

사상체질별 Short Tandem Repeat 대립유전자 빈도

박화용 · 유현주* · 구임희* · 김종열*

한국한의학연구원 신한방제제연구센터, *한국한의학연구원 체질생물학의공학연구센터

Abstract

Short Tandem Repeat Allele Frequencies in Sasang Constitution

Hwayong Park, Hyun Joo Yu*, Imhoi Ku*, Jong Yeol Kim*

Herbal Medicine Improvement Research Center

Constitutional Biology & Medical Engineering Research Center

*Korea Institute of Oriental Medicine

1. Objectives

As a basic trial for identification of Sasang constitutional gene marker, we genotyped and analysed statistical relationships of STR(short tandem repeat) alleles and its distribution in each constitution.

2. Methods

After obtaining basic constitutional data with questionnaire (QSCC II), decision of constitution was made by 3 different constitution specialists' diagnosis, and only the samples of specialists' agreement of each constitution by discussion were taken into this research. Using multiplex PCR kit, total 146 constitutional samples were amplified in 16 autosomal STR marker, genotyped, and analysed statistically. Among 16 markers, 15 were analysed in this study excluding the amelogenin marker is used for in gender identification.

3. Results and Conclusions

It is difficult to determine the relationship between constitution and STR marker as the sample size is small, however, Penta D and vWA were shown to be related statistically with constitution. It has been know that STRs has no genetic informations, however there are some recent research results showing STRs as a regulatory element, relationship between microsatellite instability and repeat number and size, and post-transcriptional signaling. STRs which is not known about its function currently, are proposed to have function and/or regulatory activities anyhow with Sasang constitution. It is believed that the results of this study can help determine and identify the markers related to Sasang Constitutional Medicien.

Key words :: Sasang constitution, STR, Genotyping.

• 접수일 2009년 3월 8일; 승인일 2009년 3월 23일
• 연구비지원기관 : 한국한의학연구원
• 교신저자 : 김종열
(우)306-811. 대전광역시 유성구 전민동 엑스포로 483.
Tel : (042)868-9489 Fax : (042)868-9480
Email : ssmed@kiom.re.kr

I. 序 論

세계 각국에는 각각 고유의 전통의학이 있으며 이를 현대화하고 과학화 하려는 노력이 최근들어 더욱 활발히 진행되고 있으며 특히 중국에서는 이미 오래 전부터 꾸준히 관련 연구를 진행하여 오고 있다. 한편 유럽에서는 천연물, 또는 유효성분 추출물을 이용하여 현대의 화학적 합성 약물을 대체하려는 연구가 활발하다. 한국의 전통 한의학 역시 오랜 역사를 지니고 있지만 이에 대한 과학적 연구는 다른 이화학적 연구 분야에 비하여 그리 많지 않은 것이 사실이고 그나마 본격적인 연구가 시작된 것은 최근의 일이다. 한의학에서도 특히 사상체질의학은 100년이 넘는 역사를 가지고 있으며 현재에도 이제마의 사상체질의학에 근거한 진단과 치료는 일반적인 진단과 치료에 비해 월등한 효과를 보이고 있는 것이 사실이다. 특히 체질별로 처방을 달리하는 사상처방에 있어서 더욱 그러하다. 체질별로 약재를 달리하고 체질특성에 따라 약재의 가감을 달리하는 체질처방은 같은 질병이라 하더라도 체질별로 다르게 약을 처방함으로써 치료 효과를 극대화 할 수 있는 것이며 임상에서도 실제 효과의 우수성은 잘 알려진 사실이다.

사상체질이란 것은 사람의 성격, 체형, 용모사기, 성질재간 등 수많은 체질관련 요소들을 望, 問, 聞, 切의 四診에 의하여 판단하는 것으로서 생명과학적으로는 표현형(phenotype)이라 할 수 있는 것들이다. 하지만 이 표현형들에 관련된 생물학적 마커나 유전자로서 알려진 것은 현재 거의 없다. 왜냐하면 그동안 관련 생물학적 연구가 많지 않았던 것이 사실이고, 또한 시도된 연구에서도 체질과의 연관성을 보여줄 수 있는 큰 유의성이 있는 성과를 거두지 못하였기 때문이기도 하다. 또한 사상체질 표현형들은 생물학적 요소들에 의해서만 결정되는 것이 아니고 환경과 교육 및 섭생 등도 체질형성에 상당부분 기여하기 때문인 것으로 생각된다. 생물학적으로 체질이라는 것은 인간이라고 하는 생명체에서 볼 수 있는 표현형(phenotype)의 범주에 속하는 것으로서 생명과학적 방법을 이용한 체질관련 마커와 유전자의 발굴 및 이의 밀바탕에 있는 유전자형(genotype)을 밝히기 위한 연구를 시도할 수 있다.

체질은 한두 개, 또는 몇 개의 유전자에 의해 결정되는 것은 아니고 수많은 유전자와 이들의 발현 정도 및 그 조합에 따라 결정되는 소위 ‘복합형질’(complex trait)이라 할 수 있는 만큼 다양한 수많은 유전자와 마커에 대하여 무수한 연구와 시도가 이루어져야만 한다. 최근에는 DNA chip과 RNA microarray와 같은 대용량 실험방법과 도구들이 개발되어있는 상태이어서 수많은 유전자와 그 발현정도를 whole genome의 수준에서 관찰할 수 있게 된 것은 매우 다행스러운 일이다.

본 연구에서는 체질관련 유전자, 또는 유전자형을 발굴하기 위한 시도 중의 하나로서 친자확인, 신원확인, 범죄자확인 등에 사용되는 유전자 감식방법에 사용되는 미세위성체(microsatellite, 또는 STR; short tandem repeat)라고 하는 마커를 이용하여 체질별 genotyping을 수행하였다¹⁾. 특히 SINE (short interspersed transposable element), 또는 LINE (long interspersed transposable element)과 마찬가지로, 이들 유전자 마커에 어떤 유전정보가 있는 것은 아니지만 최근 이와 같은 유전정보가 없는 부위들의 존재 이유에 대하여 다양한 가설과 의견들이 대두되고 있고, 현재는 우리가 알지 못하지만 어떤 중요한 역할을 할 것이라는 의견이 점점 설득력을 얻고 있는 상황에서 체질을 대상으로 STR 마커의 빈도와 분포를 관찰하여 어떠한 유의성이 있는가를 관찰하고자 하였다.

II. 材料 및 方法

체질별 DNA 시료는 체질이 확진된 사람 146명의 말초혈액으로부터 DNA 추출 kit(QIAamp DNA Blood Midi Kit, Qiagen, Hilden, Germany)를 사용하여 추출하였으며 working DNA의 최종 농도는 10ng/ μ l로 조절하여 실험에 사용하였다. 피험자는 모두 혈연관계가 없는 사람들로써 소양인 49명, 소음인 48명, 태음인 49명 등이며 회귀체질인 태양인은 시료를 구할 수 없어 본 연구에 포함하지 않았다. 체질의 진단은 우선 체질설문지 QSCC II를 사용하여 기초적인 체질관련 정보를 획득한 후 3인의 사상체질전문의에 의한 진단을 실시한 다음 논의를 거쳐 체질진단결과가 일치하는 대상을 연구에 포함하였다.

Multiplex PCR은 PowerPlex® 16 System manual을 참고하여 genomic DNA 10~20ng에 PowerPlex® 16 primer pair mix (Promega, USA)와 Gold STR buffer를 final 1×가 되도록 가하고 1 unit의 AmpliTaq Gold DNA polymerase를 더하여 최종 reaction volume 20μl를 GeneAmp PCR system 9700 (Applied Biosystems, Foster City, CA, USA)에서 시행하였다²⁾. PCR amplification의 조건은 다음과 같다. Initial denaturation은 95°C 11분, 96°C 2분 진행한 후 94°C 1분, 60°C 1분, 70°C 90초의 반응을 10 cycle, 그리고 90°C 1분, 60°C 1분, 70°C 90초의 반응을 20 cycle 반복하고 final extension은 60°C에서 30분간 반응시켰다.

반응이 끝난 PCR product는 ABI 3100 automatic sequencer(Applied Biosystems, Foster City, CA, USA)를 이용하여 분획한 다음 GeneScan과 Genotyper software (ver 3.7)를 이용하여 genotyping하고 분석하였다. ABI 3100 sequencer에 loading하기 전에 증폭된 PCR product에서 1.0μl를 취하여 formamide/internal lane standard mix(Hi-Di formamide 9.5μl + ILS600 0.5μl)와 섞어 95°C에서 3분간 denaturation시킨 다음 얼음에 급냉시킨 후 loading하였다. PowerPlex® 16 System은 microsatellite 마커 Penta E, D18S51, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, vWA, Amelogenin, Penta D, CSF1PO, D16S539, D7S820, D13S317, D5S818의 16가지로 구성되어있으며 이들 16가지의 마커가 하나의 tube 내에서 한번의 PCR 반응으로 증폭될 수 있는 시약이다. 16가지의 마커 중 성별 감별을 위한 Amelogenin을 제외한 나머지 15가지의 마커를 본 연구의 분석 대상으로 삼았다.

Ⅲ. 結 果

본 연구에 회귀체질인 태양인의 시료는 포함하지 못하였으며 나머지 소양, 소음, 태음의 세가지 체질 146명의 DNA 시료를 연구대상으로 시행한 genotyping 결과 및 allele의 빈도와 분포 및 repeat structure는 Appendix의 table 1부터 table 6에 정리하였다.

결과적으로 15가지의 STR repeat 마커들 중에서 Penta D의 경우 p value가 0.0198로서 통계적 유의성이 가장 높은 STR marker로 나타났으며, 그 다음으로는

vWA에서 0.0239의 p value를 얻었다. 나머지 13개의 marker들은 모두 p value가 0.05이상으로서 유의성이 없는 것으로 판단하였다. Penta D는 chromosome 21번의 21q22.3에 위치하고 있으며 GenBank (<http://www.ncbi.nlm.nih.gov/Genbank/>)에 13개의 repeat를 갖는 sequence (AP001752)가 등록되어 있는데 본 연구결과 세가지 모든 체질에서 9개의 repeat를 갖는 allele 9가 가장 높은 빈도(소양 0.2857, 소음 0.3646, 태음 0.3571)로 관찰되었고 체질 사이에서 가장 큰 차이를 보이는 것은 allele 8인데 소양 0.0408, 소음 0.0417, 태음 0.0102의 빈도로 태음에 비해 나머지 체질에서 4배 이상의 빈도를 보였으며 allele 12와 13에서도 체질별 빈도의 차이를 볼 수 있었다(Table 1).

두 번째로 큰 체질 유의성을 보이는(p value 0.0239) vWA는 염색체상의 위치가 12p13.31이며 von Willebrand Factor A의 40번째 intron에 존재한다. GenBank에는 18개의 repeat를 갖는 sequence(M25858)가 등록되어 있으며 본 연구에서는 17개의 repeat를 갖는 allele 17이 가장 높은 빈도로 관찰되었다. 체질별로는 소양에서 allele 18이 0.2653, 소음에서 allele 17이 0.3125, 태음에서는 allele 17이 0.3061로서 각각 가장 높은 빈도를 보였으며, 체질별로는 minor allele을 제외한 나머지 allele 14, 15, 16, 17, 18에서 유의한 차이를 보였었다(Table 2).

Penta D와 vWA에 대한 Pearson chi-square test를 통한 통계량을 보면, 우선 Penta D의 경우 allele 12는 태음체질에 대하여 상당히 낮은 연관성을 가지고 있지만 소음과 소양체질에서는 매우 큰 연관성을 가짐을 알 수 있다. 따라서 Pearson chi-square test를 통한 검정통계량은 26.906, 자유도는 14, 접근 유의확률(양측검정)은 0.02 이므로 유의수준을 0.05로 선택하였을 때 전체적으로 Penta D의 체질 연관성이 높음을 알 수 있다(Table 3). vWA의 경우는 allele 14이 소음체질에서는 연관성이 작지만 태음과 소양의 체질 연관성이 큼을 알 수 있고, 특히 태음은 allele 16과의 연관성이 강함을 알 수 있다. Chi-square test 값이 가장 크게 나타난 소양의 allele 20은 기대빈도가 5보다 작아서 그 신뢰도가 높지는 않지만 전체적으로 Pearson chi-square test의 통계량은 22.916, 자유도는 12, 접근 유의확률(양측검정)은 0.028을 얻음으로써 유의수준 0.05를 기

Table 2. Observed Allele Frequencies of D13S317, D7S820, D16S539, CSF1PO, and Penta D (χ^2 Tests are 17.8114, 16.8582, 18.8187, 13.5463, 26.9060, and p values are 0.1215, 0.2638, 0.0930, 0.1947, 0.0198, respectively).

STR Locus	Allele	소양(n=49)	소음(n=48)	태음(n=49)	Total(n=146)	Repeat Structure
Penta D	8	0.0408	0.0417	0.0102	0.0308	[AAAGA] ₈
	9	0.2857	0.3646	0.3571	0.3356	[AAAGA] ₉
	10	0.1122	0.1875	0.0918	0.1301	[AAAGA] ₁₀
	11	0.1531	0.1979	0.2245	0.1918	[AAAGA] ₁₁
	12	0.2551	0.0729	0.1633	0.1644	[AAAGA] ₁₂
	13	0.1327	0.0938	0.1327	0.1199	[AAAGA] ₁₃
	14	0.0000	0.0417	0.0204	0.0205	[AAAGA] ₁₄
	15	0.0204	0.0000	0.0000	0.0068	[AAAGA] ₁₅
D13S317	8	0.3367	0.2396	0.3265	0.3014	[TATC] ₈
	9	0.1327	0.1979	0.1735	0.1678	[TATC] ₉
	10	0.0918	0.1667	0.0612	0.1062	[TATC] ₁₀
	11	0.2551	0.2083	0.2449	0.2363	[TATC] ₁₁
	12	0.1429	0.1250	0.1837	0.1507	[TATC] ₁₂
	13	0.0408	0.0625	0.0000	0.0342	[TATC] ₁₃
	14	0.0000	0.0000	0.0102	0.0034	[TATC] ₁₄
D7S820	7	0.0102	0.0000	0.0000	0.0034	[GATA] ₇
	8	0.1122	0.1146	0.0510	0.0925	[GATA] ₈
	9	0.1020	0.0417	0.0714	0.0719	[GATA] ₉
	10	0.1633	0.1458	0.2245	0.1781	[GATA] ₁₀
	11	0.3571	0.3229	0.4286	0.3699	[GATA] ₁₁
	12	0.2041	0.3125	0.1633	0.2260	[GATA] ₁₂
	13	0.0408	0.0417	0.0306	0.0377	[GATA] ₁₃
	14	0.0102	0.0208	0.0306	0.0205	[GATA] ₁₄
D16S539	8	0.0000	0.0104	0.0000	0.0034	[GATA] ₈
	9	0.2857	0.2917	0.4286	0.3356	[GATA] ₉
	10	0.1939	0.0938	0.0918	0.1267	[GATA] ₁₀
	11	0.2449	0.3125	0.2245	0.2603	[GATA] ₁₁
	12	0.1837	0.2083	0.1939	0.1952	[GATA] ₁₂
	13	0.0918	0.0833	0.0408	0.0719	[GATA] ₁₃
	14	0.0000	0.0000	0.0204	0.0068	[GATA] ₁₄
CSF1PO	8	0.0102	0.0000	0.0000	0.0034	[AGAT] ₈
	9	0.0612	0.0417	0.0510	0.0514	[AGAT] ₉
	10	0.1735	0.3125	0.1429	0.2089	[AGAT] ₁₀
	11	0.2347	0.2083	0.2143	0.2192	[AGAT] ₁₁
	12	0.4694	0.3958	0.5000	0.4555	[AGAT] ₁₂
	13	0.0510	0.0417	0.0918	0.0616	[AGAT] ₁₃

Table 3. Observed Allele Frequencies of vWA, D8S1179, and TPOX (χ^2 tests are 23.4842, 15.0932, 12.3450 and p values are 0.0239, 0.5178, 0.1365, respectively).

STR Locus	Allele	소양(n=49)	소음(n=48)	태음(n=49)	Total(n=146)	Repeat Structure
vWA	14	0.2551	0.1875	0.1224	0.1884	TCTA[TCTG] ₄ [TCTA] ₁₁
	15	0.0306	0.0104	0.0102	0.0171	TCTA[TCTG] ₄ [TCTA] ₁₀ TCCATCTA, or TCTA[TCTG] ₃ [TCTA] ₁₁ TCCATCTA
	16	0.1633	0.1667	0.2959	0.2089	TCTA[TCTG] ₄ [TCTA] ₁₁ TCCATCTA, or TCTA[TCTG] ₃ [TCTA] ₁₂ TCCATCTA
	17	0.1735	0.3125	0.3061	0.2637	TCTA[TCTG] ₄ [TCTA] ₁₂ TCCATCTA
	18	0.2653	0.2396	0.1837	0.2295	TCTA[TCTG] ₄ [TCTA] ₁₃ TCCATCTA
	19	0.0816	0.0833	0.0816	0.0822	TCTA[TCTG] ₄ [TCTA] ₁₄ TCCATCTA
	20	0.0306	0.0000	0.0000	0.0103	TCTA[TCTG] ₄ [TCTA] ₁₅ TCCATCTA
	D8S1179	10	0.1327	0.1354	0.0816	0.1164
11		0.1224	0.1250	0.0816	0.1096	[TCTA] ₁₁
12		0.1429	0.1354	0.2041	0.1610	[TCTA] ₁₂
13		0.2041	0.2708	0.2653	0.2466	[TCTA] ₁ [TCTG] ₁ [TCTA] ₁₁
14		0.1939	0.0833	0.1735	0.1507	[TCTA] ₁ [TCTG] ₁ [TCTA] ₁₂
15		0.1633	0.1771	0.1531	0.1644	[TCTA] ₁ [TCTG] ₁ [TCTA] ₁₃
16		0.0204	0.0625	0.0408	0.0411	[TCTA] ₂ [TCTG] ₁ [TCTA] ₁₃
17		0.0102	0.0000	0.0000	0.0034	[TCTA] ₂ [TCTG] ₂ [TCTA] ₁₃
TPOX	8	0.4796	0.4167	0.4592	0.4521	[AATG] ₈
	9	0.0714	0.1146	0.1735	0.1199	[AATG] ₉
	10	0.0306	0.0208	0.0714	0.0411	[AATG] ₁₀
	11	0.3776	0.4167	0.2857	0.3596	[AATG] ₁₁
	12	0.0408	0.0313	0.0102	0.0274	[AATG] ₁₂

Table 3. Pearson χ^2 Test Stastic of Penta D and vWA.

Allele	Penta D			Allele	vWA		
	소양(n=49)	소음(n=48)	태음(n=49)		소양(n=49)	소음(n=48)	태음(n=49)
8	1.35161	0.36631	0.31760	14	2.2600	0.0004	2.3179
9	0.13531	0.24001	0.72715	15	0.2740	0.2522	1.0413
10	1.10466	2.42736	0.24107	16	3.5519	0.8198	0.9771
11	0.54671	0.01885	0.76609	17	0.6689	0.8670	3.0256
12	0.00075	4.88586	4.90636	18	0.8951	0.0429	0.5490
13	0.13375	0.54614	0.13375	19	0.0004	0.0015	0.0004
14	0.00009	2.08371	2.01370	20	1.0068	0.9863	3.9456
15	0.67123	0.65753	2.63042	-	-	-	-

Table 4. Observed Allele Frequencies of D18S51, Penta E, and D5S818 (χ^2 Tests are 27.9358, 34.9222, 16.6018, and p values are 0.3616, 0.2455, 0.1652, respectively).

STR Locus	Allele	소양(n=49)	소음(n=48)	태음(n=49)	Total(n=146)	Repeat Structure
D18S51	11	0.0102	0.0104	0.0000	0.0068	[AGAA] ₁₁
	12	0.0306	0.0625	0.1122	0.0685	[AGAA] ₁₂
	13	0.1531	0.2500	0.1939	0.1986	[AGAA] ₁₃
	14	0.3061	0.2188	0.1939	0.2397	[AGAA] ₁₄
	15	0.1939	0.1771	0.1633	0.1781	[AGAA] ₁₅
	16	0.0714	0.0729	0.1327	0.0925	[AGAA] ₁₆
	17	0.0816	0.0833	0.0612	0.0753	[AGAA] ₁₇
	18	0.0510	0.0313	0.0306	0.0377	[AGAA] ₁₈
	19	0.0612	0.0208	0.0408	0.0411	[AGAA] ₁₉
	20	0.0204	0.0208	0.0408	0.0274	[AGAA] ₂₀
	21	0.0000	0.0208	0.0204	0.0137	[AGAA] ₂₁
	22	0.0000	0.0208	0.0000	0.0068	[AGAA] ₂₂
	23	0.0204	0.0000	0.0102	0.0103	[AGAA] ₂₃
	24	0.0000	0.0104	0.0000	0.0034	[AGAA] ₂₄
Penta E	5	0.0714	0.0625	0.0918	0.0753	[AAAGA] ₅
	9	0.0102	0.0208	0.0000	0.0103	[AAAGA] ₉
	10	0.0408	0.0208	0.0306	0.0308	[AAAGA] ₁₀
	11	0.1327	0.1250	0.0306	0.0959	[AAAGA] ₁₁
	12	0.0816	0.1042	0.1224	0.1027	[AAAGA] ₁₂
	13	0.0102	0.0313	0.0408	0.0274	[AAAGA] ₁₃
	14	0.0306	0.0313	0.0714	0.0445	[AAAGA] ₁₄
	15	0.1633	0.1146	0.0816	0.1199	[AAAGA] ₁₅
	16	0.0714	0.0938	0.1224	0.0959	[AAAGA] ₁₆
	17	0.0714	0.0625	0.0204	0.0514	[AAAGA] ₁₇
	18	0.1633	0.1771	0.2857	0.2089	[AAAGA] ₁₈
	19	0.0816	0.0417	0.0612	0.0616	[AAAGA] ₁₉
	20	0.0306	0.0833	0.0306	0.0479	[AAAGA] ₂₀
21	0.0102	0.0104	0.0102	0.0103	[AAAGA] ₂₁	
22	0.0204	0.0208	0.0000	0.0137	[AAAGA] ₂₂	
23	0.0102	0.0000	0.0000	0.0034	[AAAGA] ₂₃	
D5S818	7	0.0204	0.0000	0.0000	0.0068	[AGAT] ₇
	9	0.1224	0.1146	0.1122	0.1164	[AGAT] ₉
	10	0.1020	0.2188	0.1531	0.1575	[AGAT] ₁₀
	11	0.3673	0.2917	0.2857	0.3151	[AGAT] ₁₁
	12	0.2653	0.2292	0.2143	0.2363	[AGAT] ₁₂
	13	0.1224	0.1354	0.2347	0.1644	[AGAT] ₁₃
	14	0.0000	0.0104	0.0000	0.0034	[AGAT] ₁₄

Table 6. Observed Allele Frequencies of D3S1358, TH01, and D21S11 (χ^2 Tests are 10.7007, 13.7915, 28.8084, and p values are 0.2192, 0.1825, 0.1504, respectively).

STR Locus	Allele	소양(n=49)	소음(n=48)	태음(n=49)	Total(n=146)	Repeat Structure
D3S1358	14	0.0408	0.0313	0.0816	0.0514	TCTA(TCTG) ₂ [TCTA] ₁₀
	15	0.3878	0.3333	0.3469	0.3562	TCTA(TCTG) ₃ [TCTA] ₁₁
	16	0.3265	0.2917	0.3469	0.3219	TCTA(TCTG) ₃ [TCTA] ₁₂
	17	0.2245	0.2396	0.1735	0.2123	TCTA(TCTG) ₃ [TCTA] ₁₃
	18	0.0204	0.1042	0.0510	0.0582	TCTA(TCTG) ₃ [TCTA] ₁₄
TH01	6	0.1531	0.2188	0.2245	0.1986	{AATG} ₆
	7	0.3776	0.2188	0.2959	0.2979	{AATG} ₇
	8	0.0612	0.0417	0.0204	0.0411	{AATG} ₈
	9	0.3469	0.4271	0.4286	0.4007	{AATG} ₉
	9.3	0.0204	0.0625	0.0204	0.0342	{AATG} ₆ ATG{AATG} ₃
	10	0.0408	0.0313	0.0102	0.0274	{AATG} ₁₀
D21S11	27	0.0000	0.0104	0.0306	0.0137	{TCTA} ₄ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₉
	28	0.0816	0.1042	0.0102	0.0651	{TCTA} ₄ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₀
	29	0.1633	0.1563	0.1837	0.1678	{TCTA} ₄ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₁
	29.2	0.0102	0.0000	0.0000	0.0034	{TCTA} ₅ {TCTG} ₅ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₀ TATCTA
	30	0.3980	0.4271	0.3776	0.4007	{TCTA} ₆ {TCTG} ₅ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₁
	30.2	0.0102	0.0000	0.0000	0.0034	{TCTA} ₅ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₀ TATCTA
	30.3	0.0000	0.0104	0.0204	0.0103	variant allele
	31	0.1224	0.0521	0.0918	0.0890	{TCTA} ₅ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₂
	31.2	0.0816	0.0833	0.0816	0.0822	{TCTA} ₅ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₁ TATCTA
	32	0.0000	0.0000	0.0306	0.0103	{TCTA} ₆ {TCTG} ₅ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₃
	32.2	0.0714	0.1250	0.1327	0.1096	{TCTA} ₅ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₂ TATCTA
	33.2	0.0612	0.0313	0.0408	0.0445	{TCTA} ₅ {TCTG} ₆ {TCTA} ₃ TA {TCTA} ₃ TCA{TCTA} ₂ TCCATA {TCTA} ₁₃ TATCTA

Table 7. Observed Allele Frequencies of FGA (χ^2 Test 21.6164, p Value 0.7095).

Allele	소양(n=49)	소음(n=48)	태음(n=49)	Total(n=146)	Repeat Structure
17	0.0000	0.0000	0.0102	0.0034	{TTTC ₃ TTTTTTCT{CITT} ₉ CTCC{TTCC} ₂
18	0.0306	0.0208	0.0204	0.0240	{TTTC ₃ TTTTTTCT{CITT} ₁₀ CTCC{TTCC} ₂
19	0.0510	0.0833	0.0510	0.0616	{TTTC ₃ TTTTTTCT{CITT} ₁₁ CTCC{TTCC} ₂
20	0.0918	0.0729	0.0714	0.0788	{TTTC ₃ TTTTTTCT{CITT} ₁₂ CTCC{TTCC} ₂
21	0.0714	0.1146	0.1122	0.0993	{TTTC ₃ TTTTTTCT{CITT} ₁₃ CTCC{TTCC} ₂
21.2	0.0102	0.0000	0.0000	0.0034	{TTTC ₃ TTTTTT{CITT} ₁₄ CTCC{TTCC} ₂
22	0.1837	0.1667	0.1531	0.1678	{TTTC ₃ TTTTTTCT{CITT} ₁₄ CTCC{TTCC} ₂
23	0.2143	0.2083	0.2347	0.2192	{TTTC ₃ TTTTTTCT{CITT} ₁₅ CTCC{TTCC} ₂
23.2	0.0204	0.0000	0.0000	0.0068	{TTTC ₃ TTTTTT{CITT} ₁₆ CTCC{TTCC} ₂
24	0.1735	0.2083	0.2143	0.1986	{TTTC ₃ TTTTTTCT{CITT} ₁₆ CTCC{TTCC} ₂
24.2	0.0102	0.0000	0.0306	0.0137	{TTTC ₃ TTTTTT{CITT} ₁₇ CTCC{TTCC} ₂
25	0.1327	0.0833	0.0510	0.0890	{TTTC ₃ TTTTTTCT{CITT} ₁₇ CTCC{TTCC} ₂
26	0.0102	0.0313	0.0408	0.0274	{TTTC ₃ TTTTTTCT{CITT} ₁₈ CTCC{TTCC} ₂
27	0.0000	0.0104	0.0102	0.0068	{TTTC ₃ TTTTTTCT{CITT} ₁₉ CTCC{TTCC} ₂

준으로 관찰한 결과 체질과 vWA의 연관성이 높다고 할 수 있다.

IV. 考 察

인간 유전체의 약 80~90%는 소위 ‘junk DNA’라 불리는 유전정보로서의 의미가 없는 부분으로 구성되어 있다. STR 역시 intron이나 intergenic DNA에 속하는 것들로서 수백만 년 동안의 인류 진화과정 동안 유지 되어오면서 그 repeat의 수와 형태가 점점 다양하게 변화해왔다. 그런데 이러한 repeat 들이 과연 정말로 아무런 의미가 없는 반복되어 배열된 단순 염기들에 불과한 것일까? 유전적 의미는 없다 하더라도 구조적

으로 어떤 역할을 하거나, 또는 아직 알지 못하는 어떤 기작을 통해 transcription을 regulation하는 것은 아닐까?

이와같은 repeat들에 대해서 그 존재 이유와 의미를 재해석 해보고자 하는 노력들이 최근에 진행되어 왔다. 일반적으로 DNA 복제의 slippage나 unequal crossing-over가 microsatellite에서의 새로운 mutation을 유발하는 mechanism인 것으로 알려져 있는데 이러한 event 들이 일어나는 microsatellite들 중에서도 특히 expression gene에 근접해 있는 repetitive sequence들은 진화과정에 있어 regulatory element로서 중요한 역할을 한다는 보고가 있다³⁾.

한편 mutation rate나 microsatellite instability는 homo-

geneous repeat의 크기 및 수와 매우 밀접한 상관관계가 있다는 연구결과가 있는데, Wierdl 등이 yeast에서의 tandem repeat를 대상으로 연구한 바에 의하면 micro-satellite instability rate와 single repeat addition rate는 repeat unit의 길이가 길수록 증가하며, 2개 이상의 repeat가 deletion되는 large deletion은 33bp 이하의 repeat 보다 는 주로 repeat unit가 51bp 이상되는 경우에 발생함을 관찰하기도 하였다⁴⁾.

Short tandem repeat(STR)들은 highly polymorphic하고 그 sequence는 common하며 genome 전체에 고루 random하게 산재해 있는 특성으로 인하여 종종 질병 관련 유전자의 genotyping에도 사용되고 있는데 2001년 Riley의 prostate carcinoma에 관한 연구를 보면, phosphoglycerate kinase(PGK1) gene의 STR allelic distribution이 환자와 정상인 사이에서 유의하게 차이가 있음을 알 수 있다⁵⁾.

또한 역시 같은 연구자인 Riley의 2002년 보고에 의하면 수많은 repeat들 중에서 가장 흔한 것 중의 하나가 GATA repeat인데, GenBank database를 조사하여 분석한 결과 GATA가 있거나 또는 없는 spliced variants들에게서 STR을 갖는 transcript들이 더욱 많이 관찰되었다. 이는 GATA repeat가 어떤 기능을 하고 있음을 의미하는 것이며 그 대부분의 기능은 membrane-associated protein과 관련된 mRNA에 대한 post-transcriptional signal을 encoding하는 것으로 여겨진다⁶⁾.

사상체질이라는 표현형을 결정하는, 또는 그에 관여하는 유전자나 생물학적 표현형 마커는 아직까지 구체적으로 알려진 것이 없다. 체질이라는 것 자체가 생물학적으로는 복합형질(complex trait)의 하나로 볼 수 있는데 여기에는 수많은 유전자를 포함하는 생물학적 요인과 개체발생 및 발달과정, 그리고 교육과 학습 등의 사회 환경적 요인까지 영향을 미치는 것으로 여겨진다. 한의학에서의 사상체질의학 개념에는 생물학적인 측면과 물리학적 측면 뿐 아니라 정신과적 측면, 그리고 사회인류학적 측면까지도 고려하고 있기 때문이다. 이렇게 많은 요인 중에서 생물학적 요인과 마커를 발굴하기 위하여 많은 시도가 있어왔고 지금도 진행되고 있지만 아직 뚜렷한 성과는 없다. 그러나 이 같은 노력들과 실험적 시도들에 의한 무수

한 결과들이 축적되면 결국에는 사상체질과 연관된 마커를 발굴할 수 있을 것이며 궁극적으로는 사상체질을 결정하는 유전자를 밝힐 수 있을 것으로 여겨진다.

특히 본 연구결과는 STR에 대한 연구들에서 표현형, 질병, 유전자의 기능 등과 관련된 유의한 결과들이 보고되고 있는 최근의 상황에서 사상체질이라는 복합형질을 결정하는, 또는 그 형질의 형성에 기여하는 체질유전자의 발굴에 중요한 기초적 연구결과가 될 수 있을 것으로 사료된다. 또한 체질과 관련된 유전자는 그 수가 매우 많을 것으로 생각되며 따라서 향후 생물학 뿐 아니라 타 연구 분야에서의 관련 연구결과에서 발생하는 유의한 결과와 융합된 과학적 체질진단과 해석에 유용한 지표로 활용될 수 있을 것으로 기대된다.

V. 結 論

희귀체질 태양인을 제외한 소양, 소음, 태음의 세 가지 체질 146명의 시료를 대상으로 15가지의 autosomal STR marker를 이용하여 그 빈도를 관찰하고 체질별 유의성이 있는지를 분석하였다. 이 중에서 p value가 0.05이하로서 통계적 유의성을 보이는 것은 Penta D와 vWA의 두 가지이다. 이와 같은 결과는 통계적 방법에 의한 수치상의 결과이며 이 연구의 결과에서 곧바로 체질과 연관된 유전자나 마커를 얻을 수 있는 것은 아니다. 다만 체질과의 연관 가능성을 의미하는 것으로 더욱 많은 샘플을 대상으로 연구하여 보다 더 통계적 의미가 더욱 큰 결과를 도출한 다음 이들 마커와 그 인접부위의 유전자와 sequence를 분석해 보는 것은 체질 유전자 연구에 있어 매우 큰 연구가치가 있을 것으로 사료된다.

본 연구의 목적도 역시 이와 같은 시도에서 이루어진 것이며, 위에 언급한 repeat들의 잠재적 기능이나 역할에 관련된 연구결과들과 같은 차원에서 체질유전자 발굴을 위한 연구의 기초가 될 수 있을 것으로 사료된다. STR repeat와 같이 현재에는 기능이 알려지지 않거나, 또는 우리가 알지 못하는 유전자 부위의 경우 체질유연성이 없어 보이는 무의미한 것일 수 있지만 향후 관련 연구가 많이 진행되고 또 융합·복합 연구에 의한 새로운 차원에서의 연구결과들을

기대해 볼 때 본 연구에서와 같은 기초연구 또한 매우 중요한 시도이며, 체질유전자 발굴과 생물학적 체질 규명의 밑바탕이 될 수 있을 것이다.

VI. 感謝의 글

본 논문은 한국한의학연구원 기관고유과제 ‘체질 건강수준 표준개발’ 연구비(과제번호 K08020) 지원에 의하여 수행되었음.

VII. 參考文獻

1. <http://www.cstl.nist.gov/biotech/strbase>
2. <http://www.promega.com>
3. Levinson G, Gutman GA. Slipped strand mispriming: a major mechanism for DNA sequence evolution. *Molecular Biology and Evolution*. 1987;4:203-221.
4. Wierdl M, Dominska M, Petes TD. Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics*. 1997 ; 146 : 769-779.
5. Riley DE, Krieger JN. Short tandem repeat polymorphism linkage to the androgen receptor gene in prostate carcinoma. *Cancer*. 2001 ; 92 : 2603-2608.
6. Riley DE, Krieger JN. Diverse eukaryotic transcripts suggest short tandem repeats have cellular functions. *Biochemical and Biophysical Research Communications*. 2002 ; 298 : 581-586.