Genetic influence of known Calpain-10 polymorphisms in the Korean Type II diabetes mellitus

Kim Hye Joo

Department of Medical Science

The Graduate School, Yonsei University

Genetic influence of known Calpain-10 polymorphisms in the Korean Type II diabetes mellitus

Directed by Professor Hyun Chul Lee

The Master's Thesis submitted to the Department of Medical Science, the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Master of Medical Science.

Hey Joo Kim

December 2003

This certifies that the Master's Thesis of Hye Joo Kim is approved.

[Thesis Supervisor : Hyun Chul Lee]

[Thesis Committee Member]

[Thesis Committee Member]

The Graduate School
Yonsei University

December 2003

Genetic influence of known Calpain-10 polymorphisms in the Korean Type II diabetes mellitus

Directed by Professor Hyun Chul Lee

The Master's Thesis submitted to the Department of Medical Science, the Graduate School of Yonsei University in partial fulfillment of the requirements for the degree of Master of Medical Science.

Hey Joo Kim

December 2003

Acknowledgements

대학을 졸업한지가 얼마되지 않은 것 같은데 어느덧 시간이 흘러서 대학원 졸업을 하게 되었습니다. 많은 시간이 흐르고 그 시간 속에 좋았던 기억들도 힘들었던 시간들도 있었던 것 같습니다. 그 동안 저에게 힘이 되어주셨던 분들께 이 지면을 빌어 감사함을 전하고자 합니다.

먼저 석사생활 동안에 믿어주시고 많은 지도와 가르침을 주신 이현철 교수님께 감사를 드립니다. 또한 바쁘신 와중에도 석사 논문을 지도해 주신 남문석 교수님과 이진성 교수님께 감사를 드립니다.

가장 많은 시간동안 저와 함께 생활하고 부족한 것이 많았던 저에게 실험을 가르쳐 주시고, 힘이 되어주신 우선옥 선생님께 진심으로 깊은 감사를 드립니다. 실험실 생활동안 기쁜 일, 힘든 일을 함께 나눈 혜준언니를 비롯한 실험실 분들께 감사를 드립니다.

지치고 힘들 때, 기쁜 일이 있을 때마다 함께 울고 웃어주면서 힘이 되어준 사랑하는 친구 경림이, 효현이, 지영이 그리고 바람 쐬고 싶다는 말에 대전에서 서울까지 단숨에 올라와준 희경이, 오영미 그리고 가까운 곳에서 힘들 때면 웃음을 주었던 은정이와 후배 진주를 비롯한 많은 친구들에게 진심으로 고마움을 표합니다. 늘 힘들 때 짜증과 투정을 묵묵히 받아주고 언제나 내 편이 되어주고, 옆에 있는 것만으로도 가장 큰 힘이 되어주었던 태훈이에게 진심으로 고마움을 전합니다.

연구소 생활하기도 많이 힘들 텐데 어려울 때마다 늘 신경 써주고 배려해 준 큰오빠와 늘 편안함과 용기를 주어 힘이 되어준 작은오빠에게 진심으로 고마움을 전합니다. 마지막으로 늘 변함없는 자상함과 헌신 그리고 말로 표현할 수 없는 사랑으로 버팀목이 되어주신 부모님께 감사함을 전하며 부족하지만 이 작은 결실을 바칩니다.

2003

김혜주

Contents

Abstract......1

I. Introduction......3

II. Materials and Methods......8

1. The study population......8

2. DNA extraction......9

3. Genotyping analysis......10

4. Sequencing in SNP region......13

5. Statistical Analyses......14

III. Results......15

1. Genotyping of calpain-10 in subjects......15

2. The sequencing of the PCR products......17

3. The assessment of the association of the previously described haplotypes of calpain-10 and type II diabetes......25

IV. Discussion......34

V. Conclusion......38

Reference......40

Abstract (in Korean)......45

List of Figures

Figure 1. Exon - intron organization of calpain-10..19

Figure 2. SNP-43 genotyped by using PCR and RFLP..20

Figure 3. SNP-19 genotyped by using PCR...21

Figure 4. SNP-63 genotyped by using PCR and RFLP..22

Figure 5. Sequence of calpain-10..23

List of Tables

Table 1. Clinical Characteristics of Subjects..18

Table 2. Allele frequency of Calpain-10 of SNP-43, -19, and -63........................27

Table 3. Genotype distribution of Calpain-10 of SNP-43, -19, and -63...............28

Table 4-1. Haplotype allele distribution from the results of PHASE...................29

Table 4-2. Haplotype analysis...30

Table 5-1. Haplotype allele combinations with haplotype allele 1-5....................31

Table 5-2. Haplotype combination analysis..32

Abstract

Genetic influence of known Calpain-10 polymorphisms in the Korean Type II diabetes mellitus

Hye Joo Kim

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Hyun Chul Lee)

It has been reported recently that the polymorphisms of two genes, calpain-10 and PPARγ, is associated with type II diabetes mellitus. Calpain-10 is a member of a large intracellular protease family. In Mexican-Americans and other populations, the variants of three single nucleotide polymorphisms (SNPs) -43, -19, and -63 of this ubiquitously expressed protein influence the susceptibility to type II diabetes. However, the difference in the alleles and the haplotypes attribututing to the risk was observed in different ethnic groups. This suggests the importance of determining the role of calpain-10 in various populations. The aim of this study

was to investigate the influence of calpain-10 (SNPs-43, -19, and -63) on the susceptibility to type II diabetes in Korean.

739 samples, 499 type II diabetes and 240 control individuals, were examined. Their calpain-10 genotype regarding SNPs -43, -19, and -63 was analyzed. The genotype of SNP-19 was determined by electrophoresis of the polymerase chain reaction (PCR). The genotype of SNP-43 and -63 was analyzed by the restriction fragment length polymorphism (RFLP) method. The association was determined by analyzing the difference of the frequency of the genotype, alleles, haplotype, and the combination of haplotypes in control groups and in type II diabetes patients using SAS ver 8.0.

We have observed that the frequency of alleles, genotypes, and haplotypes of the three SNPs is significantly different in two groups. The difference was observed in the 121/111 (OR=3.344 and p=0.001) and 112/112 (OR=2.381 and p=0.049) haplotype. In contrast, the 112/121 heterozygosity has been reported to be associated with the increased risk of type II diabetes mellitus in other populations.

The results of this study suggest the association of the calpain-10 heterozygotes with combinations of SNPs -43, -19, and -63 with the risk of type II diabetes mellitus in Korean population.

Key Words: type II diabetes, calpain-10, RFLP, single nucleotide polymorphism

Genetic influence of known Calpain-10 polymorphisms

in the Korean Type II diabetes mellitus

Hye Joo Kim

Department of Medical Science

The Graduate School, Yonsei University

(Directed by Professor Hyun Chul Lee)

I. Introduction

Type I and type II diabetes are multi-factorial disorders.[1] Type I diabetes mellitus is an autoimmune diseases that destroys insulin-producing cells in the pancreas. Type I diabetes accounts for 5~10% of all of diabetes cases. The major susceptibility gene of type I diabetes mellitus has been mapped to the HLA region of chromosome 6.[1, 15]

Type II diabetes mellitus is a common form of diabetes that accounts for approximately 90% of all diabetes cases. Type II diabetes mellitus affects

10~20% of adults over 45 years in many developed countries. Type II diabetes mellitus patients are estimated more than 135million people worldwide.[1, 6, 13] Type II diabetes is characterized by the defects in insulin secretion and insulin action.[1, 12, 13] The causality of Type II diabetes is complex: multiple genetic effects, metabolic factors, and environmental factors - contribute to the causality.[1, 2, 5, 13] Similarly, the combination of genetic factors and the environmental factors contribute to its development.[5, 12]

Insulin resistance is a key factor in the development of type II diabetes mellitus. Although the complex pathogenesis of insulin resistance is not completely understood, genetic factors play an important role in the development of diabetes, with some forms of the disease resulting from mutations in a single gene.[1, 6]

Genome wide scans have led to the chromosomal localization of susceptibility loci for type II diabetes in Pima Indians of North America, African-Americans, Caucasians, Japanese Americans from the American Diabetes Association Genetics of NIDDM (GEBBID) Study, Mexican-Americans, in Caucasians from Finland, France, Utah (U.S) and the U.K.[1, 4, 5, 8]

The genome-wide search for the type II diabetes gene in Mexican Americans (from Starr County, TX) localized a susceptibility gene, NIDDM1, to the region of D2S125-D2S140.[1, 4] Of these loci, only NIDDM1 has been mapped to a single

gene.[1, 5, 8] NIDDM1 may also contribute to the development of type II diabetes in other population, but its contribution in various European (German, French, Sardinian, British and Finnish), Asians (Japanese, Chinese) and native Americans (Pima Indian) appear to be less than in Mexican Americans.[1, 4, 16-22] Each study assigned the susceptibility to the different region of the genome, suggesting that different combinations of the susceptibility gene contribute to the development of type II diabetes in these populations.[1]

By the positional cloning, Horikawa et al[1], recently identified NIDDM1 as calpain-10. (CAPN-10) [1, 3, 4, 14] The variation in calpain-10 is associated with a threefold increased risk of diabetes in Northern European population.[1, 3- 5] Recently, many studies have shown that the variation in the gene encoding calpain-10 is associated with type II diabetes.[1, 2, 14] In addition, the diabetes-associated polymorphism in the calpain-10 gene is associated with the decreased level of calpain-10 mRNA in the skeletal muscle and decreased insulin resistance in non-diabetic subjects. This suggests that calpain-10 may increase the susceptibility to type II diabetes.[1, 4, 5, 14]

Calpain-10 is a member of the calpain (calcium-activated neutral proteases) family. The calpain function is not known. Calpain is expressed in many tissues, such as pancreatic islets, muscles, adipose tissues, and liver-tissues, which play a key role in controlling glucose homeostasis.[1-9] They are essential for multiple

cellular functions. For example, the processing property of these enzymes makes it possible to directly modulate the activities and structures of other proteins.[5, 15, 21] In Mexican-Americans, calpain-10 is associated with the threefolds increased risk of type II diabetes. In this population, approximately 14% are at the risk of developing diabetes. [1, 5] The combination of two haplotypes generated by variation at three single nucleotide polymorphisms (SNPs), designated UCSNPs 43, 19 and 63, accounts for the increased risk.[1, 5, 8] All these three SNPs are located within introns and are very likely to influence the gene expression.[1-3, 6, 8] The genotype most strongly associated with the evidence for linkage - that is, SNP-43 G/G (or 1/1) - was not significantly associated with increased risk of diabetes in the Mexican American patient sample.[1, 3, 7] Instead, increased risk was observed for multi site haplotypes centered on the calpain-10 gene, including site 43.[1-7, 14] The allelic association in Mexican Americans suggests that the haplotypes at the rick of developing diabetes may a subset of three variations, at site 43, 19 and 63. [1-7, 14-16] In two independent groups of Mexican Americans (OR 2.8 and 3.58), Finns (OR 2.55) and Germans (OR 4.97), the analysis of three haplotypes suggests that the 1/1-1/2-2/1 genotype (haplotypes 1-1-2 and 1-2-1) is the greatest risk of the disease. [1, 3, 6, 7] Previous studies in non-diabetic Pima Indians revealed that the homozygote at the G allele in SNP-43 has reduced muscle calpain-10 mRNA levels and insulin resistance.[1, 5, 7-9] However, the risk of

type II diabetes is not increased in the SNP-43 G/G homozygotes.[1, 3, 7-10] Similarly, in Oji-Cree Indians, the increased risk of type II diabetes was not observed in the at SNP-43 G/G homozygotes.[1, 3, 7-10] In contrast, the SNP-43 G/G homozygotes were found to have the significantly increased risk of type II diabetes in African Americans.[1, 3, 6-10] It is not clear whether these studies are in agreement or disagreement with the original findings that the risk of diabetes is conferred by the combination of three haplotypes.[1-14] It has been reported that at least these three calpain-10 variants are present.[1-17]

The aim of this study examined the association of the calpain-10 haplotypes SNPs-43, -19, and -63 with the risk of type II diabetes in Korean.

II. MATERIALS AND METHODS

1. The study population

739 unrelated individuals were recruited for this study: 499 type II diabetes patients and 240 non-diabetes controls.

A. Type II diabetes patients

All patients are treated at the Yonsei university and Inha university college of medicine. During the diagnosis, the WHO definitions and criteria were used. Patients received the standard questionnaire that contains questions regarding the onset of type II diabetes diagnosis, family history, the treatment regimen and other medical issues. Only patients diagnosed as type II diabetes clinically and are under the insulin therapy for at least two years after diagnosis were recruited. Each subject was submitted to the basic physical examination such as the information of anthropometry (height, weight) and fasting glucose (glucose, SBP, DBP, FBG, total cholesterol, HDL cholesterol, triglycerides and C-peptide levels) for the statistical analysis.

Families were ascertained for at least 2 siblings who were diagnosed with type II diabetes from 35 year to 80 year. If possible, one parent or sister and brother was known to have type II diabetes.

B. Control subjects

Control subjects are individuals with normal fasting glucose (<110mg/dl), normal random glucose (<140mg/dl) and negative family history of type II diabetes among the first degree relatives. Their age was older than 55.

2. DNA Extraction

Total genomic DNA was extracted from 5ml fresh whole blood collected in the EDTA-, Heparin- and citrate-anticogulant tubes. 15ml of Cell Lysis Solution (Promega in U.S.A) was added to a 50ml centrifuge tube. The tube containing blood was gently rocked until thoroughly mixed, inverted 5~6 times to mix, incubated for 10 minutes at room temperature (invert 2~3 times once during the incubation) to lyse red blood cells, and centrifuged at 2,000 x g for 10 minutes at room temperature. The supernatant was

discarded without disturbing the white pellet. Approximately 50~100μl of residual liquid will remain in the tube. The pellet may be contaminated with red blood cells or cell debris. If the pellet appeared to contain red blood cells, the lysing process was repeated. 5ml of Nuclei Lysis Solution (Promega, in U.S.A) was added to lyse. The white blood cells were pipetted the solution 5~6 times to lyse. If clumps of cells are visible after mixing, incubate the solution at 37°C until the clumps are disrupted. The nuclear lysate were added 1.7ml of Protein Precipitation Solution (Promega, in U.S.A) and vortexes vigorously for 10~20 seconds. Centrifuged at 2,000 x g for 10 minutes at room temperature. The supernatant was transferred to a fresh-tube, added 5ml isopropanol, and mixed gently until the mass of white thread-like DNA is formed, centrifuged at 13,000rpm for 20min, and added 500μg 70% ethanol gently mixed, and centrifuged at 13,000rpm for 10 minutes. The pellet thus obtained was air-dried for 10~15 minutes, added 500μg of DNA Rehydration Solution or D.W and incubated overnight at 4°C.

3. Genotyping Analysis

A. UCSNP-43 (nt4852, G/A)

The polymerase chain reaction was performed with 100ng of genomic DNA in 20 μl volume and the containing the upstream primer 5′-CACGCTTGCTGTGAAGTAATGC-3′ and the downstream primer 5′-CTCTGATTCCCATGGTCTGTAG-3′ in the buffer containing 1 pmol up-primer and 1 pmol down-primer, PCR premix (bioneer, in Korea). The initial denaturation step was 95°C for 5 minutes, followed by 30 cycle of 95°C for 30 sec, 56°C for 30 sec, and 72°C for 30 sec. The extension step was 72°C for 7 minutes. The resulting 144bp product was digested with Nsi I (Promega, in U.S.A) at 37°C for 15~18h. The digested products were subjected to electrophoresis through a 3% agarose gel containing 250 nmol/l ethidium bromide. Allele 1 (G/G) product remained 144bp after digested with Nsi I, whereas Allele 2 (A/A) product was cleaved into two products of 121 and 23bp.

B. UCSNP-19 (ins/del 32bp)

The polymerase chain reaction was performed 100ng of genomic DNA in 20 μl volume containing the upstream primer 5′-GTTTGGTTCTCTTCAGCGTGGAG -3′ and the downstream primer

5′-CATGAACCCTGGCAGGGTCTAAG-3′ in the buffer containing 1pmol up-primer and 1pmol down-primer, PCR premix (bioneer, in Korea). The initial denaturation step was 95°C for 5 minutes, followed by 30 cycle of 95°C for 30 sec, 57°C for 30 sec, and 72°C for 30 sec, and the extension step of 72°C for 7 minutes. The PCR products were separated on a 3% agarose gel. Allele 1 (two repeats of 32bp sequence) is 155bp, and Allele 2 (three repeats of 32bp) is 187bp, and Allele 3 (Allele 1 and Allele 2) is 155bp and 187bp.

C. UCSNP-63 (nt34288, C/T)

The polymerase chain reaction was performed with 100ng of genomic DNA in 20 μl volume containing the upstream primer 5′-AAGGGGGGCCAGGGCCTGACGGGGGTGGCG-3′ and the downstream primer 5′-AGCACTCCCAGCTCCTGATC-3′ in the buffer containing 1 pmol up-primer and 1 pmol down-primer, PCR premix (bioneer, in Korea). The initial denaturation step was 94°C for 5minutes, followed by 30 cycle of 94°C for 30 sec, 63°C for 30 sec, and 72°C for 30 sec, and the extension step was 72°C for 10 minutes. PCR products were digested with two units of Hha I (Takara, in Japan) according to the

manufacturer's instructions for 2h at 37°C. The digested products were separated onto 3% agarose gel. Allele 1 (C) was detected as a 162bp fragment and Allele 2 (T) as a 192bp fragment.

4. Sequencing in SNP region

The entire samples confirmed that SNP-43, -19, and -63 are checked in sequence by used vector cloning construct. Each region of PCR products and restriction enzyme products was cloned into pGEM-T easy vector (Promega in U.S.A). DNA generated by PCR was purified using the DNA pure PCR clean up kit (Unoclon, in U.S.A). Purified DNA was cloned in pGEM-T easy vector (Promega, in U.S.A). The cloned vectors were ligated with ligase (Takara, in Japan) for 2hours at room - temperature. After the ligation, the vector mixture carrying out transformation used DH5α cell. If clone obtained, clone was culture 3ml of LB media in 37°C. After DNA purification, EcoR I (Takara, in Japan) was confirmed. If vector had the PCR product, EcoR I was divided into vector (3.5kb) and PCR product. This constructs read T7 primer and S6 primer. Using of DNA club and cluster site made sequence affirmation.

5. Statistical Analysis

The haplotype was inferred by the expectation using the maximization method with PHASE Software (http://www.stat.washington.edu/stephens/software..html). The haplotypes were reconstructed and the recombination rate was estimated from the population data. The difference of the allelic frequencies in type II diabetes and control groups was analyzed by SAS (Statistical analysis system) ver 8.0. The OR and 95% confidence interval were calculated by logistic regression analysis. Comparison of variables between groups of genotypes was performed using chi-square method.

III. Results

1. Genotyping of calpain-10 in subjects

The clinical characteristic of the type II diabetes mellitus patients and control subjects is summarized in Table 1.

The inheritance of the specific haplotype combination defined by three single-nucleotide polymorphisms - SNP-43, -19, and -63 - was found to be associated with the three-folds increased risk. (Fig. 1) The polymorphism is in the non-coding region of calpain-10 and thus believed to increase the risk by regulating the transcription of calpain-10. The genotype of the SNP region in calpain-10 was carried out using the amplification refractory mutation system-PCR.

SNP-43 was genotyped with PCR method and the restriction fragment length polymorphism (RFLP). The modified SNP-43 primer was used because the original sequence did not have Nsi I (Promega, in U. S. A) site. The modified primer created an Nsi I restriction site in the presence of an allele and thus can be detected by the restriction enxyme digestion. The modified SNP-43 has 144base pairs (bp), which is smaller than the original primer (495bp). The digestion of Nsi I generates two DNA

fragment, 121pb and 23bp. The size of Allele 1 (G/G) was 144bp. The Allele (A/A) was 121bp and 23bp (Fig.2).

SNP-19 is the two-allele insertion/deletion (indel) polymorphism containing two or three copies of the 32bp repeated sequence. Indel-19 (simple tandem repeat insertion/deletion variant with either two or three copies of 32bp fragment) was genotyped by amplifying the repeat region and measuring the size by gel electrophoresis. The size of allele 1 containing 2 copies of 32bp was 155bp. The allele 2 containing 3 copies of 32bp was 187bp (Fig. 3).

SNP-63 was also genotyped using the PCR method and the restriction fragment length polymorphism (RFLP). The forward primer with the mismatch (5′- AAGGGGGGCCAGGGCCTGACGGGGGTGGCG-3′, with mismatch underlined) was used in addition to the unmodified reverse primer. The modified primer created the Hha I restriction site in the C allele, which is detected by the restriction enzyme digestion. The Hhal I digenstion generated the fragments, 162bp and 30bp. The size of Allele 1 (C) was 162bp and allele 2 (T) was 192bp. (Fig. 4)

2. The sequencing of the PCR products

Genomic DNA was isolated from peripheral blood-lymphocytes and was used for the genotyping of calpain-10 polymorphism. Most of PCR products amplified specific SNP regions. PCR products of the terminal region were mutated in some case. Thus, the sequence of PCR products was confirmed using pGEM - T easy vector system. (Fig. 5)

Table 1 Clinical Characteristic of Subjects

	Control Subjects	Diabetic Subjects	P value
No. of Subjects	236	454	-
Sex (M/F)	124/112	230/222	-
Age (years)	62.6 ± 5.1	53.3 ±11.0	< 0.001
BMI (kg/m^2)	23.8 ± 2.7	24.4 ± 3.5	< 0.01
SBP (mmHg)	129.7 ± 16.4	134.3 ± 19.1	< 0.001
DBP (mmHg)	79.6 ± 9.1	82.4 ± 10.5	< 0.001
Fasting Glucose (mM/l)	5.1 ±0.5	8.3 ±2.5	< 0.001
Fasting Insulin (ug/ul)	-	11.6 ± 13.3	-
HOMA ins	-	62.7 ± 79.9	-
HOMA IR	-	4.2 ±5.4	-
HbA1c (%)	-	8.4 ± 3.3	-
T-chol (mg/dl)	196.5 ± 32.9	199.0 ± 40.0	0.392
TG (mg/dl)	121.8 ± 68.9	163.8 ± 90.6	< 0.001
HDL-chol (mg/dl)	52.1 ± 12.5	45.1 ± 13.6	< 0.001

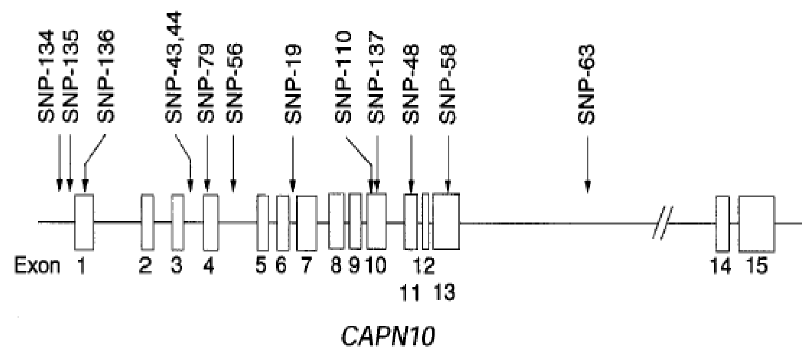


Fig. 1 Exon - intron organization of calpain-10. Exons are shown as numbered boxes together with the positions of the three SNPs that comprise the high - risk haplotype combination.

A

Forward region 5′-CACGCTTGCTGTGAAGTAATGC-3′
Reverse region 5′-CTCTGATTCCCATGGTCTGTAG-3′

B

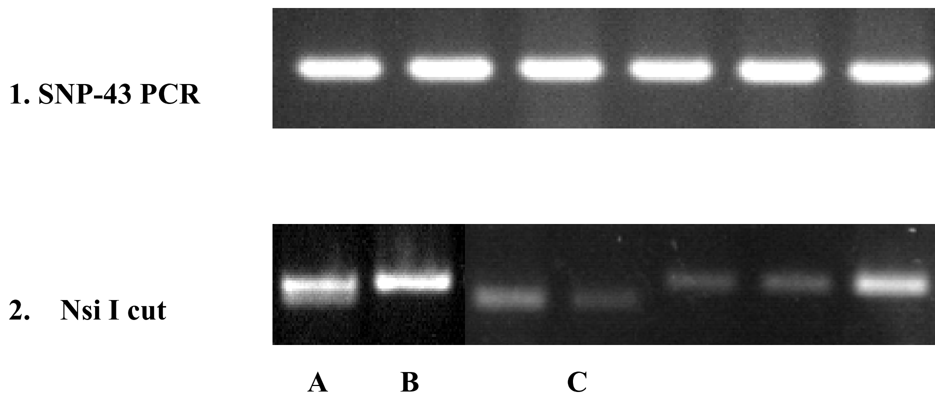


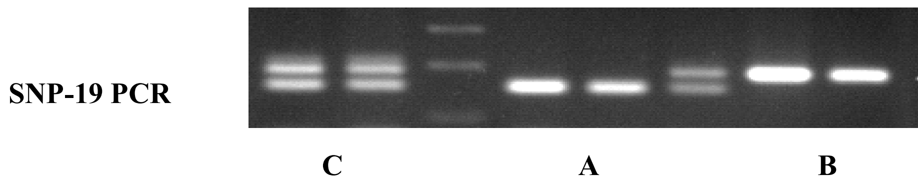
Fig. 2 SNP-43 genotyped by using PCR and RFLP. A. Modified primer for genotyping in PCR method and RFLP. The modified primer created an Nsi I restriction site in the presence of an allele, which is then detected by restriction enzyme digestion. B. These figure show genotype results in SNP-43 of calpain-10. B-1 was shown just PCR product result. Product size is 144bp. Another was shown digested Nsi I result. If product digested in Nsi I, result size confirmed 121bp and 23bp by gel electrophoresis. A: heterozygote, B: wild type, C: variant type

A

Forward region	5′-GTTTGGTTCTCTTCAGCGTGGAG -3′
Reverse region	5′-CATGAACCCTGGCAGGGTCTAAG-3′

B

SNP-19 PCR



C A B

A. Allele 1 (2 repeat of 32bp) : 155bp

B. Allele 2 (3 repeat of 32bp) : 187bp

C. Allele 3 (155bp & 187bp)

Fig. 3 SNP-19 genotyped by using PCR. A. This table box showed SNP-19 primer for carrying out PCR method. The PCR product showed a result in SNP-19 on a 3% agarose gel. Differently SNP-43 pattern, SNP-19 genotype confirmed divided bands. Allele 1 is 2 repeat of 32bp and this size is 155bp. The other pattern allele 2 is 3 repeat of 32bp and size is 187bp. Differently allele 1 and allele 2 pattern, allele 3 is show both 155pb and 187bp.

A

Forward region 5′-AAGGGGGGCCAGGGCCTGACGGGGGTGGCG-3′
Reverse region 5′-AGCACTCCCAGCTCCTGATC-3

B

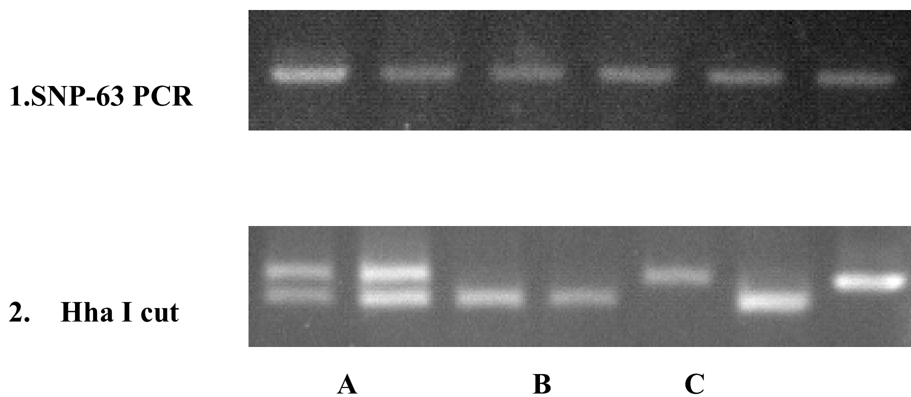


Fig. 4 SNP-63 genotyped by using PCR and RFLP. A. Modified primer for genotyping in PCR method and RFLP. The modified primer creates an Hha I restriction site in the presence of C allele, which is then detected by restriction enzyme digestion. B. These figure show genotype results in SNP-63 of calpain-10. B-1 was shown just PCR product result. Product size is 192bp. Another was shown digested Hha I result. If product digested in Hha I, result size confirmed 162bp and 30bp by gel electrophoresis. A: heterozygote, B: variant type, C: wild type

A. SNP-19 Sequence

1. Allele 1 of sequence in SNP-19

```
19   GTTTGGTTCTCTTCAGCGTGGAGAGATGATTCTGTCCCAGGAGCCGGGAGGAGGGTGATG
19P  GTTTGGTTCTCTTCAGCGTGGAGAGATGATTCTGTCCCAGGAGCCGGGAGGAGGGTGATG
19C  GTTTGGTTCTCTTCAGCGTGGAGAGATGATTCTGTCCCAGGAGCCGGGAGGAGGGTGATG
19   ATTCTGTCCCAGGAGCTGGGAGGAGGGTGGGCTTGTGGGAGGGGCTGGCTCTGTCTGTGG
19P  ATTCTGTCCCAGGAGCTGGGAGGAGGGTGGGCTTGTGGGAGGGGCTGGCTCTGTCTGTGG
19C  ATTCTGTCCCAGGAGCTGGGAGGAGGGTGGGCTTGTGGGAGGGGCTGGCTCTGTCTGTGG
19   CCGTAGCTGCTGCTTAGACCCTGCCAGGGTTCATG
19P  CCGTAGCTGCTGCTTAGACCCTGCCAGGGTTCATG
19C  CCGTAGCTGCTGCTTAGACCCTGCCAGGGTTCATG
```

2. Allele 2 of sequence in SNP-19

```
19   GTTTGGTTCTCTTCAGCGTGGAGAGATGATTCTGTCCCAGGAGCCGGGAGGAGGGTGATG
19P  GTTTGGTTCTCTTCAGCGTGGAGAGATGATTCTGTCCCAGGAGCCGGGAGGAGGGTGATG
19C  GTTTGGTTCTCTTCAGCGTGGAGAGATGATTCTGTCCCAGGAGCCGGGAGGAGGGTGATG
19   ATTCTGTCCCAG----------------------------------GAGCTGGGAGGAGGGTGG
19P  ATTCTGTCCCAGGAGCCGGGAGGAGGGTGAGATTCTGTCCCAGGAGCTGGGAGGAGGGTGG
19C  ATTCTGTCCCAGGAGCCGGGAGGAGGGTGAGATTCTGTCCCAGGAGCTGGGAGGAGGGTGG
19   GCTTGTGGGAGGGGCTGGCTCTGTCTGGCCGTAGCTGCTGCTTAGACCCTGCCAGGGTTCA
19P  GCTTGTGGGAGGGGCTGGCTCTGTCTGGCCGTAGCTGCTGCTTAGACCCTGCCAGGGTTCA
19C  GCTTGTGGGAGGGGCTGGCTCTGTCTGGCCGTAGCTGCTGCTTAGACCCTGCCAGGGTTCA
19   TG
19P  TG
19C  TG
```

B. SNP-43 of sequence

43 CACGCTTGCTGTGAAGTAATGCGTTTGAAGGTGAGGCTAAGCCTTGACTTGGTGAGGATGAGG

43N CACGCTTGCTGTGAAGTAATGCATTTGAAGGTGAGGCTAAGCCTTGACTTGGTGAGGATGAGG

43C CACGCTTGCTGCGAAGTAATGCGTTTGAAGGTGAGGCTAAGCCTTGACTTGGTGAGGATGAGG

43 AAGAAGGCAGAGGGGAGTAAAGAGGTGGGATTGAGGCAGCGGTTGGACGATTTGGGGTGCTA

43N AAGAAGGCAGAGGGGAGTAAAGAGGTGGGATTGAGGCAGCGGTTGGACGATTTGGGGTGCTA

43C AAGAAGGCAGAGGGGAGTAAAGAGGTGGGATTGAGGCAGCGGTTGGACGATTTGGGGTGCTA

43 CAGACCATGGGAATCAGAG

43N CAGACCATGGGAATCAGAG

43C CAGACCATGGGAATCAGAG

C. SNP-63 of sequence

63 AAGGGGGGCCAGGGCCTGACGGGGGTGGAGCGAGGGGGTGGGCCGCGTCTGTGCAGGCTC

63H AAGGGGGGCCAGGGCCTGACGGGGGTGGCCGCGAGGGG-TGGGCCGCGTCTCTGCAGGCTC

63C AAGGGGGGCCAGGGCCTGACGGGGGTGGAGCGAGGGG-TGGGCCGCGTCTCTGCAGGCTC

63 AAGAAGCTTCCTAAGAGGCTGGAGAGTGGAACCTTCAGGCACCACGCACTGCCTCCTCCC

63H AAGAAGCTTCCTAAGAGGCTGGAGAGTGGAACCTTCAGGCACCACGCACTGCCTCCTCCC

63C AAGAAGCTTCCTAAGAGGCTGGAGAGTGGAACCTTCAGGCACCACGCACTGCCTCCTCCC

63 TGCCCACGGTCCTGGGTTTCTCCAGATGGGGCCTTGGCCTTGGCTAGGTGTTGATCAGGA

63H TGCCCACGGTCCTGGG-TTCTCCAGATGGGGCCTTGGCCTTGGCTAGGTGTTGATCAGGA

63C TGCCCACGGTCCTGGG-TTCTCCAGATGGGGCCTTGGCCTTGGCTAGGTGTTGATCAGGA

Fig. 5 Sequence of calpain-10. SNP-43, -19, and -63 checked in sequencing by used vector cloning construct. A. SNP-19 divided in two type of allele. B. SNP-43 of sequence. C. SNP-63 of sequence. Each sequence primer indicated line. SNP site of mutation in sequence indicated shadow.

3. The assessment of the association of the previously described haplotypes of calpain-10 and type II diabetes

Horikawa et al.[1] have shown that the halotypes combination formed by SNP-43, -19 and -63 are associated with the risk of type II diabetes more than individual SNPs. The polymorphism consists of three haplotype. The allele consists of wild type and variant type. Table box showed calpain-10 of SNP regions for allele frequencies. (Table 2)

Table-3 showed the genotype distribution of the combination of SNP-43, -19 and -63 combinations in calpain-10. The analysis shows that the association of SNP-19 with type II diabetes is significant. (p=0.066) Table 4-1 showed the Calpain-10 of analysis by using PHASE software for haplotype allele distributaion and haplotype recombination. The results showed 8 groups. The genotype of 5 large groups was analyzed. The association of 3 groups was significant. (p=0.001) The association of these 3 groups with the risk of type II diabetes was analyzed by the logistic analysis. (Table 4-2) The allelic combination of haplotypes is shown in Table 5-1. 122/111 combination allele groups haven't this study of subject. Odd Ratio of the C haplotype combination (121/111) was 3.334 by the logistic analysis

adjusted by sex and age. (p=0.001) (Table 5-2) The association of L haplotype combination (112/112) was also significant. (p=0.047) The Odd Ratio was 2.381 by logistic analysis adjusted by sex and age. (p=0.047)

Table 2. Allele frequency of Calpain-10 of SNP-43, -19, and -63

Alleles					
SNP-43	G/A	GG	AA	DF	P
T2DM	53	388	13	2	0.402
Controls	34	198	4		
SNP-19	2A/3A	2A	3A	DF	P
T2DM	228	50	173	2	0.153
Controls	107	20	107		3.752
SNP-63	C/T	CC	TT	DF	P
T2DM	141	253	60	2	0.326
Controls	76	120	40		2.242

The polymorphic genotypes are: UCSNP-43, allele 1, G, allele 2, A, allele 0, heterozygote; UCSNP-19, allele 1, 2 repeats of 32bp sequence (named by "A"), allele 2, 3 repeats, allele 0, heterozygote; UCSNP-63, allele 1, C, allele 2, T, allele 0, heterozygote.

Table 3. Genotype distribution of Calpain-10 of SNP-43, -19, and -63

SNP-43	GG	AA+G/A	DF	P
T2DM	388	66	2	0.586
Controls	198	38		
SNP-19	2A+2A/3A	3A	DF	P
T2DM	281	173	2	0.066
Controls	127	107		3.369
SNP-63	CC	TT+C/T	DF	P
T2DM	253	201	2	0.222
Controls	120	116		

Table 4-1. Haplotype Allele Distributions from the result of PHASE

Haplotype	SNP-43	SNP-19	SNP-63	No of haplotypes
1	**1**	**2**	**1**	**668**
2	**1**	**2**	**2**	**118**
3	**1**	**1**	**1**	**186**
4	**1**	**1**	**2**	**287**
5	**2**	**2**	**1**	**105**
6	2	2	2	11
7	2	1	1	4
8	2	1	2	1

The genotypes of 3 SNPs are: 1: wild type, 2: variant type

Table 4-2 Haplotype analysis

1. Group by haplotype allele

Allele	**1**	**2**	**3**	**4**	**5**	**Total**
Control	243	42	38	107	33	463
T2DM	425	76	148	180	72	901
Total	668	118	186	287	105	1364

Statistic	**DF**	**Value**	**Probability**
Chi-Square	**4**	**18.78**	**0.001**

2. Logistic analysis of haplotype allele (adjusted by sex and age)

Haplotype allele	**Wald Chi-Square**	**Pr > Chi-Square**	**Odds Ratio**
2	0.03	0.870	1.034
3*	**16.21**	**0.001**	**2.227**
4	0.07	0.790	0.962
5	0.96	0.326	1.247

Allele* ; $p < 0.05$

Table 5-1. Haplotype allele combinations with haplotype allele 1-5

Haplotype-Genotypes	Haplotype allele combination
A	121/121
B	121/122
C	121/111
D	121/112
E	121/221
F	122/122
G	122/112
H	122/221
I	111/111
J	111/112
K	111/221
L	112/112
M	112/221
N	221/221

There is no case in the allele combination of 122/111.

Table 5-2 Haplotype combination analysis

1. Group by haplotype-genotype including allele 1(121)

Allele	1	2	3	4	5	Total
Control	69	11	24	52	18	174
T2DM	98	30	104	76	19	327
Total	167	41	128	128	37	501

2. Chi-Square results

Statistic	DF	Value	Probability
Chi-Square	**4**	**23.87**	**0.001**

3. Logistic analysis of haplotype-genotype (adjusted by sex and age)

Haplotype-genotype	Wald Chi-Square	Pr > Chi-Square	Odds Ratio
B	2.991	0.083	2.127
C*	**15.901**	**0.001**	**3.344**
D	**0.055**	**0.815**	**1.063**
E	0.002	0.965	1.029
F	0.063	0.802	0.829
G	0.140	0.709	1.156
H	2.432	0.119	5.882
I	2.670	0.102	3.610
J	2.069	0.150	2.463
K	1.963	0.161	2.725
L*	**3.868**	**0.049**	**2.381**
M	0.059	0.808	1.145
N	3.506	0.061	4.854

Allele* ; $p < 0.05$

IV. Discussion

Horikawa et al.[1] reported the unexpected result that in spite of being a member of the cystein protease family, calpains are cysteine proteases and are unlikely to be involved in the pathways of glucose homeostasis.[1-5] In the commentary, Permutt et al. noted that the connection between calpain-10 and glucose metabolism is far from clear.[1, 5, 9, 10] And yet, the finding that calpain-10 is a putative type II diabetes susceptibility gene was significant, because it represented the first gene associated with a common and complex disease to be identified by positional cloning.[1] However, the validation of these findings in other populations is needed to establish the role of the calpain-10 polymorphisms in susceptibility to type II diabetes.[2, 8, 12] In Pima Indians, although the homozygosity in the G allele at SNP-43 was not associated with the higher prevalence of type II diabetes, homozygote was associated with reduced calpain-10 mRNA level in the muscle and with the decreased rates of the glucose turnover that were suggestive of insulin resistance.[1, 3, 12, 14] Calpain-10 was the first identified gene involved in the type II diabetes susceptibility in Hispanics in Starr Count, Texas, by the linkage disequilibrium studies.[12, 14] The study showed the association of three SNPs in calpain-10 and type II diabetes. In this study performed tests

and the problem of a correction for multiple comparisons arises. The study had grouped subjects based on BMI, weighed age, and sex and subsequent comparisons were done. The correction was applied to the stratification analyses. Another important issue is the size of our study groups as. In the case-control genetic studies, the sample size was large enough to detect the assumed association. In this study, the sample size was large enough to detect the modest effect of calpain-10 on type II diabetes in the Korean population. This study has tested three haplotypes of calpain-10 - SNP-43, -19, and -63 - for their linkage and the association with type II diabetes. The haplotype was selected because of their prior association with type II diabetes individually or in combination, insulin resistance (SNP-43), and the transcriptional regulation of calpain-10 expression (SNP-43).

Here, the haplotypes consisting of the UCSNP-43, -19, and -63 combinations are associated with the risk of type II diabetes more than the individual SNPs. Horikawa et al.[1] observed the OR 2.80 in the 112/221 haplotype combination in 170 Mexican Americans families and the OR 3.58 in their second set of 69 Mexican American families.[1] In two other samples of white populations (German and Czech), the association of type II diabetes and calpain-10 polymorphisms was also found.[1, 6, 7] However, the association was not detected in the Polish, Chinese and Japanese population.[1, 6, 7,14]

This study evaluated the frequency of the combinations comprising these allele. This data showed that the 112/121 haplotype combination does not increase the risk of type II diabetes in Korean population. (Table 5-2)

Horikawa et al.[1] (2000) reported the OR 2.80 for the 112/121 haplotype combination in 170 Mexican - American families and the OR 3.58 in aother study with 69 Mexican - American families. In the Korean population, the association of the risk of type II diabetes and the allelic frequencies, the haplotype frequencies, or the combination of SNP-43, -19, and -63 was not detected. This study suggests that other haplotype may be associated with type II diabetes in Korean population. The haplotype frequency in Korean is different from Mexican - Americans, Europeans, and Asians.

This data show that the 121/111 haplotype combination appears to increase the risk of type II diabetes in Koreans: the OR (3.344, p=0.001) (Table 5-2) Once more, another data show that the 112/112 haplotype combination also appears to increase the risk of type II diabetes in Koreans: the OR (2.381, p=0.049) (Table 5-2) These findings, together with the observations in other populations raise the question whether calpain-10 is the diabetes susceptibility gene only in some particular populations. It is likely that evolutionary histories of individual populations have an important role in disease susceptibility loci of complex dividual effects.

Thus, the association of alleles at a specific locus to type II diabetes susceptibility will be likely to vary in different populations. The calpain-10 polymorphisms may interact with other genetic polymorphisms, whose individual effects on disease susceptibility could have been altered by differences in allele frequency among various populations. Such differences may result from the difference evolutionary history of the populations, as noted above. Alternately, the difference may arise from the interaction of the calpain-10 polymorphisms with variable environmental exposures. As a result of such varied gene- gene or gene- environment interaction, the effect of a functional variant could vary between populations, thereby leading to the variation in the disease susceptibility.

IV. Conclusion

This study data of the case-control study in the Korean population suggest the role of SNP-43, -19, and -63 of calpain-10. This study has tested three haplotypes of calpain-10 - SNP-43, -19, and -63 - for the linkage and the association with type II diabetes. The genotype was determined using polymerase chain reaction (PCR) and the restriction fragment length polymorphisms (RFLP). SNP-19 was analyzed by the PCR method. The other SNPs was analysis by the PCR method and the RFLP of PCR products. Base on the allelic type, all subjects were classified. Allele 0 represents the heterozygote of allele 1 and 2, allele 1 is wild type and allele 2 is the variant types. All data are analysed for the haplotype frequency and the haplotype recombination. This study data show that the haplotypes comprising the combination of UCSNP-43, -19, and -63 polymorphisms. Is associated with the risk of type II diabetes more than the haplotype with individual SNPs. The data show that the haplotype of the combination of 121/111 (p=0.001) and 112/112 (p=0.049) increased the risk of type II diabetes in Korean population. Taken together, this study suggests that in the population-based case-control studies, a large sample, will provide the better understanding of the contribution of calpin-10 makes to the diabetes risk in this population. This

study data show the association of the genetic variations in calpain-10 and the risk type II diabetes in Korean.

References

1. Horikawa Y, Oda N, Cox NJ, Li X, Orho-Melander M, Hara M et al. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nature Genet 2002; 26: 163-175.

2. Stumvoll M, Fritsche A, Madaus A, Stefan N, Weisser M, Machicao F et al. Functional significance of the UCSNP-43 polymorphism in the CAPN10 gene for proinsulin processing and insulin secretion in nondiabetic Germans. Diabetes 2001; 50:2161-2163.

3. Lynn S, Evans JC, White C, Frayling TM, Hattersley AT, Turnbull DM et al. Variation in the Calpain-10 gene affects blood glucose levels in the British population. Diabetes 2002; 51:247-249.

4. Garant M, Linda Kao WH, Brancati F, Coresh J, Rami TM, Hanis CL et al. SNP43 of CAPN10 and the risk of type 2 diabetes in African-Americans. Diabetes 2002; 51: 231-237.

5. Mclntyre EA, Walker M. Genetics of type 2 diabetes and insulin resistance:

knowledge from human studies. Clinical Endocrinology 2002; 57:303-311.

6. Tasi H, Sun G, Weeks DE, Kaushal R, Wolujexicz M, McGarvey ST et al. Type 2 diabetes an three calpain-10 gene polymorphisms in Samoans: No evidence of association. Am. J. Hum. Genet 2001; 69:1236-1244.

7. Fullerton SM, Bartoszewicz A, Ybazeta G, Horikawa Y, Bell GI, Kidd KK et al. Geographic and haplotype structure of candidate type 2 diabetes-susceptibility variants at the calpain-10 locus. Am. J. Hum. Genet. 2002; 70:1096-1106.

8. Elbein SC, Chu W, Ren Q, Hemphill C, Schay J, Cox NJ et al. Role of calpain-10 gene variants in familial type 2 diabetes in Caucasians. J Clin Endocrinol Metab 2002; 87(2): 650-654.

9. Orho-Melander M, Klannemark M, Svensson MK, Ridderstrale M, Lindgern C, Groop L. Variants in the calpain-10 gene predispose to insulin resistance and elevated free fatty acid levels. Diabetes 2002; 51:2658-2664.

10. Malecki MT, Moczulski DK, Klupa T, Wanic K, Cyganek K, Frey J et al. Homozygous combination of calpain-10 gene haplotypes is associated with

type 2 diabetes mellitus in a Polish population. Eur J Endocrinol 2002; 146(5): 695-9.

11. Hani EH, Hager J, Philippi A, Demenais F, Froguel P, Vionnet N. Mapping NIDDM susceptibility loci in French families. Diabetes 1997; 46: 1225-1226.

12. Horikawa Y, Oda N, Yu L, Imamura S, Fujiwara K, Makino M et al. Genetic variation in calpain-10 gene are not a major factor in the occurrence of type 2 diabetes in Japanese. J Clin Endocrinol Metab 2003; 88(1): 244-247.

13. Evans JC, Frayling TM, Cassell PG, Saker PJ, Hitman GA, Walker M et al. Studies of association between the gene for calpain-10 and type 2 diabetes mellitus in the United Kingdom. Am. J. Hum. Genet 2001; 69: 544-522.

14. Sreenan SK, Zhou YP, Otani K, Hanses PA, Currie KPM, Pan CY et al. Calpains play a role in insulin secretion and action. Diabetes 2001; 50: 2013-2020.

15. Hanis C, Boerwinkle E, Chakraborty R, Ellsworth D, Concannon P, Stirling B

et al. A genome-wide search for human non-insulin-dependent (type 2) diabetes genes reveals a major susceptibility locus on chromosome 2. Nature Genet 1996; 13: 161-166.

16. Cassell PG, Jackson AE, Evans JC, Ramachandran A, Snehalatha C, Hitman A. Haplotype combination of the calpain 10 gene polymorphisms is a determinant of abnormal glucose tolerance in South Indians (Abstract) Diabetes 2001; 50: A115.

17. Malecki MT, Klupa T, Katnrzyna AZ, Cyganek DK, Moczulski JS. Allelic variation in Calpain 10, PPAR [gamma] genes and risk of type 2 diabetes in a Polish population. (Abstract) Diabetes 2001; 50: A1000.

18. Baier LJ, Permana PA, Yang Z, Pratley RE, Hanson RL, Shen GQ et al. A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. J Clin Invest 2000; 106: R69-R73.

19. Shima Y, Nakanishi K, Odawara M, Kobayashi T, Ohta H. Association of the SNP-19 genotype 22 in the calpain-10 gene with elevated body mass index and hemoglobin A1c levels in Japanese. Clinica Chimica Acta 2003; 336:

89-96.

20. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 2001; 68: 978-989.

21. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet 2003; 73(6): 1162-9.

22. North BV, Curtis D, Cassell PG, Hitman GA, Sham PC. Assessing optimal neural network architecture for identifying disease-associated multi-marker genotypes using a permutation test, and application to calpain-10 polymorphisms associated with diabetes. Annals of Human Genetics 2003; 67: 348-356.

23. Stumvoll M, Wahl HG, Machicao F, Häring H. Insulin sensitivity of glucose disposal and lipolysis: no influence of common genetic variants in IRS-1 and CAPN-10. Diabetologia 2002; 45: 651-656.

국문요약

한국인 제 2형 당뇨병에서 알려진 Calpain-10 다형성의 유전적인 영향

김혜주

연세대학교 대학원

의과학과

(지도교수 : 이현철)

최근에 보고된 연구에서는 두 유전자의 다형현상은 최근에 제 2형 당뇨병의 복잡한 형태와 관련되었다: Calpain-10과 PPARγ. Calpain-10은 세포내의 단백질 분해효소의 큰 개체의 구성원이다. 멕시코계의 미국인과 다른 종족에서는, 어디에서나 발현되는 단백질의 세 가지의 단일염기 다형성인 -43, -19, -63의 변형이 제 2형 당뇨병을 유발되기 쉽도록 영향을 미친다는 것이 보고 되었다. 그러나, 실질적인 차이점은 그들의

기인하는 원인에서 뿐만 아니라 일배체형과 위험성 대립유전자에서 민족 특유의 집단에서 보여진다. 그러므로, 다양한 민족에서 calpain-10의 역할이 결정적으로 중요하다. 이 연구의 목적은 한국인에서 유전적으로 제 2형 당뇨병을 유발하기 쉬운 calpain-10의 단일염기 다형성 -43, -19 그리고 -63의 영향을 조사하는 것이다.

종합적으로, 739명 중 499명의 제 2형 당뇨병 환자와 240명의 대조군을 조사하였다. 이들의 그룹에서 calpain-10의 SNP-43, -19와 -63에 대한 유전자형을 확인하였다. 단일염기 다형성 -19는 크기에 의해서 agarose gel에서 중합효소연쇄반응 생성물을 전기영동을 수행함으로 확인했고, 반면에 제한효소 절편길이 다양성 방법은 두 다른 유전자 표지에 대해서 사용하였다. 그룹사이에서 대립유전자, 유전자형, 일배체형과 일배체형 조합 분포에서의 차이점은 버전 SAS 8.0에 의해서 조사되었다.

단일염기 다양성의 조사에 의해 확인된 세 곳의 유전자에서 대립유전자, 유전자형과 일배체형의 차이점은 그룹사이에서 큰 의미를 확인하게 되었다. 121/111 (OR=3.344d와 p=0.001)과 112/112 (OR=2.381와 p=0.049)의 일배체형 분포에서 차이점이 의미있게 발견되었다. 이와는 반대로, 112/121 일배체형 조합은 몇 몇의 인종에서 제 2형 당뇨병의 위험성을 증가시키는 것에 관여한다고 보고되었다.

이 연구의 결론으로 한국인에서 제 2형 당뇨병과 관련하여 단일염기 다형성 -43, 단일염기 다형성 -19와 단일염기 다형성 -63에 의해서 야

기되는 calpain-10의 다양한 일배체형 조합의 관계로 인해 당뇨병의 발병률이 증가할 수도 있다는 것을 추정할 수 있다. 이와 같은 연구를 바탕으로, 단일염기 다형성의 이상으로 야기되는 병리 생리를 연구하는데 기초가 될 것이다.

핵심 되는 말 : 제 2형 당뇨병, calpain-10, 제한효소 절편길이 다양성, 단일 염기 다양성