ACADEMIC PRESS

Molecular Genetics and Metabolism 77 (2002) 189–193

Molecular Genetics and Metabolism

www.academicpress.com

Analysis of *MGEA5* on 10q24.1–q24.3 encoding the β-O-linked *N*-acetylglucosaminidase as a candidate gene for type 2 diabetes mellitus in Pima Indians

V.S. Farook,* C. Bogardus, and M. Prochazka

Clinical Diabetes and Nutrition Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 4212, North 16th Street, Phoenix, AZ 85016, USA

Received 31 May 2002; received in revised form 21 June 2002; accepted 1 July 2002

Abstract

Several diseases including type 2 diabetes mellitus (T2DM) are associated with abnormal O-glycosylation of proteins. β-O-linked *N*-acetylglucosaminidase (O-GlcNAcase) encoded by *MGEA5* on 10g24.1–q24.3 removes N-acetylglucosamine (O-GlcNAc), and we investigated this locus in Pima Indians who have the world's highest prevalence of T2DM. We detected two variants but there was no association with parameters of insulin resistance or diabetes in ~1300 Pimas. We conclude that mutations in *MGEA5* are unlikely to contribute to T2DM in this population.

Keywords: *MGEA5*; O-GlcNAcase; O-GlcNAc; Type 2 diabetes mellitus; Candidate gene; Single nucleotide polymorphism; Chromosome 10q24; Pima Indians

1. Introduction

MGEA5 is a recently described gene coding for the meningioma-expressed antigen-5 [1]. The gene has been positioned at 10q24.l–q24.3 and its product identified as a β-O-linked *N*-acetylglucosaminidase (O-GlcNAcase), which is expressed in many tissues and catalyzes the removal of *N*-acetylglucosamine (O-GlcNAc) from glycosylated proteins [2]. Dynamic glycosylation of serine and threonine residues by O-GlcNAc is a widespread mechanism of posttranslational modification of nuclear and cytoplasmic proteins [3,4] accomplished by O-GlcNAc transferase (OGT), whereas the sugar moiety is selectively removed by O-GlcNAcase. This reversible posttranslational modification affects a variety of cytoplasmic and nuclear proteins and plays an important role that appears to be comparable with the significance of serine/threonine phosphorylation [4]. Aberrant protein glycosylation by GlcNAc has been linked to several disease states including degenerative neurological disorders, cancer and diabetes [4].

Several recent studies reported evidence that increased intracellular O-GlcNAc modification of proteins has been associated with elevated levels of extracellular glucose and glucosamine [5–7]. One study demonstrated that activation of the hexosamine pathway by glucosamine infusion induces insulin resistance by inhibition of early postreceptor insulin signaling via O-GlcNAc modification of insulin receptor substrates IRS-1 and IRS-2 [8]. In addition, O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc), which is a potential inhibitor of O-GlcNAcase, increased the levels of O-GlcNAc leading to insulin resistance and also increased glycosylation of insulin signaling proteins IRS1 and β-catenin [9]. Thus, hyperglycemia-induced O-GlcNAc modification can perturb normal signaling events required for insulin-mediated glucose homeostasis.

O-GlcNAcase is widely expressed in various tissues and cell types including the insulin-producing β-cells in pancreatic islets. The enzyme is irreversibly inhibited by the diabetes-inducing drug streptozotocin (STZ), a

* Corresponding author. Fax: 602-200-5335.
E-mail address: vfarook@mail.nih.gov (V.S. Farook).

1096-7192/02/$ - see front matter
PII: S1096-7192(02)00127-0

Table 1
Primers used for PCR amplification and comparative sequencing of O-GlcNAcase

Primer	Sequences (5′–3′)	Regions
MGEA5-5′-1F	GAG CAC AAT ATC CAG TAA AGC	5′ Region
MGEA5-5′-1R	GTA GCC CAA CTT GTA GTT TCA	
MGEA5-5′-2F	TGA AAC TAG AAG TTG GGC TAG	5′ Region
MGEA5-5′-2R	GCT TCA CTT TTC AGC TGG ATT	
MGEA5-5′-3F	TTC ATG CCT CCC TAG CTT TG	5′ Region
MGEA5-5′-3R	ATT ACA TCT GCC TAA CCC AG	
MGEA5-5′-4F	ATGAGGAATGAACGAATGATTA	5′ Region
MGEA5-5′-4R	GAA ATA GAA CAG GAG ACG CG	
MGEA5-5′-5F	CAG GCT TGT TTC GGA GGG	5′ Region
MGEA5-5′-5R	CTA AGC GGA GAG CGA GG	
MGEA5-1F1	GCA GTG CGG ATA AAC AGG AA	Exon 1
MGEA5-1R1	TGG AGA GGG CTT CAG CTC	
MGEA5-1F2	GCG CAC ACT TGG AGC TGA A	Exon 1
MGEA5-1R2	AGG ACA GAA CGT GTT ACT GC	
MGEA5-2F	AGTCACCTTGTAACGTCACC	Exon 2
MGEA5-2R	GTTCTCCTCTCTCCTTCATA	
MGEA5-3F	GACAGCCTATGAATGTTGCC	Exon 3
MGEA5-3R	ATATAACCGTTTTCTTCACAGA	
MGEA5-4F	CTCTTCAGTTAAACCTTCAGT	Exon 4
MGEA5-4R	CAAAGAAATAGGAATTTGATTTCC	
MGEA5-5F	GTTAACGCTGTAATGTTTGGAA	Exon 5
MGEA5-5R	AGAAGACTAGTTAAATGTAAAGC	
MGEA5-6F	CCCATGAGCATTTTAATTCTGA	Exon 6
MGEA5-6R	CTACTTAATTCAACTTTGACTG	
MGEA5-7F	GAACGGTTGTCGTTATTTCTA	Exon 7
MGEA5-7R	TGAAAGATCAAGTACAGGCTA	
MGEA5-8F	TTATACTGTGAAAACTAGTTGCA	Exon 8
MGEA5-8R	GACTCAACTATGTGATCTAAC	
MGEA5-9F1	CTCTAATTCTGATTATCTGTGTGA	Exon 9
MGEA5-9R1	TCTTGCATGGACATCTCTGG	
MGEA5-9F2	AGAGAGCATAGCTGAATCAAAA	Exon 9
MGEA5-9R2	CTTCAGCAGCAGGTACATTA	
MGEA5-10F	GTTAAGCTTACAGTTTTTTGAG	Exon 10
MGEA5-10R	AAATGTGCCAATTCCCTAAATG	
MGEA5-Int-1F	TCTCCAATTGTGCCAACAGG	Intron 10
MGEA5-Int-1R	CAACGAATTGAAGAGACTTGG	
MGEA5-Int-2F	AGGAATAATCTCTTCTCCTCA	Intron 10
MGEA5-Int-2R	GGCCCTTTAGGTCATTTCAG	
MGEA5-Int-3F	ACAAGCTTTCTGCTGCTTGC	Intron 10
MGEA5-Int-3R	CTTATAGCTGGATTGAAAACC	
MGEA5-Int-4F	GGTAAAAAGCACCAGCATGG	Intron 10
MGEA5-Int-4R	GCTAATTAGTTGTATCACAGG	
MGEA5-Int-5F	CCTGTGATACAACTAATTAGCT	Intron 10
MGEA5-Int-5R	GCCACCAAGAGCGCTGG	
MGEA5-Int-6F	CCAGCGCTCTTGGTGGCT	Intron 10
MGEA5-Int-6R	CGCTGCGCACGGTCAGTAT	
MGEA5-11F	TCAGTTTGTTGGACATTTGAGA	Exon 11
MGEA5-11R	GCGCAACATAAGGTTTCTCA	
MGEA5-12F	CCCCAAACTAATCCATCTCC	Exon 12
MGEA5-12R	AGGTTAAGTTTTAAGTGTGCAC	
MGEA5-13F	CAGACTGACTTGTCAGTTATA	Exon 13
MGEA5-13R	CCTCAACCTGATAAGCCTCA	
MGEA5-14F	CGTAAGTTCATGCAAGTAC	Exon 14
MGEA5-14R	AGACCATTACCATAGTATAATTC	
MGEA5-15F	GTTAGATATCTAATGCGTAACTA	Exon 15
MGEA5-15R	CCTATCTTGCCCAAATACCA	
MGEA5-16F1	CTGTTTATATTAACCATCTCGA	Exon 16
MGEA5-16R1	CCAACCAGTGAGTAGTCTCA	
MGEA5-16F2	GGATTCGAAACAAAGAAGACTA	3′ Region
MGEA5-16R2	TCATTACAGGAGACAATTGGC	
MGEA5-16F3	TGGTTGTGGCATTAGTTTAAGCT	3′ Region
MGEA5-16R3	GCACTTTTCAAC AG ATTAAGTC	
MGEA5-16F4	TTGTAGGACTTAATCTGTTGAAA	3′ Region

Table 1 (*continued*)

Primer	Sequences (5′–3′)	Regions
MGEA5-16R4	CTTTCACAATCCGGTTAAAGG	
MGEA5-16F5	CTGAATGAATGAGCACACTTTA	3′ Region
MGEA5-16R5	ACCTTCTCTCCTTTCTTCTCA	
MGEA5-16F6	TAGTGACTTCCAACCAAAGC	3′ Region
MGEA5-16R6	GAAATAGCATAGCACTTCATGA	
MGEA5-16F7	GTCACAGGTGACTGTGAAAC	3′ Region
MGEA5-16R7	ATCTGGAAGTCTTACTGGTGA	
MGEA5-3′-1F	TTTCTCCTCTTTCATGCCTCA	3′ Region
MGEA5-3′-1R	TCACTGAGACTTAGCTGCTG	

Primers were designed for all the 16 exons and also for intron 10 (which is included in splice variant *MGEA5*s) and their direction indicated by F (forward) and R (reverse).

GlcNAc analog with β-cell specific toxicity, which is used to induce diabetes in animal models [10]. It has been proposed that the accumulation of glycosylated proteins in β-cells following exposure to STZ can play a significant role in β-cell loss leading to diabetes [10].

These observations indicate that alterations of genes involved in the regulation of protein O-glycosylation (e.g. *OGT* on chromosome Xq13 coding for GlcNAc transferase, and *MGEA5* on 10q24.l–q24.3 encoding the GlcNAcase) could play a role in the genetic predisposition to diabetes by contributing to β-cell loss and/or insulin resistance. Type 2 diabetes mellitus (T2DM) is the most common form of disturbed glucose homeostasis, representing a complex, heterogeneous disease caused by a combination of genetic and environmental factors [11]. Results of recent genome-wide scans in different populations have revealed evidence for linkage of T2DM with various chromosomes, including an overlapping area on chromosome 10q in Mexican Americans [12], UK Caucasians [13], French [14], and Finns [15]. Furthermore, this area has been tentatively linked with differences in glycated hemoglobin A1C (HbA1C) levels in the Framingham Heart Study [16]. In addition, this genomic region also showed a modest linkage with 2-h plasma glucose (2hPG) measured during an oral glucose tolerance test, and also with whole body glucose disposal in response to a physiological dose of insulin (M_{low}) determined by hyperinsulinemic euglycemic clamp in non-diabetic Pima Indians [17]. Although there was no linkage of 10q with overt diabetes in a larger set of Pima families [18], 2-h glucose concentration during the oral glucose tolerance test (2-hPG), and glucose disposal during the low (40 mU/m^2 per min) insulin infusion of the glucose clamp (M_{low}) are clinical indicators of insulin sensitivity, and their impairment is a predictor of T2DM in this Native American population with the highest documented prevalence of the disease in the world [19]. Also this region is linked with Alzheimer's disease and contains the gene encoding the insulin degrading enzyme (IDE), which has the ability to degrade β-amyloid [20].

The location of *MGEA5* on 10q in the region linked either with T2DM or possibly with insulin resistance in different populations, and the potential pathophysiologic role of the enzyme in the development of abnormal glucose homeostasis through alterations in O-GlcNAc protein modification, led us to investigate the gene as a possible candidate in the Pima Indians.

2. Materials and methods

2.1. Subjects

Fifteen diabetic subjects with an early onset (before 25 years) and 15 non-diabetic controls (at least 45 years old with normal glucose tolerance) used in our study are members who have been participating in a longitudinal study of the development of Type 2 diabetes [19] which is diagnosed according to WHO criteria as described [18]. The informative polymorphism (SNP2) was analyzed in the entire population of approximately 1300 Pimas involved in our original linkage studies [18]. A smaller subset (234) of non-diabetic subjects was also used for analysis of association of the marker with parameters of insulin action (M_{low} and fasting insulin) described in [17]. This study was approved by the Institutional Review Board of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) and the Tribal Council of the Gila River Indian Community. All subjects provided written informed consent prior to participation.

2.2. DNA amplification, variant screening, sequencing, and statistical analysis

For variant detection we used three different pooled DNA samples. Each pool comprised of 10 subjects (5 diabetic and 5 non-diabetic Pimas) thus representing a total of 30 individuals were screened. The primers used for variant (SNP) screening and sequencing are given in Table 1. Overlapping PCR products of approximately 300–600 bp in length were examined for variants by denaturing high performance liquid chromatography (DHPLC) on the WAVE DNA fragment analysis system according to manufacturer's instructions (Transge-

nomic, Omaha, USA) as described before [21]. SNPs were confirmed by sequencing the corresponding PCR products from individual DNA samples. Sequencing reactions were performed with the Big Dye Terminator cycle sequencing kit (PE Applied Biosystems, Foster City, CA) supplemented with the Seq Saver buffer (Sigma, St. Louis, MO) and analyzed on an ABI 3700 Sequence Detection System (PE Applied Biosystems). Association analyses were performed using the χ^2-test.

3. Results and discussion

The *MGEA5* gene has been localized to chromosome 10q24.1–q24.3 [1] and is carried on BAC clone RP11-573E23. The gene consists of 16 exons spanning over ~34 kb [22] and coding for a 130-kDa protein. An alternatively spliced transcript (*MGEA5s*) was described recently, which consists of exons 1–10 and extends through intron 11. This transcript has an alternative stop and polyadenylation sites and codes for a 75-kDa protein [22].

We initially screened all exons plus the flanking splice sites, the entire intron 11 retained in the *MGEA5s* transcript, 2 kb upstream of exon 1, and ~2 kb of the entire region of the 3′ UTR for variants by DHPLC in three different pooled DNA samples. By using this strategy, two variants were identified and confirmed by sequencing in the 30 subjects. The variants were SNP-1 a C to T base change in the promoter using forward and reverse primers 5′-GCAGTGCGGATAAACAGGAA-3′, 5′-TGGAGAGGGCTTCAGCTC-3′, and SNP-2 a T to G base change (corresponding to dbSNP entry rs#2305194) in intron 7 with forward and reverse primers 5′-TTATACTGTGAAAACTAGTTGCA-3′, 5′-GACTCAACTATGTGATCTAAC-3′, respectively.

Based on initial sequencing in 30 subjects, SNP-1 was not informative (allele frequency 0.98/0.02) and therefore was not analyzed further. The more informative SNP-2 (allele frequency 0.6/0.4) was genotyped in a set of DNAs consisting of approximately 1300 individuals who were analyzed in our original genome linkage scan [18]. Because of the putative linkage of the area on 10q with indices of insulin resistance [19], we analyzed association of SNP2 with plasma glucose levels measured at different time points during oral glucose tolerance test, and with M_{low} in a subset of 234 non-diabetic subjects for whom these parameters were available. There was no association of SNP2 with any of these clinical measurements ($p > 0.1$). Similarly, no association of this polymorphism with T2DM or with HbAlC was detected in the entire linkage set ($p > 0.3$).

In conclusion, we investigated *MGEA5* as a positional and biological candidate gene for T2DM or insulin resistance by scanning the entire locus for variants in diabetic and non-diabetic Pima Indians and found two SNPs in this population. Although, neither of the detected SNPs was associated with diabetes or insulin resistance, identification of variants in this gene should be valuable for explorations of the locus for its pathogenic significance in T2DM in other populations, including those showing linkage of the 10q region with diabetes.

Acknowledgments

We thank the members of the Gila River Indian Community for their participation and cooperation in the ongoing longitudinal studies.

References

[1] D. Heckel, N. Comtesse, N. Brass, N. Blin, K.D. Zang, E. Meese, Novel immunogeneic antigen homologous to haluronidase in meningioma, Hum. Mol. Genet. 7 (1998) 1859–1872.
[2] Y. Gao, L. Wells, F.I. Comer, G.L. Parker, G.W. Hart, Dynamic O-glycosylation of nuclear and cytosolic proteins: Cloning and characterization of a neutral, cytosolic β-*N*-acetylglucosaminidase from human brain, J. Biol. Chem. 276 (2001) 9838–9845.
[3] L. Wells, K. Vosseller, G.W. Hart, Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc, Science 291 (2001) 237–2378.
[4] J.A. Hanover, Glycan-dependent signaling: O-linked *N*-acetylglucosamine, FASEB J. 11 (2001) 1865–1876.
[5] H. Yki-Jarvinen, A. Virkamaki, M.C. Daniels, D. McClain, W.K. Gottschal, Insulin and glucosamine infusions increase O-linked *N*-acetylglucosamine in skeletal muscle proteins in vivo, Metabolism 47 (1998) 449–455.
[6] K. Liu, A.J. Paterson, E. Chin, J.E. Kudlow, Glucose stimulates protein modification by O-linked GlcNAc in pancreatic beta cells: linkage of O-linked GlcNAc to beta cell death, Proc. Natl. Acad. Sci. USA 97 (2000) 2820–2825.
[7] I. Han, E.S. Oh, J.E. Kudlow, Responsiveness of the state of O-linked *N*-acetylglucosamine modification of nuclear pore protein p62 to the extracellular glucose concentration, Biochem. J. 350 (2000) 109–114.
[8] M.E. Patti, A. Virkamaki, E.J. Landaker, C.R. Kahn, H. Yki-Jarvinen, Activation of the hexosamine pathway by glucosamine in vivo induces insulin resistance of early postreceptor insulin signaling events in skeletal muscle, Diabetes 48 (1999) 1562–1571.
[9] K. Vosseller, L. Wells, M.D. Lane, G.W. Hart, Elevated nucleocytoplasmic glycosylation by O-GlcNAc results in insulin resistance associated with defects in *Akt* activation in 3T3-L1 adipocytes, Proc. Natl. Acad. Sci. USA. 99 (2002) 5313–5318.
[10] R.J. Konrad, I. Mikolaenko, J.F. Tolar, K. Liu, J.E. Kudlow, The potential mechanism of the diabetogenic action of streptozotocin: inhibition of pancreatic beta-cell O-GlcNAc-selective *N*-acetyl-β-D-glucosaminidase, Biochem. J. 356 (2001) 31–41.
[11] S.C. Elbein, The genetics of human noninsulin-dependent (type 2) diabetes mellitus, J. Nutr. 127 (1997) 1891–S1896.
[12] R. Duggirala, J. Blangero, L. Almasy, D.T. Dyer, K.L. Williams, R.J. Leach, P. O'Connell, M.P. Stern, Linkage of type 2 diabetes mellitus and of age at onset to a genetic location on chromosome 10q in Mexican, Americans Am. J. Hum. Genet. 64 (1999) 1127–1140.
[13] S. Wiltshire, A.T. Hattersley, G.A. Hitman, M. Walker, J.C. Levy, M. Sampson, S. O'Rahilly, T.M. Frayling, J.I. Bell, G.M.

Lathrop, A. Bennett, R. Dhillon, C. Fletcher, C.J. Groves, E. Jones, P. Prestwich, N. Simecek, P.V. Rao, M. Wishart, R. Foxon, S. Howell, D. Smedley, L.R. Cardon, S. Menzel, M.I. McCarthy, G.F. Bottazzo, A genomewide scan for loci predisposing to type 2 diabetes in a UK population (the Diabetes UK Warren 2 Repository): analysis of 573 pedigrees provides independent replication of a susceptibility locus on chromosome 1q, Am. J. Hum. Genet. 69 (2001) 553–569.

[14] N. Vionnet, E.H. Hani, S. Dupont, S. Gallina, S. Francke, S. Dotte, F.D. Matos, E. Durand, F. Lepretre, C. Lecoeur, P. Gallina, L. Zekiri, C. Dina, P. Froguel, Genomewide search for type 2 diabetes-susceptibility genes in French whites: evidence for a novel susceptibility locus for early-onset diabetes on chromosome 3q27-qter and independent replication of a type 2-diabetes locus on chromosome 1q21–q24, Am. J. Hum. Genet. 67 (2000) 1470–1480.

[15] S. Ghosh, R.M. Watanabe, T.T. Valle, E.R. Hauser, V.L. Magnuson, C.D. Langefeld, D.S. Ally, K.L. Mohlke, K. Silander, K. Kohtamaki, P. Chines, J. Balow Jr., G. Birznieks, J. Chang, W. Eldridge, M.R. Erdos, Z.E. Karanjawala, J.I. Knapp, K. Kudelko, C. Martin, A.M. Mena, A. Musick, T. Musick, C. Pfahl, R. Porter, J.B. Rayman, The Finland-United States investigation of non-insulin-dependent diabetes mellitus genetics (FUSION) study. I. An autosomal genome scan for genes that predispose to type 2 diabetes, Am. J. Hum. Genet. 67 (2000) 1174–1185.

[16] J.B. Meigs, C.I. Panhuysen, R.H. Myers, P.W. Wilson, L.A. Cupples, A genome-wide scan for loci linked to plasma levels of glucose and HbA(1c) in a community-based sample of Caucasian pedigrees: The Framingham Offspring Study, Diabetes 51 (2002) 833–840.

[17] R.E. Pratley, D.B. Thompson, M. Prochazka, L. Baier, D. Mott, E. Ravussin, H. Sakul, M.G. Ehm, D.K. Burns, T. Foroud, W.T. Garvey, R.L. Hanson, W.C. Knowler, P.H. Bennett, C. Bogardus, An autosomal genomic scan for loci linked to prediabetic phenotypes in Pima Indians, J. Clin. Invest. 101 (1998) 1757–1764.

[18] R.L. Hanson, M.G. Ehm, D.J. Pettitt, M. Prochazka, D.B. Thompson, D. Timberlake, T. Foroud, S. Kobes, L. Baier, D.K. Burns, L. Almasy, J. Blangero, W.T. Garvey, P.H. Bennett, W.C. Knowler, An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians, Am. J. Hum. Genet. 63 (1998) 1130–1138.

[19] S. Lillioja, D.M. Mott, M. Spraul, R. Ferraro, J.E. Foley, E. Ravussin, W.C. Knowler, P.H. Bennett, C. Bogardus, Insulin resistance and insulin secretory dysfunction as precursors of non-insulin-dependent diabetes mellitus. Prospective studies of Pima Indians, N. Engl. J. Med. 329 (1993) 1988–1992.

[20] C. Lendon, N. Craddock, Susceptibility gene(s) for Alzheimer's disease on chromosome 10, Trends Neurosci. 10 (2001) 557–559.

[21] J.K. Wolford, D. Blunt, C. Ballecer, M. Prochazka, High-throughput SNP detection by using DNA pooling and denaturing high performance liquid chromatography (DHPLC), Hum. Genet. 107 (2000) 483–487.

[22] N. Comtesse, E. Maldener, E. Meese, Identification of a nuclear variant of *MGEA5*, a cytoplasmic hyaluronidase and a β-*N*-acetylglucosaminidase, Biochem. Biophy. Res. Commun. 283 (2001) 634–640.