

ORIGINAL ARTICLE

Korean Circulation J 2006;36:229-235

ISSN 1738-5520

© 2006, The Korean Society of Circulation

혼합모형을 이용한 혈중 지질농도의 유전적 관련성 분석

연세대학교 의과대학 의학통계학과학교실,¹ 연세대학교 심혈관계질환 유전체 연구센터²
송기준¹ · 박찬미¹ · 임길섭¹ · 장양수² · 김동기¹

Genetic Association Analysis of Lipid Profiles Using Linear Mixed Model

Kijun Song, PhD¹, Chan Mi Park, MS¹, Kil Seob Lim, PhD¹,
Yang Soo Jang, MD, PhD² and Dong Kee Kim, PhD¹¹Department of Biostatistics and ²Cardiovascular Genome Center, Yonsei University College of Medicine, Seoul, Korea

ABSTRACT

Background and Objectives : Analyzing the association between multiple SNPs and the disease outcomes will provide new insight into the disease's etiology. However, this presents an analytic difficulty due to the large number of SNPs and the complex relationships among them. We proposed using the mixed model approach to identify the significant multi-locus genotypes and the high-order gene-to-gene interactions. **Subjects and Methods :** We described the mixed effects model and applied this approach to real world data. For the purpose of these analyses, we examine the association of four types of SNPs (AGT5, APOB, CETP3 and ACE6) with the lipid profiles and the measures related with cardiovascular disease. We used data from 672 healthy individuals (283 males and 389 females) who were without cardiovascular diseases. **Results :** The results of our analysis suggested that there were significant random genotype patterns and genotype groups according to the gender effect on the lipid profiles. In other words, there was significant variability across the genotype groups because of the effect of gender on the lipid profiles. **Conclusion :** The mixed model approach provided a flexible statistical framework for controlling potential confounding variables and for identifying a significant genetic contributions that may come about through the effects of multi-locus genotypes or through an interaction between the genotype and environmental variables (e.g. gender) with the variations in quantitative traits (e.g. lipid profiles). There were significant genetic contributions to the variability in the lipid profiles, and these were explained by the 4 SNPs described in our real data. (Korean Circulation J 2006;36:229-235)

KEY WORDS : SNPs ; Lipids ; Association.

서 론

심혈관계질환은 전세계적으로 손꼽히는 사망의 원인이며 혈중 지질농도(serum lipid profiles)는 이와 관련하여 주요한 위험 인자들로 알려져 왔다.¹⁾ 이러한 지질농도는 성별, 체질량지수(body mass index: BMI), 연령, 흡연 등의 환경적인 영향과 더불어 유전적 인자와 관련이 있는 것으로 보

고되고 있다.²⁻⁹⁾ 혈중 지질농도와 유전자간의 관련성(association)을 분석한 연구들에서 많은 단일 염기 다형성(single nucleotide polymorphisms: SNPs)들이 후보 유전자(candidate gene)로 보고되었는데, 그 연구 결과들이 반드시 일치하는 것은 아니었다.¹⁰⁻¹⁴⁾ 이러한 불일치는 연구 설계 방법, 연구 대상 집단 등의 차이에서 기인하는 것이 대부분이지만 연구의 초점을 각 다형성들의 개별적인 영향만을 파악하는데 맞추었기 때문에 발생할 수도 있다. 즉, 여러 단일 염기 다형성들간의 복잡한 조합에 따른 고차원(high-order)의 유전적 교호작용(interaction)이나 또는 유전자와 환경적(environmental) 요인들 사이의 교호작용 등을 간과했기 때문에 분석 결과가 다르게 나올 수 있다는 것이다.¹⁵⁾ 이러한 결과는 혈중 지질농도의 변화 혹은 그에 따

논문접수일 : 2005년 10월 6일

수정논문접수일 : 2006년 1월 9일

심사완료일 : 2006년 1월 24일

교신저자 : 김동기, 120-752 서울 서대문구 신촌동 134

연세대학교 의과대학 의학통계학과학교실

전화 : (02)2228-2491 · 전송 : (02)364-8037

E-mail : dkkimbios@yumc.yonsei.ac.kr

른 질병 발생 등에 대한 유전적 관련성 분석을 할 경우, 유전자간의 복잡한 교호작용과 유전자와 환경적 요인간의 교호작용 등을 반드시 고려할 필요가 있다는 것을 강하게 시사한다고 할 수 있다. 그런데, 고려하고자하는 단일 염기 다형성의 수가 많아지면 그것들 사이의 복잡한 교호작용에 의한 영향을 검증하는데에는 그 분석 방법 자체에 어려움이 발생 할 수 있다. 혈중 지질농도와 같이 양적 형질(quantitative trait)에 대한 다형성의 개별적인 관련성을 분석하기 위해 흔히 쓰이는 통계분석 방법으로 회귀분석(regression analysis) 방법이 있는데, 이것은 잠재적인 환경적 인자들의 영향을 통제하면서 유전자의 영향력을 손쉽게 분석할 수 있다는 장점을 가지고 있다. 그러나, 위에서 언급한 것처럼 많은 수의 다형성들에 대한 고차원의 유전적 교호작용을 분석할 경우, 각 다형성들의 유전형(genotype)에 대한 조합(combination)의 수가 극단적으로 많아지게 되고 그에 따른 변수의 수 또한 증가하게 되므로 주어진 표본 크기 내에서 분석이 불가능해질 수도 있다. 이에 본 연구에서는 혈중 지질농도와 같은 양적 형질에 대한 유전적 관련성을 분석함에 있어서, 많은 수의 다형성들간의 교호작용 및 환경적 인자와의 교호작용을 효율적으로 분석할 수 있는 방법으로 혼합모형(mixed model)을 이용한 방법을 제안하고자 한다. 이 방법은 여러 다형성들의 유전형의 조합에 따른 영향을 확률적 효과(random effect)로 설정하여 이러한 조합에 의한 지질농도의 변동량(variability)이 유의한지 평가하며, 아울러 이때 잠재적으로 영향을 줄 수 있는 변수들을 통제할 수 있는 융통성(flexibility)을 확보하고 있는 방법이다. 또한 본 연구에서는 실제 유전자 자료에서 서로 다른 염색체상에 존재하는 4개의 다형성들을 이용하여 혈중 지질농도에 대한 유전자들 간의 고차원적인 교호작용의 유의성을 평가하고자 한다.

대상 및 방법

본 연구에서는 연세대학교 심혈관계질환 유전체 연구센터에서 실시하는 건강검진에 참여하여 조사된 자료를 이용했는데, 이 때 심혈관계질환에 대한 위험인자를 가지고 있지 않고 또한 당뇨, 고혈압, 뇌혈관질환, 말초혈관질환 등이 없으며 이러한 질환들에 대한 가족력이 없고 아울러 지질 강하제나 호르몬 제제 등을 복용하지 않는 1,419명의 건강한 정상 성인들을 연구 대상으로 하였다. 이러한 대상들은 심전도 검사 결과 정상소견을 보이고 흉부 X-선상 심비대 소견이 없는 경우였다. 단일 염기 다형성으로서는 1번 염색체에 위치한 AGT5(Angiotensinogen M235T), 2번 염색체에 있는 APOB(Apolipoprotein B C-516T), 16번 염색체상의 CETP3(Cholesteryl ester transfer protein TAQ1B)과 17번 염색체에 존재하는 ACE6(Angiotensin Converting Enzyme Alu I/D)의 4가지를 이용했는데, 연구 대상들은 이 단일 염

기 다형성들의 유전형(genotype)을 모두 알고 있는 경우로서 최종 분석에 포함된 대상은 672명으로 구성되었다. 이들에 관한 유전형 분석을 위해 DNA는 DNA isolation kit (Gentra Genomic DNA purification kit, MN, USA)를 이용하여 5 mL 혈액 내의 백혈구 세포로부터 추출하였다. 다형성 분석은 SNP stream 25K system(Orchid Biosciences, Princeton, NJ)을 이용하여 SNP-IT(SNP-Identification Technology) assay 방법으로 단일염기확장 반응을 수행하였고, primer와 각각의 PCR 생성물을 재생시킨 후 biotin 또는 fluorescein isothiocyanate(FITC)으로 표지된 각각의 dideoxynucleotide와 반응시켜 단일염기를 확장하였다. 반응이 끝난 후 결과분석은 ELISA 판독기를 이용하여 분석하였으며 최종 유전자형은 QC Review™ program을 이용하여 확인하였다. 혈중 지질농도들로서 총 콜레스테롤, 중성지방, LDL-콜레스테롤, HDL-콜레스테롤의 수치를 이용하였다.

본 연구에서는 단일 염기 다형성들의 유전형의 조합에 따른 교호작용을 분석하기 위해 혼합모형에 기초한 방법을 적용하는데 그 내용을 간략히 설명하면 다음과 같다. 먼저, 다중의 유전소(multi-locus)에서 동일한 유전형의 조합을 갖고 있는 대상들을 하나의 유전형 집단(genotype group)으로 정의하는데, 이것은 각 개인들이 갖고 있는 유전형의 조합을 하나의 변수로 나타내어 그 조합 자체가 혈중 지질농도에 어떻게 영향을 미치고 있는지 확인하고자 하는 이유에서 설정되는 것이다.¹⁶⁾ 또한 이런 식으로 유전형 집단을 정의하는 것은 비록 각 단일 염기 다형성의 개별적인 영향이 존재하지 않을지라도 조합을 구성하고 있는 유전자들이 어떻게 상호작용을 하고 있는지에 관한 정보를 제공해 줄 수 있다는 장점을 가지고 있다. 유전형 집단을 구성하고 있는 조합의 수는 예를 들어, 두 개의 대립유전자(allele) A, B로 구성되어 있는 하나의 단일 염기 다형성이 있다고 하면 가능한 유전형의 수는 3가지(AA, AB, BB)가 되고, C, D의 대립유전자를 갖는 또 다른 다형성이 있을 경우, 이 두 가지에 따른 조합은 총 9가지(=3²; AACB, AACD, AADD, ... , BBCC, BBBD, BBDD)의 경우가 나올 수 있다. 같은 방법으로 다형성의 수가 만약 n개이면 가능한 모든 유전형의 조합의 수는 가지가 나올 수 있다. 이런 식으로 관찰된 특정한 유전형 집단들은 그러한 유전형들을 가진 일반 인구집단으로부터 확률적으로 얻어진 표본이라고 생각할 수 있고, 이를 다음의 식 (1)과 같은 혼합모형에서 확률적 효과로 설정한다.

$$Y = X\beta + Z\gamma + \varepsilon \dots\dots\dots (1)$$

위의 식 (1)은 혼합모형을 나타내는 일반적인 식으로서, Xβ 부분으로 표현되는 고정 효과(fixed effect)와 Zγ 부분으로 나타내어지는 확률적 효과가 동시에 혼합되어 표현되

는 모형이다.¹⁷⁻²⁰⁾ 여기에서 Y는 혈중 지질농도와 같은 양적 형질을 말하는 것이고, X는 성별, 연령 등과 같이 개인별 특성 변수를 나타내며 β 는 X에 대한 회귀계수(regression coefficient)이다. 또한 Z는 본 연구에서 정의한 유전형 집단 변수처럼 확률적 효과를 나타내는 변수이고 γ 는 그에 따른 회귀계수를 말하며, ε 은 모형의 오차항(error term)을 나타낸다. 위의 혼합모형에서 γ 와 ε 은 각각 정규분포(normal distribution)를 따르며 서로 간에는 독립이라고 가정한다.

본 연구에서는 이러한 혼합모형에서 고정 효과로서 연령, 체질량지수, 흡연력, 음주력, 운동 여부 등의 환경적 요인들을 고려하였고, 확률적 효과로서 유전형 집단 및 유전형 집단과 성별과의 교호작용을 포함시켰다. 혈중 지질농도에서 유전형 집단에 의한 포괄적인(overall) 효과가 존재하는 지 또는 유전형 집단과 환경적 요인간의 교호작용이 존재하는 지 분석한다. 이것은 유전형 집단에 따른 혈중 지질농도들의 분산(variance) 또는 유전형 집단과 환경적 요인간의 교호작용에 따른 공분산(covariance)이 0이 아닌지를 우도비 검정(likelihood ratio test)을 통하여 평가할 수 있다.¹⁷⁾ 만약, 검정 결과 추정된 분산 혹은 공분산이 통계학적으로 0이 아니라면 이러한 요인들이 혈중 지질농도에 영향을 미치고 있는 것으로 판단한다. 아울러 포괄적인 유전형 집단 혹은 교호작용에 의한 효과들이 통계학적으로 유의하다면 구체적으로 어떤 유전형 집단이 다른 것들과 다른 양상으로 나타나는지 분석하는데, 이것은 특별한 유전형 집단 혹은 교호작용에 대한 회귀계수 추정치가 0과 다른지를 개별적으로 검정하여 평가한다. 이러한 혼합모형을 이용하여 실제 자료를 분석하기 위해 SAS ver. 8.1(SAS Institute, Cary, NC)을 이용하였고, 이 때 p값이 0.05 미만인 경우를 통계학적으로 유의한 것으로 평가하였다.

결 과

연구 대상은 4가지의 단일 염기 다형성들의 유전형과 환경적 요인들의 정보를 모두 알고 있는 개인들로서, 남자 283명(42.1%), 여자 389명(57.9%) 총 672명을 대상으로 하였다. 분석하고자 하는 혈중 지질농도에서 총 콜레스테롤은 208.8 ± 38.5 , 중성지방은 142.3 ± 122.7 , LDL-콜레스테롤은 131.5 ± 33.5 , HDL-콜레스테롤은 48.9 ± 11.8 의 분포를 보여주고 있었다(Table 1). 전체 대상에서 환경적 요인들 중 평균 연령은 47.7 ± 10.4 세로 분포하고 있었고 체질량지수는 평균 23.6 ± 2.9 이고, 현재 흡연을 하고 있는 경우는 258명(38.4%), 현재 음주를 하는 경우는 387명(57.6%), 현재 규칙적으로 운동을 하는 경우는 236명(35.1%)의 분포를 보여주고 있었다(Table 1). 4가지의 단일 염기 다형성에서 각각의 유전형에 따른 빈도는 다음과 같다(Table 2). 유전형 집단은 4가지의 다형성에 대하여 구성할 수 있는 가능한 $81(=3^4)$ 가지 중 본 연구의 대상들에서는 실제로 총 50가지의 경우

가 관찰되었는데, 그 중에서 집단별로 적어도 10명 이상의 대상을 포함하는 경우는 21가지로서 584명(86.9%)이 포함되어 있었다(Table 3). 또한 유전형 집단에서 다형성 AGT5, APOB, CETP3, ACE6의 순서대로 유전형이 (CC, CC, GA, ID)인 경우가 72명(10.7%), (CC, CC, GG, ID)인 경우는 70명(10.4%)으로서 가장 빈번하게 나타나는 조합이었다(Table 3). 혼합모형을 이용한 분석에서 연령, 체질량지수, 흡연력, 음주력, 운동 여부의 환경적 요인들을 고정효과로서 모형에 포함시켰을 때, 유전형 집단 및 유전형 집단과 성별과의 교호작용에 대한 확률적 효과의 유의성을 평가하였는데, 그 결과는 다음과 같다(Table 4). 먼저, 총 콜레스테롤에서 포괄적인 유전형 집단의 효과(variance estimate=15.9)가 유의하게 나타났으며($p=0.024$), 성별과의 교호작용에 대한 효과(covariance estimate=1.3) 또한 유의하게 나타났다($p=0.036$). 중성지방의 경우에는 교호작용의 효과(covariance estimate=-531.2)가 유의하게 나타났고($p=0.004$), LDL-콜레스테롤에서는 유의한 효과가 존재하지 않았다. HDL-콜

Table 1. Distribution of lipid profiles and environmental factors

	Male (n=283)	Female (n=389)	Total (n=672)
Total cholesterol	207.8±40.1	209.6±37.3	208.8±38.5
Triglyceride	168.3±139.7	123.5±104.9	142.3±122.7
LDL-C	130.5±34.9	132.3±32.4	131.5±33.5
HDL-C	45.5±10.6	51.5±12.1	48.9±11.8
Age	47.9±11.1	47.6±9.9	47.7±10.4
BMI	24.1±2.7	23.3±3.1	23.6±2.9
Current smoker	233 (82.3)	25 (6.4)	258 (38.4)
Current drinker	235 (83.0)	152 (39.1)	387 (57.6)
Exercise	123 (43.5)	113 (29.1)	236 (35.1)

Values are Mean±SD and frequencies (percent). LDL-C: low density lipoprotein-cholesterol, HDL-C: high density lipoprotein-cholesterol, BMI: body mass index

Table 2. Frequency of four single nucleotide polymorphism genotypes

SNP	Genotype	Frequency (%)		
		in male (n=283)	in female (n=389)	in total (n=672)
AGT5	CC	186 (65.7)	257 (66.1)	443 (65.9)
	CT	89 (31.5)	122 (31.3)	211 (31.4)
	TT	8 (2.8)	10 (2.6)	18 (2.7)
APOB	CC	210 (74.2)	306 (78.7)	516 (76.8)
	CT	67 (23.7)	81 (20.8)	148 (22.0)
	TT	6 (2.1)	2 (0.5)	8 (1.2)
CETP3	AA	43 (15.2)	54 (13.9)	97 (14.4)
	GA	142 (50.2)	181 (46.5)	323 (48.1)
	GG	98 (34.6)	154 (39.6)	252 (37.5)
ACE6	DD	52 (18.4)	59 (15.2)	111 (16.5)
	ID	139 (49.1)	194 (49.9)	333 (49.6)
	II	92 (32.5)	136 (34.9)	228 (33.9)

SNP: single nucleotide polymorphism, AGT5: angiotensinogen M235T, APOB: apolipoprotein B C-516T, CETP3: cholesteryl ester transfer protein TAQ1B, ACE6: angiotensin converting enzyme Alu I/D

레스테롤에서는 유전형 집단의 효과(variance estimate=2.8)가 유의하게 존재하고 있었으며(p=0.043), 성별과의 교호작용에 대한 효과(covariance estimate=4.1)도 유의하게 나타났다(p=0.037). 이러한 분석결과를 토대로 구체적으로 어떠한 조합의 유전형 집단에 의한 효과가 유의하게 나타나는 지 그리고 그러한 효과가 성별에 따른 차이를 보이는 지를 개별적으로 평가하였는데 그 결과는 다음과 같다(Table 5). 여기에서 다형성 AGT5, APOB, CETP3, ACE6의 순서대로 유전형이 (CC, CC, AA, ID)인 경우 총 콜레스테롤에서 유전형 집단과 성별과의 교호작용이 유의하게 나타났는데 (coefficient estimate=-5.8, p=0.021), 이것은 이러한 유전형의 조합을 가지고 있을 때, 여자에 비해 남자에서 총 콜레스테롤 수치가 더 낮게 나타난다는 것과 관련이 있음을 의미

한다. 이런 식으로 각 혈중 지질농도별로 유의한 관련성이 존재하는 유전형 집단들을 정리해보면 먼저, 총 콜레스테롤에서는 (CC, CC, AA, ID), (CC, CC, GG, ID), (CT, CC, GG, II)의 경우, 성별과의 교호작용에 대한 회귀계수 추정치가 -5.8(p=0.021), -4.7(p=0.029), -3.2(p=0.024)로서 남자에서 총 콜레스테롤 수치가 더 낮은 것과 관련성을 보이고 있었다. 또한 유전형(CT, CT, AA, ID)에서는 반대의 경향을 보이고 있었으며(coefficient estimate=6.7, p=0.019), (CC, CT, GA, II)에서는 성별에 따른 차이는 없지만 포괄적인 관련성이 존재하고 있었다(coefficient estimate=3.1, p=0.033). 중성지방에서는 (CC, CC, AA, ID), (CC, CC, GA, II), (CC, CT, GG, ID), (CC, CT, GG, II)들에서 교호작용에 대한 회귀계수 추정치가 각각 -26.1(p=0.034), -34.3(p=0.007), -25.9(p=0.028), -41.3(p=0.017)으로서 남자에서 중성지방 수치가 더 낮은 것과 관련성을 보이고 있었으며, (CC, CC, AA, II), (CT, CT, AA, ID) 및 (CT, CT, GG, ID)에서는 그와 반대의 관련성이 유의하게 존재하고 있었다 (coefficient estimate=54.3, 43.1, 58.5, p=0.004, 0.016, 0.006). HDL-콜레스테롤에서는 (CC, CC, GG, II)와 (CT, CC, GA, ID)에서 유전형 집단의 효과에 대한 회귀계수 추정치가 각각 0.8(p=0.026), 0.9(p=0.019)로서 성별에 상관없이 포괄적인 관련성이 존재하고 있었으며, 또한 성별과의 교호작용에서는 회귀계수 추정치가 -1.7(p=0.024)과 -1.9(p=0.020)로서 남자에서 HDL-콜레스테롤 수치가 더 낮은 것과 관련성을 보이고 있었다.

Table 3. Genotype groups with at least ten observations

Genotype group pattern				Frequency (%)
AGT5	APOB	CETP3	ACE6	
CC	CC	AA	DD	11 (1.9)
CC	CC	AA	ID	18 (3.1)
CC	CC	AA	II	20 (3.4)
CC	CC	GA	DD	24 (4.1)
CC	CC	GA	ID	72 (12.3)
CC	CC	GA	II	62 (10.6)
CC	CC	GG	DD	24 (4.1)
CC	CC	GG	ID	70 (12.0)
CC	CC	GG	II	33 (5.7)
CC	CT	GA	ID	20 (3.4)
CC	CT	GA	II	17 (2.9)
CC	CT	GG	ID	33 (5.7)
CC	CT	GG	II	10 (1.7)
CT	CC	AA	ID	15 (2.6)
CT	CC	GA	DD	14 (2.4)
CT	CC	GA	ID	40 (6.9)
CT	CC	GA	II	29 (5.0)
CT	CC	GG	DD	13 (2.2)
CT	CC	GG	ID	25 (4.3)
CT	CC	GG	II	24 (4.1)
CT	CT	GA	ID	10 (1.7)

AGT5: angiotensinogen M235T, APOB: apolipoprotein B C-516T, CETP3: cholesteryl ester transfer protein TAQ1B, ACE6: angiotensin converting enzyme Alu I/D

고 찰

본 연구에서는 혈중 지질농도와 같은 양적 형질에 대한 유전적 관련성을 분석함에 있어서, 여러 개의 다형성들간의 교호작용 및 환경적 요인과의 교호작용을 효율적으로 분석할 수 있는 방법으로 혼합모형을 이용한 방법을 제안하였다. 혼합모형을 이용한 분석은 복수의 다형성에 의한 영향 및 그것들과 환경적 요인과의 교호작용을 분석하면서 동시에 잠재적으로 영향을 줄 수 있는 혼란변수(confounder)들을 보정할(adjusting)수 있는 효율적인 방법이다. 특히 다형성들의 유전형의 조합에 따른 유전형 집단을 정의하여 이를 확률적 효과로 혼합모형에 포함시킨 후, 그것들이 혈중 지질농도와

Table 4. Significance of random effects in mixed model analysis*

Lipid profiles	Variance-covariance estimates (SE) of random effects					
	Genotype group	p	Sex	p	Genotype group × sex	p
Total cholesterol	15.9 (8.1)	0.024 [†]	42.4 (20.2)	0.018 [†]	1.3 (0.7)	0.036 [†]
Triglyceride	7.2 (8.4)	0.196	1780.3 (774.1)	0.011 [†]	-531.2 (183.4)	0.004 [†]
LDL-C	1.4 (31.4)	0.482	0.1 (4.9)	0.491	-0.2 (25.1)	0.992
HDL-C	2.8 (1.6)	0.043 [†]	7.2 (4.9)	0.072	4.1 (2.3)	0.037 [†]

*: analysis was performed including variables of fixed effects (age, BMI, current smoker, current drinker, exercise), †: p<0.05. SE: standard error, LDL-C: low density lipoprotein-cholesterol, HDL-C: high density lipoprotein-cholesterol

Table 5. Specific genotype patterns with significant random effects

Genotype group pattern				Lipid profile	Genotype*	p	Interaction by sex [†]	p
AGT5	APOB	CETP3	ACE6					
CC	CC	AA	ID	TC	-2.4	0.305	-5.8	0.021
				TG	13.9	0.399	-26.1	0.034
CC	CC	AA	II	TG	-26.3	0.101	54.3	0.004
CC	CC	GA	II	TG	37.4	0.147	-34.3	0.007
CC	CC	GG	ID	TC	-1.1	0.369	-4.7	0.029
CC	CC	GG	II	HDL	0.8	0.026	-1.7	0.024
CC	CT	GA	II	TC	3.1	0.033	1.8	0.401
CC	CT	GG	ID	TG	12.2	0.361	-25.9	0.028
CC	CT	GG	II	TG	1.1	0.960	-41.3	0.017
CT	CT	AA	ID	TC	2.6	0.283	6.7	0.019
				TG	-16.9	0.261	43.1	0.016
CT	CC	GA	ID	HDL	0.9	0.019	-1.9	0.020
CT	CC	GG	II	TC	0.6	0.871	-3.2	0.024
CT	CT	GG	ID	TG	-24.2	0.128	58.5	0.006

*: estimates of regression coefficients of random genotype pattern effects, †: estimates of regression coefficients of random genotype pattern by sex (reference group: female) effects. AGT5: angiotensinogen M235T, APOB: apolipoprotein B C-516T, CETP3: cholesteryl ester transfer protein TAQ1B, ACE6: angiotensin converting enzyme Alu I/D, TC: total cholesterol, TG: triglyceride, HDL: high density lipoprotein-cholesterol

같은 양적 형질의 변동량에 얼마나 기여할 수 있는지를 평가하여 유전적 관련성이 존재하는 지를 분석할 수 있다. 이렇게 확률적 효과로서 유전형 집단을 정의하여 분석하는 것은 여러 다른 유전자내의 다형성들의 영향을 동시에 평가할 수 있다는 장점을 갖고 있고, 또한 다형성들 각각의 개별적인 효과의 존재 여부에 상관없이 유전형의 조합에 의한 교호작용의 효과를 분석할 수 있다는 장점을 가지고 있다.

본 연구에서는 실제자료를 이용한 분석에서 각기 다른 염색체상에 존재하는 네 개의 다형성들(AGT5, APOB, CETP3, ACE6)로 구성된 유전형 집단을 확률적 효과로 정의한 후, 그것들이 혈중 지질농도의 변동량에 대한 기여도를 평가하여 유전적 관련성이 존재하는 지를 혼합모형의 틀 안에서 분석하였다. 그 때 연령, 체질량지수, 흡연력, 음주력, 운동 여부 등의 환경적 요인들을 고려하여 분석을 수행하였다. 분석 결과 혈중 지질농도의 변동량에 유의하게 기여하는 구체적인 유전형의 조합에서 그 집단을 구성하고 있는 다형성을 살펴보면 먼저, AGT5에서는 유전형이 CC인 경우가 가장 많이 포함되어 있고 APOB의 경우 CC의 유전형이, CETP3에서는 GG가 가장 많이 나타나고 있었으며, ACE6에서는 ID와 II의 유전형이 대등하게 분포하고 있었다. 또한 AGT5와 APOB에서 유전형이 TT인 경우는 유의한 유전형 집단에 포함되지 않았고 ACE6에서도 유전형이 DD인 경우는 나타나지 않았다. 유의하게 나타난 유전형 집단들에서 각 조합에 해당하는 연구 대상의 수는 비교적 높은 빈도를 보이고 있었는데, 유전형이(CT, CT, AA, ID)인 경우는 5명, (CT, CT, GG, ID)인 경우에는 4명으로 상대적으로 적은 수의 대상을 포함하고 있는 집단에서도 유의하게 기여하는 것을 알 수 있었다.

본 연구에서 분석에 이용된 다형성들은 국내·외에서 기

존의 심혈관계질환과의 관련성에 관한 연구가 다형성 각각에 대하여 개별적으로 빈번하게 진행된 것들이었다.²¹⁻²⁹⁾ 먼저 ACE6(Angiotensin Converting Enzyme Alu I/D)의 경우, 김 등²¹⁾에 의하면 남자의 경우에서 총콜레스테롤 수치와의 관련성이 존재하는 것으로 보고되었다. 또한 이 등²²⁾의 연구에서는 총콜레스테롤 수치 및 허혈성 심질환과의 관련성이 존재하는 것으로 보고되었으며, 박 등²³⁾의 연구에서도 허혈성 심질환 등과의 관련성이 보고되었고, Nakai 등²⁴⁾의 연구에서는 관상동맥질환 발생의 위험 인자로 보고되었다. AGT5(Angiotensinogen M235T)의 경우에는 Araujo 등²⁵⁾이 급성 심근경색증과의 관련성이 존재함을 보고하였으며, Kumar 등²⁶⁾과 Say 등²⁷⁾은 고혈압 발생과 관련이 있음을 보고하였다. Fidani 등²⁸⁾은 CETP3(Cholesteryl ester transfer protein TAQ1B)과 허혈성 뇌졸중과의 관련성이 존재하지 않음을 보고하였다.

그런데 위의 연구들에서 각 다형성에 대한 분석결과가 반드시 일치하는 것은 아니었다. 또한 이러한 기존의 연구들은 여러 다형성들간의 복잡한 조합에 따른 유전적 교호작용이나 또는 유전자와 환경적 요인들간의 교호작용을 고려한 것이 아니라 각 다형성들의 개별적인 관련성만을 파악하는데 그 목적을 두고 있는 것이 대부분이다. 즉, 하나의 다형성에서 동일한 유전형을 갖고 있지만 다른 여러 개의 다형성들의 유전형이 어떻게 구성되어지는가에 따라 질병 발생에 기여하는 정도가 다르게 나타날 수 있다는 가능성을 고려하지 않은 것이라 할 수 있다. 이러한 사실은 본 연구의 결과를 통해서 확인할 수 있었는데, 특히 동일한 유전형 집단이라고 해도 그 기여도가 성별과 같은 특정한 환경적 요인과의 교호작용에 의해 차이를 보일 수 있다는 점을 알 수 있었다.

본 연구에서 제안된 분석 방법은 상용화된 통계분석 프로그램들(SAS, SPSS 등)을 이용하여 기존의 혼합모형의 틀 내에서 쉽게 적용할 수 있고, 환경적인 영향 등에 따른 혼란 변수들을 고려하여 분석할 수 있다는 장점을 가지고 있다. 그런데, 한정된 표본 크기에서 유전형 집단을 구성하는 조합의 수가 과도하게 많아지는 경우 즉, 유전형 집단의 수가 너무 많아서 각 집단에 속한 연구 대상의 수가 극소수로 존재하는 경우에는, 특별한 유전형 집단의 영향을 구분해 내어 파악하기가 쉽지 않게 된다. 이런 상황에서는 많은 수의 다형성들을 몇 개의 하부집합(subset)으로 묶어서 그에 따른 유전형 집단 변수를 각각 정의하여 분석에 이용할 수 있다. 예를 들어, 두 개의 대립유전자로 구성된 10개의 다형성의 정보를 동시에 이용하여 유전형 집단을 정의할 경우 가능한 총 집단의 수는 가지가 나올 수 있다. 이 경우 하나의 유전형 집단 변수만을 정의하게 되면 표본 크기가 방대하지 않을 때에는 각 조합에 속하는 연구 대상의 수가 극단적으로 작아지게 되어 개별 유전형 집단의 구체적인 효과를 측정하는 것이 불가능해질 수도 있다. 이 때 10개의 다형성에 대하여 3개 정도의 하부집합(다형성을 3개, 3개, 4개로 묶음)을 구성한 후, 그에 따른 유전형 집단 변수 3개를 정의하여 이것들을 혼합모형에서 확률적 효과로 포함시켜 분석하는 것이다. 이러한 과정에서 하부집합을 어떻게 구성할 것인가의 문제는 형질과 각 다형성들간의 개별적인 관련성에 관한 사전 정보를 이용할 수 있다.

본 연구에서 적용된 분석 방법은 양적 형질을 주 관심사로 하고 있는데, 만약 특정한 질병의 유·무 등과 같이 이분형의(binary) 질적 형질(qualitative trait)이 관심 변수라면 본 연구의 방법을 확장시킨 일반화 혼합모형(generalized linear mixed model) 등을 이용한 분석이 가능할 것이라 여겨진다.¹⁷⁾

요 약

배경 및 목적 :

양적 형질에 대한 유전적 관련성 분석을 할 경우, 유전자 간의 복잡한 교호작용과 유전자와 환경적 요인간의 교호작용 등을 고려할 필요가 있는데, 만약 단일 염기 다형성의 수가 많아지면 그것들 사이의 복잡한 구조에 의한 영향을 파악하는 데에는 분석 방법 자체에 어려움이 발생할 수 있다. 이에 본 연구에서는 이를 효율적으로 극복할 수 있는 분석 방법으로 혼합모형을 이용한 방법을 제안한다. 또한 본 연구에서는 실제 4개의 단일 염기 다형성들을 이용하여 혈중 지질농도에 대한 유전자들 간의 고차원적인 교호작용의 유의성을 평가한다.

방 법 :

본 연구에서는 혼합모형에서 고정 효과로서 연령, 성별, 체질량지수, 흡연력, 음주력, 운동 여부 등의 환경적 요인들을

고려하였고, 확률적 효과로서 유전형 집단 및 유전형 집단과 성별과의 교호작용을 포함시켰다.

단일 염기 다형성으로서는 AGT5, APOB, CETP3과 ACE6의 4가지를 이용했고, 혈중 지질농도들로서 총 콜레스테롤, 중성지방, LDL-콜레스테롤, HDL-콜레스테롤의 수치를 이용하였다. 심혈관계질환에 대한 위험인자를 가지고 있지 않은 남자 283명(42.1%), 여자 389명(57.9%), 총 672명의 건강한 정상 성인들을 대상으로 분석하였다.

결 과 :

혼합모형을 이용한 분석에서 총 콜레스테롤의 경우, 포괄적인 유전형 집단의 효과가 유의하게 나타났으며($p=0.024$), 성별과의 교호작용에 대한 효과 또한 유의하게 나타났다($p=0.036$). 중성지방의 경우에는 교호작용의 효과가 유의하게 나타났고, HDL-콜레스테롤에서는 유전형 집단의 효과가 유의하게 존재하고 있었으며($p=0.043$), 교호작용에 대한 효과도 유의하게 나타났다($p=0.037$). 구체적으로 어떠한 조합의 유전형 집단에 의한 효과가 유의하게 나타나는지 그리고 그러한 효과가 성별에 따른 차이를 보이는 지를 개별적으로 평가하였다. 총 콜레스테롤에서는 (CC, CC, AA, ID), (CC, CC, GG, ID), (CT, CC, GG, II)의 경우 남자에서 더 낮은 것과 관련성을 보이고 있었다. 중성지방에서는 (CC, CC, AA, ID), (CC, CC, GA, II), (CC, CT, GG, ID), (CC, CT, GG, II)들에서 남자에서 더 낮은 것과 관련성을 보이고 있었으며, (CC, CC, AA, II), (CT, CT, AA, ID) 및 (CT, CT, GG, ID)에서는 그와 반대의 관련성이 유의하게 존재하고 있었다. HDL-콜레스테롤에서는 (CC, CC, GG, II)와 (CT, CC, GA, ID)에서 포괄적인 유전형 집단의 효과와 성별과의 교호작용이 유의하게 존재했다.

결 론 :

본 연구에서는 혈중 지질농도와 같은 양적 형질에서 관찰되는 변동량이 다중의 단일 염기 다형성으로 구성된 유전형 집단, 혹은 환경적 요인과의 교호작용 등에 의하여 영향을 받는지 확인하기 위한 통계학적 방법으로 혼합모형을 이용했다. 혈중 지질농도에서 각 유전형 집단들은 성별과의 유의한 교호작용이 존재하고 있었고 그 양상은 유전형 집단별로 각기 다르게 나타나고 있었다.

중심 단어 : 유전자 ; 혈중 지질농도 ; 관련성.

본 논문은 보건복지부 보건의료바이오기술개발사업 중 특정센터연구지원사업(00-PJ3-PG6-GN01-0001)의 연구비 지원을 받아 이루어졌음.

REFERENCES

- 1) Murray CJ, Lopez AD. Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet* 1997;349:1269-76.
- 2) Frick MH, Elo O, Haapa K, et al. Helsinki Heart Study: primary prevention trial with gemfibrozil in middle-aged men with dyslipidemia: safety of treatment, changes in risk factors and

- incidence of coronary heart disease. *N Engl J Med* 1987;317:1237-45.
- 3) Sigurdsson G Jr, Gudnason V, Sigurdsson G, Humphries SE. Interaction between a polymorphism of apo A-I promoter region and smoking determines plasma levels of HDL and apo A-I. *Arterioscler Thromb* 1992;12:1017-22.
 - 4) Wilson PW, Anderson KM, Harris T, Kannel WB, Castelli WP. Determinants of change in total cholesterol and HDLC with age. *J Gerontol* 1994;49:M252-7.
 - 5) O'Connell BJ, Genest J Jr. High density lipoproteins and endothelial function. *Circulation* 2001;104:1978-83.
 - 6) Sacks FM, Alaupovic P, Moye LA, et al. VLDL, apolipoproteins B, CIII, and E, and risk of recurrent coronary events in the cholesterol and recurrent events (CARE) trial. *Circulation* 2000;102:1886-92.
 - 7) Kessling AM, Horsthemke B, Humphries SE. A study of DNA polymorphism around the human apolipoprotein A1 gene in hyperlipidaemic and normal individuals. *Clin Genet* 1985;28:296-306.
 - 8) Lindpaintner K, Pfeiffer MA, Kreutz R, et al. A prospective evaluation of an angiotensin-converting enzyme gene polymorphism and the risk of ischemic heart disease. *N Eng J Med* 1995;332:706-11.
 - 9) Dallinga-Thie GM, Bu XD, van Linde-Sibebius TM, Rotter JJ, Lusis AJ, de Bruin TW. Apolipoprotein A1-CIII/AIV gene cluster in familial combined hyperlipidemia: effects on LDL-cholesterol and apolipoproteins B and CIII. *J Lipid Res* 1996;37:136-47.
 - 10) Wile DB, Barbir M, Gallagher J, et al. Apolipoprotein A-I polymorphisms: frequency in patients with coronary artery disease and healthy controls and association with serum apo A-I and HDL cholesterol concentration. *Atherosclerosis* 1989;78:9-18.
 - 11) Anderson RA, Burns TL, Lee J, Swendon D, Bristow JL. Restriction fragment length polymorphisms associated with abnormal lipid levels in an adolescent population. *Atherosclerosis* 1989;77:227-37.
 - 12) Kee F, Amouyel P, Fumeron F, et al. Lack of association between genetic variations of apo A1CIII-AIV gene cluster and myocardial infarction in a sample of European male. *Atherosclerosis* 1999;145:187-95.
 - 13) Saha N, Tay JS, Low PS, Humphries SE. Guanidine to adenine (G/A) substitution in the promoter region of the apolipoprotein A-I gene is associated with elevated serum apolipoprotein A-I levels in Chinese non-smokers. *Genet Epidemiol* 1994;11:255-64.
 - 14) Barre DE, Guerra R, Verstraete R, Wang Z, Grundy SM, Cohen JC. Genetic analysis of a polymorphism in the human apolipoprotein A-I gene promoter: effect on plasma HDL-cholesterol levels. *J Lipid Res* 1994;35:1292-6.
 - 15) Glazier AM, Nadeau JH, Aitman TJ. Finding genes that underlie complex traits. *Science* 2002;298:2345-9.
 - 16) Hoehe MR, Kopke K, Wendel B, et al. Sequence variability and candidate gene analysis in complex disease: association of opioid receptor gene variation with substance dependence. *Hum Mol Genet* 2000;9:2895-908.
 - 17) McCulloch CE, Searle SR. *Generalized, linear, and mixed Models*. 1st ed. New York: John Wiley & Sons; 2001.
 - 18) Littell RC, Milliken GA, Stroup WW, Wolfinger RD. *SAS system for mixed models*. 1st ed. Cary, NC: SAS Institute Inc; 1996.
 - 19) Verbeke G, Molenberghs G. *Linear mixed models in practice: a SAS-oriented approach*. 1st ed. New York: Springer; 1997.
 - 20) Brown H, Prescott R. *Applied mixed models in medicine*. 1st ed. New York: John Wiley & Sons; 1999.
 - 21) Kim J, Shin DJ, Bae Y, et al. Association between I/D, G14480C, A22982G polymorphisms of angiotensin I-converting enzyme gene and essential hypertension in the Korean population. *Korean Circ J* 2004;34:1137-47.
 - 22) Lee MM, Kim HS, Song JM, et al. Insertion/deletion polymorphism of angiotensin I-converting enzyme gene in Korean patients with ischemic heart disease. *Korean Circ J* 1996;26:5-13.
 - 23) Park HY, Kwon HM, Kim HS, Song KS, Kim CH. An I/D polymorphism in angiotensin-converting enzyme gene in myocardial infarction. *Korean Circ J* 1996;26:465-72.
 - 24) Nakai K, Itoh C, Miura Y, et al. Deletion polymorphism of the angiotensin I-converting enzyme gene is associated with serum ACE concentration and increased risk for CAD in the Japanese. *Circulation* 1994;90:2199-202.
 - 25) Araujo MA, Menezes BS, Lourenco C, Cordeiro ER, Gatti RR, Goulart LR. The angiotensinogen gene (M235T) and the acute myocardial infarction. *Rev Assoc Med Bras* 2005;51:164-9.
 - 26) Kumar A, Li Y, Patil S, Jain S. A haplotype of the angiotensinogen gene is associated with hypertension in African Americans. *Clin Exp Pharmacol Physiol* 2005;32:495-502.
 - 27) Say YH, Ling KH, Duraisamy G, Isaac S, Rosli R. Angiotensinogen M235T gene variants and its association with essential hypertension and plasma renin activity in Malaysian subjects: a case control study. *BMC Cardiovasc Disord* 2005;5:7.
 - 28) Fidani L, Hatzitolios AI, Goulas A, Savopoulos C, Basayannis C, Kotsis A. Cholesteryl ester transfer protein TaqI B and lipoprotein lipase Ser447Ter gene polymorphisms are not associated with ischaemic stroke in Greek patients. *Neurosci Lett* 2005;384:102-5.
 - 29) Brousseau ME, Diffenderfer MR, Millar JS, et al. Effects of cholesteryl ester transfer protein inhibition on high-density lipoprotein subspecies, apolipoprotein A-I metabolism, and fecal sterol excretion. *Arterioscler Thromb Vasc Biol* 2005;25:1057-64.