Brief Genetics Report

Polymorphic Variations in the Neurogenic Differentiation-1, Neurogenin-3, and Hepatocyte Nuclear Factor-1α Genes Contribute to Glucose Intolerance in a South Indian Population

Alan E. Jackson, Paul G. Cassell, Bernard V. North, Shanti Vijayaraghavan, Susan V. Gelding, Ambady Ramachandran, Chamukuttan Snehalatha, and Graham A. Hitman

The neurogenic differentiation-1 (NEUROD1), neurogenin-3 (NEUROG3), and hepatic nuclear factor-1α (TCF1) genes are interacting transcription factors implicated in controlling islet cell development and insulin secretion. Polymorphisms of these genes (Ala45Thr [NEUROD1], Ser199Phe [NEUROG3], and Ala98Val [TCF1]) have been postulated to influence the development of type 2 diabetes. We have investigated the role and interaction between these variants using PCR/restriction fragment–length polymorphism assays in 454 subjects recruited as part of a population survey in South India. Additionally, 97 South Indian parent-offspring trios were studied. Polymorphisms of all three genes were associated with either fasting blood glucose (FBG) and/or 2-h blood glucose (BG) in either the total dataset or when restricted to a normoglycemic population. A monotonically increasing effect, dependent on the total number of risk-associated alleles carried, was observed across the whole population (P < 0.0001 for FBG and 2-h BG), raising FBG by a mean of 2.9 mmol/l and 2-h BG by a mean of 4.3 mmol/l. Similarly, an ascending number of the same risk alleles per subject increased the likelihood of type 2 diabetes (P = 0.002). In conclusion, we observed a combined effect of variations in NEUROD1, NEUROG3, and TCF1 in contributing to overall glucose intolerance in a South Indian population. Diabetes 53:2122–2125, 2004

Type 2 diabetes is a multifactorial disease with a significant genetic component. The genetic component of type 2 diabetes is postulated to be polygenic in nature; hence it is likely that a combination of genes would influence the underlying level of glucose intolerance in a population and contribute to the overall susceptibility to the disease. In a South Indian population, complex segregation analyses for type 2 diabetes have suggested that the best-fitting model included a polygenic influence on the range of glucose tolerance from normoglycemia to type 2 diabetes (65% of the model), whereas the rest was due to a major gene effect (1). This indicates that an additional approach to identify genes underlying the susceptibility to type 2 diabetes is to study the range of glucose intolerance, in addition to type 2 diabetes.

Susceptibility to disease in several monogenic forms of diabetes, such as maturity-onset diabetes of the young-3, are thought to result from defects in genes that code for a number of transcription factors. Many of these genes are potential candidate genes for type 2 diabetes because they are capable of modifying the rate of transcription of the insulin gene and are involved in β-cell developmental processes. Defects in either insulin secretion or in the formation and maintenance of β-cells could lead to both impaired glucose tolerance and type 2 diabetes. Several of these transcription factors are not only involved in a common biological pathway, but have also been shown to directly interact with each other. Therefore an alternative approach to elucidate an underlying cause of the disease may be to look for the combined effect of a selected group of these candidate genes in addition to an individual gene effect.

Three genes that are part of the β-cell differentiation transcriptional network are neurogenic differentiation-1 (NEUROD1), neurogenin-3 (NEUROG3), and hepatic nuclear factor-1α (TCF1). NEUROD1 is required for both normal islet development and the regulation of insulin gene transcription (2,3). Rare mutations in NEUROD1 have been recently characterized and associated with monogenic diabetes (4). Only one common polymorphism of NEUROD1 has been identified, Ala45Thr, which has been associated (5) with altered C-peptide secretion in type 2 diabetic subjects. NEUROG3 is an upstream “proendocrine” switch that is involved in the development of pancreatic endocrine cells by stimulating terminal differentiation genes including NEUROD1 (6,7). Therefore, NEUROG3 has a critical role...
in determining pancreatic β-cell mass and thereby may influence diabetes pathophysiology. In humans, *NEUROG3* has been located to a region on chromosome 10q in which linkage with type 2 diabetes has been demonstrated in Mexican Americans (8). The coding region of *NEUROG3* has been previously shown (9) to include only one common variant, Ser199Phe. In 249 Danish normoglycemic offspring of type 2 diabetic parents, carriers of the *NEUROG3* Ser199Phe variant showed higher fasting blood glucose (FBG) (P = 0.006), fasting serum insulin (P = 0.008), and fasting C-peptide (P = 0.006) compared with the common genotype (9).

The *TCF1* gene is the genetic determinant of the maturity-onset diabetes of the young-3 phenotype and is an important determinant of insulin gene transcription (10,11). There are three common *TCF1* missense variants resulting in amino acid substitutions, but only the Ala45Thr variant has been previously found (12) to associate with glucose intolerance. Heterozygous subjects from Denmark with *TCF1* Ala49Val and diabetes have been found to have decreased C-peptide (P = 0.004) and insulin (P = 0.03) levels compared with subjects who were normoglycemic (12,13).

The aim of the study was to determine whether the common mutations of these three transcription factors contribute individually or additively to the range of glucose tolerance and type 2 diabetes in subjects from South India. From the outset, we chose to study one common variant in each of the three genes (*NEUROD1* [Ala45Thr], *NEUROG3* [Ser199Phe], and *TCF1* [Ala98Val]) selected because they were previously shown to be associated with diabetes or glucose intolerance (5,9,12).

**NEUROD1** (Ala45Thr). No association was found between the *NEUROD1* variant and type 2 diabetes in either the 97 families (P = 0.216) or the case-control study (Tables 1 and 2). As only two subjects had the AA genotype, the AA subjects were pooled with subjects possessing the CT genotype. Single-locus regression analysis, including the covariates of waist-to-hip ratio, BMI, age, and sex, found that the *NEUROD1* Ala45 allele was an independent determinant of FBG in both the total group (normoglycemia/impaired glucose tolerance/type 2 diabetes) and those with normal glucose tolerance alone (Table 3). No association was found with 2-h blood glucose (BG), *NEUROG3* (Ser199Phe). No association was found between the *NEUROG3* variant and type 2 diabetes in either the families (P = 0.604) or the case-control study (Tables 1 and 2). Single-locus regression analysis using the same covariates as the previous locus found that the *NEUROG3* Ser199Phe allele was an independent determinant of FBG and 2-h BG in the total group only (Table 3).

**TCF1** (Ala98Val). An association was found between the *TCF1* variant and type 2 diabetes in the case-control study (Table 1 and 2). Subjects with the rare TT genotype (n = 3) were pooled with those subjects possessing the CT genotype. The families were not studied because there was insufficient power for analysis due to the low variant-allele frequency. Single-locus regression analysis using the same covariates in the model as the one previously described did not find that the *TCF1* phenotype was an independent determinant of FBG and 2-h BG in the total group. In contrast, in the normoglycemic subjects, the *TCF1* Val98 allele was an independent predictor of FBG (Table 3).

**NEUROD1, NEUROG3, and TCF1 combinations.** The potential combined allelic effect of *NEUROD1* Ala45, *NEUROG3* Phe199, and *TCF1* Val98 was sought in the 454 subjects from the urban survey, which was comprised of the case-control collection. Subjects were classified according to the number of alleles previously found to be associated with increased risk of hyperglycemia for all three loci. The results are presented in Table 4. A monotonically increasing effect on blood glucose levels, depending on the number of high-risk alleles carried per subject, was found across the whole population (P < 0.0001). For instance, those subjects carrying only one risk allele had an FBG of 4.6 mmol/l and a 2-h BG of 4.9 mmol/l, whereas those subjects with five or six risk alleles had an FBG of 7.5 mmol/l and a 2-h BG of 9.2 mmol/l. Similarly, the proportion of subjects with diabetes increased with the

### Table 1

<table>
<thead>
<tr>
<th></th>
<th><em>NEUROD1</em> (A1a45Thr)</th>
<th><em>NEUROG3</em> (Ser199Phe)</th>
<th><em>TCF1</em> (A1a98Val)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GG</td>
<td>GA</td>
<td>AA</td>
</tr>
<tr>
<td>Normal glucose tolerance</td>
<td>0.763 (241)</td>
<td>0.231 (73)</td>
<td>0.006 (2)</td>
</tr>
<tr>
<td>IGT/IFG</td>
<td>0.679 (38)</td>
<td>0.304 (17)</td>
<td>0.018 (1)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>0.866 (71)</td>
<td>0.134 (11)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>P</td>
<td>0.14*</td>
<td>0.482†</td>
<td>0.037*</td>
</tr>
</tbody>
</table>

*Genotype counts are in parentheses. *Cochran-Armitage. IFG, impaired fasting glucose; IGT, impaired glucose tolerance.*

### Table 2

<table>
<thead>
<tr>
<th></th>
<th><em>NEUROD1</em> (A1a45Thr)</th>
<th><em>NEUROG3</em> (Ser199Phe)</th>
<th><em>TCF1</em> (A1a98Val)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>Normal glucose tolerance</td>
<td>0.878 (555)</td>
<td>0.122 (77)</td>
<td>0.415 (262)</td>
</tr>
<tr>
<td>IGT/IFG</td>
<td>0.830 (93)</td>
<td>0.170 (19)</td>
<td>0.429 (48)</td>
</tr>
<tr>
<td>Type 2 diabetes</td>
<td>0.933 (153)</td>
<td>0.067 (11)</td>
<td>0.372 (61)</td>
</tr>
<tr>
<td>P</td>
<td>0.15</td>
<td>0.40</td>
<td>0.02</td>
</tr>
</tbody>
</table>

*Allele counts are in parentheses. *Cochran-Armitage. IFG, impaired fasting glucose; IGT, impaired glucose tolerance.*
number of risk alleles carried by an individual subject \((P = 0.0024)\) from 5 to 38%.

Our study highlights the possible importance of additive gene effects. Studies (14) in mice have also shown that the combined additive effects of a number of heterozygous mutations of the transcriptional factors involved in pancreatic development and insulin transcription are more important than the effects of individual loci. When the maximum number of high-risk alleles \((5 \text{ or } 6)\) is present in an individual from South India, the FBG is raised by a mean of 2.9 mmol/l and the 2-h BG by a mean of 4.3 mmol/l.

In conclusion, we have demonstrated an interaction between polymorphisms of three genes \((\text{NEUROD1, NEUROG3, and TCF1})\) involved in pancreatic β-cell development/function that underlie the polygenic component of glucose intolerance in a South Indian population.

Furthermore, the number of risk alleles present in a subject was a significant determinant of FBG, 2-h BG, and type 2 diabetes. These current studies clearly demonstrate both the polygenic nature of diabetes and the importance of studying the whole range of glucose tolerance.

### RESEARCH DESIGN AND METHODS

Informed consent was obtained from each study participant, and use of the samples for genetic studies was approved by the ethics committee of the Diabetes Research Centre, Chennai, India. DNA samples were available from 473 South Indian subjects initially recruited into a cross-sectional population-based survey of the prevalence of diabetes, all of whom underwent a 75-g oral glucose tolerance test (OGTT). 328 had normal glucose tolerance, 60 impaired glucose tolerance, and 85 type 2 diabetes. All patients known to have diabetes had their oral therapy stopped 2 days before OGTT testing, and therefore treatment did not interfere with the test results (15,16). Full genotyping for all three loci were obtained in 454 subjects, and analysis was restricted to these subjects. A second study collection consisted of 97 families, each characterized by a proband with type 2 diabetes, at least one other sibling, and both parents, all of whom were of South Indian origin. All subjects were characterized by a 75-g OGTT if their status was unknown. Genomic DNA was extracted from whole blood, and subjects were genotyped for the three polymorphic sites by PCR amplification and restriction endonuclease digestion of the PCR products (PCR/restriction fragment–length polymorphism). All digested fragments were visualized on a 2-3% agarose gel with ethidium bromide staining. \textit{NEUROD1} AlaThr results in a loss of a MboI restriction enzyme site, \textit{NEUROG3} Ser190Pro the loss of an EarIII digestion site, and \textit{TCF1}Ala98Val the loss of a HaeIII digestion site, as has been previously reported (9,12,17).

Genotype and allele distributions between the normal glucose tolerant, impaired glucose tolerant, and diabetic subgroups were assessed using the Cochran-Armitage trend test for all of the \(2 \times 3\) tables in Tables 1 and 2, while a Spearman rank correlation was used for the \(3 \times 3\) table of \textit{NEUROG3} in Table 1. The Cochran-Armitage trend test was also used to provide a one-sided \(P\) value to assess the significance in the increase in the proportion of diabetic subjects with an increasing count of risk alleles in Table 4.

Relationships between genotypes/risk alleles and quantitative traits (FBG and 2-h BG) in Tables 3 and 4 were examined by weighted linear regression in SPSS (version 10) for Windows. This allowed for the variation in FBG/2-h BG variance over genotype. Potential covariates included were age, sex, waist-to-hip ratio, and BMI, which were chosen to complement earlier studies (15,18) from South India.

The extended transmission-disequilibrium test was used to test for evidence of association between diabetes in the family trio collection, as previously described (19). There was insufficient power to perform the transmission-disequilibrium test study on \textit{TCF1} Ala98Val based on the number of informative transmissions.

### ACKNOWLEDGMENTS

This study was supported by a grant from Diabetes U.K.

### REFERENCES


### TABLE 3

Single locus regression analysis on individual loci in subjects from the South Indian urban survey.*

<table>
<thead>
<tr>
<th>Risk Alleles</th>
<th>FBG (mmol/l)</th>
<th>2-h BG (mmol/l)</th>
<th>Subjects with diabetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>4.6 ± 1.5</td>
<td>4.9 ± 1.8</td>
<td>5% (1 of 19)</td>
</tr>
<tr>
<td>Two</td>
<td>4.7 ± 0.9</td>
<td>5.2 ± 2.2</td>
<td>14% (13 of 95)</td>
</tr>
<tr>
<td>Three</td>
<td>5.4 ± 2.2</td>
<td>6.3 ± 4.2</td>
<td>17% (30 of 180)</td>
</tr>
<tr>
<td>Four</td>
<td>6.2 ± 3.7</td>
<td>7.4 ± 5.8</td>
<td>22% (33 of 147)</td>
</tr>
<tr>
<td>Five/six*</td>
<td>7.5 ± 5.3</td>
<td>9.2 ± 8.1</td>
<td>38% (5 of 13)</td>
</tr>
<tr>
<td>(P)</td>
<td>&lt;0.0001†</td>
<td>&lt;0.0001†</td>
<td>0.0024‡</td>
</tr>
</tbody>
</table>

Data are means ± SD. *Only one subject possessed all six risk alleles; †weighted linear regression with covariates of waist-to-hip ratio, BMI, age, and sex; ‡one-sided \(P\) value using a Cochran-Armitage trend test.
onset to a genetic location on chromosome 10q in Mexican Americans. *Am J Hum Genet* 64:1127–1140, 1999


