



NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

Abdullah, N.; Abdul Murad, N.; Scott, R. J.; Holliday, E. G.; Attia, J.; Oldmeadow, C.; Mohd Haniff, E. A.; Syafruddin, S. E.; Abd Jalal, N.; Ismail, N.; Ishak, M.; Jamal, R. "Characterising the genetic risk for type 2 diabetes in a Malaysian multi-ethnic cohort". Published in Diabetic Medicine Vol. 32, Issue 10, p. 1377-1384 (2015)

Available from: <http://dx.doi.org/10.1111/dme.12735>

This is the peer reviewed version of the following article: Abdullah, N.; Abdul Murad, N.; Scott, R. J.; Holliday, E. G.; Attia, J.; Oldmeadow, C.; Mohd Haniff, E. A.; Syafruddin, S. E.; Abd Jalal, N.; Ismail, N.; Ishak, M.; Jamal, R. "Characterising the genetic risk for type 2 diabetes in a Malaysian multi-ethnic cohort", Diabetic Medicine Vol. 32, Issue 10, p. 1377-1384 (2015), which has been published in final form at <http://dx.doi.org/10.1111/dme.12735>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

Accessed from: <http://hdl.handle.net/1959.13/1309627>

Characterising the genetic risk for type 2 diabetes in a Malaysian multi-ethnic cohort

Running head: Type 2 diabetes genetics in Malaysia

N. Abdullah ^{1,2}, N. A. Abdul Murad ², J. Attia ^{3,4}, C. Oldmeadow ^{3,4}, E. A. Mohd Haniff ², S. E. Syafruddin ², N. Abd Jalal ², N. Ismail ², M. Ishak ², R. Jamal ², R.J. Scott ^{1,5}, E. G. Holliday ^{3,4}

¹ School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Newcastle, New South Wales, Australia

² UKM Medical Molecular Biology Institute (UMBI), Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

³ Clinical Research Design, IT and Statistical Support (CReDITSS) Unit, Hunter Medical Research Institute, Newcastle, New South Wales, Australia

⁴ Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, Faculty of Health, University of Newcastle, Newcastle, New South Wales, Australia

⁵ Hunter Area Pathology Service, John Hunter Hospital, Newcastle, New South Wales, Australia

Corresponding Authors:

Elizabeth G. Holliday

Mailing address:

CReDITSS, HMRI

C/- University Drive

Callaghan NSW 2308

Australia

E-mail: Liz.Holliday@newcastle.edu.au

Telephone: +61(2)40420508

Facsimile: +61(2)40420039

Rahman Jamal

Mailing address:

UKM Medical Molecular Biology Institute (UMBI)

Jalan Yaacob Latif, Bandar Tun Razak, Cheras, 56000 Kuala Lumpur, Malaysia

E-mail: rahmanj@ppukm.ukm.edu.my

Telephone: +60(3)91459000

Facsimile: +60(3)91717185

Word count: 2,903

Funding Sources

NA is supported by Ministry of Higher Education of Malaysia and Universiti Kebangsaan Malaysia. EGH is supported by a Fellowship (100071) from the Australian Heart Foundation and National Stroke Foundation. The Malaysian Cohort Study is funded by a top-down grant from the National Biotechnology Division, Ministry of Science, Technology and Innovation (MOSTI), Malaysia (ER-05-01-02-MEB001).

Conflict of Interest

We declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Novelty statement

- This is the first large-scale genetic study of type 2 diabetes in a Malaysian population
- We assessed 62 SNPs previously associated with type 2 diabetes for their association with disease in 2,392 type 2 diabetes cases and 2,594 controls of Malay, Chinese and Indian ancestry
- Seven individual SNPs were associated with type 2 diabetes after multiple testing adjustment
- We observed highly significant excess in concordance of allelic effect directions between Malaysian and previously studied populations ($P=1 \times 10^{-8}$)
- A genetic risk score including the 62 SNPs showed strong association in the Malaysian sample ($P=2 \times 10^{-16}$) and explained 1-1.7% of disease variance

Abstract

Aims: While genome-wide association studies (GWAS) have identified numerous type 2 diabetes risk variants across diverse populations, the Malaysian population remains unstudied to date. We characterised the association of known type 2 diabetes risk variants in Malaysian subjects of Malay, Chinese and Indian ancestry from The Malaysian Cohort project.

Methods: Using the MetaboChip array, 1,604 Malays (722 cases, 882 controls), 1,654 Chinese (819 cases, 835 controls) and 1,728 Indians (851 cases, 877 controls) were genotyped. First, 62 candidate SNPs previously associated with type 2 diabetes were assessed for association via logistic regression within ancestral groups, and then across ancestral groups via meta-analysis. Second, estimated odds ratios were assessed for excess directional concordance with previously studied populations. Third, a genetic risk score (GRS) aggregating allele dosage across the candidate SNPs was tested for association within and across ancestral groups.

Results: After Bonferroni correction, 7 individual SNPs were associated with type 2 diabetes in the combined Malaysian sample. We observed a highly significant excess in concordance of effect directions between Malaysian and previously studied populations. The GRS was strongly associated with type 2 diabetes in all Malaysian groups, explaining from 1.0 to 1.7% of total type 2 diabetes risk variance.

Conclusion: This study suggests substantial overlap of the genetic risk alleles underlying type 2 diabetes in Malaysian and other populations.

Introduction

Type 2 diabetes is continuing to grow in incidence and prevalence worldwide. In 2013, there were 382 million people worldwide living with diabetes and this number is projected to escalate by 55%, particularly in low and middle income countries (1). In Southeast Asia alone, the incidence of type 2 diabetes is projected to increase by 71% by 2035 (1).

According to data from International Diabetes Federation (IDF), Malaysia has the highest comparative prevalence of type 2 diabetes among Asian countries in 2013 (1). Malaysia is a multi-ethnic country whose population of 28.3 million (2) includes three major ancestral groups: Malays (~63%); Chinese (~25%); and Indian (~7%). The prevalence of type 2 diabetes between Malaysian populations appears to differ among the three major groups, with Asian Indians having the highest prevalence (25% to 28%), followed by Malays (17% to 19%) and the lowest apparent prevalence in Chinese (9% to 14%) (3).

Type 2 diabetes has a substantial genetic component and genome-wide association studies (GWAS) have identified more than 100 individual genetic variants associated with the condition. However, the majority of type 2 diabetes GWAS have been conducted in populations of European ancestry, which contribute only a fraction of total human genetic variation. Recent studies in broader populations show the importance of extending the population base of GWAS for type 2 diabetes. Benefits include the potential discovery of novel risk alleles due to population allele frequency differences or population specificity of risk alleles, improved fine-mapping due to population differences in linkage disequilibrium (4) and the ability to characterise transferability and consistency of risk alleles across populations (4). For a range of known risk alleles, initial analyses suggest population overlap of individually associated variants and consistency of allelic effects, both in direction and magnitude (4, 5). The largest trans-ethnic GWAS of type 2 diabetes included 26,488 cases and 83,964 controls from previously published type 2 diabetes GWAS samples of European,

East Asian, South Asian, Mexican and Mexican American ancestry (4). It detected seven new susceptibility loci and a significant excess in directional consistency of risk alleles across populations, indicating relevance of established risk loci across diverse ancestral groups.

To date, there are no published type 2 diabetes GWAS from the Malaysian population, in spite of the high comparative type 2 diabetes prevalence in this country. We conducted the first large-scale genetic study of type 2 diabetes using samples selected from The Malaysian Cohort project (3). Using genotype data generated using the Illumina MetaboChip array, we sought to characterise the association of known type 2 diabetes loci in Malaysian samples of Malay, Chinese and Indian ancestry. Our study had three principal aims: i) to assess the association of individual, previously reported type 2 diabetes risk variants with type 2 diabetes within and across Malaysian ancestry groups; ii) to assess evidence for excess concordance in the directional effect of type 2 diabetes risk alleles between previously studied and the Malaysian population, and; iii) to test genetic risk scores that combine information across multiple SNPs for association with type 2 diabetes in Malaysian groups.

Methods

Data sources and study samples

This was a nested case-control study. Type 2 diabetes cases and controls of Malay, Chinese and Indian ancestry were selected from the Malaysian Cohort (TMC), a prospective population-based cohort including 106,527 volunteers aged between 35 and 70 years. Subjects were recruited between April 2006 and September 2012 from regions across Malaysia (3). Comprehensive baseline measurements included fasting plasma glucose (FPG). For the current study, samples with FPG exceeding 7.5 mmol/L (or 126 mg/dL) were classified as type 2 diabetes, while controls had FPG <5.5 mmol/L (or 99 mg/dL).

A total of 1,604 Malays (722 cases, 882 controls), 1,654 Chinese (819 cases, 835 controls) and 1,728 Indians (851 cases, 877 controls) were selected for genotyping. For selection, ethnicity was defined using the self-reported ethnicity of the subject and their family for the three preceding generations. All relevant ethics approvals for The Malaysia Cohort were approved by the institutional review and ethics board of the Universiti Kebangsaan Malaysia, in accordance with the declaration of Helsinki. All subjects gave written, informed consent for participation in the study.

Genotyping and quality control

Samples were genotyped at the UKM Medical Molecular Biology Institute, Kuala Lumpur, Malaysia using the MetaboChip array (Illumina Inc, USA). This custom genotyping array includes 196,725 variants from loci previously implicated in cardiometabolic disease traits, and provides a high-throughput, cost-effective approach to genotyping SNPs previously associated with type 2 diabetes (6). Genotype calling was performed using Illumina GenomeStudio software with default quality score (GenCall) thresholds of ≥ 0.3 and ≥ 0.25 for

overall SNPs and individual genotypes, respectively. Manual quality control (QC) of genotype data was performed using PLINK (7). From the full set of genotyped SNPs, we first excluded SNPs with minor allele frequency (MAF) <0.01 , missingness >0.05 , or with significant deviation from Hardy-Weinberg equilibrium ($P < 10^{-5}$) in any of the three ancestral groups. In addition, cluster plots were visually checked for all selected candidate SNPs (see below) to ensure clear separation of genotype clusters (8). We then excluded samples with missingness >0.05 , outlying heterozygosity (± 8 SD from the mean), discrepant clinical and genotypic gender, accidental duplication, or cryptic relatedness (IBS sharing proportion >0.1875 ; midway between second and third-degree relatives). Genetic ancestry was assessed by principal components analysis (PCA) using reference data from the Singaporean Genome Variation Project (SGVP) and EIGENSTRAT software (9). The SGVP was used due to high similarity between the Singaporean and Malaysian populations. The SGVP includes reference genotype data for 89 Singaporean Malays, 96 Singaporean Chinese and 83 Singaporean Indians. Malaysian Cohort samples not clustering with their specified ancestral group (± 6 SD from the cluster mean on the first two principal components) were removed.

After performing SNP- and sample-level quality control, we performed logistic regression of case-control status against allelic dose across all remaining SNPs within each ancestral group. These analyses were performed with sequential adjustment for up to ten principal components (PCs) to calculate genomic inflation factors (λ_{GC}) and inform decisions about PC inclusion in candidate SNP association models, in order to minimise λ_{GC} .

Candidate SNP Selection

More than 40 GWAS studies of type 2 diabetes and its complications have been published and listed in the online Catalogue of published Genome-Wide Association Studies (10).

Using the Catalogue and a recent comprehensive review of type 2 diabetes genetic

associations (11), we identified SNPs previously showing genome-wide significant association ($P < 5 \times 10^{-8}$) with type 2 diabetes. Identified SNPs were selected for testing in our Malaysian sample if they were: a) present on the Metabochip array, and; b) passed quality control in at least two of the three Malaysian population groups. For type 2 diabetes-associated loci containing multiple associated SNPs, we selected a single lead SNP. The final set of candidate SNPs were in approximate linkage equilibrium, with all pairwise $r^2 < 0.5$ based on linkage disequilibrium in HapMap Chinese/Japanese combined reference data (CHB/JPT) (12).

Statistical Analyses

Association of each candidate SNP with type 2 diabetes was first assessed separately within ancestral groups using a logistic model assuming an additive allelic effect on the log-odds scale. Principal components of ancestry were included as covariates as indicated for each ancestral group. If the genomic inflation factor (λ_{GC}) exceeded 1, the standard errors of estimated SNP allele coefficients were further adjusted via genomic control, based on the observed λ_{GC} for the relevant ancestry group (13). Association summary statistics from each ancestral group were combined via inverse-variance weighted, fixed-effects meta-analysis using METAL (14). Heterogeneity of allelic effects was assessed using Cochran's Q Statistic. The experiment-wide significance level was derived via Bonferroni correction for the number of lead SNPs assessed via meta-analysis. Quantile-quantile (Q-Q) plots were also generated to visually assess enrichment for true associations.

To supplement individual association tests, the full set of candidate SNPs was assessed for the extent of concordance of allelic effect direction with previous studies, as previously described (4). This was performed both within and across the three ancestral groups. The observed proportion of directionally concordant SNPs was compared with that expected by

chance, with the null proportion being 0.5 for tests within individual ancestral groups, and 0.125 ($0.5 \times 0.5 \times 0.5$) for the meta-analysis analysis across groups, for which concordance across all groups was required. Observed and expected proportions were compared using a binomial test.

To assess the association evidence aggregated across candidate SNPs, a genetic risk score (GRS) was constructed. The GRS was formed as the weighted sum of reference alleles for each candidate SNP, with weights specified as the log odds ratio (beta coefficient) reported in the original publication. If multiple studies had reported genome-wide significant association of a SNP, we used the effect estimate from the largest study. Scoring was performed using PLINK. Association of the GRS with type 2 diabetes was assessed within each ancestral group via logistic regression. The proportion of case-control variance explained by the score was estimated using Nagelkerke's pseudo R^2 . Association evidence for the GRS was combined across ancestral groups via fixed-effects meta-analysis. Association testing was performed using Stata (15).

Secondary analysis

As a secondary analysis, we performed association analyses of all MetaboChip SNPs passing quality control. Logistic regression within ethnic groups and meta-analysis of results across the three ethnic groups were performed as described for candidate SNPs.

Results

After quality control, 4077 samples remained: 1,323 Malays samples (600 cases, 723 controls), 1344 Chinese samples (654 cases, 690 controls) and 1,410 Indians sample (708 cases, 702 controls). The Malay, Chinese and Indian groups were clearly separated on the first two ancestry principal components (**Supplementary Figure 1**) and each clustered closely with its respective Singaporean group. Based on PCA results and observed genomic inflation factors, the first 3 principal components were selected for inclusion as model covariates in both Malay and Indian groups, to minimise the λ_{GC} . Using logistic models including PCs as covariates, the observed inflation factors for Malay and Indian groups were 1.069 and 1.029, respectively. No principal components were necessary in Chinese, for which the unadjusted λ_{GC} was <1 (0.977).

Candidate SNP association tests

Of the identified 188 SNPs previously showing genome-wide association with type 2 diabetes, 97 had data available in at least two of the three ancestral groups. This set included several clusters of SNPs within a single locus. After selecting a single lead SNP for each locus, 62 SNPs remained. Based on Bonferroni correction for 62 SNPs, the pre-specified, adjusted significance threshold was $\alpha = 0.05/62 = 8.06 \times 10^{-4}$. Power to detect associated SNPs was calculated (16), assuming an additive model, perfect linkage disequilibrium (LD) between risk and marker alleles and an adjusted significance threshold of $\alpha=0.000806$. For a genetic risk ratio of 1.2, we had 38%, 72% and 85-89% power to identify risk alleles with frequency 0.1, 0.2 and 0.3-0.5 respectively. For a true risk ratio of 1.1, power was low, ranging from 4% to 19% across allele frequencies.

Association results for all SNPs, both within ancestral groups and in the meta-analysis across groups, are shown in **Supplementary Table 1**. Of the 62 SNPs, 7 reached $P < 8.06 \times 10^{-4}$

(**Table 1**) and 10 reached a nominal significance threshold of $P < 0.05$ (**Supplementary Table 1**) in the meta-analysis across groups. The SNPs reaching $P < 8.06 \times 10^{-4}$ were rs10965250 within *CDKN2A* ($P = 3 \times 10^{-5}$), rs4607517 within *GCK* ($P = 6 \times 10^{-5}$), rs7903146 within *TCF7L2* ($P = 2 \times 10^{-4}$), rs9939609 within *FTO* ($P = 2 \times 10^{-4}$), rs12970134 within *MC4R* ($P = 3 \times 10^{-4}$), rs11708067 within *ADCY5* ($P = 4 \times 10^{-4}$), and rs1801282 within *PPARG* ($P = 7 \times 10^{-4}$). Variants reaching nominal significance were rs1801214 in *WFS1* ($P = 5 \times 10^{-3}$), rs6931514 in *CDKAL1* ($P = 2 \times 10^{-3}$), rs3802177 in *SLC30A8* ($P = 7 \times 10^{-3}$), rs2796441 in *TLE1-FAM75D5* ($P = 0.03$), rs1111875 in *HHEX - EXOC6* ($P = 1 \times 10^{-3}$), rs6583826 in *IDE - RPL11P4* ($P = 0.02$), rs174550 in *FADS1* ($P = 1 \times 10^{-3}$), rs1552224 in *ARAP1* ($P = 0.01$), rs7177055 in *HMG20A-LINGO1* ($P = 0.02$) and rs8042680 in *PRC1; LOC100507118* ($P = 0.04$). Within individual groups, 7 of 62 SNPs reached a nominal significance threshold of $P < 0.05$ in Malays, 8 of 58 reached $P < 0.05$ in Chinese and 9 of 62 SNPs reached $P < 0.05$ in Indians (**Supplementary Table 1**). Quantile-quantile (Q-Q) plots representing the P -value distribution for association of the 62 SNPs showed considerable deviation from the distribution expected under the null hypothesis, both within groups and in the meta-analysis (**Supplementary Figure 2**). This suggests considerable enrichment for true associations, in spite of relatively few SNPs reaching the adjusted significance threshold.

Concordance in allelic effect directions between Malay and other populations

Within individual groups, we observed evidence of a significant excess in concordance of allelic effect directions with previously reported values. Of 62 SNPs with data available in Malays, 45 (72.6 %) showed effect directions consistent with previous studies, compared to the 50% expected by chance (binomial $P = 3.37 \times 10^{-5}$). Similarly, in Chinese, 45 of 58 SNPs (77.6%) showed concordant effects (binomial $P = 2.31 \times 10^{-7}$) and in Indians 47 of 62 (75.8%) were directionally consistent with previous findings (binomial $P = 1.04 \times 10^{-6}$). In results from the meta-analysis, 83.9% of SNPs showed a consistent summary effect direction

with that previously reported (52 of 62 SNPs; binomial $P=1.90 \times 10^{-13}$) (see also **Figure 1**). Of these 52 SNPs showing the previously reported effect direction, 33 (63.5%) showed an attenuated magnitude of effect compared to the original publication, significantly more than the 50% expected by chance (binomial $P=0.02$).

From the meta-analyses of results for individual SNPs, 56 SNPs had data available for all three ancestral groups. Of these, 28 (50.0%) showed consistent effect directions across all ancestral groups, significantly more than the proportion expected by chance (12.5%; binomial $P=9.97 \times 10^{-9}$). These results are consistent with the QQ-plots of SNP P -values, indicating enrichment for true associations among the selected candidate SNPs. They also show that effect directions for type 2 diabetes risk alleles in Malay groups are both relatively homogenous between groups and consistent with results in other ancestral populations.

Association of genetic risk scores

The genetic risk score (GRS) included data from 62 candidate SNPs in the Malay and Indian groups and 58 SNPs in Chinese. The GRS showed significant and consistent association with type 2 diabetes in all ancestral groups (Malay: $P = 4.91 \times 10^{-8}$; Chinese: $P = 1.35 \times 10^{-8}$ and Indian: $P = 4.71 \times 10^{-6}$), reaching a higher level of significance in the meta-analysis across groups ($P = 2.2 \times 10^{-16}$) (**Table 2**). The estimated proportion of type 2 diabetes risk variance explained by the GRS was 1.7% in Chinese, 1.6% in Malays, and 1.0% in Indians. There was no evidence of heterogeneity of the GRS effect across ancestral groups (Cochran's $I^2=0.0\%$; $P=0.39$) (**Figure 2**). The effect direction of the GRS was consistent with prior evidence both within and across ancestral groups, with risk scores reflecting a higher burden of previously reported risk alleles also associating with increased type 2 diabetes risks in Malaysian groups.

Secondary analysis

Supplementary Figure 3 shows the Manhattan plot of MetaboChip-wide meta-analysis results for 106,701 SNPs passing quality control in at least two of the three ancestral groups. The genomic inflation factor was 1.054. We calculated power to detect associated SNPs assuming an additive model (16), perfect linkage disequilibrium (LD) between risk and marker alleles with genome-wide significance threshold of $\alpha = 5 \times 10^{-8}$. For a true risk ratio of 1.2, power was low, ranging from 1% to 19% across allele frequencies. For a risk ratio of 1.1, power was 0 across all allele frequencies. As expected due to the modest sample size and small effect size of known T2D genetic risk variants, no SNPs reached $P < 5 \times 10^{-8}$ in the MetaboChip-wide analysis. The strongest associations were observed in the *FTO* gene on chromosome 16 ($P = 3.4 \times 10^{-6}$), and on chromosomes 7, 12 and 13 (including variants within the *OGDH* and *DDX56* genes). A total of 5 markers reached $P < 1 \times 10^{-5}$ (**Supplementary Table 2**).

Discussion

To our knowledge, this represents the first detailed genetic study of type 2 diabetes in Malaysia. We characterised the effect of previously reported, type 2 diabetes-associated variants at more than 60 loci in the three largest Malaysian ancestral groups: Malays, Chinese and Indians. Meta-analyses across groups identified SNPs in seven loci reaching significance after multiple-testing adjustment, at the *TCF7L2*, *CDKN2A*, *FTO*, *PPARG*, *GCK*, *MC4R* and *ADCY5* loci. In addition, 10 additional SNPs reached nominal significance in *WFS1*, *CDKAL1*, *SLC30A8*, *TLE1-FAM75D5*, *HHEX - EXOC6*, *IDE - RPL11P4*, *FADS1*, *ARAP1*, *HMG20A-LINGO1* and *PRC1; LOC100507118*. The majority of these genes are involved in biological pathways influencing diabetes pathophysiology, including pancreatic beta-cell development/function, insulin availability, glucose utilisation, fatty acid concentrations and obesity. While these variants were each initially identified in European ancestry populations, they have each also shown association in broader global populations, including groups of South-Asian and/or East Asian ancestry (17-23). This study confirms their additional involvement in type 2 diabetes in Malaysia.

In this Malaysian sample, we were unable to confirm individual association for many genetic variants previously associated with type 2 diabetes. A likely explanation was insufficient statistical power to identify variants with small individual effect. Power was reduced first by our modest sample size relative to earlier studies by large, international consortia. Second, for the majority of tested variants, estimated odds ratios in our Malaysian sample were small, generally ranging from 1.0 to 1.2. Indeed, we observed a significant excess of variants with smaller effect size in the Malaysian sample compared with the original study. A tendency for lower effect sizes has also been reported in studies of similar populations from Singapore (17, 24). This may be due to the phenomenon known as “winner’s curse”, or upward bias of effect estimates in the initial reporting study. Alternatively, smaller effect sizes may reflect lower

linkage disequilibrium between assessed and underlying causal variants in these South-East Asian populations. Regardless of the cause, attenuated odds ratios will diminish power to detect trait-variant association. Notably, for the seven variants showing significant association with type 2 diabetes, estimated odds ratios were relatively large, ranging from 1.2 to 1.4.

Notwithstanding limited power for testing individual variants, QQ-plots revealed an excess of nominally associated variants compared to chance expectation. Formal tests also showed a significantly elevated number of SNPs whose estimated effect direction was consistent with earlier studies. This suggests that many of the assessed SNPs may well influence type 2 diabetes risks in the Malaysian population and could demonstrate more significant association in larger samples.

The composite genetic risk score also demonstrated highly significant association with type 2 diabetes, both within individual groups and in the meta-analysis across groups, with all scores having an effect direction consistent with earlier studies. This further supports the relevance of many previously-reported type 2 diabetes risk variants in the Malaysian population.

Despite this apparent transferability of type 2 diabetes risk alleles into Malaysia, the genetic risk score explained less than 2% of overall type 2 diabetes risk in any individual group. We do acknowledge that our study assessed association for SNPs representing 97 of an identified 188 variants previously showing genome-wide association with type 2 diabetes. The effect of including SNPs representing the additional 91 variants is unknown, but would likely produce higher estimates of explained variance. Nevertheless, if the additional 91 variants explain a similar, additional proportion of risk, the rapidly escalating type 2 diabetes prevalence in Malaysia (1) seems unlikely to result solely from common genetic variants. Recent environmental changes in dietary patterns and physical activity may contribute more substantially. In addition, lifestyle-related factors such BMI, waist-hip circumference or

dietary intake of fats/sugar may interact with genetic risk alleles to further elevate type 2 diabetes risk in the Malaysian population. Future studies of low frequency variants or epigenetic modifications may also reveal genetic factors influencing the rising prevalence of type 2 diabetes in south-east Asian populations.

Funding Sources

NA is supported by Ministry of Higher Education of Malaysia and Universiti Kebangsaan Malaysia. EGH is supported by a Fellowship (100071) from the Australian Heart Foundation and National Stroke Foundation. The Malaysian Cohort Study is funded by a top-down grant from the National Biotechnology Division, Ministry of Science, Technology and Innovation (MOSTI), Malaysia (ER-05-01-02-MEB001).

Conflict of Interest

We declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Acknowledgments

We thank all UKM Medical Molecular Biology Institute (UMBI) and The Malaysian Cohort staff members and research assistants. The voluntary participation of all the subjects is greatly appreciated.

References

1. International Diabetes Federation. IDF Diabetes Atlas 6th Edition 2013; 6:[169 p.].
2. Population Distribution and Basic Demographic Characteristics 2010 [Internet]. Department of Statistics Malaysia,. 2011 [cited 12 May 2014].
3. Jamal R. SZ, S.Z., Kamaruddin M.A., Jalal A.N, Ismail N., Kamil N.M, Abdullah N., Baharudin N., Hussin N.H, Othman H., Mahadi N. M., The Malaysian Cohort Group,. Cohort profile: The Malaysian Cohort (TMC) project: a prospective study of non-communicable diseases in a multi-ethnic population. *The International Journal of Epidemiology*. 2014;9.
4. Replication DIG, Meta-analysis C. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. *Nature genetics*. 2014;46(3):234-44.
5. Saxena R, Elbers CC, Guo Y, Peter I, Gaunt TR, Mega JL, et al. Large-scale gene-centric meta-analysis across 39 studies identifies type 2 diabetes loci. *American journal of human genetics*. 2012;90(3):410-25.
6. Voight BF, Kang HM, Ding J, Palmer CD, Sidore C, Chines PS, et al. Correction: The Metabochip, a Custom Genotyping Array for Genetic Studies of Metabolic, Cardiovascular, and Anthropometric Traits. *PLoS genetics*. 2013;9(4).
7. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*. 2007;81(3):559-75.
8. Tindall EA, Petersen DC, Nikolaysen S, Miller W, Schuster SC, Hayes VM. Interpretation of custom designed Illumina genotype cluster plots for targeted association studies and next-generation sequence validation. *BMC research notes*. 2010;3:39.
9. Teo YY, Sim X, Ong RT, Tan AK, Chen J, Tantoso E, et al. Singapore Genome Variation Project: a haplotype map of three Southeast Asian populations. *Genome research*. 2009;19(11):2154-62.
10. A Catalog of Published Genome-Wide Association Studies [Internet]. National Human Genome Research Institute. 2014 [cited April 2014]. Available from: www.genome.gov/gwastudies.
11. Abdullah N. AJ, Oldmeadow C., Scott R.J. & Holliday E.G. The architecture of risk for type 2 diabetes: understanding Asia in the context of global findings. *International Journal of Endocrinology*. 2014;2014:21.
12. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008;24(24):2938-9.
13. Devlin B, Roeder K. Genomic control for association studies. *Biometrics*. 1999;55(4):997-1004.
14. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*. 2010;26(17):2190-1.
15. Sterne JAC BM, Egger M. Meta-analysis in Stata. 2001. In: *Systematic Reviews in Health Care* [Internet]. London: BMJ Publishing Group; [347-69].
16. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nature genetics*. 2006;38(2):209-13.
17. Chen Z, Pereira MA, Seielstad M, Koh WP, Tai ES, Teo YY, et al. Joint effects of known type 2 diabetes susceptibility loci in genome-wide association study of Singapore Chinese: the Singapore Chinese Health Study. *PloS one*. 2014;9(2):e87762.
18. Rasmussen-Torvik LJ, Guo X, Bowden DW, Bertoni AG, Sale MM, Yao J, et al. Fasting glucose GWAS candidate region analysis across ethnic groups in the Multiethnic Study of Atherosclerosis (MESA). *Genetic epidemiology*. 2012;36(4):384-91.
19. Hara K, Fujita H, Johnson TA, Yamauchi T, Yasuda K, Horikoshi M, et al. Genome-wide association study identifies three novel loci for type 2 diabetes. *Human molecular genetics*. 2014;23(1):239-46.

20. Li H, Kilpelainen TO, Liu C, Zhu J, Liu Y, Hu C, et al. Association of genetic variation in FTO with risk of obesity and type 2 diabetes with data from 96,551 East and South Asians. *Diabetologia*. 2012;55(4):981-95.
21. Xi B, Takeuchi F, Chandak GR, Kato N, Pan HW, Consortium A-TD, et al. Common polymorphism near the MC4R gene is associated with type 2 diabetes: data from a meta-analysis of 123,373 individuals. *Diabetologia*. 2012;55(10):2660-6.
22. Chan KH, Niu T, Ma Y, You NC, Song Y, Sobel EM, et al. Common genetic variants in peroxisome proliferator-activated receptor-gamma (PPARG) and type 2 diabetes risk among Women's Health Initiative postmenopausal women. *The Journal of clinical endocrinology and metabolism*. 2013;98(3):E600-4.
23. Rees SD, Hydrie MZ, O'Hare JP, Kumar S, Shera AS, Basit A, et al. Effects of 16 genetic variants on fasting glucose and type 2 diabetes in South Asians: ADCY5 and GLIS3 variants may predispose to type 2 diabetes. *PLoS one*. 2011;6(9):e24710.
24. Sim X, Ong RT, Suo C, Tay WT, Liu J, Ng DP, et al. Transferability of type 2 diabetes implicated loci in multi-ethnic cohorts from Southeast Asia. *PLoS genetics*. 2011;7(4):e1001363.
25. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nature genetics*. 2010;42(7):579-89.
26. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, Jackson AU, et al. New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nature genetics*. 2010;42(2):105-16.
27. Perry JR, Voight BF, Yengo L, Amin N, Dupuis J, Ganser M, et al. Stratifying type 2 diabetes cases by BMI identifies genetic risk variants in LAMA1 and enrichment for risk variants in lean compared to obese cases. *PLoS genetics*. 2012;8(5):e1002741.
28. Morris AP, Voight BF, Teslovich TM, Ferreira T, Segre AV, Steinthorsdottir V, et al. Large-scale association analysis provides insights into the genetic architecture and pathophysiology of type 2 diabetes. *Nature genetics*. 2012;44(9):981-90.

Table 1: Association results for candidate SNPs showing significant association with type 2 diabetes in the meta-analysis across Malaysian groups

Lead SNP ¹	Mapped Gene(s)	Region	RA ²	OA ³	Malays		Chinese		Indians		Meta-analysis		
					OR	P	OR	P	OR	P	OR	P ⁴	EA ⁵
					(95% CI)		(95% CI)		(95% CI)		(95% CI)		
rs10965250(25)	<i>CDKN2B</i>	9p21.3	G	A	0.82	0.02 ^a	0.79	3x10 ^{-3a}	0.81	0.04 ^a	0.81	3x10 ⁻⁵	A ^b
					(0.66-0.98)		(0.68-0.92)		(0.61-1.01)		(0.73-0.89)		
rs4607517(26)	<i>GCK - YKT6</i>	7p13	A	G	1.25	0.06	1.58	1x10 ^{-5a}	1.01	0.91	1.30	6x10 ⁻⁵	A ^b
					(1.01-1.49)		(1.29-1.94)		(0.78-1.25)		(1.15-1.48)		
rs7903146(25)	<i>TCF7L2</i>	10q25.2	T	C	1.21	0.19	1.08	0.72	1.37	2x10 ^{-4a}	1.30	2x10 ⁻⁴	A ^b
					(0.93-1.49)		(0.70-1.68)		(1.21-1.53)		(1.13-1.49)		
rs9939609(27)	<i>FTO</i>	16q12.2	A	T	1.32	1x10 ^{-3a}	1.16	0.18	1.16	0.06	1.22	2x10 ⁻⁴	A ^b
					(1.15-1.49)		(0.93-1.45)		(1.01-1.32)		(1.10-1.35)		
rs12970134(28)	<i>MC4R</i>	18q21.32	A	G	1.19	0.12	1.17	0.13	1.28	3x10 ^{-3a}	1.22	3x10 ⁻⁴	A ^b
					(0.97-1.41)		(0.96-1.44)		(1.12-1.44)		(1.10-1.36)		
rs11708067(26)	<i>ADCY5</i>	3q21.1	A	G	0.69	0.05	N/A ^c	N/A ^c	0.74	2x10 ^{-3a}	1.38	4x10 ⁻⁴	A ^b
					(0.32-1.06)				(0.54-0.93)		(1.15-1.64)		
rs1801282(4)	<i>PPARG</i>	3p25.2	A	G	0.77	0.2	0.93	0.78	0.67	2x10 ^{-3a}	1.28	7x10 ⁻⁴	A ^b
					(0.36-1.18)		(0.57-1.52)		(0.41-0.92)		(1.03-1.58)		

¹ SNP previously associated with type 2 diabetes or fasting plasma glucose at genome-wide significance ($P < 5 \times 10^{-8}$), with original reference. ² Risk allele from previous study. ³ Other allele from previous study. ⁴ $P < 8.1 \times 10^{-4}$, incorporating adjustment for testing 62 independent SNPs. ⁵ Effect allele from meta-analysis.

^a Denotes SNPs reaching P -value < 0.05 in individual ancestral groups

^b Denotes same effect direction as previously reported

^c Data not available due to MAF < 0.01

Table 2: Association between the genetic risk score and type 2 diabetes within ancestral groups and in the meta-analysis across groups

Study	N _{SNPs}	<i>P</i> -value	Effect direction	Pseudo R ²
Malays	62	4.91×10^{-8}	+	1.6%
Chinese	58 ^a	1.35×10^{-8}	+	1.7%
Indians	62	4.71×10^{-6}	+	1.0%
Meta-analysis	62	2.2×10^{-16}	+	

^a Number reduced owing to 4 SNPs with MAF < 0.01 in Chinese

Figure Legends

Figure 1: Bivariate plots comparing odds ratios observed in each Malaysian ancestral group and the meta-analysis across groups, with those previously published. (A) Malays, (B) Chinese, (C) Indians, (D) Combined meta-analysis.

Figure 2: Forest plot showing association of the genetic risk score in the meta-analysis across Malaysian groups.