

# Genome-Wide Association Scan Shows Genetic Variants in the *FTO* Gene Are Associated with Obesity-Related Traits

Angelo Scuteri<sup>1,2</sup>, Serena Sanna<sup>3,4</sup>, Wei-Min Chen<sup>3</sup>, Manuela Uda<sup>4</sup>, Giuseppe Albai<sup>4</sup>, James Strait<sup>2</sup>, Samer Najjar<sup>2</sup>, Ramaiah Nagaraja<sup>2</sup>, Marco Orrù<sup>4,5</sup>, Gianluca Usala<sup>4</sup>, Mariano Dei<sup>4</sup>, Sandra Lai<sup>4</sup>, Andrea Maschio<sup>4</sup>, Fabio Busonero<sup>4</sup>, Antonella Mulas<sup>4</sup>, Georg B. Ehret<sup>6</sup>, Ashley A. Fink<sup>6</sup>, Alan B. Weder<sup>7</sup>, Richard S. Cooper<sup>8</sup>, Pilar Galan<sup>9,10</sup>, Aravinda Chakravarti<sup>6</sup>, David Schlessinger<sup>2\*</sup>, Antonio Cao<sup>4</sup>, Edward Lakatta<sup>2</sup>, Gonçalo R. Abecasis<sup>3\*</sup>

1 Unità Operativa Geriatria, Istituto per la Patologia Endocrina e Metabolica, Rome, Italy, 2 Gerontology Research Center, National Institute on Aging, Baltimore, Maryland, United States of America, 3 Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, United States of America, 4 Istituto di Neurogenetica e Neurofarmacologia, Consiglio Nazionale delle Ricerche, Cittadella Universitaria di Monserrato, Monserrato, Cagliari, Italy, 5 Unità Operativa Semplice Cardiologia, Divisione di Medicina, Presidio Ospedaliero Santa Barbara, Iglesias, Italy, 6 Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, United States of America, 7 Department of Internal Medicine, University of Michigan School of Medicine, Ann Arbor, Michigan, United States of America, 8 Department of Preventive Medicine and Epidemiology, Loyola Stritch School of Medicine, Chicago, Illinois, United States of America, 9 Institut Scientifique et Technique de la Nutrition et de l'Alimentation, Paris, France, 10 INSERM, U557 (UMR INSERM/INRA/CNAM), Paris, France

**The obesity epidemic is responsible for a substantial economic burden in developed countries and is a major risk factor for type 2 diabetes and cardiovascular disease. The disease is the result not only of several environmental risk factors, but also of genetic predisposition. To take advantage of recent advances in gene-mapping technology, we executed a genome-wide association scan to identify genetic variants associated with obesity-related quantitative traits in the genetically isolated population of Sardinia. Initial analysis suggested that several SNPs in the *FTO* and *PFKF* genes were associated with increased BMI, hip circumference, and weight. Within the *FTO* gene, rs9930506 showed the strongest association with BMI ( $p = 8.6 \times 10^{-7}$ ), hip circumference ( $p = 3.4 \times 10^{-8}$ ), and weight ( $p = 9.1 \times 10^{-7}$ ). In Sardinia, homozygotes for the rare “G” allele of this SNP (minor allele frequency = 0.46) were 1.3 BMI units heavier than homozygotes for the common “A” allele. Within the *PFKF* gene, rs6602024 showed very strong association with BMI ( $p = 4.9 \times 10^{-6}$ ). Homozygotes for the rare “A” allele of this SNP (minor allele frequency = 0.12) were 1.8 BMI units heavier than homozygotes for the common “G” allele. To replicate our findings, we genotyped these two SNPs in the GenNet study. In European Americans ( $N = 1,496$ ) and in Hispanic Americans ( $N = 839$ ), we replicated significant association between rs9930506 in the *FTO* gene and BMI ( $p$ -value for meta-analysis of European American and Hispanic American follow-up samples,  $p = 0.001$ ), weight ( $p = 0.001$ ), and hip circumference ( $p = 0.0005$ ). We did not replicate association between rs6602024 and obesity-related traits in the GenNet sample, although we found that in European Americans, Hispanic Americans, and African Americans, homozygotes for the rare “A” allele were, on average, 1.0–3.0 BMI units heavier than homozygotes for the more common “G” allele. In summary, we have completed a whole genome-association scan for three obesity-related quantitative traits and report that common genetic variants in the *FTO* gene are associated with substantial changes in BMI, hip circumference, and body weight. These changes could have a significant impact on the risk of obesity-related morbidity in the general population.**

Citation: Scuteri A, Sanna S, Chen W, Uda M, Albai G, et al. (2007) Genome-wide association scan shows genetic variants in the *FTO* gene are associated with obesity-related traits. *PLoS Genet* 3(7): e115. doi:10.1371/journal.pgen.0030115

## Introduction

There is a worldwide epidemic of obesity and type 2 diabetes across all age groups, especially in industrialized countries [1]. In the United States alone, over two-thirds of the population has a body mass index (BMI) of 25 kg/m<sup>2</sup> or greater and is thus overweight [2,3]. Being overweight is a well-established risk factor for many chronic diseases, such as type 2 diabetes, hypertension, and cardiovascular events [4], and increases in BMI are associated with higher all-cause mortality [5,6]. The economic cost attributable to obesity in the United States has been estimated to be as high as \$100 billion/yr [7], and includes not only direct health care costs but also the cost of lost productivity in affected individuals [8].

Individual susceptibility to obesity is thought to be determined by interactions between an individual's genetic

**Editor:** Greg Barsh, Stanford University School of Medicine, United States of America

**Received:** April 16, 2007; **Accepted:** May 31, 2007; **Published:** July 20, 2007

A previous version of this article appeared as an Early Online Release on May 31, 2007 (doi:10.1371/journal.pgen.0030115.eor).

This is an open-access article distributed under the terms of the Creative Commons Public Domain declaration which stipulates that, once placed in the public domain, this work may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose.

**Abbreviations:** AA, African American; BMI, body mass index; CEU, Utah residents with ancestry from northern and western Europe; EA, European American; FDR, false-discovery rate; HA, Hispanic American; LD, linkage disequilibrium; YRI, Yoruba in Ibadan, Nigeria

\* To whom correspondence should be addressed. E-mail: schlessingerd@grc.nia.nih.gov (DS); goncalo@umich.edu (GRA)

© These authors contributed equally to this work.

## Author Summary

Although twin and family studies have clearly shown that genes play a role in obesity, it has proven quite difficult to identify the specific genetic variants involved. Here, we take advantage of recent technical and methodological advances to examine the role of common genetic variants on several obesity-related traits. By examining >4,000 Sardinians, we show that a specific genetic variant, rs9930506, and other nearby variants on human Chromosome 16 are associated with body mass index, hip circumference, and total body weight. The variants overlap *FTO*, a gene with poorly understood function. Further studies of the region may implicate new biological pathways affecting susceptibility to obesity. We also show that the association is not restricted to Sardinia but is also seen in independent samples of European Americans and Hispanic Americans. This finding is particularly important because obesity is associated with increased risk of cardiovascular disease and diabetes.

make-up and behavior and the environment. Thus, the increased prevalence of obesity likely reflects the exposure of genetically susceptible individuals to unhealthy secular trends in environmental and behavioral factors, such as diet and exercise [9]. In industrialized countries, between 60%–70% of the variation in obesity-related phenotypes appears to be heritable [10,11].

The traditional approach for mapping disease genes relies on linkage mapping followed by progressive fine-mapping of candidate linkage peaks [12]. While the approach has been extremely successful at identifying genes that predispose carriers to rare Mendelian disorders [13], it has met only limited success when applied to complex traits such as obesity. We have taken advantage of recent advances in genotyping technology that enable detailed assessment of entire genomes [14,15]. These advances have already allowed the identification of genes that influence quantitative variation in heart disease-related phenotypes [16] and of susceptibility genes for age-related macular degeneration [17], inflammatory bowel disease [18], and type 2 diabetes [19].

We recruited and phenotyped 6,148 individuals, male and female, ages 14–102 y, from a cluster of four towns in the Lanusei Valley in the Sardinian province of Ogliastra [20]. By studying an isolated population, we expected to increase the genetic and environmental homogeneity of our sample, increasing power [21,22]. Our cohort included >30,000 relative pairs and represents >60% of the population eligible for participation in the study; a detailed account of the family structures we examined is available elsewhere [20]. We took advantage of the relatedness among individuals in our sample to substantially reduce study costs [23]. Specifically, because our sample includes many large families, we reasoned that genotyping a relatively small number of markers in all individuals would allow us to identify shared haplotype stretches within each family. We could then genotype a subset of the individuals in each family at higher density to characterize the haplotypes in each stretch and impute missing genotypes in other individuals in the family [23,24].

For the analyses presented here, we genotyped 3,329 individuals using the Affymetrix 10,000 SNP Mapping Array and we genotyped an additional 1,412 individuals using the Affymetrix 500,000 SNP Mapping Array Set. The genotyped

individuals were selected to represent the largest families in our sample, without respect to phenotype. The high-density arrays were generally used to genotype both parents and one child (in larger sibships) or just the parents (in smaller sibships); the lower density arrays were used to genotype everyone else. Except when parents and offspring were genotyped in the same family, we tried to ensure that individuals genotyped with the high-density array were only distantly related to one another. For the 2,893 individuals that were genotyped with the 10,000 SNP arrays only, we used a modified version of the Lander-Green algorithm [25,26] to probabilistically infer missing genotypes [24]. Our approach for estimating missing genotypes is implemented in MERLIN (<http://www.sph.umich.edu/csg/abecasis/MERLIN/>) and described in detail elsewhere [24]. Our initial analysis focused on evaluating the additive effects of 362,129 SNPs (Table S1) that passed quality control checks [27,28]. The remaining SNPs failed quality checks (~2.9% of SNPs failed checks for data completeness, Hardy-Weinberg equilibrium, and Mendelian incompatibilities) or had a minor allele frequency of <5% (~25.7% of SNPs had low minor allele frequencies).

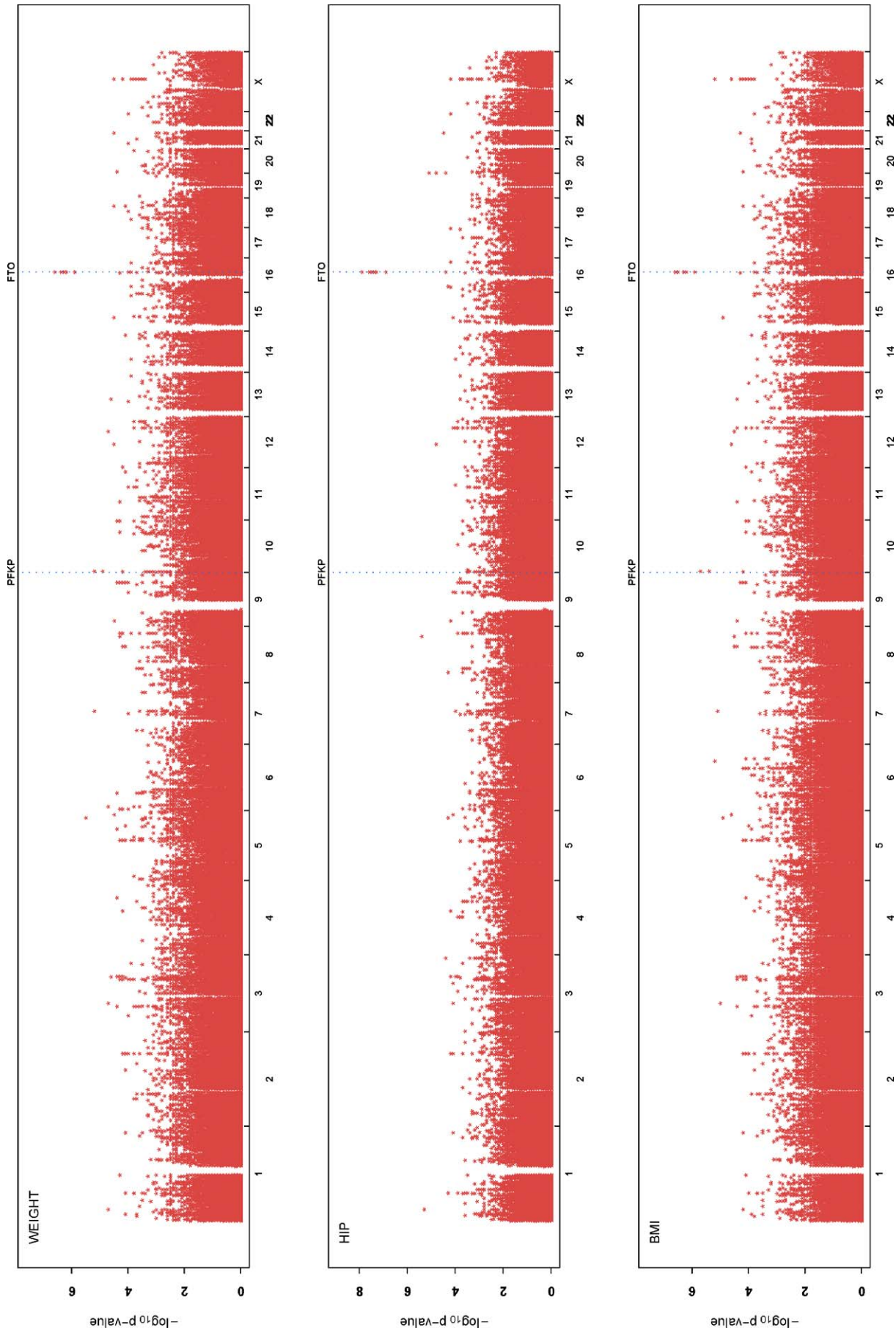
## Results

We tested 362,129 SNPs for association with three obesity-related quantitative traits (BMI, hip circumference, and weight). Height was included as a covariate in analysis of hip circumference and weight. In addition, we included age and sex as covariates in every analysis. The genomic control parameter [29] for our initial analysis of each trait ranged from 1.07 to 1.09, indicating that our estimated test statistics might be slightly inflated. This is likely due to unaccounted-for distant relationships among the sampled individuals. All results presented in our tables have been adjusted using the method of genomic control [29]. After adjustment, we observed no significant excess of results exceeding liberal significance thresholds. For example, the proportion of test statistics that were significant at  $\alpha = 0.001$  was 0.00098.

Results of our initial association analysis are summarized in Figure 1 and in Table 1. We used the false-discovery rate (FDR) to select a small set of very promising trait SNP associations for rapid replication. Using an FDR [30] of 20% highlighted a small set of SNPs for each trait. This set include the top eight SNP association results for hip circumference and weight (FDR = 0.013 and FDR = 0.16, respectively) and the top nine SNP association results for BMI (FDR = 0.20).

Eight of the SNPs listed in Table 1 overlap among the three traits. In particular, SNP rs9930506 and a cluster of nearby SNPs on Chromosome 16 show strong association with BMI ( $p = 8.6 \times 10^{-7}$ ), hip circumference ( $p = 3.4 \times 10^{-8}$ ) and weight ( $p = 9.1 \times 10^{-7}$ ). Two of the associated SNPs in the cluster, rs9939609 and rs9926289, fall within an intronic region where sequence is strongly conserved across species. For comparative purposes, using a conservative Bonferroni correction aimed at an overall type I error rate of 0.05 (one false positive per 20 genome-scans), would result in a significance threshold of  $1.4 \times 10^{-7}$ .

This cluster of SNPs on Chromosome 16 overlaps the *FTO* [31] gene, an extremely large gene whose exons span >400kb (Figure 2). KIAA1005, a gene of unknown function, also maps nearby. The *FTO* gene has not been previously implicated in obesity, but it maps to a region where linkage to BMI has been



**Figure 1.** Negative Log of  $p$ -Value for Single Marker Association Analysis with Three Obesity-Related Traits. Locations of *PFKP* and *FTO* genes are highlighted. doi:10.1371/journal.pgen.0030115.g001

**Table 1.** Markers Showing Strongest Evidence for Association

Trait	SNP	Chromosome	Position	Allele	Frequency	Effect (s.d.)	H <sup>2</sup>	p-Value	FDR	
<b>BMI</b>	rs9930506	16	52387966	A	0.54	-0.132	1.34%	8.6E-07	0.09	
	rs8050136	16	52373776	C	0.54	-0.129	1.28%	1.1E-06	0.09	
	rs1121980	16	52366748	G	0.55	-0.128	1.25%	1.4E-06	0.09	
	rs7193144	16	52368187	T	0.54	-0.127	1.24%	1.5E-06	0.09	
	rs9940128	16	52358255	G	0.55	-0.127	1.24%	1.6E-06	0.09	
	rs9939973	16	52358069	G	0.55	-0.127	1.24%	1.6E-06	0.09	
	rs9939609	16	52378028	T	0.54	-0.126	1.22%	1.8E-06	0.09	
	rs9926289	16	52378004	C	0.54	-0.123	1.17%	1.8E-06	0.15	
	rs6602024	10	3145237	G	0.88	-0.196	1.26%	4.9E-06	0.20	
	rs7907949	10	3138056	G	0.90	-0.198	1.12%	1.1E-05	0.40	
	<b>Hip</b>	rs9930506	16	52387966	A	0.54	-0.157	1.63%	3.4E-08	0.006
		rs8050136	16	52373776	C	0.54	-0.152	1.53%	6.8E-08	0.006
		rs9939973	16	52358069	G	0.55	-0.152	1.51%	8.5E-08	0.006
		rs9940128	16	52358255	G	0.55	-0.152	1.51%	8.5E-08	0.006
		rs1121980	16	52366748	G	0.55	-0.152	1.51%	8.8E-08	0.006
rs7193144		16	52368187	T	0.54	-0.150	1.48%	1.0E-07	0.006	
rs9939609		16	52378028	T	0.54	-0.149	1.45%	1.3E-07	0.007	
rs9926289		16	52378004	C	0.54	-0.145	1.38%	1.4E-07	0.013	
rs965670		8	120666727	C	0.95	0.214	0.63%	7.9E-06	0.31	
rs1188445		1	30832969	G	0.91	-0.217	0.98%	1.1E-05	0.40	
<b>Weight</b>		rs9930506	16	52387966	A	0.54	-0.118	1.34%	9.1E-07	0.10
		rs1121980	16	52366748	G	0.55	-0.116	1.28%	1.2E-06	0.10
		rs8050136	16	52373776	C	0.54	-0.115	1.27%	1.2E-06	0.10
		rs9940128	16	52358255	G	0.55	-0.115	1.26%	1.5E-06	0.10
		rs9939973	16	52358069	G	0.55	-0.115	1.25%	1.6E-06	0.10
	rs7193144	16	52368187	T	0.54	-0.114	1.25%	1.6E-06	0.10	
	rs9939609	16	52378028	T	0.54	-0.113	1.22%	2.2E-06	0.11	
	rs9926289	16	52378004	C	0.54	-0.110	1.17%	3.7E-06	0.16	
	SNP_A-2284869	5	162186248	T	0.85	-0.142	1.00%	9.5E-06	0.38	
	rs6965526	7	85760188	A	0.65	-0.105	0.96%	1.6E-05	0.58	

All positions refer to the May 2004 genome assembly. The effect is measured in standard deviation units (so that an effect of -0.132 indicates that each additional copy of the allele decreases trait values by 0.132 standard deviations on average). The H<sup>2</sup> column indicates the proportion of the trait variance that could be explained by the SNP. The FDR column estimates the proportion of false positives incurred by declaring a particular SNP significant (so that an FDR of 0.09 indicates that declaring all SNPs with a smaller p-value significant is expected to result in about 9% false-positive findings).

s.d., standard deviation.

doi:10.1371/journal.pgen.0030115.t001

reported in two previous genome-wide linkage scans (LOD = 3.2 in the Framingham Heart Study [32] and LOD = 2.2 in the families with white ancestry from the Family Blood Pressure Program [33]). Furthermore, a syndrome that results from deletion of this region of Chromosome 16q includes obesity as one of its features [34].

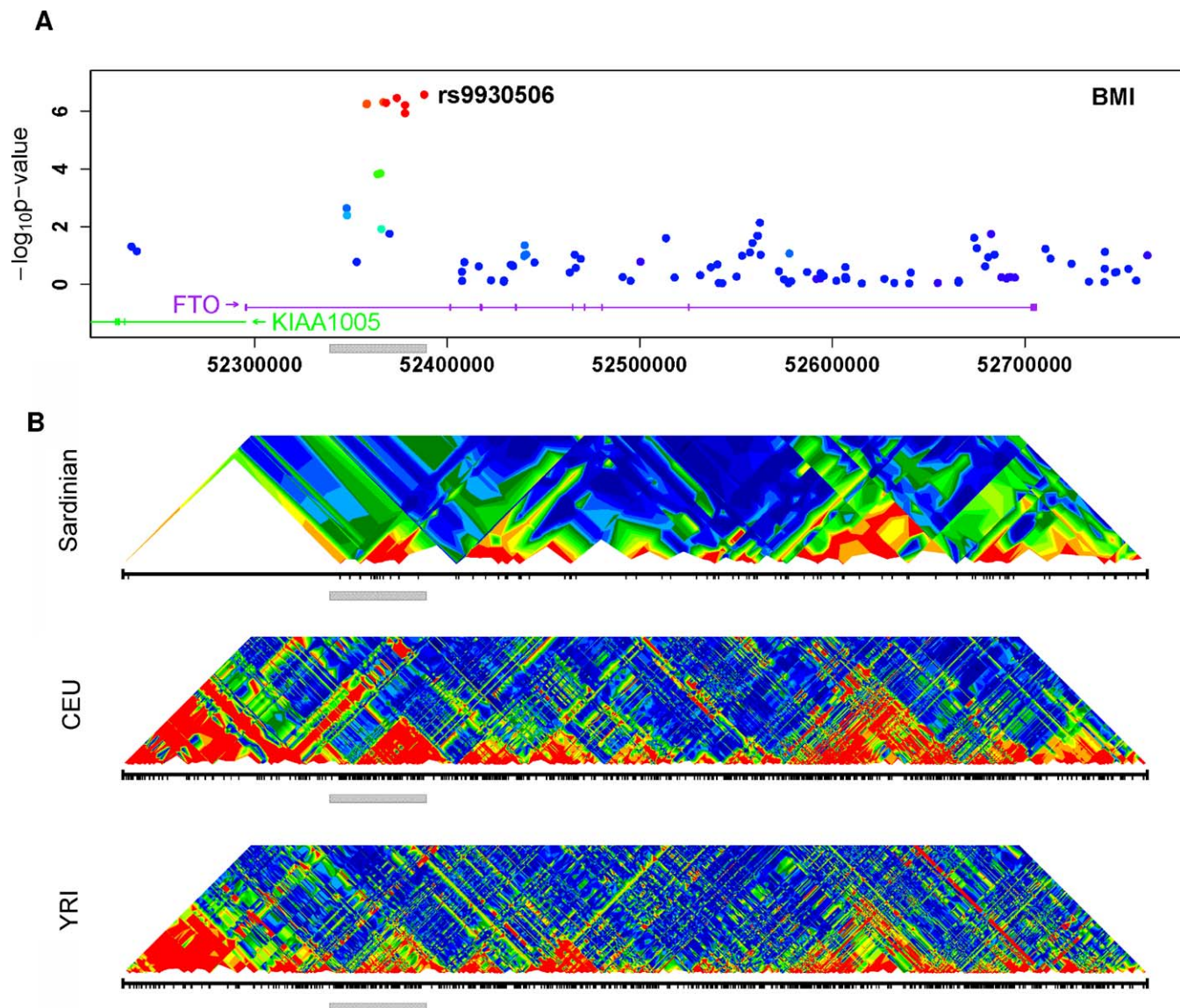
Although multiple SNPs within *FTO* show evidence for association, these do not point to multiple independently associated SNPs—rather, it is likely they are all in disequilibrium with the same causal variant(s). In a sequential analysis in which we selected the best SNP for each trait and then conditioned on it to successively select the next best SNP, only one *FTO* SNP was selected (results presented in Table S2). This result is consistent with the fact that the SNPs fall in a region of strong linkage disequilibrium, both in Sardinia and in the HapMap (Figure 2B).

Our FDR analysis of BMI selected one additional SNP outside this cluster, rs6602024 (Figure 3). This SNP maps to Chromosome 10 and shows association with BMI ( $p = 4.9 \times 10^{-6}$ ), weight ( $1.6 \times 10^{-5}$ ), and hip circumference ( $p = 0.00047$ ). The SNP maps to the platelet-type phosphofruktokinase (*PFKP*) gene, which acts as a major rate-limiting enzyme in glycolysis, converting D-fructose-6-phosphate to fructose-1,6-bisphosphate [35]. Alterations in the structure or regulation

of *PFKP* could alter the balance between glycolysis and glycogen production, ultimately leading to obesity.

Table 2 shows the phenotypic effects associated with each of the two SNPs in our sample. Because rs9930506 is more common, it shows more significant association despite being associated with smaller phenotypic effects (the two homozygotes differ, on average, by ~1.5 BMI units). A rarer polymorphism, such as rs6602024, impacts only a smaller proportion of the population and shows less significant association, despite a larger difference between homozygote means (which differ, on average, by ~2.9 BMI units). In each case, a more accurate estimate of the effect is provided by the regression model with age, sex, and (where appropriate) height as covariates. In a study, such as ours, that estimates effect sizes for many SNPs, statistical fluctuation means that some estimates will be slightly high and others will be low. SNPs that reach statistical significance are likely to include those for which effect size estimates are inflated (this is the winner's curse phenomenon) [36], and thus we proceeded to replicate our top association signals in additional large samples.

To further investigate the association between rs9930506 and rs6602024 and obesity-related traits, we genotyped these SNPs in the GenNet study [37]. The study includes a series of families recruited through probands with elevated blood



**Figure 2.** Association Results and LD Patterns in Region Surrounding the *FTO* Gene

(A) Summary of the association between SNPs in the region and BMI. The SNP showing strongest association (rs9930506) is highlighted. Other SNPs are colored according to their degree of disequilibrium with rs9930506 ranging from high (red), to intermediate (green), to low (blue). Transcripts are indicated at the bottom of the graph, with an arrow indicating transcript direction.

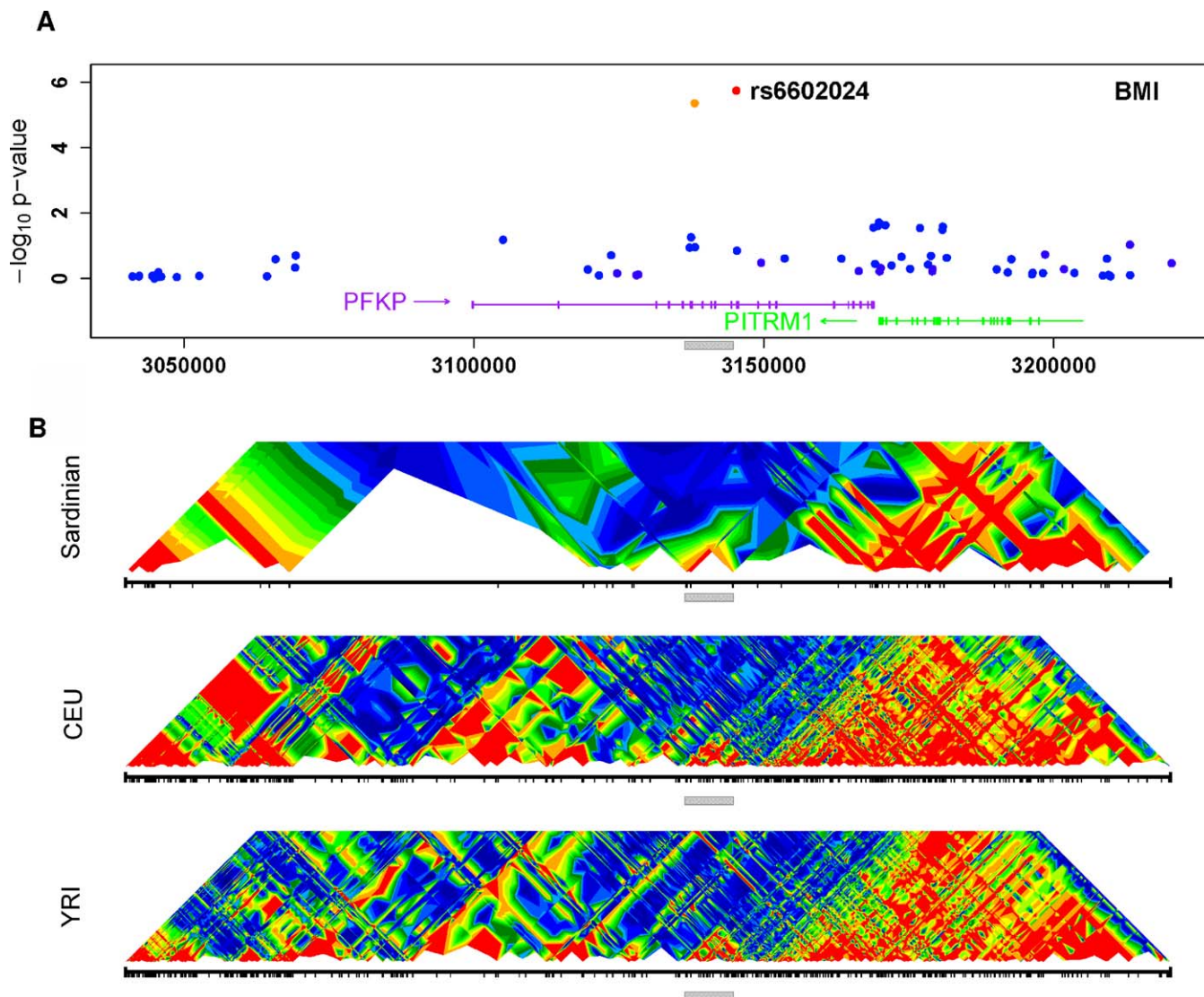
(B) Summary of the patterns of disequilibrium in the region in Sardinia and in two of the HapMap populations (CEU and YRI) [55]. The grey bar marks the region of association and facilitates comparisons between the panels.

doi:10.1371/journal.pgen.0030115.g002

pressure. The families included in this analysis comprise 3,467 individuals in total (1,101 African Americans [AA] in 369 families, 839 Hispanic Americans [HA] in 223 families, and 1,496 European Americans [EA] in 457 families). Overall, individuals in GenNet are heavier than those in our original Sardinian sample. Nevertheless, our findings strongly confirm evidence for association between rs9930506 and the three BMI-related traits (weight, hip circumference, and BMI). Specifically, rs9930506 showed association with all three traits among EA and HA in the GenNet study (meta-analysis of the EA and HA samples results in a  $p$ -value between 0.0005 and 0.001, depending on trait; see Table 3). The association is significant and in the same direction as in our original sample. The allele frequencies are also similar in all three

samples, with a frequency of 0.46 in our Sardinian sample for allele “G” of rs9930506 and of 0.44 and 0.33 in the GenNet EA and HA samples, respectively. In the GenNet sample, homozygotes for the two rs9930506 alleles differ in weight by  $\sim 1.0$  BMI units on average.

We also examined the relationship between rs9930506 and the three traits in AA, but did not observe evidence for association within that group. In AA, allele “G” of marker rs9930506 has a somewhat lower frequency of 0.21. In addition, AA show quite distinct patterns of linkage disequilibrium (LD) and thus it is not surprising that the association does not replicate. For example, in the HapMap sample of Utah residents with ancestry from northern and western Europe (CEU), the eight SNPs that show association



**Figure 3.** Association Results and LD Patterns in Region Surrounding the *PFKF* Gene

(A) Summary of the association between SNPs in the region and BMI. The SNP showing strongest association (rs6602024) is highlighted. Other SNPs are colored according to their degree of disequilibrium with rs6602024, ranging from high (orange) to low (blue). Transcripts are indicated at the bottom of the graph, with an arrow indicating transcript direction.

(B) Summary of the patterns of disequilibrium in the region in Sardinia and in two of the HapMap populations (CEU and YRI) [55]. The grey bar marks the region of association and facilitates comparisons between the panels. doi:10.1371/journal.pgen.0030115.g003

with obesity-related traits in our sample are strongly associated with each other and tag a total of 38 different variants ( $r^2 > 0.80$ ). In contrast, in the HapMap Yoruba in Ibadan, Nigeria (YRI) the strength of LD in the region is greatly reduced such that rs9930506 is not in strong LD ( $r^2 < 0.3$ ) with any of the other Chromosome 16 SNPs that show association in Sardinia.

In an attempt to fine-map association in the region, we decided to genotype the region of strong association in greater detail. In general, the study of samples from AA participants can afford an opportunity to fine-map association signals and even facilitate identification of the causal variants [38]. As noted above, a total of 38 different variants are in LD ( $r^2 > 0.8$ , HapMap CEU) with the eight SNPs that are associated with obesity-related traits in our Sardinian sample. We selected an additional seven SNPs in the region to

tag these 38 variants in samples with reduced LD. Together with rs9930506, these seven variants capture the other 30 SNPs with  $r^2 > 0.58$  (average  $r^2 = 0.87$ , HapMap YRI). The results are summarized in Table 4 and show that, whereas all the variants show association in EA and HA, none of the variants shows association in AA. One possible explanation is that obesity in AA has a different genetic architecture. Alternatively, it is possible that because some of the variants are quite common in EA and HA but rare in AA, much larger sample sizes will be required to adequately gauge their effects (for example, rs1421085 and rs3751812 have minor allele frequencies  $>0.25$  in these first two populations, but  $<0.11$  in AA).

In contrast to rs9930506, we did not replicate association between SNP rs6602024 in the *PFKF* gene and the three obesity-related traits. The “A” allele was rare in all

**Table 2.** Effects Associated with the rs9930506 and rs6602024 SNPs

Gene/SNP	Genotype Class	Frequency	Genotype Means		
			Weight (kg)	Hip (cm)	BMI (kg/m <sup>2</sup> )
<b>FTO/rs9930506</b>	G/G	0.21	69.8	101.0	27.9
	G/A	0.50	67.3	99.7	26.9
	A/A	0.29	66.7	98.3	26.4
	Estimated additive effect <sup>a</sup>	0.46 (G)	1.636	1.349	0.662
<b>PFKP/rs6602024</b>	A/A	0.01	75.6	103.2	30.6
	A/G	0.21	68.0	100.2	27.4
	G/G	0.77	67.2	99.2	26.7
	Estimated additive effect <sup>a</sup>	0.12 (A)	2.194	1.296	0.906

<sup>a</sup>Effect was estimated in a variance component model including sex, age, age<sup>2</sup>, and (except for BMI) height as covariates.  
doi:10.1371/journal.pgen.0030115.t002

populations, with a frequency of 0.12 in our Sardinian sample, 0.11 in the HA and EA GenNet subsamples and 0.25 in the AA GenNet subsample. The results are summarized in Table 5 and show that, although homozygotes for the rare “A” allele at rs6602024 were on average heavier by ~1.0–3.0 BMI units than homozygotes for the “G” allele at the SNP, these homozygotes were rare and, overall, there was no significant association. Corroborating evidence that *PFKP* and rs6602024 are associated with BMI is the observation that a region of ~120 kb including the *Pfkfb* gene has been implicated in a mouse model of obesity [39] (see Discussion). A definite assessment of the impact of *PFKP* on obesity-related quantitative traits in human populations will likely require examination of much larger sample sizes.

Our genotyping results also hint at the possible importance in Sardinia of other genes previously investigated as candidates influencing obesity and related traits (Tables S3–S5). When we evaluated evidence for association across previously identified candidate genes, we observed a small excess of nominally significant *p*-values. (We tested 837 candidate SNPs in 74 candidate genes against three traits

and found that 145 tests were significant at  $p < 0.05$ , corresponding to 5.8% of the 2,511 tests. We observed no such excess when the whole genome was considered.) Among the interesting candidates that show association in our sample are the two adiponectin receptor genes [40] *ADIPOR1* (best single SNP *p*-value = 0.013, 0.027, and 0.016 for BMI, hip circumference, and weight) and *ADIPOR2* (best *p*-values = 0.018, 0.019, 0.013) and the lipoprotein lipase gene, *LPL* [41] (best *p*-values = 0.014, 0.006, 0.018). Nevertheless, all the association signals observed in any of these previous candidate genes are far less significant than those in *FTO* or *PFKP*.

## Discussion

*FTO* association provides an example of how genome-wide association studies can point to previously unsuspected candidate genes. An interstitial deletion overlapping the region produces human syndromic obesity [34] and a hint that the gene might be involved in stress responses stems from the observation that it is down-regulated when the heat

**Table 3.** Replication of Association between rs9930506/*FTO*/G Allele and Obesity-Related Traits

Sample	Average BMI			BMI		Weight		Hip	
	A/A	A/G	G/G	Associated Effect (s.d.)	<i>p</i> -Value	Associated Effect (s.d.)	<i>p</i> -Value	Associated Effect (s.d.)	<i>p</i> -Value
Sardinia	26.4(4.1)	26.9(4.5)	27.9(5.1)	+0.132	0.00000086	+0.118	0.00000091	+0.157	0.00000034
EA <sup>a</sup>	28.9(6.6)	28.6(6.2)	30.5(7.0)	+0.080	0.03	+0.071	0.03	+0.106	0.006
HA <sup>a</sup>	28.4(6.0)	28.8(6.3)	29.7(7.1)	+0.122	0.03	+0.104	0.03	+0.096	0.08
AA <sup>a</sup>	30.1(8.1)	30.8(8.5)	30.4(8.7)	−0.004	0.9	−0.006	0.5	−0.009	0.4
Combined GenNet EA and HA (but excluding Sardinia)	—	—	—	—	0.001	—	0.001	—	0.0005
Combined GenNet EA, HA, and AA (but excluding Sardinia)	—	—	—	—	0.007	—	0.015	—	0.012

The table summarizes the effect of allele “G” for rs9930506 in the original sample and in each of the replication samples examined. The allele has a frequency of 0.46 in our Sardinian sample and of 0.44, 0.33, and 0.21 in the EA, HA, and AA GenNet subsamples, respectively. Average BMI and standard deviation (inside parenthesis) is reported for each genotype class. The effect size (in standard deviation units, after normalizing the data and adjusting for covariates) and significance of the observed association is reported for each sample. Note that, although the effect observed in Sardinia is slightly larger than in the replication samples when measured in standard deviation units, the replication samples also appear to show greater variability in trait values overall (as seen by the larger standard deviations within each genotype class). The combined *p*-values are one-sided and test for an effect that is in the same direction as that seen in Sardinia.

<sup>a</sup>GenNet Family Blood Pressure Study.

s.d., standard deviation.

doi:10.1371/journal.pgen.0030115.t003

**Table 4.** Fine-Mapping Results for *FTO* Region in GenNet Sample

Trait	Marker	Allele	EA BMI				HA BMI				AA BMI			
			Frequency	Effect	Standard Error	p-Value	Frequency	Effect	Standard Error	p-Value	Frequency	Effect	Standard Error	p-Value
BMI	rs1421085	T	0.58	-0.122	0.039	0.002	0.75	-0.155	0.055	0.005	0.92	-0.062	0.084	0.5
	rs9937053	G	0.58	-0.103	0.038	0.007	0.65	-0.095	0.051	0.06	0.58	0.050	0.047	0.3
	rs8043757	A	0.59	-0.120	0.039	0.002	0.71	-0.118	0.053	0.03	0.57	0.033	0.045	0.5
	rs3751812	G	0.59	-0.115	0.039	0.003	0.74	-0.142	0.055	0.01	0.89	-0.023	0.076	0.8
	rs9923233	G	0.55	-0.122	0.039	0.002	0.71	-0.112	0.053	0.03	0.51	0.061	0.044	0.2
	rs9932754	T	0.55	-0.088	0.038	0.02	0.71	-0.117	0.052	0.03	0.79	-0.045	0.051	0.4
Hip	rs9922619	G	0.56	-0.089	0.038	0.02	0.68	-0.107	0.052	0.04	0.82	-0.002	0.056	1
	rs1421085	T	0.58	-0.143	0.039	0.0002	0.75	-0.068	0.049	0.03	0.92	-0.025	0.082	0.8
	rs9937053	G	0.58	-0.125	0.038	0.0009	0.65	-0.117	0.053	0.2	0.58	0.038	0.045	0.4
	rs8043757	A	0.59	-0.143	0.038	0.0002	0.71	-0.002	0.052	0.1	0.57	0.019	0.044	0.7
	rs3751812	G	0.59	-0.138	0.039	0.0004	0.74	-0.105	0.053	0.05	0.89	-0.011	0.074	0.9
	rs9923233	G	0.55	-0.145	0.038	0.00014	0.71	-0.081	0.052	0.12	0.51	0.047	0.043	0.3
Weight	rs9932754	T	0.55	-0.111	0.038	0.003	0.71	-0.096	0.051	0.06	0.79	-0.037	0.05	0.5
	rs9922619	G	0.56	-0.113	0.038	0.003	0.68	-0.083	0.051	0.1	0.82	-0.004	0.054	0.9
	rs1421085	T	0.58	-0.104	0.035	0.003	0.75	-0.141	0.049	0.004	0.92	-0.059	0.084	0.5
	rs9937053	G	0.58	-0.089	0.034	0.008	0.65	-0.087	0.045	0.06	0.58	0.051	0.047	0.3
	rs8043757	A	0.59	-0.100	0.034	0.003	0.71	-0.107	0.048	0.03	0.57	0.031	0.046	0.5
	rs3751812	G	0.59	-0.095	0.034	0.006	0.74	-0.128	0.049	0.01	0.89	-0.018	0.076	0.8
Weight	rs9923233	G	0.55	-0.101	0.034	0.003	0.71	-0.103	0.048	0.03	0.51	0.061	0.045	0.2
	rs9932754	T	0.55	-0.074	0.034	0.03	0.71	-0.105	0.047	0.02	0.79	-0.044	0.051	0.4
	rs9922619	G	0.56	-0.075	0.034	0.03	0.68	-0.095	0.047	0.04	0.82	0.001	0.056	1

Additional SNPs in the region that shows strong association in Sardinia and in the HapMap CEU sample were examined in an attempt to fine-map association within the GenNet AA sample. The frequency of each allele examined, its associated effect (in standard deviation units), and corresponding *p*-value are summarized.  
doi:10.1371/journal.pgen.0030115.t004

shock response transcription factor *Hif1* is inhibited [42]. Because the gene has no recognizable functional domains and has not been studied in detail in experimental models, no putative function can be currently imputed. The fact that *FTO* is associated not only with BMI but also with hip circumference and weight is consistent with previous analyses of heritability in our cohort [20]. The analyses suggested that 80% of the genetic variance of these traits is determined by common loci (individually, the traits have heritabilities between ~30%–45%). Although the three traits examined here are correlated (all pairwise correlations were >0.73), it is important to note that apart from the SNPs that overlap *FTO*,

other strongly associated SNPs differed among the traits (see Tables 1 and S2).

In contrast to *FTO*, *PFKP* is a critical enzyme within the well-studied pathway of glucose metabolism but, to our knowledge, has not been previously implicated in obesity in humans. *PFKP* is one of the three phosphofructokinase subunit proteins that show partially overlapping patterns of expression and form hetero-tetramers in diverse cells and tissues. The subunits are encoded by different genes. One form is highly expressed in muscle (*PFKM*); a second, in liver (*PFKL*); and the third, *PFKP*, is the only form in platelets and is also highly expressed in subregions of the brain [42]. None

**Table 5.** Replication of Association between rs6602024/*PFKP*/A Allele and Obesity-Related Traits

Sample	Average BMI			BMI		Weight		Hip	
	G/G	A/G	A/A	Effect	p-Value	Effect	p-Value	Effect	p-Value
Sardinia	26.7(4.5)	27.3(4.6)	30.6(4.9)	0.196	.0000049	0.167	.000016	0.113	0.00047
EA <sup>a</sup>	28.9 (6.4)	29.5(6.9)	32.2(6.8)	0.071	0.3	0.075	0.2	0.083	0.2
HA <sup>a</sup>	28.8(6.4)	28.4(6.0)	31.2(7.8)	0.000	0.9	0.008	1	-0.034	0.6
AA <sup>a</sup>	29.9(8.3)	30.9(8.4)	30.9(8.4)	0.055	0.3	0.042	0.4	0.071	0.2
Combined GenNet EA and HA (but excluding Sardinia)	—	—	—	—	0.225	—	0.152	—	0.238
Combined GenNet EA, HA and AA (but excluding Sardinia)	—	—	—	—	0.094	—	0.094	—	0.098

The table is analogous to Table 3, but focuses on allele “A” for the rs6602024 SNP. The allele has a frequency of 0.12 in our Sardinian sample, 0.11 in the EA and HA GenNet subsamples, and 0.25 in the AA GenNet subsample. We did not find a significant additive effect for this allele in the replication samples. However, note that homozygotes for the “A” allele are consistently heavier than individuals with a “G/G” genotype. There are 72 such homozygotes in the replication sample (nine EA, ten HA, and 53 AA).

<sup>a</sup>GenNet Family Blood Pressure Study.  
doi:10.1371/journal.pgen.0030115.t005



of the forms has been previously implicated in obesity in humans, although *PFKM* is mutated in some cases of impaired glycogen synthesis (glycogen storage disease VII; see Online Mendelian Inheritance in Man, <http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=232800>) [35]. It is of considerable interest that compared to the other isozymes, *PFKP* has lower affinity for fructose-6-phosphate and decreased inhibition by ATP [43]. Consequently, *PFKP* is the most stringently regulated, responding to small changes at typical metabolic levels of effectors [44]. Genetic variants in the enzyme could thus adjust the rate of glycolysis, shifting the balance of metabolism between gluconeogenesis and glucose assimilation—a possible step in the etiology of obesity. Additionally, it is intriguing that in mice a locus associated with obesity has been mapped to a 127-kb interval that includes *Pfkfb* [39]. The mouse locus shows strong evidence of interaction with diet, with different effects in mice fed high-fat and low-fat diets. One possibility is that greater homogeneity of diet in Sardinia facilitated mapping, but made replication in other populations more difficult.

How significant are the associations observed? The replication of the *FTO* association in two different populations indicates that it is likely important not only in Sardinia, but in many different populations. In contrast, the failure to replicate the *PFKP* association in other populations suggests that (a) the association we identified may refer to rarer, population-specific variants; (b) the effects of the locus may depend on genetic or environmental background; or (c) the association identified in our original sample is due to the statistical fluctuations inherent in testing hundreds of thousands of SNPs. As for the public health impact of the observed associations, a 1-unit increment in BMI has been associated with an 8% increase in the risk of coronary heart disease [45] and excess weight in middle life is associated with increased overall risk of death [46]. Thus, the alleles reported here, which shift BMI by 1–1.5 units, have effects that are not only statistically significant but could also have important health consequences. Furthermore, apart from the direct contribution of these gene variants, they provide an entrée to the analysis of genes and pathways that contribute additionally, and open new routes to possible eventual intervention.

*Note:* After completing this manuscript, we became aware of additional evidence that supports our report of association between *FTO* and obesity-related traits. First, genotyping of 1,780 individuals from the SUVIMAX study [47,48] replicated association of allele rs9930506 with increased BMI ( $p = 0.006$ ). Combined evidence from SUVIMAX, GenNet EA, and GenNet HA resulted in a replication  $p$ -value of  $1.5 \times 10^{-5}$ . In addition, two other large independent studies also show association of SNPs in *FTO* with increased BMI [49,50]. Genotyping of the SUVIMAX sample did not provide evidence for association between rs6602024 and BMI.

## Materials and Methods

**Study sample.** We recruited and phenotyped 6,148 individuals, male and female, ages 14–102 y, from a cluster of four towns in the Lanusei Valley [20]. During physical examination of each individual, a blood sample was collected (for DNA extraction) and anthropometric traits were recorded. Here, we report analyses of hip circumference, weight, and the derived quantity BMI (which is calculated from a combination of height and weight). Genotyping was carried out using the Affymetrix 10K and 500K chips (<http://affymetrix.com/>) using

standard protocols. Summary assessments of genotype data quality are provided in the Results section and in Table S1.

To follow up on SNPs rs9930506 and rs6602024, we genotyped and examined the association between these two SNPs and BMI, hip circumference, and body weight in the GenNet study. The study comprises 3,467 individuals in total, recruited between 1995 and 2004 (1,101 AA, 839 HA, and 1,496 EA). Individuals were recruited at two field centers: EA were recruited from Tecumseh, Michigan, and AA and HA were recruited from Maywood, Illinois. Participants were recruited from families starting from a proband with high blood pressure. DNA was available for 3,205 individuals (968 AA, 824 HA, and 1471 EA). SNP genotyping was performed using the 5'-nuclease-based assay (TaqMan; ABI, <http://www.appliedbiosystems.com/>) analyzed on an ABI Prism 7900 Real Time PCR System. Within each ethnic group, genotype completeness rates exceed 98% and there was no evidence for deviation from Hardy-Weinberg equilibrium ( $p > 0.05$ ).

**Statistical analysis.** To ensure adequate control of type I error rates, we applied an inverse normal transformation to each trait prior to analysis [20]. The inverse normal transformation reduces the impact of outliers and deviations from normality on statistical analysis. The transformation involves ranking all available phenotypes, transforming these ranks into quantiles and, finally, converting the resulting quantiles into normal deviates. We included sex, age, and age<sup>2</sup> as covariates in all analysis. Height was significantly associated with weight and hip circumference and was included as an additional covariate in analysis of those traits. We fitted a simple regression model to each trait and used a variance component approach to account for correlation between different observed phenotypes within each family. For individuals who had genotype data available, we coded genotypes as 0, 1, and 2 (depending on the number of copies of the allele being tested). For individuals with missing genotype data, we used the Lander-Green algorithm to estimate an expected genotype score (between 0 and 2) for each individual [24]. Briefly, to estimate each genotype score we first calculate the likelihood of the observed genotype data. Then, we instantiate each missing genotype to a specific value and update the pedigree likelihood. The ratio of the two likelihoods gives a posterior probability that the instantiated genotype is true, conditional on all available data. Due to computational constraints, we divided large pedigrees into subunits with “bit-complexity” of 19 or less (typically, 20–25 individuals) before estimating missing genotypes.

Our analytical approach considers all observed or estimated genotypes (rather than focusing on alleles transmitted from heterozygous parents) and thus is not immune to effects of population stratification. In homogenous populations, this type of analysis is expected to be more powerful [51,52]. To adjust for the effects of population structure and cryptic relatedness among sampled individuals, we used the genomic control method to adjust our test statistics for each trait separately [29]. FDRs were calculated with R's `p.adjust()` procedure using the method of Benjamini and Hochberg [30]. Since the initial analysis often identified clusters of nearby SNPs that all showed similar levels of association, we also carried out a sequential stepwise analysis. In this analysis, we selected the best SNP for each trait, and then conditioned on it to successively select the next best SNP. This sequential analysis can help identify regions with multiple independent association signals. The stepwise analysis was repeated for five rounds.

**Candidate gene analysis.** We selected 74 candidate genes previously tested for association with obesity in humans [53]. For each gene, we first evaluated the ability of the Affymetrix SNPs to tag common SNPs (MAF > 0.05) within  $\pm 5$  kb of the gene ( $r^2 > 0.50$  or  $r^2 > 0.80$ ) using the HapMap CEU database [54]. We then evaluated evidence for association using all Affymetrix SNPs within each gene as well as neighboring Affymetrix SNPs that could be used to improve coverage ( $r^2 > 0.5$ ). For each gene, we report coverage statistics as well as the SNP that showed strongest evidence for association.

We selected 74 genes that were previously targeted in associations studies aiming to identify genetic determinants of obesity in humans [53]: *ACE*, *ACTN*, *ADIPOQ*, *ADIPOR1*, *ADIPOR2*, *ADRB1*, *ADRB2*, *AGER*, *AHSG*, *APOA2*, *APOA4*, *APOA5*, *AR*, *BDNF*, *CASQ1*, *COL1A1*, *COMT*, *CRP*, *CYP11B2*, *DIO1*, *ENPP1*, *ESR1*, *ESR2*, *FABP2*, *FOXO2*, *GAD2*, *GFPT1*, *GHRHR*, *GNAS*, *GNB3*, *GPR40*, *H6PD*, *HSD11B1*, *HTR2C*, *ICAM1*, *IGF1*, *IGF2*, *IL6*, *IL6R*, *KCNJ11*, *KL*, *LEP*, *LEPR*, *LIPC*, *LPL*, *LTA*, *MC4R*, *MCHR1*, *MKKS*, *MTHFR*, *MTTP*, *NMB*, *NOS3*, *NPY*, *NPY2R*, *NR0B2*, *NTRK2*, *PARD6A*, *PLIN*, *PPARG*, *PPARGCIA1*, *PRDM2*, *PTPN1*, *PYY*, *RETN*, *SCD*, *SELE*, *SERPINE1*, *TAS2R38*, *TNF*, *UCP1*, *UCP2*, *UCP3*, and *VDR*. We did not consider genes associated with drug-induced body weight gain or mitochondrial genes [53].

The following genes have previously been investigated for their

role in obesity and related traits but are not well tagged by SNPs in the Affymetrix array: *ADRB3*, *DRD4*, *INS*, and *APOE*.

## Supporting Information

### Table S1. Genotype Data for Sardinian Cohort

Found at doi:10.1371/journal.pgen.0030115.st001 (47 KB DOC).

### Table S2. Results of Stepwise Analysis to Identify Independent Risk Alleles

To generate this table, we first sought the most significantly associated allele in the genome. We then added this allele to our baseline model and repeated our genome scan to identify the next associated SNPs.

Found at doi:10.1371/journal.pgen.0030115.st002 (55 KB DOC).

### Table S3. Tag SNP That Shows Strongest Association with BMI for Each Previously Identified Candidate Gene

The first column indicates the name of a previously identified candidate. The second column indicates the number of SNPs in our Affymetrix arrays that are either in the gene or constitute the best available tag ( $r^2 > 0.5$ ) for a genic SNP. The next column indicates the number of HapMap SNPs within  $\pm 5$  kb of the gene and the proportion of these that are covered at  $r^2 > 0.50$  or  $r^2 > 0.80$ . The next columns indicate the SNP that showed strongest association in our analysis, the  $p$ -value, the tested allele and its frequency, and the estimated additive effect. The last column corresponds to the FDR incurred when all tested SNPs are considered and this test is declared significant.

Found at doi:10.1371/journal.pgen.0030115.st003 (162 KB DOC).

### Table S4. Tag SNP That Shows Strongest Association with Hip Circumference for Each Previously Identified Candidate Gene

The first column indicates the name of a previously identified candidate. The second column indicates the number of SNPs in our Affymetrix arrays that are either in the gene or constitute the best available tag ( $r^2 > 0.5$ ) for a genic SNP. The next column indicates the number of HapMap SNPs within  $\pm 5$  kb of the gene and the proportion of these that are covered at  $r^2 > 0.50$  or  $r^2 > 0.80$ . The next columns indicate the SNP that showed strongest association in our analysis, the  $p$ -value, the tested allele and its frequency, and the estimated additive effect. The last column corresponds to the FDR incurred when all tested SNPs are considered and this test is declared significant.

Found at doi:10.1371/journal.pgen.0030115.st004 (163 KB DOC).

### Table S5. Tag SNP That Shows Strongest Association with Weight for Each Previously Identified Candidate Gene

The first column indicates the name of a previously identified candidate. The second column indicates the number of SNPs in our

Affymetrix arrays that are either in the gene or constitute the best available tag ( $r^2 > 0.5$ ) for a genic SNP. The next column indicates the number of HapMap SNPs within  $\pm 5$  kb of the gene and the proportion of these that are covered at  $r^2 > 0.50$  or  $r^2 > 0.80$ . The next columns indicate the SNP that showed strongest association in our analysis, the  $p$ -value, the tested allele and its frequency, and the estimated additive effect. The last column corresponds to the FDR incurred when all tested SNPs are considered and this test is declared significant.

Found at doi:10.1371/journal.pgen.0030115.st005 (163 KB DOC).

## Acknowledgments

We warmly thank Monsignore Piseddu, Bishop of Ogliastra; Mayor Enrico Lai and his administration in Lanusei for providing and furnishing the clinic site; the mayors of Ilbono, Arzana, and Elini; the head of the local Public Health Unit Ar1; and the residents of the towns for their volunteerism and cooperation. We also thank Harold Spurgeon and Paul Pullen for invaluable help with equipment and readings, and Michele Evans and Dan Longo for helpful discussions. Finally, we thank Mark Lathrop for genotyping of the SUVIMAX samples and for critical advice.

**IRB Approval.** The study, including the protocols for subject recruitment and assessment, the informed consent for participants (and Assent Forms for those 14–18 y old), and the overall analysis plan were reviewed and approved by IRB boards for the Istituto di Neurogenetica e Neurofarmacologia (INN; Cagliari, Italy), the MedStar Research Institute (responsible for intramural research at the National Institutes of Aging, Baltimore, Maryland, United States of America), and for the University of Michigan (Ann Arbor, Michigan, United States of America).

**Author contributions.** S. Najjar, G. B. Ehret, A. Chakravarti, D. Schlessinger, A. Cao, E. Lakatta, and G. R. Abecasis conceived and designed the experiments. M. Uda, G. Albai, M. Orrú, G. Usala, M. Dei, S. Lai, A. Maschio, F. Busonero, A. Mulas, G. B. Ehret, and A. G. Fink performed the experiments. S. Sanna, W.-M. Chen and G. Albai analyzed the data. S. Najjar, R. Nagaraja, A. B. Weder, R. S. Cooper, P. Galan, and A. Cao contributed reagents/materials/analysis tools. A. Scuteri, S. Sanna, J. Strait, D. Schlessinger, E. Lakatta, and G. R. Abecasis wrote the paper.

**Funding.** This work was supported by the Intramural Research Program of the National Institute on Aging, NIH. The SardinIA (“Progenia”) team was supported by Contract NO1-AG-1-2109 from the National Institute on Aging. The efforts of SS, WC, and GRA, were supported in part by contract 263-MA-410953 from the National Institute on Aging to the University of Michigan and by research grants HG02651 and HL084729 from the National Institutes of Health (to GRA).

**Competing interests.** The authors have declared that no competing interests exist.

## References

- Zimmet P, Alberti KG, Shaw J (2001) Global and societal implications of the diabetes epidemic. *Nature* 414: 782–787.
- Kuczmarski RJ, Flegal KM, Campbell SM, Johnson CL (1994) Increasing prevalence of overweight among US adults. *The National Health and Nutrition Examination Surveys, 1960 to 1991. JAMA* 272: 205–211.
- Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, et al. (2006) Prevalence of overweight and obesity in the United States, 1999–2004. *JAMA* 295: 1549–1555.
- Must A, Spadano J, Coakley EH, Field AE, Colditz G, et al. (1999) The disease burden associated with overweight and obesity. *Jama* 282: 1523–1529.
- Calle EE, Thun MJ, Petrelli JM, Rodriguez C, Heath CW Jr. (1999) Body-mass index and mortality in a prospective cohort of U.S. adults. *N Engl J Med* 341: 1097–1105.
- Katzmarzyk PT, Janssen I, Ardern CI (2003) Physical inactivity, excess adiposity and premature mortality. *Obes Rev* 4: 257–290.
- Wolf AM, Colditz GA (1998) Current estimates of the economic cost of obesity in the United States. *Obes Res* 6: 97–106.
- Allison DB, Zannolli R, Narayan KM (1999) The direct health care costs of obesity in the United States. *Am J Public Health* 89: 1194–1199.
- Hill JO, Peters JC (1998) Environmental contributions to the obesity epidemic. *Science* 280: 1371–1374.
- Comuzzie AG, Allison DB (1998) The search for human obesity genes. *Science* 280: 1374–1377.
- Segal NL, Allison DB (2002) Twins and virtual twins: Bases of relative body weight revisited. *Int J Obes Relat Metab Disord* 26: 437–441.
- Collins FS (1992) Positional cloning: Let's not call it reverse anymore. *Nat Genet* 1: 3–6.
- Botstein D, Risch N (2003) Discovering genotypes underlying human phenotypes: Past successes for Mendelian disease, future approaches for complex disease. *Nat Genet* 33 Suppl: 228–237.
- Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 6: 95–108.
- Wang WY, Barratt BJ, Clayton DG, Todd JA (2005) Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet* 6: 109–118.
- Arking DE, Pfeufer A, Post W, Kao WH, Newton-Cheh C, et al. (2006) A common genetic variant in the NOS1 regulator NOS1AP modulates cardiac repolarization. *Nat Genet* 38: 644–651.
- Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, et al. (2005) Complement factor H polymorphism in age-related macular degeneration. *Science* 308: 385–389.
- Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, et al. (2006) A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 314: 1461–1463.
- Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445: 881–885.
- Pilia G, Chen WM, Scuteri A, Orru M, Albai G, et al. (2006) Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* 2: e132. doi:10.1371/journal.pgen.0020132
- Wright AF, Carothers AD, Pirastu M (1999) Population choice in mapping genes for complex diseases. *Nat Genet* 23: 397–404.

22. Peltonen L, Palotie A, Lange K (2000) Use of population isolates for mapping complex traits. *Nat Rev Genet* 1: 182–190.
23. Burdick JT, Chen WM, Abecasis GR, Cheung VG (2006) In silico method for inferring genotypes in pedigrees. *Nat Genet* 38: 1002–1004.
24. Chen WM, Abecasis GR (2007) Family Based Association Tests for Genome Wide Association Scans. *Am J Hum Genet*. In press.
25. Lander ES, Green P (1987) Construction of multilocus genetic linkage maps in humans. *Proc Natl Acad Sci U S A* 84: 2363–2367.
26. Abecasis GR, Cherny SS, Cookson WO, Cardon LR (2002) Merlin—rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 30: 97–101.
27. Wigginton JE, Abecasis GR (2005) PEDSTATS: Descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics*.
28. Wigginton JE, Cutler DJ, Abecasis GR (2005) A note on exact tests of Hardy-Weinberg equilibrium. *Am J Hum Genet* 76: 887–883.
29. Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics* 55: 997–1004.
30. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Stat Methodol* 57: 289–300.
31. Peters T, Ausmeier K, Ruther U (1999) Cloning of Fatso (Fto), a novel gene deleted by the Fused toes (Ft) mouse mutation. *Mamm Genome* 10: 983–986.
32. Geller F, Dempfle A, Gorg T (2003) Genome scan for body mass index and height in the Framingham Heart Study. *BMC Genet* 4 Suppl 1: S91.
33. Wu X, Cooper RS, Borecki I, Hanis C, Bray M, et al. (2002) A combined analysis of genomewide linkage scans for body mass index from the National Heart, Lung, and Blood Institute Family Blood Pressure Program. *Am J Hum Genet* 70: 1247–1256.
34. Stratakis CA, Lafferty A, Taymans SE, Gafni RI, Meck JM, et al. (2000) Anisomastia associated with interstitial duplication of Chromosome 16, mental retardation, obesity, dysmorphic facies, and digital anomalies: Molecular mapping of a new syndrome by fluorescent in situ hybridization and microsatellites to 16q13 (D16S419–D16S503). *J Clin Endocrinol Metab* 85: 3396–3401.
35. Vora S (1983) Isozymes of human phosphofruktokinase: Biochemical and genetic aspects. *Isozymes Curr Top Biol Med Res* 11: 3–23.
36. Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN (2003) Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. *Nat Genet* 33: 177–182.
37. Thiel BA, Chakravarti A, Cooper RS, Luke A, Lewis S, et al. (2003) A genome-wide linkage analysis investigating the determinants of blood pressure in whites and African Americans. *Am J Hypertens* 16: 151–153.
38. McKenzie CA, Abecasis GR, Keavney B, Forrester T, Ratcliffe PJ, et al. (2001) Trans-ethnic fine mapping of a quantitative trait locus for circulating angiotensin I-converting enzyme (ACE). *Hum Mol Genet* 10: 1077–1084.
39. Ehrlich TH, Hrbek T, Kenney-Hunt JP, Pletscher LS, Wang B, et al. (2005) Fine-mapping gene-by-diet interactions on Chromosome 13 in a LG/J x SM/J murine model of obesity. *Diabetes* 54: 1863–1872.
40. Damcott CM, Ott SH, Pollin TI, Reinhart LJ, Wang J, et al. (2005) Genetic variation in adiponectin receptor 1 and adiponectin receptor 2 is associated with type 2 diabetes in the Old Order Amish. *Diabetes* 54: 2245–2250.
41. Fisher RM, Maily F, Peacock RE, Hamsten A, Seed M, et al. (1995) Interaction of the lipoprotein lipase asparagine 291→serine mutation with body mass index determines elevated plasma triacylglycerol concentrations: a study in hyperlipidemic subjects, myocardial infarction survivors, and healthy adults. *J Lipid Res* 36: 2104–2112.
42. Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, et al. (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc Natl Acad Sci U S A* 101: 6062–6067.
43. Nakajima H, Raben N, Hamaguchi T, Yamasaki T (2002) Phosphofruktokinase deficiency; past, present and future. *Curr Mol Med* 2: 197–212.
44. Hannemann A, Jandrig B, Gaunitz F, Eschrich K, Bigl M (2005) Characterization of the human P-type 6-phosphofruktokinase gene promoter in neural cell lines. *Gene* 345: 237–247.
45. Li TY, Rana JS, Manson JE, Willett WC, Stampfer MJ, et al. (2006) Obesity as compared with physical activity in predicting risk of coronary heart disease in women. *Circulation* 113: 499–506.
46. Adams KF, Schatzkin A, Harris TB, Kipnis V, Mouw T, et al. (2006) Overweight, obesity, and mortality in a large prospective cohort of persons 50 to 71 years old. *N Engl J Med* 355: 763–778.
47. Hercberg S, Galan P, Preziosi P, Bertrais S, Mennen L, et al. (2004) The SU.VI.MAX Study: A randomized, placebo-controlled trial of the health effects of antioxidant vitamins and minerals. *Arch Intern Med* 164: 2335–2342.
48. Hercberg S, Preziosi P, Briancon S, Galan P, Triol I, et al. (1998) A primary prevention trial using nutritional doses of antioxidant vitamins and minerals in cardiovascular diseases and cancers in a general population: the SU.VI.MAX study—design, methods, and participant characteristics. *Supplementation en Vitamines et Mineraux Antioxydants. Control Clin Trials* 19: 336–351.
49. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, 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.
50. Dina C, Meyre D, Gallina S, Durand E, Korner A, et al. (2007) Variation in FTO contributes to childhood obesity and severe adult obesity. *Nat Genet* 39: 724–726.
51. Fulker DW, Cherny SS, Sham PC, Hewitt JK (1999) Combined linkage and association analysis for quantitative traits. *Am J Hum Genet* 64: 259–267.
52. Abecasis GR, Cardon LR, Cookson WOC (2000) A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 66: 279–292.
53. Rankinen T, Zuberi A, Chagnon YC, Weisnagel SJ, Argyropoulos G, et al. (2006) The human obesity gene map: the 2005 update. *Obesity (Silver Spring)* 14: 529–644.
54. The International HapMap Consortium (2005) The International HapMap Project. *Nature* 437: 1299–1320.
55. Abecasis GR, Cookson WOC (2000) GOLD - Graphical overview of linkage disequilibrium. *Bioinformatics* 16: 182–183.