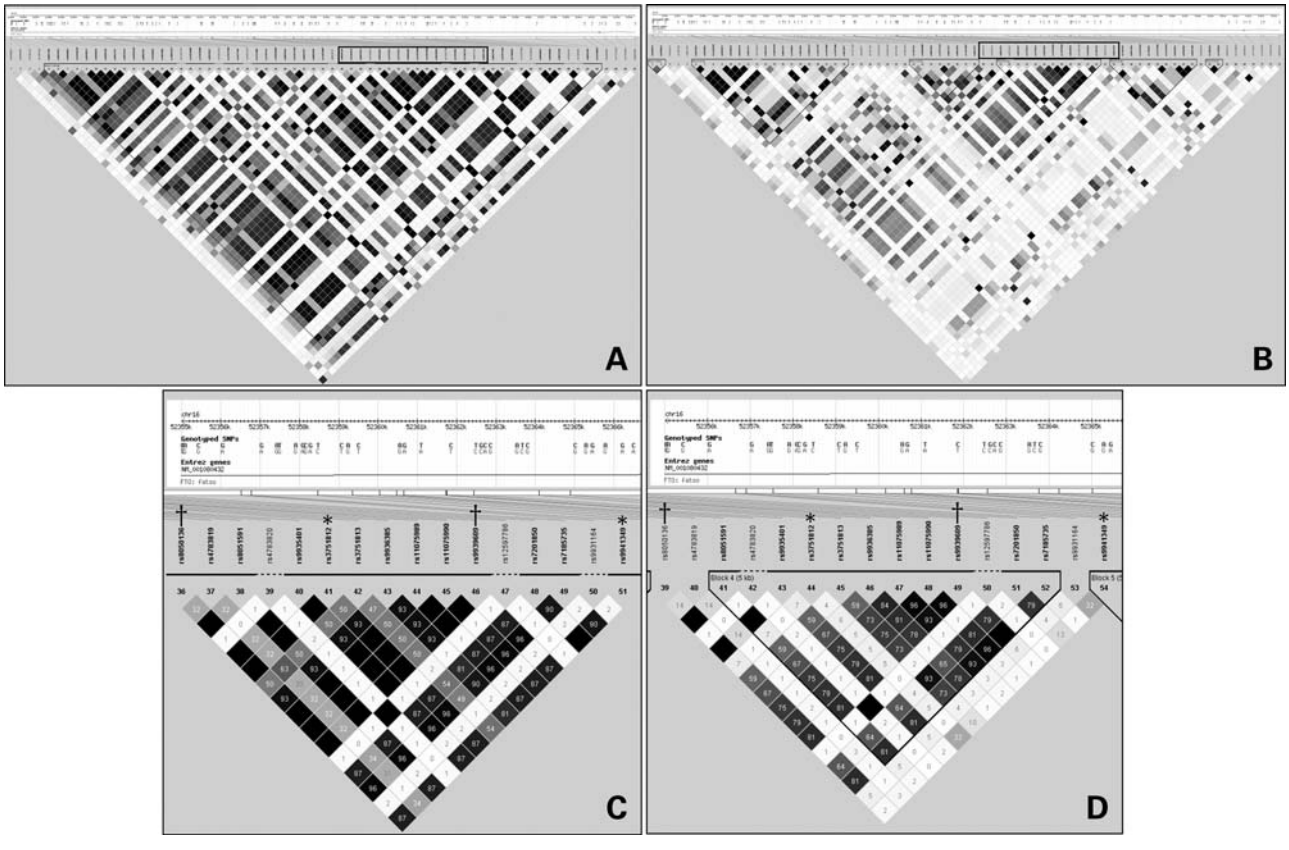

Figure 1. Linkage disequilibrium and haplotype block structure in intron 1 of *FTO*. (**A**) and (**B**) represent the CEU, and YRI populations from HapMap phase 2 data. Pairwise correlation among SNPs (r^2) are shown. The black box highlights the region targeted for fine mapping and is shown in (**C**) and (**D**) for CEU and YRI, respectively. Dagger represents index SNP identified in populations of European ancestry. Asterisks represent the two variants identified through our fine-mapping study.

both strongly correlated with the lead SNPs in the HapMap CEU sample ($r^2 = 0.90–1.0$) but not in the HapMap YRI sample ($r^2 = 0.02–0.07$) or in JHS ($r^2 = 0.03–0.11$) (19). In the MEC, none of the SNPs initially genotyped were found to be associated with BMI after correction for multiple comparisons, but additional genotyping confirmed the association at rs3751812 as well as at rs9941349 (Table 2).

Further testing in additional cohorts of largely African ancestry provided additional evidence for association with these SNPs (Table 3 and Supplementary Material, Table S1). A meta-analysis of the four cohorts in which BMI was analyzed as a continuous trait (JHS, MEC, Spanish Town, and Maywood) revealed a signal of association for rs3751812 ($P = 2.58 \times 10^{-6}$; Table 3). A comparable association was seen for rs9941349 ($P = 3.61 \times 10^{-6}$; Table 3). The estimated effect size per allele at rs3751812 was 0.0121 log BMI units and 0.0093 log BMI units for rs9941349. I^2 values were 0% ($P = 0.70$) for rs3751812 and 22.7% ($P = 0.27$) for rs9941349 indicating low heterogeneity between the cohorts. Although the Jamaican Gene × Environment Panel (G×E) was not included in the overall meta-analysis because BMI was analyzed as a dichotomous trait in this case–control

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

Table 2. SNPs genotyped for fine mapping

	SNP	Position	Minor allele	Major allele	MAF	*N*	Effect size (log BMI units)	Standard error	*P*-value (nominal)
JHS	rs1421085	52358455	C	T	0.103	4199	0.0120	0.0033	0.0003
	rs1558902	52361075	A	T	0.084	4021	0.0151	0.0035	2.10E−05
	rs17817288	52365265	G	A	0.400	4194	0.0059	0.0021	0.0044
	rs9924817	52365425	A	G	0.024	4212	−0.0160	0.0073	0.0289
	rs16952525	52367515	G	A	0.078	4092	0.0077	0.0038	0.0418
	rs8054237	52373366	A	G	0.269	4106	−0.0058	0.0024	0.0175
	rs3751812	52375961	T	G	0.099	4205	0.0120	0.0034	0.0004
	rs9939609[a]	52378028	A	T	0.469	4042	0.0018	0.0022	0.4125
	rs9941349	52382989	T	C	0.180	4212	0.0106	0.0026	5.85E−05
	rs9931494	52384680	G	C	0.176	4203	0.0117	0.0027	1.34E−05
	rs17817964	52385567	T	C	0.103	4208	0.0117	0.0033	0.0004
	rs9933040	52388368	A	T	0.209	4209	0.0072	0.0024	0.0030
	rs9922708	52388647	T	C	0.199	4175	0.0074	0.0027	0.0061
MEC	rs8050136[a]	52373776	A	C	0.436	3452	0.0042	0.0041	0.2990
	rs3751812	52375961	T	G	0.117	3451	0.0148	0.0061	0.0158
	rs9941349	52382989	T	C	0.191	3418	0.0113	0.0051	0.0272
	rs1861867	52406062	T	C	0.166	3440	−0.0126	0.0054	0.0207

Shown are significant results and previously reported signals. Position is on chromosome 16, relative to NCBI B36 assembly, dbSNP b126. Effect direction is with respect to the minor allele. JHS and MEC are as in Table 1. MAF, minor allele frequency.
[a]Reported signals in European-derived populations.

Table 3. Meta-analysis across studies with continuous traits

	Variable	rs3751812	rs9941349
JHS	MAF	0.099	0.180
	Effect	0.0120	0.0106
	SE	0.0034	0.0026
	P-value	0.0004	5.85E−05
MEC	MAF	0.117	0.191
	Effect	0.0148	0.0113
	SE	0.0061	0.0051
	P-value	0.0158	0.0272
SPT	MAF	0.078	0.173
	Effect	0.0075	0.0020
	SE	0.0060	0.0044
	P-value	0.2139	0.6443
G×E	MAF	0.083	0.173
	OR	1.117	1.403
	SE	0.1971	0.1524
	P-value	0.5761	0.0264
Maywood	MAF	0.111	0.167
	Effect	0.0208	0.0161
	SE	0.0110	0.0080
	P-value	0.0587	0.0446
Combined[a]	Effect	0.0121	0.0093
	SE	0.0026	0.0020
	P-value	2.58E−06	3.61E−06

Shown are the results for rs3751812 and rs9941349 from the samples and from a meta-analysis. The samples are as in Table 1. The effect direction is in respect to the minor allele and is in log BMI units. MAF, minor allele frequency; Effect, effect estimate; SE, standard error; OR, odds ratio.
[a]Combined analysis does not include G×E in the meta-analysis.

cohort, the results for rs3751812 and rs9941349 trended in the same directions in this sample (Table 3).

To help assess the effects of possible confounding by ancestry in these admixed populations, we estimated global African and European ancestry in both the JHS and a subset of the MEC, and global and local ancestry in JHS. The association evidence in the subset of individuals with ancestry estimates was not diminished by adjusting for ancestry, suggesting that the results were not false positives due to admixture (Supplementary Material, Table S2). We also observed nominally significant association when using family-based tests of association in the subset of related samples in JHS for both rs3751812 and rs9941349 and in the family-based Maywood cohort for rs9941349 (rs3751812 approached significance in this sample) (Table 2 and Supplementary Material, Table S1). Furthermore, the effect in the African Americans is unlikely to be due to the presence of European chromosomes in this population, because in this case we would expect to see equal association among all of the variants that are strongly correlated in European populations, whereas in our data we saw strong differences in association among this group of variants. Thus, these data provide the strongest evidence to date that variation in the *FTO* locus influences BMI in African-ancestry populations.

We next attempted to discern which of the two associated SNPs, rs3751812 and rs9941349, were more closely correlated with the biologically relevant variant at the locus (assuming that there is one such variant responsible for the observed association signal). We focused on these two SNPs because (i) they had the strong signals of association (ii), their associations with BMI remained significant when conditioning on other individual SNPs (except for rs3751812 or rs9941349 or their proxies) and (iii) other SNPs in the region were not strongly associated after conditioning on either of these two SNPs (Supplementary Material, Tables S3 and S4). We performed conditional analyses of rs3751812 and rs9941349 in the JHS, MEC, Spanish Town (SPT), G×E and the Maywood cohort, testing for the association of each SNP when using the other SNP as a covariate (Supplementary Material, Table S3). A meta-analysis of these conditional analyses in the samples with continuous traits (JHS, MEC, SPT and Maywood) showed that the association signal at SNP rs3751812 remained nominally significant after conditioning on the effect of SNP rs9941349 (overall $P =$

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

Table 4. Conditional analysis on rs3751812 and rs9941349

	SNP	Effect estimate	Standard error	*P*-value	Effect estimate	Standard error	*P*-value	Effect estimate	Standard error	*P*-value
		No conditioning			Conditioned on rs3751812			Conditioned on rs9941349		
JHS	rs3751812	0.0120	0.0034	0.0004				0.0050	0.0051	0.3243
	rs9941349	0.0106	0.0026	5.85E−05	0.0077	0.0037	0.0343			
MEC	rs3751812	0.0148	0.0061	0.0158				0.0122	0.0090	0.1736
	rs9941349	0.0113	0.0051	0.0272	0.0032	0.0075	0.6723			
SPT	rs3751812	0.0075	0.0060	0.2139				0.0096	0.0078	0.2198
	rs9941349	0.0020	0.0044	0.6443	−0.0024	0.0057	0.6725			
G×E	rs3751812	0.1102	0.1971	0.5761				−0.3876	0.2845	0.1731
	rs9941349	0.3383	0.1524	0.0264	0.5407	0.2196	0.0138			
Maywood	rs3751812	0.0208	0.0110	0.0587				0.0125	0.0139	0.3686
	rs9941349	0.0161	0.0080	0.0446	0.0107	0.0102	0.2926			
Combined[a]	rs3751812	0.0121	0.0026	2.58E−06				0.0078	0.0037	0.0354
	rs9941349	0.0093	0.0020	3.61E−06	0.0050	0.0027	0.0678			

Shown are the results for rs3751812 and rs9941349 before and after conditioning on the two SNPs. Samples are as in Table 1. The effect estimates are in log BMI units and are in reference to the minor allele.
[a]Does not include G×E in the meta-analysis.

0.0354; Table 4). However, the association signal at rs9941349 was nearly significant when conditioning on rs3751812 (overall $P = 0.0678$; Table 4). We therefore do not feel that this analysis can demonstrate with confidence which of rs3751812 and rs9941349 are more tightly correlated with a true causal variant, and may even suggest that neither of these two variants are themselves causal.

We also performed a haplotype analysis to try to assess which of these two SNPs might be more tightly correlated with a causal variant at the locus (see Materials and Methods). Because LD is strong between these two SNPs ($D' = 91–100$), there are three major haplotypes. We designated the three major haplotypes as H–H, H–L and L–L, corresponding to the BMI-increasing (H) or decreasing (L) alleles at rs9941349 and rs3751812, respectively. The H–H, H–L and L–L haplotypes (T–T, T–G, and C–G) accounted for 9.6, 8.0 and 82.3% of haplotypes in the JHS population, 11.5, 7.7 and 80.5% of haplotypes in the MEC population, 7.7, 9.5 and 82.8% of haplotypes in the SPT population, 8.3, 9.0 and 82.7% of haplotypes in the G×E population and 7.2, 7.9 and 84.4% of haplotypes in the Maywood cohort. In this analysis, the H–L haplotype was associated with a higher BMI than the L–L haplotype ($P = 0.009$), suggestive of an effect of the rs9941349 SNP (or a variant in LD with this SNP) independent of rs3751812; the H–H and H–L haplotypes were not significantly different, although the data were directionally consistent with the H–H haplotype having a stronger effect on increasing BMI than the H–L haplotype (Table 5). Thus, the conditional analysis showed a nominally significant effect of rs3751812, and the haplotype analysis showed a nominally significant effect of rs9941349; together these results suggest that neither variant has completely captured the signal of association at *FTO*, but are consistent with both variants being strongly correlated with a causal variant at this locus.

DISCUSSION

Our findings strongly indicate that there is an association between variation in the *FTO* region and BMI in African-derived populations. Because the variants showing strong association are tightly correlated in European-ancestry populations with the originally reported variants, we doubt that these represent a statistical fluctuation. Furthermore, the association does not appear to be a false positive due to admixture in these populations. Thus, these data provide additional and compelling evidence that the association with *FTO* discovered in European-derived populations extends into individuals with predominantly African ancestry.

Variation at this locus accounted for 0.11–0.56% and 0.02–0.61% of variance in log BMI by rs3751812 and rs9941349, respectively, for our population of subjects with African ancestry and continuous traits. The average proportion of variance in BMI in our samples explained by rs3751812 was 0.31% and by rs9941349 was 0.30%. The variance explained at this locus in our samples is thus similar to but perhaps slightly less than that estimated in populations of European ancestry, where 0.34% of the variation in BMI was explained by rs9939609 at this locus (5).

Because our association was observed with variants that are in strong LD with SNPs shown to be associated in European populations, it is likely that the associations in African-derived and European-derived populations arise at least largely from the same causal variant(s). We can clearly rule out rs9939609 and many other variants that are in LD with this variant in Europeans (including the previously proposed rs8050136). However, we cannot confidently distinguish between rs3751812, rs9941349 or their close correlates in African populations. Furthermore, the finding that each of rs3751812 and rs9941349 is nominally associated with BMI even after accounting for the effect of the other variant raises the possibility that a third variant, partially correlated with both rs3751812 and rs9941349, could be causal at this locus. Alternatively, one of these variants could be causal, and one or more additional causal variants (partially captured by the other variant) could also influence BMI. Complete resequencing of this region (such as will be provided by the 1000 Genomes Project) and comprehensive testing of these variants in

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

Table 5. Test for differences between haplotypes

	JHS			MEC			SPT			G×E			Maywood			Combined[a]		
	Effect	SE	*P*-value	Effect	SE	*P*-value	Effect	SE	*P*-value	Effect	SE	*P*-value	Effect	SE	*P*-value	Effect	SE	*P*-value
H–H versus H–L	0.003	0.005	0.578	0.003	0.010	0.748	0.005	0.009	0.539	−0.349	0.301	0.247	0.011	0.015	0.437	0.004	0.004	0.306
H–L versus L–L	0.009	0.004	0.017	0.009	0.008	0.280	0.000	0.006	0.966	0.415	0.234	0.076	0.015	0.010	0.148	0.008	0.003	0.009

H–H, H–L and L–L indicate haplotypes of rs9941349 and rs3751812, respectively, where 'H' refers to the allele associated with increased BMI and 'L' the allele associated with decreased BMI (see text). The strategy for haplotype analysis is depicted in Figure 2. Effect, SE, *P*-value: results from meta-analysis of the regression coefficients (effect estimates and standard error) for the comparisons of the haplotypes (see Material and Methods for details). Effect direction is in reference to the minor allele and is in log BMI units. Samples are as described in Table 1.
[a]Combined analysis does not include G×E in the meta-analysis.

large populations of African ancestry will be needed to localize a causal variant with more certainty.

Other studies have previously examined the influence of the *FTO* locus on BMI in populations with African ancestry. Grant *et al*. (16) were the first to report rs3751812 as associated with obesity in a case–control study of African-American children, although the association with BMI as a continuous trait in 1962 children was non-significant. Scuteri *et al*. (7) did not find a significant association of rs3751812 with BMI in subjects with African ancestry, although they stated that their study was under powered to detect modest effects with SNPs that have relatively low minor allele frequencies in African Americans, such as rs3751812. In addition, work conducted by Hennig *et al*. (18) found no associations for several SNPs including rs3751812 and rs9931494 at this locus with weight-for-height z-scores, in a Gambian population. Although the study was moderately well powered ($n = 2200$), the difference in study populations, including strong environmental differences, possible differences in genetic background and leanness of the Gambian population, may explain the null results. Indeed, the effect of *FTO* has been proposed to be modified by physical activity (20,21), so it is perhaps plausible that effect sizes may have been smaller in the Gambian individuals. Song *et al*. (17) did not detect an association with BMI among a subset of ∼1100 African-American women for rs9939609 in a cohort of post-menopausal women, although, in a separate family-based study, Wing *et al*. (22) found an association with rs8050136 ($P = 0.01$) and rs9939609 ($P = 0.01$), but not rs3751812, with BMI using variance components analysis in a sample of 581 African-American subjects. Finally, Thorleifsson *et al*. (6) saw no evidence of association of rs3751812 with BMI in a study of 1160 African Americans. Despite these inconsistencies in the published literature, our study, with nearly 10 000 subjects, provides convincing evidence of association, and suggests that, as in studies of European-Americans, large samples are needed to obtain evidence of association with BMI in African Americans. Furthermore, the effect size, at least for these variants, appears to be similar to or even smaller than that seen in European-Americans, perhaps in part explaining the previous difficulty in demonstrating the association.

Our data provide additional evidence that the *FTO* locus influences BMI in African Americans and that fine mapping in African-ancestry populations can be useful in narrowing the list of potentially causal variants. Such fine-mapping studies will be even more useful as complete catalogues of common variants become available, but will also require very large sample sizes to distinguish between closely correlated variants. Coordinated efforts between investigators who are studying individuals of recent African ancestry will likely be required to attempt fine mapping at additional loci, especially as the effect sizes of these loci are substantially smaller than seen at *FTO*. Indeed, at loci with very small effect sizes, the number of currently available samples of recent African ancestry may be sufficient to narrow the possible universe of biologically relevant variants at associated loci, but may not be sufficient to narrow the association to a single variant, unless the functional variant happens not to have close proxies in African populations.

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

MATERIALS AND METHODS

Samples

Baseline characteristics of the samples are described in Table 1.

The Jackson Heart Study (JHS). JHS is a population-based longitudinal study comprising of African Americans from the Jackson, Mississippi metropolitan area. Participants provided medical and social history, and physical and biochemical data were collected. The participants have a high prevalence of diabetes, hypertension, obesity and related cardiovascular disorders (23).

We studied a data set of 4539 individuals, of which 1464 individuals were in 293 families. Individuals were removed if there were phenotype data missing or a genotyping success rate of less than 85%. After checking for Mendelian errors, one individual was removed due to inconsistency with the pedigree information. Four individuals had BMI greater than four standard deviations from the mean of log-transformed BMI normalized by Z-scores, and were removed from the data set. In total, 4217 individuals had genotype and phenotype data suitable for the study.

The Multiethnic Cohort (MEC). The MEC is a prospective cohort study that includes 215 251 men and women from five racial-ethnic groups in Hawaii and California (mainly Los Angeles): Japanese-Americans, European Americans, Native Hawaiians from Hawaii, African Americans and Latinos from California. Cohort members were identified through the Department of Motor Vehicles drivers' license files and, for African Americans, in addition through Health Care Financing Administration files. Cohort members were between 45 and 75 years old at baseline (1993–1996). Information about risk factors for cancer and other chronic diseases was ascertained at baseline through a 26-page questionnaire. Beginning in 2002, we established a large bio-repository of blood and urine samples from approximately 70 000 MEC participants.

In this study, we studied a sample of 3556 African-American men and women with self-reported BMI who were included as cases and controls in ongoing genetic studies of breast cancer, prostate cancer, hypertension and type 2 diabetes in the MEC (24–26). Individuals were removed if they were non-informative for at least 3 of the 11 SNPs. In total there were 3482 individuals that met these criterion.

Spanish Town (SPT). SPT is part of a larger project, the International Collaborative Study of Hypertension in Blacks, which was previously described in detail by Cooper *et al*. (27). For SPT, a stratified random sample of unrelated Jamaicans between 25 and 74 years was recruited in and around Spanish Town, an urban area neighboring the capital city of Kingston. In this study, we genotyped successfully in 1407 individuals, after removing individuals with no BMI phenotype or were non-informative for at least two of the six SNPs.

The Jamaican Gene × Environment panel (G×E). G×E is part of a larger survey, which examined gene by environment interactions in blood pressure for adults between 25 and 74 years old (27,28). Participants were unrelated individuals from Kingston, Jamaica and their BMI values were either in the top or bottom third of the BMI distribution for the Jamaican population. QC criteria were similar to those for SPT and 938 individuals had suitable phenotype and genotype data.

The Maywood Study. African-American DNA samples were obtained from a larger cohort of families enrolled in studies of blood pressure at Loyola University in Maywood, IL, USA. The survey enrolled a representative random sample of the population between the ages of 18–74, regardless of obesity phenotype. QC criteria were similar to those for JHS. There were 775 people from 311 families for family-based association studies with suitable genotype and phenotype data.

For all studies, informed consent was obtained, and the parent studies were all approved by local institutional review boards. The genetic study protocols were approved by the institutional review boards at Children's Hospital, Boston and at the University of Southern California.

Ancestry. Ancestry information was available for the majority of the JHS and a portion of the MEC samples. Continental (European and African) ancestry was estimated by genotyping ancestry informative markers. For JHS, global and local ancestry estimates are described in detail in Nalls *et al*. (29). Of the 4217 JHS individuals, we have ancestry information for 4094 people. For the MEC, principal components were computed in EIGENSTRAT using a panel of >1400 markers that had been genotyped among 1879 subjects (of the 3482, 54%) in previous studies (30).

SNP selection

The variants reported to be associated with obesity were in a 51 kb region of strong LD in the CEU HapMap panel, within intron 1 (*chr16*: 52,355,066 to 52,406,062). Initial studies in the JHS and MEC samples were performed in parallel. In the JHS, we chose a set of SNPs designed to capture variation in African Americans across a larger 245 kb region (52,181,324 to 52,406,062) which included the LD block containing the index signal as well as the neighboring *RPGRIP1L* gene. Tag SNPs covering these regions were selected using Tagger (31) as implemented in Haploview (32). We enriched this set for possible proxies of the original signal at rs9939609 and rs1421085 by choosing all SNPs with $r^2 \geq 0.5$ with the original SNPs in HapMap CEU population and then tagging these SNPs with an $r^2 \geq 0.9$ in HapMap YRI population using Tagger.

In the MEC, SNP selection for fine mapping the defined LD block containing the index signal was conducted using the Tagger function in Haploview in the YRI HapMap population. A pairwise r^2 cutoff of 0.80 resulted in 34 SNPs that characterize the risk region. Utilizing the 'force include' option in Haploview, we excluded rs8050136 and rs9939609, as well as nine SNPs that had previously been genotyped in an African-American population and had not been associated with BMI (7). Of the remaining SNPs, nine were excluded with a MAF in CEU $<5\%$, and six SNPs were excluded for

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

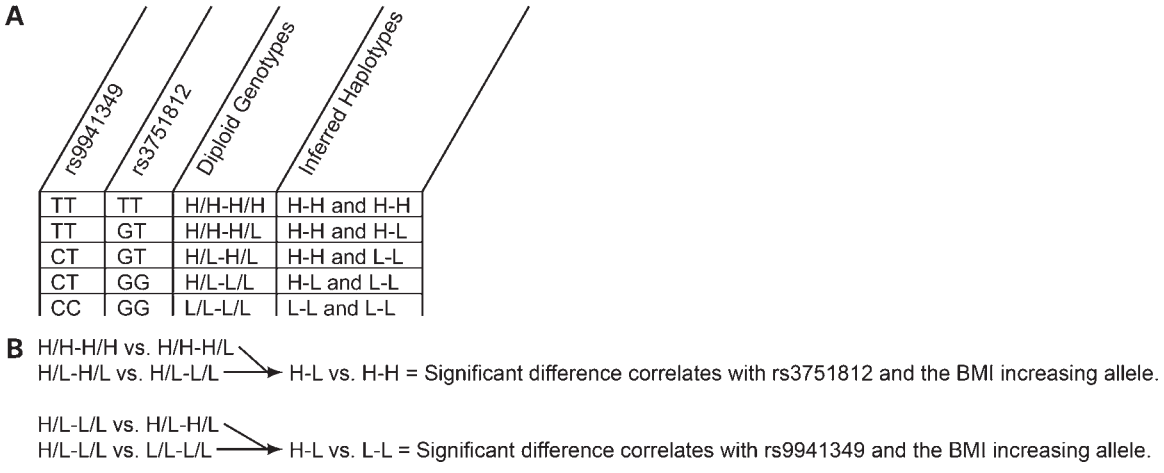

Figure 2. Test for differences between haplotypes. There were three major haplotypes in the population, T–T, T–G and C–G for rs9941349 and rs3751812, respectively. (**A**) Individuals were assigned haplotypes based on their diploid genotype data where 'H' refers to the allele associated with increased BMI ('T' for both SNPs) and 'L' refers to the allele associated with decreased BMI ('C' for rs9941349 and 'G' for rs3751812). Because of the low frequency of the H–L haplotype, the H/H–L/L genotype was assigned to H–H and L–L. (**B**) The H–L haplotype was compared to the H–H and L–L haplotypes by comparing groups of individuals that had one haplotype in common but differed at their other haplotype. For the H–L and H–H haplotype, the group comparisons were H–H and H–L versus H–H and H–H as well as H–L and L–L versus H–H and L–L. For the H–L and L–L haplotype, the two group comparisons were H–H and H–L versus H–H and L–L in addition H–L and L–L versus L–L and L–L. For each group comparison, the results were meta-analyzed.

having a pairwise correlation (r^2) <0.10 with rs8050136 and rs9930506. This resulted in eight tag SNPs (rs4784323, rs7206790, rs8047395, rs8055197, rs17817288, rs8057044, rs8044769 and rs1861867) that were included in our fine-mapping analysis for this region of the *FTO* gene. In the MEC, SNPs rs3751812, rs9941349 and rs1652525 were later genotyped to attempt to replicate the association signals identified from the JHS.

Statistical analysis

For JHS, SPT and Maywood, we generated residuals (logBMIz) from the BMI phenotype (log-transformed), with age and gender as covariates. In JHS and Maywood, association testing was conducted using MERLIN v1.1.2 (33) to account for a small family-based component of the sample. Global and local estimates of the percentage African ancestry were also included as covariates in the model for the subset of individuals in JHS. In the G×E cohort, individuals were largely ascertained from the top and bottom tertiles of the BMI distribution, so BMI was analyzed as a dichotomous trait (<25 versus $>27\ kg/m^2$). For the SPT and G×E studies, association testing and conditional analysis were done using the PLINK v1.05 software package (34).

In the MEC, genotype allele frequencies were assessed for deviation from Hardy–Weinberg equilibrium using a χ^2 tests with a *P*-value cutoff of 0.001; one SNP, rs8044769, was removed from the analysis on this basis. BMI was log-transformed to approximate univariate normality before conducting the analysis. A general linear model adjusting for age, sex, disease status (in the MEC as cases and controls were included from many disease studies) and, when available, global ancestry (from principal components analysis) was used in assessing SNP main effects. Bonferroni correction for the number of SNPs initially selected for fine mapping (9) was used to correct for multiple testing. Conditional analyses were performed by testing the association of a SNP, adjusting for an index SNP in the same multivariate regression model. SAS v. 9.2 was used for all statistical analyses.

Meta-analyses included only those studies with quantitative trait values, which consist of JHS, MEC, SPT and Maywood but not G×E to test the effects of rs3751812 and rs9941349, as well for the conditional analysis across JHS and MEC to assess whether residual signals lie in this region. We combined regression coefficients and standard errors from each study using the freely available program METAL (5). We tested for genetic heterogeneity between studies used in the meta-analysis by calculating I^2 values from Cochran's Q statistics as computed in METAL (5).

We also conducted haplotype analyses in all samples to assess whether rs3751812 or rs9941349 were more strongly associated with BMI. For each of these two SNPs, the allele associated with increased BMI was labeled 'H' and the other allele 'L.' We phased the genotype data using Haploview and discovered that there were three major haplotypes, (H–H, H–L and L–L corresponding to T–T, T–G and C–G for rs9941349 and rs3751812, respectively), accounting for 95–100% of all of the haplotypes in the five populations. The strong LD and the relative rarity of the H–L haplotype allowed us to assign haplotypes confidently based on diploid genotype data and thereby compare the H–L haplotype in turn with the H–H and L–L haplotypes. We identified groups of individuals who had one haplotype in common but differed at their other haplotype (Fig. 2). For a comparison of the H–L and H–H haplotype, there were two group comparisons: (H/H–H/L versus H/H–H/H) and (H/L–L/L versus H/L–H/L). For a comparison, the H–L and L–L haplotypes, the two group comparisons were (H/H–H/L versus H/L–H/L) and (H/L–L/L versus L/L–L/L). For each comparison, we performed a meta-analysis of the two group comparisons. If the H–L haplotype differs significantly from the H–H

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

haplotype, we could infer that the second SNP in the haplotype is more strongly correlated with the variant that influences BMI; similar conclusions about the first SNP can be drawn if H–L differs substantially from L–L.

Genotyping

Genotyping for JHS, SPT, G×E and Maywood were done at Children's Hospital, Boston using Sequenom iPLEX genotyping (Sequenom, Inc, San Diego, CA, USA). Forty-four candidate SNPs were genotyped in JHS, and key SNPs were genotyped in the SPT, G×E and Maywood studies to improve power. All passing SNPs had genotype success rates of >90%.

Genotyping of SNPs in MEC samples was conducted at USC using the TaqMan allelic discrimination assay (Applied Biosystems, Foster City, CA, USA). Approximately 5% quality control repeat samples were included to evaluate assay reproducibility. The concordance was ≥99% for all SNPs examined. All SNPs had call rates ≥97% in the MEC sample. Primers and probes and reaction conditions can be provided upon request for Sequenom and TaqMan.

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGEMENTS

We are extremely grateful to all the participants in each of the studies that contributed to this work. We would also like to thank all of the past and present staff of the Hirschhorn laboratory for comments, ideas and discussions, the past and present staff at the JHS, MEC, ERU and TMRU, UWI, Kingston, Jamaica, and the Maywood Study, who collaborated in the recruitment efforts.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the National Institutes of Health (R01DK075787 to J.N.H., R01HL074166 to X.Z., CA63464 to B.H., CA54281 to L.N.K.); and the National Cancer Institute Centers for Transdisciplinary Research on Energetics and Cancer (U54CA116848).

REFERENCES

1. Li, S. and Loos, R.J. (2008) Progress in the genetics of common obesity: size matters. *Curr. Opin. Lipidol.*, **19**, 113–121.
2. Mohlke, K.L., Boehnke, M. and Abecasis, G.R. (2008) Metabolic and cardiovascular traits: an abundance of recently identified common genetic variants. *Hum. Mol. Genet.*, **17**, R102–R108.
3. Lindgren, C.M., Heid, I.M., Randall, J.C., Lamina, C., Steinthorsdottir, V., Qi, L., Speliotes, E.K., Thorleifsson, G., Willer, C.J., Herrera, B.M. *et al.* (2009) Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. *PLoS Genet.*, **5**, e1000508.
4. Meyre, D., Delplanque, J., Chevre, J.C., Lecoeur, C., Lobbens, S., Gallina, S., Durand, E., Vatin, V., Degraeve, F., Proenca, C. *et al.* (2009) Genome-wide association study for early-onset and morbid adult obesity identifies three new risk loci in European populations. *Nat. Genet.*, **41**, 157–159.
5. Willer, C.J., Speliotes, E.K., Loos, R.J., Li, S., Lindgren, C.M., Heid, I.M., Berndt, S.I., Elliott, A.L., Jackson, A.U., Lamina, C. *et al.* (2009) Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nat. Genet.*, **41**, 25–34.
6. Thorleifsson, G., Walters, G.B., Gudbjartsson, D.F., Steinthorsdottir, V., Sulem, P., Helgadottir, A., Styrkarsdottir, U., Gretarsdottir, S., Thorlacius, S., Jonsdottir, I. *et al.* (2009) Genome-wide association yields new sequence variants at seven loci that associate with measures of obesity. *Nat. Genet.*, **41**, 18–24.
7. Scuteri, A., Sanna, S., Chen, W.M., Uda, M., Albai, G., Strait, J., Najjar, S., Nagaraja, R., Orru, M., Usala, G. *et al.* (2007) Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.*, **3**, e115.
8. Loos, R.J., Lindgren, C.M., Li, S., Wheeler, E., Zhao, J.H., Prokopenko, I., Inouye, M., Freathy, R.M., Attwood, A.P., Beckmann, J.S. *et al.* (2008) Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat. Genet.*, **40**, 768–775.
9. Frayling, T.M., Timpson, N.J., Weedon, M.N., Zeggini, E., Freathy, R.M., Lindgren, C.M., Perry, J.R., Elliott, K.S., Lango, H., Rayner, N.W. *et al.* (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*, **316**, 889–894.
10. Heard-Costa, N.L., Zillikens, M.C., Monda, K.L., Johansson, A., Harris, T.B., Fu, M., Haritunians, T., Feitosa, M.F., Aspelund, T., Eiriksdottir, G. *et al.* (2009) NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. *PLoS Genet.*, **5**, e1000539.
11. McCarthy, M.I. and Hirschhorn, J.N. (2008) Genome-wide association studies: potential next steps on a genetic journey. *Hum. Mol. Genet.*, **17**, R156–R165.
12. Cooper, R.S., Tayo, B. and Zhu, X. (2008) Genome-wide association studies: implications for multiethnic samples. *Hum. Mol. Genet.*, **17**, R151–R155.
13. Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F., Ward, R. *et al.* (2001) Linkage disequilibrium in the human genome. *Nature*, **411**, 199–204.
14. Dina, C., Meyre, D., Gallina, S., Durand, E., Korner, A., Jacobson, P., Carlsson, L.M., Kiess, W., Vatin, V., Lecoeur, C. *et al.* (2007) Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat. Genet.*, **39**, 724–726.
15. Stratigopoulos, G., Padilla, S.L., LeDuc, C.A., Watson, E., Hattersley, A.T., McCarthy, M.I., Zeltser, L.M., Chung, W.K. and Leibel, R.L. (2008) Regulation of Fto/Ftm gene expression in mice and humans. *Am. J. Physiol. Regul. Integr. Comp. Physiol.*, **294**, R1185–R1196.
16. Grant, S.F., Li, M., Bradfield, J.P., Kim, C.E., Annaiah, K., Santa, E., Glessner, J.T., Casalunovo, T., Frackelton, E.C., Otieno, F.G. *et al.* (2008) Association analysis of the FTO gene with obesity in children of Caucasian and African ancestry reveals a common tagging SNP. *PLoS ONE*, **3**, e1746.
17. Song, Y., You, N.C., Hsu, Y.H., Howard, B.V., Langer, R.D., Manson, J.E., Nathan, L., Niu, T., Tinker, L.F. and Liu, S. (2008) FTO polymorphisms are associated with obesity but not diabetes risk in postmenopausal women. *Obesity (Silver Spring)*, **16**, 2472–2480.
18. Hennig, B.J., Fulford, A.J., Sirugo, G., Rayco-Solon, P., Hattersley, A.T., Frayling, T.M. and Prentice, A.M. (2009) FTO gene variation and measures of body mass in an African population. *BMC Med. Genet.*, **10**, 21.
19. Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M. *et al.* (2007) A second generation human haplotype map of over 3.1 million SNPs. *Nature*, **449**, 851–861.
20. Andreasen, C.H., Stender-Petersen, K.L., Mogensen, M.S., Torekov, S.S., Wegner, L., Andersen, G., Nielsen, A.L., Albrechtsen, A., Borch-Johnsen, K., Rasmussen, S.S. *et al.* (2008) Low physical activity accentuates the effect of the FTO rs9939609 polymorphism on body fat accumulation. *Diabetes*, **57**, 95–101.
21. Jonsson, A., Renstrom, F., Lyssenko, V., Brito, E.C., Isomaa, B., Berglund, G., Nilsson, P.M., Groop, L. and Franks, P.W. (2009) Assessing the effect of interaction between an FTO variant (rs9939609) and physical activity on obesity in 15,925 Swedish and 2,511 Finnish adults. *Diabetologia*, **52**, 1334–1338.

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012

22. Wing, M.R., Ziegler, J., Langefeld, C.D., Ng, M.C., Haffner, S.M., Norris, J.M., Goodarzi, M.O. and Bowden, D.W. (2009) Analysis of FTO gene variants with measures of obesity and glucose homeostasis in the IRAS Family Study. *Hum. Genet.*, **125**, 615–626.
23. Taylor, H.A. Jr (2005) The Jackson Heart Study: an overview. *Ethn. Dis.*, **15**, S6-1-3.
24. Haiman, C.A., Hsu, C., de Bakker, P.I., Frasco, M., Sheng, X., Van Den Berg, D., Casagrande, J.T., Kolonel, L.N., Le Marchand, L., Hankinson, S.E. *et al.* (2008) Comprehensive association testing of common genetic variation in DNA repair pathway genes in relationship with breast cancer risk in multiple populations. *Hum. Mol. Genet.*, **17**, 825–834.
25. Deo, R.C., Patterson, N., Tandon, A., McDonald, G.J., Haiman, C.A., Ardlie, K., Henderson, B.E., Henderson, S.O. and Reich, D. (2007) A high-density admixture scan in 1,670 African Americans with hypertension. *PLoS Genet.*, **3**, e196.
26. Haiman, C.A., Patterson, N., Freedman, M.L., Myers, S.R., Pike, M.C., Waliszewska, A., Neubauer, J., Tandon, A., Schirmer, C., McDonald, G.J. *et al.* (2007) Multiple regions within 8q24 independently affect risk for prostate cancer. *Nat. Genet.*, **39**, 638–644.
27. Cooper, R., Rotimi, C., Ataman, S., McGee, D., Osotimehin, B., Kadiri, S., Muna, W., Kingue, S., Fraser, H., Forrester, T. *et al.* (1997) The prevalence of hypertension in seven populations of west African origin. *Am. J. Public Health*, **87**, 160–168.
28. Chiang, C.W.K., Gajdos, Z.K.Z., Korn, J.M., Kuruvilla, F.G., Butler, J.L., Hackett, R., Guiducci, C., Nguyen, T.T., Wilks, R., Forrester, T. *et al.* (2010) Rapid assessment of genetic ancestry in populations of unknown origin by genome-wide genotyping of pooled samples. *PLoS Genet.*, **6**, e1000866.
29. Nalls, M.A., Wilson, J.G., Patterson, N.J., Tandon, A., Zmuda, J.M., Huntsman, S., Garcia, M., Hu, D., Li, R., Beamer, B.A. *et al.* (2008) Admixture mapping of white cell count: genetic locus responsible for lower white blood cell count in the Health ABC and Jackson Heart studies. *Am. J. Hum. Genet.*, **82**, 81–87.
30. Hunter, D.J., Riboli, E., Haiman, C.A., Albanes, D., Altshuler, D., Chanock, S.J., Haynes, R.B., Henderson, B.E., Kaaks, R., Stram, D.O. *et al.* (2005) A candidate gene approach to searching for low-penetrance breast and prostate cancer genes. *Nat. Rev. Cancer*, **5**, 977–985.
31. de Bakker, P.I., Yelensky, R., Pe'er, I., Gabriel, S.B., Daly, M.J. and Altshuler, D. (2005) Efficiency and power in genetic association studies. *Nat. Genet.*, **37**, 1217–1223.
32. Barrett, J.C., Fry, B., Maller, J. and Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, **21**, 263–265.
33. Abecasis, G.R., Cherny, S.S., Cookson, W.O. and Cardon, L.R. (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.*, **30**, 97–101.
34. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.

Downloaded from http://hmg.oxfordjournals.org/ at Bibliotheque LSH Periodiques on January 4, 2012