A Genome-Wide Association Study of the Metabolic Syndrome in Indian Asian Men

Delilah Zabaneh*, David J. Balding**

Department of Epidemiology and Public Health, Imperial College London, London, United Kingdom

Abstract

We conducted a two-stage genome-wide association study to identify common genetic variation altering risk of the metabolic syndrome and related phenotypes in Indian Asian men, who have a high prevalence of these conditions. In Stage 1, approximately 317,000 single nucleotide polymorphisms were genotyped in 2700 individuals, from which 1500 SNPs were selected to be genotyped in a further 2300 individuals. Selection for inclusion in Stage 1 was based on four metabolic syndrome component traits: HDL-cholesterol, plasma glucose and Type 2 diabetes, abdominal obesity measured by waist to hip ratio, and diastolic blood pressure. Association was tested with these four traits and a composite metabolic syndrome phenotype. Four SNPs reaching significance level \( p < 5 \times 10^{-7} \) and with posterior probability of association \( > 0.8 \) were found in genes CETP and LPL, associated with HDL-cholesterol. These associations have already been reported in Indian Asians and in Europeans. Five additional loci harboured SNPs significant at \( p < 10^{-6} \) and posterior probability \( > 0.5 \) for HDL-cholesterol, type 2 diabetes or diastolic blood pressure. Our results suggest that the primary genetic determinants of metabolic syndrome are the same in Indian Asians as in other populations, despite the higher prevalence. Further, we found little evidence of a common genetic basis for metabolic syndrome traits in our sample of Indian Asian men.

Introduction

The metabolic syndrome is the combination of most or all of: raised plasma glucose, abdominal obesity, dyslipidemia, and high blood pressure [1]. People affected by the metabolic syndrome are at an increased risk of coronary heart disease and type 2 diabetes (T2D), which are large and rapidly-increasing causes of illness and death globally. Indian Asians have a high prevalence of the metabolic syndrome compared with Europeans, and metabolic syndrome traits are highly heritable in Indian Asians (\( h^2 \) between 0.27 and 0.53 [2]). It is thought that the syndrome results from a complex interplay of genetic and environmental factors, and genetic variants underlying metabolic traits have been identified at several loci in European-origin populations. Little is known about whether the same genetic mechanisms trigger metabolic disturbances in Indian Asians as in Europeans, nor whether there are genetic mechanisms that are common across metabolic syndrome traits in Indian Asians. Our study was intended to investigate these two questions.

We screened common SNPs for association with metabolic syndrome and four of its component traits in a sample of Indian Asian men. Metabolic syndrome traits vary substantially between men and women; to reduce heterogeneity we included only men in this study. The four traits included three quantitative traits (diastolic blood pressure DBP, waist-hip ratio WHR, and HDL cholesterol), and one binary (presence or absence of T2D). We also tested a metabolic syndrome phenotype that was created from individual metabolic traits. We used a two-stage design, with independent “top and tail” sample selection and different genotyping platforms in each stage. Overall, about 1500 SNPs, primarily selected from the Stage 1 results, were genotyped in approximately 5000 individuals in the two stages combined. We report \( p \)-values of association under an additive model, after adjustment for covariates (see Materials and Methods). Although familiar, \( p \)-values suffer from problems of interpretation and the difficulty of combining signals under different genetic models [3]. We therefore also report the posterior probability of association (PPA) for a 4:1 weighting of additive and general genetic models (see Materials and Methods). We assumed a prior probability of association of \( 10^{-4} \) for each trait, which corresponds to a cautious assumption that only around 300 kb of the genome is in high linkage disequilibrium (LD) with a causal variant. The PPA is a directly interpretable measure of weight of evidence for association, and gives due emphasis to additive genetic models while also allowing strong, non-additive signals of association to be taken into account.

Results

Characteristics of individuals selected for genotyping in each stage are shown in Table 1. Quantile-quantile (Q-Q) and signal intensity plots for Stage 1 results are in Figure S1. In summary, after combining data from stages 1 and 2 (Figure 1), four SNPs at two loci were strongly associated with HDL-cholesterol \( (p < 5 \times 10^{-7}, \text{PPA}>0.8) \). In addition, a TCF7L2 SNP had a PPA of almost 0.7 for T2D \( (p = 7 \times 10^{-7}) \). A further four SNPs were
associated with HDL or DBP at $p<10^{-6}$ and PPA>0.5. These results are further described below and in Table 2, and a list of all SNPs significant at $p<10^{-5}$ from the combined analysis is in Table S1.

Two SNPs in the FTO gene, which has a well-established association with obesity (e.g. [4]), were genotyped in our study (rs9982563 and rs9973132), and showed only weak association with WHR ($p = 4.6 \times 10^{-3}$ and $4.4 \times 10^{-3}$). No SNP showed strong association with the compound metabolic syndrome phenotype. The largest PPA for a SNP associated with the metabolic syndrome as defined by the IDF [1] was 0.08 ($p = 6.8 \times 10^{-6}$) at rs12957347, about 180 kb upstream from gene PMAIP1 (phorbol-12-myristate-13-acetate-induced protein) for which no associations have previously been reported [5], and 288 kb downstream from MC4R (Melanocortin 4 Receptor). Rs12957347 is correlated with rs12970134 reported in [6] ($r^2 = 0.7$) which is 155 kb downstream of MC4R, and gave a p-value = $2.4 \times 10^{-7}$ from the Endoglin (ENG) gene. In our study it showed suggestive evidence of association (PPA = 0.57), each copy of the rare allele reducing DBP by 1.19 mmHg (95% CI: 0.71 to 1.67). An association with DBP is biologically plausible: ENG encodes a type I membrane glycoprotein and is part of the TGF-beta receptor complex. It is crucial for maintaining vascular integrity and has a role in the development of the cardiovascular system [16]. Its expression is regulated during heart development [5]. A large meta-analysis did not show association with variants within or near ENG in Europeans [17]. Although that study included 12,000 Indian Asians, only 12 SNPs were genotyped in these individuals, none of them near ENG.

**Discussion**

The metabolic syndrome and its components are a major health concern, particularly in Indian Asians. The GWAS approach has met with some success in dyslipidemia, type 2 diabetes and obesity phenotypes [4,7,15,17], with most studies to date being conducted in Europeans. Our study has further confirmed a number of previously reported associations, in some cases for the first time in Indian Asians, and identified some novel suggestive associations requiring further confirmation.

The metabolic syndrome consists of a number of phenotypes that tend to co-occur, raising the question of whether or not they have common genetic mechanisms [18,19]. A number of definitions for the metabolic syndrome have been developed over the years, including those proposed by IDF, NCEP ATPIII or WHO [1,20,21]. We chose the IDF definition, which is the most recent and incorporates ethnicity by providing different criteria for the metabolic syndrome in different ethnic groups [1]. Most published associations for the metabolic syndrome are only with individual component phenotypes, or in some cases with multiple phenotypes but not matching any of the above definitions. Joy et al. [22] reviewed a large number of genetic association and linkage studies for the metabolic syndrome using all definitions, and concluded that these studies have not provided any confirmed associations. Our results for Indian Asians also found no evidence for common genetic mechanisms underlying the metabolic syndrome, despite its high prevalence in this population, and the

Table 1. Characteristics of genotyped Indian Asian men.

| Description             | Stage 1 (N = 2684) | Stage 2 (N = 2020) | Combined (N = 4560) |
|-------------------------|-------------------|-------------------|-------------------|
|                         | Mean (SD) or prevalence | Range | Mean (SD) or prevalence | Range | Mean (SD) or prevalence | Range |
| Age (yrs)               | 50.0 (11.0)       | 35.0–74.8         | 47.5 (10.7)       | 35.0–74.8 | 49.9 (10.9)       | 35.0–74.8 |
| CHD (%)                 | 8%                | -                 | 8%                | -          | 8%                | -          |
| Hypertension (%)        | 33%               | -                 | 33%               | -          | 33%               | -          |
| T2D (%)                 | 25%               | -                 | 25%               | -          | 25%               | -          |
| Cholesterol med (%)     | 18%               | -                 | 17%               | -          | 18%               | -          |
| SBP (mmHg)              | 134.2 (20.6)      | 87–231            | 133.8 (19.9)      | 89–244     | 134.1 (20.4)      | 87–244    |
| DBP (mmHg)              | 82.6 (12.1)       | 53–132            | 83.0 (11.9)       | 55–149     | 82.8 (12.0)       | 53–149    |
| WHR                     | 0.97 (0.07)       | 0.55–1.29         | 0.97 (0.07)       | 0.68–1.41  | 0.97 (0.07)       | 0.55–1.41  |
| HDL-cholesterol (mmol/l)| 1.22 (0.31)       | 0.57–3.28         | 1.22 (0.33)       | 0.25–4.85  | 1.22 (0.32)       | 0.25–4.85  |
| Glucose (mmol/l)        | 6.03 (2.18)       | 2.60–21.90        | 6.04 (2.20)       | 2.00–21.40 | 6.04 (2.19)       | 2.00–21.9  |
| Metabolic syndrome (IDF)| 47.0%             | -                 | 49.0%             | -          | 48.0%             | -          |

Prevalence of the metabolic syndrome is for the selected sample and is not representative of the general population.

| Description | Mean (SD) or prevalence | Range |
|-------------|-------------------------|-------|
| Age         | 50.0 (11.0)             | 35.0–74.8 |
| CHD (%)     | 8%                      | - |
| Hypertension (%) | 33%              | - |
| T2D (%)     | 25%                     | - |
| Cholesterol med (%) | 18%             | - |
| SBP (mmHg)  | 134.2 (20.6)            | 87–231 |
| DBP (mmHg)  | 82.6 (12.1)             | 53–132 |
| WHR         | 0.97 (0.07)             | 0.55–1.29 |
| HDL-cholesterol (mmol/l) | 1.22 (0.31) | 0.57–3.28 |
| Glucose (mmol/l) | 6.03 (2.18)    | 2.60–21.90 |
| Metabolic syndrome (IDF) | 47.0% | - |

Diabetic blood pressure. SNP rs7865146 is located<3 kb from the Endoglin (ENG) gene. In our study it showed suggestive evidence of association (PPA = 0.57), each copy of the rare allele reducing DBP by 1.19 mmHg (95% CI: 0.71 to 1.67). An association with DBP is biologically plausible: ENG encodes a type I membrane glycoprotein and is part of the TGF-beta receptor complex. It is crucial for maintaining vascular integrity and has a role in the development of the cardiovascular system [16]. Its expression is regulated during heart development [5]. A large meta-analysis did not show association with variants within or near ENG in Europeans [17]. Although that study included 12,000 Indian Asians, only 12 SNPs were genotyped in these individuals, none of them near ENG.

**Discussion**

The metabolic syndrome and its components are a major health concern, particularly in Indian Asians. The GWAS approach has met with some success in dyslipidemia, type 2 diabetes and obesity phenotypes [4,7,15,17], with most studies to date being conducted in Europeans. Our study has further confirmed a number of previously reported associations, in some cases for the first time in Indian Asians, and identified some novel suggestive associations requiring further confirmation.

The metabolic syndrome consists of a number of phenotypes that tend to co-occur, raising the question of whether or not they have common genetic mechanisms [18,19]. A number of definitions for the metabolic syndrome have been developed over the years, including those proposed by IDF, NCEP ATPIII or WHO [1,20,21]. We chose the IDF definition, which is the most recent and incorporates ethnicity by providing different criteria for the metabolic syndrome in different ethnic groups [1]. Most published associations for the metabolic syndrome are only with individual component phenotypes, or in some cases with multiple phenotypes but not matching any of the above definitions. Joy et al. [22] reviewed a large number of genetic association and linkage studies for the metabolic syndrome using all definitions, and concluded that these studies have not provided any confirmed associations. Our results for Indian Asians also found no evidence for common genetic mechanisms underlying the metabolic syndrome, despite its high prevalence in this population, and the
Figure 1. Genome-wide association results of the combined Stage 1 and Stage 2 analysis.
doi:10.1371/journal.pone.0011961.g001
Table 2. SNPs with p-value $< 10^{-6}$ and PPA $> 0.5$.
binary phenotype with scores 3–5 as cases and scores 0–2 as controls. This did not reveal any new results and is therefore not discussed here.

Genotyping

The Stage 1 samples were genotyped by DeCode, Iceland, using the Illumina 300HumanHap Bead Chip, which includes 317,963 SNPs of which 308,942 are autosomal. Stage 2 samples were genotyped using a custom array and Illumina Golden Gate technology at Imperial College, Hammersmith Hospital. Due to a problem with genotyping, a random set of 500 samples selected for Stage 2 were not genotyped, and analyses proceeded without these samples. This left a total of 2274 samples genotyped at 1370 SNPs. A summary of the genotyping for both stages is in Tables S4 and S5, and Text S2.

Statistical analysis

I. Quality control. Phenotypes and genotype data were subjected to rigorous quality control procedures, detailed in Text S2. In summary, individuals and SNPs were excluded based on genotyping quality, Hardy-Weinberg equilibrium and relatedness. In Stage 1, we investigated the effect of population structure through principal component (PC) analysis. The first two PCs are shown in Figure 2, and indicate a complex pattern of population structure that has some correlation with religious affiliation (Figure 2 top) and language (Figure 2 bottom). Among the first fifteen PCs, only the first four were significantly associated with any phenotype and these four were used to adjust for population structure. Bayes PCs were not available to adjust for population structure. Bayes factors (BFs) were calculated using both additive and general models. For the quantitative traits, the additive model assumed a linear trend in trait mean and constant trait variance with increasing minor allele count (0, 1, or 2), while the general model allowed any changes in both mean and variance over genotypes.

II. Association analysis. In Stage 1, association of 10^5 SNPs with each of the primary phenotypes and the metabolic syndrome phenotype were tested in PLINK [23], using logistic regression for T2D and linear regression for the other traits. For each SNP, both additive and dominant genetic models were tested. To maximise power, each individual was included in the analysis of each phenotype, irrespective of the reason for selecting that individual. Perhaps because of this and our adjustments, despite the top-and-tail selection we found that standard regression p-values based on the Gaussian statistical distribution showed correct type 1 error, as evidenced by good adherence of observed quantiles to their null expectations over all but the upper tail of the distribution (Figure S1).

We selected 1536 SNPs for the Stage 2 Golden Gate custom array, based on a number of conditions: 1433 were significant under the additive model, either at \( p < 10^{-3} \) for one of the four primary phenotypes, or at \( p < 10^{-5} \) for the two metabolic syndrome phenotype, or at \( p < 10^{-5} \) under the recessive model for any phenotype. An additional 103 SNPs were selected based on being proxies \( (r^2 > 0.9) \) for a top-ranked SNP, or for a SNP with low design score, or based on candidate SNPs from the literature. SNPs with \( p > 10^{-5} \) were excluded if they had high LD \( (r^2 > 0.9) \) with a genotyped SNP, or low design score. We tested association of Stage 2 SNPs with each of the primary phenotypes and the metabolic syndrome phenotype.

Data from both stages were pooled and association analysis was carried out on the combined data, in the same way as for Stage 1, including testing for recessive and dominant models, except that PCs were not available to adjust for population structure. Bayes factors (BFs) were calculated using both additive and general models. For the quantitative traits, the additive model assumed a linear trend in trait mean and constant trait variance with increasing minor allele count (0, 1, or 2), while the general model allowed any changes in both mean and variance over genotypes.

For the binary traits (T2D and the two definitions of the metabolic syndrome), the general model is described as BF\(_\text{r}\) in the supplementary material of [3]. Full R code for the quantitative trait BFs is given in Text S3. We used the default prior parameters, which imply that a 95% interval for the mean effect size is approximately \( \pm 0.4 \) phenotype standard deviations, while under the general model the phenotypic variance was allowed to vary over genotypes by about 5%. The two Bayes factors were combined using Bayes theorem to generate the PPA, giving a 4:1 weight in favour of the additive model, and a prior probability of association of 10^{-5} at each SNP. See Text S3 for the full R code that includes all parameter values for all four Bayes factors.

Supporting Information

Text S1
Found at: doi:10.1371/journal.pone.0011961.s001 (0.01 MB DOC)

Text S2
Found at: doi:10.1371/journal.pone.0011961.s002 (0.02 MB DOC)
**Text S3**
Found at: doi:10.1371/journal.pone.0011961.s003 (0.02 MB DOC)

**Table S1**
Found at: doi:10.1371/journal.pone.0011961.s004 (0.03 MB DOC)

**Table S2**
Found at: doi:10.1371/journal.pone.0011961.s005 (0.01 MB DOC)

**Table S3**
*Although individuals were selected based on the top and bottom 500 ranked samples, some extra criteria were used in the selection process as set out in the table. These criteria were applied to the raw measurements, whereas selection of the “top” and “tail” was carried out on adjusted traits as described in the methods section.

Found at: doi:10.1371/journal.pone.0011961.s006 (0.01 MB DOC)

**Table S4**
Found at: doi:10.1371/journal.pone.0011961.s007 (0.01 MB DOC)

**Table S5**
*In the combined analyses of stages 1 and 2, the “Total” numbers applied to the quantitative traits (HDL, WHR, DBP and quantitative metabolic syndrome) and the Cases and Controls numbers applied to T2D and binary metabolic syndrome.

Found at: doi:10.1371/journal.pone.0011961.s008 (0.02 MB DOC)

**Table S6**
Found at: doi:10.1371/journal.pone.0011961.s009 (0.01 MB DOC)

**Figure S1**
Found at: doi:10.1371/journal.pone.0011961.s010 (0.08 MB TIF)

**Acknowledgments**

The authors would like to thank JS Kooner and J. Chambers for providing data, and JS Kooner, J. Chambers, P Elliott and J Scott for contributing to the design of the study.

**Author Contributions**

Analyzed the data: DZ. Wrote the paper: DZ DB. Contributed to the design of the study: DZ DB.

**References**

1. Alberti KGMM, Zimmet P, Shaw J (2005) The metabolic syndrome – a new worldwide definition. Lancet 366 (9491): 1059–1062.
2. Zabaneh D, Chambers JC, Elliott P, Scott J, Balding DJ, et al. (2009) Heritability and genetic correlations of insulin resistance and component phenotypes in Asian Indian families using a multivariate analysis. Diabetologia 52 (12): 2582–2589.
3. Stephens M, Balding DJ (2009) Bayesian statistical methods for genetic association studies. Nature Reviews Genetics 10:681–690.
4. Freyling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, et al. (2007) A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science 316 (5826): 809–814.
5. Yu W, Clyne M, Wu L, Yesupriya A, Gwinn M, et al. (2009) Genopedia, HuGE Navigator. Ref Type: Catalog.
6. Chambers JC, Elliott P, Zabaneh D, Zhang WH, Li Y, et al. (2008) Common genetic variation near MC4R is associated with waist circumference and insulin resistance. Nature Genetics 40 (6): 716–718.
7. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Munro K, et al. (2009) Common variants at 30 loci contribute to polygenic dyslipidemia. Nature Genetics 41 (1): 56–65.
8. Sabatti C, Service SK, Harikainen AI, Pouta A, Ripatti S, et al. (2009) Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. Nature Genetics 41 (1): 35–46.
9. Salonen JT, Uimar L, Aalto JM, Pirskanen M, Klockenon J, et al. (2007) Type 2 diabetes whole-genome-association study in four populations: The DiaGen consortium. American Journal of Human Genetics 81 (2): 338–345.
10. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PIW, et al. (2007) Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 316 (5829): 1331–1336.
11. Scott L, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, et al. (2007) A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 316 (5829): 1314–1318.
12. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, et al. (2007) A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 445 (7130): 881–886.
13. Steinthorsdottir V, Thorleifsson G, Reynisdottir I, Benediktsson R, Jonasdottir T, et al. (2007) A variant in CDKAL1 influences insulin response and risk of type 2 diabetes. Nature Genetics 39 (6): 770–775.
14. Timpson NJ, Lindgren CM, Weedon MN, Randall J, Ouwerkerk WH, et al. (2009) Adiposity-Related Heterogeneity in Patterns of Type 2 Diabetes Susceptibility Observed in Genome-Wide Association Data. Diabetes 58 (2): 565–570.
15. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nature Genetics 40(5): 638–645.
16. Leibn F, Deckers M, Bertolino P, Ten Dijke P (2005) TGF-beta receptor function in the endothelium. Cardiovasc Res 65: 599–607.
17. Newton-Cheh C, Johnson T, Gueva V, Tobin MD, Bochud M, et al. (2009) Genome-wide association study identifies eight loci associated with blood pressure. Nature Genetics 41 (6): 666–676.
18. Benjamini, Sorrensen TIA, Schousboe K, Fenger M, Vischer PM, et al. (2007) Are there common genetic and environmental factors behind the endophenotypes associated with the metabolic syndrome? Diabetologia 50 (5): 881–888.
19. Spj gren M, Lyssenko V, Jonsson A, Berglund G, Nilsson P, et al. (2008) The search for putative unifying genetic factors for components of the metabolic syndrome. Diabetologia 51 (12): 2242–2251.
20. Executive Summary of the third report of the national cholesterol education program (NCEP) Expert Panel on Detection, evaluation and Treatment of high blood cholesterol in Adults (Adult Treatment Panel III) 2001. JAMA 285: 2486–2497.
21. WHO: Definition, diagnosis, and classification of diabetes mellitus and its complications. Report of a WHO Consultation, Available at http://www.staff.ncl.ac.uk/philip.home/who_dmc.htm MetabSyndDef. Accessed Feb 2009.
22. Joy T, Lahiry P, Pollex RL, Hegele RA (2008) Genetics of the metabolic syndrome. Current Diabetes Reports 8: 141–148.
23. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 81 (3): 559–575.
24. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LK, et al. (2008) Newly identified loci that influence lipid concentrations and risk of coronary artery disease. Nature Genetics 40(2): 161–169.
25. Hiura Y, Chen C, Kokubo Y, Okamura T, Morisaki T, et al. (2009) Identification of genetic markers associated with high-density lipoprotein-cholesterol by genome-wide screening in a Japanese population: the Suata study. Circ J 73(6): 1119–1126.
26. Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, et al. (2009) Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nature Genetics 41(1): 47–53.
Author/s:
Zabaneh, D; Balding, DJ

Title:
A Genome-Wide Association Study of the Metabolic Syndrome in Indian Asian Men

Date:
2010-08-04

Citation:
Zabaneh, D; Balding, DJ, A Genome-Wide Association Study of the Metabolic Syndrome in Indian Asian Men, PLOS ONE, 2010, 5 (8)

Persistent Link:
http://hdl.handle.net/11343/52567