The Insulin Gene Variable Number Tandem Repeat Class I/III Polymorphism Is in Linkage Disequilibrium With Birth Weight but Not Type 2 Diabetes in the Pima Population

Robert S. Lindsay, Robert L. Hanson, Chris Wiedrich, William C. Knowler, Peter H. Bennett, and Leslie J. Baier

The insulin gene variable number tandem repeat (*INS*-VNTR) is proposed to exert pleiotropic genetic effects on birth weight and diabetes susceptibility. In our study, we examined the influence of a polymorphism in tight linkage disequilibrium with *INS*-VNTR (−23*Hph*1) on birth weight and type 2 diabetes in the Pima population. A parent-offspring "trio" design was used to assess parent-of-origin effects and population stratification. The presence of the –23*Hph*1 T-allele was associated with lower birth weight (n = 192; −140 g per copy of the T-allele; P = 0.04), even after adjustment for effects of population stratification (P = 0.03). The effects of paternally transmitted T-alleles were greater than those of maternally transmitted alleles (paternally transmitted: –250 g, P = 0.05; maternally transmitted: –111 g, P = 0.43), but this difference was not statistically significant (P = 0.50). The −23*Hph*1 T-allele was associated with an increased prevalence of type 2 diabetes (P = 0.009), which family-based association analysis suggested was attributable to population structure (P = 0.04) without significant evidence of linkage disequilibrium between diabetes prevalence and genotype (P = 0.86). Thus allelic variation of the *INS* gene is associated with lower birth weight and increased prevalence of type 2 diabetes. Significant linkage disequilibrium was found between –23*Hph*1 and birth weight but not type 2 diabetes, an observation that supports a potential functional role of *INS* polymorphisms in the regulation of birth weight. *Diabetes* 52:187–193, 2003

Low birth weight is associated with increased risk of type 2 diabetes (1). It has been suggested that this may be because of pleiotropic effects of one or more genes to both increase susceptibility to type 2 diabetes and lower birth weight (2). The insulin gene (*INS*) has been proposed as a potential candidate for such effects, as insulin is a promoter of growth in utero and insulin secretion and action are central to the pathogenesis of type 2 diabetes (2). Studies examining genetic variations of *INS* in relation to disease development have concentrated on the *INS*-VNTR (variable number of tandem repeats), a polymorphic region ~0.5 kb upstream of the insulin gene, comprising a variable number of repeats of an oligonucleotide sequence (ACAGGGGT(G/C)(T/C)GGGG) (3). Depending on the number of repeats, *INS*-VNTR is described as class I (26–63 repeats, average 570 bp), II (64–140 repeats, average 1,200 bp), or III (141–209 repeats, average 2,200 bp) (3). Polymorphism of the *INS*-VNTR is associated with changes in insulin expression in the pancreas (4,5) and thymus (6), and the presence of the class I allele is associated with increased risk of type 1 diabetes in Caucasian populations (7).

Results of studies examining the association of *INS*-VNTR and type 2 diabetes have been inconsistent (7); however, a collation of five of these studies (carried out in Caucasians) suggested that class III *INS*-VNTR conferred a modest increase in risk of type 2 diabetes (7). More recently, a large study using family-based association methods reported no significant overall association of *INS*-VNTR with type 2 diabetes (8). The latter study did, however, find an excess transmission of class III *INS*-VNTR from fathers to diabetic probands, which would suggest that parent-of-origin effects, such as imprinting, might be at work (8). Similarly, an increase in the risk of polycystic ovarian syndrome has been reported with paternal transmission of the class III allele (9). The potential for parent-of-origin effects is biologically plausible; *INS*-VNTR lies in close proximity to genes known to be imprinted in humans, notably *INS* itself (generally biallelically expressed in pancreas, but imprinted in the fetal yolk sac) (4,10) and IGF-2 (11). Furthermore, if there is a genetic connection between lower birth weight and type 2 diabetes, then imprinting effects should be considered. The majority of known imprinted genes influence growth in early life or placental development (11). In contrast to studies of type 2 diabetes, studies of association of birth weight to *INS*-VNTR have not taken account of potential imprinting effects. Higher birth weight associated with the class III allele has been reported (12); however, this effect was not found in all subjects, but restricted to those changing weight centiles by <0.67 SD score between birth and 2 years.

From the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, Arizona.

Address correspondence to Dr. Robert L. Hanson, National Institute of Diabetes and Digestive and Kidney Diseases, 1550 East Indian School Rd., Phoenix, AZ 85014. E-mail: rhanson@phx.niddk.nih.gov.

Received for publication 11 June 2002 and accepted in revised form 23 September 2002.

LD, linkage disequilibrium; OGTT, oral glucose tolerance test; SNP, single nucleotide polymorphism; VNTR, variable number tandem repeat; WHO, World Health Organization.

The Pima Indians of Arizona have a marked propensity to type 2 diabetes and obesity (13,14). In Pima Indians, both lower and higher birth weight are associated with increased risk of type 2 diabetes (15). Paternal diabetes is also associated with lower birth weight of offspring (16), an observation that supports the presence of a pleiotropic genetic effect influencing birth weight and type 2 diabetes in this population and suggests a particular importance for paternally derived alleles in this relationship (16). In addition, linkage analyses in the Pima Indians indicate an important genetic contribution to birth weight, with evidence that an imprinted, paternally expressed gene on chromosome 11 influences birth weight (17); however, the peak of linkage (LOD 4.1) was 70–80 cM distant from *INS*.

Results of a previous study of a small group of Pima Indians did not support a relation of *INS*-VNTR to type 2 diabetes susceptibility (18). In the current study, we wished to assess the relation of *INS*-VNTR to birth weight, type 2 diabetes, and adiposity, allowing for the potential imprinting effects. To achieve this, we examined –23*Hph*1, an A/T polymorphism within the *Hph*1 restriction endonuclease site 23 bp upstream from the start site of the insulin gene. The polymorphism is in tight linkage disequilibrium with *INS*-VNTR in Caucasian populations (7) and has been used extensively in studies addressing variations of *INS*-VNTR. We used a parent-offspring trio design because, first, where an association is present, it allows an assessment of whether this arises secondary to linkage disequilibrium or to other effects, such as population stratification (19); and second, the presence of parent-of-origin effects can be addressed. We tested the following hypotheses: *1*) –23*Hph*1 polymorphisms are associated with birth weight with or without imprinting effects, and *2*) –23*Hph*1 polymorphisms are associated with type 2 diabetes or BMI with or without imprinting effects.

RESEARCH DESIGN AND METHODS

Subjects and phenotypes. The subjects in this study were participants in the longitudinal study of health in the Pima population whose heritage is at least half Pima or Tohono O'odham or a mixture of these two closely related groups. Members of the community over age 5 years are invited to participate in biennial research examinations, including an oral glucose tolerance test (OGTT) after a 75-g glucose load. Diabetes was diagnosed by World Health Organization (WHO) criteria (20). Owing to our interest in imprinting effects, we used a "trio" design, selecting families for whom DNA was available from both parents and one or more offspring and whose offspring had been examined at least once for the presence of type 2 diabetes. A total of 660 subjects were included, comprising 418 children from 130 nuclear families (median 3 siblings, range 1–11). No half-siblings were included. For 18 subjects, children of one family were parents in another family; those individuals were also included.

Birth weight and gestational age were ascertained by a review of medical records. Subjects were included if they were singleton births, both their birth weight and their gestational age were available, and their gestational age was ≥33 weeks. Given that the presence of maternal diabetes influences birth weight because of environmental effects of maternal hyperglycemia on the fetus in utero (21) and, further, that a small secular increase in birth weight is described in this population (22), adjustments for maternal diabetes and year of birth were made. Children were considered offspring of diabetic mothers (DIAB) if their mother had been diagnosed as having type 2 diabetes according to WHO criteria (20), either before or during the index pregnancy; offspring of nondiabetic mothers (NON) if their mother was known to have a nondiabetic OGTT after the index pregnancy; offspring of mothers who were possibly diabetic (POSS), if maternal diabetes had been diagnosed postnatally without an intervening nondiabetic OGTT after the birth of the child, thus rendering the exact time of onset of maternal diabetes uncertain; and offspring of unknown status (UNK), where maternal diabetes status had not been assessed. Before analysis, birth weights were adjusted for sex, gestational age (as linear and quadratic terms), and birth year for all analyses, using data from 4,076 births in the Pima population by means of linear regression and either with or without further adjustment for maternal diabetes status as detailed (with indicator variables for the different categories representing 246 DIAB, 2,943 NON, 304 POSS, and 583 UNK subjects). Birth weights (mean ± SD) for the 192 offspring included in the present genetic analyses were NON 3.7 ± 0.5 kg (n = 136), POSS 3.9 ± 0.5 kg (n = 35), DIAB 4.0 ± 0.6 kg (n = 19), and UNK 3.3 ± 0.1 kg (n = 2).

Informed consent for participation in the study was obtained from all subjects, and ethical approval was received from both the National Institutes of Health and the Gila River Indian Community.

Sequencing of the –23*Hph*1 single nucleotide polymorphism in Pima subjects. The A to T polymorphism that disrupts the *Hph*1 site CCATCT at nucleotide 2,401 of the insulin gene (accession #V00565) was sequenced in the 660 members of parent-child "trios" described above and in 77 individuals previously typed for *INS*-VNTR (18,23). The PCR primers forward 5′-TCA GAAGAGGCCATCAAGCAGGT-3′ and reverse 5′-CGCACAGGTGTTGGTTCA CAAA-3′ were used to amplify a 574-bp fragment from genomic DNA, and the forward primer was further used for cycle sequencing. Sequencing was performed using the Big Dye Terminator (Applied Biosystems) on an automated DNA capillary sequencer (Model 3700; Applied Biosystems).

Analysis. Evidence of linkage disequilibrium between –23*Hph*1 and *INS*-VNTR was assessed using the EH algorithm (24), which was also used to estimate haplotype frequencies of the two polymorphisms.

The association of –23*Hph*1 genotype to traits was examined by a modification of the approach of Fulker et al. (25). It is important to note that in our design, the association of marker and trait could potentially be influenced by family membership (statistical nonindependence of siblings) and population stratification. A series of models were therefore assessed.

In model 1, "simple association," the additive influence of the genotype on a trait was assessed by modeling a genotype score (calculated from the number of T-alleles) against the trait. To account for family membership, the relation of genotype to trait was examined in a mixed model in which membership of nuclear family was modeled as a random effect. This random effect accounted for nonindependence among siblings, but the results of this model may still be confounded by population stratification.

In model 2, "linkage disequilibrium" (LD), the hypothesis of joint linkage and association of polymorphisms against trait, was examined. Association effects were partitioned into between- and within-family components (25). The between-family component is potentially influenced by population stratification, whereas the within-family effect assesses the presence of joint linkage and association (linkage disequilibrium). In brief, the genotype score for each subject is broken down into orthogonal between-family (b; derived from expected number of T-alleles for offspring conditional upon parental genotype) and within-family scores (w; derived from the deviation of offspring genotype score from the between-family score). Between- and within-family scores are then modeled (with additional random effect of nuclear family) against the trait.

In model 3, "imprinting (Imp) LD," the hypothesis of linkage and association of polymorphisms inherited from either father or mother against trait, was examined. Between-family scores were calculated as above. The within-family component was further broken down into terms reflecting paternal (w_{pa}) or maternal (w_{ma}) origin of inherited T-alleles. Thus w_{pa} was calculated from the number of T-alleles inherited from the father minus the expected number given paternal genotype (1/2 paternal genotype score). The maternal within-family component was similarly derived. It is important to note that w_{pa} and w_{ma} cannot be calculated when all three members of the trio (mother, father, and offspring) are heterozygous, as the source T-allele of the offspring is ambiguous; where this was the case, the subjects were excluded from model 3. In addition, subjects can only be informative when either the father or the mother is heterozygous. Scores b, w_{pa}, and w_{ma} are then modeled (with additional random effect of nuclear family) against the trait. The hypothesis that T-alleles exert different influences on trait depending on the parent of origin was assessed by deriving a P value (P_{diff}) by two-tailed t test of the difference of the β coefficients of w_{pa} and w_{ma} divided by its standard error (derived from the standard errors of w_{pa} and w_{ma} and their covariance).

The above framework was used for analyzing offspring birth weight and BMI (kg/m^2). Although the analysis of each trait with all three models entailed some degree of multiple testing, the models were highly correlated with one another, as any observed association was partitioned into different sources. Thus P values are presented without correction for multiple testing. Because the presence of type 2 diabetes may influence BMI (26), data from examinations at or after diagnosis of type 2 diabetes were excluded from this analysis. BMI was analyzed using *1*) the measurement from the earliest adult (age >18 years) examination and *2*) one measure each from five age ranges (ages 5–9, 10–14, 15–19, 20–29, and 30–39 years, truncated at age 40 years because of the

TABLE 1
Frequency of diabetes and birth weight by –23 *Hph*1 genotype

	A/A	A/T	T/T	Total
n (%)				
Offspring + both parents	458 (69.4)	179 (27.2)	23 (3.5)	660
Offspring only	293 (70.1)	113 (27.0)	12 (2.9)	418
Sex (M/F)				
Offspring + both parents	212/245	72/117	13/10	660
Offspring only	126/167	45/68	7/5	418
Number with diabetes (%)				
Offspring + both parents	255 (55.6)	115 (64.3)	16 (69.6)	386
Offspring only	140 (47.8)	59 (52.2)	6 (50)	205
Age of diabetes onset (years)				
Offspring + both parents	33.7 ± 11.8	33.3 ± 12.1	31.4 ± 8.3	386
Offspring only	28.1 ± 9.1	26.9 ± 9.8	29.1 ± 9.8	205
Age at last exam (years)				
Offspring + both parents	40.9 ± 15.6	39.4 ± 16.5	36.6 ± 14.3	660
Offspring only	32.9 ± 10.6	30.1 ± 11.0	28.7 ± 13.2	418
Birth weight*				
n (%)	133 (69.3)	54 (28.1)	5 (2.6)	192
kg	3.73 ± 0.48	3.64 ± 0.52	3.62 ± 0.50	

Data are *n*, *n* (%), or means ± SD. *Birth weight (for offspring only) adjusted for gestational age, sex, year of birth, and maternal diabetes status; birth weight was not available for all subjects.

low number of observations in nondiabetic individuals older than age 40 years). For subjects who had been examined more than once within an age group, only the earliest examination was included.

The association of genotype with prevalence of type 2 diabetes was examined by logistic regression within the above framework, using diabetes status at the last examination. A generalized estimating equation model was used, with the inclusion of "repeated" term to allow for covariance among members of the same nuclear family. In addition, to explore the influences of genotype on disease prevalence at different ages, the prevalence for examinations in four age ranges (ages 10–19, 20–29, 30–39, and 40–49 years) was also examined. In general, because of the requirement for parental genotype, models 2 and 3 were applied to offspring only. By contrast, model 1 (simple association with adjustment for nuclear family membership) could be applied to offspring alone and offspring plus parents. Model 1 is presented for all offspring plus parents and offspring alone in the analysis of type 2 diabetes, but not birth weight (very few parental birth weights were available) or BMI (no substantial difference in results).

RESULTS

Linkage disequilibrium of –23*Hph*1 and *INS*-VNTR. In all, 77 unrelated individuals (40 I/I homozygotes, 35 I/III heterozygotes, and 2 III/III homozygotes) for whom DNA was available and who had previously been typed for *INS*-VNTR (18) were genotyped at –23 *Hph*1. *INS*-VNTR and –23*Hph*1 were in strong linkage disequilibrium ($P < 0.0001$), with the occurrence of haplotypes *INS*-VNTR class I/–23*Hph*1 A and *INS*-VNTR class III/–23*Hph*1 T being greater than expected under equilibrium, as previously described in Caucasian populations (7). An estimated 6.0% of haplotypes were discordant (2.0% class I/T and 4.0% class III/A) and 94% were concordant (73% class I/A and 21% class III/T).

Relation of –23 *Hph*1 polymorphism to diabetes. A higher proportion of heterozygotes and T/T homozygotes had developed diabetes than A/A homozygotes (Table 1). This difference approached significance in all subjects (parents and offspring combined: χ^2, $P = 0.08$), but not in offspring alone (χ^2, $P = 0.7$). In keeping with this, logistic regression of all subjects (model 1; $P = 0.009$) (Table 2) showed a significant association of the presence of the T-allele with type 2 diabetes. Analysis of offspring alone displayed a nonsignificant trend ($P = 0.23$) in the same direction. An important finding was that where between-

TABLE 2
Influence of –23*Hph*1 on prevalence of type 2 diabetes

	Cases/subjects (families)	Genotype effect	OR (95% CI)	*P*	P_{diff}
All subjects (Model 1: Association)	386/660 (130)	Additive	1.53 (1.11–2.11)	0.009	—
Offspring					
Model 1: Association	205/418 (130)	Additive	1.27 (0.86–1.85)	0.23	—
Model 2: LD	205/418 (130)	Between	1.84 (1.02–3.32)	0.04	—
		Within family	0.95 (0.53–1.70)	0.86	—
Model 3: Imp LD	197/403 (129)*	Between	1.99 (1.12–3.52)	0.02	
		Within family maternal	0.90 (0.36–2.23)	0.83	0.83
		Within family paternal	1.04 (0.40–2.67)	0.93	

Effect of simple association of genotype (model 1) and models examining linkage disequilibrium without (model 2) or with (model 3) influences of parent of origin of T-allele. OR is given per additional T-allele (model 1 and between-family effects in models 2 and 3) or per additional T-allele above that expected given parental genotype (within-family effects in models 2 and 3). The joint hypothesis of linkage and association is examined by "within-family" effects. The presence or absence of diabetes was modeled using logistic regression with fixed effects of sex, age at examination, birth year, and effect of nuclear family in all models. *Number is reduced as inheritance of offspring T-allele in triple heterozygote trios (mother, father, and child) cannot be defined.

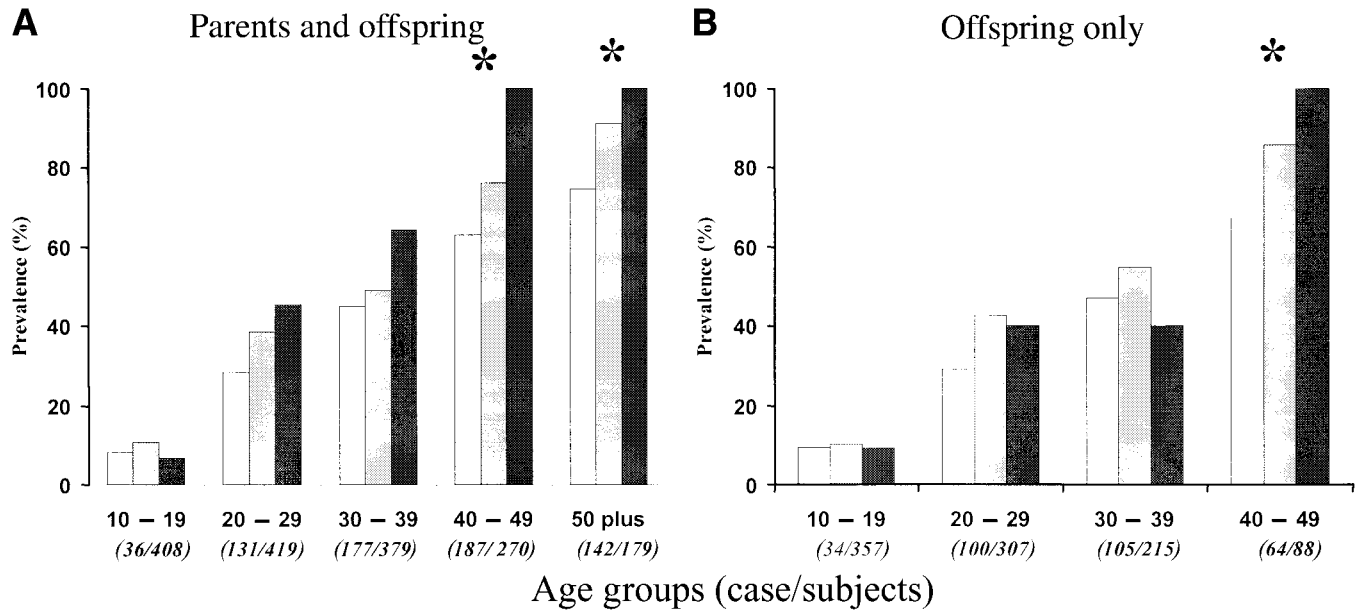

FIG. 1. The −23 *Hph*1 genotype and prevalence of type 2 diabetes in Pima Indians. Data from all 660 members of the parent-offspring trios (*A*) and 418 offspring alone (*B*) are shown. □, AA; ▨, AT; ■, TT. **P* < 0.05 for effect of genotype in logistic regression with age at examination, sex, and year of birth as additional predictors.

and within-family effects were modeled in all offspring either without (model 2) or with parent-of-origin effects (model 3), there was no evidence of linkage disequilibrium between –23*Hph*1 and diabetes. Contrasting with this, between-family effects were significant in both models (model 2, $P = 0.04$; model 3, $P = 0.02$), in keeping with the presence of population stratification. The influence of the –23*Hph*1 genotype on diabetes prevalence was examined in subgroups classified based on the age at examination. The simple association (model 1) of –23*Hph*1 genotype and diabetes prevalence was most apparent in the older age groups (Fig. 1). In offspring, significant linkage disequilibrium between the –23*Hph*1 genotype and diabetes prevalence was not observed in any of the age groups (model 2; data not shown). Potential evidence of imprinting effects (model 3; $P_{diff} < 0.05$) was seen only in the youngest age group, with a protective effect of paternally transmitted T-alleles (age 10–19 years; odds ratio [OR] 0.05, 95% confidence interval [CI] 0.01–0.23; $P = 0.001$ for effect of paternal alleles; $P_{diff} = 0.0005$), although this group contained relatively few cases (34 of 342 subjects).

The results were not substantially altered by including terms for maternal diabetes or the fraction of Pima heritage in the above models (data not shown). Further, analysis of diabetes restricted to those of full Pima/Tohono O'odham heritage (153 events in 311 subjects from 96 families) confirmed an increased risk of type 2 diabetes associated with the –23*Hph*1 T-allele (model 1; OR for effect of presence of T-allele 1.54 [CI 0.94–2.54]; $P = 0.06$), again without evidence of linkage disequilibrium (model 2; OR for within-family effect, 1.02 [CI 0.54–1.92], $P = 0.9$; for between-family effect, 3.25 [CI 1.39–7.61], $P = 0.006$).

Relation of –23 *Hph*1 polymorphism to birth weight. The presence of the T-allele was significantly associated (model 1; $P = 0.04$) (Table 3) with lower birth weight (adjusted for sex, gestational age, birth year, and maternal diabetes status). Family-based association analysis (model 2) supported significant linkage disequilibrium between

TABLE 3
Influence of –23*Hph*1 on birth weight

Number of subjects (nuclear families)	Number informative	Genotypes AA	AT	TT	Model	Effect	β	*P*
192 (92)	192	133	54	5	1	T-allele	–140	0.046
192 (92)	87	42	42	3	2	Between family	–61	0.67
						Within-family	–190	0.03
187 (92)*	82	133	49	5	3	Between-family	–39	0.73
						Paternal (w_{pa})	–250	0.05
						Maternal (w_{ma})	–111	0.43

Effect of simple association of genotype (model 1) and models examining linkage disequilibrium without (model 2) or with (model 3) influences of parent of origin of T-allele. The joint hypothesis of linkage and association is examined by "within family" effects. β is given as effect of each additional T-allele (model 1 and between-family effects in models 2 and 3) or per additional T-allele above that expected, given parental genotype (within family effects, models 2 and 3). All models include random effect of nuclear family. Birth weight is adjusted for sex, gestational age at birth, birth year, and maternal diabetes. *Number is reduced as inheritance of offspring T-allele in triple heterozygote trios (mother, father, and child) cannot be defined.

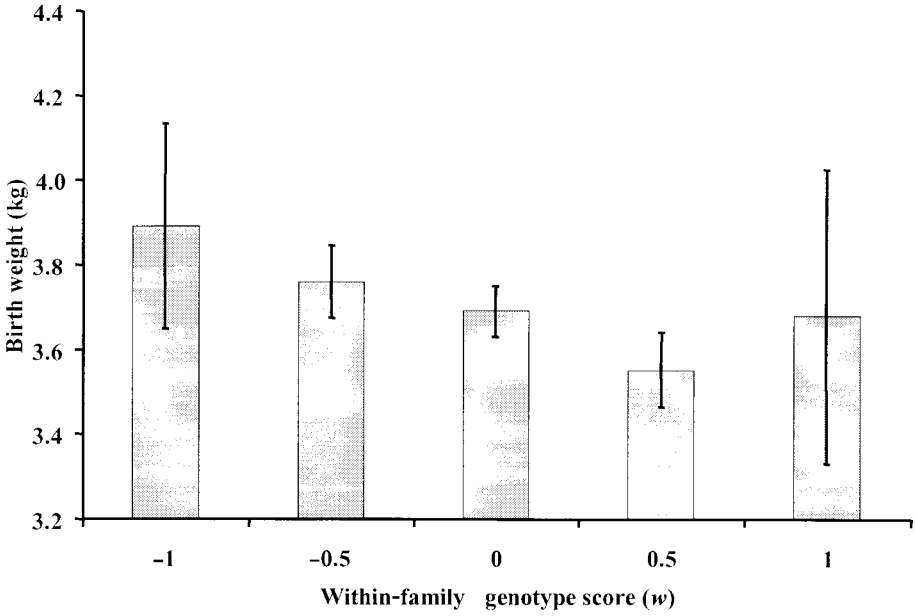

FIG. 2. Effect of transmission of T-alleles within family on birth weight. Birth weights (means ± SE) were adjusted for gestational age, sex, year of birth, maternal diabetes, between-family association, and nuclear family membership (as a random effect). The within-family association score (w) is derived from the deviation of the offspring genotype score (number of −23*Hph*1 T-alleles) from average parental genotype. For example, for two heterozygous parents, a child having genotype AA has $w = -1$, with AT has $w = 0$, and with TT has $w = 1$. The total number of offspring was 192 (n = 4, 39, 110, 37, and 2 for $w = -1$, −0.5, 0, 0.5 and 1, respectively). $P = 0.03$ for effect of w on birth weight.

the T-allele and lower birth weight ($P = 0.03$) (Fig. 2). The influence of paternally transmitted T-alleles showed a greater effect than that of maternally transmitted alleles (–250 vs. −111, respectively), but this difference was not statistically significant ($P_{diff} = 0.50$).

Analysis of birth weight unadjusted for the presence of maternal diabetes showed a similar trend in both the simple association (model 1; difference in birth weight associated with one allele difference in number of T-alleles –116 g; $P = 0.11$) and the model examining linkage disequilibrium (model 2; within-family effect –178 g; $P = 0.047$). Again, no significant parent-of-origin effect was observed (data not shown for model 3; $P_{diff} = 0.69$).

We examined whether the presence of maternal diabetes might influence the relation of –23*Hph*1 to birth weight. The number of offspring from mothers who had diabetes during pregnancy was small ($n = 19$). A subanalysis of subjects depending on time of onset of maternal diabetes after birth of the child suggested that the negative influence of the T-allele was greater when mothers had developed diabetes within a few years after the birth of their child (Table 4). In these groups, the influence of paternally transmitted T-alleles also tended to be greater (e.g., offspring of mothers with diabetes before or up to 5 years after birth: $n = 45$, β coefficient w_{pa} –862 vs. w_{ma} 233, $P_{diff} = 0.04$; offspring of mothers with diabetes before or up to 10 years after birth: $n = 80$, β coefficient w_{pa} –596 vs. w_{ma} 17, $P_{diff} = 0.07$).

Relation of –23 *Hph*1 polymorphism to BMI. No simple association (model 1) of BMI to genotype was observed for all subjects with an examination in adulthood ($P = 0.77$) or in the subgroup analysis by age (data not shown). Similarly, there was no evidence of linkage disequilibrium between −23*Hph*1 and BMI in all adult subjects either without (model 2; effect of w, $P = 0.58$) or allowing for parent-of-origin effects (model 3; w_{pa}, $P = 0.75$, w_{ma}, $P = 0.93$). No evidence of association of −23*Hph*1 and BMI secondary to population stratification was observed (*b*, model 2 or 3; data not shown).

DISCUSSION

Our data indicated that presence of a T-allele at –23*Hph*1 is associated with lower birth weight. The −23*Hph*1 polymorphism itself is not thought to be functional, but has been used as a more easily defined indicator of *INS*-VNTR polymorphism. This reflects the tight linkage disequilibrium between −23*Hph*1 and *INS*-VNTR class

TABLE 4
Within-family association of –23*Hph*1 and birth weight according to maternal diabetes

Interval	Diabetes status of mother	n	w	P
Before birth of child	Diabetes	19	–91	0.84
	No diabetes	136	–187	0.09
Before or up to 5 years after birth of child	Diabetes	45	–347	0.12
	No diabetes	104	–143	0.27
Before or up to 10 years after birth of child	Diabetes	80	–321	0.03
	No diabetes	73	–89	0.46
Before or up to 15 years after birth of child	Diabetes	107	–299	0.01
	No diabetes	58	–103	0.40

Diabetes status is defined as "diabetes" when mother was diagnosed either before or up to 5, 10, or 15 years after birth of the child, and as "no diabetes" when mother was examined and found not to be diabetic after these intervals (e.g., after birth of child). Total numbers are less than those in Table 4, as this classification excludes some mothers (e.g., where diabetes has not been diagnosed, but examination has not taken place at a suitable interval after birth of the child). Birth weight is adjusted for sex, gestational age, and year of birth. β is given as effect of each T-allele above that expected given parental genotype (within-family effect, model 2).

I/III in Caucasian populations (7). In the Pima Indians, we have confirmed that −23*Hph*1 and *INS*-VNTR are in linkage equilibrium, albeit with a haplotype discordance rate of 6%, which is higher than that previously observed in Caucasians (0.023%) (7). Nevertheless, −23*Hph*1 typing gives a reliable indication of *INS*-VNTR status in this population.

INS-VNTR, unlike −23*Hph*1, does appear to be a functional polymorphism. Transfection of rodent β-cell lines with reporter constructs of the *INS* promoter and either class I or class III *INS*-VNTR results in alteration of the level of transcription (27,28). The class III allele is associated with a reduction of insulin transcription in the fetal (4) and adult pancreas (5). The observation that *INS*-VNTR acts as a binding site to the transcription factor Pur1 (27) forms a potential basis of these functional relationships (29).

We have presented evidence of significant linkage disequilibrium between birth weight and the *INS* gene –23*Hph*1 polymorphism, with the T-allele being associated with lower birth weight. Although our data did not allow us to determine whether this association was attributable to a direct effect of the *INS*-VNTR or to linkage disequilibrium with a nearby functional polymorphism, the former hypothesis is biologically plausible. Insulin promotes growth in utero (2). Therefore, if the class III allele results in reduced pancreatic expression of insulin, it in turn might lead to reduced growth in utero. This is also in keeping with the suggestion that obese children who are class I/I homozygotes have increased insulin secretion in response to their obesity (30). *INS*-VNTR class III has also been associated with decreased expression of IGF-2 in placenta (31), forming a further potential mechanism whereby this polymorphism might result in altered fetal growth. Although such mechanisms show potential, it is important to ascertain whether the associations we observed are present in other populations. At least in Caucasians, little effect of *INS*-VNTR has been observed, and even a positive effect has been seen in selected subgroups, although only simple associations were examined (12).

Fetal insulin production is greatly increased in the presence of maternal diabetes during pregnancy, and is therefore of particular importance in the offspring of mothers with abnormal glucose metabolism during pregnancy (32). Consequently, we examined influences of the –23*Hph*1 polymorphism when maternal diabetes was present during or detected shortly after pregnancy (in the latter case, it being likely that abnormal glucose metabolism was present during pregnancy). We found that the relationship of birth weight to –23*Hph*1 appeared greater in these subgroups, and indeed that there was the suggestion of a greater effect of paternally transmitted T-alleles. As we do not have detailed information regarding the glycemia of mothers during pregnancy to ascertain whether an interaction between maternal glycemia and fetal genotype might have taken place, we must interpret these results with caution.

We did not find evidence of linkage disequilibrium between *INS*-VNTR and type 2 diabetes, in keeping with a previous report in a Caucasian population (8). Overall, we observed a significant association of the –23*Hph*1 T-allele with increased prevalence of type 2 diabetes, a finding also reported in collated results of smaller series (7). Our results, however, suggested that in this population, this association reflects population stratification rather than a functional association. One potential source of such population stratification is admixture with other ethnic groups. Using self-reported estimates of Pima/Tohono O'odham heritage (which correspond well with genetic estimates of admixture) (33), we found that our results were consistent, even when analysis was restricted to those of full Pima/Tohono O'odham heritage. This suggests that ethnic admixture, at least as represented by Pima versus non-Pima heritage, is not the source of the association of –23*Hph*1 genotype and type 2 diabetes. It would also suggest that the Pima population is heterogeneous both in terms of diabetes risk and –23*Hph*1 genotype, perhaps as a result of an older population admixture in the Pima, although this remains entirely speculative.

Both *INS* (in the fetal yolk sac) (10) and *IGF2* are imprinted with expression from paternally derived alleles (11). In addition, results of a previous study in Caucasians (8) suggest that paternal transmission of the class III *INS*-VNTR allele is associated with increased risk of type 2 diabetes. Apart from the results of subgroup analyses (which are relatively small and introduce potential problems of multiple testing), we found no clear evidence of a parent-of-origin effects of the –23*Hph*1 locus on diabetes or birth weight.

As part of this investigation, we examined the hypothesis that *INS*-VNTR might provide a genetic explanation for the association of low birth weight and type 2 diabetes; more specifically, that the presence of *INS*-VNTR class III would be associated with lower birth weight and increased risk of type 2 diabetes. Taken together, our results only partly support this hypothesis. Although –23*Hph*1 T, and by inference *INS*-VNTR class III, does appear to be associated with lower birth weight, we did not find significant evidence to support a functional effect of the class III allele on diabetes susceptibility.

Class I *INS*-VNTR has also been associated with an increase in childhood obesity, an effect particularly associated with paternally transmitted class I allele (34). This association has not been apparent in adults (8) and was not apparent in our series.

We used the –23*Hph*1 polymorphism as a means of assessing the *INS*-VNTR. Although supported by linkage disequilibrium between –23*Hph*1 and *INS*-VNTR, our data suggest that a small number (~6%) of haplotypes will be misclassified, thereby potentially weakening the association if *INS*-VNTR is indeed of functional importance to variations in birth weight. Of greater concern is the potential oversimplification of variations of *INS*-VNTR by categorization into class I and class III. This is convenient for studies such as ours and allows comparison with previous population and functional studies. Nevertheless, it is important to note that *INS*-VNTR is extremely polymorphic; subclassifications based on the size of *INS*-VNTR have been proposed (5), and sequence variation within the repeat sequences is also present, with potential functional consequences. For example, it has been reported that certain subclasses of the class I *INS*-VNTR are protective against type 1 diabetes (class I generally being a risk factor) (5). The reasons for this remain unclear and

perhaps await a clearer understanding of the functional relationship of *INS*-VNTR to the transcription of *INS* and other genes before this can be fully explored at the population level.

In conclusion, we have presented evidence that the −23*Hph*1 T-allele of the *INS* gene (and by implication class III *INS*-VNTR) is associated with both lower birth weight and prevalence of type 2 diabetes. Furthermore, our assessment with family-based association analysis supported the presence of linkage disequilibrium with birth weight. These findings suggest a potential functional relationship of polymorphisms in INS with birth weight, but suggest the relationship with type 2 diabetes, if any, is modest.

ACKNOWLEDGMENTS

We thank the members of the Gila River Indian Community for their continued support and participation, and the staff of the Diabetes and Arthritis Epidemiology Section for conducting this study.

REFERENCES

1. Barker DJ: *Fetal and Infant Origins of Adult Disease*. London, British Medical Journal, 1992
2. Hattersley AT, Tooke JE: The fetal insulin hypothesis: an alternative explanation of the association of low birth weight with diabetes and vascular disease. *Lancet* 353:1789–1792, 1999
3. Bell GI, Selby MJ, Rutter WJ: The highly polymorphic region near the human insulin gene is composed of simple tandemly repeating sequences. *Nature* 295:31–35, 1982
4. Vafiadis P, Bennett ST, Colle E, Grabs R, Goodyer CG, Polychronakos C: Imprinted and genotype-specific expression of genes at the IDDM2 locus in pancreas and leucocytes. *J Autoimmun* 9:397–403, 1996
5. Bennett ST, Lucassen AM, Gough SC, Powell EE, Undlien DE, Pritchard LE, Merriman ME, Kawaguchi Y, Dronsfield MJ, Pociot F: Susceptibility to human type 1 diabetes at IDDM2 is determined by tandem repeat variation at the insulin gene minisatellite locus. *Nat Genet* 9:284–292, 1995
6. Pugliese A, Zeller M, Fernandez A Jr, Zalcberg LJ, Bartlett RJ, Ricordi C, Pietropaolo M, Eisenbarth GS, Bennett ST, Patel DD: The insulin gene is transcribed in the human thymus and transcription levels correlated with allelic variation at the INS VNTR-IDDM2 susceptibility locus for type 1 diabetes. *Nat Genet* 15:293–297, 1997
7. Bennett ST, Todd JA: Human type 1 diabetes and the insulin gene: principles of mapping polygenes. *Annu Rev Genet* 30:343–370, 1996
8. Huxtable SJ, Saker PJ, Haddad L, Walker M, Frayling TM, Levy JC, Hitman GA, O'Rahilly S, Hattersley AT, McCarthy MI: Analysis of parent-offspring trios provides evidence for linkage and association between the insulin gene and type 2 diabetes mediated exclusively through paternally transmitted class III variable number tandem repeat alleles. *Diabetes* 49:126–130, 2000
9. Bennett ST, Todd JA, Waterworth DM, Franks S, McCarthy MI: Association of insulin gene VNTR polymorphism with polycystic ovary syndrome. *Lancet* 349:1771–1772, 1997
10. Moore GE, Abu-Amero SN, Bell G, Wakeling EL, Kingsnorth A, Stanier P, Jauniaux E, Bennett ST: Evidence that insulin is imprinted in the human yolk sac. *Diabetes* 50:199–203, 2001
11. Efstratiadis A: Genetics of mouse growth. *Int J Dev Biol* 42:955–976, 1998
12. Dunger DB, Ong KK, Huxtable SJ, Sherriff A, Woods KA, Ahmed ML, Golding J, Pembrey ME, Ring S, Bennett ST, Todd JA: Association of the INS VNTR with size at birth. ALSPAC Study Team: Avon Longitudinal Study of Pregnancy and Childhood. *Nat Genet* 19:98–100, 1998
13. Knowler WC, Pettitt DJ, Saad MF, Bennett PH: Diabetes mellitus in the Pima Indians: incidence, risk factors and pathogenesis. *Diabetes Metab Rev* 6:1–27, 1990
14. Knowler WC, Pettitt DJ, Saad MF, Charles MA, Nelson RG, Howard BV, Bogardus C, Bennett PH: Obesity in the Pima Indians: its magnitude and relationship with diabetes. *Am J Clin Nutr* 53:1543S–1551S, 1991
15. McCance DR, Pettitt DJ, Hanson RL, Jacobsson LT, Knowler WC, Bennett PH: Birth weight and non-insulin dependent diabetes. Thrifty genotype, thrifty phenotype, or surviving small baby genotype? *BMJ* 308:942–945, 1994
16. Lindsay RS, Dabelea D, Roumain J, Hanson RL, Bennett PH, Knowler WC: Type 2 diabetes and low birth weight: the role of paternal inheritance in the association of low birth weight and diabetes. *Diabetes* 49:445–449, 2000
17. Lindsay RS, Kobes S, Knowler WC, Hanson RL: Genome-wide linkage analysis assessing parent-of-origin effects in the inheritance of birth weight. *Hum Genet* 110:503–509, 2002
18. Knowler WC, Pettitt DJ, Vasquez B, Rotwein PS, Andreone TL, Permutt MA: Polymorphism in the 5′ flanking region of the human insulin gene: relationships with noninsulin-dependent diabetes mellitus, glucose and insulin concentrations, and diabetes treatment in the Pima Indians. *J Clin Invest* 74:2129–2135, 1984
19. Spielman RS, Ewens WJ: The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet* 59:983–989, 1996
20. World Health Organization: *Diabetes Mellitus: Report of a WHO Study Group*. Geneva, World Health Org., 1985 (Tech. Rep. ser., no. 727)
21. Pettitt DJ, Knowler WC, Baird HR, Bennett PH: Gestational diabetes: infant and maternal complications of pregnancy in relation to third-trimester glucose tolerance in the Pima Indians. *Diabetes Care* 3:458–464, 1980
22. Lindsay RS, Hanson RL, Bennett PH, Knowler WC: Secular trends in birth weight, BMI, and diabetes in the offspring of diabetic mothers. *Diabetes Care* 23:1249–1254, 2000
23. Rotwein PS, Chirgwin J, Province M, Knowler WC, Pettitt DJ, Cordell B, Goodman HM, Permutt MA: Polymorphism in the 5′ flanking region of the human insulin gene: a genetic marker for non-insulin-dependent diabetes. *N Engl J Med* 308:65–71, 1983
24. Terwilliger J, Ott J: *Handbook for Human Genetic Linkage*. Baltimore, Johns Hopkins University Press, 1994
25. Fulker DW, Cherny SS, Sham PC, Hewitt JK: Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 64:259–267, 1999
26. Looker HC, Knowler WC, Hanson RL: Changes in BMI and weight before and after the development of type 2 diabetes. *Diabetes Care* 24:1917–1922, 2001
27. Kennedy GC, German MS, Rutter WJ: The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. *Nat Genet* 9:293–298, 1995
28. Lucassen AM, Screaton GR, Julier C, Elliott TJ, Lathrop M, Bell JI: Regulation of insulin gene expression by the IDDM associated, insulin locus haplotype. *Hum Mol Genet* 4:501–506, 1995
29. Lew A, Rutter WJ, Kennedy GC: Unusual DNA structure of the diabetes susceptibility locus IDDM2 and its effect on transcription by the insulin promoter factor Pur-1/MAZ. *Proc Natl Acad Sci U S A* 97:12508–12512, 2000
30. Le Stunff C, Fallin D, Schork NJ, Bougneres P: The insulin gene VNTR is associated with fasting insulin levels and development of juvenile obesity. *Nat Genet* 26:444–446, 2000
31. Paquette J, Giannoukakis N, Polychronakos C, Vafiadis P, Deal C: The INS 5′ variable number of tandem repeats is associated with IGF2 expression in humans. *J Biol Chem* 273:14158–14164, 1998
32. Freinkel N: Banting Lecture 1980: Of pregnancy and progeny. *Diabetes* 29:1023–1035, 1980
33. Williams RC, Knowler WC, Pettitt DJ, Long JC, Rokala DA, Polesky HF, Hackenberg RA, Steinberg AG, Bennett PH: The magnitude and origin of European-American admixture in the Gila River Indian Community of Arizona: a union of genetics and demography. *Am J Hum Genet* 51:101–110, 1992
34. Le Stunff C, Fallin D, Bougneres P: Paternal transmission of the very common class I INS VNTR alleles predisposes to childhood obesity. *Nat Genet* 29:96–99, 2001