Supplementary Material

Methods for data collection of cardiovascular and metabolic outcomes

Full details of laboratory assays for fasting glucose and total cholesterol and blood pressure measurement have been described previously. [1] Diabetes was defined as having a fasting plasma glucose of >7.0 mmol/l or having been diagnosed previously with diabetes by a doctor.

MX Methods

In order to correct for the effect of substructure we performed within family tests of association as parameterized by the Fulker Model. [2] Briefly the test maximizes the natural log of the following likelihood:

with respect to the vector of expected means $\boldsymbol{\mu}_i$ and covariance $\sum_i$ for family *i* where k = 2 is the number of siblings measured in family *i*, $\boldsymbol{y}_i$ is a vector of observed scores for individuals in family *i*, and *M* is the number of independent families. The test for association is modelled in the means (fixed effects) part of the model where the expected value for each individual is parameterized as a function of the genotype at the locus under study:

where μ is an overall grand mean, g_{ij} reflects allelic dosage for individual j of sib-pair i (i.e. -1, 0, and 1) at the marker under study, and β_a is a regression coefficient quantifying the degree of association. The Fulker test partitions the allelic dosage into orthogonal between family (b_i) and within family (w_{ij}) components of $g_{ij.} = b_i + w_{ij}$ (see also Table 2 in Fulker et al. [2] for a complete definition) and a regression coefficient is estimated for each:

Since sibling pairs must share the same genetic ancestry (by definition), comparing the full model against a restricted model where the within family component is constrained to zero yields a test of association which is robust to the effects of population stratification. Asymptotically twice the difference in log-likelihood between the models is distributed as a chi-square statistic with one degree of freedom.

The non-independence between siblings is modelled in the covariance (random effects) part of the model:

where σ_q^2 is the additive genetic variance due to the putative quantitative trait locus, σ_a^2 is the (residual) polygenic additive genetic variance, and σ_e^2 is the unique environmental variance.

$\hat{\pi}_{ij}$ denotes the estimated proportion of alleles shared IBD at the marker locus by siblings *i* and *j* and was estimated by the Merlin program.[3]

We extended the fixed effects part of the basic Fulker model to estimate a regression coefficient parameterizing the interaction between the between family genetic effects and a measured environmental variable (β_{be}), and a coefficient reflecting the interaction between the within family genetic effects and a measured environmental variable (β_{we}). The coefficients for these regression parameters were derived by multiplying the coefficient for the between families component with the environmental variable (i.e.), and the within families component with the environmental variable respectively:

This full model was compared against a reduced model where β_{we} was constrained to zero to yield a test of gene by environment interaction which is robust to the effect of population stratification. Using this framework we tested for interaction between *FTO* and rural/urban location, coded as 0/1. All analyses were performed using the Mx software package. [4]

Reference List

1. Ebrahim S, Kinra S, Bowen L, Andersen E, Ben-Shlomo Y, Lyngdoh T*, et al.* The effect of rural-to-urban migration on obesity and diabetes in India: a cross-sectional study. *PLoS Med* 2010; **7(4)**:e1000268.

2. Fulker DW, Cherny SS, Sham PC, Hewitt JK. Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 1999; **64(1)**:259-267.

3. Abecasis GR, Cardon LR, Cookson WO. A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 2000; **66(1)**:279-292.

4. Neale MC (1997) Statistical modelling. Department of Psychiatry, Medical College of Virginia, Richmond, VA.

Supplementary Tables

Table S1 Genotype frequencies in sample of unrelated individuals from the whole study

	SNP number		Genotype Frequencies N (%)		Minor Allele Frequency	HWE p value
FTO	rs9939609	TT	TA	AA		
		1504(44.7)	1511 (44.9)	350 (10.4)	0.33	0.31
MC4R	rs12970134	GG	GA	AA		
		1,470 (43.2)	1,534 (45.0)	403 (11.8)	0.36	0.94
MC4R	rs17782313	TT	TC	CC		
		1,412 (41.7)	1,544 (45.6)	431(12.7)	0.34	0.79

Table S2 Genotype frequencies in sample of unrelated individuals from Lucknow

	SNP number		Genotype Frequencies N (%)		Minor Allele Frequency	HWE p value
FTO	rs9939609	TT	TA	AA		
		447 (45.4)	415(42.1)	123(12.5)	0.34	0.09
MC4R	rs12970134	GG	GA	AA		
		398(40.1)	467 (47.1)	127(12.8)	0.36	0.63
MC4R	rs17782313	TT	TC	CC		
		385(39.1)	464(47.2)	135(13.7)	0.37	0.84

Table S3 Genotype frequencies in sample of unrelated individuals from Nagpur

	SNP number		Genotype Frequencies N (%)		Minor Allele Frequency	HWE p value
FTO	rs9939609	TT	TA	AA		
		462(47.6)	422(43.5)	87(9.0)	0.31	0.54
MC4R	rs12970134	GG	GA	AA		
		428(44.0)	445(45.8)	99(10.2)	0.33	0.31
MC4R	rs17782313	TT	TC	CC		
		401(41.4)	452(46.7)	115(11.9)	0.35	0.53

Table S4 Genotype frequencies in sample of unrelated individuals from Hyderabad

	SNP number		Genotype Frequencies N (%)		Minor Allele Frequency	HWE p value
FTO	rs9939609	TT	TA	AA		
		331(43.2)	345(45.0)	90(11.8)	0.34	1.00
MC4R	rs12970134	GG	GA	AA		
		428(44.0)	331(42.8)	87(11.24)	0.33	0.46
MC4R	rs17782313	TT	TC	CC		
		349(45.3)	336(43.6)	86(11.2)	0.33	0.74

Table S5 Genotype frequencies in sample of unrelated individuals from Bangalore

	SNP number		Genotype Frequencies N (%)		Minor Allele Frequency	HWE p value
FTO	rs9939609	TT	TA	AA		
		264(41.1)	329(51.2)	50(7.8)	0.33	0.0001
MC4R	rs12970134	GG	GA	AA		
		288(43.1)	291(43.5)	90(13.5)	0.35	0.24
MC4R	rs17782313	TT	TC	CC		
		277(41.7)	292(44.0)	95(14.3)	0.36	0.21

Table S6 Demographic and cardiovascular and metabolic outcomes in the study population stratified by obesity

	Males					Females				
	BMI≤25kg/m^2		BMI>25 kg/m^2		P[a]	BMI≤25 kg/m^2		BMI>25 kg/m^2		P[a]
N	2678		1246			1568		1283		
% Rural	49.6		25.7		<0.001	37.7		18.4		<0.001
% Diabetic	5.7		10.6		<0.001	2.9		10.5		<0.001
Age (years)	39.8	(0.2)	44.9	(0.2)	<0.001	36.8	(0.3)	43.4	(0.2)	<0.001
Total METS (hr/day)	40.07	(0.1)	38.4	(0.1)	<0.001	38.2	(0.1)	37.3	(0.1)	<0.001
Dietary Fat intake (g/day)	87.6	(0.8)	92.7	(1.1)	<0.001	73.5	(0.8)	76.3	(0.9)	0.02
Systolic Blood pressure (mmHg)	121.9	(0.3)	129.5	(0.5)	<0.001	115.3	(0.4)	123.9	(0.5)	<0.001
Total cholesterol (mmol/l)	4.5	(0.02)	4.9	(0.03)	<0.001	4.6	(0.03)	5.0	(0.03)	<0.001
Fasting glucose (mmol/l)[b]	5.0	(4.6, 5.5)	5.2	(4.8,5.9)	<0.001	4.9	(4.5,5.3)	5.2	(4.7,5.7)	<0.001

Data presented as mean (standard errors) for continuous outcomes and percentages for binary outcomes

[a] P values from linear regression (continuous outcomes) and logistic regression (for binary outcomes) with robust standard errors to account for sibling pairs

[b] Median and interquartile range presented, p value from linear regression with log transformed outcome

Table S7 Associations of rs12970134 with age, sex adjusted Z-scores of obesity traits

	Coeff[a]	95% CI		P
BMI	0.03	(-0.03,	0.09)	0.33
WHR	-0.003	(-0.07,	0.06)	0.92
Waist circumference	0.03	(-0.03,	0.09)	0.29
Weight	0.04	(-0.02,	0.10)	0.15
Hip circumference	0.05	(-0.01,	0.11)	0.09
Body fat	0.04	(-0.02,	0.10)	0.16

[a] Coefficents represent SD change per minor allele

Table S8 Age, sex adjusted associations of SNPs in *FTO* and *MC4R* genes with obesity

		OR[a]	95% CI		P
FTO	rs9939609	1.08	(0.91,	1.28)	0.39
MC4R	rs17782313	1.19	(1.00,	1.40)	0.05
MC4R	rs12970134	1.13	(0.96,	1.37)	0.15

[a] Odds ratio (OR) from mixed effects logistic regression. OR represents change per minor allele. Obesity defined as BMI>25kg/m^2

Table S9 Interactions between obesity SNPs and sex in associations with obesity traits

	rs9939609			rs17782313			rs12970134		
	Coeff[a]	95% CI	P	Coeff[a]	95% CI	P	Coeff[a]	95% CI	p
BMI	0.09	(-0.04, 0.23)	0.18	-0.03	(-0.16, 0.11)	0.69	-0.05	(-0.18, 0.08)	0.48
WHR	-0.04	(-0.18, 0.10)	0.61	0.02	(-0.12, 0.16)	0.75	0.01	(-0.13, 0.14)	0.90
Waist circumference	0.06	(-0.08, 0.20)	0.39	-0.01	(-0.15, 0.12)	0.86	-0.02	(-0.16, 0.11)	0.75
Weight	0.12	(-0.01, 0.26)	0.07	-0.02	(-0.15, 0.11)	0.72	-0.06	(-0.19, 0.07)	0.36
Hip circumference	0.11	(-0.03, 0.24)	0.12	-0.04	(-0.17, 0.09)	0.57	-0.05	(-0.18, 0.08)	0.45
Body fat	0.08	(-0.06, 0.21)	0.27	0.01	(-0.12, 0.15)	0.84	0.004	(-0.13, 0.13)	0.96

[a] Coefficients represent differences in age, sex adjusted SD scores per minor allele in females compared to men

Table S10 Interactions between rs12970134 and rural/urban dwelling in associations with obesity phenotypes

	Coeff[a]	95% CI		p
BMI	0.03	(-0.10	0.17)	0.66
WHR	-0.04	(-0.18	0.10)	0.61
Waist circumference	-0.01	(-0.15	0.13)	0.89
Weight	0.05	(-0.08	0.19)	0.43
Hip circumference	0.02	(-0.11	0.16)	0.74
Body fat	-0.03	(-0.16	0.10)	0.66

[a] Coefficients represent differences in age, sex adjusted SD scores per minor allele in urban compared to rural dwellers

Table S11 Interactions of *FTO* with tertiles of physical activity in associations with BMI and weight

		BMI			Weight		
	Tertile	Coeff[a]	95% CI	P	Coeff[a]	95% CI	P
Total METS (hr/day)	1						
	2	0.08	(-0.09, 0.24)	0.37	0.08	(-0.09, 0.24)	0.36
	3	0.01	(-0.15, 0.17)	0.92	-0.01	(-0.17, 0.15)	0.90
Time spent in MVPA (min/day)	1						
	2	-0.02	(-0.18, 0.15)	0.84	0.02	(-0.14, 0.19)	0.79
	3	-0.05	(-0.21, 0.11)	0.56	-0.09	(-0.25, 0.07)	0.27
METS from MVPA (hr/day)	1						
	2	-0.02	(-0.19, 0.15)	0.81	0.01	(-0.15, 0.18)	0.87
	3	-0.06	(-0.23, 0.10)	0.44	-0.09	(-0.25, 0.07)	0.25
Dietary fat intake	1						
	2	-0.01	(-0.15, 0.17)	0.90	0.05	(-0.11, 0.21)	0.55
	3	-0.02	(-0.18, 0.14)	0.81	-0.002	(-0.16, 0.15)	0.98

Abbreviations: METS: Metabolic equivalent tasks, MVPA: Moderate to vigorous physical activity

[a] Coefficients represent differences in age, sex adjusted SD scores per minor allele by tertile of physical activity or dietary fat intake

Table S12 MX analyses: Association between obesity SNPs and traits

Trait	SNP	Beta(int)[a]	Chi square	P value
BMI	rs12970134	.0114	.157	.692
BMI	rs17782313	.0272	.884	.347
BMI	FTO	.0779	7.040	.008
Body fat	rs12970134	.0228	.661	.416
Body fat	rs17782313	.0344	1.463	.227
Body fat	FTO	.0146	.249	.618
Hip	rs12970134	.0328	1.384	.239
Hip	rs17782313	.0491	3.028	.082
Hip	FTO	.0492	2.907	.088
Waist	rs12970134	.0150	.253	.615
Waist	rs17782313	.0215	.506	.477
Waist	FTO	.0424	1.957	.163
Weight	rs12970134	.0244	.783	.376
Weight	rs17782313	.0420	2.263	.133
Weight	FTO	.0850	8.948	.003
WHR	rs12970134	-.0141	.200	.655
WHR	rs17782313	-.0206	.414	.520
WHR	FTO	.0079	.061	.806

[a] Coefficients represent SD change per copy of minor allele

Table S13 MX analyses: Interactions between rural/urban location and genetic variants in their associations with obesity traits

Trait	SNP	Beta(int)[a]	Chi square	P value
BMI	rs12970134	.0315	0.205	.651
BMI	rs17782313	-.0036	0.003	.959
BMI	FTO	.0727	1.015	.314
Body fat	rs12970134	-.0250	0.130	.718
Body fat	rs17782313	-.0499	0.509	.476
Body fat	FTO	.0620	.735	.391
Hip	rs12970134	.0236	.117	.732
Hip	rs17782313	-.0057	.007	.935
Hip	FTO	.1115	2.408	.121
Waist	rs12970134	-.0101	.020	.886
Waist	rs17782313	-.0332	.213	.644
Waist	FTO	.0922	1.581	.209
Weight	rs12970134	.0553	.650	.420
Weight	rs17782313	.0203	.085	.770
Weight	FTO	.1525	4.571	.033
WHR	rs12970134	-.0345	.227	.633
WHR	rs17782313	-.0291	.156	.693
WHR	FTO	.003	.002	.968

[a] Coefficients represent difference in urban compared to rural dwellers in per minor allele effect