Evaluating the association between type-2 diabetes and polymorphisms at +1127 INS–PstI and +3580 IGF-II–MspI locus

Sarita Bagadi, Jatin K. Nagpal, Rashmi Khadapkar and Bibhu R. Das*
Research and Development Division, SRL-Ranbaxy Limited, Clinical Reference Laboratories, Mumbai 400 093, India

Type-2 diabetes is a heterogeneous, multifactorial disease with a strong genetic background. The objective of the study was to analyse the relationship between polymorphisms present in the human insulin gene (INS) because of its impact on glucose homeostasis and its immediate neighbouring IGF-II gene and genetic susceptibility to type-2 diabetes. This is a study on the Indian population involving screening of INS and IGF-II candidate genes for nucleotide variation associated with type-2 diabetes. Our study population comprised fifty diabetic subjects and forty age-matched non-diabetic controls. Restriction fragment length polymorphism was employed to study +1127T/C polymorphism in INS gene and +3580G/A polymorphism in IGF-II gene. Biochemical analysis revealed hypertriglyceridaemia in diabetic group. It was observed that T allele frequency in INS +1127 gene was significantly high in normals having high triglyceride level. No correlation was observed between susceptibility to type-2 diabetes and +3580 IGF-II and +1127 INS polymorphism. Moreover, homozygous TT was found to be the predominant INS genotype in Indian population, in contrast to CC which was predominant in African American, Caucasian and Hispanic populations. We propose here that our study be used as a base for future comprehensive studies exploring the inter-relationships among different genetic variations at VNTRs–INS–IGF-II and predisposition to diabetes.

Keywords: IGF-II polymorphism, INS polymorphism, type-2 diabetes.

Type-2 diabetes is the most common form of diabetes mellitus constituting 90% of the diabetic population. The prevalence of diabetes for all age-groups worldwide was estimated to be 2.8% in 2000 and 4.4% in 2030. The total number of people with diabetes is projected to rise from 171 million in 2000 to 366 million in 2030. India, China and the United States are among the top three countries estimated* to have the highest number of people with diabetes (i.e. 31.7, 20.8 and 17.7 millions respectively) in 2000 and (79.4, 42.3 and 30.3 millions respectively) in 2030. Epidemiological studies among migrant Asian Indians in many countries showed higher prevalence of type-2 diabetes compared with the host populations and other migrant ethnic groups². A national survey³ of diabetes conducted in six major cities in India in the year 2000 showed that the prevalence of diabetes in urban adults was 12.1%. Type-2 diabetes is a multifactorial disease with both a genetic component and important non-genetic component(s), which interact in order to precipitate the diabetic phenotype⁴. Although much is known about environmental factors contributing to type-2 diabetes, such as obesity, sedentary lifestyle, smoking and certain drugs, much less is known about the genetic predisposition which is required for the onset of the more common form(s) of the disease. Identification of genes predisposing individuals to develop type-2 diabetes will facilitate early diagnosis and thereby lead to effective treatment and intervention strategies⁵. An absolute insulin deficiency in type-1 diabetes and a relative (rather than absolute) insulin deficiency in type-2 diabetes could be the result of defects in the insulin gene. Cloning of insulin gene has facilitated molecular–genetic studies, which include identification of multiple DNA sequence polymorphisms detected with restriction fragment length polymorphisms (RFLPs) and sequencing. The insulin gene (INS) has been mapped to the short arm of chromosome-11 adjacent to the insulin-like growth factor II (IGF-II) and tyrosine hydroxylase genes. The combined RFLPs for the insulin, IGF-II and tyrosine hydroxylase (TH) genes make this a highly informative locus for genetic studies in diabetes⁶.

The variable number of tandem repeats (VNTRs) that lie immediately adjacent to the 5′ promoter region of INS are believed to have a direct effect on INS regulation. In addition to VNTRs, there are 19 non-coding SNP markers (polymorphisms) in the TH–INS–IGF-II region on chromosome 11p15.5, which are in tight linkage disequilibrium with each other and with the VNTRs, such that they constitute two major haplotypes. In such a region of tight linkage disequilibrium, assaying for one marker would generally provide genotype information of all the others. We assayed for two of these SNP markers, +1127 T/C polymorphism of INS and +3580 G/A polymorphism of IGF-II in type-2 diabetes. The +1127 INS–PstI marker is located in the 3′ untranslated region (UTR) of INS; the UTR regions of the preproinsulin mRNA have recently been demonstrated to play a crucial role in regulating insulin production. The 3′-UTR of INS acts cooperatively with the 5′-UTR and markedly increases glucose-induced proinsulin biosynthesis. Therefore, polymorphisms at +1127 INS–PstI, although located in an untranslated region⁷, may have functional effect on the expression of INS. IGF-I and IGF-II are single-chain polypeptide having 62% homology with proinsulin. IGFs are regulators of processes like growth and metabolism⁸–⁹. IGF-1 and IGF-II also contribute to pancreatic β-cell growth and development by regulating β-cell replication, renewal and apoptosis¹⁰. Deregulation of the balance between β-cell renewal and apoptosis due to alterations in IGF levels is potentially of

*For correspondence. (e-mail: brdas@srlranbaxy.co.in)
RESEARCH COMMUNICATIONS

great importance in the development of glucose intolerance. Polymorphisms in INS and IGF-II have been extensively studied and shown to be implicated in type-1 diabetes\textsuperscript{15-17}. But few studies have been done evaluating their significance in type-2 diabetes\textsuperscript{18-20}. In the present study we have examined the physiological significance, if any, of these two polymorphisms in type-2 diabetic patients of India.

A total of fifty diabetic subjects (equal number of men and women) with mean age 49.82 yrs, who came for clinical testing at SRL-Ranbaxy, Mumbai were included in the present study. The criterion for diagnosis of diabetes mellitus was based on the recommendations of the American Diabetes Association’s modified form of the WHO report, 1985 (fasting glucose $\geq 126$ mg/dl and 2 h post load glucose value $\geq 200$ mg/dl after 75 g oral glucose intake)$^{17}$. Forty age-matched subjects with no clinical evidence of diabetes mellitus and biochemically showing normal glucose tolerance (fasting plasma glucose $< 110$ mg/dl and 2 h post load glucose $< 140$ mg/dl) were selected randomly from corporate health check-up samples by filling up questionnaires. Biochemical evaluation included determination of fasting plasma glucose, 2 h post glucose load after 75 g oral glucose intake and lipid profile (total cholesterol, triglycerides, high density lipoprotein (HDL) and low density lipoprotein (LDL)). The fasting glucose and 2 h post glucose load determination were carried out using fluoridated plasma, while for lipid estimation 12 h fasting serum sample was used. Plasma glucose levels were determined by hexokinase-glucose-6-phosphate dehydrogenase method on Dade Dimension AR Analyser (Dade Behring, USA). A lipid profile (total cholesterol, triglycerides and HDL) was measured spectrophotometrically on Dade Dimension AR Analyser, while LDL was calculated using Friedwald formula.

Genomic DNA was extracted from EDTA anticoagulated whole peripheral blood using Flexi Gene DNA system (Qiagen, GmbH, Germany) and 10 $\mu$l of this was used for PCR amplification. PCR amplification was performed in 50 $\mu$l volumes, containing 1 $\mu$g genomic DNA, in a reaction mix containing 1X PCR buffer, 200 $\mu$M dNTPs, 2.5 mM MgCl$_2$, 50 mM KCl, 1 $\mu$M each of sense and antisense primers and 1 unit of Taq DNA polymerase (Invitrogen, Carlsbad, CA). Primers used to amplify the genomic regions were INS+1035 5'–GGG TCC CCT GCA GAA GGC TGG CA–3' and INS+1597 5'–CTC CCT CCA CAG GGA CTC CAT C–3' for insulin gene and IGF-II-MspF 5'–CCA CCC CTT CTG GGA AGC TAA AAG–3' and IGF-II-MspR 5'–GCG CCC TCG CTC CTC CAG GAA TGG ACA–3' for IGF-II gene. PCR amplification involved an initial denaturation step at 94°C for 5 min, followed by 35 cycles at 94°C for 30 s, 55°C for 45 s and 72°C for 45 s, with final elongation step at 72°C for 7 min. The INS amplicon was more refractory to amplification, being a GC-rich region. Hence, we used Hot Star Taq DNA Polymerase (Qiagen, GmbH, Germany), with proof-reading activity to obtain reliable amplification. Amplified DNA (15 $\mu$l) was subjected to overnight digestion with 10 U of enzyme using the manufacturer’s recommendations. Digested products were size fractionated on 3% agarose gel and visualized by UV-induced ethidium fluorescence. The restriction fragments of the alleles are as follows:

<table>
<thead>
<tr>
<th>Gene</th>
<th>A allele</th>
<th>C allele</th>
<th>G allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>INS</td>
<td>562 bp</td>
<td>470 bp + 92 bp</td>
<td></td>
</tr>
<tr>
<td>IGF-II</td>
<td>122 bp + 118 bp</td>
<td>122 bp + 84 bp + 34 bp.</td>
<td></td>
</tr>
</tbody>
</table>

Data were represented as mean $\pm$ SD in case of continuous variables. Hardy–Weinberg equilibrium was assessed by comparison of observed genotype frequencies, with expected genotype frequencies inferred from observed allele frequencies. Standard contingency tables were used to calculate odds ratio and $\chi^2$ analysis was performed to generate P values. The odds ratio (95% confidence interval, CI), was used as a measure of association between genotypes and diabetes. Unpaired t-test was used in case of continuous variable.

The age-matched diabetic group (50 cases) and normal group (40 controls) were compared for several biochemical parameters (Table 1). Three subjects in the diabetic group had been excluded for calculation of mean LDL levels, as their triglyceride levels were above 400 mg/dl. RFLP analysis was performed to determine the distribution of genotypes at +3580 IGF-II-MspI locus (Figure 1a) and +1127 INS-PstI locus (Figure 1b). The percentage distribution of the three IGF-II genotypes AA, AG and GG in diabetic group was 20, 50 and 30% respectively, whereas in the normal group these frequencies were 20, 42.5 and 37.5% (Table 2). According to the Hardy–Weinberg equilibrium, the frequency for A allele and G allele was observed to be 0.45 and 0.55 respectively, in the diabetic group in comparison to 0.41 and 0.59 in the normal group (Table 2). When the biochemical parameter of triglyceride level was taken into consideration (cut-off value of 170 mg/dl)$^{22,23}$, there was no significant difference in the frequency of A and G alleles in the diabetic group having triglyceride level $\geq 170$ mg/dl as well as $\leq 170$ mg/dl. The frequency of A allele was found to be significantly higher in normal groups having triglyceride values $\geq 170$ mg/dl (0.58) compared to normal groups having triglycerides values $\leq 170$ mg/dl (0.34). However, the odds ratio of 0.36 indicates that A allele of IGF-II is not a risk factor for high triglyceride levels (Table 3).

The frequencies of the three INS genotypes CT, TT and CC in the diabetic group were 24, 70 and 6% respectively, whereas in the normal group were 25, 72.5 and 2.5% (Table 4). Hardy–Weinberg equilibrium gives the frequency of C and T alleles to be 0.18 and 0.82 in diabetic group compared to 0.15 and 0.85 in normal group (Table 4). With reference to triglyceride levels (cut-off value of 170 mg/dl), the frequency of C and T alleles in the diabetic group...
having triglyceride levels >170 mg/dl was 0.24 and 0.76 in comparison to allele frequency of 0.3 and 0.7 in diabetic group having triglyceride levels <170 mg/dl (Table 5). When the normal group was categorized with respect to high level (>170 mg/dl) and low triglyceride levels (<170 mg/dl), there was significant difference in the frequencies of T (0.88 vs 0.68) and C alleles (0.13 vs 0.32). Odds ratio of 3.3 indicates that T allele may be a major risk factor for elevated triglyceride levels in normals, but this needs to be confirmed by increasing sample size in the normal group (Table 5).

Despite several epidemiological studies in India showing a rising prevalence of diabetes, there is a striking lack of genetic studies pertaining to type-2 diabetic Indian patients. A comprehensive national study by Ramachandran et al.\(^3\) indicates towards a large pool of subjects with impaired glucose tolerance that has a high risk of conversion to diabetes. The present pilot project attempts to understand the involvement of SNPs in IGF-II and INS promoters and its implications in predisposition and diagnosis. After reviewing the literature it was observed that various SNPs exist in the INS–IGF-II region having implications in obesity, type-1 diabetes, polycystic ovary syndrome, hypertension and recently, various cancers also. We selected two SNP markers of +3580 IGF-II–MspI and +1127 INS–PvuI that have shown significant role in susceptibility to few cancers like prostate\(^24\); colorectal cancer\(^25\), etc. for evaluating their use as susceptibility markers in type-2 diabetes.
diabetes. A recent study by Sanchez-Corona et al. analyzed eight different polymorphisms in INS, INSR and IRS1 genes in the Mexican population. Among the eight polymorphisms analyzed, the PstI polymorphism in INS was significantly associated with hypertriglyceridaemia (>169.5 mg/dl). However, there are few reports showing that a locus at 11p18.5 confers susceptibility to insulin-dependent diabetes mellitus in North Indians. Here, we have evaluated the significance of VNTR–INS–IGF-II locus as a surrogate marker to determine individuals susceptibility to develop type-2 diabetes. We clarify that this was a pilot project carried out with comparatively small sample size. The objective of the project was to screen two SNPs showing relevance to type-2 diabetes risk, to get some idea of the genotype distribution pattern for these polymorphisms, before replicating them in large population studies.

Several new findings emerged from our study exploring the biochemical parameters and genotype distribution in IGF-II and INS genes in type-2 diabetes patients. Comparison of biochemical parameter revealed that fasting glucose and 2 h post glucose load were quite high in the diabetic group (P < 0.05) than the normal group. The high triglycerides in patient group confirmed diabetic triglyceridaemia in this group. It was observed that T allele at +1127 INS–PstI locus was associated with high triglyceride accumulation risk in normal group. These observations need to replicate in large sample size in order to use T allele at +1127 INS–PstI as a surrogate marker to individuals susceptibility for elevated triglycerides in Indian ethnic group. Similar observations were made by Sanchez-Corona et al., where PstI polymorphism in the INS gene was shown to be associated with hypertriglyceridaemia (triglycerides >169.5 mg/dl). Thus, apparently the lipid metabolism is hampered and affected in diabetic people compared to normals. We observed no significant association between +3580 IGF-II and +1127 INS polymorphism with type-2 diabetes in our study population. In case of IGF-II polymorphisms, genotype frequencies (AA, AG, GG) were comparable to other populations, viz. African Americans, Caucasians and Hispanic. However, homozygous TT was the predominant INS genotype (70%) in the Indian population compared to African American, Caucasian and Hispanic population, where the predominant genotype was CC (72, 61 and 79% respectively).

The screening of various candidate genes for nucleotide variations associated with type-2 diabetes is a key component of diabetes genetic research. The significance of the genotype distribution within the important SNPs in the Indian population cannot be overemphasized. Hence,
more comprehensive studies covering other important SNPs in key genes pertaining to glucose metabolism are planned to be carried out.


Received 10 August 2004; revised accepted 3 March 2005

Molecular characterization of a Pao-like long terminal repeat retrotransposon, Tamy in saturniid silkworm Antheraea mylitta

Ananta K. Ghosh, Abhik Datta, B. Mahendran and S. C. Kundu*

Department of Biotechnology, Indian Institute of Technology, Kharagpur 721 302, India

A long terminal repeat (LTR) retrotransposon named Tamy, was obtained by screening Antheraea mylitta sub-genomic DNA library with PCR amplified partial fibroin gene sequence as probe. Tamy was 8387 nucleotides long with 1305 nucleotides of long LTRs at its 5′ and 3′-ends having features characteristic of a functional LTR retrotransposon. Starting from its N-terminus, nucleic acid binding motif (Cys), protease, reverse transcriptase (RT), RNAaseHI and integrase domains were present in sequential order. RT domain in Tamy showed high homology with Pao-like retrotransposable elements. This is a report of a Pao-like LTR retrotransposon in A. mylitta genome.

Keywords: Antheraea mylitta, Pao-like retrotransposons, reverse transcriptase, Tasar silkworm.

The genomes of eukaryotic organisms that display several multi-copy DNAs dispersed randomly in the genome, is referred to as mobile genetic elements or transposable ele-

For correspondence. (e-mail: kundu@biij.iitkgp.ernet.in)