

MAP2K3 is associated with body mass index in American Indians and Caucasians and may mediate hypothalamic inflammation

Li Bian, Michael Traurig, Robert L. Hanson, Alejandra Marinelarena, Sayuko Kobes, Yunhua L. Muller, Alka Malhotra, Ke Huang, Jessica Perez, Alex Gale, William C. Knowler, Clifton Bogardus and Leslie J. Baier*

Diabetes Molecular Genetics Section and Diabetes Epidemiology and Clinical Research Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institute of Health, Phoenix, AZ 85004, USA

Received January 16, 2013; Revised May 31, 2013; Accepted June 19, 2013

To identify genes that affect body mass index (BMI) in American Indians who are predominately of Pima Indian heritage, we previously completed a genome-wide association study in 1120 American Indians. That study also included follow-up genotyping for 9 SNPs in 2133 additional subjects. A comprehensive follow-up study has subsequently been completed where 292 SNPs were genotyped in 3562 subjects, of which 128 SNPs were assessed for replication in 3238 additional subjects. In the combined subjects ($n = 6800$), BMI associations for two SNPs, rs12882548 and rs11652094, approached genome-wide significance ($P = 6.7 \times 10^{-7}$ and 8.1×10^{-7} , respectively). Rs12882548 is located in a gene desert on chromosome 14 and rs11652094 maps near *MAP2K3*. Several SNPs in the *MAP2K3* region including rs11652094 were also associated with BMI in Caucasians from the GIANT consortium ($P = 10^{-2} - 10^{-5}$), and the combined P -values across both American Indians and Caucasian were $P = 10^{-4} - 10^{-9}$. Follow-up sequencing across *MAP2K3* identified several paralogous sequence variants indicating that the region may have been duplicated. *MAP2K3* expression levels in adipose tissue biopsies were positively correlated with BMI, although it is unclear if this correlation is a cause or effect. *In vitro* studies with cloned *MAP2K3* promoters suggest that *MAP2K3* expression may be up-regulated during adipogenesis. Microarray analyses of mouse hypothalamus cells expressing constitutively active *MAP2K3* identified several up-regulated genes involved in immune/inflammatory pathways and a gene, *Hap1*, thought to play a role in appetite regulation. We conclude that *MAP2K3* is a reproducible obesity locus that may affect body weight via complex mechanisms involving appetite regulation and hypothalamic inflammation.

INTRODUCTION

Heritable factors are estimated to explain 40–70% of the inter-individual variance in body weight (1). The Pima Indians of Arizona have an extremely high prevalence of obesity, and body mass index (BMI) is highly heritable (2). We previously conducted a genome-wide association study (GWAS) to identify variation associated with BMI in 1120 Pima Indians (3). This report included replication data on 9 SNPs in 2133 subjects; among the 9 SNPs, 5 had nominal evidence for

replication, rs17612333, rs9381282, rs11652094, rs1418029 and rs4811346, but none of the associations achieved genome-wide statistical significance (3). In this prior report, the association for rs11652094 with maximum BMI in the GWAS ($n = 1120$) and replication sample ($n = 2133$) were $P = 2.21 \times 10^{-5}$ and 0.13, respectively, adjusted for age, gender, individual admixture and birth year (3). In the present study, we report a more comprehensive follow-up of the GWAS in a larger sample of American Indians ($n = 6800$) and in this study, the

*To whom correspondence should be addressed at: Diabetes Molecular Genetics Section, NIDDK, NIH, 455 North 5th street, Phoenix, AZ 85004, USA. Tel: +1 6024406589; Fax: +1 6022534140; Email: lbaier@phx.niddk.nih.gov

association of rs11652094 with BMI approaches genome-wide significance ($P = 8.1 \times 10^{-7}$). Rs11652094 is located upstream of *MAP2K3*; therefore, this gene was further analyzed by genotyping additional SNPs located in the *MAP2K3* region, assessing replication of the BMI associations in Caucasians, and investigating potential functional links with obesity.

RESULTS

GWAS SNPs near *MAP2K3* show strong associations with BMI in 6800 American Indians

SNPs from our prior 1 M GWAS (Affymetrix Human SNP array 6.0) were genotyped in two population-based samples of American Indians who are part of a longitudinal study of the Gila River Indian Community (Table 1) (3). Individuals living in this community are predominately full-heritage Pima Indian or mixed-heritage American Indian whose heritage is, on average, 1/2 Pima and 3/4 American Indian including other tribes. A subset of the American Indians from the longitudinal study ($n = 538$) have also participated in inpatient metabolic testing and are informative for detailed metabolic traits in addition to longitudinal measures of BMI. To identify SNPs that are associated with BMI on a population level and to assess the consistency of their association with other measures of body composition (e.g. identify SNPs that are associated with both BMI and percent body fat), 292 SNPs from the original GWAS were genotyped in Sample

1 ($n = 3562$) that included full-heritage Pima Indians from the longitudinal study and all of the subjects who were informative for metabolic traits regardless of heritage. All of the subjects analyzed in our prior GWAS ($n = 1120$) are included in Sample 1. SNPs ($n = 128$) with the strongest associations for BMI in Sample 1 were assessed for replication in Sample 2 ($n = 3238$) which includes all remaining DNA samples from our longitudinal study. Subjects in Sample 2 are of mixed American Indian heritage, and there is no overlap between Sample 1 and Sample 2. Of the 128 SNPs analyzed for association with BMI in all 6800 subjects (Sample 1 + Sample 2), 71 SNPs have a P -value < 0.05 (Supplementary Material, Table S1). SNPs that tag the top five regions ($P < 5 \times 10^{-5}$) are also shown in Table 2.

The association of two SNPs, rs12882548 and rs11652094, with BMI approach genome-wide significance when analyzed in 6800 American Indians (Supplementary Material, Tables S1 and S2). The intergenic SNP rs12882548 provided the strongest association ($P = 6.7 \times 10^{-7}$). Whole genome sequence data which has recently become available for 35 Pima Indians (data not shown) identified 23 SNPs that are in linkage disequilibrium (LD) with rs12882548. The LD block covers a relatively small region (26 kb) located 180 kb downstream of *NOVA1* and 100 kb upstream of *MIR4037* (data not shown). The SNP, rs11652094, with the second strongest association with BMI ($P = 8.1 \times 10^{-7}$) maps upstream (telomeric) of *MAP2K3*. Three additional GWAS SNPs, rs7217373, rs7217403 and rs9901404, that are tagged by rs11652094 ($r^2 > 0.97$), were also among the SNPs with the strongest associations with BMI (Supplementary Material, Table S1 and Fig. S1A). Analysis of the LD pattern between rs11652094, rs7217373, rs7217403, rs9901404 and 30 additional GWAS SNPs spanning a 300 kb region showed that the SNPs with the strongest BMI associations all mapped to the same LD block (Supplementary Material, Fig. S1B). *MAP2K3* is the only gene located within this LD block and was therefore selected for further genotyping and functional studies.

Table 1. Longitudinally studied American Indians analyzed for BMI

Sample	<i>n</i> (M/F)	Mean age (years) \pm SD ^a at max BMI (range)	Mean max BMI (kg/m ²) \pm SD
Adults			
Sample 1	3562 (1521/2038)	36.1 \pm 13.3 (15.0–81.7)	37.4 \pm 8.6 ^b
Sample 2	3238 (1441/1797)	28.9 \pm 11.8 (15.0–85.3)	34.6 \pm 8.8 ^b
Childhood			
Sample 1	2404 (1077/1327)	13.9 \pm 4.0 (5.0–19.9)	27.1 \pm 6.5 ^c
Sample 2	2920 (1336/1584)	13.7 \pm 3.8 (5.0–19.9)	27.3 \pm 7.0 ^c

^aRefers to the mean age at which the highest BMI was recorded.

^bFor adults, BMI is the maximum BMI recorded at age > 15 years.

^cFor the children, BMI is the maximum age and sex adjusted Z-score at age < 20 years. For presentation, Z-scores are converted to BMI units using mean and standard deviation of 12-year-old Pima females.

MAP2K3 sequencing

MAP2K3 has two major transcripts, *MAP2K3* variant A (*MAP2K3-A*) and *MAP2K3* variant B (*MAP2K3-B*) (RefSeq Genes track, UCSC Genome Browser; Supplementary Material,

Table 2. SNPs with the strongest associations with BMI in the combined Sample 1 and Sample 2 ($n = 6800$)

Chr	SNP	Risk/ non-risk	Sample 1 ($n = 3562$)			Sample 2 ($n = 3238$)			Sample 1 + Sample 2			Nearest gene
			Freq risk	Beta ^a	P -value ^b	Freq risk	Beta ^a	P -value ^b	Freq risk	Beta ^a	P ^b	
14	rs12882548	G/A	0.13	0.037	2.0×10^{-6}	0.13	0.025	5.7×10^{-3}	0.13	0.030	6.7×10^{-7}	<i>NOVA1</i> (200 kb)
17	rs11652094	G/C	0.73	0.024	3.9×10^{-5}	0.65	0.022	1.2×10^{-3}	0.69	0.022	8.1×10^{-7}	<i>MAP2K3</i> (16 kb)
6	rs4715352	C/T	0.31	0.029	7.9×10^{-7}	0.45	0.013	0.037	0.38	0.021	1.3×10^{-6}	<i>GSTA5</i> (intron 2)
6	rs9295592	A/G	0.69	0.020	7.7×10^{-4}	0.64	0.022	1.2×10^{-3}	0.67	0.020	8.3×10^{-6}	<i>NRSNI, DCDC2</i> (600 kb)
6	rs12216336	G/C	0.88	0.032	1.1×10^{-4}	0.78	0.023	3.5×10^{-3}	0.83	0.026	9.1×10^{-6}	<i>HLA-DOA</i> (4 kb)

Freq, frequency.

SNPs representing five regions with the strongest associations for BMI in the combined Sample 1 + Sample 2 ($n = 6800$). A complete list of SNPs is given in Supplementary Material, Table S1. The P -values in the GWAS subjects alone, which are a subset of Sample 1, are also shown in Supplementary Material, Table S1.

^a β -Values represent the effect on the logarithmic scale per copy of the risk allele.

^b P -values were adjusted for age, sex, birth year, family membership and American Indian heritage.

Fig. S1B) and the latter encodes 29 additional amino acids at the N-terminus that may possibly function as a nuclear localization signal (4). To identify novel potentially functional variants within *MAP2K3*, the exons, exon–intron boundaries and promoter regions for both *MAP2K3*-A and *MAP2K3*-B along with the promoter region and first 3 exons of an alternative *MAP2K3* transcript (uc002gyu.1) from a previous version of the UCSC Genes build (build 36) were sequenced in 24 unrelated Pima Indians. No missense variants were identified in these 24 subjects. A total of 102 variants were identified, of which 45 were heterozygous in all 24 subjects (data not shown). The high density and abnormal heterozygosity of these polymorphisms are suggestive of large copy number variations or segmental duplications. The genomic sequence encompassing the entire *MAP2K3* region plus an additional 28 kb upstream of *MAP2K3* (the 28 kb region is part of the LD block containing the GWAS lead SNPs, Supplementary Material, Fig. S1B) was queried using the NCBI BLAST program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The BLAST results identified two overlapping segmental duplications located on separate unplaced contigs (segmental duplication 1, contig NW_001839455.1 and segmental duplication 2, contig NW_001841142.1). The duplications span most of the *MAP2K3* region (~19 and 10 kb, respectively, Supplementary Material, Fig. S2A) and are >96% homologous to *MAP2K3*. There are also two confirmed segmental duplications spanning most of the *MAP2K3* genes in the Segmental Duplication Database [WSSD Duplication Track, UCSC Human Genome Browser, 2004 assembly (5–7)] (Supplementary Material, Fig. S2B). There is a high density of dbSNPs mapping to the duplicated regions (Supplementary Material, Fig. S2A) suggesting that a large number of database SNPs assigned to *MAP2K3* are not true polymorphisms but instead are mismatches from paralogous copies of the duplicated regions, i.e. they are paralogous sequence variants (PSVs) (6). The variant calling programs used for our recently obtained whole genome sequence data for 35 Pima Indians similarly identified a large number of SNPs heterozygous (potential PSVs) in all of the subjects in the duplicated region spanning from the middle of intron 1 (rs66859001) to the 3'-UTR (rs62055368) (data not shown). No *MAP2K3* missense mutations were identified in these 35 subjects.

SNP validation

Due to the likelihood of PSVs resulting from segmental duplications across *MAP2K3*, several SNPs in the duplicated regions were validated prior to genotyping. Seventy-five SNPs selected from our sequencing data and the prior GWAS were chosen for validation. SNPs located in the promoter regions and near or within exons were prioritized (Supplementary Material, Table S2). Regions containing putative SNPs were PCR amplified and sequenced using *MAP2K3* and segmental duplication specific primers in 96 unrelated Pima subjects. The 75 putative SNPs located in the duplicated regions along with 12 SNPs in the non-duplicated 5' region of *MAP2K3* were then grouped into five categories: (i) SNPs located in the non-duplicated 5' region of *MAP2K3*; (ii) validated SNPs in the *MAP2K3* region and monomorphic (with consistent allele) in the duplicated regions; (iii) polymorphic in at least one duplicated region but monomorphic in *MAP2K3*; (iv) monomorphic (with opposite

allele) in both *MAP2K3* and the duplicated regions; and (v) polymorphic in both the *MAP2K3* and duplicated regions (Supplementary Material, Table S2). Genotypes for 17 true common SNPs (minor allele frequency >0.02) from groups 1 and 2 were then used to identify three tag SNPs (pair-wise $r^2 \geq 0.80$), including rs11652094 (lead GWAS SNP), rs10468608 and rs12602109 (Fig. 1).

Association results for the *MAP2K3* tag SNPs

As done previously with rs11652094, the other two tag SNPs, rs10468608 and rs12602109, were genotyped for association analysis with BMI in Sample 1 and Sample 2. These subjects are part of the longitudinal study and are informative for maximum BMI recorded during adulthood (age >15 years), maximum BMI recorded during childhood (defined as the highest age and sex standardized Z-score achieved before the age of 20 years), and type 2 diabetes status. All three tag SNPs are associated with maximum BMI measured during adulthood and childhood in Sample 1, but only rs11652094 and rs10468608 replicated in Sample 2 (Table 3). Combining Sample 1 and Sample 2 provided the strongest associations with BMI for rs11652094 and rs10468608 (Table 3). However, none of the three tag SNPs was significantly associated with type 2 diabetes (Supplementary Material, Table S3). A subset of the non-diabetic subjects ($n = 538$) included in Sample 1 had undergone detailed metabolic testing which included measures of body composition. In these individuals, the obesity risk alleles for rs11652094 and rs10468608 were also associated with increased fat mass, percent body fat, waist circumference and waist to thigh ratio (Table 4).

Replication of the BMI associations in Caucasians

GWAS data from Caucasian studies that were included in the GIANT meta-analysis ($n \approx 123\,800$) are available for rs11652094, rs10468608 and four other SNPs, rs72173731, rs72174032, rs2001651 and rs9901404, tagged by rs11652094 in the full-heritage Pima Indians (8). To get a clear comparison between American Indian and Caucasian data, the four redundant SNPs were also directly genotyped in Sample 1 and Sample 2 (combined American Indians, Table 5). In the GIANT meta-analysis, rs11652094 and rs2001651 provided the best evidence for replication (P -values = 2.1×10^{-4} and 6.8×10^{-5} , respectively, Table 5). The two SNPs are in LD in both the full-heritage Pima Indians and HapMap CEU Caucasians ($r^2 = 0.80$ and 0.79 , respectively). The frequencies of the risk alleles for rs11652094 (G allele) and rs2001651 (C allele) are higher in American Indians compared with the Caucasians from the HapMap CEU population (0.68/0.69 and 0.33/0.35, respectively, Table 5). For comparability of the effect estimates, the American Indian data were analyzed with a normalizing transformation of the ranks of BMI comparable with that used in the GIANT study, and the effect estimate for rs11652094 was statistically significantly larger in American Indians ($\beta = 0.089$) than in Caucasians ($\beta = 0.019$, $P = 0.0004$ for difference in effect size by Cochran's Q statistic). When the P -values for the American Indian and Caucasian associations were combined to test the null hypothesis of no association between the SNPs and BMI, there were strong associations with rs11652094 and

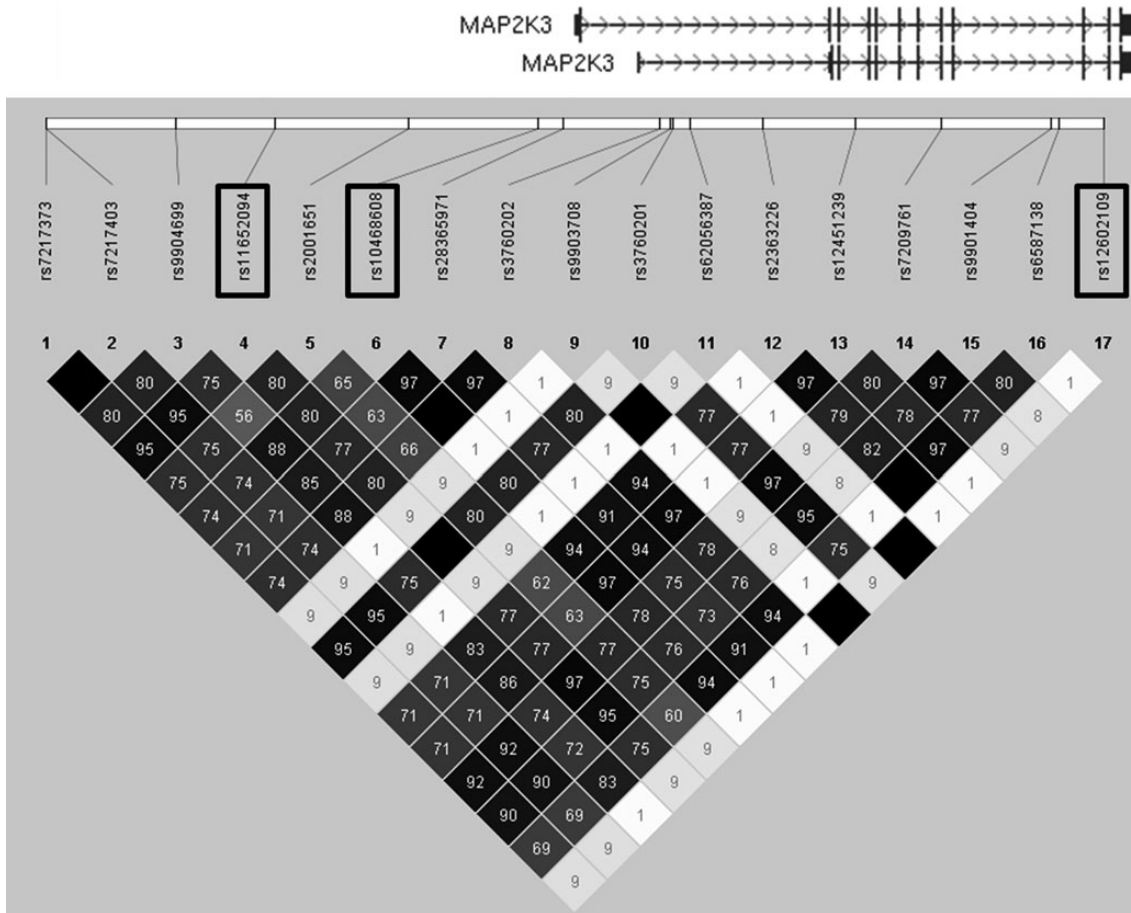


Figure 1. Relative positions and LD plot for the 17 common SNPs across the *MAP2K3* locus in Pima Indians. LD is shown as r^2 and the three tag SNPs are highlighted by boxes.

rs2001651 displaying the lowest P -values (3.6×10^{-9} and 5.5×10^{-9} , respectively, Table 5).

Adipose *MAP2K3* expression is correlated with obesity

Primers specific for *MAP2K3-A* and *MAP2K3-B* transcripts were designed to screen cDNA isolated from human hypothalamus, adipocytes, preadipocytes, skeletal muscle, liver, pancreas and islets. Both *MAP2K3* transcripts were present in the tissues examined. *MAP2K3-B* was highly expressed in all of the tissues, while *MAP2K3-A* had higher expression levels in skeletal muscle and liver (data not shown). To determine whether the segmental duplications of *MAP2K3* are expressed, adipose cDNA isolated from Pima Indians was used to screen for the presence of PSVs in the *MAP2K3* mRNA transcripts. Non-specific primers that amplified both *MAP2K3* and the paralogous regions were designed to sequence across exons 3 through 5 and part of exon 12 which are duplicated on segmental duplications 1 and 2, respectively. The PSVs identified in genomic DNA were not present in the same individual's *MAP2K3* transcript (Supplementary Material, Fig. S3). The transcripts contained only the *MAP2K3* alleles (Supplementary Material, Table S2), suggesting that the paralogous *MAP2K3* regions located on the segmental duplications are not expressed

in human adipose tissue. *MAP2K3* expression levels were determined by quantitative real-time PCR (qRT-PCR) in adipose tissue biopsies from 78 non-diabetic Pima Indians. Adjusted for age and sex, *MAP2K3* expression levels were positively correlated with both BMI and percent body fat (P -values = 0.0004 and 0.006; $r = 0.40$ and 0.31, respectively, Fig. 2). However, rs11652094 which was associated with BMI and percent body fat in Pima Indians was not associated with *MAP2K3* expression levels in adipose tissue ($P > 0.05$; data not shown); therefore, it is unclear whether the increased expression of *MAP2K3* in adipose tissue is a cause or consequence of obesity.

MAP2K3 gene expression is induced during 3T3-L1 preadipocyte differentiation

To determine whether *MAP2K3* could potentially play a role in adipogenesis, we examined *MAP2K3-A* and *MAP2K3-B* promoter activities separately during the initial stage of 3T3-L1 preadipocyte differentiation. *MAP2K3-A* and *MAP2K3-B* promoter reporter vectors were transfected into 3T3-L1 preadipocytes, the cells were treated with standard differentiation media and luciferase activities were measured 48 h post-transfection. Luciferase levels for both *MAP2K3-A* and *MAP2K3-B* promoter vectors were significantly increased for

Table 3. Association results for the three tag SNPs in the *MAP2K3* region with maximum BMI recorded during adulthood and childhood in the longitudinally studied subjects

SNP	Sample	Age ^a	Risk/non-risk	Freq risk allele	Mean BMI (kg/m ²) ± SD			β^b	P-value ^c
					Risk/risk	Risk/non-risk	Non-risk/non-risk		
rs11652094	Sample 1	Adult BMI	G/C	0.73	38.0 ± 8.7	37.0 ± 8.7	35.3 ± 7.3	0.024	3.9 × 10 ⁻⁵
		Childhood BMI			27.6 ± 6.6	26.8 ± 6.5	26.1 ± 6.4	0.11	7.5 × 10 ⁻⁴
	Sample 2	Adult BMI		0.65	35.7 ± 9.2	34.1 ± 8.4	33.0 ± 8.8	0.022	1.2 × 10 ⁻³
		Childhood BMI			28.0 ± 7.1	27.1 ± 6.8	25.9 ± 6.0	0.098	1.4 × 10 ⁻³
	Sample 1 + 2	Adult BMI		0.69	37.0 ± 8.9	35.6 ± 8.7	33.9 ± 8.4	0.022	8.1 × 10 ⁻⁷
		Childhood BMI			27.8 ± 6.9	27.0 ± 6.7	26.0 ± 6.2	0.099	1.0 × 10 ⁻⁵
rs10468608	Sample 1	Adult BMI	C/T	0.77	37.7 ± 8.7	37.1 ± 8.6	35.6 ± 7.4	0.018	4.6 × 10 ⁻³
		Childhood BMI			27.4 ± 6.5	26.9 ± 6.6	26.7 ± 6.5	0.077	0.031
	Sample 2	Adult BMI		0.75	35.2 ± 9.0	34.0 ± 8.4	33.1 ± 8.7	0.022	2.0 × 10 ⁻³
		Childhood BMI			27.7 ± 7.0	27.1 ± 6.8	26.0 ± 5.9	0.087	8.9 × 10 ⁻³
	Sample 1 + 2	Adult BMI		0.76	36.6 ± 8.9	35.6 ± 8.7	34.2 ± 8.2	0.018	1.2 × 10 ⁻⁴
		Childhood BMI			27.5 ± 6.8	27.0 ± 6.7	26.3 ± 6.2	0.076	1.8 × 10 ⁻³
rs12602109	Sample 1	Adult BMI	G/A	0.96	37.6 ± 8.6	35.1 ± 7.8	31.7 ± 5.6	0.056	3.6 × 10 ⁻⁵
		Childhood BMI			27.3 ± 6.6	25.1 ± 5.7	21.3 ± 1.9	0.29	8.1 × 10 ⁻⁵
	Sample 2	Adult BMI		0.92	34.7 ± 8.8	33.7 ± 8.6	33.2 ± 11.6	0.005	0.72
		Childhood BMI			27.4 ± 6.9	26.9 ± 6.7	26.0 ± 6.6	0.014	0.80
	Sample 1 + 2	Adult BMI		0.94	36.3 ± 8.8	34.2 ± 8.4	33.0 ± 10.6	0.024	0.013
		Childhood BMI			27.4 ± 6.8	26.4 ± 6.5	25.3 ± 6.3	0.09	0.044

^aAdult BMI is the highest BMI recorded from an exam at age > 15 years. Childhood BMI is the highest age and sex adjusted Z-score from an exam at age < 20 years. For presentation, the Z-scores were sex and age (female, 12 year) standardized to a BMI scale.

^b β -Values represent the effect on the logarithmic scale per copy of the risk allele.

^cP-values were adjusted for age, sex, birth year, family membership and American Indian heritage.

Table 4. Association of rs11652094 and rs10468608 with obesity-related traits among American Indians who had been metabolically phenotyped

	rs11652094 (G/C)			P-value	rs10468608 (C/T)			P-value
	GG	GC	CC		CC	CT	TT	
Male/Female (n)	171/123	124/73	18/23	—	194/130	108/76	12/13	—
Age (years)	26.9 ± 6.0	26.6 ± 6.3	26.7 ± 6.0	—	27.1 ± 6.0	26.4 ± 6.4	25.6 ± 5.8	—
Fat free mass (kg) ^a	63.6 ± 13.3	61.4 ± 12.3	56.0 ± 12.0	9.0 × 10 ⁻³	63.3 ± 13.1	61.5 ± 12.6	56.0 ± 12.4	0.15
Fat mass (kg) ^a	33.6 ± 13.9	28.4 ± 12.6	26.5 ± 14.3	2.0 × 10 ⁻⁴	32.8 ± 13.9	29.6 ± 12.7	24.2 ± 16.1	0.01
Percent body fat ^a	33.6 ± 8.2	30.6 ± 8.5	31.7 ± 9.8	1.0 × 10 ⁻³	33.0 ± 8.2	31.6 ± 8.5	30.0 ± 10.6	0.04
Waist circumference (cm) ^a	43.2 ± 7.0	40.7 ± 6.5	38.9 ± 6.7	9.0 × 10 ⁻⁵	42.8 ± 7.0	41.2 ± 6.6	37.9 ± 7.1	9.0 × 10 ⁻³
Waist/thigh ^b	1.66 ± 0.17	1.61 ± 0.15	1.56 ± 0.17	2.0 × 10 ⁻⁴	1.65 ± 0.17	1.61 ± 0.16	1.54 ± 0.16	3.0 × 10 ⁻³

Data are given as raw (unadjusted) mean ± SD. P-values are given for an additive model and are adjusted for the following covariates.

^aAge, sex, family membership and American Indian heritage.

^bAge, sex, percent body fat, family membership, and American Indian heritage.

the 3T3-L1 cells treated with the differentiation media compared with the non-induced cells (approximately 3- and 5-fold respectively, Fig. 3A and B), suggesting that *MAP2K3* is up-regulated during adipogenesis.

Downstream targets of MAP2K3

To gain a better understanding of the genes that may function downstream of *MAP2K3*, microarray analyses were done in 3T3-L1 mouse preadipocytes and hypothalamus cells using Affymetrix GeneChip Mouse Gene 1.0 ST arrays to identify genes whose expression levels are changed in response to over-expressing *MAP2K3*. Plasmids that express constitutively active *MAP2K3*-A (*MAP2K3*-A-EE) and *MAP2K3*-B (*MAP2K3*-B-EE) under the control of the CMV promoter were transfected into 3T3-L1 mouse preadipocytes (non-induced and induced), N41 mouse embryonic hypothalamus cells and 2/28 mouse adult hypothalamus cells. Global gene expression levels were

compared between cells transfected with the *MAP2K3* expression plasmids and corresponding cells transfected with an empty control plasmid that only contains the CMV promoter.

Non-induced 3T3-L1 mouse preadipocytes

For 3T3-L1 preadipocytes examined without induction of differentiation, the microarray analysis identified 77 (25 up-regulated/52 down-regulated) and 88 (21 up-regulated/67 down-regulated) differentially expressed gene/expressed sequence tag (est) probe sets with a ≥ 1.5 -fold difference compared with the empty vector control for *MAP2K3*-A-EE and *MAP2K3*-B-EE, respectively (Supplementary Material, Table S4, sheets 1 and 2). Based on gene ontology functional annotation, pathways with the highest enrichment scores for *MAP2K3*-A-EE involve immune/inflammatory response and nucleosome/chromatin assembly. For *MAP2K3*-B-EE, the pathways with the highest enrichment scores include immune/inflammatory response, peptidase inhibitor activity, apoptosis, regulation of cell proliferation and

Table 5. Association and meta-analysis for the SNPs in the *MAP2K3* region with BMI in the combined sample of American Indians and Caucasians from the GIANT study

SNP	Risk/ non-risk	Combined American Indians ^a			Caucasians (GIANT)			Meta-analysis Fixed Effects		Heterogeneity		Stouffer	
		Freq risk	β	<i>P</i> -value	Freq risk	β	<i>P</i> -value	β	<i>P</i> -value	<i>I</i> ² (%)	<i>P</i> -value	Z	<i>P</i> -value
rs7217373	C/G	0.69	0.020	9.7×10^{-6}	0.34	0.016	1.4×10^{-3}	0.020	2.6×10^{-5}	91.3	6.9×10^{-4}	5.35	8.9×10^{-8}
rs7217403	C/T	0.69	0.021	4.0×10^{-6}	0.34	0.016	1.6×10^{-3}	0.020	2.8×10^{-5}	92.1	3.8×10^{-4}	5.41	6.2×10^{-8}
rs11652094	G/C	0.69	0.022	8.1×10^{-7}	0.33	0.019	2.1×10^{-4}	0.024	1.8×10^{-6}	92.0	4.1×10^{-4}	5.90	3.6×10^{-9}
rs2001651	C/A	0.69	0.020	5.6×10^{-6}	0.35	0.020	6.8×10^{-5}	0.024	7.7×10^{-7}	89.8	1.7×10^{-3}	5.85	5.5×10^{-9}
rs10468608	C/T	0.76	0.018	1.2×10^{-4}	0.65	0.009	7.1×10^{-2}	0.013	7.2×10^{-3}	90.2	1.4×10^{-3}	3.94	8.1×10^{-5}
rs9901404	A/G	0.70	0.021	2.5×10^{-6}	0.53	0.013	1.1×10^{-2}	0.017	1.9×10^{-4}	91.8	4.7×10^{-4}	4.83	1.4×10^{-6}

^aCombined American Indians = Sample 1 + Sample 2. The β estimates are given per copy of the risk allele. For GIANT, β estimates were calculated based on an inverse Gaussian transformation of the ranks of BMI and represent the effect per copy of the allele in SD units. For comparison, β estimates for American Indians were also calculated based on an inverse Gaussian transformation (SD units); thus these estimates and *P*-values differ slightly from those presented elsewhere. The β estimate for the fixed effects meta-analysis across American Indians and GIANT is derived by the inverse variance method; *I*² represents the percentage of the total variation between American Indians and Caucasians attributable to heterogeneity. Stouffer's method represents a test based on combining the *P*-values for association across American Indians and Caucasians. Freq, frequency.

oxidation/reduction. A more detailed gene ontology list is shown in Supplementary Material, Table S5. Several genes known to be involved in adipogenesis and lipid metabolism were also differentially expressed. For example, overexpression of *MAP2K3* resulted in the down-regulation of *Lcn2*, *Cyp7b1*, *Igfbp3*, *Cebpb* and *Cebpd*.

Induced 3T3-L1 mouse preadipocytes

We also examined the effects of overexpressing *MAP2K3*-A-EE and *MAP2K3*-B-EE on global gene expression during the initial stage of 3T3-L1 preadipocyte differentiation. After the transfections, the 3T3-L1 cells were treated with standard differentiation media for 48 h. For the induced 3T3-L1 preadipocytes, the microarray analysis identified 196 (59 up-regulated/137 down-regulated) and 107 (45 up-regulated/62 down-regulated) differentially expressed gene/est probe sets with a ≥ 1.5 -fold difference compared with the empty vector control for *MAP2K3*-A-EE and *MAP2K3*-B-EE, respectively (Supplementary Material, Table S4, sheets 3 and 4). Functional pathways with the highest enrichment scores for *MAP2K3*-A-EE include immune/inflammatory response, ribonucleoprotein complex, mitochondrion, protein disulfide oxidoreductase activity and ribonucleotide binding. Examples of functional pathways for *MAP2K3*-B-EE involve immune/inflammatory response, and peptidase inhibitor activity. The analysis also identified a number of up-regulated genes (*Acot2*, *CD36*, *Ces2g* and *Klf9*) and down-regulated genes (*Adipoq*, *Bnip3*, *Fabp5* and *Pex11a*) involved in adipogenesis and lipid metabolism. Three of the genes, *Adipoq*, *Bnip3* and *Pex11a*, were grouped into the fat cell/adipocyte differentiation gene ontology category shown in Supplementary Material, Table S5.

N41 mouse embryonic hypothalamus cells

The microarray analysis identified 255 (191 up-regulated/64 down-regulated) and 115 (93 up-regulated/22 down-regulated) differentially expressed gene/est probe sets with a ≥ 1.5 -fold difference for *MAP2K3*-A-EE and *MAP2K3*-B-EE, respectively (Supplementary Material, Table S4, sheets 5 and 6). For *MAP2K3*-A-EE, some of the significantly over-represented functional pathways involve immune/inflammatory response, antigen processing and presentation and ribonucleotide binding.

Functional pathways identified for *MAP2K3*-B-EE also involve immune/inflammatory response, antigen processing and presentation and ribonucleotide binding. One of the inflammatory response genes, *Myd88*, modestly up-regulated (1.6-fold) by *MAP2K3*-A-EE is a component of the toll-like receptor 4 (TLR4) signaling pathway. The TLR4 signaling pathway is thought to play a role in food intake in response to hypothalamic inflammation (9). A second gene, *Hap1*, that was up-regulated (3.1-fold) by *MAP2K3*-A-EE has also been shown to be involved in appetite regulation (10). Therefore, *Myd88* and *Hap1* were selected for qRT-PCR validation. QRT-PCR was performed using RNA isolated from three independent samples transfected with the *MAP2K3*-A-EE plasmid including the sample used for the microarray analysis. *Myd88* expression levels were modestly up-regulated in all three samples, 2.1- (microarray sample), 2.5- and 2.0-fold (data not shown). *Hap1* expression levels were also up-regulated in all three samples with fold changes of 8.1 (microarray sample), 8.06 and 4.2 (data not shown).

2/28 Mouse adult hypothalamus cells

For the mouse adult hypothalamus cell line, the microarray analysis identified 85 (71 up-regulated/14 down-regulated) and 149 (126 up-regulated/23 down-regulated) differentially expressed gene/est probe sets with a ≥ 1.5 -fold difference for *MAP2K3*-A-EE and *MAP2K3*-B-EE, respectively (Supplementary Material, Table S4, sheets 7 and 8). Over-represented gene ontology terms for *MAP2K3*-A-EE involve immune/inflammatory response and ribonucleotide binding. For *MAP2K3*-B-EE, the category with highest enrichment score was ribonucleoprotein complex. There were also several non-coding small nucleolar RNAs (snoRNAs) that were up-regulated in response to overexpressing *MAP2K3*-B-EE. Most of the snoRNAs are predicted to modify ribosomal RNA via methylation and pseudouridylation.

Genes differentially expressed in both the 3T3-L1 and hypothalamus cell lines

Genes that were differentially expressed in more than one cell type in response to overexpressing *MAP2K3* are shown in Supplementary Material, Table S6, sheets 1 and 2. For the

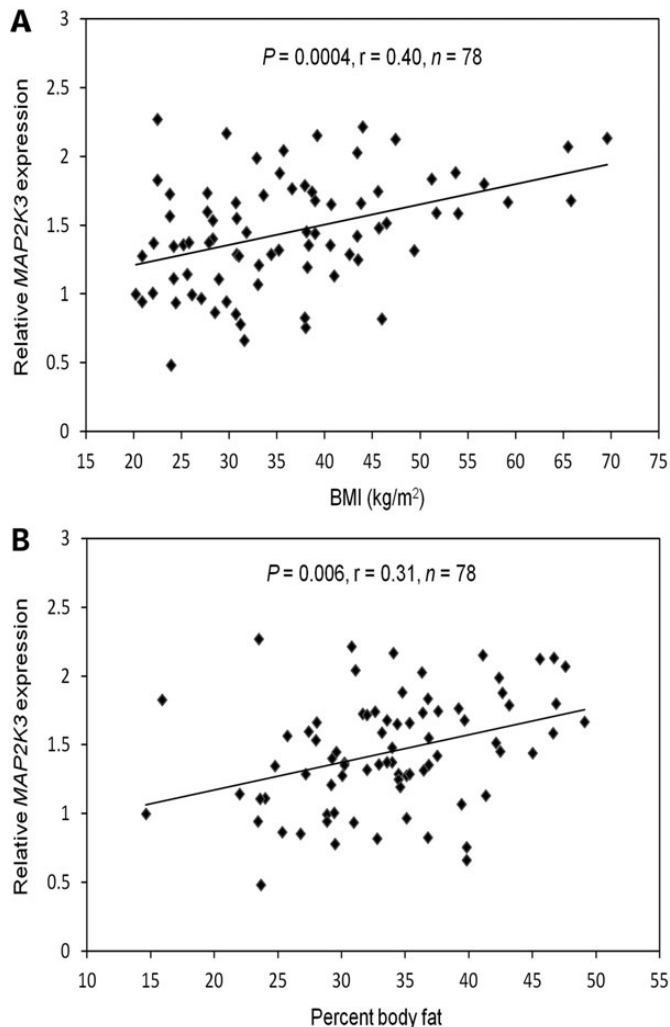


Figure 2. Positive correlation of *MAP2K3* gene expression levels in adipocytes with (A) body mass index (BMI) and (B) percent body fat. Prior to the analysis, relative *MAP2K3* expression levels were logarithmically transformed to approximate the normal distribution. *P*-values were adjusted for age and sex.

MAP2K3-A-EE plasmid, 50 genes were differentially expressed in both the 3T3-L1 preadipocytes and mouse hypothalamus cells. Most of these genes were shared between the induced 3T3-L1 and hypothalamus cells, where the direction of the expression was consistent in both cell types (Supplementary Material, Table S6, sheet 1). For the *MAP2K3*-B-EE plasmid, 16 genes were differentially expressed in both the 3T3-L1 and hypothalamus cell lines; however, most of these genes were shared between the non-induced 3T3-L1 preadipocytes and hypothalamus cells, and the direction of expression was primarily discordant between the cell types (Supplementary Material, Table S6, sheet 2).

DISCUSSION

This study provides information for follow-up genotyping of 71 SNPs from our prior GWAS to identify BMI loci in American

Indians (3). The strongest association was with an intergenic SNP (rs12882548, $P = 6.7 \times 10^{-7}$) located in a 300 kb gene desert on chromosome 14. LD patterns suggest that this signal is localized to a small region (26 kb) ~180 kb downstream of *NOVA1* and 100 kb upstream of *MIR4037* (data not shown), but to date, this region has not been linked to any gene or function. Rs12882548 was not associated with BMI in publically available data from the GIANT study of Caucasians ($P = 0.13$) (8). In contrast, we were able to connect a SNP with a comparably strong association (rs11652094, $P = 8.1 \times 10^{-7}$) to the *MAP2K3* locus and variants in this gene were also associated with BMI in the GIANT study (8). The effect of the risk alleles in *MAP2K3* is significantly stronger in the American Indians when compared with the Caucasians. In the context of such heterogeneity between ethnic groups, the *P*-value-based meta-analysis likely provides the most appropriate test of the combined null hypothesis of no association. Using this analysis, genome-wide significance is achieved for SNPs rs11652094 and rs2001651. The smaller effect in Caucasians may reflect a lower LD between the GWAS lead SNP, rs11652094, and an unidentified causative SNP in this region. Comparison of the LD pattern for the *MAP2K3* region between the Pima Indians and HapMap CEU population shows that in general the LD in this region is lower in Caucasians (Supplementary Material, Fig. S1B and C). Alternatively, the heterogeneous genetic nature of BMI could also account for some SNPs having a larger effect in one population compared with another. For example, we have observed that most of the widely-replicated Caucasian GWAS SNPs for BMI, with the exception of those in *FTO* and *TMEM18*, have weak or non-significant associations in Pimas (Baier *et al.*, unpublished data). Conversely, common SNPs in *SIM1* which show reproducible associations with BMI in both of the American Indian population-based samples studied here are not associated with BMI in French Caucasians (11) but rare SNPs in this gene have subsequently been identified that segregate with severe obesity in Caucasians (12).

The process of adipogenesis is regulated by various MAP kinase signaling cascades such as the p38 MAP kinase pathway (13). Although the function of p38 MAP kinase in adipocyte differentiation is not yet fully understood, it has been shown that p38 MAP kinase may both stimulate (13–17) and inhibit (13,18–20) adipogenesis. The closely related protein kinases, MAP2K3 and MAP2K6, function as specific upstream activators of the four p38 MAP kinase isoforms (α , β , γ and δ) by phosphorylating the TGY motif located in the activation loops (21,22). Engelman *et al.* (23) demonstrated that constitutive expression of *MAP2K6* stimulated p38 MAP kinase-dependent adipogenesis and the accumulation of lipids in 3T3-L1 cells. Our *in vivo* observations of a positive correlation between *MAP2K3* expression levels in adipose biopsies and the BMI of the donor, and our *in vitro* data showing higher luciferase levels for the *MAP2K3* promoter vectors in 3T3-L1 cells treated with differentiation media suggest that *MAP2K3* could also potentially be involved in adipogenesis. Similar to MAP2K6, higher expression of MAP2K3 may dysregulate genes involved in adipogenesis and lipid metabolism. This is supported by the microarray results for the induced 3T3-L1 preadipocytes showing that the expression levels of genes involved in adipogenesis are potentially affected by the overexpression of *MAP2K3*. For example, overexpression of *MAP2K3* led to a

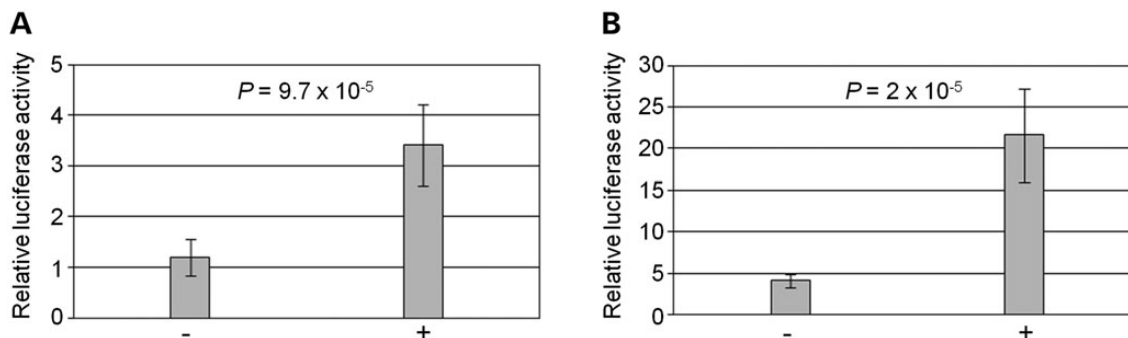


Figure 3. Relative luciferase activities for (A) *MAP2K3*-A promoter and (B) *MAP2K3*-B promoter reporter vectors in 3T3-L1 preadipocytes treated without (–) and with (+) preadipocyte differentiation media. Luciferase assays were done in replicates of six and the results are shown as means \pm SD.

modest increase in *Klf9* (*Kruppel-like factor 9*) a transcription factor that has been shown to be a positive regulator of adipogenesis by regulating the transcription of *Ppar γ* (24). Additional studies are needed to verify the effect of overexpressing *MAP2K3* on *Klf9* expression.

Obesity is the result of an imbalance between energy intake and energy expenditure. We found no evidence that SNPs associated with BMI in the *MAP2K3* region were associated with energy expenditure in 399 subjects who had been studied in our human respiratory chamber (data not shown); therefore, it is possible that *MAP2K3* may have an effect on energy intake. The microarray analysis for the N41 hypothalamus cells identified an up-regulated gene, *Hap1* (*Huntingtin-associated protein 1*), that has been shown to regulate food intake. Four groups have reported that the reduction in *Hap1* gene expression resulted in decreased food intake and lower body weights in mice (25–28). For example, Sheng *et al.* (26) showed that *Hap1* interacts with the GABA_A receptor and that reducing *Hap1* expression in mouse hypothalamus lead to a decrease in GABA_A receptor activity resulting in decreased food intake. *Hap1* is also involved in intracellular trafficking of Bdnf (brain-derived neurotrophic factor), a neurotrophin known to be involved with food intake (29,30).

Inflammatory response was one of gene ontology categories consistently over-represented in our microarray analysis. Obesity and type 2 diabetes have been linked to chronic low-grade inflammation (31,32). In two prior studies comparing global gene expression in adipose tissue and preadipocytes isolated from obese and non-obese Pima Indians, the biological pathway with the most differentially expressed genes was immune/inflammation response (33,34). However, it remains unclear whether the inflammatory response is the cause or effect of obesity. It is well documented that MAP kinases including *MAP2K3* mediate inflammatory responses (35,36), and *MAP2K3* can be activated by proinflammatory cytokines such as interleukin-1 and tumor necrosis factor (37). Our microarray results are consistent with *MAP2K3* playing a role in inflammatory pathways in both adipose and hypothalamus tissues; and, there are reports suggesting that hypothalamic inflammation can lead to leptin and insulin resistance which in turn affects appetite regulation (38,39). One of the inflammatory response genes, *Myd88*, modestly up-regulated in response to constitutively active *MAP2K3*-A in the mouse hypothalamus cells has been shown to be component of a TLR4 hypothalamic

inflammatory signaling pathway that may affect food intake (9). Brain-specific *Myd88* knockout mice were shown to be protected from high-fat diet (HFD) induced obesity and leptin resistance (40). The *Myd88* deficient mice had a slight decrease in food intake compared with the control mice, suggesting that the reduction in food intake may be the mechanism protecting these mice from HFD-induced weight gain (40). Although the thought of *MAP2K3* having an effect on food intake via hypothalamic inflammation is intriguing, more work will be required to address this hypothesis.

From a technical perspective, accurate genotyping across *MAP2K3* was complicated by nucleotide mismatches between homologous regions of *MAP2K3* and two *MAP2K3* segmental duplications (i.e. PSVs). This observation stresses the importance of using Hardy Weinberg criteria as a quality control measure for genotype data. In addition, we found that public databases have incorrectly assigned several SNPs to the *MAP2K3* locus. Some of the SNPs are actually PSVs, while others map to the segmental duplications found on the unplaced contigs NW_001839455.1 and NW_001841142.1 (Supplementary Material, Table S2).

In summary, variation in *MAP2K3* is reproducibly associated with BMI in two large population-based samples of American Indians, as well as Caucasians from the GIANT consortium. These variants have a notable effect size on both adult and childhood BMI. The frequency of the risk allele is higher in American Indians when compared with mixed-heritage American Indians and Caucasians. Our microarray analyses show that *MAP2K3* may play a role in hypothalamic inflammation, and it also identified a gene, *Hap1*, that was up-regulated in response to expression of constitutively active *MAP2K3*. Both hypothalamic inflammation and *Hap1* can affect food intake; however, further studies are needed to prove a causal relationship between *MAP2K3* genotype and appetite regulation.

MATERIALS AND METHODS

Subjects and phenotyping

Characteristics for the American Indians included in this study are provided in Table 1. The subjects are part of a longitudinal study of the etiology of type 2 diabetes in the Gila River Indian Community in Central Arizona where most of the residents are of Pima heritage (2). Individuals who are not full-heritage

Pima Indian are, on average, 1/2 Pima and 3/4 American Indian which includes other tribes. Most (91%) of the subjects in Sample 1 are full-heritage Pima Indian (all 538 individuals informative for metabolic traits were also included in this sample and a few of these individuals are not full-heritage Pima). In contrast, none of the subjects in Sample 2 was full-heritage Pima. Since many of the longitudinally studied subjects had multiple exams in which BMI was measured, the maximum BMI recorded from any single exam after the age of 15 years (adult) was analyzed. For selected variants, BMI in childhood was also analyzed, which was defined as the maximum age and sex adjusted Z-score. For the 538 individuals who had been metabolically characterized in our Clinical Research Center (all were non-diabetic when studied), body composition was estimated by underwater weighing or by total body dual energy X-ray absorptiometry (DPX-1; Lunar Radiation, Madison, WI, USA) (41). All studies were approved by the Institutional Review Board of the National Institute of Diabetes and Digestive and Kidney Diseases. Data from the GIANT consortium were used to assess replication in Caucasians (8).

SNP identification, genotyping methods and quality control

Sequencing of *MAP2K3* in 24 Pima Indians was performed using a Big Dye Terminator Sequencing kit (Applied Biosystems, Foster City, CA, USA) on an automated DNA capillary sequencer (model 3730XL, Applied Biosystems). Whole genome sequence data for 35 Pima Indians was generated at Complete Genomics Inc. (Mountain View, CA, USA). Genotyping was done by allelic discrimination using Taqman genotyping assays (Applied Biosystems), SNPlex genotyping System 48-plex (Applied Biosystems) or the Illumina BeadXpress System (Illumina, San Diego, CA, USA). All genotypic data met our quality control criteria which require a lack of deviation from Hardy–Weinberg equilibrium at $P < 0.001$ and a discrepancy rate of $\leq 1.0\%$ in blind duplicates (330 and 100 blind duplicates for Sample 1 and Sample 2, respectively). All SNPs had a genotyping call rate $\geq 95\%$ in Sample 1 and Sample 2 except for two SNPs in Sample 2 (rs12216336 and rs17612333) which had call rates of 91 and 94%, respectively (Supplementary Material, Table S1).

Statistical analysis

Statistical analyses were performed using SAS software version 9.1 (SAS Institute, Cary, NC, USA). Linear regression was used to assess the association of continuous traits with genotype (assuming an additive model). The generalized estimating equation procedure was used to account for family membership and to adjust for other confounding variables (42). The individual estimate of European admixture was also used as a covariate and these estimates were derived by the method of Hanis *et al.* (43) from 45 markers with large difference in allele frequency between populations (44). Because the analyses for the GIANT study were conducted using a normalizing inverse Gaussian transformation of the ranks of BMI, we similarly transformed the American Indian data for meta-analysis with GIANT. The inverse variance method was used to estimate the common effect size across American Indians and Caucasians and Cochran's Q statistic was used to evaluate heterogeneity

(45). Stouffer's method was also used to test the null hypothesis of no association between the variant and BMI across the American Indian and GIANT populations (46). Stouffer's method combines the P -values across studies and does not require the estimation of a common effect; it assesses only consistency in the direction of effect across populations. In the presence of heterogeneity in effect sizes, Stouffer's method can provide a more powerful test of association than the inverse variance method (47). The LD pattern was determined using Haploview (version 4.2). Tag SNPs were selected using the Tagger algorithm (48) with a pair-wise $r^2 \geq 0.80$ taken as indicative of redundancy.

Tissue profiling for *MAP2K3-A* and *MAP2K3-B*

Specific forward primers located in *MAP2K3-A* exon 1 (5-CAGTTGGCCCGTGTG AGGAGA-3) and *MAP2K3-B* exon 1 (5-AGTCCTCTAGATTAGTCTCCA-3) and a common reverse primer located in both transcripts (5-TCCAGGGCCC GCACGATAGAC-3) were used to screen cDNA from the following human tissues: hypothalamus (BD Marathon-Ready cDNA, BD Bioscience/Clontech, Mountain View, CA, USA); brain, skeletal muscle, liver and pancreas (BD Human MTC Multiple Tissue cDNA Panel, BD Bioscience/Clontech); pancreatic islets (kindly provided by Dr Lorella Marcelli at the Joslin Diabetes Center); and adipocytes and preadipocytes isolated from Pima Indians. Adipocytes and preadipocytes were obtained from subcutaneous abdominal adipose tissue biopsies as previously described (33,49) and total RNA was extracted using an RNeasy Mini Kit (Qiagen, Valencia, CA, USA). To remove any residual DNA, the purified RNA was treated with DNase using RNase-free DNase Set (Qiagen). First-strand cDNA was synthesized from the adipocyte and preadipocyte total RNA using a BD Advantage RT-for-PCR Kit (BD Bioscience/Clontech) following the manufacturer's instructions.

MAP2K3 quantitative real-time PCR in adipocytes

MAP2K3 gene expression levels were quantified using a TaqMan Gene Expression Assay (assay ID: Hs00177127_m1, Applied Biosystems) on an ABI PRISM 7700 Real-Time PCR System (Applied Biosystems). Each sample was run in triplicate and *MAP2K3* transcript levels were normalized using beta actin (assay ID: Hs99999903_m1, Applied Biosystems). The relative *MAP2K3* expression levels were determined by the $\Delta\Delta C_t$ method according to the manufacturer's protocol (Applied Biosystems).

MAP2K3 promoter vector and reporter assays

To construct the *MAP2K3* promoter reporter vectors, the *MAP2K3-A* and *MAP2K3-B* promoter regions were amplified by PCR using human genomic DNA. The promoter regions (variant A, chr17:21,190,566–21,191,461 and variant B, chr17: 21,187,213–21,188,230, NCBI build 37.1, UCSC Genome Browser) were cloned into the promoterless PGL3-basic vector (Promega, Madison, WI, USA). 3T3-L1 preadipocytes were purchased from ZenBio and maintained in 3T3-L1 Preadipocyte Medium (ZenBio, Research Triangle Park, NC, USA). Transfections were done with Lipofectamine

LTX and PLUS Reagent (Life Technologies, Grand Island, NY, USA) following the manufacturer's instructions using serum and antibiotic free medium. To control for transfection efficiency, cotransfections were done using the pGL4.74[hRLuc/TK] plasmid (Promega) and normalized to the *Renilla* luciferase expression. Following the transfections, the 3T3-L1 preadipocytes were treated with 3T3-L1 Differentiation Media (ZenBio, Research Triangle Park). Forty-eight hours post-transfection, the cells were washed with phosphate buffer solution (PBS) and cell lysates were prepared using the lysis buffer from the Dual-Luciferase Reporter Assay System (Promega). Luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega). Luciferase activity is expressed as the relative activity and is shown as the means \pm standard deviation (SD) for six replicates.

Constitutively active *MAP2K3-A* and *MAP2K3-B* expression plasmids

The expression plasmid containing the constitutively active form of *MAP2K3-B* (S218E/T222E, *MAP2K3-B-EE*) was created by site-directed mutagenesis (QuikChange II Site-Directed Mutagenesis Kit, Agilent Technologies) using a wild-type *MAP2K3-B* expression vector purchased from Origene. Constitutively active *MAP2K3-B* was created using two site-directed mutagenesis thermal cyclings. For the first mutagenesis, T222 was converted to E222 using the following primers: forward, 5'-GCTACTTGGTGGACTCTGTGGCCAAGGAGATGGATGCCGGC-3' and reverse, 5'-GCCGGCATCCATCTCCTTGGCCACAGAGTCCACCAAGTAGC-3'. S218 was then converted to E218 using the following primers: forward, 5'-GCTACTTGGTGGACGAGGTGGCCAAGGATGATGGATGCCGGC-3' and reverse, 5'-GCCGGCATCCATCTCCTTGGCCACCTCG TCCACCAAGTAGC-3'. Because the 5'-UTR of *MAP2K3-B* is GC rich, both thermal cycling reactions contained 5% DMSO. Underlined letters in the primers indicate the nucleotides that were changed. The constitutively active form of *MAP2K3-A* (S189E/T193E, *MAP2K3-A-EE*) was constructed by replacing exon 1 of *MAP2K3-B-EE* with exon 1 (5'-UTR) of *MAP2K3-A*. The only difference between *MAP2K3-A* and *MAP2K3-B* is an alternative exon 1.

Microarray analysis

N41 mouse embryonic and 2/28 mouse adult hypothalamus cells were purchased from CELLutions Biosystems Inc. (Toronto, ON, Canada). 3T3-L1 preadipocytes (ZenBio, Research Triangle Park) were maintained in 3T3-L1 Preadipocyte Medium (ZenBio, Research Triangle Park) and the N41 and 2/28 mouse hypothalamus cells were grown in DMEM containing 10% fetal bovine serum (ATCC). To induce differentiation, the 3T3-L1 preadipocytes were treated with 3T3-L1 Differentiation Media (ZenBio, Research Triangle Park). The 3T3-L1, N41 and 2/28 cells were transfected with plasmids containing constitutively active *MAP2K3-A*, constitutively active *MAP2K3-B* and an empty control vector. The expression plasmids containing constitutively active *MAP2K3-A* and *MAP2K3-B* are under the control of the CMV promoter, while the empty control plasmid (pCMV6-XL4, Origen, Rockville, MD, USA) contains only the CMV promoter. Transfections were done with

Lipofectamine LTX and Plus Reagent (Life Technologies) following the manufacturer's instructions using serum and antibiotic free medium. Total RNA was isolated from non-induced and induced 3T3-L1, N41 mouse embryonic hypothalamus and 2/28 mouse adult hypothalamus cells transfected with either the *MAP2K3* constitutively active (*MAP2K3-A-EE*, *MAP2K3-B-EE*) or empty control (pCMV6-XL4, Origene) plasmids. Briefly, total RNA was extracted using an RNeasy Mini Kit (Qiagen) and the purified RNA was treated with DNase using RNase-free DNase Set (Qiagen) to remove any DNA. For the microarray analysis, cDNA was generated from the total RNA samples using the Ambion WT Expression Kit (Applied Biosystems) and the cDNA was prepared for hybridization using the GeneChip WT Terminal Labeling and Hybridization Kit (Affymetrix, Santa Clara, CA, USA) following the manufacturer's instructions. The fragmented and labeled cDNA was then hybridized to GeneChip Mouse Gene 1.0 ST arrays (Affymetrix) for 16–18 h at 45°C and the arrays were stained and washed using an Affymetrix GeneChip Fluidics Station 450. The arrays were scanned using an Affymetrix GeneChip Scanner and gene expression data were analyzed using Affymetrix Expression Console version 1.0 software. The differentially expressed genes were classified into gene ontology terms using the web-based DAVID Functional Annotation Tool, David Bioinformatics Resources 6.7 (50,51). The microarray results for *Myd88* and *Hap1* were confirmed by qRT-PCR using TaqMan Gene Expression Assays (assay IDs: Mm00440338_m1 and Mm00468825_m1, Applied Biosystems) on an ABI Prism 7900HT-Fast Real-Time PCR System (Applied Biosystems). Each sample was run in triplicate and expression levels were normalized using beta actin (assay ID: Mm01205647_g1, Applied Biosystems). Relative expression levels for *Myd88* and *Hap1* were determined by the $\Delta\Delta C_t$ method according to the manufacturer's protocol (Applied Biosystems).

SUPPLEMENTARY MATERIAL

Supplementary Material is available at *HMG* online.

ACKNOWLEDGMENTS

We thank all the participants from the Gila River Indian Community. We also thank the GIANT consortium for providing us with early access to their data and making their large databases publicly available for the benefit of all investigators. We also thank Vicky Ossowski for her help with the microarray analysis.

Conflict of Interest statement. None declared.

FUNDING

This work was supported by the intramural research program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. Funding to pay the Open Access publication charges for this article was provided by The Intramural Program of NIDDK, NIH.

REFERENCES

- Barsh, G.S., Farooqi, I.S. and O'Rahilly, S. (2000) Genetics of body-weight regulation. *Nature*, **404**, 644–651.
- Knowler, W.C., Bennett, P.H., Hamman, R.F. and Miller, M. (1978) Diabetes incidence and prevalence in Pima Indians: a 19-fold greater incidence than in Rochester, Minnesota. *Am. J. Epidemiol.*, **108**, 497–505.
- Malhotra, A., Kobes, S., Knowler, W.C., Baier, L.J., Bogardus, C. and Hanson, R.L. (2011) A genome-wide association study of BMI in American Indians. *Obesity*, **19**, 2102–2106.
- Han, J., Wang, X., Jiang, Y., Ulevitch, R.J. and Lin, S. (1997) Identification and characterization of a predominant isoform of human MKK3. *FEBS Lett.*, **10**, 19–22.
- Bailey, J.A., Gu, Z., Clark, R.A., Reinert, K., Samonte, R.V., Schwartz, S., Adams, M.D., Myers, E.W., Li, P.W. and Eichler, E.E. (2002) Recent segmental duplications in the human genome. *Science*, **297**, 1003–1007.
- Speliotes, E.K. and Willer, C.J. (2010) Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.*, **42**, 937–948.
- Lumeng, C.N. and Saltiel, A.R. (2011) Inflammatory links between obesity and metabolic disease. *J. Clin. Invest.*, **121**, 2111–2117.
- Woods, S.C. and Seeley, R.J. (2006) Hap1 and GABA: thinking about food intake. *Cell Metab.*, **3**, 388–390.
- Traurig, M., Mack, J., Hanson, R.L., Ghossaini, M., Meyre, D., Knowler, W.C., Kobes, S., Froguel, P., Bogardus, C. and Baier, L.J. (2009) Common variation in SIM1 is reproducibly associated with BMI in Pima Indians. *Diabetes*, **58**, 1682–1689.
- Stutzmann, F., Ghossaini, M., Couturier, C., Marchand, M., Vatin, V., Corset, L., Lecoq, C., Balkau, B., Horber, F., Driscoll, D.J. *et al.* (2009) Loss-of-function mutations in SIM1 cause a specific form of Prader-Willi-like syndrome (abstract). *Diabetologia*, **52**, S104.
- Bost, F., Aouadi, M., Caron, L. and Binétruy, B. (2005) The role of MAPKs in adipocyte differentiation and obesity. *Biochimie*, **87**, 51–56.
- Aouadi, M., Jager, J., Laurent, K., Gonzalez, T., Cormont, M., Binétruy, B., Le Marchand-Brustel, Y., Tanti, J-F. and Bost, F. (2007) p38MAP kinase activity is required for human primary adipocyte differentiation. *FEBS Lett.*, **581**, 5591–5596.
- Engelman, J.A., Lisanti, M.P. and Scherer, P.E. (1998) Specific inhibitors of p38 mitogen-activated protein kinase block 3T3-L1 adipogenesis. *J. Biol. Chem.*, **273**, 32111–32120.
- Patel, N.G., Holder, J.C., Smith, S.A., Kumar, S. and Eggo, M.C. (2003) Differential regulation of lipogenesis and leptin production by independent signaling pathways and rosiglitazone during human adipocyte differentiation. *Diabetes*, **52**, 43–50.
- Maekawa, T., Jin, W. and Ishii, S. (2010) The role of ATF-2 family transcription factors in adipocyte differentiation: antiobesity effects of p38 inhibitors. *Mol. Cell. Biol.*, **30**, 613–625.
- Wang, X.Z. and Ron, D. (1996) Stress-induced phosphorylation and activation of the transcription factor CHOP (GADD153) by p38 MAP Kinase. *Science*, **272**, 1347–1349.
- Aouadi, M., Laurent, K., Prot, M., Le Marchand-Brustel, Y., Binétruy, B. and Bost, F. (2006) Inhibition of p38MAPK increases adipogenesis from embryonic to adult stages. *Diabetes*, **55**, 281–289.
- He, Z., Zhu, H.H., Bauler, T.J., Wang, J., Ciaraldi, T., Alderson, N., Li, S., Raquil, M.A., Ji, K., Wang, S. *et al.* (2012) Nonreceptor tyrosine phosphatase Shp2 promotes adipogenesis through inhibition of p38 MAP kinase. *Proc. Natl Acad. Sci. USA*, Epub ahead of print.
- Gerits, N., Kostenko, S. and Moens, U. (2007) In vivo functions of mitogen-activated protein kinases: conclusions from knock-in and knock-out mice. *Transgenic Res.*, **16**, 281–314.
- Remy, G., Risco, A.M., Inesta-Vaquera, F.A., Gonzalez-Teran, B., Sabio, G., Davis, R.J. and Cuenda, A. (2010) Differential activation of p38MAPK isoforms by MKK6 and MKK3. *Cell Signal*, **22**, 660–667.
- Engelman, J.A., Berg, A.H., Lewis, R.Y., Lin, A., Lisanti, M.P. and Scherer, P.E. (1999) Constitutively active mitogen-activated protein kinase kinase 6 (MKK6) or salicylate induces spontaneous 3T3-L1 adipogenesis. *J. Biol. Chem.*, **274**, 35630–35638.
- Pei, H., Yao, Y., Yang, Y., Liao, K. and Wu, J.R. (2011) Krüppel-like factor KLF9 regulates PPAR γ transactivation at the middle stage of adipogenesis. *Cell Death Differ.*, **18**, 315–327.
- Chan, E.Y., Nasir, J., Gutekunst, C.A., Coleman, S., Maclean, A., Maas, A., Metzler, M., Gertsenstein, M., Ross, C.A., Nagy, A. and Hayden, M.R. (2002) Targeted disruption of Huntingtin-associated protein-1 (Hap1) results in postnatal death due to depressed feeding behavior. *Hum. Mol. Genet.*, **11**, 945–959.
- Sheng, G., Chang, G.Q., Lin, J.Y., Yu, Z.X., Fang, Z.H., Rong, J., Lipton, S.A., Li, S.H., Tong, G., Leibowitz, S.F. and Li, X.J. (2006) Hypothalamic huntingtin-associated protein 1 as a mediator of feeding behavior. *Nat. Med.*, **12**, 526–533.
- Lin, Y.F., Xu, X., Cape, A., Li, S. and Li, X.J. (2010) Huntingtin-associated protein-1 deficiency in orexin-producing neurons impairs neuronal process extension and leads to abnormal behavior in mice. *J. Biol. Chem.*, **285**, 15941–15949.
- Niu, S.N., Huang, Z.B., Wang, H., Rao, X.R., Kong, H., Xu, J., Li, X.J., Yang, C. and Sheng, G.Q. (2011) Brainstem Hap1-Ahi1 is involved in insulin-mediated feeding control. *FEBS Lett.*, **585**, 85–91.
- Wu, L.L., Fan, Y., Li, S., Li, X.J. and Zhou, X.F. (2010) Huntingtin-associated protein-1 interacts with pro-brain-derived neurotrophic factor and mediates its transport and release. *J. Biol. Chem.*, **285**, 5614–5623.
- Rosas-Vargas, H., Martínez-Ezquerro, J.D. and Bienvu, T. (2011) Brain-derived neurotrophic factor, food intake regulation, and obesity. *Arch. Med. Res.*, **42**, 482–494.
- Sun, S., Ji, Y., Kersten, S. and Qi, L. (2012) Mechanisms of inflammatory responses in obese adipose tissue. *Annu. Rev. Nutr.*, **21**, 261–286.
- Johnson, A.R., Milner, J.J. and Makowski, L. (2012) The inflammation highway: metabolism accelerates inflammatory traffic in obesity. *Immunol. Rev.*, **249**, 218–238.
- Lee, Y.H., Nair, S., Rousseau, E., Allison, D.B., Page, G.P., Tataranni, P.A., Bogardus, C. and Permana, P.A. (2005) Microarray profiling of isolated abdominal subcutaneous adipocytes from obese vs non-obese Pima Indians: increased expression of inflammation-related genes. *Diabetologia*, **48**, 1776–1783.
- Nair, S., Lee, Y.H., Rousseau, E., Cam, M., Tataranni, P.A., Baier, L.J., Bogardus, C. and Permana, P.A. (2005) Increased expression of inflammation-related genes in cultured preadipocytes/stromal vascular cells from obese compared with non-obese Pima Indians. *Diabetologia*, **48**, 1784–1788.
- Kyriakis, J.M. and Avruch, J. (2012) Mammalian MAPK signal transduction pathways activated by stress and inflammation: a 10-year update. *Physiol. Rev.*, **92**, 689–737.
- Dong, C., Davis, R.J. and Flavell, R.A. (2002) MAP Kinases in the immune response. *Annu. Rev. Immunol.*, **20**, 55–72.
- Dérjard, B., Raingeaud, J., Barrett, T., Wu, I., Han, J., Ulevitch, R.J. and Davis, R.J. (1995) Independent human MAP kinase signal transduction pathways defined by MEK and MKK isoforms. *Science*, **267**, 682–685.
- Araújo, E.P., Torsoni, M.A. and Velloso, L.A. (2010) Hypothalamic inflammation and obesity. *Vitam. Horm.*, **82**, 129–143.
- Thaler, J.P., Choi, S.J., Schwartz, M.W. and Wise, B.E. (2010) Hypothalamic inflammation and energy homeostasis: resolving the paradox. *Front. Neuroendocrinol.*, **31**, 79–84.
- Kleinridders, A., Schenten, D., Konner, A.C., Belgardt, B.F., Mauer, J., Okamura, T., Wunderlich, F.T., Medzhitov, R. and Brüning, J.C. (2009) Myd88 signaling in the CNS is required for development of fatty acid-induced leptin resistance and diet-induced obesity. *Cell Metab.*, **10**, 249–259.
- Norman, R.A., Tataranni, P.A., Pratley, R., Thompson, D.B., Hanson, R.L., Prochazka, M., Baier, L., Ehm, M.G., Sakul, H., Foroud, T. *et al.* (1998) Autosomal genomic scan for loci linked to obesity and energy metabolism in Pima Indians. *Am. J. Hum. Genet.*, **62**, 659–668.
- Zeger, S.L. and Liang, K.Y. (1986) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, **42**, 121–130.
- Hanis, C.L., Chakraborty, R., Ferrell, R.E. and Schull, W.J. (1986) Individual admixture estimates: disease associations and individual risk of diabetes and gallbladder disease in Mexican-Americans in Starr County, Texas. *Am. J. Phys. Anthropol.*, **70**, 433–441.
- Tian, C., Hinds, D.A., Shigeta, R., Adler, S.G., Lee, A., Pahl, M.V., Silva, G., Belmont, J.W., Hanson, R.L., Knowler, W.C. *et al.* (2007) A genome-wide single nucleotide polymorphism panel for Mexican American admixture mapping. *Am. J. Hum. Genet.*, **80**, 1014–1023.
- Petitti, D.B. (2000) Statistical methods in meta-analysis. In Petitti, D.B. (ed.), *Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine*. Oxford University Press, Oxford, pp. 94–118.

44. Stouffer, S.A., Suchman, E.A., Devinney, L.C., Star, S.A. and Williams, R.M. (1949) How the volumes were produced. In: *The American Soldier, Adjustment to Army Life*. Princeton University Press, Princeton, Vol. 1, pp. 45.
45. Hedges, L.V., Cooper, H. and Bushman, B.J. (1992) Testing the null hypothesis in meta-analysis: a comparison of combined probability and confidence interval procedures. *Psychol. Bull.*, **111**, 188–194.
46. 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.
47. Permana, P.A., Nair, S., Lee, Y.H., Luczy-Bachman, G., Vozarova de Courten, B. and Tataranni, P.A. (2004) Subcutaneous abdominal preadipocyte differentiation in vitro inversely correlates with central obesity. *Am. J. Physiol. Endocrinol. Metab.*, **286**, E958–E962.
48. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
49. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
50. She, X., Jiang, Z., Clark, R.A., Liu, G., Cheng, Z., Tuzun, E., Church, D.M., Sutton, G., Halpern, A.L. and Eichler, E.E. (2004) Shotgun sequence assembly and recent segmental duplications within the human genome. *Nature*, **431**, 927–930.
51. Bailey, J.A., Yavor, A.M., Massa, H.F., Trask, B.J. and Eichler, E.E. (2001) Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res.*, **11**, 1005–1017.