

厚生労働科学研究費補助金

創薬基盤推進研究事業

生活習慣と遺伝子型による2型糖尿病発症リスク
予測法の開発

平成20年度 総括・分担研究報告書

主任研究者 原 一雄

平成21(2009)年 4月

目 次

I. 総括研究報告		
生活習慣と遺伝子型による2型糖尿病発症リスク予測法の開発		12
原 一雄 -----	3	
II. 分担研究報告		
1. 新規2型糖尿病感受性遺伝子の機能解明		
門脇 孝 -----	8	
2. 日本人2型糖尿病感受性遺伝子の同定		
山内 敏正 -----	11	
III. 研究成果の刊行に関する一覧表	-----	15
IV. 研究成果の刊行物・別刷	-----	16

厚生労働科学研究費補助金（創薬基盤推進研究事業）

総括研究報告書

生活習慣と遺伝子型による2型糖尿病発症リスク予測法の開発

主任研究者 原 一雄 東京大学医学部附属病院 講師

| 3

研究要旨：本研究は、日本人で2型糖尿病発症リスクを上昇させる感受性遺伝子とその多型を明らかにすることで、糖尿病の1次予防を効率的に行うハイリスク者スクリーニング法開発のための基礎的データを得ることを目的としている。既に、罹患同胞対法による全ゲノム解析と候補遺伝子アプローチによって、AMP キナーゼ $\alpha 2$ サブユニット遺伝子、HNF4 α 遺伝子が日本人における2型糖尿病感受性遺伝子であること、欧米人で明らかになった2型糖尿病感受性遺伝子について日本人における意義を検討し、TCF7L2 遺伝子、HHEX 遺伝子が日本人においても糖尿病感受性遺伝子として一定の役割を担っていることを明らかにした。また、染色体 11 番の領域で有意に2型糖尿病と相関する新規の遺伝子多型を見出したが、当該遺伝子の欠損マウスについてその表現型を解析したところ、インスリン分泌低下を示し、当該遺伝子がグルコース応答性のインスリン分泌に関与していることが判明した。これまでに判明した複数の糖尿病感受性遺伝子多型の情報を統合して判定することによって、2型糖尿病の高リスク者をスクリーニング出来る可能性が示唆された。

分担研究者

門脇 孝 東京大学医学部附属病院
教授

山内 敏正 東京大学医学部附属病院
特任准教授

戸辺 一之 富山大学医学部附属病院
教授

抵抗性に関わっているのかを、当該遺伝子改変動物の表現型を解析することで明らかにする。また、諸外国で2型糖尿病感受性遺伝子として報告された遺伝子に、日本人での糖尿病易罹患性における意義を検討する。以上により、2型糖尿病発症リスク予測法開発の基礎的データを得る。

A. 研究目的

2型糖尿病はインスリン分泌低下にインスリン抵抗性が加わって発症する多因子病である。本研究では、インスリン分泌不全によって2型糖尿病を発症させやすくしている感受性遺伝子と、インスリン抵抗性によって2型糖尿病を発症させやすくしている感受性遺伝子を同定する。機能が未知の新規2型糖尿病感受性遺伝子については、インスリン分泌不全に関わっているのか、インスリン

B. 研究方法

(1)日本人2型糖尿病感受性遺伝子の同定：罹患同胞対法による全ゲノム解析によって同定した9箇所(1p36-p32, 2q34, 3q26-q28, 6p23, 7p22-p21, 9p, 11p13-p12, 15q13-q21, 20q12-q13)の日本人2型糖尿病感受性遺伝子座についてインスリン作用あるいはインスリン分泌に関与することが分かっている機能的候補遺伝子に

ついてSNPを利用した相関解析によって日本人における意義を検討する。また、欧米から2型糖尿病感受性遺伝子として報告されたHHEX遺伝子などの遺伝子についてSNPを利用した相関解析を行い日本人における意義を検討した。また、HOMA(homeostasis model assessment)によるインスリン抵抗性指標あるいはインスリン分泌指標と遺伝子多型との相関を検討し、日本人における意義を検討した。

(2) 機能未知の新規2型感受性遺伝子の同定：SNPによる相関解析によって統計学的に2型糖尿病と相関しても、当該遺伝子の機能が明らかにすることが、2型糖尿病遺伝素因における役割の解明には必須である。本研究ではSNPによる相関解析によって新たに2型糖尿病感受性遺伝子であることが推測された遺伝子について遺伝子欠損マウスの表現型を解析することによって、当該遺伝子がインスリン分泌に関与しているのか、インスリン抵抗性に関与しているのかを明らかにする。また、ヒト組織を利用した発現解析によって、SNPの遺伝子型が当該遺伝子の発現に影響を与えているか否かについて検討を行う。

(倫理面への配慮)平成15年7月30日に厚生労働省によって策定された「臨床研究に関する倫理指針」を遵守して研究を遂行する。その具体的な配慮として、臨床研究を実施するに当たり、被験者の個人情報の保護のために、本研究で提供される試料はすべて個人識別情報(カルテ番号、名前、住所など)を除き、連結可能匿名化した上で解析に利用される。連結可能のための対応表は他の一切の

コンピューターと切り離されたコンピューターに専用のIDとパスワードによって厳重に保管される。また、当該コンピューターは民間警備会社によるセキュリティシステムによって守られ、不特定多数の者の出入りができない専用の部屋に設置されるため個人情報の保護については厳格に行っている。

C. 研究結果

(1) 日本人2型糖尿病感受性遺伝子の同定：欧米人で2型糖尿病感受性遺伝子として報告されたHHEX遺伝子上の3つのSNPは互いに連鎖不平衡にあり、いずれも有意に2型糖尿病と相関した(rs5015480, オッズ比[OR]=1.46 [95% CI 1.20-1.77], $p=2.0 \times 10^{-4}$; rs7923837, [OR]=1.40 [95% CI 1.17-1.68], $p=2.0 \times 10^{-4}$; rs1111875, (OR=1.30 [95% CI 1.11-1.52], $p=0.0013$)。HHEX遺伝子と近隣のKIF11遺伝子、IDE遺伝子を含む267kbの範囲において、27個のtag SNPに関して連鎖不平衡マップを作成したところ、これらの3遺伝子がひとつの連鎖不平衡ブロックに含まれていた欧米人と異なり、日本人ではKIF11, IDE遺伝子を含むブロックと、HHEXを含むブロックの2つが互いに独立していた。よって日本人における検討から、この範囲における2型糖尿病感受性遺伝子はHHEX遺伝子であることが示された。また、FTOrs8050136, CDKAL1rs7756992, CDKN2Brs10811661, SCL30A8rs13266634多型も多重検定であることを考慮しなければ2型糖尿病と有意な相関を示した(FTO, OR=1.22 [95% CI 1.03-1.46], $p=$

0.025; CDKAL1, OR=1.20 [95% CI 1.03-1.39], p=0.017; CDKN2B, OR=1.22 [95% CI 1.05-1.41], p=0.0076; SCL30A8, OR = 1.19 [95% CI 1.03-1.37], p=0.016)。このうち CDKAL1rs7756992 と CDKN2Brs10811661 多型は HOMAによるインスリン分泌指標と有意な相関を示し、リスクアリル保持者ではインスリン分泌が有意に低下していた。HOMA-IR によるインスリン抵抗性指標と有意な相関を示す遺伝子多型は存在しなかった。欧米人においては肥満を介して2型糖尿病感受性を発揮している肥満遺伝子として報告されているFTO 遺伝子多型は、日本人においてはBMIとの相関は認められなかった。

(2) 機能未知の新規2型感受性遺伝子の同定:日本人2型糖尿病感受性遺伝子座のうち、染色体11番(11p13-p12)の領域は、日本人を対象とした他のグループによる全ゲノム解析によっても一致して糖尿病との連鎖が報告され (J Hum Genet 2004 49:629-34)日本人に特異的な2型糖尿病感受性遺伝子座として極めて有望な領域である。そこで本領域について網羅的なSNPによる相関解析を行ったところ、複数のDNAパネルで2型糖尿病と相関する遺伝子多型を見出した。本遺伝子が糖代謝における役割について報告がなかったものの、公開データベース(NCBIなど)から本遺伝子が糖代謝関連臓器で実際に発現していることが判明し有望であると考えられた。そこで本遺伝子の欠損マウスの表現型を解析することによって、本遺伝子の機能について個体レベルで検討を行った。本遺伝子欠損マウスは

OGTT (Oral Glucose Tolerance Test:経口ブドウ糖負荷試験)におけるインスリン分泌能が野生型に比べて低下していること、高脂肪食負荷時の膵島面積が野生型に比べて小さいことから、本遺伝子がインスリン分泌低下を来す2型糖尿病感受性遺伝子であることが示唆された。同時に、ヒトから同意を得て採取された脂肪組織の遺伝子発現解析によって、本遺伝子多型の糖尿病のリスクを上昇させるアリル保持者は非保持者に比較して有意に本遺伝子の発現が低下していることが判明した。

D. 考察

HHEX 遺伝子など欧米で報告された2型糖尿病感受性遺伝子は概ね日本人においても2型糖尿病感受性遺伝子として一定の役割を担っているものと考えられた。但し、TCF7L2 遺伝子は欧米人では糖尿病のリスクを上昇させるアリルの頻度が高く、糖尿病の遺伝素因全体に占める役割が大きいのに比べて、日本人ではその頻度が低いことから役割はそれほど大きくないと考えられた。新たに2型糖尿病との相関を見出した染色体11番上の遺伝子はグルコース応答性のインスリン分泌に一定の役割を担っており、当該遺伝子内にあるSNPの遺伝子型によって発現が低下するためにインスリン分泌不全を来して糖尿病を発症しやすくなることが示唆された。

更に、これまでに明らかになった2型糖尿病感受性遺伝子多型の糖尿病リスクを上昇させるアリル(リスクアリル)の個数が多いほど糖尿病リスクは有意に上昇するものの、リスクアリルの保持数に年齢、性別、BMI(Body mass

index)などの臨床指標も合わせた方が、リスクアレルの個数のみよりも、より正確に糖尿病のリスクを推定できることが判明した。

E. 結論

個々の2型糖尿病感受性遺伝子単独では糖尿病リスクの上昇は30%程度であり、それ自体で糖尿病を予測する感度・特異度は高くない。しかしながら本研究や内外の研究で明らかになった複数の感受性遺伝子に、今後も明らかにされる糖尿病感受性遺伝子、年齢・性別・BMIなどの臨床情報を組み合わせることによって糖尿病リスクがある程度の正確性を持って予測できることが期待される。日本における糖尿病罹患者数の多さや今後の増加率を考慮すると、今後この分野の発展がますます望まれる。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

(1) Miyake K, Yang W, Hara K, Yasuda K, Horikawa Y, Osawa H, Furuta H, Ng MC, Hirota Y, Mori H, Ido K, Yamagata K, Hinokio Y, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Wang HY, Tanahashi T, Nakamura N, Takeda J, Maeda E, Yamamoto K, Tokunaga K, Ma RC, So WY, Chan J C, Kamatani N, Makino H, Nanjo K, Kadowaki T, Kasuga M. Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong

evidence of the association. *J Hum Genet.* 54(4): 236-41, 2009

(2) Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, Hirota Y, Mori H, Jonsson A, Sato Y, Yamagata K, Hinokio Y, Wang HY, Tanahashi T, Nakamura N, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Takeda J, Maeda E, Shin HD, Cho YM, Park KS, Lee HK, Ng MC, Ma RC, So WY, Chan JC, Lyssenko V, Tuomi T, Nilsson P, Groop L, Kamatani N, Sekine A, Nakamura Y, Yamamoto K, Yoshida T, Tokunaga K, Itakura M, Makino H, Nanjo K, Kadowaki T, Kasuga M. Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus. *Nature Genetics* 40(9):1092-7, 2008

(3) Unoki H, Takahashi A, Kawaguchi T, Hara K, Horikoshi M, Andersen G, Ng DP, Holmkvist J, Borch-Johnsen K, Jørgensen T, Sandbaek A, Lauritzen T, Hansen T, Nurbaya S, Tsunoda T, Kubo M, Babazono T, Hirose H, Hayashi M, Iwamoto Y, Kashiwagi A, Kaku K, Kawamori R, Tai ES, Pedersen O, Kamatani N, Kadowaki T, Kikkawa R, Nakamura Y, Maeda S. SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations. *Nature Genetics* 40(9):1098-102, 2008

(4) Wang G, Watanabe M, Imai Y, Hara K, Manabe I, Maemura K, Horikoshi M, Kohro T, Amiya E, Sugiyama T, Fujita T, Kadowaki T, Yamazaki T, Nagai R. Genetic variations of Mrf-2/ARID5B confer risk of coronary atherosclerosis in the Japanese population. *Int Heart J.* 49(3):313-27, 2008

(5) Miyake K, Horikawa Y, Hara K, Yasuda K, Osawa H, Furuta H, Hirota Y, Yamagata K, Hinokio Y, Oka Y, Iwasaki N, Iwamoto Y, Yamada Y, Seino Y, Maegawa H, Kashiwagi A, Yamamoto K, Tokunaga K, Takeda J, Makino H, Nanjo K, Kadowaki T, Kasuga M. Association of TCF7L2 polymorphisms with susceptibility to type 2 diabetes in 4,087 Japanese subjects. *J Hum Genet.* 53(2):174-80, 2008

2. 学会発表

(1) Horikoshi M, Hara K, Ito C, Shojima N, Genomic Disorders 2008 Wellcome Trust Conference Centre, Cambridge UK, March, 2008

(2) Horikoshi M, Hara K, Shojima N, Ito C et al. Variations in the *TCF7L2* and *HHEX* genes independently confer Type 2 Diabetes Susceptibility on Japanese Population.

68th Scientific Sessions of the American Diabetes Association. 1146-P San Francisco, USA. June, 2008

H. 知的財産権の出願・登録状況

1. 特許出願

なし

2. 実用新案登録

なし

3. その他

なし

厚生労働科学研究費補助金（創薬基盤推進研究事業）

総括研究報告書

新規2型糖尿病感受性遺伝子の機能解明

分担研究者 門脇 孝 東京大学医学部附属病院 教授

18

研究要旨：糖尿病は細小血管症によるQOL(生活の質)の低下と動脈硬化促進による心筋梗塞・脳卒中発症のリスク増大を介して健康寿命を短縮している最大の問題の一つである。本研究は、日本人の2型糖尿病遺伝素因を解明し2型糖尿病の1次予防を効率的に行うためのハイリスク者スクリーニング法開発の基礎的データを得ることを目的としている。9ヶ所の日本人糖尿病感受性遺伝子座の中で染色体11番(11p13-p12)の領域については網羅的相関解析によって一貫して有意に2型糖尿病と相関する遺伝子多型を見出した。本遺伝子の糖代謝における役割は全く不明であったため、当該遺伝子の欠損マウスの表現型の解析により個体レベルで本遺伝子の機能について検討したところ、グルコース応答性のインスリン分泌に関与していることが明らかになった。

A.研究目的

2型糖尿病はインスリン分泌低下にインスリン抵抗性が加わって発症する多因子病である。本研究では、インスリン分泌不全によって2型糖尿病を発症させやすくしている感受性遺伝子と、インスリン抵抗性によって2型糖尿病を発症させやすくしている感受性遺伝子を同定する。機能が未知の新規2型糖尿病感受性遺伝子については、インスリン分泌不全に関わっているのか、インスリン抵抗性に関わっているのかを、当該遺伝子改変動物の表現型を解析することで明らかにする。また、諸外国で2型糖尿病感受性遺伝子として報告された遺伝子に、日本人での糖尿病易罹患性における意義を検討する。以上により、2型糖尿病発症リスク予測法開発の基礎的データを得る。

B.研究方法

全ゲノム解析によってマップした2型糖尿病感受性遺伝子座のうち、染色体1

番、3番、20番以外の6箇所の染色体領域に含まれるSNPのうち、国際HAPMAP projectによるハプロタイプブロック・ハプロタイプ標識SNPの情報によって疾患感受性遺伝子を効率よく同定することができるSNPを選択する。3つの独立した糖尿病・糖代謝正常者のDNAパネル(既に文書による同意を得てDNAを抽出済み)で一貫して2型糖尿病と相関を示すSNPを同定し、日本人2型糖尿病感受性遺伝子としての意義を検討する。機能が既知であるか否かに関わらず日本人2型糖尿病感受性遺伝子座に含まれるSNPを網羅的に解析するため、機能の未知な糖尿病感受性遺伝子も同定することが出来る。そのような機能未知であるが、3つのDNAパネルで一貫して2型糖尿病と相関を示す多型を含む遺伝子について遺伝子欠損マウスの表現型を詳細に解析し、当該遺伝子がインスリン分泌不全に関与するのかインスリン抵抗性に関与するのかなどについて明らかにする。

(倫理面への配慮)「ヒトゲノム・遺伝子解析研究のための倫理指針」「疫学研究に関する倫理指針」「臨床研究に関する倫理指針」を遵守し、書面による本研究への同意が得られた者のみを対象として研究を遂行する。更に指針の策定を受けて本学に設置された倫理審査委員会の承認を得て(承認番号 421)研究計画書に従い研究を実施する。具体的な配慮の項目としては以下の通り。①研究で使用される試料に関して:既に提供されているDNAなどの試料は、遺伝子解析研究での利用を明示した上で同意が文書で得られているもののみを研究対象としており問題はない。本研究で新規に提供される試料は、全て指針に従って採取される。②個人情報の保護に関して:本研究で提供される試料はすべて、個人識別情報(カルテ番号,名前,住所など)を管理する個人識別情報管理者により匿名化される。連結可能のための対応表は、個人識別情報管理者のもとで厳重に管理され、個人の情報を処理するコンピューターは他の一切のコンピューターと切り離され、個人識別情報管理者によって保管される。更に、個人識別情報管理者以外は、連結可能匿名化番号と個人識別情報との連結が不可能なようにする。東大病院では主任研究者の原一雄により個人情報匿名化システムが確立している(J.Hum.Genet.48:327-330, 2003) ③予測される試料提供者に対する危険や不利益に関して:試料提供は主として前腕の静脈からの採血によっており身体的危険はほとんどないといっている。また提供された試料は解析に先立って速やかに匿名化されるので、試料等提供者の尊厳と人権は十分に保護されていると考えられる。また、本研究

は多因子病としての2型糖尿病の感受性遺伝子を対象としており、単一遺伝子病の様に必ずしも遺伝カウンセリングが必要となるケースは少ないと思われるが、予想外の遺伝病の存在が明らかになった場合の予測される病態や予後、治療方針などにつき、結果の開示に伴い本人および家族に医学的あるいは心理的問題を生じる可能性があり、遺伝カウンセリングなどの体制を準備している。

19

C.研究結果

罹患同胞対による全ゲノム解析で同定した9箇所の日本人2型糖尿病感受性遺伝子座のうち染色体11番(11p13-p12)の領域は、日本人を対象とした他のグループによる全ゲノム解析によっても一致して糖尿病との連鎖が報告され(J Hum Genet 2004 49:629-34)、日本人に特異的な2型糖尿病感受性遺伝子座として極めて有望な領域である。そこで本領域について網羅的なSNPによる相関解析を行ったところ、2型糖尿病と一貫して相関する遺伝子多型を同定した。本遺伝子の生体内での機能については殆ど明らかにされていなかったため、統計学的には本遺伝子が2型糖尿病感受性遺伝子であると考えられたものの、生物学的にみても本遺伝子が糖脂質代謝に関与していることの確認が必要であった。そこで本遺伝子の欠損マウスの解析により本遺伝子の機能について個体レベルで検討を行った。本遺伝子欠損マウスはOGTT(Oral Glucose Tolerance Test:経口ブドウ糖負荷試験)におけるインスリン分泌能が野生型に比べて低下していること、高脂肪食負荷時の膵島面積が野生型に比べて小さいことから、本遺伝子がインス

リン分泌低下を来す2型糖尿病感受性遺伝子であることが示唆された。同時に、ヒトから同意を得て採取された脂肪組織の遺伝子発現解析によって、本遺伝子多型の糖尿病のリスクを上昇させるアリル保持者は非保持者に比較して有意に本遺伝子の発現が低下していることが判明した。

D. 考察

新たに2型糖尿病との相関を見出した染色体11番上の遺伝子はグルコース応答性のインスリン分泌に一定の役割を担っており、当該遺伝子内にあるSNPの遺伝子型によって発現が低下するためにインスリン分泌不全を来して糖尿病を発症しやすくなることが示唆された。

E. 結論

新たに2型糖尿病との相関を見出した染色体11番上の遺伝子について、どのような経路を利用してグルコース応答性のインスリン分泌に関与しているかを今後明らかにしていく。

F. 健康危険情報

特になし

G. 研究発表

1. 論文発表

1: Horikoshi M, Hara K, Ito C, Shojima N, Nagai R, Ueki K, Froguel P, Kadowaki T. Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia* 50(12):2461-6, 2007

2: Ochi M, Osawa H, Hirota Y, Hara K, Tabara Y, Tokuyama Y, Shimizu I,

Kanatsuka A, Fujii Y, Ohashi J, Miki T, Nakamura N, Kadowaki T, Itakura M, Kasuga M, Makino H. Frequency of the G/G genotype of resistin single nucleotide polymorphism at -420 appears to be increased in younger -onset type 2 diabetes. *Diabetes*. 56(11):2834-8, 2007.

3: Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, Kadowaki T. A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* 50(4):747-51, 2007

H. 知的財産権の出願・登録状況

(ア) 特許出願

なし

(イ) 実用新案登録

なし

(ウ) その他

なし

厚生労働科学研究費補助金（ヒトゲノム・再生医療等研究事業）
分担研究報告書

生活習慣と遺伝子型による2型糖尿病発症リスク 予測法の開発
—インスリン抵抗性に関する糖尿病感受性遺伝子の同定—
分担研究者 山内 敏正 東京大学医学部附属病院 特任准教授

| 11

研究要旨：2型糖尿病は心筋梗塞・脳卒中発症のリスク増大によって国民の健康をおびやかしている。本研究は、先に明らかにした日本人2型糖尿病感受性遺伝子座について一塩基多型による相関解析を行い、2型糖尿病遺伝素因の全貌を解明する。既に、日本人の糖尿病感受性遺伝子座として同定した9ヶ所の染色体領域のうち染色体1番の領域（1p36-p32）についてはAMPキナーゼ α 2サブユニット遺伝子がインスリン抵抗性を惹起して2型糖尿病を起こしやすくしている感受性遺伝子であること、染色体20番の領域（20q12-q13）はHNF4 α 遺伝子がインスリン分泌不全を来たして2型糖尿病を発症しやすくしている2型糖尿病感受性遺伝子であることを報告した。また、欧米から2型糖尿病感受性遺伝子として報告されたHHEXが日本人においても一定の役割を担っていることを明らかにした。

A. 研究目的

2型糖尿病はインスリン分泌低下にインスリン抵抗性が加わって発症する多因子病である。本研究では、インスリン分泌不全によって2型糖尿病を発症させやすくしている感受性遺伝子と、インスリン抵抗性によって2型糖尿病を発症させやすくしている感受性遺伝子を同定する。機能が未知の新規2型糖尿病感受性遺伝子については、インスリン分泌不全に関わっているのか、インスリン抵抗性に関わっているのかを、当該遺伝子改変動物の表現型を解析することで明らかにする。また、諸外国で2型糖尿病感受性遺伝子として報告された遺伝子に、日本人での糖尿病易罹患性における意義を検討する。以上により、2型糖尿病発症リスク予測法開発の基礎的データを得る。

B. 研究方法

全ゲノム解析によってマップした2型糖尿病感受性遺伝子座のうち、染色体1番、3番、20番以外の6箇所の染色体領域に含まれるSNPのうち、国際HAPMAP projectによるハプロタイプブロック・ハプロタイプ標識SNPの情報によって疾患感受性遺伝子を効率よく同定することができるSNPを選択する。3つの独立した糖尿病・糖代謝正常者のDNAパネル（既に文書による同意を得てDNAを抽出済み）で一貫して2型糖尿病と相関を示すSNPを同定し、日本人2型糖尿病感受性遺伝子としての意義を検討する。日本人2型糖尿病感受性遺伝子座に含まれる遺伝子のうちインスリン作用に関与している機能的にも候補となる遺伝子についてはより重点的・詳細にSNPのタイピングを行い、日本人2型糖尿病感受性遺伝子としての意義を明らかにしていく。

（倫理面への配慮）「ヒトゲノム・遺

伝子解析研究のための倫理指針」「疫学研究に関する倫理指針」「臨床研究に関する倫理指針」を遵守し、書面による本研究への同意が得られた者のみを対象として研究を遂行する。更に指針の策定を受けて本学に設置された倫理審査委員会の承認を得て(承認番号 421)研究計画書に従い研究を実施する。具体的な配慮の項目としては以下の通り。①研究で使用される試料に関して：既に提供されている DNA などの試料は、遺伝子解析研究での利用を明示した上で同意が文書で得られているもののみを研究対象としており問題はない。本研究で新規に提供される試料は、全て指針に従って採取される。②個人情報の保護に関して：本研究で提供される試料はすべて、個人識別情報(カルテ番号、名前、住所など)を管理する個人識別情報管理者により匿名化される。連結可能のための対応表は、個人識別情報管理者のもとで厳重に管理され、個人の情報を処理するコンピューターは他の一切のコンピューターと切り離され、個人識別情報管理者によって保管される。更に、個人識別情報管理者以外は、連結可能匿名化番号と個人識別情報との連結が不可能なようにする。東大病院では主任研究者の原一雄により個人情報匿名化システムが確立している(J.Hum.Genet.48:327-330, 2003)。③予測される試料提供者に対する危険や不利益に関して：試料提供は主として前腕の静脈からの採血によっており身体的危険はほとんどないといつてよい。また提供された試料は解析に先立って速やかに匿名化されるので、試料等提供者の尊厳と人権

は十分に保護されていると考えられる。また、本研究は多因子病としての2型糖尿病の感受性遺伝子を対象としており、単一遺伝子病の様に必ずしも遺伝カウンセリングが必要となるケースは少ないと思われるが、予想外の遺伝病の存在が明らかになった場合の予測される病態や予後、治療方針などにつき、結果の開示に伴い本人および家族に医学的あるいは心理的問題を生じる可能性があり、遺伝カウンセリングなどの体制を準備している。

| 12

C. 研究結果

欧米人で2型糖尿病感受性遺伝子として報告された HHEX 遺伝子上の3つの SNP は互いに連鎖不平衡にあり、いずれも有意に2型糖尿病と関連した(rs5015480, オッズ比[OR]=1.46 [95% CI 1.20-1.77], $p=2.0 \times 10^{-4}$; rs7923837, [OR]=1.40 [95% CI 1.17-1.68], $p=2.0 \times 10^{-4}$; rs1111875, (OR=1.30 [95% CI 1.11-1.52], $p=0.0013$)。HHEX 遺伝子と近隣の KIF11 遺伝子、IDE 遺伝子を含む 267kb の範囲において、27 個の tag SNP に関して連鎖不平衡マップを作成したところ、これらの3遺伝子がひとつの連鎖不平衡ブロックに含まれていた欧米人と異なり、日本人では KIF11, IDE 遺伝子を含むブロックと、HHEX を含むブロックの2つが互いに独立していた。よって日本人における検討から、この範囲における2型糖尿病感受性遺伝子は HHEX 遺伝子であることが示された。また、FTOrs8050136, CDKAL1rs7756992, CDKN2Brs10811661, SCL30A8rs132666

34 多型も多重検定であることを考慮しなければ2型糖尿病と有意な相関を示した(FTO, OR=1.22 [95% CI 1.03-1.46], p=0.025; CDKAL1, OR=1.20 [95% CI 1.03-1.39], p=0.017; CDKN2B, OR=1.22 [95% CI 1.05-1.41], p=0.0076; SCL30A8, OR=1.19 [95% CI 1.03-1.37], p=0.016)。このうち CDKAL1rs7756992 と CDKN2B rs10811661 多型はHOMAによるインスリン分泌指標と有意な相関を示し、リスクアリル保持者ではインスリン分泌が有意に低下していた。HOMA-IRによるインスリン抵抗性指標と有意な相関を示す遺伝子多型は存在しなかった。欧米人においては肥満を介して2型糖尿病感受性を発揮している肥満遺伝子として報告されているFTO 遺伝子多型は、日本人においてはBMIとの相関は認められなかった。

D. 考察

HHEX 遺伝子など欧米で報告された2型糖尿病感受性遺伝子は概ね日本人においても2型糖尿病感受性遺伝子として一定の役割を担っているものと考えられた。但し、TCF7L2 遺伝子は欧米人では糖尿病のリスクを上昇させるアリルの頻度が高く、糖尿病の遺伝素因全体に占める役割が大きいのに比べて、日本人ではその頻度が低いことから役割はそれほど大きくないと考えられた。新たに2型糖尿病との相関を見出した染色体11番上の遺伝子はグルコース応答性のインスリン分泌に一定の役割を担っており、当該遺伝子内にあるSNPの遺伝子型によって発現が低下するため

にインスリン分泌不全を来して糖尿病を発症しやすくなることが示唆された。

E. 結論

欧米から報告された2型糖尿病感受性遺伝子は概ね日本人においても2型糖尿病のリスクを上昇させるが、集団全体に与える影響は異なっている。

F. 健康危険情報

特になし

G. 研究発表

- 1: Horikoshi M, Hara K, Ito C, Shojima N, Nagai R, Ueki K, Froguel P, Kadowaki T. Variations in the HHEX gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabetologia*.50(12):2461-6, 2007
- 2: Ochi M, Osawa H, Hirota Y, Hara K, Tabara Y, Tokuyama Y, Shimizu I, Kanatsuka A, Fujii Y, Ohashi J, Miki T, Nakamura N, Kadowaki T, Itakura M, Kasuga M, Makino H. Frequency of the G/G genotype of resistin single nucleotide polymorphism at -420 appears to be increased in younger-onset type 2 diabetes. *Diabetes*. 56(11):2834-8, 2007.
- 3: Horikoshi M, Hara K, Ito C, Nagai R, Froguel P, Kadowaki T. A genetic variation of the transcription factor 7-like 2 gene is associated with risk of type 2 diabetes in the Japanese population. *Diabetologia* 50(4):747-51, 2007

2. 学会発表

なし

H.知的財産権の出願・登録状況

1. 特許出願
なし
2. 実用新案登録
なし
3. その他
なし

II. 研究成果の刊行一覧

雑誌

発表者氏名	論文タイトル名	発表誌名	巻	ページ	出版年 15
Miyake K, Yang W, Hara K, Yasuda K et al	Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong evidence of the association.	<i>J Hum Genet</i>	54	236-41	2009
Yasuda K, Miyake K, Horikawa Y, Hara K et al	Variants in KCNQ1 are associated with susceptibility to type 2 diabetes mellitus.	<i>Nature Genetics</i>	40	1092-7	2008
Unoki H, Takahashi A, Kawaguchi T, Hara K et al	SNPs in KCNQ1 are associated with susceptibility to type 2 diabetes in East Asian and European populations	<i>Nature Genetics</i>	40	1098-102	2008
Wang G, Watanabe M, Imai Y, Hara K et al	Genetic variations of Mrf-2/ARID5B confer risk of coronary atherosclerosis in the Japanese population.	<i>Int Heart J</i>	49	313-27	2008

ORIGINAL ARTICLE

Construction of a prediction model for type 2 diabetes mellitus in the Japanese population based on 11 genes with strong evidence of the association

Kazuaki Miyake¹, Woosung Yang², Kazuo Hara³, Kazuki Yasuda⁴, Yukio Horikawa⁵, Haruhiko Osawa⁶, Hiroto Furuta⁷, Maggie CY Ng⁸, Yushi Hirota¹, Hiroyuki Mori¹, Keisuke Ido^{2,20}, Kazuya Yamagata^{9,21}, Yoshinori Hinokio¹⁰, Yoshitomo Oka¹⁰, Naoko Iwasaki¹¹, Yasuhiko Iwamoto¹¹, Yuichiro Yamada^{12,22}, Yutaka Seino^{12,23}, Hiroshi Maegawa¹³, Atsunori Kashiwagi¹³, He-yao Wang^{4,24}, Toshihito Tanahashi¹⁴, Naoto Nakamura¹⁵, Jun Takeda⁵, Eiichi Maeda², Ken Yamamoto¹⁶, Katsushi Tokunaga¹⁷, Ronald CW Ma⁸, Wing-Yee So⁸, Juliana CN Chan⁸, Naoyuki Kamatani¹⁸, Hideichi Makino⁶, Kishio Nanjo⁷, Takashi Kadowaki³ and Masato Kasuga^{1,19}

Prediction of the disease status is one of the most important objectives of genetic studies. To select the genes with strong evidence of the association with type 2 diabetes mellitus, we validated the associations of the seven candidate loci extracted in our earlier study by genotyping the samples in two independent sample panels. However, except for *KCNQ1*, the association of none of the remaining seven loci was replicated. We then selected 11 genes, *KCNQ1*, *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGFBP2*, *SLC30A8*, *HHEX*, *GCKR*, *HNF1B*, *KCNJ11* and *PPARG*, whose associations with diabetes have already been reported and replicated either in the literature or in this study in the Japanese population. As no evidence of the gene–gene interaction for any pair of the 11 loci was shown, we constructed a prediction model for the disease using the logistic regression analysis by incorporating the number of the risk alleles for the 11 genes, as well as age, sex and body mass index as independent variables. Cumulative risk assessment showed that the addition of one risk allele resulted in an average increase in the odds for the disease of 1.29 (95% CI=1.25–1.33, $P=5.4 \times 10^{-53}$). The area under the receiver operating characteristic curve, an estimate of the power of the prediction model, was 0.72, thereby indicating that our prediction model for type 2 diabetes may not be so useful but has some value. Incorporation of data from additional risk loci is most likely to increase the predictive power.

Journal of Human Genetics (2009) 54, 236–241; doi:10.1038/jhg.2009.17; published online 27 February 2009

Keywords: gene–gene interaction; genome-wide association study; prediction model; single nucleotide polymorphism (SNP); type 2 diabetes mellitus

¹Division of Diabetes, Metabolism and Endocrinology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Kobe, Japan; ²Clinical Genome Informatics Center, Kobe University Graduate School of Medicine, Kobe, Japan; ³Department of Metabolic Diseases, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; ⁴Department of Metabolic Disorder, Research Institute, International Medical Center of Japan, Tokyo, Japan; ⁵Division of Molecule and Structure, Department of Diabetes and Endocrinology, Gifu University School of Medicine, Gifu, Japan; ⁶Department of Molecular and Genetic Medicine, Ehime University Graduate School of Medicine, Ehime, Japan; ⁷First Department of Medicine, Wakayama Medical University, Wakayama, Japan; ⁸Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Shatin, Hong Kong; ⁹Department of Metabolic Medicine, Graduate School of Medicine, Osaka University, Osaka, Japan; ¹⁰Division of Molecular Metabolism and Diabetes, Tohoku University Graduate School of Medicine, Sendai, Japan; ¹¹Department of Medicine, Diabetes Center, Tokyo Women's Medical University, Tokyo, Japan; ¹²Department of Diabetes and Clinical Nutrition, Kyoto University School of Medicine, Kyoto, Japan; ¹³Division of Endocrinology and Metabolism, Department of Medicine, Shiga University of Medical Science, Shiga, Japan; ¹⁴Division of Genetic Information, Institute for Genome Research, University of Tokushima, Tokushima, Japan; ¹⁵Department of Endocrinology and Metabolism, Kyoto Prefectural University of Medicine, Graduate School of Medical Sciences, Kyoto, Japan; ¹⁶Department of Molecular Genetics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Japan; ¹⁷Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; ¹⁸Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science, Tokyo Women's Medical University, Tokyo, Japan and ¹⁹Research Institute, International Medical Center of Japan, Tokyo, Japan

Correspondence: Dr M Kasuga, Research Institute, International Medical Center of Japan, 1-21-1 Toyama, Shinjuku-ku, Tokyo 162-8655, Japan.

E-mail: kasuga@ri.imcj.go.jp

²⁰Current address: Information Center for Medical Sciences, Tokyo Medical and Dental University, Tokyo, Japan.

²¹Current address: Faculty of Medical and Pharmaceutical Sciences, Department of Medical Biochemistry, Kumamoto University, Kumamoto, Japan.

²²Current address: Department of Internal Medicine, Akita University School of Medicine, Akita, Japan.

²³Current address: Kansai Electric Power Hospital, Osaka, Japan.

²⁴Current address: Shanghai Institute of Materia Medica, Chinese Academy of Science, Shanghai, China.

Received 22 December 2008; revised 25 January 2009; accepted 5 February 2009; published online 27 February 2009

INTRODUCTION

Genome-wide association studies (GWASs) have identified novel susceptibility genes for type 2 diabetes mellitus in Caucasians.^{1–5} *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8* and *HHEX* have been widely replicated as susceptibility genes for type 2 diabetes in Asian populations^{6–12} as well as in populations of European ancestry.^{13,14} We recently identified *KCNQ1* as a novel susceptibility gene, as well as seven other candidate susceptibility loci in a multistage GWAS for type 2 diabetes in the Japanese population, in which a total of 1612 cases and 1424 controls and 100 000 single nucleotide polymorphisms (SNPs) were included.¹⁵ *KCNQ1* was found to confer risk of type 2 diabetes with a relatively large effect size in Asian populations (odds ratio (OR) for Japanese, Chinese and Korean individuals of 1.42),¹⁵ which was similar to that demonstrated earlier for *TCF7L2* in the Japanese population.⁶

Follow-up of GWASs includes analysis of second-tier genes, meta-analysis for specific populations, as well as analysis of gene–gene or gene–environment interactions. A large-scale meta-analysis¹⁶ and an analysis of gene–gene interaction for susceptibility genes¹⁷ have been performed for type 2 diabetes in populations of European ancestry.

In this study, we attempted to confirm in independent subject panels of Japanese and Hong Kong Chinese individuals the associations of the seven candidate susceptibility loci that we identified in addition to *KCNQ1* in our GWAS of type 2 diabetes.¹⁵ However, as described in this article, we failed to replicate the associations of the seven loci with diabetes. We then attempted to extract genes with strong evidence of the associations with diabetes, and selected 11 genes, including *KCNQ1*. As we did not detect any gene–gene interaction between the 11 genes, we then attempted to construct a prediction model for this disease by using the data from the 11 genes, as well as age, gender and body mass index (BMI) as independent variables to obtain a comprehensive understanding of the genetic background of diabetes in the Japanese population.

MATERIALS AND METHODS

Validation of the results from a multistage GWAS in the Japanese population

Study subjects. We assembled two independent subject panels for our replication study: replication-Japanese and replication-Chinese. The 1000 cases and 1000 controls for the replication-Japanese panel were recruited by the Study Group of the Millennium Genome Project for Diabetes Mellitus. The inclusion criteria for diabetic patients were (i) an age at disease onset of 30–60 years and (ii) the absence of antibodies to GAD. Types of diabetes other than type 2 were excluded on the basis of clinical data. The criteria for controls included (i) an age of >50 years, (ii) no past history of a diagnosis of diabetes and (iii) an HbA_{1c} content of <5.8%.

For the replication-Chinese panel, subjects of southern Han Chinese ancestry, who resided in Hong Kong, were recruited. The cases consisted of 1416 individuals with type 2 diabetes selected from the Prince of Wales Hospital Diabetes Registry;^{5,18} 626 of these subjects had early-onset diabetes (age at diagnosis of <40 years) and a positive family history, whereas the remaining 790 patients were randomly selected from the registry. Patients with classic type 1 diabetes with acute ketotic presentation or a continuous requirement for insulin within 1 year of diagnosis were excluded. The controls consisted of 1577 subjects with normal glucose tolerance (fasting plasma glucose concentration of <6.1 mmol l⁻¹); 596 of these individuals were recruited either from the general population participating in a community-based screening program for cardiovascular risk or from hospital staff, whereas the remaining 981 subjects were recruited from a population-based screening program for cardiovascular risk in adolescents.¹⁹ The clinical characteristics of the subjects in each panel are summarized in Supplementary Table 1A. The study protocol was approved by the local ethics committee of each institution. Written informed consent was obtained from each subject.

Study design and statistical analysis. For the validation of the results from our earlier multistage GWAS,¹⁵ seven SNPs (rs2250402, rs2307027, rs3741872, rs574628, rs2233647, rs3785233 and rs2075931) were genotyped in the two panels either by sequence-specific primer–PCR analysis followed by fluorescence correlation spectroscopy²⁰ or by real-time PCR analysis with TaqMan probes (Applied Biosystems, Foster City, CA, USA). Differences in allele frequency between cases and controls for each SNP were evaluated by χ^2 with one degree of freedom. Meta-analysis was performed by the Mantel–Haenszel method (fixed-effects models) with the 'meta' package of the R-Project (<http://www.r-project.org>). A *P*-value of <0.05 was considered statistically significant.

Examination of gene–gene interaction and construction of a prediction model

Study subjects. In total, 2424 cases and 2424 controls of the Japanese population obtained by combining the second and third screening panels in our original study¹⁵ and the replication-Japanese panel of this study were included in this analysis (analysis-panel). The criteria for the second and third screening panels were described in the earlier report.¹⁵ The clinical characteristics of the subjects are summarized in Supplementary Table 1B.

Selection of the loci included in this study. Prediction of the phenotypes on the basis of genetic polymorphisms should include the genetic data from the loci with strong evidence of the association. Starting from 15 genes described in earlier reports, we selected 11 genes with strong evidence of the association on the basis of the data in the literature and on the results of the replication experiments in this study. Process of the selection of the 11 genes will be described in detail in Results.

Statistical methods. Multiplicative gene–gene interaction was evaluated for each pair of the 11 genes using an interaction term in addition to the terms for the pair of the genes in the logistic regression model. The genotypes for each locus were coded by 0, 1 and 2. Correction for multiple testing was performed by Bonferroni's method.

As there was no evidence for the presence of gene–gene interactions, we attempted to construct a phenotype prediction model by incorporating the number of risk alleles for the 11 loci as an independent variable in addition to age, gender and BMI. The Cochran–Armitage test was used to examine the trend of the increase in the odds by increasing the number of the risk alleles. To construct a prediction model, the log of odds was expressed by the linear combination of the independent variables. Coefficients for the variables were estimated by the logistic regression analysis after making disease (cases) or nondisease (controls) as the dependent variable. Using the coefficients estimated by the logistic regression analysis, we constructed a phenotype prediction model. To evaluate the prediction model, receiver operating characteristic (ROC) curves²¹ for the sensitivity and specificity of the prediction model with or without adjustment for age, sex and BMI were generated, and the area under the curve (AUC) was calculated from the ROC curve.

RESULTS

Validation of the results from a multistage GWAS in the Japanese population

We identified earlier 10 loci associated with type 2 diabetes by three-stage GWAS starting from 100 000 SNPs. Among the 10 loci, 3 SNPs were located in an intron of *KCNQ1*, and the association of this gene with diabetes was confirmatory.¹⁵ To validate the other seven loci for the association with type 2 diabetes, we analyzed them in two independent replication panels of Japanese and Han-Chinese individuals (Table 1, Supplementary Table 2). Only one SNP, rs2250402, which is located in *EIF2AK4*, was found to be significantly associated in the replication-Japanese panel (*P*=0.039, OR=1.17, 95% CI=1.01–1.36). However, neither this SNP (*P*=0.41, OR=1.05) nor any of the other six SNPs showed such an association in the replication-Chinese panel. Meta-analyses for these SNPs showed that rs2307027 in *KRT4* and rs3785233 in *A2BP1* yielded *P*-values of <0.05 and ORs between 1.12 and 1.13 (Table 1). When the original second and third screening

Table 1 Association study for the candidate susceptibility genes for type 2 diabetes selected by multistage screening in the Japanese population

SNP ID	Chr	Gene	Risk allele	Panel	RAF (DM)	RAF (NC)	P	OR	95% CI
rs2250402	15	EIF2AK4	C	Replication-Japanese	0.23	0.20	0.04	1.17	1.01–1.36
				Replication-Chinese	0.24	0.23	0.41	1.05	0.93–1.19
				Meta-analysis			0.05	1.10	1.00–1.20
rs2307027	12	KRT4	C	Replication-Japanese	0.18	0.17	0.17	1.12	0.95–1.32
				Replication-Chinese	0.14	0.13	0.16	1.11	0.96–1.29
				Meta-analysis			0.05	1.12	1.00–1.25
rs3741872	12	FAM60A	C	Replication-Japanese	0.25	0.24	0.18	1.11	0.96–1.28
				Replication-Chinese	0.23	0.22	0.21	1.08	0.96–1.22
				Meta-analysis			0.07	1.09	0.99–1.20
rs574628	20	ANGPT4	G	Replication-Japanese	0.60	0.61	0.46	0.95	0.84–1.08
				Replication-Chinese	0.65	0.65	0.59	1.03	0.93–1.15
				Meta-analysis			0.96	1.00	0.92–1.08
rs2233647	6	SPDEF	G	Replication-Japanese	0.86	0.87	0.70	0.97	0.81–1.16
				Replication-Chinese	0.94	0.93	0.54	1.07	0.87–1.31
				Meta-analysis			0.90	1.01	0.88–1.16
rs3785233	16	A2BP1	C	Replication-Japanese	0.18	0.16	0.19	1.12	0.95–1.32
				Replication-Chinese	0.13	0.12	0.10	1.14	0.97–1.34
				Meta-analysis			0.04	1.13	1.01–1.27
rs2075931	1	Intergenic	A	Replication-Japanese	0.67	0.66	0.85	1.01	0.89–1.16
				Replication-Chinese	0.73	0.74	0.27	0.94	0.84–1.05
				Meta-analysis			0.48	0.97	0.89–1.06

Abbreviations: Chr, chromosome; OR, odds ratio for risk allele frequency.

Assignment of risk alleles was based on the original study.¹⁵ Numbers of cases versus control subjects in the replication-Japanese and replication-Chinese panels were 1000 versus 1000 and 1416 versus 1577, respectively. RAF (DM) and RAF (NC) denote risk allele frequencies in cases and controls, respectively. *P* values were calculated for allele frequency. Meta-analysis was performed by the Mantel-Haenszel method (fixed-effects models). *P* values for the test of heterogeneity among panels joined in the Mantel-Haenszel tests were all >0.05.

panels were included in the meta-analyses, these two loci, as well as the SNPs in *EIF2AK4* (rs2250402) and *FAM60A* (rs3741872), gave *P*-values of <0.001 and ORs between 1.15 and 1.18 (Supplementary Table 3). However, the *P*-values did not reach the proposed significance of GWAS ($=5 \times 10^{-7}$).

Selection of polymorphisms for the prediction model

To construct a reliable prediction model for diabetes, polymorphisms with strong evidence of association should be used. From the previous literature, we selected 15 genes (including one intergenic marker), that is, *SLC30A8*, *HHEX*, *LOC387761*, *EXT2*, *CDKN2A/B*, *GCKR*, *IGF2BP2*, *CDKAL1*, *FTO*,^{1–5} *TCF7L2*,²² *KCNJ11*,²³ *PPARG*,²⁴ *WFS1*,²⁵ *HNF1B*²⁶ and *KCNQ1*,¹⁵ as candidate genes to be included in both gene–gene interaction analysis and construction of a prediction model. Starting from 23 SNPs in these 15 genes, we selected 11 SNPs in 11 genes according to the following process. There is sufficient evidence of the associations of *KCNQ1* and *TCF7L2* genes with diabetes as supported by replication studies in the Japanese population.^{6,15,27} In addition, *SLC30A8*, *HHEX*, *CDKN2A/B*, *IGF2BP2* and *CDKAL1* associated with the disease in the European population were found in our earlier study to be associated with the disease in the Japanese population as well.^{7–9}

To further extract genes with strong evidence of the association with diabetes, we attempted to replicate the associations reported earlier using our own data (analysis panel with 2424 cases and 2424 controls). For the 19 SNPs in *SLC30A8*, *HHEX*, *LOC387761*, *EXT2*, *CDKN2A/B*, *GCKR*, *IGF2BP2*, *CDKAL1*, *FTO*, *TCF7L2*, *KCNJ11*, *PPARG* and *KCNQ1*, we extracted genotyping data from our earlier studies^{6–9,15,27–29} and, if necessary, genotyped additional subjects to obtain a data set for 2424 cases and 2424 controls of the Japanese population (analysis panel). The SNPs in *WFS1* (rs6446482, rs734312)

and *HNF1B* (rs7501939, rs4430796) were genotyped for this study in the same individuals. SNPs with *P*-values for the test of deviation from the Hardy–Weinberg equilibrium of <0.01 were excluded for further analysis. When two SNPs were located in the same genomic region, the one with the lower *P*-value for the association test was selected for further analysis. *GCKR*, for which we earlier reported the marginal association with type 2 diabetes,⁷ was found to be associated with the disease in this enlarged Japanese panel ($P=1.7 \times 10^{-5}$; Supplementary Table 4). *KCNJ11* and *PPARG*, which have been included in the genes associated with diabetes in Caucasians, showed marginal associations ($P=0.066$ and $P=0.075$, respectively; Supplementary Table 4) in our panel. Two SNPs in *WFS1* and two SNPs in *HNF1B* were newly genotyped in the analysis panel. Although no association was apparent between *WFS1* and type 2 diabetes, both SNPs in *HNF1B* exhibited *P*-values of <0.05 (Supplementary Table 4). From these data, we included 11 SNPs in 11 genes as described above for the source of genotype data to be analyzed in both the examination of gene–gene interaction and the prediction of phenotypes.

Gene–gene interaction

We evaluated multiplicative gene–gene interaction for each pair of the 11 loci as described in Materials and methods. Two combinations, rs1801282 (*PPARG*) \times rs1470579 (*IGF2BP2*) (nominal $P=0.0025$) and rs1801282 \times rs3802177 (*SLC30A8*) (nominal $P=0.018$), showed *P*-values of less than 0.05 (Supplementary Figure 1). However, these *P*-values were not significant when Bonferroni's correction for multiple testing was applied (significance level, $0.05/55=9.1 \times 10^{-4}$). Although *PPARG* and *IGF2BP2* are located on the same chromosome (3p25 and 3q28, respectively), it is unlikely that loci on different arms of the same chromosome show significant linkage disequilibrium. *SLC30A8* is located on a different chromosome (8q24.11) from

PPARG. The reason why nominal *P*-values of these combinations showed less than 0.05 may be because of the low minor allele frequency of rs1801282.

Cumulative risk assessment for type 2 diabetes on the basis of susceptibility genes

As there was no evidence of gene–gene interaction between 11 SNPs of 11 genes, *SLC30A8*, *HHEX*, *CDKN2A/B*, *GCKR*, *IGF2BP2*, *CDKAL1*, *TCF7L2*, *KCNJ11*, *PPARG*, *KCNQ1* and *HNF1B*, they were included in the prediction model as independent variables with the additive effect (additive effect in the liability and multiplicative effect in the odds) without interaction terms. Effective numbers of cases and controls whose genotypes for the 11 loci were successfully obtained were 2316 and 2370, respectively. The Cochran–Armitage trend test gave a *P*-value of 4.7×10^{-56} for the trend in the increase in the odds for cases relative to controls with an increasing number of risk alleles for the 11 susceptibility loci (Supplementary Table 5). We then estimated ORs for type 2 diabetes in subjects with different numbers of risk alleles on the basis of the multiplicative model by logistic regression analysis with adjustment for age, sex and BMI. The ORs for type 2 diabetes in subjects with 7–18 risk alleles in comparison with those harboring 0–6 risk alleles are shown in Figure 1. An increase of one risk allele resulted in an average increase in the odds of 1.29 (95% CI=1.25–1.33, $P=5.4 \times 10^{-53}$, logistic regression analysis).

To predict disease status for type 2 diabetes in a given individual, we constructed a prediction model on the basis of the number of risk alleles or the liability value calculated from the number of risk alleles as well as age, sex and BMI. The coefficients to calculate the liability value were estimated with the logistic regression model. To estimate the predictive power of the model, we generated ROC curves as described in Materials and methods. The AUC was 0.63 when only the number of risk alleles was used for the prediction. When age, sex and BMI were also included, the AUC increased to 0.72 (Figure 2). Meanwhile, an AUC value for the ROC curve based on only age, sex and BMI was 0.68, which was better than that based on only the number of risk alleles (data now shown). The model incorporating age, sex and BMI as well as the number of risk alleles thus showed moderate power for the prediction of type 2 diabetes. The best

accuracy was 0.66 at the threshold between non-diabetic and diabetic status of 0.52 (non-diabetic status=0, diabetic status=1), for which the specificity and the sensitivity were 0.71 and 0.61, respectively.

DISCUSSION

By the validation of the results from our multistage GWAS, we detected only marginal associations of *EIF2AK4*, *KRT4* and *A2BP1* with type 2 diabetes in meta-analyses with two subject panels of Japanese or Chinese individuals. Relations of *KRT4* (keratin 4 gene) and *A2BP1* (ataxin-2-binding protein 1 gene, also known as *FOX1*) to glucose or lipid metabolism are unknown. Deletion of *EIF2AK4* (eukaryotic translation initiation factor 2 alpha kinase 4 gene, also known as *GCN2*) in mice resulted in liver steatosis during leucine deprivation as a result of unrepressed expression of lipogenic genes.³⁰ The functionally related gene, *EIF2AK3* (also known as *PERK* or *PEK*), has been shown to cause diabetes mellitus both in humans (Wolcott–Rallison syndrome, OMIM604032) and in rodent models.^{31,32} Taken together, *EIF2AK4* may be a good candidate for the diabetes susceptibility gene. The sample size required for a statistical power of 0.80 with equal numbers of cases and controls is 10 505 when the frequency of the risk allele, OR and type I error probability are assumed to be 0.20, 1.10 (the value for *EIF2AK4* in the meta-analysis in Table 1) and 0.05, respectively. Further studies of these genes in other Asian populations as well as in other ethnic groups are needed for confirmation of their association with type 2 diabetes. Given this uncertainty, we did not include these genes in the assessments of cumulative risk and gene–gene interaction.

Among tens of type 2 diabetes susceptibility genes identified by recent GWASs in Caucasians, the associations of six genes, that is, *TCF7L2*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8* and *HHEX*, have been replicated in Asian populations as well as in populations of European ancestry. A recent meta-analysis in Japanese subjects also supported the associations.¹² In this study, we performed replication study, and, on the basis of the results, we added five more genes, that is, *KCNJ11*, *PPARG*, *GCKR*, *KCNQ1* and *HNF1B*, for the cumulative risk assessment for type 2 diabetes. Thus, the SNPs of *HNF1B*, which were earlier associated with type 2 diabetes in Chinese as well as in Caucasians,²⁶ showed the association with the disease in the Japanese

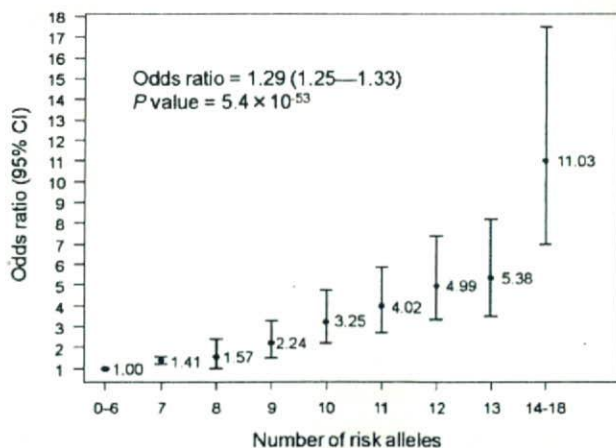


Figure 1 Odds ratios for subjects with different numbers of risk alleles for 11 susceptibility loci for type 2 diabetes. The cumulative effect of the 11 loci on type 2 diabetes was tested by counting the number of risk alleles associated with type 2 diabetes with a logistic regression model with adjustment for age, sex and BMI. The ORs for subjects with each number of risk alleles are expressed relative to individuals with 0–6 risk alleles.

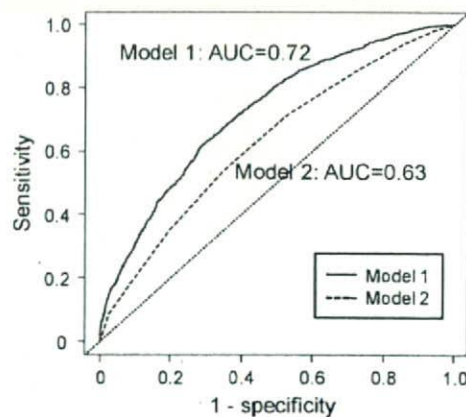


Figure 2 ROC curves for the prediction model on the basis of the number of risk alleles for 11 susceptibility loci for type 2 diabetes. The prediction model for type 2 diabetes was constructed using the logistic regression model, and ROC curves for the model were generated. In model 1, the number of risk alleles was used as an independent variable together with age, sex and BMI as covariates, whereas only the number of risk alleles was used as an independent variable in model 2.

population in this study. In addition, the C allele of rs780094 in *GCKR* was associated with increased risk of type 2 diabetes in this study, which is consistent with a recent study in Caucasians.³³ The associations of *KCNJ11* and *PPARG* with diabetes were marginal in this study; however, they were included for the prediction model, as the associations were replicated in some studies of Caucasians.

Our gene-gene interaction analysis showed no significant interaction for any of the 55 possible pairs of genes when corrected for multiple testing. When the significance level was set at 0.05, two pairs were judged to be significant. However, such gene-gene interactions were not supported from the functional point of view. A large-scale study may provide more convincing evidence for such interactions.

As no confirmatory evidence for gene-gene interaction was observed, we treated the 11 genes as independent variables in the prediction model. The addition of one risk allele was estimated to increase the odds by an average of 1.29 according to the multiplicative model. This value is similar to that (1.24) estimated for type 2 diabetes in Caucasians.¹⁷ Two earlier cumulative risk assessments for type 2 diabetes in Asian populations with relatively small numbers of associated loci yielded values of 1.17 and 1.24 for the fold increase in risk for each additional risk allele.^{11,34} In our prediction model for type 2 diabetes, the AUC for the ROC curve was lower than that in the earlier study¹⁷ based on 15 loci in Caucasians (0.72 and 0.86, respectively). However, the number of loci in our study (11 loci) was lower than that in the study for Caucasians. The inclusion of additional loci in our model should improve its ability to predict type 2 diabetes in Asian populations. Several reports of the prediction of type 2 diabetes using ~18 loci were recently described for populations of European ancestry.³⁵⁻³⁸ A prediction based on 18 loci gave an AUC value of 0.80 for the ROC curve,³⁵ whereas the corresponding values for a population-based prospective study were 0.68,³⁶ 0.615³⁷ and 0.75.³⁸ They concluded that genetic variations associated with diabetes had a small effect on the ability to predict the development of type 2 diabetes as compared with clinical characteristics alone. In fact, the AUC value (0.72) based on both the genetic variations and the clinical characteristics was slightly better than that based on only the clinical characteristics (0.68). We admit that the evidence of the association with diabetes is a little weaker for *KCNJ11* and *PPARG* in the Japanese population than for the other nine genes. If *KCNJ11* and *PPARG* were excluded from the analysis, the AUC for the ROC curve in the prediction model incorporating age, sex and BMI remained unchanged at 0.72, probably because of the relatively large effects of *KCNQ1* and *TCF7L2*.

Finally, our prediction model for type 2 diabetes achieved limited success even though it has some value. Given that GWASs for diabetes in Asians have not been as extensive as those in Caucasians, many risk loci for diabetes in Asians remain most likely to be undiscovered. Considering that the average increase in OR conferred by each additional risk allele was similar between Caucasians and Japanese, incorporation of data from additional risk loci is most likely to increase the predictive power.

ACKNOWLEDGEMENTS

We thank all the subjects who joined this project; Sumio Sugano and Shoji Tsuji for support and helpful discussion throughout the project; and Megumi Yamaoka-Sageshima for technical assistance. This work was supported by a grant from the Program for Promotion of Fundamental Studies in Health Sciences of the Pharmaceuticals and Medical Devices Agency (PMDA) of Japan; a grant from the National Institute of Biomedical Innovation (NIBIO) of Japan; grants from the Ministry of Health, Labour and Welfare of Japan; a Grant-in-Aid for Scientific Research on Priority Areas (C), 'Medical Genome Science

(Millennium Genome Project)', 'Applied Genomics', and 'Comprehensive Genomics', from the Ministry of Education, Culture, Sports, Science and Technology of Japan; and a grant from the Cooperative Link of Unique Science and Technology for Economy Revitalization (CLUSTER, Tokushima, Japan). The Hong Kong diabetes case-control study was supported by the Hong Kong Research Grants Committee Central Allocation Scheme CUHK 1/04C.

- Sladek, R., Rocheleau, G., Rung, J., Dina, C., Shen, L., Serre, D. et al. A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* **445**, 881-885 (2007).
- Saxena, R., Voight, B. F., Lyssenko, V., Burt, N. P., de Bakker, P. I., Chen, H. et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331-1336 (2007).
- Zeggini, E., Weedon, M. N., Lindgren, C. M., Frayling, T. M., Elliott, K. S., Lango, H. et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* **316**, 1336-1341 (2007).
- Scott, L. J., Mohlke, K. L., Bonnycastle, L. L., Willer, C. J., Li, Y., Duren, W. L. et al. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* **316**, 1341-1345 (2007).
- Steinthorsdottir, V., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Jonsdottir, T., Walters, G. B. et al. A variant in *CDKAL1* influences insulin response and risk of type 2 diabetes. *Nat. Genet.* **39**, 770-775 (2007).
- Miyake, K., Horikawa, Y., Hara, K., Yasuda, K., Osawa, H., Furuta, H. et al. Association of *TCF7L2* polymorphisms with susceptibility to type 2 diabetes in 4087 Japanese subjects. *J. Hum. Genet.* **53**, 174-180 (2008).
- Horikawa, Y., Miyake, K., Yasuda, K., Enya, M., Hirota, Y., Yamagata, K. et al. Replication of genome-wide association studies of type 2 diabetes susceptibility in Japan. *J. Clin. Endocrinol. Metab.* **93**, 3136-3141 (2008).
- Horikoshi, M., Hara, K., Ito, C., Shojima, N., Nagai, R., Ueki, K. et al. Variations in the *HHEX* gene are associated with increased risk of type 2 diabetes in the Japanese population. *Diabe* **50**, 2461-2466 (2007).
- Furukawa, Y., Shimada, T., Furuta, H., Matsuno, S., Kusuyama, A., Doi, A. et al. Polymorphisms in the *IDE-KIF11-HHEX* gene locus are reproducibly associated with type 2 diabetes in a Japanese population. *J. Clin. Endocrinol. Metab.* **93**, 310-314 (2008).
- Omori, S., Tanaka, Y., Takahashi, A., Hirose, H., Kashiwagi, A., Kaku, K. et al. Association of *CDKAL1*, *IGF2BP2*, *CDKN2A/B*, *HHEX*, *SLC30A8*, and *KCNJ11* with susceptibility to type 2 diabetes in a Japanese population. *Diabetes* **57**, 791-795 (2008).
- Ng, M. C., Park, K. S., Oh, B., Tam, C. H., Cho, Y. M., Shin, H. D. et al. Implication of genetic variants near *TCF7L2*, *SLC30A8*, *HHEX*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2* and *FTO* in type 2 diabetes and obesity in 6719 Asians. *Diabetes* **57**, 2226-2233 (2008).
- Tabara, Y., Osawa, H., Kawamoto, R., Onuma, H., Shimizu, I., Miki, T. et al. Replication study of candidate genes associated with type 2 diabetes based on genome-wide screening. *Diabetes* **58**, 493-498 (2009)10.2337/db07-1785.
- Grarup, N., Rose, C. S., Andersson, E. A., Andersen, G., Nielsen, A. L., Albrechtsen, A. et al. Studies of association of variants near the *HHEX*, *CDKN2A/B*, and *IGF2BP2* genes with type 2 diabetes and impaired insulin release in 10705 Danish subjects: validation and extension of genome-wide association studies. *Diabetes* **56**, 3105-3111 (2007).
- Cauchi, S., Proença, C., Choquet, H., Gaget, S., De Graeve, F., Marre, M. et al. 2008 Analysis of novel risk loci for type 2 diabetes in a general French population: the D.E.S.I.R. study. *J. Mol. Med.* **86**, 341-348 (2008).
- Yasuda, K., Miyake, K., Horikawa, Y., Hara, K., Osawa, H., Furuta, H. et al. Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. *Nat. Genet.* **40**, 1092-1097 (2008).
- Zeggini, E., Scott, L. J., Saxena, R., Voight, B. F., Marchini, J. L., Hu, T. et al. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat. Genet.* **40**, 638-645 (2008).
- Cauchi, S., Meyre, D., Durand, E., Proença, C., Marre, M., Hadjadj, S. et al. Post genome-wide association studies of novel genes associated with type 2 diabetes show gene-gene interaction and high predictive value. *PLoS ONE* **3**, e2031 (2008).
- Yang, X., So, W. Y., Kong, A. P., Ho, C. S., Lam, C. W., Stevens, R. J. et al. Development and validation of stroke risk equation for Hong Kong Chinese patients with type 2 diabetes: the Hong Kong Diabetes Registry. *Diabetes Care* **30**, 65-70 (2007).
- Ozaki, R., Qiao, Q., Wong, G. W., Chan, M. H., So, W. Y., Tong, P. C. et al. Overweight, family history of diabetes and attending schools of lower academic grading are independent predictors for metabolic syndrome in Hong Kong Chinese adolescents. *Arch. Dis. Child.* **92**, 224-228 (2007).
- Bannai, M., Higuchi, K., Akasaka, T., Furukawa, M., Yamaoka, M., Sato, K. et al. Single-nucleotide-polymorphism genotyping for whole-genome-amplified samples using automated fluorescence correlation spectroscopy. *Anal. Biochem.* **327**, 215-221 (2004).
- Sing, T., Sander, D., Beerewinkel, N. & Lengauer, T. ROC: visualizing classifier performance in R. *Bioinformatics* **21**, 3940-3941 (2005).
- Grant, S. F., Thorleifsson, G., Reynisdottir, I., Benediktsson, R., Manolescu, A., Sainz, J. et al. Variant of transcription factor 7-like 2 (*TCF7L2*) gene confers risk of type 2 diabetes. *Nat. Genet.* **38**, 320-323 (2006).