Previously Associated Type 2 Diabetes Variants May Interact With Physical Activity to Modify the Risk of Impaired Glucose Regulation and Type 2 Diabetes

A Study of 16,003 Swedish Adults

Ema C. Brito,[1] Valeriya Lyssenko,[2] Frida Renström,[1] Göran Berglund,[3] Peter M. Nilsson,[3] Leif Groop,[2] and Paul W. Franks[1,2]

OBJECTIVE—Recent advances in type 2 diabetes genetics have culminated in the discovery and confirmation of multiple risk variants. Two important and largely unanswered questions are whether this information can be used to identify individuals most susceptible to the adverse consequences of sedentary behavior and to predict their response to lifestyle intervention; such evidence would be mechanistically informative and provide a rationale for targeting genetically susceptible subgroups of the population.

RESEARCH DESIGN AND METHODS—Gene × physical activity interactions were assessed for 17 polymorphisms in a prospective population-based cohort of initially nondiabetic middle-aged adults. Outcomes were *1*) impaired glucose regulation (IGR) versus normal glucose regulation determined with either fasting or 2-h plasma glucose concentrations (n = 16,003), *2*) glucose intolerance (in mmol/l, n = 8,860), or *3*) incident type 2 diabetes (n = 2,063 events).

RESULTS—Tests of gene × physical activity interactions on IGR risk for 3 of the 17 polymorphisms were nominally statistically significant: *CDKN2A/B* rs10811661 ($P_{interaction}$ = 0.015), *HNF1B* rs4430796 ($P_{interaction}$ = 0.026), and *PPARG* rs1801282 ($P_{interaction}$ = 0.04). Consistent interactions were observed for the *CDKN2A/B* ($P_{interaction}$ = 0.013) and *HNF1B* ($P_{interaction}$ = 0.0009) variants on 2-h glucose concentrations. Where type 2 diabetes was the outcome, only one statistically significant interaction effect was observed, and this was for the *HNF1B* rs4430796 variant ($P_{interaction}$ = 0.0004). The interaction effects for *HNF1B* on IGR risk and incident diabetes remained significant after correction for multiple testing ($P_{interaction}$ = 0.015 and 0.0068, respectively).

CONCLUSIONS—Our observations suggest that the genetic predisposition to hyperglycemia is partially dependent on a person's lifestyle. ***Diabetes*** **58:1411–1418, 2009**

From the [1]Genetic Epidemiology and Clinical Research Group, Department of Public Health and Clinical Medicine, Section for Medicine, Umeå University Hospital, Umeå, Sweden; the [2]Department of Clinical Sciences-Diabetes and Endocrinology, Clinical Research Center, Malmö University Hospital, Lund University, Malmö, Sweden; and the [3]Department of Medicine, Malmö University Hospital, Lund University, Malmö, Sweden.

Corresponding author: Paul W. Franks, paul.franks@medicin.umu.se.

Received 21 November 2008 and accepted 8 March 2009.

Published ahead of print at http://diabetes.diabetesjournals.org on 26 March 2009. DOI: 10.2337/db08-1623.

Recent advances in high-throughput genotyping methods have facilitated the discovery and confirmation (1–7) of multiple common genetic risk factors for type 2 diabetes. The notion of using genetic information for disease prevention is predicated on the assumption that genetic risk can be offset with drug or lifestyle intervention. Thus, studies that explore this possibility are integral to the process of translating the results of genetic association studies into preventive practice. Furthermore, because interaction effects modify the extent to which genetic risk is conveyed, with the risk varying in magnitude across the spectrum of environmental exposure, information on gene × environment interactions may help improve the sensitivity and specificity of genetic prediction models (8).

In a recent report from the Diabetes Prevention Program, 10 of the previously associated type 2 diabetes risk polymorphisms were assessed in the context of a clinical trial of drug or lifestyle intervention for diabetes risk reduction (9). In that study, a single nucleotide polymorphism (SNP) at the cyclin-dependent kinase inhibitor 2A/B (*CDKN2A/B*) locus (rs10811661) was shown to modify the effects of lifestyle intervention on diabetes risk reduction, such that the lifestyle intervention slowed the progression to type 2 diabetes to a greater extent in those carrying the previously reported high-risk genotype at rs10811661 compared with those with the lower-risk genotypes.

In the current study, we aimed to determine whether the effects of 17 previously associated type 2 diabetes gene variants on the risk of impaired glucose regulation (IGR) or incident type 2 diabetes are modified by physical activity. The study was undertaken in an ethnically homogeneous prospective population-based cohort study of ~16,000 initially nondiabetic middle-aged adults from Sweden.

RESEARCH DESIGN AND METHODS

A preventive case-finding program called the Malmö Preventive Project was started in 1974 at the Department of Preventive Medicine, University Hospital, Malmö, Sweden. The aim was to screen large strata of the adult population in the southern Swedish city of Malmö to find individuals at high risk of developing chronic diseases who might benefit from preventive interventions. Participants were invited to participate in a general health screening that included a physical examination, a self-administered questionnaire on lifestyle habits, and fasting and postchallenge blood sampling for laboratory tests. Between 1974 and 1992, a total of 22,444 male and 10,902 female subjects attended the screening program, with an overall attendance rate of 71%. All participants gave written informed consent to take part in the study, which

was approved by the local research ethics committees of Lund University and Malmö University Hospital.

Follow-up for incident type 2 diabetes. A detailed description of the case ascertainment and other follow-up procedures is given elsewhere (10,11). Briefly, all known prevalent cases of type 2 diabetes were excluded from the cohort. Incident diabetes (n = 2,063 events) was ascertained using information collected from hospital records detailing a clinical diagnosis of type 2 diabetes or a fasting plasma glucose value >7.0 mmol/l at one of the Malmö Preventive Project follow-up examinations.

Clinical measurements. The clinical examination included measurements of height (m) and weight (kg) as previously described (11). Blood samples were drawn at 0, 40, and 120 min of an oral glucose tolerance test for measurements of blood glucose using a hexokinase method, the details of which have previously been described in detail (12). BMI was calculated as the weight in kilograms divided by height in meters squared (kg/m^2). IGR was defined as an elevated 2-h glucose concentration of 7.8–11.0 mmol/l (or an elevated fasting value of 5.6–6.9 mmol/l if the 2-h value was unavailable), whereas normal glucose regulation (NGR) was defined as a 2-h glucose concentration <7.8 mmol/l (or a fasting value <5.6 mmol/l if the 2-h value was unavailable).

Measurement of physical activity. Physical activity was assessed using a self-administered computer-based questionnaire. The current study spanned a number of years, and during that time the questions used to assess physical activity were changed. The different questions used throughout the survey are shown below. Therefore, we constructed a variable where individuals were classified as "physically active" if they had answered "yes" when responding to one of the following questions used to assess physical activity throughout the duration of the survey: *1*) Do you walk or cycle to and from work? *2*) Do you walk or cycle for recreation during weekdays? *3*) Do you walk or cycle for recreation during weekend days? *4*) Do you undertake at least 3 h per week of structured physical exercise? *5*) Do you walk to work or do yard work? or *6*) Do you perform light structured physical exercise each week? Conversely, those who answered "no" to these questions were classified as "physically inactive." In this report, the physical activity variable was associated with glucose concentrations and diabetes incidence (see the RESULTS section) in a manner consistent with prior studies using objective assessment methods (13), indicating that the questionnaire used here correctly classifies individuals into physically active and inactive subgroups.

Genotyping. All polymorphisms have previously been examined for association with incident type 2 diabetes in this cohort (10,14,15). Data on the association between *MTNR1B* (rs10830963) and fasting glucose also have previously been reported (15). A Plasmid Maxi Kit (Qiagen) was used to extract DNA from white blood cells. For the rs7903146, rs1801282, rs5219, rs7754840, and rs10811661 variants, genotyping was performed using the matrix-assisted laser desorption–ionization time-of-flight mass spectrometry method on a MassARRAY platform (Sequenom). For rs4430796, rs4402960, rs10010131, rs1111875, rs864745, rs12779790, rs7961581, rs7578597, rs4607103, rs10923931, and rs10830963, genotyping was performed using an allelic discrimination assay-by-design approach on an ABI 7900 (Applied Biosystems). The rs13266634 variant was genotyped using an allele-specific assay (KASPar; KBioscience). The genotyping success rate and accuracy (11% of samples were regenotyped using Sequenom) exceeded 95 and 98.7%, respectively. As previously described, the concordance rates for the vast majority of SNPs exceeded 99% (10,15). All variants fulfilled Hardy-Weinberg expectations ($P \geq 0.001$).

Statistical analysis. Statistical analyses were conducted using SAS software (version 9.1; SAS Institute, Cary, NC). Hardy-Weinberg equilibrium was assessed using Haploview version 4.0 (http://www.broad.mit.edu/mpg/haploview). Statistical power was calculated with Quanto (version 1.2.3) (16) using parameters described in the legend of Table 3. Logistic regression adjusted for age, sex, and BMI was used to determine the strength of the relationship between physical activity and IGR. Differences between means were calculated using independent sample, two-sided Student's t tests to establish whether age, BMI, and glucose levels differed between active and inactive individuals. A likelihood ratio test with 1 degree of freedom was used to calculate the difference for categorical traits. For models where glucose tolerance was expressed as a binary variable (i.e., NGR vs. IGR), logistic regression was used, including interaction terms for genotype × physical activity. Generalized linear models were used to test associations and interactions (also incorporating an interaction term for genotype × physical activity) for each of the gene variants and with 2-h glucose concentrations as the dependent variable. Cox proportional hazards regression models were used to test main genetic effects (genes and physical activity separately) and gene × physical activity interactions on type 2 diabetes incidence. A time-dependent interaction term was fitted to the model to ensure that the proportionality assumption of Cox models was fulfilled, which it was ($P_{\text{interaction}} > 0.05$). All regression models were adjusted for age, sex, and BMI and assumed an additive mode of genetic inheritance. Adjustments for multiple statistical comparisons across the 17 polymorphisms were made for each of the three dependent variables using the Holm procedure (17). This procedure requires that the probability statistics for each hypothesis test is placed in a rank-ordered list, with the highest P value appearing first. The total number of P values within this list represents the denominator. The least significant nominal P value is divided by 1, the second least significant by 2, and so on until all P values within the list have been corrected. This correction method was applied to each dependent variable separately.

TABLE 1
Participant characteristics stratified by level of physical activity (n = 16,003)

Variable	Physically inactive	Physically active	P
n	3,455	12,548	—
Sex (M/F)	2,287/1,168	8,115/4,433	0.097
Baseline age (years)	44.7 ± 7.3	45.7 ± 6.8	<0.0001
Baseline BMI (kg/m^2)	24.6 ± 3.7	24.2 ± 3.1	<0.0001
Baseline fasting glucose (mmol/l)	4.87 ± 0.48	4.81 ± 0.49	<0.0001
Baseline 2-h glucose (mmol/l)*	5.67 ± 1.54	5.64 ± 1.44	0.480
Baseline glucose regulation (NGR/IGR)	2,585/870	9,601/2,947	0.038
Developed diabetes (no/yes)	2,958/497	10,982/1,566	0.003

Data are N or means ± SD. Differences between means were calculated using a two-sided, independent sample Student's t test. Differences between proportions were tested using a likelihood ratio test with 1 degree of freedom. *Available in subsamples of n = 1,647 (physically inactive) and n = 7,217 (physically active).

RESULTS

Participant characteristics for 16,003 initially nondiabetic middle-aged whites are shown in Table 1. To indirectly quantify the validity of the physical activity assessment instrument, we tested the association between physical activity and IGR or diabetes risk after adjustment for age, sex, and BMI. In these models, physical inactivity conveyed a 1.15-fold increased risk (95% CI 1.05–1.26, P = 0.002) of IGR and a 1.41-fold increased risk (1.29–1.57, $P <$ 0.0001) of type 2 diabetes.

The following section describes the results for the models focusing on main genetic effects and gene × physical activity interactions. Unless otherwise stated, all P values are unadjusted for multiple comparisons.

Genetic association models (main effects). The associations between each of the polymorphisms with IGR risk and 2-h glucose concentrations are shown in Table 2, ranked by the P value for the test of gene × physical activity interaction. *SLC30A8* (rs13266634), *TCF7L2* (rs7903146), *CDKAL1* (rs7754840), *NOTCH2* (rs10923931), *KCNJ11* (rs5219), *IGFBP2* (rs4402960), *JAZF1* (rs864745), *HHEX* (rs1111875), *MTNR1B* (rs10830963), and *TSPAN8* (rs7961581) all showed evidence for association with IGR risk and/or 2-h glucose concentrations in directions consistent with previous reports. No statistical association was observed between the *CDKN2A/B* rs10811661 variant and 2-h glucose (β = 0.02 mmol/l per copy of the major allele, P = 0.43) or the risk of having IGR (vs. NGR; odds ratio [OR] 1.08 per copy of the major allele, 95% CI 0.99–1.15, P = 0.08). Similarly, no statistical association was observed for the *HNF1B* rs4430796 variant and 2-h glucose (β = 0.02 mmol/l per copy of the minor allele, P = 0.65) or IGR risk (OR 1.02 per copy of the minor allele, 95% CI 0.97–1.08, P = 0.36). The associations between each of

TABLE 2
Tests of association and gene × physical activity interaction for 17 polymorphisms on IGR risk (vs. NGR) and 2-h glucose levels (mmol/l)

	Main effects (OR or β-coefficient) and *P* values (for main effects and interactions)							
	Overall		$P_{interaction}$		Physically inactive		Physically active	
Nearest gene (variant)	IGR risk (OR, *P* value)	2-h glucose (β-coefficient, *P* value)	IGR risk	2-h glucose	IGR risk (OR, *P* value)	2-h glucose (β-coefficient, *P* value)	IGR risk (OR, *P* value)	2-h glucose (β-coefficient, *P* value)
CDKN2A/B (rs10811661)*	1.08, 0.084	0.02, 0.432	0.015	0.013	0.91, 0.197	−0.12, 0.064	1.12, 0.0075	0.06, 0.070
HNF1B (rs4430796)	1.02, 0.361†	0.02, 0.651†	0.026	0.0009	0.92, 0.126	−0.13, 0.005	1.06, 0.066	0.04, 0.056
PPARG (rs1801282)	1.01, 0.744	0.01, 0.790	0.041	0.776	0.88, 0.097	−0.01, 0.877	1.05, 0.221	0.01, 0.717
SLC30A8 (rs13266634)	1.04, 0.169	0.02, 0.055	0.134	0.560	0.96, 0.540	0.01, 0.798	1.06, 0.055	0.05, 0.047
WFS1 (rs10010131)	1.03, 0.231	−0.01, 0.651	0.140	0.082	0.96, 0.456	−0.08, 0.087	1.05, 0.082	0.01, 0.748
TCF7L2 (rs7903146)	1.10, 0.001	0.06, 0.012	0.141	0.600	1.01, 0.858	0.03, 0.538	1.13, 0.0005	0.06, 0.012
ADAMTS9 (rs4607103)	1.03, 0.351	−0.02, 0.420	0.155	0.784	0.94, 0.420	−0.01, 0.886	1.05, 0.138	−0.02, 0.405
CDKAL1 (rs7754840)	1.06, 0.024	0.04, 0.049	0.222	0.715	1.14, 0.033	0.02, 0.659	1.04, 0.165	0.05, 0.051
CAMK1D (rs12779790)	1.03, 0.306	0.01, 0.633	0.238	0.753	0.96, 0.565	−0.00, 0.978	1.05, 0.143	0.02, 0.565
NOTCH2 (rs10923931)	1.11, 0.013	0.04, 0.190	0.313	0.281	1.20, 0.042	0.12, 0.129	1.09, 0.092	0.03, 0.471
KCNJ11 (rs5219)	1.04, 0.144	0.04, 0.047	0.391	0.974	1.09, 0.148	0.04, 0.391	1.03, 0.388	0.04, 0.072
THADA (rs7578597)*	1.04, 0.357	−0.03, 0.511	0.472	0.484	0.98, 0.858	0.02, 0.774	1.06, 0.249	−0.03, 0.358
IGFBP2 (rs4402960)	1.10, 0.002	0.02, 0.012	0.627	0.871	1.12, 0.057	0.06, 0.289	1.09, 0.014	0.06, 0.025
JAZF1 (rs864745)*	1.04, 0.154	0.05, 0.008	0.830	0.435	1.05, 0.384	0.08, 0.078	1.04, 0.253	0.04, 0.034
HHEX (rs1111875)	1.08, 0.005	0.05, 0.024	0.862	0.802	1.08, 0.220	0.03, 0.485	1.09, 0.010	0.05, 0.031
MTNR1B (rs10830963)	1.18, <0.0001	0.06, 0.009	0.912	0.300	1.20, 0.004	0.10, 0.067	1.18, <0.0001	0.05, 0.048
TSPAN8 (rs7961581)	1.04, 0.141	0.02, 0.053	0.989	0.857	1.04, 0.502	0.06, 0.286	1.04, 0.210	0.04, 0.091

P values are unadjusted for multiple statistical comparisons. Effect estimates are expressed as ORs (IGR vs. NGR) or β-coefficient (mmol/l of 2-h glucose) per copy of the risk allele at each locus. Data are adjusted for age, sex, and BMI and are ranked by *P* value for the test of gene × physical activity interaction on IGR risk. *The major allele is shown as the risk allele. In all other cases the minor allele is the risk allele. †This result has previously been reported (14). Fasting glucose and type 2 diabetes data also have previously been reported in this cohort for *MTNR1B* (15). Associations with type 2 diabetes have been reported for all other SNPs (10).

TABLE 3
Statistical power calculations to detect gene × physical activity interactions on IGR or type 2 diabetes risk

Nearest gene	Variant	Alleles	MAF	Power for IGR (%)	Power for type 2 diabetes (%)
CDKN2A/B	rs10811661	T/C	0.16	0.49–0.99	0.30–0.94
HNF1B	rs4430796	G/A	0.46	0.84–0.99	0.63–0.99
PPARG	rs1801282	C/G	0.14	0.44–0.99	0.29–0.92
SLC30A8	rs13266634	C/T	0.32	0.78–0.99	0.58–0.99
WFS1	rs10010131	G/A	0.43	0.86–0.99	0.63–0.99
TCF7L2	rs7903146	C/T	0.26	0.72–0.99	0.53–0.99
ADAMTS9	rs4607103	C/T	0.23	0.67–0.99	0.45–0.99
CDKAL1	rs7754840	G/C	0.31	0.78–0.99	0.57–0.99
CAMK1D	rs12779790	A/G	0.19	0.58–0.99	0.38–0.98
NOTCH2	rs10923931	G/T	0.10	0.31–0.95	0.20–0.84
KCNJ11	rs5219	G/A	0.38	0.82–0.99	0.61–0.99
THADA	rs7578597	T/C	0.10	0.30–0.95	0.16–0.78
IGFBP2	rs4402960	G/T	0.30	0.77–0.99	0.56–0.99
JAZF1	rs864745	A/G	0.49	0.85–0.99	0.64–0.99
HHEX	rs1111875	G/A	0.41	0.83–0.99	0.62–0.99
MTNR1B	rs10830963	C/G	0.28	0.75–0.99	0.53–0.99
TSPAN8	rs7961581	T/C	0.26	0.71–0.99	0.51–0.99

Environmental risk (OR) for IGR = 1.15 and for type 2 diabetes = 1.41 (as shown in the RESULTS section). Genetic risk (OR) is taken from Table 2 and Lyssenko et al. (10,15). Background population disease risk was set at 5%. Detectable interaction effect sizes (R_{GE}) are 1.25–1.50 per copy of the minor allele. Two-sided $P_{interaction} = 0.00294$ (i.e., $P_{interaction} = 0.05$ corrected for 17 tests). Power is for log-additive genetic models. Alleles are coded as major/minor and are from the National Center for Biotechnology Information Build 36.3 (HapMap-CEU). MAF, minor allele frequency.

the gene variants and type 2 diabetes have previously been described in detail (10,15).

Gene × physical activity interaction models. Statistical power calculations are shown in Table 3 for all variants. As shown in Table 2, no statistical evidence of interaction emerged for the majority of the polymorphisms tested. The remaining section focuses on the variants for which the uncorrected interaction terms were statistically significant ($P_{interaction} < 0.05$) for at least one of the three outcomes of interest in this report.

***HNF1B* (rs4430796).** The gene × physical activity interaction terms for *HNF1B* (rs4430796) on IGR risk and on 2-h glucose concentrations were both statistically significant ($P_{interaction} = 0.026$ and 0.0009, respectively), and the latter remained significant after correction for multiple hypothesis testing ($P_{interaction} = 0.015$). Although not statistically significant, the minor A allele at *HNF1B* rs4430796 tended to be associated with lower risk of IGR (OR 0.92 per allele, 95% CI 0.82–1.03, $P = 0.13$) in physically inactive individuals ($n = 3,335$) and with increased risk (1.06, 1.00–1.12, $P = 0.066$) in physically active individuals ($n = 12,015$). In a concordant manner, the minor A allele was associated with lower 2-h glucose levels ($\beta = -0.13$ mmol/l per allele, $P = 0.005$) in physically inactive individuals ($n = 1,619$), with a contrasting effect (0.04, $P = 0.056$) in physically active individuals ($n = 6,981$) (Fig. 1). The Cox proportional hazards regression model testing the gene × baseline physical activity interaction term on type 2 diabetes as the outcome yielded an uncorrected P value of 0.0004 (corrected $P_{interaction} = 0.0068$). The nature of the interaction reflected those observed in the glucose regulation models, where the minor allele was associated with decreased risk of diabetes in sedentary individuals (hazard rate ratio [HRR] = 0.85; 95% CI 0.74–0.96; $P = 0.011$), with a contrasting effect in physically active individuals (HRR = 1.10; 95% CI 1.03–1.18; $P = 0.007$). To illustrate this interaction, cumulative incidence curves were plotted stratified by genotype and level of physical activity (see Fig. 2).

***CDKN2A/B* (rs10811661).** The interaction terms for the *CDKN2A/B* rs10811661 polymorphism and physical activity, with IGR risk or 2-h glucose concentrations as outcomes, were both nominally statistically significant ($P_{interaction} = 0.015$ and 0.013, respectively). Though not statistically significant, the minor allele at rs10811661 tended to be associated with higher risk of IGR (OR 1.10 per allele, 95% CI 0.95–1.28, $P = 0.20$) in physically inactive individuals ($n = 3,468$), but it was protective of IGR risk (0.89, 0.82–0.97, $P = 0.0075$) in active individuals ($n = 12,525$). Similarly, in physically inactive individuals ($n = 1,618$), the rs10811661 minor allele tended to be associated with elevated 2-h glucose concentrations ($\beta = 0.12$ mmol/l per allele, $P = 0.064$), whereas in physically active individuals ($n = 6,680$), a contrasting effect was observed (-0.06, $P = 0.070$), although neither was statistically significant. No statistical evidence of interaction was observed where type 2 diabetes was the outcome ($P_{interaction} = 0.57$).

***PPARG* (rs1801282).** The interaction between the *PPARG* rs1801282 polymorphism and physical activity on

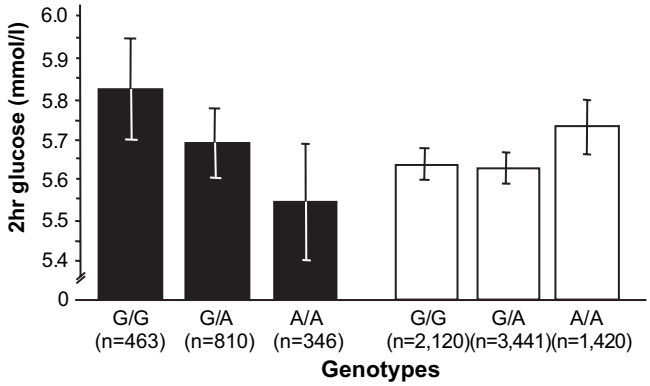

FIG. 1. Interaction between the *HNF1B* rs4430796 variant and physical activity on 2-h glucose levels in 8,600 Swedish middle-aged men and women. ■, physically inactive; □, physically active. Data are means adjusted for age and sex. Error bars are 95% CIs. $P_{interaction} = 0.0009$.

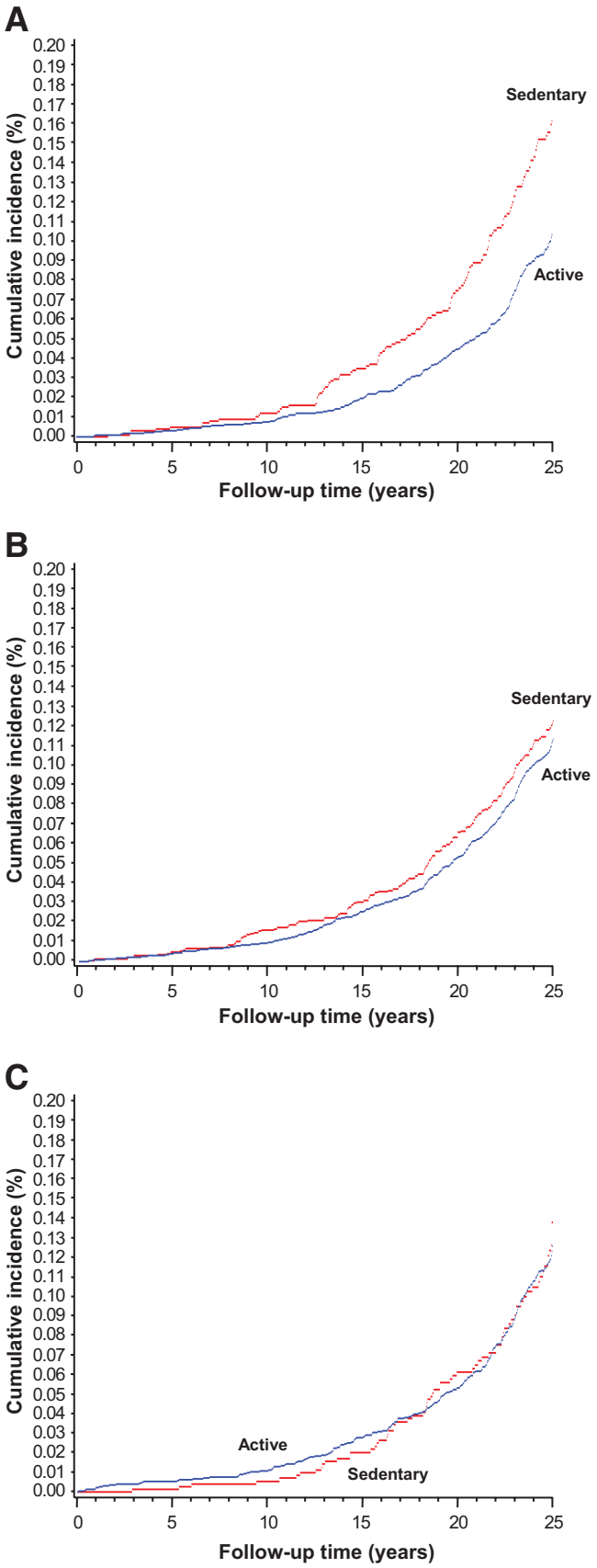

FIG. 2. Type 2 diabetes cumulative incidence plots stratified by level of physical activity and genotype at the *HNF1B* rs4430796 locus. *A*: G/G. *B*: G/A. *C*: A/A. Cumulative incidences are calculated within each genotype group. Follow-up is truncated at the median duration for the cohort (25 years).

IGR risk was nominally statistically significant ($P_{interaction}$ = 0.04). In inactive individuals, the minor allele appeared to be protective of IGR (OR 0.88 per allele, 95% CI 0.75–1.02, P = 0.097), whereas this effect was reversed in active individuals (1.05, 0.97–1.15, P = 0.22), albeit neither stratified effect was statistically significant. No evidence of statistical interaction was observed for 2-h glucose concentrations ($P_{interaction}$ = 0.78) or incident diabetes ($P_{interaction}$ = 0.67).

Tests of gene × BMI interaction on IGR risk. It is possible that the interactions reported above might be driven by interactions between the gene variants and BMI because both physical activity and BMI are related (13). Therefore, we tested age- and sex-adjusted BMI interaction terms for each SNP with the dependent variable being IGR vs. NGR. In these models, only one term, for *ADAMTS9* rs4607103, was nominally statistically significant ($P_{interaction}$ = 0.024). The BMI interaction terms for the *CDKN2A/B* rs10811661 ($P_{interaction}$ = 0.973), *HNF1B* rs4430796 ($P_{interaction}$ = 0.763), and *PPARG* rs1801282 ($P_{interaction}$ = 0.724) variants did not approach statistical significance. The inclusion or exclusion of BMI as a covariate in the SNP × physical activity interaction models made no material difference to the interaction results, and no novel interaction effects emerged.

DISCUSSION

A growing number of gene variants have been convincingly shown to influence type 2 diabetes risk (1–3,18,19). In the current study, the risk of glucose intolerance and type 2 diabetes conveyed by one of these variants (*HNF1B* rs4430796) appeared to be offset in people reporting physically active lifestyles. These putative interactions were observed in ~16,000 people, of whom >2,000 developed type 2 diabetes during a median follow-up period of 24.5 years. Consistent interaction effects emerged in the subgroup of nearly 9,000 individuals in whom 2-h glucose concentrations in response to an oral glucose tolerance test had been recorded. The uncorrected interaction effect for a second variant at *CDKN2A/B* (rs10811661) was nominally statistically significant when baseline IGR or 2-h glucose was the outcome. The nature of these effects is consistent with those reported in a recent clinical trial of intensive lifestyle intervention (9). However, when type 2 diabetes was the outcome, the gene × physical activity interaction term was not statistically significant in this cohort. A borderline uncorrected statistical interaction on IGR risk for a third variant at the *PPARG* gene (rs1801282) was also observed, but this was not apparent within the subcohort in which 2-h glucose concentrations were available, nor was an interaction evident when type 2 diabetes was the outcome. The absence of a statistical interaction in the cohort with 2-h glucose concentrations may be a consequence of lower statistical power owing to the smaller sample size, but inadequate power is unlikely to explain the lack of effect where diabetes was the outcome. For this locus, numerous prior reports of gene-lifestyle interaction exist, although most have focused on gene-nutrient interaction effects (20).

HNF1B (also known as *TCF2*) encodes a homeobox transcription factor that controls cell proliferation and differentiation in the kidney, pancreas, liver, and genital tract tissues. Rare monogenic forms of maturity-onset diabetes of the young (MODY5) are caused by *HNF1B* mutations (21). In the initial reports of association be-

tween the *HNF1B* rs4430796 polymorphism and type 2 diabetes (22), the major G allele was associated with decreased risk, as observed in physically active individuals in the current study. However, any beneficial effect of this allele is essentially lost, and possibly reversed, in the physically inactive individuals studied here. From Fig. 2, it is apparent that the initially strong inverse association between physical activity and diabetes incidence diminishes in magnitude with each additional copy of the minor *HNF1B* allele. This result suggests that the protective effects of physical activity on diabetes risk may be attenuated in carriers of the rs4430796 minor allele. Interestingly, the same variant was associated in the opposite direction with prostate cancer risk in the initial report (22). Thus, it would be important to determine whether gene × physical activity interactions are also relevant for that disease and, if so, how lifestyle intervention influences prostate cancer risk in carriers of the different rs4430796 genotypes. No prior studies to our knowledge have reported evidence of gene × physical activity interactions at the *HNF1B* rs4430796 locus. Thus, unlike the interactions for *CDKN2A/B* and *PPARG*, the prior probability that the *HNF1B* interaction tests would be statistically significant was low; therefore, that these tests remained statistically significant after correction for multiple testing strengthens the credibility of our observations.

The *CDKN2A* and *CDKN2B* genes localize to neighboring genomic regions. Both genes function as tumor supressors via inhibition of cyclin-dependent kinases (23). The *CDKN2A/B* genes are highly expressed in adipocytes and pancreatic islets (4). The rs10811661 polymorphism, which maps to a region between *CDKN2A* and *CDKN2B*, was initially identified through genome-wide association studies of type 2 diabetes risk (1,2,18,19). Several (7,24–27), but not all (28), subsequent studies found strong support for an association with type 2 diabetes at this locus in a range of ethnic groups. Consistent with our observations in physically active individuals, the major T allele was associated with increased risk in prior reports (1,7,18,19,24–27). In one of the follow-up studies (7), the authors reported moderately elevated 2-h glucose concentrations in nondiabetic T allele carriers. The validity of the *CDKN2A/B* × physical activity interaction observed here is supportive of a recent report from the Diabetes Prevention Program (9), where the effect of the rs10811661 polymorphism on diabetes risk was abolished with intensive lifestyle intervention. These data tend to suggest that whereas they are protected from IGR or type 2 diabetes when physically active, carriers of the minor allele at rs10811661 are, relative to the major allele homozygotes, at increased risk of these conditions when leading sedentary lifestyles.

The third polymorphism for which we observed weak evidence of statistical interaction is *PPARG* rs1801282 (encoding a proline-to-alanine amino acid substitution at codon 12). Although the data supporting gene × physical activity interactions at the rs1801282 locus are tenuous, the variant is a strong biological and statistical candidate for gene-nutrient (i.e., fatty acids) interactions (20,29). Hence, because dietary patterns and physical activity tend to coalesce within free-living populations, our observation of gene × physical activity interaction may be attributable to dietary correlates of physical activity, rather than direct biological interactions between *PPARG* and physical activity—unfortunately, measures of dietary fat intake were unavailable for these analyses. Nonetheless, it is important to stress that the statistical evidence of the interaction between *PPARG* rs1801282 and physical activity is of nominal statistical significance for IGR risk, and no interaction was observed on 2-h glucose concentrations or type 2 diabetes incidence. Therefore, the tentative interaction for IGR risk may be attributable to chance.

This is the first study to our knowledge to assess gene-lifestyle interactions across a wide range of previously associated type 2 diabetes gene variants. It is also the largest study of gene-lifestyle interaction on prediabetic and diabetic traits to date and, hence, one of the most well powered to test related hypotheses. To illustrate this point, we calculated the power to detect statistically significant interactions for the loci of interest (Table 3). Although it is very difficult at this stage to estimate the magnitude of interaction effects because so few well-designed interaction studies exist from which to garner this information, a conservative estimate is that they might range from 1.25 to 1.50. As Table 3 shows for IGR and type 2 diabetes as outcomes, our study was well powered to detect interactions of $\geq$1.50 per allele.

A limitation of this report, as with all existing large-scale epidemiological investigations, is that physical activity was assessed using questionnaires, which are prone to error and bias. This will have detrimentally affected power to detect interactions owing to a degree of misclassification of the environmental exposure. However, our measure of physical activity is inversely correlated with IGR and type 2 diabetes risk in this cohort, which is consistent with the direction of association reported in clinical trials involving exercise as part of a program of intensive lifestyle intervention (30,31) or epidemiological studies where physical activity was objectively assessed (13), instilling confidence that the questionnaire correctly classified the behavior of physical activity. Moreover, we purposefully studied people free from diabetes at baseline to minimize the extent to which changes in behavior and cognition attributable to disease labeling and treatment might bias the reporting of physical activity levels by study participants. In the future, large studies may emerge where physical activity has been objectively assessed. In this event, it may be possible to model genetic interaction terms where the different components of physical activity are examined. This is an important objective in furthering our understanding of how behavioral and genetic predispositions interact to influence disease risk, but this is beyond the scope of the current study.

Given that our study is likely to be adequately powered to detect realistic interaction effects, the absence of statistical interactions for the majority of variants examined here supports the idea that the most strongly associated SNPs derived from existing type 2 diabetes genome-wide association studies do not, by and large, exert their effects through interactions with physical activity. To some extent, this is unsurprising because existing genome-wide association studies are insensitive to detecting variants with heterogeneous effects, owing to the strict significance thresholds used to determine genome-wide statistical significance. Thus, the most strongly associated variants in genome-wide association studies are likely to be those with the least heterogeneous effects, which by definition are likely to exclude many variants that exert their effects through interactions with environmental (or other genetic) exposures.

In summary, in this study we examined putative gene × physical activity interactions for 17 confirmed type 2

diabetes risk variants on measures of glucose intolerance and incident type 2 diabetes. Our findings, in combination with other previously published observations (9), suggest that impairments in glucose regulation caused by variants at the *HNF1B* and *CDKN2A/B* loci may be attenuated by physical activity. In the event that these interactions are independently replicated, they may shed light of the mechanisms through which physical inactivity and genetic variants influence the risk of type 2 diabetes; whether this information will prove useful for the targeted prevention of type 2 diabetes remains to be determined.

ACKNOWLEDGMENTS

This study was funded in part by grants from the Påhlssons Foundation, the Swedish Heart-Lung Foundation, and the Swedish Diabetes Association. E.C.B. was supported by a studentship from the Portuguese Foundation for Science and Technology. P.W.F. was supported in part by funding from Västerbotten's Health Authority (Regional Authority for Clinical Eduction and Research [ALF] strategic appointment 2006-2009). Some of this work was conducted by P.W.F. while a Visiting Scientist at Harvard School of Public Health and the Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School (supported by the Swedish Research Council and the Royal Physiological Society of Lund).

This study was also funded by Novo Nordisk. No other potential conflicts of interest relevant to this article were reported.

We thank those who participated in this study. We are also grateful for the technical assistance of the staff of the Clinical Research Center in Malmö for genotyping assistance and the statistical advice of Kathleen Jablonski.

REFERENCES

1. Saxena R, Voight BF, Lyssenko V, Burtt NP, de Bakker PI, Chen H, Roix JJ, Kathiresan S, Hirschhorn JN, Daly MJ, Hughes TE, Groop L, Altshuler D, Almgren P, Florez JC, Meyer J, Ardlie K, Bengtsson Bostrom K, Isomaa B, Lettre G, Lindblad U, Lyon HN, Melander O, Newton-Cheh C, Nilsson P, Orho-Melander M, Rastam L, Speliotes EK, Taskinen MR, Tuomi T, Guiducci C, Berglund A, Carlson J, Gianniny L, Hackett R, Hall L, Holmkvist J, Laurila E, Sjogren M, Sterner M, Surti A, Svensson M, Svensson M, Tewhey R, Blumenstiel B, Parkin M, Defelice M, Barry R, Brodeur W, Camarata J, Chia N, Fava M, Gibbons J, Handsaker B, Healy C, Nguyen K, Gates C, Sougnez C, Gage D, Nizzari M, Gabriel SB, Chirn GW, Ma Q, Parikh H, Richardson D, Ricke D, Purcell S. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. Science 2007;316:1331–1336
2. Scott LJ, Mohlke KL, Bonnycastle LL, Willer CJ, Li Y, Duren WL, Erdos MR, Stringham HM, Chines PS, Jackson AU, Prokunina-Olsson L, Ding CJ, Swift AJ, Narisu N, Hu T, Pruim R, Xiao R, Li XY, Conneely KN, Riebow NL, Sprau AG, Tong M, White PP, Hetrick KN, Barnhart MW, Bark CW, Goldstein JL, Watkins L, Xiang F, Saramies J, Buchanan TA, Watanabe RM, Valle TT, Kinnunen L, Abecasis GR, Pugh EW, Doheny KF, Bergman RN, Tuomilehto J, Collins FS, Boehnke M. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. Science 2007;316:1341–1345
3. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat Genet 2008;40:638–645
4. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, Timpson NJ, Perry JR, Rayner NW, Freathy RM, Barrett JC, Shields B, Morris AP, Ellard S, Groves CJ, Harries LW, Marchini JL, Owen KR, Knight B, Cardon LR, Walker M, Hitman GA, Morris AD, Doney AS; Wellcome Trust Case Control Consortium (WTCCC), McCarthy MI, Hattersley AT. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. Science 2007;316:1336–1341
5. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A, Stefansson K. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. Nat Genet 2006;38:320–323
6. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, Daly A, Lango H, Frayling TM, Neumann RJ, Sherva R, Blech I, Pharoah PD, Palmer CN, Kimber C, Tavendale R, Morris AD, McCarthy MI, Walker M, Hitman G, Glaser B, Permutt MA, Hattersley AT, Wareham NJ, Barroso I. Common variants in WFS1 confer risk of type 2 diabetes. Nat Genet 2007;39:951–953
7. Grarup N, Rose CS, Andersson EA, Andersen G, Nielsen AL, Albrechtsen A, Clausen JO, Rasmussen SS, Jorgensen T, Sandbaek A, Lauritzen T, Schmitz O, Hansen T, Pedersen O. Studies of association of variants near the *HHEX*, *CDKN2A/B*, and *IGF2BP2* genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects: validation and extension of genome-wide association studies. Diabetes 2007;56:3105–3111
8. Lu Q, Elston RC. Using the optimal receiver operating characteristic curve to design a predictive genetic test, exemplified with type 2 diabetes. Am J Hum Genet 2008;82:641–651
9. Moore AF, Jablonski KA, McAteer JB, Saxena R, Pollin TI, Franks PW, Hanson RL, Shuldiner AR, Knowler WC, Altshuler D, Florez JC. Extension of type 2 diabetes genome-wide association scan results in the Diabetes Prevention Program. Diabetes 2008;57:2503–2510
10. Lyssenko V, Jonsson A, Almgren P, Pulizzi N, Isomaa B, Tuomi T, Berglund G, Altshuler D, Nilsson P, Groop L. Clinical risk factors, DNA variants, and the development of type 2 diabetes. N Engl J Med 2008;359:2220–2232
11. Berglund G, Nilsson P, Eriksson KF, Nilsson JA, Hedblad B, Kristenson H, Lindgarde F. Long-term outcome of the Malmo preventive project: mortality and cardiovascular morbidity. J Intern Med 2000;247:19–29
12. Blitzen P-O, Melander A, Schersten B. How to screen for diabetes. Acta Endocrinol 1984;105:37–41
13. Franks PW, Ekelund U, Brage S, Wong MY, Wareham NJ. Does the association of habitual physical activity with the metabolic syndrome differ by level of cardiorespiratory fitness? Diabetes Care 2004;27:1187–1193
14. Holmkvist J, Almgren P, Lyssenko V, Lindgren CM, Eriksson KF, Isomaa B, Tuomi T, Nilsson P, Groop L. Common variants in maturity-onset diabetes of the young genes and future risk of type 2 diabetes. Diabetes 2008;57:1738–1744
15. Lyssenko V, Nagorny CL, Erdos MR, Wierup N, Jonsson A, Spegel P, Bugliani M, Saxena R, Fex M, Pulizzi N, Isomaa B, Tuomi T, Nilsson P, Kuusisto J, Tuomilehto J, Boehnke M, Altshuler D, Sundler F, Eriksson JG, Jackson AU, Laakso M, Marchetti P, Watanabe RM, Mulder H, Groop L. Common variant in MTNR1B associated with increased risk of type 2 diabetes and impaired early insulin secretion. Nat Genet 2009;41:82–88
16. Gauderman WJ. Sample size requirements for matched case-control studies of gene-environment interaction. Stat Med 2002;21:35–50
17. Holm S. A simple sequentially rejective Bonferroni test procedure. Scand J Stat 1979;6:65–70
18. Sladek R, Rocheleau G, Rung J, Dina C, Shen L, Serre D, Boutin P, Vincent D, Belisle A, Hadjadj S, Balkau B, Heude B, Charpentier G, Hudson TJ, Montpetit A, Pshezhetsky AV, Prentki M, Posner BI, Balding DJ, Meyre D, Polychronakos C, Froguel P. A genome-wide association study identifies novel risk loci for type 2 diabetes. Nature 2007;445:881–885
19. The Wellcome Trust Case-Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678
20. Franks PW, Mesa JL, Harding AH, Wareham NJ. Gene-lifestyle interaction on risk of type 2 diabetes. Nutr Metab Cardiovasc Dis 2007;17:104–124
21. Bellanne-Chantelot C, Chauveau D, Gautier JF, Dubois-Laforgue D, Clauin S, Beaufils S, Wilhelm JM, Boitard C, Noel LH, Velho G, Timsit J. Clinical

spectrum associated with hepatocyte nuclear factor-1beta mutations. Ann Intern Med 2004;140:510–517
22. Gudmundsson J, Sulem P, Steinthorsdottir V, Bergthorsson JT, Thorleifsson G, Manolescu A, Rafnar T, Gudbjartsson D, Agnarsson BA, Baker A, Sigurdsson A, Benediktsdottir KR, Jakobsdottir M, Blondal T, Stacey SN, Helgason A, Gunnarsdottir S, Olafsdottir A, Kristinsson KT, Birgisdottir B, Ghosh S, Thorlacius S, Magnusdottir D, Stefansdottir G, Kristjansson K, Bagger Y, Wilensky RL, Reilly MP, Morris AD, Kimber CH, Adeyemo A, Chen Y, Zhou J, So WY, Tong PC, Ng MC, Hansen T, Andersen G, Borch-Johnsen K, Jorgensen T, Tres A, Fuertes F, Ruiz-Echarri M, Asin L, Saez B, van Boven E, Klaver S, Swinkels DW, Aben KK, Graif T, Cashy J, Suarez BK, van Vierssen Trip O, Frigge ML, Ober C, Hofker MH, Wijmenga C, Christiansen C, Rader DJ, Palmer CN, Rotimi C, Chan JC, Pedersen O, Sigurdsson G, Benediktsson R, Jonsson E, Einarsson GV, Mayordomo JI, Catalona WJ, Kiemeney LA, Barkardottir RB, Gulcher JR, Thorsteinsdottir U, Kong A, Stefansson K. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. Nat Genet 2007;39:977–983
23. Gil J, Peters G. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: all for one or one for all. Nat Rev Mol Cell Biol 2006;7:667–677
24. Kirchhoff K, Machicao F, Haupt A, Schafer SA, Tschritter O, Staiger H, Stefan N, Haring HU, Fritsche A. Polymorphisms in the TCF7L2, CDKAL1 and SLC30A8 genes are associated with impaired proinsulin conversion. Diabetologia 2008;51:597–601
25. Ng MC, Park KS, Oh B, Tam CH, Cho YM, Shin HD, Lam VK, Ma RC, So WY, Cho YS, Kim HL, Lee HK, Chan JC, Cho NH. Implication of genetic variants near *TCF7L2*, *SLC30A8*, *HHEX*, *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, and *FTO* in type 2 diabetes and obesity in 6,719 Asians. Diabetes 2008;57:2226–2233
26. Omori S, Tanaka Y, Takahashi A, Hirose H, Kashiwagi A, Kaku K, Kawamori R, Nakamura Y, Maeda S. Association of *CDKAL1*, *IGF2BP2*, *CDKN2A/B*, *HHEX*, *SLC30A8*, and *KCNJ11* with susceptibility to type 2 diabetes in a Japanese population. Diabetes 2008;57:791–795
27. Wu Y, Li H, Loos RJ, Yu Z, Ye X, Chen L, Pan A, Hu FB, Lin X. Common variants in *CDKAL1*, *CDKN2A/B*, *IGF2BP2*, *SLC30A8*, and *HHEX/IDE* genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population. Diabetes 2008;57:2834–2842
28. Herder C, Rathmann W, Strassburger K, Finner H, Grallert H, Huth C, Meisinger C, Gieger C, Martin S, Giani G, Scherbaum WA, Wichmann HE, Illig T. Variants of the PPARG, IGF2BP2, CDKAL1, HHEX, and TCF7L2 genes confer risk of type 2 diabetes independently of BMI in the German KORA studies. Horm Metab Res 2008;40:722–726
29. Deeb SS, Fajas L, Nemoto M, Pihlajamaki J, Mykkanen L, Kuusisto J, Laakso M, Fujimoto W, Auwerx J. A Pro12Ala substitution in PPAR-gamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. Nat Genet 1998;20:284–287
30. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Nathan DM; the Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346:393–403
31. Tuomilehto J, Lindström J, Eriksson JG, Valle TT, Hämäläinen H, Ilanne-Parikka P, Keinänen-Kiukaanniemi S, Laakso M, Louheranta A, Rastas M, Salminen V, Uusitupa M; Finnish Diabetes Prevention Study Group. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med 2001;344:1343–1350