

AMP-kinase $\alpha 2$ subunit gene *PRKAA2* variants are associated with total cholesterol, low-density lipoprotein-cholesterol and high-density lipoprotein-cholesterol in normal women

N J Spencer-Jones, D Ge, H Snieder, U Perks, R Swaminathan, T D Spector, N D Carter, S D O'Dell



This article is available free on JMG online via the **JMG Unlocked** open access trial, funded by the Joint Information Systems Committee. For further information, see <http://jmg.bmjournals.com/cgi/content/full/42/2/97>



J Med Genet 2006;43:936–942. doi: 10.1136/jmg.2006.041988

Background: 5'-AMP-activated protein kinase (AMPK) inactivates critical enzymes in fatty acid and cholesterol synthesis. We hypothesised that the serum lipid profile may be influenced by genetic variation in the AMPK catalytic $\alpha 2$ subunit.

Method: We examined association of 5 tagging SNPs (tSNPs) in the *PRKAA2* gene with serum lipids in 2777 normal Caucasian females (mean age 47.4 ± 12.5 years).

Results: All tSNPs were associated with total- and LDL-cholesterol, ($p < 0.001$ to 0.034), explaining variances of 0.13–0.59% and 0.11–0.55% respectively. One haplotype (frequency 34.7%) showed lower total- and LDL-cholesterol compared with the most common haplotype (frequency 45.7%) ($p \leq 0.001$), explaining 0.78% of total- and 0.75% of LDL-cholesterol. Another haplotype (frequency 10.5%) was significantly associated with lower HDL-cholesterol ($p = 0.005$), explaining 0.59% of variance.

Conclusions: *PRKAA2* gene variants are significantly associated with serum lipoproteins in a large sample of normal female Caucasians.

Key points

- Association study of five tagging single-nucleotide polymorphisms (tSNPs) in the 5'-AMP-activated protein kinase $\alpha 2$ subunit gene *PRKAA2* in 2777 normal Caucasian women (mean (SD) age 47.4 (12.5) years).
- All tSNPs were significantly associated with total cholesterol and low-density lipoprotein (LDL)-cholesterol ($p < 0.001$ to 0.034).
- One haplotype was associated with lower total cholesterol and LDL-cholesterol (all $p \leq 0.001$) and another with lower high-density lipoprotein-cholesterol ($p = 0.005$), compared with the common haplotype.

AMPK is a heterotrimeric complex comprising a catalytic subunit (α) and two regulatory subunits (β and γ). In humans, there are isoforms of all three subunits, encoded by separate genes. Compared with the wild type, AMPK $\alpha 2$ -knockout mice exhibit increased body weight and fat mass when exposed to a high-fat diet,⁷ and have impaired insulin sensitivity.⁸ Expression of a constitutively active mutant AMPK α subunit in insulin-resistant human HepG2 cells has been shown to decrease intracellular triglyceride and cholesterol.⁹

We hypothesised that regulation of the lipid profile, affecting body fat deposition and sensitivity of target organs to insulin, may be potentially mediated by differential expression or activity of the AMPK $\alpha 2$ subunit. No population studies have examined variations in the AMPK $\alpha 2$ subunit gene *PRKAA2* with respect to serum lipids or measures of body fat. In a recent study, Sun *et al*¹⁰ found no association between *PRKAA2* tagging single-nucleotide polymorphisms (tSNPs) and type 2 diabetes in five independent Caucasian samples, nor with serum glucose and insulin in normal subjects. In another recent study, Horikoshi *et al*¹¹ found no independent associations between *PRKAA2* single-nucleotide polymorphisms (SNPs) and type 2 diabetes in Japanese subjects, but found one associated haplotype, which was

Known as the “fuel gauge of the cell”,¹ 5'-AMP-activated protein kinase (AMPK), is activated by an increase in the cellular AMP:ATP ratio after ATP depletion. Once activated, AMPK switches on ATP-generating (catabolic) pathways and switches off ATP-consuming (anabolic) pathways, allowing the cell to restore its energy balance.² In addition to its response to acute changes in energy levels in individual cells, AMPK also plays an important part in the regulation of whole-body energy metabolism (see Kahn *et al*³ for review). For example, in the liver, AMPK phosphorylates and inactivates both acetyl-coenzyme A carboxylase and 3-hydroxy-3-methyl-glutaryl-coenzyme A (HMG-CoA) reductase.⁴ Acetyl-coenzyme A carboxylase catalyses the biosynthesis of malonyl-CoA, which is the initial substrate for fatty acid synthesis and an inhibitor of carnitine palmitoyl transferase-1, the rate-limiting step for mitochondrial fatty acid oxidation.⁵ HMG-CoA reductase is the rate-limiting enzyme of cholesterol synthesis.⁶ AMPK activity therefore inhibits fatty acid and cholesterol biosynthesis and promotes fatty acid oxidation, thereby opposing intracellular lipid accumulation and development of insulin resistance in non-adipose tissues.

Abbreviations: AMPK, 5'-AMP-activated protein kinase; Apo, apolipoprotein; GEE, generalised estimating equation; HDL, high-density lipoprotein; HMG-CoA, 3-hydroxy-3-methyl-glutaryl-coenzyme A; LDL, low-density lipoprotein; MAF, minor allele frequency; PCR, polymerase chain reaction; SNP, single-nucleotide polymorphism; tSNP, tagging single-nucleotide polymorphism

replicated in three case-control groups. Neither study measured serum lipids. Recently, Xu *et al*¹² found that rare homozygotes of an intronic SNP in the regulatory AMPK $\gamma 2$ subunit gene *PRKAG2* had significantly higher serum concentrations of triglycerides and total cholesterol than common allele carriers ($n = 290$; $p = 0.01$). In Chinese people, therefore, it seems that AMPK $\gamma 2$ polymorphisms might be associated with lipid metabolism. The $\gamma 2$ isoform has broad tissue distribution, but the only known functional mutations interfere with AMP activation of AMPK and cause hypertrophic cardiomyopathy associated with Wolff-Parkinson-White syndrome.¹³

We have used a multistep design in a large study sample to examine association of *PRKAA2* variants with adiposity and lipid phenotypes, on a gene-wide level.¹⁴ tSNPs effectively capture information of all common variants by taking into account patterns of linkage disequilibrium across the gene.^{15, 16} We tested their effect, individually and as haplotypes, on a range of variables, including anthropometry, serum lipids and body fat distribution, in a sample of Caucasian female twins ($n = 2771$, mean (SD) age 47.4 (12.5) years) and established strong associations across the gene with serum low-density lipoprotein (LDL)-cholesterol and total cholesterol levels. We have attempted a fine mapping analysis using the tSNPs to predict potential associations between the remaining (ungenotyped) SNPs and serum cholesterol to suggest potential functional sites.¹⁵

METHODS

Subjects

The Twins UK Registry comprises unselected, mostly female volunteers ascertained from the general population through national media campaigns in the UK.¹⁷ Means and ranges of

quantitative phenotypes in Twins UK are normally distributed and similar to the age-matched general population in the UK.¹⁸ Informed consent was obtained from all participants before they entered the studies, which were approved by the local research ethics committee. Table 1 shows the characteristics of the subjects. Sixteen subjects using lipid-lowering agents were excluded for the analysis of lipid profiles.

Zygoty, body composition and biochemical analyses

Zygoty in the Twins UK sample was determined by a standardised questionnaire and confirmed by DNA fingerprinting. Serum leptin concentration determined after an overnight fast, body fat measures determined by dual x ray absorptiometry and fasting insulin and glucose measurements including oral glucose tolerance test were as described previously.¹⁹ Blood samples were collected for determination of fasting lipids after a minimum 8-h overnight fast. Serum was stored at -40°C until analysed using a Cobas Fara machine (Roche Diagnostics, Lewes, UK).²⁰ A colorimetric enzymatic method was used to determine total cholesterol, triglycerides and high-density lipoprotein (HDL)-cholesterol levels. The HDL-cholesterol level was measured after precipitation from chylomicron, LDL and very low-density lipoprotein particles by magnesium and dextran sulphate. Apolipoprotein (Apo)A1 and ApoB were assayed by an immunoturbidometric method. The Friedewald equation²¹ was used to calculate LDL-cholesterol levels in subjects with triglycerides ≤ 4.52 mmol/l.

Genotyping for SNP validation and tSNP selection

We used the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/SNP/>) and Children's Hospital Informatics Program Bioinformatics (<http://snpper.chip.org>) databases to choose SNPs for validation and tSNP selection in our cohort. A total of 27 SNPs validated in the database populations were selected for confirmation in the Twins UK cohort, by genotyping 16 random unrelated subjects, using polymerase chain reaction (PCR) and restriction digest. Primers and PCR conditions are given in supplementary table A (available at <http://jmg.bmjournals.com/supplemental>). SNPs that were heterozygotic in at least one sample were further genotyped in 94 unrelated subjects for tSNP selection. Five tSNPs (rs1124900, rs2796516, rs2746342, rs2796498 and rs1418442) were selected. These cover all common *PRKAA2* SNPs (minor allele frequency (MAF) $>6\%$) on the HapMap database (HapMap Data Rel#20/phase II Jan 2006; <http://www.hapmap.org>), with $R_L^2 > 95\%$.¹⁵ Of the 34 HapMap SNPs, eight were not polymorphic and six had MAF $<6\%$. Twenty with MAF $>6\%$ included our five tSNPs.

Genotyping in cohorts

Of the five tSNPs selected, three (rs2796516, rs2746342 and rs2796498) were genotyped by Pyrosequencing (Biotage, Uppsala, Sweden). Two SNPs refractory to genotyping by Pyrosequencing (rs1124900 and rs1418442) were genotyped in the complete cohort by KBiosciences (Hoddesdon, Hertfordshire, UK) using the KASPar system. This is a fluorescence-based allele-specific PCR with improved robustness and discriminating power over conventional amplification refractory mutation systems (<http://www.kbioscience.co.uk/chemistry/chemistry-intro.htm>).

One microplate (94 samples) for each SNP was genotyped by restriction digest as a check on the calls generated by KASPar and complete correspondence was found. Genotyping accuracy for all SNPs was assessed by inclusion of duplicates (pairs of monozygotic twins) in the arrays and negative controls (water blanks) on each plate. Primers and PCR

Table 1 Characteristics of subjects

Variable	n	Mean (SD)
Age (years)*	2771	47.4 (12.5)
Postmenopausal (%)	2449	47.7
Obesity-related variables		
Leptin (ng/ml)	2771	16.5 (11.9)
BMI (kg/m ²)	2755	24.7 (4.4)
Weight (kg)	2756	65.3 (11.8)
Waist (cm)	2704	78.4 (10.2)
Total fat (kg)	2715	23.4 (8.8)
Total fat (%)	2674	35.6 (8.0)
Central fat (kg)	2672	1.33 (0.73)
Central fat (%)	2672	31.1 (11.5)
Lipid profile		
Total cholesterol (mmol/l)	2595	5.56 (1.25)
LDL-cholesterol (mmol/l)	2453	3.46 (1.15)
HDL-cholesterol (mmol/l)	2604	1.55 (0.39)
Triglyceride (mmol/l)	2483	1.27 (0.80)
ApoA1 (g/l)	2431	1.70 (0.34)
ApoB (g/l)	2449	1.17 (0.36)
Insulin sensitivity†		
Fasting glucose (mmol/l)	1123	4.39 (0.46)
Fasting insulin ($\mu\text{U/ml}$)	1009	6.26 (4.40)
HOMA	1005	1.24 (0.89)
2-h glucose (mmol/l)	739	5.18 (1.10)
2-h insulin ($\mu\text{U/ml}$)	739	34.2 (25.4)
SiM	739	88.5 (68.8)

Apo, apolipoprotein; BMI, body mass index; HDL, high-density lipoprotein; HOMA, homeostasis model assessment; LDL, low-density lipoprotein; SiM, sensitivity to insulin measure.

Number of subjects (846 monozygotic, 1925 dizygotic) with leptin data and genotype data on at least one single-nucleotide polymorphism.

†Non-fasting subjects, patients with either type 1 or type 2 diabetes, patients on any drugs for diabetes and subjects with fasting glucose >7.8 mmol/l or 2 h glucose >11.1 mmol/l were all excluded.

conditions for SNP genotyping in the full cohort by Pyrosequencing are given in supplementary table B.

Selection of tSNPs and prediction

We used two approaches to identify an optimal subset of tSNPs as described by Stram *et al*¹⁶ and Chapman *et al*.¹⁵ The main difference is that *tagsnps*¹⁶ is based on the prediction of extended haplotypes (in this case based on 21 SNPs) from the marker haplotypes (in this case based on five tSNPs), whereas *htSNP2*¹⁵ is based on the prediction of single SNP loci. Chapman *et al*'s¹⁵ method can also be used to predict the association of the ungenotyped SNPs with the trait, which can be useful for fine mapping. Furthermore, results of a multilocus test including all tSNPs is given, which provides a global or gene-wide significance level.²² Full details of these methods are given elsewhere.¹⁹

Association analyses

Association analyses in the Twin cohort were carried out using generalised estimating equations (GEE),²² which allows for the relatedness between twins and yields unbiased standard errors and p values. Each of the SNPs was analysed separately and followed up by haplotype analyses. For individual SNP association analyses, we first carried out a 2-df overall test of genotypic association. In the presence of a significant association, additive, dominant and recessive models (all 1 df) were further tested to find the best mode of inheritance. Age and menopausal status were included as covariates in the models. Details of our approach to test the association of statistically inferred haplotypes with continuous traits have been described previously.²³ We adapted Zaykin *et al*'s²⁴ haplotype trend regression for the analysis of related subjects, such as twins, by replacing the linear regression with the GEE procedure. The probabilities of haplotype pairs were estimated by PHASE V.2.0 software.²⁵ Obesity-related variables included leptin, weight, body mass index, waist, total fat mass, % total fat, central fat mass and % central fat. Factor analysis was used to combine strongly correlated indices of obesity into two measures: one for general obesity (serum leptin, body mass index, weight, total fat mass and % total fat) and the other for central obesity (waist, central fat mass and % central fat). To reduce the likelihood of identifying false-positive associations, we analysed individual obesity-related variables only for those tSNPs showing significant results for at least one of the combined scores. We used two indices of insulin sensitivity: homeostasis model assessment (HOMA), based on fasting insulin and glucose data, and sensitivity insulin measure (SiM), based on both fasting and 2-h data, as described previously.²⁶ Serum lipids included total cholesterol, LDL-cholesterol and HDL-cholesterol, triglycerides, and ApoA1 and ApoB.

Preliminary analyses were carried out using STATA V.8. Where needed, phenotypic variables were log transformed to obtain better approximations of the normal distribution before analysis. The Hardy–Weinberg equilibrium was tested by a χ^2 test with 1 df in one twin of each pair chosen at random to prevent inflated significance. Pairwise linkage disequilibrium coefficients in the tSNP test sample were calculated using GOLD and reported as D' and r^2 .²⁷

RESULTS

SNP validation and tSNP selection

Of the 27 SNPs validated in the National Center for Biotechnology Information and Children's Hospital Informatics Program Bioinformatics database populations that we tested in 16 subjects, four were not polymorphic (rs6671760, rs2051040, rs12749128 and rs2179761) and two occurred at frequencies of <6% (rs857155 and rs3738568). In all, 21 SNPs were polymorphic and were genotyped in 94

subjects for tSNP selection. Figure 1 shows the positions of these 21 SNPs. All were located in non-coding regions, and their genotype frequencies were consistent with Hardy–Weinberg proportions. The SNPs had MAFs >5% and showed strong pairwise linkage disequilibrium (supplementary table C). The presence of strong linkage disequilibrium throughout the gene suggested the feasibility of tSNP selection, facilitating great savings in time and costs.

The inferred haplotypes of the 21 SNPs ($n = 94$) and tSNP selection based on haplotype prediction (R_H^2) and single SNP prediction (R_L^2) are shown in supplementary table D. Of the 14 haplotypes present, 2 were common (32.5% and 42.8%), 4 had a frequency of 2.4–7.9% and the remainder showed a frequency of $\leq 1.2\%$. Five tSNPs (rs1124900, rs2796516, rs2746342, rs2796498 and rs1418442) were selected by *htSNP2*¹⁵ based on single SNP prediction (R_L^2), with the criterion $R_L^2 \geq 0.80$, and by *tagsnps*¹⁶ based on haplotype prediction (R_H^2). This set of tSNPs accurately predicted both common haplotypes and unmeasured loci. The minimum value of R_H^2 was 0.93 (for haplotype 2) and that of R_L^2 was 0.90 (for rs11206887).

The genotype, allele and haplotype distributions of the five tSNPs in the whole cohort based on one monozygotic and both dizygotic twins genotyped for each pair are shown in supplementary table D. The total number of subjects genotyped for each polymorphism varied slightly and was somewhat lower than 2402, the total number of twins genotyped (ie, one monozygotic and both dizygotic twins of each pair). This was due to unsuccessful amplification of the target sequences for some samples. None of the loci showed deviation from the Hardy–Weinberg equilibrium. The same three most common haplotypes observed in the 94 subjects (supplementary table D) were also identified by the five tSNPs, although the actual estimates of haplotype frequencies and corresponding order varied somewhat between the full cohort and the tSNP selection sample ($n = 94$).

Tests of association of phenotypes with individual tSNP genotypes

Table 2 presents the results of individual SNP association analyses using GEEs.²² All SNPs were significantly associated with LDL-cholesterol and total cholesterol, with p values ranging from <0.001 to 0.034. The five tSNPs explained between 0.11% and 0.55% of variance in LDL-cholesterol, and between 0.13% and 0.59% of variance in total cholesterol. Additionally, rs1124900 was significantly associated with ApoB, rs2796516 with ApoB and serum triglyceride, and rs2796498 with HDL-cholesterol. For the combined central and general obesity scores, we observed a significant association only between rs2796516 and the central obesity score ($p = 0.043$). Follow-up analyses of the individual obesity variables for this SNP yielded significant associations with % total fat, % central fat and central fat mass (table 2).

Tests of association of phenotypes with tSNP haplotypes

Table 3 shows the frequencies for haplotypes of rs1124900–rs2796516–rs2746342–rs2796498–rs1418442 and results of haplotype association tests. Initially, we ran PHASE V.2.0 five times and used output from the run with the best average value for the goodness of fit. We also ran PHASE several times and found that the solutions in terms of haplotype frequencies were virtually identical for our data each time. Therefore, even the rarest two haplotype frequencies (2.3%) that we included in our haplotype analyses are meaningful. Also, because we have such a large sample size ($n = 2254$, 4508 haplotypes), a haplotype frequency of 2.3% still translates to about 104 haplotypes. Furthermore, we ran the haplotype analyses again after including the two rare

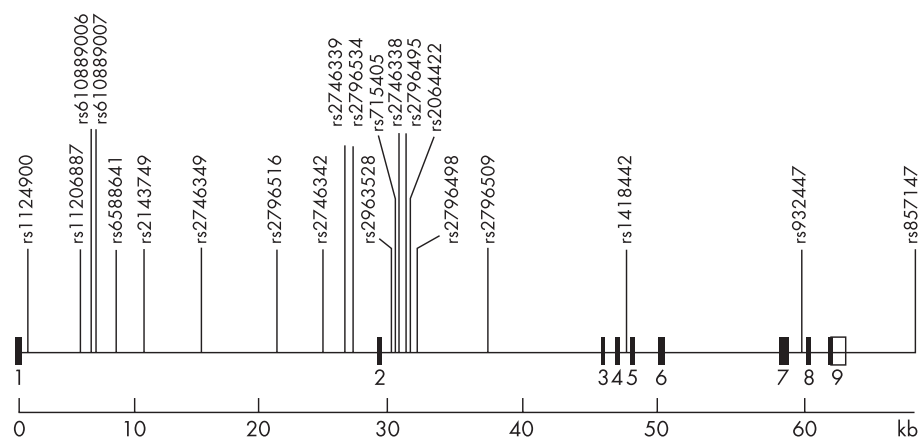


Figure 1 Genomic map of the *PRKAA2* gene with locations of the 21 validated single-nucleotide polymorphisms genotyped in 94 subjects. Solid boxes, exons; open boxes, untranslated regions.

haplotypes of 2.3% in the rest group, but results did not change. Haplotype probabilities were included in the analytical approach as described previously.²³ Haplotype analyses showed results similar to those of the individual SNP associations, with haplotype 2 (frequency 34.7%) showing lower LDL-cholesterol and total serum cholesterol ($p \leq 0.001$) compared with the most common haplotype 1 (frequency 45.7%). The β coefficient for haplotype 2 on LDL-cholesterol was -0.235 and on total cholesterol was -0.257 —that is, LDL-cholesterol level was 0.235 mmol/l (0.21 SD) and total cholesterol level was 0.257 mmol/l (0.21 SD) lower for protective haplotype 2 homozygotes compared with common

haplotype 1 homozygotes. The explained percentages of variance in LDL-cholesterol and total cholesterol by *PRKAA2* haplotypes were 0.75% and 0.76%, respectively.

Haplotype 3 (frequency 10.5%) showed significantly lower HDL-cholesterol compared with common haplotype 1 ($p = 0.005$). HDL-cholesterol levels were 0.111 mmol/l (0.29 SD) lower for risk haplotype 3 homozygotes compared with common haplotype 1 homozygotes. The explained percentage of variance in HDL-cholesterol by *PRKAA2* haplotypes was 0.59%. Overall p values of haplotype tests were significant for total-cholesterol ($p = 0.002$), LDL-cholesterol ($p = 0.01$) and HDL-cholesterol ($p = 0.05$).

Table 2 Association of *PRKAA2* tagging single-nucleotide polymorphisms with obesity-related phenotypes and lipid profile

tSNP	Phenotypes	Number			Mean (SD)			Genetic model	Var (%)	p Value
		11/12/22	11	12	22	11	12			
rs1124900	Total cholesterol (mmol/l)	650/1137/551	5.73 (1.24)	5.52 (1.26)	5.50 (1.21)	Dominant	0.59	<0.001		
	LDL-cholesterol (mmol/l)	612/1074/523	3.61 (1.14)	3.41 (1.15)	3.42 (1.12)	Dominant	0.55	<0.001		
	ApoB (g/l)	608/1084/509	1.20 (0.36)	1.17 (0.35)	1.16 (0.36)	Dominant	0.17	0.022		
rs2796516	Total cholesterol (mmol/l)	1464/518/31	5.58 (1.26)	5.47 (1.20)	6.07 (1.53)	Recessive	0.13	0.017		
	LDL-cholesterol (mmol/l)	1379/487/30	3.46 (1.15)	3.36 (1.10)	3.88 (1.44)	Recessive	0.11	0.034		
	Triglyceride (mmol/l)	1395/494/31	1.24 (0.74)	1.27 (0.90)	1.57(1.18)	Recessive	0.20	0.027		
	ApoB (g/l)	1364/496/27	1.17 (0.36)	1.14 (0.34)	1.32 (0.47)	Codominant	0.29	0.018		
	Total fat (%)	1501/543/33	35.5 (7.9)	34.6 (7.9)	37.1 (9.3)	Dominant	0.11	0.046		
	Central fat (kg)	1513/549/33	1.32 (0.7)	1.24 (0.69)	1.64 (1.06)	Codominant	0.26	0.033		
	Central fat (%)	1513/549/33	31.1 (11.4)	29.8 (11.3)	33.3 (12.2)	Dominant	0.10	0.021		
rs2746342	Total cholesterol (mmol/l)	861/1119/302	5.65 (1.26)	5.51 (1.24)	5.44 (1.25)	Additive	0.32	0.002		
	LDL-cholesterol (mmol/l)	818/1063/287	3.53 (1.16)	3.43 (1.14)	3.31 (1.15)	Additive	0.36	0.003		
rs2796498	Total cholesterol (mmol/l)	652/1156/455	5.68 (1.26)	5.52 (1.25)	5.41 (1.18)	Additive	0.58	<0.001		
	LDL-cholesterol (mmol/l)	614/1089/429	3.53 (1.16)	3.41 (1.15)	3.32(1.08)	Additive	0.37	0.004		
	HDL-cholesterol (mmol/l)	656/1161/458	1.59 (0.40)	1.55 (0.39)	1.53(0.39)	Dominant	0.32	0.01		
rs1418442	Total cholesterol (mmol/l)	651/1120/509	5.51 (1.23)	5.51 (1.25)	5.74 (1.27)	Recessive	0.29	0.001		
	LDL-cholesterol (mmol/l)	618/1054/480	3.42 (1.12)	3.41 (1.16)	3.63 (1.16)	Recessive	0.38	0.001		

HDL, high-density lipoprotein; LDL, low-density lipoprotein; tSNP, tagging single-nucleotide polymorphisms; Var, variance.

Table 3 Association of *PRKAA2* haplotypes (frequency >2%) with lipid profile

Haplotype	Freq. (%)	ApoB			Total cholesterol			LDL- cholesterol			HDL-cholesterol		
		β (SE)	p Value	Variance explained (%)	β (SE)	p Value	Variance explained (%)	β (SE)	p Value	Variance explained (%)	β (SE)	p Value	Variance explained (%)
11112*	45.7	—	—	—	—	—	—	—	—	—	—	—	—
21221	34.7	-0.038 (0.019)	0.034	—	-0.257 (0.071)	<.001	—	-0.235 (0.069)	0.001	—	-0.006 (0.025)	NS	—
22121	10.5	-0.022 (0.029)	NS	—	-0.082 (0.111)	NS	—	-0.025 (0.107)	NS	—	-0.111 (0.039)	0.005	—
11211	2.3	0.032 (0.058)	NS	—	0.337 (0.229)	NS	—	0.248 (0.221)	NS	—	0.0421 (0.080)	NS	—
12111	2.3	0.037 (0.061)	NS	—	0.157 (0.244)	NS	—	-0.177 (0.239)	NS	—	0.067 (0.0834)	NS	—
Overall	—	—	NS	—	—	0.002	0.76	—	0.01	0.75	—	0.05	0.59

Apo, apolipoprotein; Freq, frequency.

*The most common haplotype with which the others were compared.

Predictive associations with ungenotyped SNPs

The approach we used to select an optimal set of tSNPs¹⁵ can also predict which of those SNPs not typed in the full cohort might also show association with the trait. Figure 2 shows the fine mapping results (including the ungenotyped SNPs) based on prediction of single SNP loci from the five tSNPs genotyped in the Twins UK cohort. Associations of both tSNPs and predicted values for unmeasured SNPs with ApoB, HDL-cholesterol, LDL-cholesterol and total cholesterol under the additive genetic model are indicated by $-\log(P)$. Except for three SNPs, the tSNP rs2796516 and the ungenotyped SNPs rs11206887 and rs10889007, consistent associations (all $p < 0.01$) with LDL-cholesterol and total cholesterol levels were observed. For ApoB, a similar but

weaker association pattern (all $p < 0.05$) was observed. Predicted associations for body fat were found to be uninformative and are not shown. Multilocus (or gene-wide) tests, including all tSNPs, were significant for total cholesterol ($p = 0.004$) and LDL-cholesterol ($p = 0.04$).

DISCUSSION

AMPK is a metabolic “master switch”, phosphorylating key target proteins that control flux through metabolic pathways of hepatic ketogenesis, cholesterol synthesis, lipogenesis, triglyceride synthesis, adipocyte lipolysis and skeletal muscle fatty acid oxidation.³ We hypothesised that genetic variation in the AMPK $\alpha 2$ subunit gene *PRKAA2* is associated with regulation of the lipid profile, affecting body fat deposition

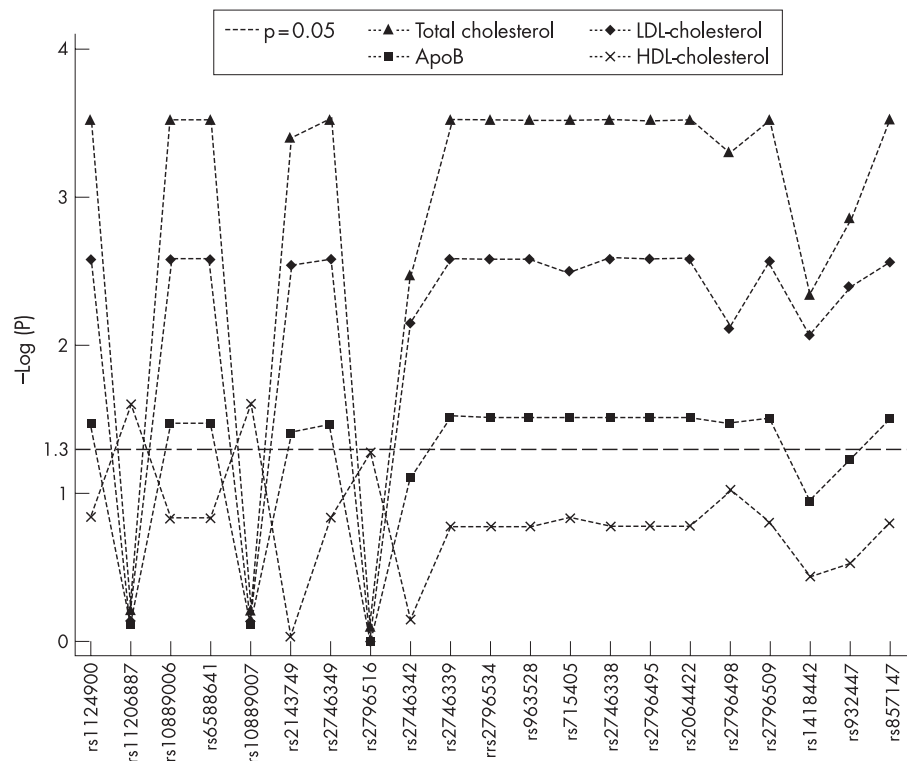


Figure 2 Associations of tagging single-nucleotide polymorphisms (tSNPs) and predicted values for unmeasured SNPs and tSNPs with apolipoprotein (Apo)B, low-density lipoprotein (LDL)-cholesterol, high-density lipoprotein (HDL)-cholesterol and total cholesterol under the additive genetic model. Associations of tSNPs and predicted values for unmeasured SNPs with ApoB, LDL-cholesterol, HDL-cholesterol and total cholesterol under the additive genetic model are indicated by $-\log(P)$. Points above the dotted line representing $-\log(P) 0.05$ represent SNP associations found, or predicted to be significant at the $p = 0.05$ level.

and sensitivity of target organs to insulin, potentially mediated by differential expression or activity of the AMPK $\alpha 2$ subunit. To this end, we have examined association of a set of five tSNPs spanning the AMPK catalytic $\alpha 2$ subunit gene *PRKAA2*, with a range of phenotypes representing body fat, insulin sensitivity and lipids. Minor associations were observed with ApoB, triglyceride, HDL-cholesterol and measures of body fat, and none with the insulin-sensitivity parameters. Therefore, in this study sample, common variation in *PRKAA2* does not seem to be responsible for changes in effects on body weight through mediation of the anorexigenic effect of leptin by down regulation of AMPK in the hypothalamus,²⁷ or for any associated effects on insulin sensitivity via stimulation of AMPK by leptin in muscle.²⁸ However, all tSNPs were most strongly associated with LDL-cholesterol and total cholesterol ($p < 0.001$ to 0.034). A protective haplotype 2 was identified, showing lower LDL-cholesterol and total serum cholesterol ($p \leq 0.001$) compared with the most common haplotype 1. Importantly, global (ie, gene-wide) p values provided by both haplotype analyses and multilocus tests were significant for total cholesterol and LDL-cholesterol.

We have established these associations only in a sample of women, but our findings in these twin subjects can be considered as representative of the UK female population as a whole.¹⁸ We have previously found few differences between twins and singletons in the population generally, the only indication being that monozygotic twins had a slightly lower weight and a smaller variance for weight than dizygotic twins and singletons.¹⁸ This study had 80% ($\alpha = 0.05$) power to detect a biallelic quantitative trait locus, explaining as little as 0.5% of the variance.

In a recent study, Sun *et al*¹⁰ tested the association of nine *PRKAA2* tSNPs for association with type 2 diabetes in five different Caucasian samples, the largest of which comprised 1189 discordant sibs, and found no associations. They also tested glucose and insulin sensitivity in 756 controls without effect. In another study, Horikoshi *et al*¹¹ tested association of six *PRKAA2* tSNPs in 192 type 2 diabetes cases and 272 control Japanese samples and found no independent SNP associations, but they identified an associated haplotype, which was replicated in two other case-control studies. The minor allele of rs2051040, not among our tSNPs, was associated with the disease. The absence of single SNP associations found in all samples, with type 2 diabetes, insulin sensitivity or glucose, accords with our finding of absence of associations with insulin or glucose in 2771 normal Caucasian women. Neither Sun *et al* nor Horikoshi *et al* investigated the association with serum lipids.

Sun *et al* chose a set of tSNPs different from ours and from that of Horikoshi *et al*. Horikoshi *et al* included one of our tSNPs, rs1418442, in their set. There are several reasons for this, but as the gene lies in a region of high linkage disequilibrium, there is considerable diversity of choice in achieving equal coverage. Firstly, our tSNPs span 65.6 kb rather than the 100 kb tagged by Sun *et al*, which extends from 27 kb upstream of the transcription start site to 12 kb downstream of the 3' untranslated region. Horikoshi *et al* included three new SNPs discovered by them, the total spanning from 1.4 kb upstream of the gene to the 3' untranslated region. We selected no SNPs within the 1 kb region 5' of the start site that we investigated, having failed to validate any in our test sample, and the last was approximately 5 kb 3' of the gene. We chose a higher MAF ($> 6\%$) than either group considering the appropriate power to detect expected effects. Sun *et al* chose nine tSNPs from initially 73 db SNP candidates genotyped in the HapMap CEPH reference panel of 30 parent-offspring trios. Of these, 37 were retained and a further three candidates were added from the HapMap project. Hirokoshi *et al* genotyped 10 SNPs in 192

subjects and selected 6 tSNPs. We genotyped 27 dbSNP candidates in 96 unrelated people drawn from Twins UK and applied the methods of Chapman *et al*¹⁵ and Stram *et al*¹⁶ to select five tSNPs as described. We confirmed that our set of tSNPs cover all SNPs with MAF $> 6\%$ in the CEPH genotype data of the Phase II HapMap (January 2006) release, with $R^2 > 95\%$; so our coverage of the *PRKAA2* gene is as comprehensive as current data allow. In regions of high linkage disequilibrium such as this, we find that if Haploview/Tagger is run twice for the same region, using the same parameters, different tSNPs will be reported but coverage on both occasions remains almost the same. Hence, the tSNP sets chosen by the other groups and ourselves can be considered equal to the task of tagging common functional variants.

Strong linkage disequilibrium across the gene allowed us to effectively capture common variation by the selected set of tSNPs, making it unlikely that we may have missed any major variants that are either causal or in strong linkage disequilibrium with a causal locus. All of our associated tSNPs are intronic and are most likely to be linkage disequilibrium markers, although an increasing number of intronic regulatory variants are emerging. From our predictive analysis (fig 2), we showed fairly consistent association along the entire gene, with all but three SNPs providing signal. This is a consequence of strong linkage disequilibrium in the region and, although not providing informative fine-mapping peaks to identify functional alleles, this procedure does indicate a strong relationship between *PRKAA2* variants and serum cholesterol levels. A rare functional variant with large effect is unlikely, because our tSNPs tag only common SNPs (MAF $> 6\%$) satisfactorily. A *cis*-interacting variant is possible, especially as the intronic SNPs could be regulatory.

HMG-CoA reductase, the rate-limiting enzyme of cholesterol synthesis,⁶ is a highly regulated enzyme. Phosphorylation and down regulation of its activity by AMPK⁴ and reversal by dephosphorylation mediated mainly by protein phosphatase 2A²⁹ are responsible for rapid regulation of the enzyme's activity. Recently, it has been found that cholesterol 7 α -hydroxylase, the rate-limiting enzyme in the conversion of cholesterol to bile acids, is also phosphorylated and down regulated by AMPK in HepG2 cells.³⁰ Thus, cholesterol synthesis and disposal are coregulated via the phosphorylation of key enzymes by AMPK.

The association between five individual *PRKAA2* tSNPs and a haplotype with lower LDL-cholesterol and total serum cholesterol ($p \leq 0.001$) suggests that genetic variation in *PRKAA2* may influence the ability of AMPK to phosphorylate and down regulate HMG-CoA reductase. One possibility is that one or more variants responsible for more efficient phosphorylation and deactivation of the enzyme by AMPK could be found on the *PRKAA2* haplotype 2 background. Identification of functional AMPK variants capable of reducing serum cholesterol levels could have potential therapeutic value.



Supplementary tables A, B and C are available at <http://jmg.bmjournals.com/supplemental>.

Authors' affiliations

N J Spencer-Jones, S D O'Dell, Nutrition Food and Health Research Centre, King's College London, London, UK
D Ge, H Snieder, Georgia Prevention Institute, Department of Pediatrics, Medical College of Georgia, Augusta, Georgia, USA
U Perks, R Swaminathan, Department of Clinical Chemistry, St Thomas' Hospital, King's College London, London, UK
T D Spector, Twin Research and Genetic Epidemiology Unit, St Thomas' Hospital, London, UK

N D Carter, Division of Clinical Developmental Sciences, St George's University of London, London, UK

Funding: This study was funded by the British Heart Foundation Project grant number PG/04/028 and Wellcome Trust Project grant number 073142. The Twin Research and Genetic Epidemiology Unit received support from the Wellcome Trust, Arthritis Research Campaign, the Chronic Disease Research Foundation and the European Union 5th Framework Programme Genom EU twin number QL2-CT-2002-01254.

Competing interests: None.

NJS-J and DG are joint first authors.

Correspondence to: S O'Dell, Nutrition Food and Health Research Centre, King's College London, Franklin-Wilkins Building, 150 Stamford Street, London SE1 9NH, UK; sandra.o'dell@kcl.ac.uk

Received 24 February 2006

Revised 25 May 2006

Accepted 5 June 2006

Published Online First 26 June 2006

REFERENCES

- 1 **Hardie DG**, Carling D. The AMP-activated protein kinase - fuel gauge of the mammalian cell? *Eur J Biochem* 1997;**246**:259-73.
- 2 **Hardie DG**, Scott JW, Pan DA, Hudson ER. Management of cellular energy by the AMP-activated protein kinase system. *FEBS Lett* 2003;**546**:113-20.
- 3 **Kahn BB**, Alquier T, Carling D, Hardie DG. AMP-activated protein kinase: ancient energy gauge provides clues to modern understanding of metabolism. *Cell Metab* 2005;**1**:15-25.
- 4 **Hardie DG**. Regulation of fatty acid and cholesterol metabolism by the AMP-activated protein kinase. *Biochim Biophys Acta* 1992;**1123**:231-8.
- 5 **Kim KH**. Regulation of acetyl-CoA carboxylase. *Curr Top Cell Regul* 1983;**22**:143-76.
- 6 **Goldstein JL**, Brown MS. Regulation of the mevalonate pathway. *Nature* 1990;**343**:425-30.
- 7 **Villena JA**, Viollet B, Andreelli F, Kahn A, Vaulont S, Sul HS. Induced adiposity and adipocyte hypertrophy in mice lacking the AMP-activated protein kinase-alpha2 subunit. *Diabetes* 2004;**53**:2242-9.
- 8 **Viollet B**, Andreelli F, Jorgensen SB, Perrin C, Geloan A, Flamez D, Mu J, Lenzner C, Baud O, Bennoun M, Gomas E, Nicolas G, Wojtaszewski JF, Kahn A, Carling D, Schuit FC, Birnbaum MJ, Richter EA, Burcelin R, Vaulont S. The AMP-activated protein kinase alpha 2 catalytic subunit controls whole-body insulin sensitivity. *J Clin Invest* 2003;**111**:91-8.
- 9 **Zang M**, Zuccollo A, Hou X, Nagata D, Walsh K, Herscovitz H, Brecher P, Ruderman NB, Cohen RA. AMP-activated protein kinase is required for the lipid-lowering effect of metformin in insulin-resistant human HepG2 cells. *J Biol Chem* 2004;**279**:47898-905.
- 10 **Sun MW**, Lee JY, de Bakker PI, Burt NP, Almgren P, Rastam L, Tuomi T, Gaudet D, Daly MJ, Hirschhorn JN, Altshuler D, Groop L, Florez JC. Haplotype structures and large-scale association testing of the 5' AMP-activated protein kinase genes PRKAA2, PRKAB1, and PRKAB1 with type 2 diabetes. *Diabetes* 2006;**55**:849-55.
- 11 **Horikoshi M**, Hara K, Ohashi J, Miyake K, Tokunaga K, Ito C, Kasuga M, Nagai R, Kadowaki T. A polymorphism in the AMPK α 2 subunit gene is associated with insulin resistance and type 2 diabetes in the Japanese population. *Diabetes* 2006;**55**:919-23.
- 12 **Xu M**, Li X, Wang JG, Du P, Hong J, Gu W, Zhang Y, Ning G. Glucose and lipid metabolism in relation to novel polymorphisms in the 5'-AMP-activated protein kinase gamma2 gene in Chinese. *Mol Genet Metab* 2005;**86**:372-8.
- 13 **Blair E**, Redwood C, Ashrafian H, Oliveira M, Broxholme J, Kerr B, Salmon A, Ostman-Smith I, Watkins H. Mutations in the gamma(2) subunit of AMP-activated protein kinase cause familial hypertrophic cardiomyopathy: evidence for the central role of energy compromise in disease pathogenesis. *Hum Mol Genet* 2001;**10**:1215-20.
- 14 **Neale BM**, Sham PC. The future of association studies: gene-based analysis and replication. *Am J Hum Genet* 2004;**75**:353-62.
- 15 **Chapman JM**, Cooper JD, Todd JA, Clayton DG. Detecting disease associations due to linkage disequilibrium using haplotype tags: a class of tests and the determinants of statistical power. *Hum Hered* 2003;**56**:18-31.
- 16 **Stram DO**, Haiman CA, Hirschhorn JN, Altshuler D, Kolonel LN, Henderson BE, Pike MC. Choosing haplotype-tagging SNPs based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the Multiethnic Cohort Study. *Hum Hered* 2003;**55**:27-36.
- 17 **Spector TD**, MacGregor AJ. The St. Thomas' UK Adult Twin Registry. *Twin Res* 2002;**5**:440-3.
- 18 **Andrew T**, Hart DJ, Snieder H, de Lange M, Spector TD, MacGregor AJ. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res* 2001;**4**:464-77.
- 19 **Spencer-Jones NJ**, Snieder H, Wang X, Miller CSA, Spector TD, Carter ND, O'Dell SD. Protein tyrosine phosphatase-1B gene *PTPN1*: selection of tagging SNPs and association with body fat, insulin sensitivity and the metabolic syndrome in a normal female population. *Diabetes* 2005;**54**:3296-304.
- 20 **Middelberg RP**, Spector TD, Swaminathan R, Snieder H. Genetic and environmental influences on lipids, lipoproteins, and apolipoproteins: effects of menopause. *Arterioscler Thromb Vasc Biol* 2002;**22**:1142-7.
- 21 **Friedewald WT**, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972;**18**:499-502.
- 22 **Tréguët D-A**, Ducimetière P, Tiret L. Testing association between candidate-gene markers and phenotype in related individuals, by use of estimating equations. *Am J Hum Genet* 1997;**61**:189-99.
- 23 **Dong Y**, Zhu H, Wang X, Dalageorgou C, Carter N, Spector TD, Snieder H. Obesity reveals an association between blood pressure and the G-protein beta3 subunit gene: a study of female dizygotic twins. *Pharmacogenetics* 2004;**14**:419-27.
- 24 **Zaykin DV**, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 2002;**53**:79-91.
- 25 **Stephens M**, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 2003;**73**:1162-9.
- 26 **Abecasis GR**, Cookson WO. GOLD: graphical overview of linkage disequilibrium. *Bioinformatics* 2000;**16**:182-3.
- 27 **Minokoshi Y**, Alquier T, Furukawa N, Kim YB, Lee A, Xue B, Mu J, Foufelle F, Ferre P, Birnbaum MJ, Stuck BJ, Kahn BB. AMP-kinase regulates food intake by responding to hormonal and nutrient signals in the hypothalamus. *Nature* 2004;**428**:569-74.
- 28 **Minokoshi Y**, Kim YB, Peroni OD, Fryer LG, Muller C, Carling D, Kahn BB. Leptin stimulates fatty-acid oxidation by activating AMP-activated protein kinase. *Nature* 2002;**415**:339-43.
- 29 **Gaussin V**, Skarlas P, Ching YP, Hardie DG, Hue L. Distinct type-2A protein phosphatases activate HMGCoA reductase and acetyl-CoA carboxylase in liver. *FEBS Lett* 1997;**413**:115-18.
- 30 **Stroup D**, Ramsaran JR. Cholesterol 7 alpha-hydroxylase is phosphorylated at multiple amino acids. *Biochem Biophys Res Commun* 2005;**329**:957-65.