

ORIGINAL ARTICLE

Genetic-risk assessment of GWAS-derived susceptibility loci for type 2 diabetes in a 10 year follow-up of a population-based cohort study

Min Jin Go¹, Young Lee¹, Suyeon Park^{1,2}, Soo Heon Kwak³, Bong-Jo Kim¹ and Juyoung Lee¹

To date, genome-wide meta-analyses have identified genetic susceptibility to type 2 diabetes mellitus (T2D) predominantly in populations of European ancestry. However, comprehensive genetic-risk assessment based on previous GWAS loci has not been fully tested in non-European populations. To evaluate whether a genetic-risk score (GRS) could improve T2D-risk prediction in the Korean population, a GRS (GRS-55) was constructed by summing 55 risk alleles based on the 1000 Genomes imputation in the Korean Association Resource study (T2D cases = 1042 and controls = 2943 at baseline). We also constructed another GRS (GRS-19) based on nominal significance and consistent direction of effect. In mean difference tests, the mean value of the GRS-19 was significantly higher in T2D cases than in controls at baseline examination. In a model adjusted for area, age, sex and body mass index, weighted GRS-19 was found to be associated with enhanced effect sizes of T2D risk under consistent C-statistics. In addition, we confirmed cumulative risk effects on incidence rates of T2D, fasting plasma glucose and glycated hemoglobin (HbA1c) levels in a longitudinal 10 year of follow-up study. These findings highlight that a genotype score comprised of 19 common variants contributed to T2D-risk prediction in the Korean population. Further multi-locus epistatic interactions may provide the possibility to improve risk prediction in C-statistics for discrimination or reclassification.

Journal of Human Genetics (2016) 61, 1009–1012; doi:10.1038/jhg.2016.93; published online 21 July 2016

INTRODUCTION

Type 2 diabetes mellitus (T2D) is a chronic multi-factorial disease underlying polygenic predisposition and environmental risk factors. The first wave of genome-wide association studies (GWAS) has led to the discovery of T2D susceptibility loci.¹ Most T2D loci so far been identified were derived from populations of European ancestry with a few East Asian groups. Recent trans-ethnic fine-mapping identifies additional population-specific signals that increase the trait variance explained.² However, identification of the genetic components of T2D has been limited by genetic heterogeneity across populations of ethnically diverse ancestries.

As an integrative analysis approach to T2D-risk prediction, genetic-risk scores (GRS) can be an efficient and effective method of constructing genome-wide risk measurements from GWAS findings.^{3–5} Recently, genetic-risk assessment studies for T2D have been reported by evaluating the predictive value of cumulative genetic scores.^{6,7} However, predictive values of GWAS-derived genetic scores have been affected by ethnic-specific determinants of genotype frequencies, phenotypic effect sizes and disease incidence.^{8–10} Further development and improvement of a multi-single-nucleotide polymorphism (SNP) GRS led to our better understanding for complex

disease prediction and prevention in an independent ethnic population. In this study, we evaluated genetic contributions of a GRS for T2D susceptibility in the Korean population.

MATERIALS AND METHODS

Study participants

The participants for this study were selected from the Korean Association Resource study. The Korean Association Resource study was collected through the Korean Genome Epidemiologic Study project. Total 10 038 people in the 40–69 years old living in Ansong and Ansan was started in 2001–2002. For more information about study were already reported.^{11,12} On the basis of the criteria of WHO (World Health Organization) diagnosis guideline, 1042 subjects were included as T2D cases according to the following criteria: (1) past-medical and family history of T2D; (2) fasting plasma glucose ≥ 7 mmol l⁻¹ or plasma glucose 2-h after ingestion of 75 gm oral glucose ≥ 11.1 mmol l⁻¹; and (3) age of disease onset ≥ 40 years. The inclusion criteria of normal controls ($n = 2943$) were as follows: (1) no past-medical and family history of T2D; and (2) fasting plasma glucose < 5.6 mmol l⁻¹ and plasma glucose 2-h after ingestion of 75 gm oral glucose < 7.8 mmol l⁻¹. Procedures were in accordance with institutional guidelines and approved by an institutional review committee. Written informed consent was obtained from all study participants. The clinical characteristics of the subjects are shown in Table 1.

¹Division of Structural and Functional Genomics, The Center for Genome Science, Korea National Institute of Health, KCDC, Osong, Republic of Korea; ²Department of Biostatistics, Soonchunhyang University, College of Medicine, Seoul, Republic of Korea and ³Department of Internal Medicine, Seoul National University Hospital, Seoul, Republic of Korea

Correspondence: Dr J Lee, Division of Structural and Functional Genomics, Center for Genome Science, Korea National Institute of Health, Chuncheonbuk-do, Osong 363-951, Republic of Korea.

E-mail: jylee@cdc.go.kr

Received 16 March 2016; revised 2 June 2016; accepted 9 June 2016; published online 21 July 2016

SNP imputation and GRS construction

Imputation of genotypes to the 1000 Genomes phase I integrated variant call set release (version 3) in NCBI build 37 (hg19) as a reference panel was carried out using the IMPUTE (high imputation quality: proper info >0.5). Of these, we dropped SNPs with a posterior probability score <0.90, low genotype information content (info <0.5), Hardy–Weinberg equilibrium (HWE) ($P < 1 \times 10^{-7}$), minor allele frequency (MAF) <0.01, and SNP missing rate >0.1. To construct GRSs using previous T2D susceptibility loci, we considered 72 GWAS loci from the GWAS catalog (<https://www.ebi.ac.uk/gwas/>). In addition to the SNP (rs2233580 in *PAX4*) specific to East Asians, we selected 54 SNPs having the same direction of effect as the original reports. A total of 55 SNPs were available in this study (Supplementary Table 1). Of the GRS-55 remaining loci, we further constructed the GRS-19 using 19 SNPs showing nominal significance and consistent direction of effect in the Korean population. On the basis of the Bonferroni threshold (significance, $0.05/69 = 0.0007$), four variants (rs4402960, rs7756992, rs10811661 and rs1111875) remain significant after correction for multiple testing (data not shown). To maximize the statistical power of a genetic prediction model, all SNPs with a nominal *P*-value <0.05 were included in the following risk score analysis. The weighted GRS was calculated by multiplying the number of risk alleles for each individual by each β -coefficient obtained from the association statistics. In addition, the cumulative number of risk alleles was scored using an additive model (0 for homozygous of non-risk alleles, one for heterozygous of alleles and two for homozygous of the risk alleles for the effect allele).

Table 1 Descriptive characteristics of T2D case and control subjects at baseline

Variable	T2D	Control	P-value
<i>n</i>	1042	2943	
Sex (M/F)	539/503	1355/1588	0.0016
Age (year)	56.37 ± 8.58	51.06 ± 8.57	<0.0001
Body mass index (kg m ⁻²)	25.48 ± 3.29	24.13 ± 2.95	<0.0001
Waist circumference (cm)	86.48 ± 8.22	81.45 ± 8.59	<0.0001
Systolic blood pressure (mm Hg)	122.28 ± 17.85	113.4 ± 15.95	<0.0001
Diastolic blood pressure (mm Hg)	76.89 ± 11.11	73.24 ± 10.92	<0.0001
HDL cholesterol (mg dl ⁻²)	42.6 ± 9.71	45.25 ± 9.8	<0.0001
Total cholesterol (mg dl ⁻²)	200.19 ± 42.32	185.62 ± 33.42	<0.0001
Triglycerides (mg dl ⁻²)	214.46 ± 149.57	140.77 ± 78.47	<0.0001
Fasting plasma glucose (mmol l ⁻²)	6.91 ± 2.56	4.47 ± 0.37	<0.0001

Abbreviations: HDL, high-density lipoprotein; T2D, type 2 diabetes mellitus. Data are means ± s.d.

Table 2 Comparison of the mean number of risk alleles in the GRS-55 and the GRS-19

	GRS-19	P-value	GRS-55	P-value
T2D case	17.16 ± 2.72	2.16×10^{-26}	57.38 ± 4.47	1.00×10^{-28}
Control	16.11 ± 2.74		55.64 ± 4.32	
	WGRS-19	P-value	WGRS-55	P-value
T2D case	3.12 ± 0.50	1.75×10^{-28}	11.70 ± 0.72	5.32×10^{-27}
Control	2.92 ± 0.50		11.42 ± 0.72	

Abbreviations: GRS, genetic-risk score; SNP, single-nucleotide polymorphism; T2D, type 2 diabetes mellitus; WGRS, weighted GRS. The difference between cases and controls for each GRS are represented by *P*-values. Data are means ± s.d. The number of risk alleles for each SNP was multiplied by each beta-coefficient obtained from the association analysis for T2D.

Statistical analysis

In performing association tests, SNPs were analyzed with R (2.15.1) software package, PLINK (<http://pngu.mgh.harvard.edu/~purcell/plink/>) and SAS programs (version 9.1; SAS institute Inc, Cary, NC, USA). T2D associations were tested by logistic regression analysis in an additive genetic model (1-d.f.) with or without adjustment for age, sex and body mass index as covariates. Quantitative traits for fasting plasma glucose and glycated hemoglobin (HbA1c) levels were tested by multiple linear regression analysis.

RESULTS

The baseline participant characteristics in the Korean Association Resource study (*n* = 3985) are listed in Table 1. According to T2D diagnostic criteria, we selected T2D cases (*n* = 1042) and normal controls (*n* = 2943), respectively. Demographic and clinical variables in T2D individuals were predominantly higher than the controls. Of non-diabetic subjects at baseline examination, 651 participants were found to be incident T2D cases (cumulative incidence proportion: 12.35%) in a prospective 10-year follow-up study. To evaluate comprehensive risk assessment for previous GWAS loci, we constructed a GRS-55 by summing 55 risk alleles available from the 1000 Genomes-based imputation data. In addition, a GRS-19 was constructed using 19 association signals based on nominal significance and consistent direction of effect in the Korean population.

In mean difference tests between the GRS-19 and the GRS-55, mean value of weighted GRS-19 (wGRS-19) was significantly higher than that of the wGRS-55 (wGRS-19: $P = 1.75 \times 10^{-28}$ and wGRS-55: $P = 5.32 \times 10^{-27}$) (Table 2). The distribution of the number of risk alleles in T2D cases shifted to the right (Figure 1a). Although there was

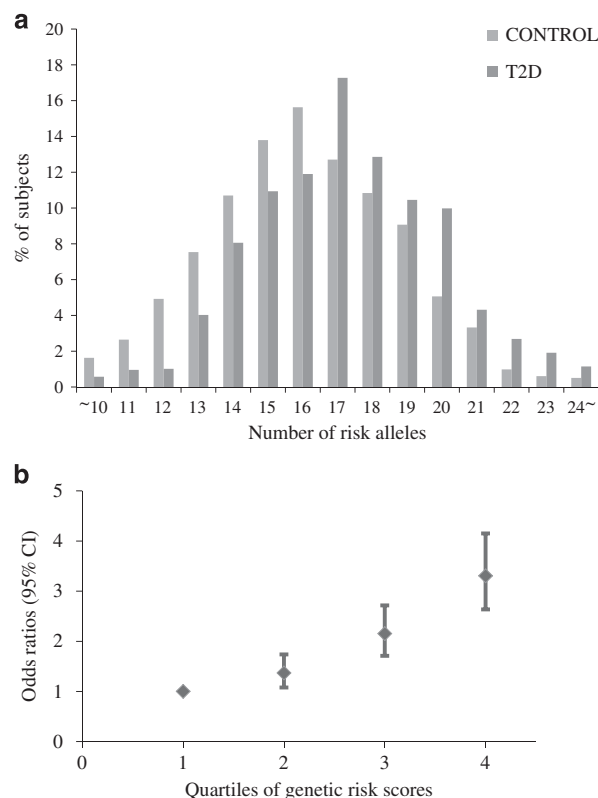


Figure 1 Distribution of genotype score and cumulative effect of the wGRS-19. (a) Distribution of the number of risk alleles. Red and blue bars indicate T2D cases and controls, respectively. (b) Quartile-based odd ratios for T2D risk alleles. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

Table 3 Associations of the GRS-55 and the GRS-19 with T2D risk

	OR per risk allele		ROC-AUC
	(95% CI)	P-value	(GRS+area, age, sex and BMI)
GRS-55	1.11 (1.09–1.13)	3.41×10^{-27}	0.76 (0.75–0.78)
GRS-19	1.18 (1.15–1.21)	1.37×10^{-28}	0.76 (0.75–0.78)
WGRS-55	2.02 (1.78–2.28)	2.00×10^{-28}	0.77 (0.75–0.78)
WGRS-19	2.60 (2.22–3.04)	6.11×10^{-32}	0.77 (0.75–0.78)

Abbreviations: AUC, area under curve; BMI, body mass index; CI, confidence interval; GRS, genetic-risk score; OR, odds ratio; ROC, receiver operating characteristic. Association analyses were adjusted for geographic area, age, sex and BMI. The number of risk alleles for each SNP was multiplied by each β -coefficient obtained from the association analysis for T2D.

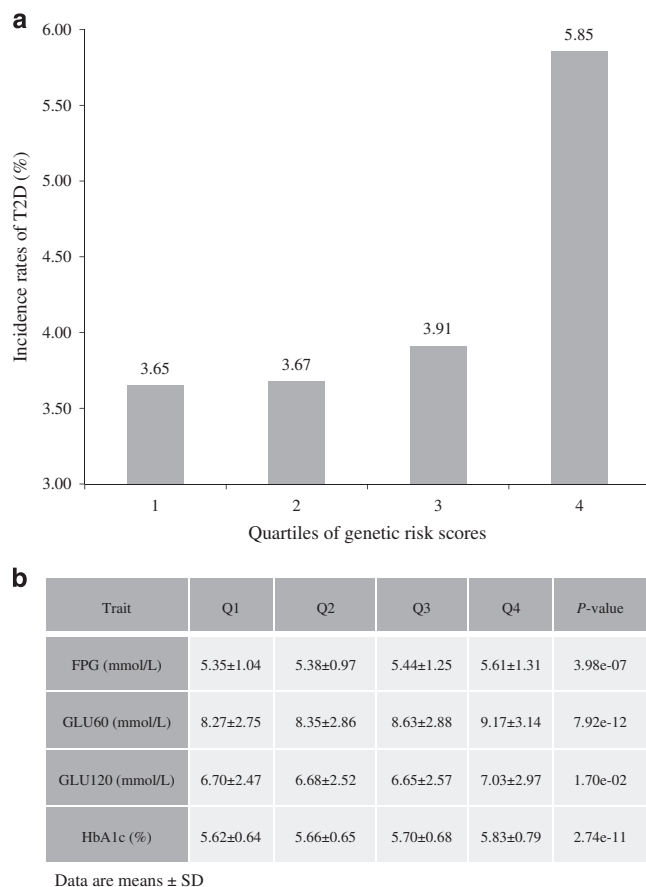


Figure 2 Cumulative incidence rates of T2D over 10 years of follow-up. (a) The 10 year cumulative incidence rates of T2D according to the quartile-based genotype-risk score. (b) The 10 year cumulative incidence rates of FPG and glycated hemoglobin (HbA1c) levels according to the quartile-based genotype-risk score. FPG, fasting plasma glucose. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

no difference in C-statistics between stratified groups, the wGRS-19 was significantly associated with increased T2D risk (wGRS-19: $P=6.11 \times 10^{-32}$, odds ratio per risk allele=2.60, 95% confidence interval=2.22–3.04; wGRS-55; $P=2.00 \times 10^{-28}$, odds ratio per risk allele=2.02, 95% confidence interval=1.78–2.28) after controlling for area, age, sex and body mass index (Table 3). In a weighted quartile analysis, an increased number of risk allele was associated with an increased odds ratio for T2D risk (Figure 1b).

To assess improvement in risk prediction of the wGRS-19, we further evaluated cumulative incidence rates of T2D, fasting plasma

glucose and glycated hemoglobin (HbA1c) levels in a longitudinal 10 year follow-up study. Incidence of T2D in wGRS-19 is associated with increase in the highest quartile (Figure 2a). The mean values for fasting plasma glucose, GLU60, GLU120 and HbA1c traits were calculated by quartile-based genotype-risk scores only in non-diabetic control subjects. These findings demonstrate that the wGRS-19 is associated with consistently additive risk effects in glycemic status (Figure 2b). However, compared with those of fasting glucose levels ($n=5907$), the number of valid observations (non-missing) is relatively small for calculating the P -values in 2 h glucose levels ($n=5107$), indicating a marginal improvement.

DISCUSSION

Conventional GWASs focused on the effect of individual SNPs have been typically limited by disease causality and complexity underlying etiological heterogeneity.^{13–15} In recent years, polygenic-risk scores have generated much interest for assessing the explanatory power of risk variants in the clinical management and prevention of T2D.¹⁶ To evaluate population differences in cumulative risk allele load for T2D, we conducted a comprehensive genetic-risk assessment for previously reported T2D loci in the Korean population.

As a novel functional locus carrying a missense mutation, imputed SNP rs2233580 in the *PAX4* gene is specific to the pathogenesis of T2D in East Asian populations.^{17–19} In SNP prediction analyses using functional assessment tools,²⁰ rs2233580 was significantly associated with potentially damaging or deleterious functional impacts (FS score=0.98). Regulatory motif analyses demonstrated *PAX4* enrichment for a SNP rs163184 in the *KCNQ1* gene.²¹ Furthermore, we investigated whether risk alleles of the GRS-19 were population-specific in ancestral super-populations (African (AFR), American (AMR), East Asian (ASN) and European (EUR)) of the 1000 Genomes Project data.²¹ In contrast to completely monomorphic in AFR, AMR and EUR, the SNP (rs2233580) has ASN-specific low-risk allele frequency (RAF=0.09). In addition, the majority of GRS-19 variants were consistently found to be associated with differential risk allele distributions: relatively higher (9 loci, RAF ranged from 0.42 to 0.88) and relatively lower (four loci, RAF=0.28–0.64) than other populations (AFR, AMR and EUR) (Supplementary Table 2). These findings indicate that different fractions in cumulative risk allele loads of a GRS are associated with heterogeneities of allele frequency, effect size and population structure in genetic architecture underlying T2D risk.^{8,22}

Recently, cross-sectional case-control studies using a GWAS-derived GRS have identified population genetic signatures to T2D susceptibility in Japanese,²³ Chinese²⁴ and Singapore²⁵ populations. However, there still remain specific fractions in cumulative risk allele load of a GRS between ethnic groups in East Asia (Supplementary Figure 1 and Supplementary Table 3). Only an SNP (rs10811661) in *CDKN2A/B* was found to be associated with T2D risk as a shared variant. Notably, a SNP (rs7903146) in *TCF7L2* known as a European-derived T2D susceptibility variant was not incorporated into a risk assessment model in this study. A genetic-risk evaluation in African Americans was similar to observations in European-derived populations, which is primarily driven by the rs7903146 variant of *TCF7L2*.²⁶ Given substantial differences in MAF and LD (indirect mapping) between different ethnic groups, constructing a GRS based on population-specific signals might be important in improving the accuracy of risk prediction models.

This study has some limitations. Despite the modest genetic effects of GWAS loci so far identified, our study focused on common risk factors. It still has limited power to detect an effect of low-frequency alleles or rare variants. In addition, we did not account for epigenetic

modifications and wide spread environmental factors in terms of missing heritability.^{27–30} A recent study reported that a GRS comprising 14 central obesity loci is associated with increased T2D risk.³¹ Obesity is both a confounder and an intermediate factor for T2D development.³² Their seeming inseparability has led to multiple confounding effects on phenotypic variance. Nonetheless, a major strength of our study is that we examined a comprehensive genetic-risk assessment associated with changes in T2D risk and T2D-related traits over 10-years of follow-up.

In conclusion, we suggest that the GRS-19 was significantly associated with an increased risk of fasting plasma glucose, glycated hemoglobin and incident T2D in a prospective 10-year epidemiological cohort study in Korea. Further integrated analytical approaches on epistatic effects might expand new possibilities to improve genetic-risk prediction in evidence-based healthcare and public health.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by grants from Korea Centers for Disease Control and Prevention (4845-301, 4851-302 and 4851-307) and an intramural grant from the Korea National Institute of Health (2013-NG73002-00).

Author contributions: MJG, JL and B-JK conceived and designed the study. MJG, YL, SP and SHK contributed to analysis and interpretation of data. MJG drafted the article and revised. All authors approved the final versions to be published.

- McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P. *et al.* Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* **9**, 356–369 (2008).
- Wu, Y., Waite, L. L., Jackson, A. U., Sheu, W. H., Buyske, S., Absher, D. *et al.* Trans-ethnic fine-mapping of lipid loci identifies population-specific signals and allelic heterogeneity that increases the trait variance explained. *PLoS Genet.* **9**, e1003379 (2013).
- Vassy, J. L. & Meigs, J. B. Is genetic testing useful to predict type 2 diabetes? *Best. Pract. Res. Clin. Endocrinol. Metab.* **26**, 189–201 (2012).
- Talmud, P. J., Cooper, J. A., Morris, R. W., Dudbridge, F., Shah, T., Engmann, J. *et al.* Sixty-five common genetic variants and prediction of type 2 diabetes. *Diabetes* **64**, 1830–1840 (2015).
- Shigemizu, D., Abe, T., Morizono, T., Johnson, T. A., Borevich, K. A., Hirakawa, Y. *et al.* The construction of risk prediction models using GWAS data and its application to a type 2 diabetes prospective cohort. *PLoS ONE* **9**, e92549 (2014).
- Vassy, J. L., Hivert, M. F., Porneala, B., Dauriz, M., Florez, J. C., Dupuis, J. *et al.* Polygenic type 2 diabetes prediction at the limit of common variant detection. *Diabetes* **63**, 2172–2182 (2014).
- Meigs, J. B., Shrader, P., Sullivan, L. M., McAteer, J. B., Fox, C. S., Dupuis, J. *et al.* Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N. Engl. J. Med.* **359**, 2208–2219 (2008).
- Keaton, J. M., Cooke Bailey, J. N., Palmer, N. D., Freedman, B. I., Langefeld, C. D., Ng, M. C. *et al.* A comparison of type 2 diabetes risk allele load between African Americans and European Americans. *Hum. Genet.* **133**, 1487–1495 (2014).
- Manolio, T. A. Bringing genome-wide association findings into clinical use. *Nat. Rev. Genet.* **14**, 549–558 (2013).
- Klimentidis, Y. C., Wineinger, N. E., Vazquez, A. I. & de Los Campos, G. Multiple metabolic genetic risk scores and type 2 diabetes risk in three racial/ethnic groups. *J. Clin. Endocrinol. Metab.* **99**, E1814–E1818 (2014).
- Kim, Y. J., Go, M. J., Hu, C., Hong, C. B., Kim, Y. K., Lee, J. Y. *et al.* Large-scale genome-wide association studies in East Asians identify new genetic loci influencing metabolic traits. *Nat. Genet.* **43**, 990–995 (2011).
- Cho, Y. S., Go, M. J., Kim, Y. J., Heo, J. Y., Oh, J. H., Ban, H. J. *et al.* A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat. Genet.* **41**, 527–534 (2009).
- Klimentidis, Y. C., Zhou, J. & Wineinger, N. E. Identification of allelic heterogeneity at type-2 diabetes loci and impact on prediction. *PLoS ONE* **9**, e113072 (2014).
- Phani, N. M., Guddattu, V., Bellampalli, R., Seenappa, V., Adhikari, P., Nagri, S. K. *et al.* Population specific impact of genetic variants in KCNJ11 gene to type 2 diabetes: a case-control and meta-analysis study. *PLoS ONE* **9**, e107021 (2014).
- Manchia, M., Cullis, J., Turecki, G., Rouleau, G. A., Uher, R. & Alda, M. The impact of phenotypic and genetic heterogeneity on results of genome wide association studies of complex diseases. *PLoS ONE* **8**, e76295 (2013).
- Peter, I., McCaffery, J. M., Kelley-Hedgepeth, A., Hakonarson, H., Reis, S., Wagenknecht, L. E. *et al.* Association of type 2 diabetes susceptibility loci with one-year weight loss in the look AHEAD clinical trial. *Obesity (Silver Spring)* **20**, 1675–1682 (2012).
- Cho, Y. S., Chen, C. H., Hu, C., Long, J., Ong, R. T., Sim, X. *et al.* Meta-analysis of genome-wide association studies identifies eight new loci for type 2 diabetes in east Asians. *Nat. Genet.* **44**, 67–72 (2011).
- Ma, R. C., Hu, C., Tam, C. H., Zhang, R., Kwan, P., Leung, T. F. *et al.* Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. *Diabetologia* **56**, 1291–1305 (2013).
- Sakai, K., Imamura, M., Tanaka, Y., Iwata, M., Hirose, H., Kaku, K. *et al.* Replication study for the association of 9 East Asian GWAS-derived loci with susceptibility to type 2 diabetes in a Japanese population. *PLoS ONE* **8**, e76317 (2013).
- Lee, P. H. & Shatkay, H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res.* **36**, D820–D824 (2008).
- Ward, L. D. & Kellis, M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res.* **44**, D877–D881 (2016).
- Bennet, L., Groop, L., Lindblad, U., Agardh, C. D. & Franks, P. W. Ethnicity is an independent risk indicator when estimating diabetes risk with FINDRISC scores: a cross sectional study comparing immigrants from the Middle East and native Swedes. *Prim. Care Diabetes* **8**, 231–238 (2014).
- Imamura, M., Shigemizu, D., Tsunoda, T., Iwata, M., Maegawa, H., Watada, H. *et al.* Assessing the clinical utility of a genetic risk score constructed using 49 susceptibility alleles for type 2 diabetes in a Japanese population. *J. Clin. Endocrinol. Metab.* **98**, E1667–E1673 (2013).
- Tam, C. H., Ho, J. S., Wang, Y., Lam, V. K., Lee, H. M., Jiang, G. *et al.* Use of net reclassification improvement (NRI) method confirms the utility of combined genetic risk score to predict type 2 diabetes. *PLoS ONE* **8**, e83093 (2013).
- Chen, Z., Pereira, M. A., Seielstad, M., Koh, W. P., Tai, E. S., Teo, Y. Y. *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* **9**, e87762 (2014).
- Cooke, J. N., Ng, M. C., Palmer, N. D., An, S. S., Hester, J. M., Freedman, B. I. *et al.* Genetic risk assessment of type 2 diabetes-associated polymorphisms in African Americans. *Diabetes Care* **35**, 287–292 (2012).
- Villegas, R., Goodloe, R. J., McClellan, B. E. Jr, Boston, J. & Crawford, D. C. Gene-carbohydrate and gene-fiber interactions and type 2 diabetes in diverse populations from the National Health and Nutrition Examination Surveys (NHANES) as part of the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) study. *BMC Genet.* **15**, 69 (2014).
- Che, R. & Moutsier-Reif, A. A. Evaluation of genetic risk score models in the presence of interaction and linkage disequilibrium. *Front. Genet.* **4**, 138 (2013).
- Knuppel, S., Meidtnr, K., Arregui, M., Holzhutter, H. G. & Boeing, H. Joint effect of unlinked genotypes: application to type 2 diabetes in the EPIC-Potsdam case-cohort study. *Ann. Hum. Genet.* **79**, 253–263 (2015).
- Dauriz, M., Porneala, B. C., Guo, X., Bielak, L. F., Peyser, P. A., Durant, N. H. *et al.* Association of a 62 variant type 2 diabetes genetic risk score with markers of subclinical atherosclerosis: a transethnic, multicenter study. *Circ. Cardiovasc. Genet.* **8**, 507–515 (2013).
- Huang, T., Qi, Q., Zheng, Y., Ley, S. H., Manson, J. E., Hu, F. B. *et al.* Genetic predisposition to central obesity and risk of type 2 diabetes: two independent cohort studies. *Diabetes Care* **38**, 1306–1311 (2015).
- Robiou-du-Pont, S., Bonnefond, A., Yengo, L., Vaillant, E., Lobbens, S., Durand, E. *et al.* Contribution of 24 obesity-associated genetic variants to insulin resistance, pancreatic beta-cell function and type 2 diabetes risk in the French population. *Int. J. Obes. (Lond)* **37**, 980–985 (2013).

Supplementary Information accompanies the paper on Journal of Human Genetics website (<http://www.nature.com/jhg>)