

Common variants in the *TCF7L2* gene and predisposition to type 2 diabetes in UK European Whites, Indian Asians and Afro-Caribbean men and women

Steve E. Humphries · David Gable · Jackie A. Cooper · Helen Ireland · Jeffrey W. Stephens · Steven J. Hurel · Ka Wah Li · Jutta Palmen · Michelle A. Miller · Francesco P. Cappuccio · Robert Elkeles · Ian Godslan · George J. Miller · Philippa J. Talmud

Received: 26 May 2006 / Accepted: 11 August 2006
© Springer-Verlag 2006

Abstract Common variants of *TCF7L2*, encoding a β -cell-expressed transcription factor, are strongly associated with increased risk of type 2 diabetes (T2D). We examined this association using both prospective and case-control

Electronic supplementary material Supplementary material is available in the online version of this article at <http://dx.doi.org/10.1007/s00109-006-0108-7> and is accessible for authorized users.

S. E. Humphries (✉) · D. Gable · J. A. Cooper · H. Ireland · K. W. Li · J. Palmen · P. J. Talmud
Department of Medicine, Centre for Cardiovascular Genetics, British Heart Foundation Laboratories, Rayne Building, Royal Free and University College Medical School, 5 University Street, London WC1E 6JF, UK
e-mail:rmhaseh@ucl.ac.uk

J. W. Stephens
Diabetes Research Group, The Medical School, University of Wales Swansea, Singleton, Park, Swansea SA2 8PP, UK

S. J. Hurel
Department of Diabetes and Endocrinology, UCL Hospitals, Mortimer Street, London W1T 3AA, UK

M. A. Miller · F. P. Cappuccio
Clinical Sciences Research Institute, UHCW Campus, Warwick Medical School, Clifford Bridge Road, Coventry CV2 2DX, UK

R. Elkeles · I. Godslan
Endocrinology and Metabolic Medicine, Imperial College London and St Mary's Hospital, London, UK

G. J. Miller
Medical Research Council Cardiovascular Group, Department of Environmental and Preventive Medicine, Wolfson Institute of Preventive Medicine, Charterhouse Square, London EC1M 6BQ, UK



STEVE E. HUMPHRIES received his Ph.D. in Biochemistry and Molecular Biology from the University of Glasgow. He is the British Heart Foundation Professor of Cardiovascular Genetics in the Department of Medicine, University College London, UK and the CEO of the London IDEAS Genetics Knowledge Park. His research focuses on the genetic determinants of CHD traits and the development of DNA tests to identify those at risk of CHD.



PHILIPPA J. TALMUD received her Ph.D. in Microbial Genetics from University College London, UK. She is currently a Professor of Cardiovascular Genetics in the Department of Medicine at University College London. Her research interests include the study of gene: environment interactions in complex disorders such as coronary heart disease and diabetes.

designs. A total of 2,676 healthy European white middle-aged men from the prospective NPHSII (158 developed T2D over 15 years surveillance) were genotyped for two intronic SNPs [rs 7903146 (IVS3C>T) and rs12255372 (IVS4G>T)] which showed strong linkage disequilibrium (D' =0.88, p <0.001; R^2 =0.76, p <0.001). The IVS5T allele

frequency was 0.28 (95% CI 0.27–0.29) and 0.33 (0.28–0.39) in healthy and T2D, respectively ($p=0.04$). Compared to CC men, CT and TT men had an adjusted [for age, body mass index, systolic blood pressure, triglyceride and C-reactive protein levels] hazard ratio for T2D of 1.65 (1.13–2.41) and 1.87 (0.99–3.53), respectively, $p<0.01$. The population attributable fraction for diabetes risk was 17%. In 1459, European white T2D men and women (60% male), T allele frequency was 0.36 (0.34–0.38), and compared to NPHSII healthy men the OR for T2D for the CT and TT subjects was 1.43 (1.24–1.65) and 2.11 (1.69–2.63), respectively $p=<0.0001$. A similar effect was observed in 919 T2D Indian Asians [OR=1.50 (1.14–1.99) and 1.64 (1.03–2.63) $p=0.003$] and 385 Afro-Caribbean subjects [OR=1.25 (0.90–1.75) and 1.32 (0.74–2.33) $p=0.17$] compared to non-diabetic ethnically matched subjects from South London. Weaker associations were found for the IVS4G>T in all studies. Linkage disequilibrium between the two SNPs was high in Indian Asians ($D'=0.94$), but much weaker in Afro-Caribbeans ($D'=0.17$) and haplotype frequencies differed markedly in this group. These results extend previous observations to other ethnic groups, and strongly confirm that *TCF7L2* genotype is a major risk factor for development of T2D.

Keywords Type 2 diabetes · Risk prediction · *TCF7L2* · South Indian · Afro-Caribbean

Introduction

Whilst the current world-wide prevalence of type 2 diabetes (T2D) was estimated to be 2.8% in 2000 [1], it is predicted to increase to epidemic proportions in the coming decades, primarily due to lifestyle changes, particularly obesity. In the United Kingdom, there are over 1.4 million men and women with T2D and this is forecasted to increase by 20–30% by the year 2035 [2]. In addition to a strong environmental element, the existence of an underlying genetic component to T2D risk is supported by twin studies, family studies and the widely different T2D prevalence across ethnic groups [3]. By comparison to UK European whites with an estimated prevalence of 2.4%, the rate in UK Indian Asians and Afro-Caribbeans is three- to sixfold higher [1]. This disparity cannot be explained by cultural and lifestyle differences alone and raise the possibility that part of the ethnic difference could be due to the presence of different genes, or different allele frequencies of predisposing genes, in these groups.

Localization of a susceptibility locus on chromosome 2q37 and subsequent positional cloning analyses suggested that calpain-10 (*CAPN10*) was associated with T2D [4]. However, other studies found inconsistent associations and

only meta-analysis has provided convincing support for the involvement of *CAPN10* variants in T2D risk [5]. Three additional susceptibility loci have been identified for T2D on chromosome 5q34, 10q and 12q in the Icelandic population [6]. In a recent paper, Grant et al. [7] reported that the chromosome 10q25.2 locus contains a gene determining T2D risk called transcription factor 7-like 2 gene (*TCF7L2*). *TCF7L2* is a high mobility group (HMG) box-containing transcription factor which is part of the Wnt signalling pathway. Using a microsatellite, DG10S478, within exon 3 of *TCF7L2*, Grant et al. [7] showed a highly significant association with T2D in an Icelandic case control cohort ($p<2.1\times 10^{-9}$) which was replicated in Danish and American cohorts. Heterozygous and homozygous carriers of the at-risk alleles had a relative risk of T2D of 1.45 and 2.41, respectively, corresponding to a population attributable risk (PAF) of 21%, although the causative mutation has not been identified. Two intronic single nucleotide polymorphism (SNPs) within *TCF7L2* were subsequently identified, rs7903146 (IVS3C>T) and rs12255372 (IVS4G>T), which replicated the strong association of DG10S478 with T2D risk.

In this study, we have examined the association of IVS3C>T and IVS4G>T with future risk of T2D by prospective cohort analysis and compared the risk association in T2D cohorts of European white, Indian Asian and Afro-Caribbean origin, with non-diabetic ethnically-matched subjects.

Materials and methods

Study groups

The samples used and their numbers by ethnic group and study are shown in Supplementary Table 1.

The prospective Second Northwick Park Heart Study

From April 1989 to April 1994, 3,012 healthy European white men, aged 50–64 years, registered with nine primary care practices in the United Kingdom, were recruited for prospective surveillance [8]. The study was approved by the institutional ethics committees and performed in accordance with the declaration of Helsinki. All subjects gave written informed consent. To be eligible, subjects had to be free of unstable angina, myocardial infarction or evidence of silent infarction, coronary surgery, aspirin or anticoagulant therapy, cerebrovascular disease, malignancy (except skin cancer other than melanoma), or any condition precluding informed consent. Weight, height and blood pressure measurements were recorded and venous blood samples were collected for plasma and DNA analysis.

Participants were recalled annually for 5 years for interview and repeat venous blood collection. Self-report by questionnaire was used to identify 69 subjects requiring insulin or oral hypoglycaemics as cases at baseline. New cases were identified up to the end of 2005 by practice note search for physician diagnosed and treated T2DM according to current national guidelines.

Wandsworth Heart and Stroke Study

This is a population-based study drawn from general practices within the former Wandsworth Health Authority in South London [9]. By design, the three ethnic groups, European white ($n=524$) those of African descent ($n=549$) and those of Indian Asian origin ($n=505$) were of similar sample size with comparable numbers of men and women, aged between 40–59. Ethnic group was recorded at the time of interview, based on the answers to a combination of questions including place and country of birth, language, religion, history of migration and parental country of birth. Those with a clinical diagnosis of T2D or who were identified as T2D on the basis of the OGTT (according to WHO criteria [10]) (European white $n=153$, Indian Asian $n=193$, Afro-Caribbean $n=218$). The protocol was approved by the local ethics committee and all participants gave their informed consent.

University College London Diabetes and Cardiovascular Study and the Ealing Diabetes Study of Coagulation

Between 2001–2002, 1,014 consecutive subjects were recruited from the diabetes clinic at University College London Hospitals NHS Trust (UCLH) into University College London Diabetes and Cardiovascular Study (UDACS) [11], and 938 from Ealing Hospital into Ealing Diabetes Study of Coagulation (EDSC) [12]. Patients completed a questionnaire with details of age, ethnicity, smoking habit, fasting status, duration of diabetes, family history of diabetes, history of heart attack, stroke, current medication and other clinical details. Blood was collected for plasma and DNA analysis. Blood pressure, body mass index (BMI), glucose, HbA_{1c}, cholesterol, HDL, LDL, triglyceride (TG), urea, creatinine, albumin to creatinine ratio, potassium and proteinuria were measured. All patients classified as T2D according to WHO criteria [10] were included in the current analysis. For this and the PREDICT study, type 2 diabetes was distinguished from type 1 diabetes using criteria in common use in the UK, based on the basis of the mode of presentation and history. T2D subjects were those presenting with diabetes at 40 years or over, and not requiring insulin therapy within 12 months of diagnosis. Recruitment into UDACS comprised of 600 European white (59% male), 107 Indian

Asians (74% male) and 74 Afro-Caribbeans (66% male). In EDSC, there were 331 European white (60% male), 503 Indian Asians (56% male) and 93 Afro-Caribbeans (61% male). Ethical approval was obtained from UCL/UCLH and Ealing Hospital ethics committees. Full details have been published elsewhere [11, 12].

PREDICT

This is a prospective study of patients with established T2D, recruited from routine diabetes clinics at several London hospitals to assess the predictive value of coronary artery calcification, measured by electron beam tomography for coronary heart disease (CHD) events and stroke [13]. Four hundred European whites (61% men) and 116 Indian Asians (70% men) were considered for the present analysis. Patients were assessed at their local clinics using a standardised protocol. Exclusion criteria were: current or past history of CHD on the basis of clinical history alone without recourse to exercise testing, ECG or angiography, congestive heart failure, uncontrolled hypertension (blood pressure on treatment >160/95 mmHg), serious medical disorders likely to limit life expectancy (e.g. cancer) or requiring extensive medical treatment and command of English insufficient to provide informed consent. Patients were accepted irrespective of treatment status. For those receiving insulin, T2D was diagnosed on the basis of presentation and history. This study had approval from all relevant local ethics committees.

DNA extraction and genotyping

DNA was extracted from whole blood using the salting out method [14]. Genotyping for IVS3C>T (rs7903146) and INV4G>T (rs12255372) was performed using TaqMan technology [Applied Biosciences (ABI), Warrington UK]. Reactions were performed on 384-well microplates and analysed using ABI TaqMan 7900HT software. Primers and MGB probes are detailed in Supplementary Table 2.

Statistical analysis

Analysis was performed using ‘Intercooled STATA’ (version 8.2, STATA, TX, USA). Baseline characteristics for all studies were transformed to a normal distribution as appropriate and differences were tested using analysis of covariance. For the three T2D studies, trait data were combined by ethnic group or genotype after adjusting for study. For the prospective study, results are presented as hazard ratios (HR) obtained from Cox regression models with their corresponding 95% confidence interval (CI), adjusted for age and practice (recruitment site), and with further adjustment as described in the results section.

Genotype effects were tested using an additive model. Differences in allele frequencies were assessed both with and without the inclusion of baseline diabetics. Where results were further adjusted and for the interaction with obesity, 69 baseline diabetics were excluded from the analysis. Haplotypes were inferred by use of a maximum likelihood algorithm implemented in the THESIAS program [15, 16]. The comparison for the T2D cohorts with *non-diabetic* controls is presented as odds ratios (OR) obtained from logistic regression models. Frequencies were compared by Chi-squared test. Population attributable fraction (PAF) (%) was estimated as $pd \times (HR-1/HR)$, where pd is the proportion of cases carrying the rare allele and confidence intervals were based on asymptotic approximations [17]. Statistical significance was taken as $p < 0.05$.

Results

Second Northwick Park Heart Study

Genotype was obtained for *TCF7L2* IVS3C>T and IVS4G>T in 2745 Second Northwick Park Heart Study (NPHSII) men (of whom 69 had T2D at baseline) (98.9% in whom DNA was available) and the baseline characteristics of those who did and did not develop T2D over the 15 year follow-up period are presented in Table 1. At baseline, the 158 men who went on to develop T2D had, as expected, significantly higher blood pressure, higher BMI, a greater proportion with a BMI over 30 kg/m², a higher proportion of current smokers, and higher levels of plasma lipids, fibrinogen and C-reactive protein (CRP). For both SNPs genotype, distribution was as expected from Hardy Weinberg proportions,

Table 1 Baseline characteristics and their association with the development of type 2 diabetes in NPHSII (excluding and including those with T2D at baseline)

	No diabetes (2,518)	With diabetes (158)	HR (95% CI) ^a	<i>p</i> -value
Age (years)	56.0 (3.4)	56.3 (3.4)	1.17 (0.92–1.48)	0.21
SBP (mmHg) ^b	136.7 (18.7)	142.1 (19.2)	1.34 (1.14–1.57)	<0.0001
DBP (mmHg)	84.5 (11.3)	86.4 (11.3)	1.14 (0.97–1.34)	0.11
BMI (kg/m ²) ^b	26.1 (3.3)	28.5 (3.6)	2.03 (1.75–2.36)	<0.0001
Obesity (% >30 kg/m ²)	12.3 (310)	33.1 (52)	3.68 (2.63–5.15)	<0.0001
Smoking (% current)	28.1 (707)	34.8 (55)	1.50 (1.08–2.09)	0.02
Cholesterol (mmol/l)	5.73 (1.01)	5.90 (0.98)	1.21 (1.04–1.41)	0.02
Triglyceride (mmol/l) ^b	1.76 (0.92)	2.25 (1.07)	1.56 (1.34–1.81)	<0.0001
Fibrinogen (g/l) ^b	2.71 (0.51)	2.76 (0.53)	1.14 (0.98–1.34)	0.09
CRP (mg/l) ^b	2.96 (3.45)	4.01 (4.31)	1.34 (1.13–1.59)	0.001
CRP excluding >20 mg/l	2.44 (2.46)	3.32 (3.13)	1.42 (1.17–1.72)	<0.0001
<i>TCF7L2</i> IVS3C>T				
CC	1,295 (52.0)	67 (43.0)	1.00	
CT	1,001 (40.2)	74 (47.4)	1.47 (1.05–2.06)	0.06
TT	197 (7.9)	15 (9.6)	1.43 (0.81–2.51)	
T allele (95% CI)	0.280 (0.27–0.29)	0.333 (0.28–0.39)		0.04
<i>TCF7L2</i> IVS4G>T				
GG	1,307 (52.5)	66 (42.3)	1.00	
GT	994 (39.9)	78 (50.0)	1.57 (1.13–2.19)	0.03
TT	188 (7.6)	12 (7.7)	1.25 (0.68–2.33)	
T allele (95% CI)	0.275 (0.26–0.29)	0.327 (0.28–0.38)		0.05
<i>TCF7L2</i> IVS3C>T				
CC	1,295 (52.0)	91 (40.6) ^c	1.00	
CT	1,001 (40.2)	111 (49.6) ^c	1.61 (1.22–2.14)	0.002
TT	197 (7.9)	22 (9.8) ^c	1.58 (0.99–2.52)	
T allele (95% CI)	0.280 (0.27–0.29)	0.346 (0.30–0.39) ^c		0.003
<i>TCF7L2</i> IVS4G>T				
GG	1,307 (52.5)	94 (42.2) ^c	1.00	
GT	994 (39.9)	109 (48.9) ^c	1.52 (1.14–2.00)	0.01
TT	188 (7.6)	20 (9.0) ^c	1.48 (0.91–2.40)	
T allele (95% CI)	0.275 (0.26–0.28)	0.327 (0.28–0.38) ^c		0.008

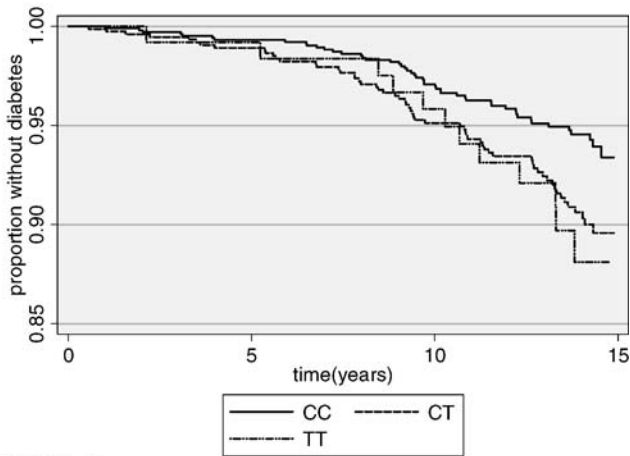
CRP measurements made after diabetes was recorded are excluded from the analysis ($N=2$).

^a Age and practice adjusted hazard ratio for 1 SD increase in all variables except smoking (current: none), obesity (>30:<30) and age (5 year increase)

^b Geometric mean (approximate SD)

^c Including baseline diabetics (227)

a) IVS3C>T



b) IVS4G>T

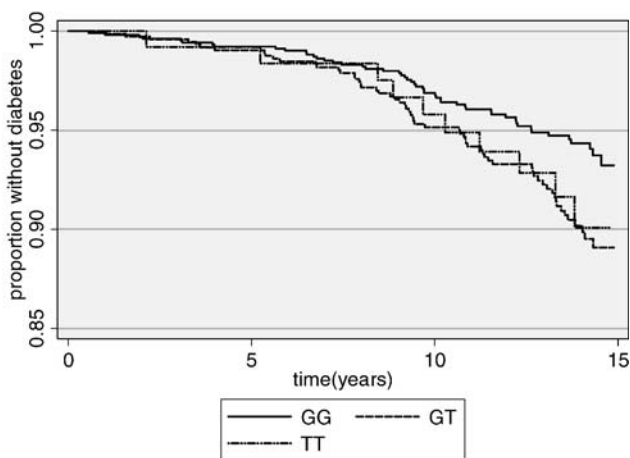


Fig. 1 Kaplan Meier plot of development of T2D by *TCF7L2* genotypes in NPHS men. **a** For the IVS3C>T SNP: Compared to the CC individuals, individuals with the CT genotype had a hazard ratio of 1.65 (1.13–2.41) and those with the TT genotype 1.87 (0.99–3.53) $p < 0.01$, after adjustment for age, practice, CRP, BMI, SBP and TG. **b** For the IVS4G>T SNP: Compared to the GG individuals, individuals with the GT genotype had a hazard ratio of 1.72 (1.18–2.50) and those with the TT genotype 1.67 (0.84–3.31) $p < 0.01$, after adjustment for age, practice, CRP, BMI, SBP and TG

and as shown in Table 1, the frequencies of the rare alleles were significantly higher in those with T2D (whether those with diabetes at baseline were excluded or included). Figure 1 shows Kaplan Meier plots for the development of T2D by the two SNPs, showing an early and continuing divergence of risk. The possibility that *TCF7L2* genotype might be having its impact on development of T2D through diabetes risk factors was considered, but individuals homozygous for the rare allele tended to have lower levels of diabetes risk factors (Supplementary Table 3a,b). After adjustment for age, general practice recruitment, plasma triglyceride, CRP, systolic blood pressure and BMI, compared to the IVS3 CC genotype, subjects with the CT genotype had a hazard ratio of 1.65 (1.13–2.41) and those with the TT genotype 1.87 (0.99–3.53) $p < 0.01$ for linear trend. For the IVS4G>T genotype, these values were 1.72 (1.18–2.50) for the GT group and 1.67 (0.84–3.31) for the TT group ($p < 0.01$).

Because obesity is a major determinant of development of T2D, the possibility of interaction between these SNPs and BMI was examined. As shown in Fig. 2, men with a BMI greater than 30 kg/m² had no increased risk associated with IVS3 genotype with a hazard ratio per T allele of 0.77 (0.45–1.32) after adjustment for age, general medical practice, CRP and TG. For individuals with a lower BMI, the risk increased as BMI decreased, and for those below 25 kg/m², the HR per T allele was 2.23 (1.15–4.31). The interaction between genotype and BMI was statistically significant ($p = 0.001$). Similar results (Fig. 2 legend) were seen for interaction between the IVS4G>T genotype and BMI (interaction value $p = 0.04$). Overall, in this sample, the population attributable fraction for development of T2D associated with obesity was 23% (95% CI 0.14–0.30), whilst that associated with IVS3C>T was 16.6% (95% CI 0.01–0.30), and for the IVS4G>T was 18.6% (0.03–0.32), with the combination for both genotypes being 19.9% (0.03–0.34). This suggests that this genotype is having as large an impact on development of diabetes as obesity in this group of middle-aged UK subjects.

Fig. 2 Hazard ratio for T2D by BMI for each IVS3 T allele carried. For IVS4 T allele the HRs were for BMI > 30 kg/m² 0.79 (0.46–1.38), 28–30 kg/m² 1.91 (1.08–3.37), 25–27.9 kg/m² 1.42 (0.89–2.28), < 25 kg/m² 2.28 (1.17–4.45). BMI genotype interaction $p = 0.04$

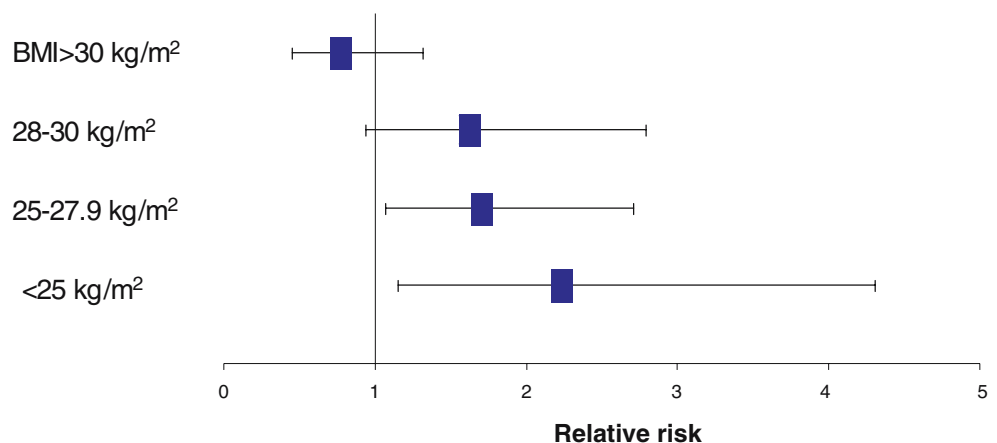


Table 2 Haplotype association with T2D (NPHSII)

IVS3C>T IVS4G>T	Non-diabetic frequency	Diabetic frequency	HR ^a (95% CI)
CG	0.698	0.637	1.00
CT	0.023	0.014	0.70 (0.32–1.55) $p=0.38$
TG	0.030	0.023	0.94 (0.50–1.76) $p=0.85$
TT	0.250	0.325	1.40 (1.13–1.73) $p=0.002$
Global p value			$p<0.01$

^a Age adjusted

Information from the two SNPs was combined into inferred haplotypes to attempt to see if a risk haplotype could be identified. The linkage disequilibrium (LD) between the two SNPs was high with a D' of 0.88 ($p<0.001$), $R^2=0.76$, $p<0.001$. As shown in Table 2, all four possible haplotypes were observed, with only the haplotype defined by both rare alleles being associated with significant effects on risk, and a HR per haplotype of 1.40 (1.13–1.73). This risk was not affected by adjustment for baseline recorded plasma triglyceride levels or BMI. Thus, middle-aged men homozygous for the IVS3T/IVS4T haplotype (~6% of the population) would be predicted to have an increased risk of developing T2D over 15 years of approximately 196% compared to men homozygous for the common haplotype.

T2D case-control studies

To obtain maximal power, T2D subjects recruited for three different studies in London (UDACS, EDSC and PREDICT) were combined. The overall characteristics of these study groups are very similar ([12] and data not shown). The characteristics of these subjects by ethnic group are presented in Table 3. These are typical of T2D patients, with the expected differences in the prevalence of diabetes risk factors in subjects from the different ethnic groups (an earlier age of onset and a lower BMI and prevalence of obesity in the Indian Asians). For European whites, the NPHSII T2D-free group was used for the comparison, whilst subjects of Indian Asians and Afro-Caribbeans ethnicity were recruited from the Wandsworth Heart and Stroke Study. The characteristics of these subjects have been published previously [9]. Using standard WHO criteria, subjects with T2D at baseline were identified and results combined with those from the T2D clinic groups (see Supplementary Table 1). Tables 4, 5 and 6 show the genotype distribution, allele frequency, and ORs for T2D in the three ethnic groups separately. There was no evidence of significant genotype or allele frequency differences when subjects of different ethnic groups from the three studies were compared (all $p>0.2$, data not shown). The allele frequency of the IVS3C>T is essentially the same across ethnic groups (all contrasts $p>0.1$), whilst for IVS4G>T, the

Table 3 Baseline characteristics by ethnic group in UDACS, EDSC and PREDICT combined

	European white ($N=1331$)	Black ($N=167$)	S. Indians ($N=726$)
Age (years)	64.9 (10.9)	61.6 (11.2)	58.7 (10.9)
Age of onset (years)	54.9 (12.2)	50.8 (10.6)	48.1 (11.2)
Duration of diabetes (years) ^a	8 (4–14)	10 (5–15)	10 (5–16)
SBP (mmHg) ^b	137.1 (19.0)	142.5 (20.6)	134.2 (20.1)
DBP (mmHg)	77.5 (10.8)	79.5 (12.1)	76.2 (11.0)
BMI (kg/m ²) ^b	29.4 (5.5)	28.4 (4.9)	27.4 (4.4)
Obesity (% >30 kg/m ²)	46.6 (340)	33.3 (31)	25.9 (160)
Smoking (% current)	26.2 (343)	13.9 (23)	8.9 (64)
Cholesterol (mmol/l)	4.96 (1.07)	4.99 (1.04)	4.87 (1.12)
Triglyceride (mmol/l) ^b	1.83 (1.03)	1.44 (0.77)	1.92 (1.02)
CRP (mg/l) ^b	1.78 (1.60)	1.31 (1.56)	1.92 (1.95)
HbA _{1c} (%)	8.01 (1.65)	8.40 (2.07)	8.49 (1.59)

^a Median (inter-quartile range)^b Geometric mean (approximate SD)

frequency in the Indian Asians and Afro-Caribbeans was modestly but not statistically significantly lower ($p=0.055$), and higher ($p=0.046$), respectively. For the IVS3 SNP, the rare allele frequency in T2Ds was higher in all groups compared to non-diabetic controls, with this difference being statistically significant in European whites and Indian Asians. For European whites, the OR for each additional T allele was 1.45 (1.31–1.59) $p<0.0001$, and for the Indian Asians the respective OR was 1.37 (1.11–1.68) $p=0.003$. For the Afro-Caribbean group, in whom numbers were smallest, this effect was 1.19 (0.93–1.52) $p=0.17$. There was no evidence for heterogeneity of effect on risk associated with this genotype between ethnic groups ($p=0.34$ for heterogeneity). These estimates were essentially unchanged when analyses were confined to clinic subjects only. (Supplementary Table 4a–c) There was no evidence for a significant difference in this OR between men and women (p for interaction in European whites between gender and genotype for IVS3=0.98, for IVS4=0.51, data not shown). For the IVS4 SNP, the effect sizes were essentially similar in the European whites and Indian Asians, but by contrast, in the Afro-Caribbeans, there was no evidence for increased risk in those carrying the rare allele, with significant evidence for heterogeneity of effect on risk associated with this genotype between ethnic groups ($p=0.007$ for heterogeneity).

For both SNPs, there was again evidence that individuals homozygous for the rare allele had lower levels of classical T2D risk factors. Due to lack of heterogeneity of genotype effect amongst the three ethnic groups, data could be combined and those with IVS3 TT genotype had the lowest plasma TG, 1.66 mmol/l (0.94) vs 1.84 mmol/l (1.00), $p<0.001$, and modestly lower CRP ($p=0.05$). Similar effects

Table 4 Genotype distribution, allele frequency and odds ratio for T2D in European white

	NPHSII	T2D	OR (95% CI)	<i>p</i> value
IVS3C>T				
CC	1,295 (52.0)	601 (41.2)	1.00	<i>p</i> <0.0001
CT	1,001 (40.2)	665 (45.6)	1.43 (1.25–1.64)	
TT	197 (7.9)	193 (13.2)	2.11 (1.69–2.63)	
Allele frequency; (95% CI)	0.280; (0.267–0.292)	0.360; (0.343–0.378)	CT/TT:CC; 1.54 (1.35–1.76)	<i>p</i> <0.0001
IVS4G>T				
GG	1,307 (52.5)	636 (43.3)	1.00	<i>p</i> <0.0001
GT	994 (39.9)	662 (45.1)	1.37 (1.19–1.57)	
TT	188 (7.6)	170 (11.6)	1.86 (1.48–2.33)	
Allele frequency; (95% CI)	0.275; (0.263–0.288)	0.341; (0.324–0.359)	GT/TT:GG; 1.45 (1.27–1.65)	<i>p</i> <0.0001

were seen in IVS4T>G for plasma TG levels ($p=0.03$) (Supplementary Table 4a,b).

Haplotypes for the two SNPs were inferred in subjects from all three ethnic groups. The LD in Indian Asians was similar to that in European whites ($D'=0.94$, $p<0.001$, $R^2=0.68$, $p<0.001$) whilst in Afro-Caribbeans, LD was much weaker ($D'=0.17$, $p=0.61$, $R^2=0.03$, $p=0.71$). In the T2D European whites, as shown in Table 7, IVS3T-IVS4G and IVS3T-IVS4T haplotypes were associated with significantly higher T2D risk, an effect which was not materially changed by adjustment for recorded plasma triglyceride levels (not shown). In Indian Asians (Table 8), these same two haplotypes were also associated with significantly elevated risk. In Afro-Caribbeans (Table 9), the relative frequency of the four haplotypes was markedly different, with the IVS3T-IVS4T being the rarest, and IVS3C-IVS4T and IVS3T-IVS4G being present at much higher frequencies than in European whites. The IVS3T-IVS4G haplotype was associated with the highest HR [1.22 (0.88–1.70)] but this effect was not statistically significant ($p=0.24$).

Discussion

These data have confirmed the importance of variants in the *TCF7L2* gene in the risk of development of T2D, in both

healthy middle-aged men and in patients with T2D. In the prospective study, the population attributable fraction for the combined effect of both variants is large (20%), in keeping with results from previous studies [7]. This effect is much larger than other genes so far studied in NPHSII [18] and is of similar magnitude to the effect of obesity (BMI>30 kg/m²) in this study which was 22.3%. Our results extend previous observations to subjects of South Indian origin, in whom the frequency of the risk alleles is similar and the effect size on risk of T2D is comparable. The frequency of the risk variants was also similar in Black subjects, but our study was underpowered to demonstrate unequivocally that these variants are having a similar effect in this ethnic group, although the data suggest that this is likely.

The mechanism by which these *TCF7L2* variants influence development of T2D is currently unclear. In NPHSII, men homozygous for the risk-raising alleles in general had lower baseline levels of diabetes risk factors, including BMI, plasma TG and CRP. Despite this, men carrying one or more of either of these alleles had about 40% higher risk of developing T2D over the 15 years of follow-up. After adjustment for these diabetes risk factors, compared to those homozygous for the IVS3C allele, there was a step-wise increase in the HR for those carrying one or two IVS3T alleles (1.65 and 1.87, respectively), although because of the small sample size, the confidence intervals for these two

Table 5 Genotype distribution, allele frequency and odds ratio for T2D in Indian Asians

	WHSS	T2D	OR (95% CI)	<i>p</i> value
IVS3C>T				
CC	163 (54.3)	366 (43.7)	1.00	<i>p</i> =0.003
CT	111 (37.0)	375 (44.8)	1.50 (1.14–1.99)	
TT	26 (8.7)	96 (11.5)	1.64 (1.03–2.63)	
Allele frequency; (95% CI)	0.272; (0.236–0.309)	0.339; (0.316–0.362)	CT/TT:CC; 1.53 (1.17–2.00)	<i>p</i> =0.002
IVS4G>T				
GG	178 (58.9)	423 (50.3)	1.00	<i>p</i> =0.02
GT	104 (34.4)	350 (41.6)	1.42 (1.07–1.87)	
TT	20 (6.6)	68 (8.1)	1.43 (0.84–2.43)	
Allele frequency; (95% CI)	0.238; (0.205–0.274)	0.289; (0.267–0.311)	GT/TT:GG; 1.42 (1.09–1.85)	<i>p</i> =0.01

Table 6 Genotype distribution, allele frequency and odds ratio for T2D in Afro-Caribbeans

	WHSS	Combined	OR (95% CI)	<i>p</i> value
IVS3C>T				
CC	161 (51.8)	141 (45.9)	1.00	<i>p</i> =0.17
CT	124 (39.9)	136 (44.3)	1.25 (0.90–1.75)	
TT	26 (8.4)	30 (9.8)	1.32 (0.74–2.33)	
Allele frequency; (95% CI)	0.283; (0.248–0.320)	0.319; (0.282–0.358)	CT/TT:CC; 1.26 (0.92–1.73)	<i>p</i> =0.15
IVS4G>T				
GG	141 (45.1)	156 (51.3)	1.00	<i>p</i> =0.36
GT	148 (47.3)	120 (39.5)	0.73 (0.53–1.02)	
TT	24 (7.7)	28 (9.2)	1.05 (0.58–1.90)	
Allele frequency; (95% CI)	0.313; (0.277–0.351)	0.289; (0.245–0.348)	GT/TT:GG; 0.78 (0.57–1.07)	<i>p</i> =0.12

estimates overlap considerably. It has previously been reported that the effect of this genotype is essentially co-dominant [7], and our data confirm this finding. In support of this, in the group of European white T2D subjects, the risk associated with having one and two copies of the IVS3 SNP was 1.43 and 2.11, respectively, with a similar co-dominant pattern seen in the IVS4 variant. In the Indian Asian subjects, although the effect of carrying either the IVS3 or IVS4 rare variant was to raise risk by 53 and 42%, respectively, there was not strong evidence for a co-dominant effect. The sample size in this study is, however, much smaller, and confidence intervals overlapped. Data in the Afro-Caribbean subjects must be interpreted with caution, but the results are compatible with the IVS3 SNP having essentially the same effect on risk as in the other ethnic groups (no statistically significant evidence for heterogeneity of genotype effect across ethnic groups). Although because of the small sample size, this effect was of borderline statistical significance. Power calculations indicate that a sample size of roughly twice that available to us would be required for the effects seen to be statistically significant. It is interesting to note that the IVS4 SNP was not showing any evidence of increased risk in Black subjects.

We have previously reported that common variation in the genes for uncoupling proteins (*UCP2* and *UCP3*) show additive effects with obesity in the NPHSII men [18], and the observation that the *TCF7L2* variants appear to have their

major effect on prospective risk of T2D in non-obese subjects is of interest. The statistical evidence for the genotype–obesity interaction is due to the lower than expected risk in IVS3T carriers who are obese, compared to the risk seen in non-obese IVS3T carriers. One plausible explanation for this apparent paradox would be that a proportion of obese IVS3T carriers develops T2D prematurely, and were therefore ineligible for recruitment into studies such as NPHSII where diabetes as baseline was an exclusion criterion. Studies with earlier recruitment and longer follow-up would be required to explore this suggestion.

Haplotype analysis

In the original report [7], the authors recommended using both the IVS3 and IVS4 SNPs which in combination could reproduce the strong association they had found with a hypervariable marker in *TCF7L2*. Although all four possible haplotypes were seen unambiguously in all three ethnic groups, in both Whites and Indian Asians the frequency of the IVS3C/IVS4T and IVS3T/IVS4G haplotypes were very low (below 5% in all groups). In NPHSII, significant risk was only found associated with the IVS3T/IVS4T haplotype. However, in the case–control comparisons both in European whites and in Indian Asians, both IVS3T/IVS4G and IVS3T/IVS4T haplotypes showed significantly elevated risk. There was much lower LD between these two SNPs in the Afro-Caribbean subjects, and only

Table 7 Combined T2D European whites

IVS3C>T IVS4G>T	Non-diabetic frequency	Diabetic frequency	OR ^a (95% CI)
CG	0.698	0.623	1.00
CT	0.023	0.019	0.97 (0.68–1.37) <i>p</i> =0.86
TG	0.030	0.036	1.42 (1.09–1.86) <i>p</i> =0.01
TT	0.250	0.322	1.44 (1.29–1.59) <i>p</i> <0.0001
Global <i>p</i> value			<i>p</i> <0.0001

^a Age adjusted**Table 8** Combined T2D Indian Asians

IVS3C>T IVS4G>T	Non-diabetic frequency	Diabetic frequency	OR (95% CI)
CG	0.721	0.649	1.00
CT	0.005	0.011	2.51 (0.74–8.57) <i>p</i> =0.14
TG	0.040	0.063	1.71 (1.11–2.62) <i>p</i> =0.01
TT	0.233	0.276	1.31 (1.06–1.63) <i>p</i> =0.01
Global <i>p</i> value			<i>p</i> =0.006

Table 9 Combined T2D Afro-Caribbeans

IVS3C>T IVS4G>T	Non-diabetic frequency	Diabetic frequency	OR (95% CI)
CG	0.499	0.486	1.00
CT	0.220	0.196	0.90 (0.64–1.27) $p=0.54$
TG	0.182	0.225	1.22 (0.88–1.70) $p=0.24$
TT	0.098	0.093	0.96 (0.61–1.51) $p=0.86$
Global p value			$p=0.39$

the IVS3T/IVS4G haplotype showed increased risk in this group (albeit non-statistically significant in this small sample). Grant et al. [7] reported that they were unable to identify a single functional sequence change on the *TCF7L2* risk haplotype by extensive sequencing, and the data in this study suggest that both the IVS3T/IVS4G or IVS3T/IVS4T haplotypes warrant further studies.

Study limitations

There are several limitations to these findings which need to be considered. In the prospective study, the number of individuals who have so far developed T2D is relatively small, and these results need to be confirmed in other studies. It is also possible that there may be a “catch-up” effect, in that individuals with low risk *TCF7L2* genotypes may develop T2D at a faster rate as they become older, when other risk factors are driving the development of disease. In the case–control comparison, to provide maximal power, data have been combined from different studies, which raises the possibility of the results being confounded by between-study heterogeneity. We believe this is extremely unlikely to be the case because the overall characteristics of the T2D studies were very similar [12] and genotype and allele distributions were non-significantly different between the three studies. For the European white subjects, allele frequencies were significantly different between healthy subjects and patients in all T2D studies separately, and for Indian Asians in the larger EDSC and PREDICT Studies, compared to non-diabetic Indian Asians in WHSS. For the smaller number of Afro-Caribbean subjects, the findings reported in this study should only be considered as preliminary.

Molecular mechanism of effect

The lack of association of *TCF7L2* genotype with T2D risk factors, and the higher genotype risk in non-obese subjects seen in this study confirms/suggests that the mechanism of the risk effect may be novel. *TCF7L2* is one of a family of HMG box transcription factors that have been identified as down-

stream effectors of the Wnt signalling pathway. Aberrant Wnt signalling has been linked to familial and sporadic colon cancers [19] but it is a novel candidate pathway for the development of T2D. The mechanism responsible for the *TCF7L2* risk association in T2D is currently not understood, but it has been speculated [7] that either the IVS5 and IVS6 variants themselves, or other sequence changes in LD with these variants, may influence the known pattern of isoforms expressed due to alternative splicing [20], and thus down- (or up-) regulate *TCF7L2* expression and thus Wnt signalling.

There are a number of ways by which changes to Wnt signalling could increase the risk of disease. The *TCF7L2* knockout mouse dies within 24 h due to inadequate intestinal epithelium, but with no other reported organ abnormalities [21], whilst misexpression of Wnts leads to significant abnormal development of foregut structures, including the pancreas [22]. Thus, disruption in the pathway could reduce both pancreatic size and number of insulin secreting cells. It has been suggested [7] that low levels of *TCF7L2* expression could cause a failure to develop adequate gut endocrine tissue, which would lead to disturbance of enteroendocrine function, which in turn could lead to T2D. The gut secretes a number of hormones which have an insulin-sensitizing effect such as glucagon like peptide 1 (GLP-1), a product of the proglucagon gene, and *TCF7L2* can influence the expression of the proglucagon gene [23]. GLP-1 both enhances pancreatic function and insulin sensitivity and is currently in clinical trials as a treatment for diabetes.

Prospective gene associations are overall more powerful and less open to confounding by population sub-structure than those of case–control design [24]. The prospective confirmation of the *TCF7L2* gene effect, and thus of Wnt signalling as an important candidate in the pathogenesis of T2D, opens up a new pathway for investigation. The higher risk of T2D in subjects from the Indian subcontinent could, in theory, be explained by a higher frequency of certain genotypes predisposing to insulin resistance, obesity, reduced insulin secretion or a combination of these [25]. However, the similar frequency of the *TCF7L2* SNPs examined in this study in all ethnic groups studied suggests that this gene is not contributing to this ethnic risk difference. The absence of heterogeneity of effect on risk in different ethnic groups, if confirmed, makes these genotypes more useful as genetic risk markers than others reported to date [26, 27]. The variants and haplotypes are common. The molecular identification of the causal defect(s) and a specific treatment targeted at carriers of these variants would have a significant public health impact in people with or at risk of T2D.

Acknowledgements We thank our clinical colleague Dr. Hugh Mathur for supporting the recruitment of the EDSC patients, and the following general practices who collaborated in the study: The Surgery, Aston Clinton, Upper Gordon Road, Camberley; The Health Centre, Carnoustie; Whittington Moor Surgery, Chesterfield; The

Market Place Surgery, Halesworth; The Health Centre, Harefield; Potterells Medical Centre, North Mymms; Rosemary Medical Centre, Parkstone, Poole; The Health Centre, St. Andrews. NPHSII was supported by the UK Medical Research Council, the US National Institutes of Health (grant NHLBI 33014) and Du Pont Pharma, Wilmington, USA. EDSC was supported by the Coronary Thrombosis Trust. JWS was supported by a clinical training fellowship from Diabetes UK (BDA: RD01/0001357). DRG and JAC and SEH are supported by the British Heart Foundation (FS/04/012:RG2005/014). This work was carried out in part with support from the Department of Health and the Department of Trade and Industry for the IDEAS Genetics Knowledge Park. A list of the WHSS Group is given elsewhere [10]. The study has received support from the former Wandsworth and South Thames Regional Health Authorities, NHS R&D Directorate, British Heart Foundation, former British Diabetic Association and The Stroke Association. Financial support for the PREDICT Study was provided by the British Heart Foundation and the Tompkins Foundation. IFG is supported by the Heart Disease and Diabetes Research Trust. Financial contributions were also received from AstraZeneca UK, Fournier Pharmaceuticals, and TakedaUK. The contribution of other members of the PREDICT Study group [13] is gratefully acknowledged including A. Dunlop and A. Widdowson.

References

- Oldroyd J, Banerjee M, Heald A, Cruickshank K (2005) Diabetes and ethnic minorities. *Postgrad Med J* 81:486–490
- Bagust A, Hopkinson PK, Maslove L, Currie CJ (2002) The projected health care burden of type 2 diabetes in the UK from 2000 to 2060. *Diabet Med* 19(Suppl 4):1–5
- Gloyn AL, McCarthy MI (2001) The genetics of type 2 diabetes. *Best Pract Res Clin Endocrinol Metab* 15:293–308
- Horikawa Y, Oda N, Cox NJ et al (2000) Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nat Genet* 26:163–175
- Weedon MN, Schwarz PE, Horikawa Y et al (2003) Meta-analysis and a large association study confirm a role for calpain-10 variation in type 2 diabetes susceptibility. *Am J Hum Genet* 73:1208–1212
- Reynisdottir I, Thorleifsson G, Benediktsson R et al (2003) Localization of a susceptibility gene for type 2 diabetes to chromosome 5q34–q35.2. *Am J Hum Genet* 73:323–335
- Grant SF, Thorleifsson G, Reynisdottir I et al (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38:320–323
- Miller GJ, Bauer KA, Barzegar S et al (1995) The effects of quality and timing of venepuncture on markers of blood coagulation in healthy middle-aged men. *Thromb Haemost* 73:82–86
- Cappuccio FP, Cook DG, Atkinson RW, Strazzullo P (1997) Prevalence, detection, and management of cardiovascular risk factors in different ethnic groups in south London. *Heart* 78:555–563
- Alberti KG, Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 15:539–553
- Stephens JW, Hurel SJ, Acharya J, Humphries SE (2004) An interaction between the interleukin-6 –174G>C gene variant and urinary protein excretion influences plasma oxidative stress in subjects with type 2 diabetes. *Cardiovasc Diabetol* 3:2
- Flavell DM, Ireland H, Stephens JW et al (2005) Peroxisome proliferator-activated receptor alpha gene variation influences age of onset and progression of type 2 diabetes. *Diabetes* 54:582–586
- Elkeles RS, Feher MD, Flather MD et al (2004) The association of coronary calcium score and conventional cardiovascular risk factors in type 2 diabetic subjects asymptomatic for coronary heart disease (The PREDICT Study). *Diabet Med* 21:1129–1134
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215
- Tregouet DA, Escolano S, Tiret L, Mallet A, Golmard JL (2004) A new algorithm for haplotype-based association analysis: the Stochastic-EM algorithm. *Ann Hum Genet* 68:165–177
- Tregouet DA, Tiret L (2004) Cox proportional hazards survival regression in haplotype-based association analysis using the Stochastic-EM algorithm. *Eur J Hum Genet* 12:971–974
- Greenland S, Drescher K (1993) Maximum likelihood estimation of the attributable fraction from logistic models. *Biometrics* 49:865–872
- Gable DR, Stephens JW, Cooper JA, Miller GJ, Humphries SE (2005) Variation in the UCP2-UCP3 Gene cluster predicts the development of type 2 diabetes in the Northwick park Heart study II. *Diabetes* 55:1504–1511
- Wong NACS, Pignatelli M (2002) b-Catenin—a linchpin in colorectal carcinogenesis? *Am J Pathol* 160:389–401
- Duval A, Rolland S, Tubacher E, Bui H, Thomas G, Hamelin R (2000) The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. *Cancer Res* 60:3872–3879
- Korinek V, Barker N, Moerer P et al (1998) Depletion of epithelial stem cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet* 19:379–383
- Heller RS, Dichmann DS, Jensen J et al (2002) Expression patterns of Wnts, Frizzleds, sFRPs, and misexpression in transgenic mice suggesting a role for Wnts in pancreas and foregut pattern formation. *Dev Dyn* 225:260–270
- Yi F, Brubaker PL, Jin T (2005) TCF-4 mediates cell type-specific regulation of proglucagon gene expression by beta-catenin and glycogen synthase kinase-3beta. *J Biol Chem* 280:1457–1464
- Humphries SE, Hawe E, Dhamrait S, Miller GJ, Talmud PJ (2003) In search of genetic precision. *Lancet* 361:1908–1909
- Carulli L, Rondinella S, Lombardini S, Canedi P, Loria P, Carulli N (2005) Diabetes, genetics and ethnicity. *Aliment Pharmacol Ther* 22(S2):16–19
- Abate N, Carulli L, Cabo-Chan A, Chandalia M, Snell PG, Grundy SM (2003) Genetic polymorphism PC-1 K1221Q and ethnic susceptibility to insulin resistance. *J Clin Endocrinol* 88:5297–5934
- Rahda V, Vimalaswaran KS, Babu HN et al (2006) Role of genetic polymorphism peroxisome proliferator-activated receptor-gamma2 Pro12Ala on ethnic susceptibility to diabetes in South-Asian and Caucasian subjects: evidence for heterogeneity. *Diabetes Care* 29:1046–1051